A Thermal, Mechanical, and Materials Framework for a Shape Memory Alloy Heat Engine for Thermal Management and Energy Recovery

Maria Chikhareva
University of Central Florida

Part of the Materials Science and Engineering Commons
Find similar works at: https://stars.library.ucf.edu/etd2020
University of Central Florida Libraries http://library.ucf.edu

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2020- by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation
https://stars.library.ucf.edu/etd2020/1881
A THERMAL, MECHANICAL, AND MATERIALS FRAMEWORK FOR A SHAPE MEMORY ALLOY HEAT ENGINE FOR THERMAL MANAGEMENT AND ENERGY RECOVERY

by

MARIA CHIKHAREVA
Specialist Degree, Bauman Moscow State Technical University, 2013

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science
in the Department of Materials Science and Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Spring Term
2023

Major Professor: Raj Vaidyanathan
ABSTRACT

Shape memory alloy (SMA) heat engines possess an inherent property of sensing a change in temperature, performing work, and rejecting heat through the shape memory effect resulting from a temperature-induced phase transformation. This work presents a framework for the design and implementation of an SMA-based Stirling heat engine for maximum power or speed incorporating and combining mechanical, thermal, and material aspects.

The motivation for this work comes from the growing need for reliable thermal management and energy recovery in both ground and space applications of interest to NASA and commercial space companies. In future lunar and Mars missions, an SMA heat engine can be used in cryogenic liquefaction, densification, and zero boil-off (ZBO) systems for advanced spaceport applications. Terrestrial applications include energy recovery by transforming the excessive or rejected heat into mechanical work.

Mechanical aspects were addressed by performing force balances in the SMA element and focused on the resulting stress distribution. Thermal aspects were addressed by considering the temperature distribution and the heat transfer rate between the SMA element and the heat source or the heat sink. Materials issues were addressed by considering principles of classical thermodynamics and the roles of internal and external stresses especially the hysteresis between the forward and reverse phase transformation and the stored elastic strain energy. The role of microstructure through composition, precipitates present, variant interfaces and training are also emphasized. The aforementioned aspects were combined to present a figure of merit to aid in the design and implementation of an SMA Stirling heat engine. The implementation presented considers
a low temperature, low hysteresis R-phase transformation in the NiTiFe system for rejecting heat from a system. Support from the Fulbright Program is gratefully acknowledged.
Dedicated to my parents.
ACKNOWLEDGMENTS

"Education is a slow-moving but powerful force. It may not be fast enough or strong enough to save us from catastrophe, but it is the strongest force available for that purpose and in its proper place, therefore, is not at the periphery, but at the center of international relations”

J. William Fulbright

Having the mind of an engineer I cannot always call myself an eloquent person. I wish I could find the words to express my deepest gratitude and appreciation to everyone who supported me in my academic journey at the University of Central Florida.

First, I would like to thank my advisor, Professor Raj Vaidyanathan, who guided me not only in writing my thesis but also encouraged my professional growth. The first class I took at UCF was EMA 5104 which was by far the best class I have ever taken in eight years of college education. I knew instantly I wanted him to be my advisor. He introduced the Shape Memory Alloys’ world to me which I fell in love with. Dr. Raj has provided invaluable constructive feedback throughout the duration of my work with him. I want to thank him for the support he gave me when my whole world collapsed with the beginning of the war in Ukraine. His guidance has been crucial for my personal growth and my career aspirations. He went above and beyond being my mentor. Dr. Raj never let me forget my achievements and all the hard work I put into everything I do. In moments of doubt, Dr. Raj was the one to steady my nerves and restore my self-confidence. I am forever thankful.
Second, I would like to acknowledge the support of Dr. Akihiro Kushima, who was my committee member. Thank you for your belief in me and the time you spent writing recommendation letters for me. I treasure your knowledge and expertise in TEM that you shared with me as one of your students.

Third, I would like to express my immense gratitude to Dr. Jayan Thomas, who agreed to serve in my thesis committee before we even met in person. Thank you for your interest in my topic, thought-provoking questions and ideas.

My sincere gratitude also goes to Dr. Jiyu Fang and Pamela Ross for their help and support. Thank you for answering my endless questions, believing in me and guiding me.

I could not have done it without my family and friends who were always there for me while being thousands of miles away. I hope I make them proud.

Last but not the least, graduate school would not have been possible without the Fulbright scholarship. I also want to express my gratitude to the community of the Department of Materials Science and Engineering at the University of Central Florida. It was a pleasure to be part of it.
TABLE OF CONTENTS

ABSTRACT ..................................................................................................................... iii

ACKNOWLEDGMENTS ............................................................................................ vi

TABLE OF CONTENTS .............................................................................................. viii

LIST OF FIGURES .................................................................................................... xii

LIST OF VIDEOS ..................................................................................................... xiv

LIST OF TABLES ...................................................................................................... xv

LIST OF ABBREVIATIONS ......................................................................................... xvi

CHAPTER ONE: INTRODUCTION .................................................................. 1

1.1. Motivation ....................................................................................................... 1

1.2. Organization ................................................................................................... 3

CHAPTER TWO: LITERATURE REVIEW ......................................................... 4

2.1. Shape Memory Alloys .................................................................................... 4

2.1.1. Martensitic Transformation .................................................................. 4

2.1.2. Phase Transformations in NiTi Shape Memory Alloys ......................... 5

2.1.3. Alloying of NiTi ...................................................................................... 6

2.1.4. Shape Memory Effect ............................................................................. 7

2.1.5. Superelastic Effect .................................................................................. 9

2.1.6. Relationship between Superelasticity and Shape Memory Effect ....... 10
2.2. SMA heat engines ........................................................................................ 11

2.2.1. Pulley engines ....................................................................................... 12

2.2.2. Field engines ....................................................................................... 13

2.2.3. Crank engines ..................................................................................... 14

2.2.4. Reciprocating engines ...................................................................... 16

2.2.5. Miscellaneous engines ..................................................................... 16

2.3. Heating and cooling methods .................................................................. 17

2.3.1. Thermal radiation, gas/air, cold of space ........................................ 17

2.3.2. Hot/cold water ................................................................................... 17

CHAPTER THREE: PRINCIPAL OF SMA STIRLING ENGINE ..................... 18

CHAPTER FOUR: THERMAL AND MECHANICAL ANALYSES ............... 32

4.1. Mechanical analyses ............................................................................... 32

4.1.1. SMA element shape and dimensions .............................................. 32

4.1.2. Number of SMA elements .............................................................. 37

4.1.3. External friction .............................................................................. 37

4.2. Thermal analyses ................................................................................... 37

4.2.1. Temperature profile ....................................................................... 37

4.2.2. Temperature difference between the heat source and the heat sink ... 38

CHAPTER FIVE: THERMODYNAMIC ANALYSES ................................ 41
LIST OF FIGURES

Figure 1. Schematic diagram representing martensitic stress-strain (M), austenitic stress-strain curves for the SIM mode and slip deformation mode (A). ................................................................. 10
Figure 2. Schematic diagram representing the regions of SME, SE, and slip .................. 11
Figure 3. Synchronized SMA pulley engine [15] ........................................................................................................... 12
Figure 4. Unsynchronized SMA pulley engine [16] ........................................................... 13
Figure 5. Field SMA engine [17] ..................................................................................... 14
Figure 6. Axis center crank SMA engine [13] ................................................................. 15
Figure 7. Offset center crank SMA engine [12] ............................................................... 15
Figure 8. SMA reciprocating engine [18] ......................................................................... 16
Figure 9. Work cycle diagram of Stirling engine ............................................................. 18
Figure 10. Work cycle diagram of SMA Stirling engine .................................................. 19
Figure 11. Unsynchronized pulley type SMA Stirling engine .......................................... 20
Figure 12. Initial conditions of the wire shown in an idealized martensite stress-strain curve .................................................................................................................................. 21
Figure 13. Bending stress distribution across the wire diameter ................................. 21
Figure 14. Balance between circumferential force and tight side and slack side forces .. 23
Figure 15. Centrifugal force acting on the wire at high speed ........................................ 25
Figure 16. The total stress distribution profile while in the operating mode ................. 26
Figure 17. Wire temperature prior to heat source entrance and after exit ...................... 29
Figure 18. Wire temperature at the point of first contact of the wire with the cold pulley (hot section of the wire) and the last point of contact (cold section) .............................. 30
Figure 19. Schematic of the differential scanning calorimetry response of a NiTi-based shape memory alloy ................................................................. 44

Figure 20. Schematic of SMA work output during a thermal cycle, following Ref. [3, 22] ........................................................................................................................................... 46

Figure 21. Schematic of transformational hysteresis in single crystal Cu - 14Al - 2.5Ni. (a) single interface (b) multiple interfaces, following Ref. [5] ................................................................. 51

Figure 22. Schematic of DSC curves for (a) transformation without stored elastic strain energy (b) transformation with stored elastic strain energy following Ref. [5] ................ 53

Figure 23. Self-accommodated martensite structure, following Ref. [11] ............................. 55

Figure 24. Tensile and compression deformation of martensitic SMA, following Ref. [25] ........................................................................................................................................... 58

Figure 25. Tensile and compression deformation of SE SMA, following Ref. [25] ...... 60

Figure 26. Schematic of hysteresis in stress space for samples with and without induced plastic deformation, following Ref. [27] ........................................................................................................ 61

Figure 27. Typical actuator response of SMA. ................................................................. 63

Figure 28. Schematic Differential scanning Calorimetry results for the B2 ↔ B19’ and B2 ↔ R-phase transformations ........................................................................................................ 66

Figure 29. The martensitic transformation in NiTi ............................................................. 72

Figure 30. (a) SMA wire stress dependence on temperature (b) four wire positions where power is generated ........................................................................................................ 78
LIST OF VIDEOS

Video 1. Operation of a commercial SMA Stirling engine with an unsynchronized configuration

.................................................................................................................................................. 28
# LIST OF TABLES

Table 1 ................................................................................................................................. 54

Table 2 ................................................................................................................................. 57
LIST OF ABBREVIATIONS

A     Austenite
As    Austenite start
Af    Austenite finish
DSC   Differential scanning calorimetry
M     Martensite
Ms    Martensite start
Md    Martensite desist
Mf    Martensite finish
NASA  National Aeronautics and Space Administration
Rs    R-phase start
Rf    R-phase finish
SE    Superelasticity
SIM   Stress induced martensite
SMA   Shape memory alloy
TWSME Two-way shape memory effect
CHAPTER ONE: INTRODUCTION

1.1. Motivation

Currently, NASA pursues technology and innovation to advance the human exploration of space. Key priorities include Mars and moon explorations. However, one of the main limitations in deep space exploration is storage of rocket fuels at strictly defined temperatures [1]. The lunar and Martian outposts development could incorporate in-situ resource utilization approaches and thermal management is needed for zero boil-off control during day/night cycles. Liquified oxygen (LOX) and liquified methane (LCH₄) should be maintained, while the temperature of the space environment varies in broad ranges. Other potential needs include operations where variable heat transfer is required, including residual propellant scavenging, chill down of equipment, and long-term storage of ascent module propellants.

The heat leakage on the moon occurs through radiation. The temperatures are extreme and during the daytime can reach 400 K (127°C). When the sun goes down, temperatures can dip down to 40 K (-233°C) [2].

On Mars, the operational thermal devices need to be enclosed in vacuum to prevent convective heat transfer. In accordance with NASA MARS InSight Mission, the average temperature on Mars is 213 K (-60°C), although it varies from 120K (-153°C) near the poles during the Winter to 310 K (36°C) at midday near the equator.

Devices for thermal control are an integral part in achieving these goals for deep space exploration. Conventional thermal switches such as thermostatic elements and wax actuators require a sensor or actuating mechanism which makes the switch configuration complicated and less reliable. However, Shape Memory Alloy (SMA) thermal control
devices do not have such drawbacks. The actuation phenomenon in these “smart” materials is accomplished by inherent material properties based on a solid-solid phase transformation. This allows SMA thermal control devices to act as a sensor and an actuator together, which in turn increases the reliability of the process, simplifies the mechanism design, and reduces its dimensions, which are especially important in aerospace applications [2]. Moreover, SMAs have a significantly high force output-to-weight ratio in comparison with traditional materials [3].

Regardless of the potential of SMA thermal devices to expedite space exploration, they are not limited to extra terrestrial applications. Earth’s applications may include all types of heat rejection systems in conditions where sufficient difference in temperatures exists.

This work considers an SMA Stirling heat engine as a type of actuator used for the purpose of extracting mechanical energy from available differences in temperature. The principle of engine operation is based on the property of an SMA to reversibly change its shape under different temperatures (or stresses) by means of rearrangement atoms from austenite to martensite and vice versa. Previous work was concentrated on estimating the maximum work output and thermal efficiency of the engine and did not take into consideration the rate of engine operation [3].

The intent of this work is twofold. First, is to establish a framework for designing and implementing an SMA heat engine combining material, mechanical, and heat transfer aspects. This will help in developing a systematic approach for customizing SMAs for reliable heat transfer in a demanding environment. Second is to evaluate overall efficiency by examining maximizing torque or maximizing angular velocity.
1.2. Organization

The thesis is organized as follows: Chapter 2 provides an overview of SMAs and presents SMA heat engine conceptual designs. Chapter 3 discusses the operational principle of the traditional belt drive and an SMA Stirling engine with an unsynchronized configuration. Chapter 4 analyzes thermal and mechanical aspects which affect the performance of the SMA heat engine. Chapter 5 addresses the energy balance of the martensitic transformations by considering principles of classical thermodynamics. Chapter 6 discusses the microstructure of the material and the roles of internal and external stresses, hysteresis between the forward and reverse phase transformation, and preferred orientation in the evolving texture as well as mathematically describes the transformation. Chapter 7 discusses the results of the work and develops a Figure of Merit for quantitative analysis of SMA performance for maximum speed and maximum torque as engine outputs. In Chapter 8 case study for heat rejection at high speed is considered. Finally, Chapter 9 presents the conclusions and results and lists the outline for future development of SMA heat engines.
CHAPTER TWO: LITERATURE REVIEW

There are two sections in this chapter: a section on Shape Memory Alloys and a section on Shape Memory Heat engines. The Shape Memory Alloys section reviews the general martensitic transformation and main properties of the alloys associated with a thermoelastic martensitic transformation, as well as the strong material dependence on materials composition. The Shape Memory Heat engines section considers the main principles of the engine operation and compares different configurations.

2.1. Shape Memory Alloys

Shape memory alloys are a class of materials that have the ability to “remember” and recover their original shape upon heating, even after large deformation. This ability of the material to recover a significant amount of deformation is the result of a solid-solid martensitic transformation. In NiTi alloys, the transformation occurs between a high temperature cubic austenite phase and a low temperature monoclinic martensite phase or, under certain conditions, through an intermediate trigonal R-phase. The reversible martensitic phase transformation in shape memory alloys results in the shape memory and superelastic effects.

2.1.1. Martensitic Transformation

In general, the solid-solid martensitic transformation is a diffusionless transformation. A low temperature martensite phase is formed upon cooling from a higher temperature austenite phase by atoms rearranging. [3].

When martensite transformations are discussed, they can be classified into two categories: thermoelastic and non-thermoelastic. In case of a thermoelastic transformation,
the driving force for the transformation is very small, the interface between parent and martensite is very mobile, and the transformation is crystallographically reversible (martensite reverts to austenite in the original orientation). In the case of a non-thermoelastic transformation, where the driving force is large, the interface between phases is immobile once the martensite plate grows to some critical size, and the reverse transformation occurs by the renucleation of the parent phase. Therefore, the reverse transformation is not crystallographically reversible [4]. It is known that the shape memory effect (SME) and superelasticity (SE) are generally characteristic of a thermoelastic transformation [5]. The martensitic transformation is associated with a shape change and therefore, a large strain arises around the martensite plate in the austenitic matrix. However, twinning is an energetically favorable structure which reduces the strain around the martensitic particle.

2.1.2. Phase Transformations in NiTi Shape Memory Alloys

The solid-solid phase transformation in NiTi occurs between a high temperature parent or austenite phase, having a B2 cubic structure and a low temperature martensite phase with a B19’ monoclinic structure. However, under certain conditions [3], a multi-stage martensite transformation through an intermediate R-phase is possible. R-phase in NiTi has a trigonal crystal structure and in general its introduction can be achieved by coldworking, precipitation hardening or by the addition of ternary elements such Fe [3]. The main characteristics of any phase transformation are the transformation temperatures and hysteresis associated with it.

Each phase transition in a polycrystalline sample starts and finishes at different temperatures. Therefore, the transformation temperatures are specified for each phase and
denoted as $M_s$ for martensite start; $M_f$ for martensite finish; $A_s$ for austenite start; $A_f$ for austenite finish; $R_s$ for R-phase start and $R_f$ for R-phase finish.

The transformations upon cooling (forward, A→M) and heating (reverse, M→A) occur through different paths and exhibit a hysteresis. The hysteresis mainly arises because of the interfacial friction associated with the phase transformation and results in irreversible energy losses during the transformation [4]. It is considered in detail from an energetic perspective in Chapter 5.

2.1.3. Alloying of NiTi

The transformation temperatures in SMAs are highly sensitive to the chemical composition of the alloy [3]. Therefore, changing the stoichiometric composition of the alloy or its alloying can modify the transformation temperatures and therefore modify its application for different environments. In binary NiTi alloy, the highest $A_f$ is at 50 at % Ti-50 at % Ni which comes to about 120 °C. The general trend is that the shape recovery temperature decreases with increasing Ni content and does not change with decreasing Ni [4].

Addition of such elements as Fe (substitution for Ni in NiTi) separates the transformation temperature ranges for B2→R and R →B19’. The substitution of Fe or Co for Ni lowers the martensitic transformation temperatures. Pt, Ph, Hf, and Zr are elements that raise the transformation temperatures [4], [6]-[7].

In addition to shifting the transformational temperatures, some alloying additions can change the hysteresis width. For example, in Ti-Ni-Cu alloys, with increase of Cu the
transformation hysteresis decreases. In contrast, addition of Nb broadens the transformation hysteresis [4].

2.1.4. Shape Memory Effect

The SME is an inherent property of SMAs related to the material's ability to regain its original shape after applied deformation in the low temperature phase upon heating to the high temperature phase [8]. The effect is possible due to a reversible thermoelastic martensitic phase transformation and is driven by temperature changes.

In NiTi it was discovered by Buehler et al. in 1963 at the Naval Ordinance Laboratory (NOL) and later came to be commonly known as “NITINOL” [8].

While different alloys such as Cu-Al-Zn, Fe-Mn-Si, and others [3] exhibit the SME, NiTi-based alloys are the most widely used in commercial applications due to their superior mechanical properties and the ability to generate high recovery stresses up to 600 MPa and accommodate large recoverable strains up to 8% [3]-[4].

In contrast to the inherent SME, where only the shape of the austenite phase is remembered, it is possible to “train” the material to remember the shape of martensite too. In general, the process to induce the so-called two-way shape memory effect (TWSME) consists of thermomechanical cycling through the transformation region to acquire the martensitic shape [4].

2.1.4.1. TWSME

The process of inducing TWSME usually starts with heat treating the raw material to memorize a certain shape as a hot shape as in case of one-way shape memory effect. It is obtained by constraining the material in the desired shape, then heating it in a furnace at
an appropriate temperature for a particular time, with subsequent quenching in water or air. After this, a cold shape can be induced in the sample.

There are several methods to achieve it [9]: 1) one-time martensite deformation, 2) thermomechanical cycling treatment, and 3) re-heat treatment.

In the one-time martensite deformation method, the material is severely deformed beyond the material strain limit for complete recovery. When the sample is heated to the austenite phase, extensive deformation will not allow the material to recover the original shape, because of the presence of plastic strain introduced by a high level of deformation. This plastic deformation seriously deteriorates the memory of austenite. However, when the material is cooled again to martensite, the sample will move partially toward the overdeformed shape [3]. In NiTi the maximum strain of 4.1% was achieved by a deformation strain of 13.3% beyond the recoverable limit of NiTi (8%) [10]. Therefore, when TWSME is developed by the method of single-step deformation of martensite, the compromise between the two-way recoverable strain and the loss of inherent hot shape should be considered.

Another treatment to get TWSME is thermomechanical cycling. The process is a simple repetition of shape memory cycles between austenite and preferentially shaped martensite until the two-way behavior is developed. There are several basic thermomechanical cycling training methods, known as shape memory cycling (cooling below $M_f$ – deforming – heating above $A_f$, and the repetition of all steps), constrained cycling of deformed martensite (cooling below $M_f$ – deformation and constraining – heating above $A_f$, cycling from above $A_f$ to below $M_f$), pseudoelastic cycling (loading and
unloading the parent phase above $A_f$ under no slip condition) and combined shape memory and pseudoelastic cycling.

The reheat treatment is a procedure of adding another heat treatment on the previous treatment, but with the sample fixed in a different shape from that used in the first treatment. The method results in the memory of a cold shape close to the original shape defined in the first heat treatment and the memory of a hot shape is close to the new shape defined in the second heat treatment.

Combinations of different methods are currently used to save time of training and improve the fatigue life of the sample.

Although TWSME is an attractive technology, the fact that the force for the shape change in the forward transformation is extremely small compared to that in the reverse transformation should be considered [11].

2.1.5. Superelastic Effect

In contrast to the SME, where the driving force is thermal, superelastic effect (or superelasticity, SE) is driven by mechanical force. SE in SMAs occurs when a material is deformed above $A_f$ but below the temperature $M_d$ at which critical stress to induce martensite is greater than stress required to move the dislocations [3]. In the $(A_f - M_d)$ temperature range the martensite is not stable, but applied stress can stabilize it. When the stress is removed, the martensite becomes unstable again, and therefore, the material reverts to the austenite shape. Schematic relationships between SME, SE, and slip in stress-strain space for austenite and martensite are shown in Figure 1.
The mobile twin boundaries in the martensitic structure can easily move under applied stress. It determines the low yield strength of the martensite compared to austenite. However, only a certain amount of the martensitic deformation can be accommodated by this movement (twinning stress-plateau region in martensite curve in Figure 1). When the martensite achieves a certain amount of this twin movement it will again deform elastically and then yield plastically.

2.1.6. Relationship between Superelasticity and Shape Memory Effect

Both SME and SE can be observed in the same material, depending on the test temperature. Figure 2 shows the relationships among SME, SE, and slip in stress-temperature space.
SME occurs below $A_s$, followed by heating above $A_f$, and as long as critical stress for slip is high, SE occurs above $A_f$, where martensite is unstable in the absence of stress. In the interval $(A_s - A_f)$ both phenomena occur partially [4].

2.2. SMA heat engines

SMA heat engines are based on the principle that SMAs can generate force under applied temperature change. The SMA engines are very promising candidates for thermal management applications since they were first introduced in 1973 by Ridgway Banks [12], but yet to be implemented on a large scale due to the lack of a systematic approach in the comprehensive analysis of the engine incorporating materials, mechanical, and heat transferring aspects. An SMA element is the key component in such engines and plays the role of a solid-state version of the internal combustion engine with a piston and a cylinder. A stress difference between the high and low temperature states of the SMA element generates engine power [13].
The SMA heat engines can be categorized into five main types based on the principle of operation: 1) pulley based engines; 2) field engines; 3) crank engines; 4) reciprocating engines; 5) miscellaneous engines [14].

2.2.1. Pulley engines

Pulley engines operate based on friction between the pulley and the continuous element fixed in a loop around them. Current engine designs are susceptible to losing power when the loops slip on the pulleys. Continuous SMA elements such as belts or wires are used in the SMA pulley engines as driving mechanisms.

The pulley engine category can be subcategorized into engines of synchronized and unsynchronized designs. The main difference between the two subcategories is the dependent and independent rotation of the pulleys.

In the synchronized design, the pulleys rotate in a fixed agreement. The synchronized pulley motion is achieved by using timing chains, belts, gears, and other similar components. Example of the synchronized pulley engine with timing belt is shown in Figure 3.

In the unsynchronized pulley engine design, the pulleys rotate independently of one another. Example of the engine with pulleys of different diameters is shown in Figure 4.
The only link between the two pulleys is the SMA belt loop. The engines of this subcategory generally operate in bending, extracting work when a mechanically bent wire straightens upon heating.

![Diagram of Unsynchronized SMA pulley engine](image)

Figure 4. Unsynchronized SMA pulley engine [16]

### 2.2.2. Field engines

Engines of the field category operate against a recovery force such as a magnetic or gravitational field. For example, in the case of a gravitational field, when the SMA spring is heated to a critical transformation temperature, it contracts and increases the mass distribution density in the affected section. This causes a change in net distribution of gravitational force acting on the element and in the field engine design depicted in Figure 5 it forces the engine to rotate.
2.2.3. Crank engines

Crank engines convert the reciprocating linear motion of an SMA actuator into rotational motion by eccentrically connecting the actuating element to the output shaft. Crank engines can be performed mainly in two different configurations: axis center and offset center.

The axis center type engine is shown in Figure 6. The engine consists of two crankshafts and several SMA springs attached to crankshafts at different locations. Synchronizing gears ensure engine rotation in one direction. As the lengths of the springs are less than the distance between two crankshafts, the springs are in an extended state. When the SMA spring is in contact with a heat source it contracts due to the phase transformation and activates rotation of the shaft.
Figure 6. Axis center crank SMA engine [13]

An example of the offset center engine is shown in Figure 7. Ridgway M Bank invented the current design in 1973. The engine has a flywheel with 20 NiTi wire loops attached to the spoke of an offset wheel hub. The eccentricity causes the wires to be stretched more on one side of the wheel. When the wires enter hot water, they contract and straighten. Contact with cold water reverts the wires to their original shape. This makes the wheel rotate continuously.

Figure 7. Offset center crank SMA engine [12]
The general advantage of the offset crank engines in comparison with the pulley design is in the use of less space and material.

2.2.4. Reciprocating engines

Engines of reciprocating design operate linearly in a back-and-forth fashion. The biggest challenge with reciprocating engines is thermal cycling of the SMA element without physically moving it between the heat source and heat sink. Often external pumps are used to circulate hot and cold fluids over the wires. The engine presented in Figure 8 has two SMA wires attached to a lever arm at the top and fixed at the bottom. When one of the SMA elements is transferred to the austenite phase, it contracts. Another element at this moment is extended in the martensite condition. By switching hot and cold temperatures the engine produces reciprocating motion.

![Figure 8. SMA reciprocating engine [18]](image)

2.2.5. Miscellaneous engines

Engines of any other designs which do not fall into category 2.2.1 through category 2.2.4 categories, are referred to as miscellaneous engine design. This includes a swashplate, sequential, “Novel”, and many others.
2.3. Heating and cooling methods

Although operating temperature range is a main factor in the selection of alloys for an SMA heat engine, heat exchange mechanisms will affect engine operation and design. Different mediums and methods for cooling and heating SMA elements can be used.

2.3.1. Thermal radiation, gas/air, cold of space

Gas and air can be used for both heating and cooling. When the element is cooled with air of ambient temperature, a blowing fan can be employed to increase the cooling rate. There is an example [19] where exhaust hot gas from a vehicle and surrounding cold air were used as a heat source and heat sink respectively.

Heat from the sun or a laser can also be used to activate the reverse martensitic transformation in an SMA element while the cold of the space can be used for cooling (forward transformation). The advantage of using heat radiation and the cold of the space is the absence of drag.

2.3.2. Hot/cold water

Thermal conductivity of water is more than twenty times higher than of the air, even though both depend on temperature [20]. Therefore, transformation in water in contrast to air will require less time.
CHAPTER THREE: PRINCIPAL OF SMA STIRLING ENGINE

The chapter considers the stress state of the ordinary belt drive together with the temperature dependent transformation stress of the SMA wire. The entire load path for the SMA heat engine is described and the total stress state of the SMA wire from an unsynchronized engine pulley configuration is presented.

One type of heat engine is a Stirling engine which employs four steps in its reversible cycle: two steps are of constant volume (or constant length) of a working medium such as gas during which the pressure (or force) of the gas is changed as a result of an external change in temperature [21].

The original configuration of the Stirling engine consists of a gas as a working medium. During the cycle, a gas is first compressed at low temperature in Step I. Then by raising the temperature to a higher value $T_H$, in Step II, the force is increased at constant volume. In Step III the gas is allowed to expand at $T_H$, and finally, in Step IV, the temperature is lowered to $T_L$ with a decrease in pressure [21]. The full cycle of the Stirling engine is shown in Figure 9.

Figure 9. Work cycle diagram of Stirling engine
When an SMA element is used as a force generating element in a Stirling engine, the ability of the engine to perform a substantial amount of work is based on the intrinsic property of the SMA element to recover stress under restrained conditions. The main difference in operation between an SMA Stirling engine and a conventional engine is the direction of response of the working material. Instead of a hydrostatically applied pressure on a gas, in Step 1 a load is applied to the SMA element in the martensite state (at low temperature $T_L$) which causes an increase in length due to detwinning of the martensite. An increase in temperature in Step 2 up to high temperature $T_H$ causes the martensite to begin to revert to the austenite phase. The accomplished transformation results in a decrease in both force and length in Step 3. Step 4 reverts austenite to martensite and then the cycle repeats again.

![Figure 10. Work cycle diagram of SMA Stirling engine](image)

The forward and reverse transformations can produce relatively large changes in combinations of length and force as the temperature changes, thus allowing the SMA element to perform work [21].
3.1. Principal design and operation

Commercially available an SMA Stirling engine with an unsynchronized configuration presented in Figure 11 is considered in the framework for designing and implementing an SMA heat engine. It consists of two pulleys of different radii (radius of “hot” pulley $R_h$, and radius of “cold” pulley $R_c$) and a wire loop made of NiTi. The wire has been shape set to a straight shape and then fixed in a loop.

![Unsynchronized pulley type SMA Stirling engine](image)

Figure 11. Unsynchronized pulley type SMA Stirling engine

In the current engine configuration, forces are transmitted from the input pulley (hot) to the wire and then from the wire to the output pulley (cold). Initially the wire is in its initial pre-deformed condition which is slightly above the twinning stress-plateau region in the martensite stress–strain curve as shown in Figure 12. Note that the initial temperature of the SMA wire is below $M_f$. 


In the static state (no rotation, constant temperature) only a preload $F_{init}$ and bending force acting from the pulleys exist in the wire. $F_{init}$ is constant along the wire and in the static state results in the tensile stress $\sigma_{init}$:

$$\sigma_{init} = \frac{4 F_{init}}{\pi d^2} \tag{1}$$

The macroscopic strains in the SMA wire due to bending are considered as a function of pulley radii and wire diameter. Figure 13 shows the bending stress distribution across the wire diameter. The compressive strain inside the wire is equal in magnitude to the tensile strain. Note the neutral axis running in-between the outer and inner diameters.
The bending strain in the wire around the hot pulley $\varepsilon_h$ is higher than the strain around the cold pulley $\varepsilon_c$ and is given by:

$$
\varepsilon_h = \frac{d}{2R_h + d} \quad (2)
$$

$$
\varepsilon_c = \frac{d}{2R_c + d} \quad (3)
$$

Therefore, the recovery stress for $\varepsilon_h$ is larger than $\varepsilon_c$ even at the same wire temperature.

The strains lead to tensile stresses in the outer and to compressive stresses in the inner wire diameter. The maximum bending stress at the outer (maximum tension) and inner (maximum compression) edges of the wire are:

$$
\sigma_{bh} = E_b \varepsilon_h = E_b \cdot \frac{d}{2R_h + d} \quad (4)
$$

$$
\sigma_{bc} = E_b \varepsilon_c = E_b \cdot \frac{d}{2R_c + d} \quad (5)
$$

where $\sigma_{bh}$, $\sigma_{bc}$, and $E_b$ are maximum bending stresses in the wire at the hot and cold pulleys and flexural or bending modulus of the wire respectively.

In the operating mode, circumferential force $F_c$ is generated at the hot pulley by the input torque $T_h$ or the generated power $P_h$ and the angular speed of the pulley $\omega_h$ as well as its radius $R_h$:

$$
F_c = \frac{T_h}{R_h} = \frac{P_h}{\omega_h R_h} \quad (6)
$$

This circumferential force $F_c$ results in different forces on two wire sections and forms a tight and a slack wire side with acting forces $F_S$ and $F_T$ in them respectively.
Associated tension stresses in tight and slack sides respectively:

\[
\sigma_T = \frac{4 \cdot F_T}{\pi d^2} \tag{7}
\]

\[
\sigma_S = \frac{4 \cdot F_S}{\pi d^2} \tag{8}
\]

As can be seen from Figure 14, the difference between the tight side force \( F_T \) and the slack side force \( F_S \) corresponds to the transmitting circumferential force \( F_c \):

\[
F_c = F_T - F_S \tag{9}
\]

Note that Figure 14 is not to scale.

At the cold pulley, the circumferential force generates torque \( T_c \) with respect to the cold pulley radius:

\[
T_c = F_c \cdot R_c \tag{10}
\]
The circumferential force $F_c$ is transmitted at the interface between the wire and the pulleys by frictional forces. Therefore, frictional forces should be sufficiently high to ensure that the circumferential force can be transmitted. If the frictional force is lower than the circumferential transmitting force, then “adhesion” is not able to transfer the circumferential force from the hot pulley to the wire or from the wire to the cold pulley. If the wire slips the full circumferential force cannot be transmitted.

The final tension in both the slack side and the tight side of the wire depends on its initial tension. The circumferential force $F_c$ associated with $T_h$, results in the wire force in the tight side to increase to $F_T$ and the slack side force to decrease by the same extent to $F_S$:

$$F_T = F_{init} + \frac{F_c}{2}$$
$$F_S = F_{init} - \frac{F_c}{2}$$

If the engine operates at low speed, the forces at constant torque $T_h$ can be considered to be independent of $\omega_h$. However, at high speeds, centrifugal forces act on the wire section which run around the pulleys. These centrifugal forces $\Delta F$ are trying to pull the wire outwards and lift it off the pulley. It causes a reduction in the frictional force between the wire and pulley and ultimately can cause slippage.
Centrifugal tension in the wire due to centrifugal forces is given by:

$$\sigma_{cf} = \frac{4 \cdot \Delta F}{\pi d^2}$$  \hspace{1cm} (13)

Centrifugal forces are only generated in the regions where the wire bends around the pulleys and act radially away from the pulleys as shown in Figure 15. However, due to phase transformation and the resulting different stiffnesses of the hot and cold wire sections, there is observable deflection in the slack side of the wire.

In the case of the described pulley engine configuration, only tensile stresses will be crucial as these have a pulling effect in addition to the circumferential forces and the centrifugal forces. Therefore, the wire is subjected to maximum tensile stress in the outer diameter. The resulting stresses acting on the wire are presented in Figure 16 and the
maximum possible stress $\sigma_{max}$ is calculated as a sum of all acting stresses in accordance with Equation 14.

$$\sigma_{max} = \sigma_{init} + \sigma_b + \sigma_{cf} \quad (14)$$

When the wire exchanges heat with the source, the temperature of which is above $A_f$, the transformation from martensite to austenite is initiated which causes the SMA to recover strain under a constraint. The stiffness of the SMA increases and it returns to the shape set straight shape. In the process as the wire is mechanically bent around the pulleys, it generates torque $T_h$ to rotate the hot pulley. When the wire passes the heat source it cools down in a heat sink zone to the temperature $M_f$ or below it. As the wire cools down the
stress relaxes. However, the stress in the wire due to the pulley mechanics remains. This cycle repeats to result in an engine.
The engine operating between hot water and air of ambient temperature is shown in Video 1.

Video 1. Operation of a commercial SMA Stirling engine with an unsynchronized configuration
The difference in temperature between the cold and hot sections of the SMA wire introduces differences in wire tension, which in turn creates a moment $T_h$ and rotates the pulley. Consider one revolution of the cold pulley.

![Diagram of wire tension and pulley rotation](image)

Figure 17. Wire temperature prior to heat source entrance and after exit

Then in accordance with Figure 17, the torque $T_h$ generated by the hot pulley per revolution of the cold pulley is:

$$T_h = i (M_{exit} - M_{entr}) = \frac{R_c}{R_h} (M_{exit} - M_{entr}) \quad (15)$$

where $i = \frac{\omega_h}{\omega_c} = \frac{R_c}{R_h}$ is the transmission ratio, $\omega_h, \omega_c, M_{entr}$ and $M_{exit}$ are angular speeds of the hot and the cold pulleys, bending moments acting over the wire cross section in the positions prior to entering the heat source and upon exiting, respectively.

The moments $M_{entr}$ and $M_{exit}$ act over the cross section of the SMA element and can be expressed by using a section modulus $W_p$ and bending stresses, $\sigma_{entr}$ and $\sigma_{exit}$ at the point of entry and exit, respectively:

$$M_{entr} = W_p \sigma_{entr} \quad (16)$$

$$M_{exit} = W_p \sigma_{exit} \quad (17)$$
where the section modulus $W_{p} = \frac{\pi d^3}{32}$

Then

$$T_h = \frac{R_c \pi d^3}{R_h \frac{32}{32}} (\sigma_{exit} - \sigma_{entr}) \quad (18)$$

Engine output power results from the recovery stress of the bent wire around the hot pulley. The biasing work in the current configuration is the work done to bend the wire over the cold pulley as well as the work to bend the wire over the hot pulley when the wire is in the martensitic state (before transformation).

Consider the torque on the cold pulley per one revolution in accordance with Figure 18:

$$T_c = M_{c,h} - M_{c,c} = \frac{\pi d^3}{32} (\sigma_{c,h} - \sigma_{c,c}) \quad (19)$$

![Figure 18. Wire temperature at the point of first contact of the wire with the cold pulley (hot section of the wire) and the last point of contact (cold section)](image-url)
where $M_{c,h}, M_{c,c}$ and $\sigma_{c,h}, \sigma_{c,c}$ are moments and stresses at points of the first contact of the wire with the cold pulley (hot section of the wire) and the last contact (cold section), respectively.

Then the output torque $T_{out}$ of the engine per revolution of the cold pulley is expressed as follows:

$$T_{out} = R_c \frac{R_h}{R_h} (M_{exit} - M_{entr}) - (M_{c,h} - M_{c,c}) = \frac{\pi d^3}{32} \left( \frac{R_c}{R_h} (\sigma_{exit} - \sigma_{entr}) - (\sigma_{c,h} - \sigma_{c,c}) \right)$$

(20)

For maximum torque output, it is efficient to heat part of the wire which is in contact with the pulley of smaller diameter, as the smaller pulley yields larger macroscopic wire strain.
CHAPTER FOUR: THERMAL AND MECHANICAL ANALYSES

Mechanical and thermal aspects are closely related and therefore are considered within one chapter. Mechanical aspects are addressed by performing force balances in the SMA element and focused on the resulting stress distribution and design features. Thermal aspects are addressed by considering the temperature distribution and the heat transfer rate between the SMA element and the heat source or the heat sink.

4.1. Mechanical analyses

The resulting stress distribution in an SMA element is a key principle which makes the operation of SMA engines possible. The stress state in the SMA wire of the unsynchronized pulley heat engine is described in Chapter 3. In addition to the stress distribution, mechanical aspects also include design specifics such as SMA element shape and dimensions, number of SMA elements, and external friction.

4.1.1. SMA element shape and dimensions

The SMA elements can be of any form such as springs, wires, beams, rings, belts etc. In addition to static and dynamic strength requirements from Chapter 3, the length and cross section of the SMA element can be estimated based on the required engine output and the overall allowable dimensions.

4.1.1.1. Distance between pulleys and wire length

The distance between pulleys $D$ in the configuration presented in Figure 11 sets the engine size and the wire length. Based only on geometrical parameters of the engine, the initial total length of the wire $L$ is equal to:
\[ L = \pi (R_c + R_h) + 2 \sqrt{D^2 + (R_c - R_h)^2} \]  

(21)

The final total length should be derived from design criteria in Equation 21 and heat transfer conditions in such a way, that after leaving the heat source zone, the heated segment of the SMA wire (at temperature \( \geq A_f \)) should have enough time to be cooled down to temperature \( M_f \) in the heat sink before starting the next heating cycle.

4.1.1.2. Wire diameter

SMA elements of larger diameter produce more force but take longer time to transfer heat (mechanical heat transfer). For example, an engine using a wire of larger diameter will generate larger torque in accordance with Equation 20 but will run slower than the same engine using a thinner wire because the thinner elements ensure faster transformation.

Angular speed (or revolutions per minute (RPM)) of the driving pulley (hot) can be estimated from Figure 17 as:

\[ \omega_h = \frac{\varphi}{t} \]  

(22)

where \( \varphi \) and \( t \) are the angle, corresponding to the contact area of the wire with the heat source and the required time for the point on the wire to travel through the distance associated with \( \varphi \), respectively.

The heat stored by the wire \( Q_{\text{wire}} \) is assumed to equal the heat received from the source \( Q_{\text{in}} \):

\[ Q_{\text{wire}} = Q_{\text{in}} \]  

(23)
Then, by using Newton’s law of heating/cooling, the heat transfer rate between the SMA element and the heat source or the heat sink \[20\] is:

\[
\frac{dT}{d\tau} = \frac{h}{\rho c_p} \cdot \frac{A}{V} \Delta T
\]  

(24)

where \( h, A, V, \rho, c_p, d\tau, \) and \( \Delta T \) are the heat-transfer coefficient, characteristic area and volume, density, specific heat, heat transfer time, and the characteristic temperature difference, respectively. It should be noted that \( h \) cannot be defined until \( A, V \) and \( \Delta T \) are stipulated.

Expressing \( A \) and \( V \) through the wire diameter \( d \) and length \( l_h \) of the element exposed to heating gives:

\[
\frac{dT}{d\tau} = \frac{h}{\rho c_p} \cdot \frac{\pi d l_h}{\frac{\pi d^2}{4} l_h} \cdot \Delta T = \frac{h}{\rho c_p} \cdot \frac{4}{d} \Delta T
\]  

(25)

Applying Equation 25 to calculate the heat transfer rate between the SMA element and the heat source (refer to Figure 17) results in:

\[
\frac{dT}{d\tau} = \frac{h_1}{\rho c_p} \cdot \frac{4}{d} (T_{source} - T_{wire})
\]  

(26)

where \( T_{source} \) and \( T_{wire} \) are the temperature of the heat source and the temperature of the wire in heat source for time \( d\tau \), respectively, \( h_1 \) is the heat-transfer coefficient for area \( A_h \) of the wire exposed to the heating and temperature difference \( (T_{source} - T_{wire}) \).

Rearranging Equation 20 for time required to heat \( d\tau \) gives:

\[
d\tau = \frac{\rho c_p d}{4h_1} \cdot \frac{dT}{(T_{source} - T_{wire})}
\]  

(27)
Considering the heating of the wire from the initial temperature \( T_{\text{entr}} = M_f \) (prior to entering the heat source) to the transformation temperature \( A_f \) in the heat source:

\[
\tau = \frac{\rho \cdot c_p \cdot d}{4h_1} \int_{M_f}^{A_f} \frac{dT}{(T_{\text{source}} - T_{\text{wire}})} = -\frac{\rho \cdot c_p \cdot d}{4h_1} \ln \left( \frac{T_{\text{source}} - A_f}{T_{\text{source}} - M_f} \right) \quad (28)
\]

Therefore, the maximum angular speed of the hot pulley can be found by substituting Equation 28 into Equation 22 under the assumption that \( t = \tau \):

\[
\omega_{\text{max}} = -\frac{4 \cdot h_1 \cdot \varphi}{\rho \cdot c_p \cdot d \cdot \ln \left( \frac{T_{\text{source}} - A_f}{T_{\text{source}} - M_f} \right)} \quad (29)
\]

With each revolution the wire experiences mechanical (bending around the pulley), thermal \((A_f - M_f)\), and transformational \((M-A)\) deformations. The higher the speed of engine operation and the smaller the pulleys, the shorter engine service life due to fatigue can be expected.

Similar calculations for the heat transfer rate and time between the SMA element and the heat sink give:

\[
\frac{dT}{d\tau} = \frac{h_2}{\rho c_p} \cdot \frac{4}{d} \cdot (T_{\text{sink}} - T_{\text{wire}}) \quad (30)
\]

where \( T_{\text{sink}} \) and \( T_{\text{wire}} \) are the temperature of heat sink and the temperature of the wire in heat sink for the time \( d\tau \), respectively, \( h_2 \) is the heat-transfer coefficient for the postulated area of the wire exposed to cooling \( A_c \) and temperature difference \((T_{\text{sink}} - T_{\text{wire}})\).

Considering the temperature of the wire after exiting the heat source \( T_{\text{exit}} = A_f \) gives:
$$\tau = \frac{\rho c_p}{4h} \int_{A_f}^{M_f} \frac{dT}{(T_{sink} - T_{wire})} = -\frac{\rho c_p}{4h} \ln \left( \frac{T_{sink} - M_f}{T_{sink} - A_f} \right)$$

(31)

For the temperature range, which is of interest in the subject framework, the heat transfer coefficient can be considered constant: \( h_1 = h_2 = h \).
4.1.2. Number of SMA elements

The need for getting a high power-high torque engine may be achieved by the engine design with an increased number of SMA elements. The use of an engine with several SMA elements will proportionally increase the generated torque from a single element described by Equation 18 under assumption of infinite heat source.

4.1.3. External friction

As it was pointed out earlier, the circumferential force is transmitted by frictional forces between the pulleys and the wire. Therefore, frictional forces should be large enough to ensure the required “adhesion” while the generated torque should overcome this friction to guarantee the rotation of the engine. Therefore, frictional forces and coefficient of the external friction between the wire and the pulleys should be considered in engine design.

4.2. Thermal analyses

Thermal aspects consider heat transfer mechanisms, SMA thermal conductivity and the temperature difference between the heat source and the heat sink. The SMA wire temperature profile is derived as a function of the time and position to describe the change in the wire state.

4.2.1. Temperature profile

Since shape memory behavior is a function of temperature, it is convenient for the SMA heat engine design and output to predict the temperature distribution along the SMA element as a function of time $t$ and position $x$ under different thermal conditions.

In each cycle, the SMA element experiences two transformations: reverse and forward. The general equation for the conduction of heat in solids [20]:

\[ \ldots \]
\[ \rho \cdot c_p \frac{dT}{dt} = k \nabla^2 T = k \frac{d^2T}{dx^2} \]  

(32)

or rearranging for the change in temperature with time \( \frac{dT}{dt} \):

\[ \frac{dT}{dt} = k \frac{d^2T}{\rho \cdot c_p \cdot dx^2} \]  

(33)

where \( k \) is thermal conductivity.

Thermal conductivity of material reflects the relative ease or difficulty of heat transfer through material. It is an intrinsic property of the material and depends on bonding type and structure of an alloy as well as on the temperatures at which the material operates, i.e., temperatures of heat and source.

Considering that \( T = T(t, x) \), then:

\[ \frac{dT}{dt} = \frac{\partial T}{\partial t} + \frac{\partial T}{\partial x} \frac{\partial x}{\partial t} \]  

(34)

Equating Equation 27 with Equation 28 and rearranging it for \( \frac{\partial T}{\partial t} \):

\[ \frac{\partial T}{\partial t} = \frac{k}{\rho \cdot c_p} \frac{d^2T}{dx^2} = \frac{\partial T}{\partial x} v_x \]  

(35)

\[ \text{where } v_x = \frac{\partial x}{\partial t}. \]

4.2.2. Temperature difference between the heat source and the heat sink

Any sufficient temperature difference is thermal energy can drive an SMA Stirling engine. The input heat required for the transformation should be sufficient to heat the SMA from its initial temperature (\( T_{wire} \leq M_f \)) to the temperature \( A_f \) or slightly above it. The minimum amount of heat required for the transformation for SMAs in both loaded and unloaded conditions is estimated in Chapter 5. However, having a heat source at a
temperature much higher than $A_f$ will increase the heat transfer rate and will decrease the transformation time (Equation 28).

Consider two cases of the temperature difference between the SMA element and the heat source: $T_{source} \gg A_f$ and $T_{source} \geq A_f$.

4.2.2.1. Heat source temperature is much higher than austenite finish: $T_{source} \gg A_f$

In accordance with Equation 29 heating of the SMA wire at $T_{source} \gg A_f$, will drive the engine at higher speeds, and will lower its life of service due to fatigue. The influence of higher speed on service life is described in Section 4.1.1.2.

Moreover, the heat source ($T_{source} \gg A_f$) may result in extensive heating of the SMA (temperature of the wire in the position upon exiting the heat source is higher than $A_f$) and will cause the need for additional cooling to ensure the forward transformation from austenite to martensite (Equation 31). Additional cooling can be achieved by:

1) Increasing the distance, D, between the hot and cold pulleys
2) Using a heat sink for faster heat transfer. For example, thermal conductivity k of water is more than 20 times higher than of the air, even though both depends on temperature [20]. Therefore, cooling in water in contrast to cooling in air will require less time.

4.2.2.2. Heat source temperature is slightly above the austenite finish: $T_{source} \geq A_f$

In accordance with Equation 29, the low-temperature heat source requires longer time to heat up the wire, to transform the martensite to austenite. If the contact time between the wire and the heat source is not sufficient for the SMA element to accomplish the
reversible transformation, the material will need more time to absorb the heat required for the transformation. Therefore, the angular speed will decrease. In the case of $T_{source} > A_f$, the engine makes a self-correction by lowering the angular speed to increase the contact time with the heat source.
CHAPTER FIVE: THERMODYNAMIC ANALYSES

This chapter considers the intrinsic material properties of the SMA element including thermal conductivity, specific heat, nucleation mechanism, grain size, internal friction between interfaces during the reversible martensitic transformation and depends mainly on SMA composition and thermomechanical history. Aspects related to hysteresis are also considered.

The input heat \( Q_{in} \) should be adequate to drive the transformation. The required heat \( H_{in} \) can be estimated based on the material properties from:

\[
Q_{in} = Q_s + \Delta H_{net} + W_{SMA}
\]

(36)

where \( Q_s, \Delta H_{net} \) and \( W_{SMA} \) are the sensible heat, stress-free transformation enthalpy change, and SMA work output, respectively. Change in any of these terms will affect the total amount of the required heat.

5.1. Sensible heat

Sensible heat \( Q_s \) is the amount of heat required to change the temperature of the wire from the martensite finish temperature \( M_f \) to the austenite finish temperature \( A_f \). In addition to the temperature range, the amount of sensible heat depends on material density \( \rho \) and specific heat \( c_p \):

\[
Q_s = \rho \cdot c_p \cdot (A_{f}^{a_1} - M_{f}^{a_2})
\]

(37)

The difference in specific heat of austenite \( c_{pA} \) and martensite \( c_{pM} \) phases is negligible, therefore \( c_{pM} = c_{pA} = c_p \) is used in this work. \( A_{f}^{a_1} \) and \( M_{f}^{a_2} \) are the temperatures of the austenite finish transformation under applied stress \( \sigma_1 \) and the temperature of martensite finish transformation under applied stress \( \sigma_2 \), respectively.
Due to the thermoelastic nature of the martensitic transformation in SMAs, external stress results in an increase in the transformation temperatures. In accordance with the Clausius - Clapeyron equation all the transformation temperatures are affected by stress in the same way [3]:

$$\frac{d\sigma}{dT} = -\frac{\Delta H_{net}}{\varepsilon T}$$

(38)

where \(\sigma, T, \varepsilon, \) and \(\Delta H_{net}\) are the uniaxial stress, temperature, transformation strain, and enthalpy of the transformation (or latent heat), respectively. \(\frac{d\sigma}{dT}\) is called the stress rate.

Consider the \(A_f\) change due to the applied stress \(\sigma_1\):

$$\frac{d\sigma}{dT} = \frac{\sigma_1 - 0}{A_f^{\sigma_1} - A_f} = \frac{\Delta H_{net}}{\varepsilon T}$$

(39)

where \(A_f\) is the finish temperature of the austenite transformation under no applied stress.

From Equation 39:

$$\left(A_f^{\sigma_1} - A_f\right)\Delta H_{net} = \sigma_1\varepsilon T$$

(40)

Therefore, the austenite finish transformation temperature under applied stress \(\sigma_1\) can be given by:

$$A_f^{\sigma_1} = A_f + \frac{\sigma_1\varepsilon T}{\Delta H_{net}} = A_f + \frac{\sigma_1}{\frac{d\sigma}{dT}}$$

(41)

Similar calculations for the martensite finish transformation temperature under applied stress \(\sigma_2\) give:

$$M_f^{\sigma_2} = M_f + \frac{\sigma_2\varepsilon T}{\Delta H_{net}} = M_f + \frac{\sigma_2}{\frac{d\sigma}{dT}}$$

(42)

where \(M_f\) is the temperature of the martensite finish transformation under no applied stress.
Substituting Equation 41 and Equation 42 in Equation 37 gives:

\[ Q_s = \rho \cdot c_p \left( A_f^{e_f} - M_f^{e_f} \right) = \rho \cdot c_p \left( A_f + \frac{\sigma_1}{d\sigma/dT} - M_f + \frac{\sigma_2}{d\sigma/dT} \right) = \]

\[ = \rho \cdot c_p \left( A_f - M_f + \frac{\sigma_1}{d\sigma/dT} - \frac{\sigma_2}{d\sigma/dT} \right) \]

\[ = \rho \cdot c_p \left( A_f - M_f + \frac{\sigma_1 - \sigma_2}{d\sigma/dT} \right) \quad (43) \]

where \((A_f - M_f)\) is the difference between the stress-free austenite finish and martensite finish temperatures or the hysteresis in the transformation.

5.2. Stress-free transformation enthalpy change of the reverse endothermic transformation

The nonlinear temperature response of SMAs in the differential scanning calorimetry (DSC) curves during heating and cooling is a result of the phase transformation which corresponds to endothermic and exothermic reactions. Figure 19 shows the schematic of the DSC response of a NiTi-based shape memory alloy. Transformation upon heating is endothermic and the transformation upon cooling is exothermic. Therefore, change in the transformation enthalpy of the reverse endothermic transformation should be taken into account for estimation of the required heat \(Q_{in}\).
Figure 19. Schematic of the differential scanning calorimetry response of a NiTi-based shape memory alloy

For a thermoelastic transformation, the stress-free enthalpy change is estimated as follows [5]:

$$\Delta H_{net} = \Delta H_{ch} + \Delta H_{el}(T) + \Delta H_i(T)$$  \hspace{1cm} (44)$$

where $\Delta H_{ch}$, $\Delta H_{el}$, and $\Delta H_i$ are chemical, elastic, and internal interface enthalpies changes respectively.

Consider separately the contribution of each enthalpy change:

1) $\Delta H_{ch}$ is the main contribution to the overall energetics in Equation 44. It can be measured through a single interface experiment (no elastic energy stored as transformational shape change takes place against atmospheric air pressure), i.e., as it was obtained for a Cu - 14Al - 2.5Ni single crystal transformation through gradient cooling and heating in [5]. $\Delta H_{ch}$ is insensitive to the transformation temperature for a given composition.
2) $\Delta H_{el}$ is about an order of magnitude smaller than $\Delta H_{ch}$. In contrast to $\Delta H_{ch}$, $\Delta H_{el}$ introduces a temperature dependance to $\Delta H_{net}$, because, as was mentioned earlier, the stored elastic strain energy increases as the transformation proceeds on cooling. $\Delta H_{el}$ can be calculated independent from experimental data as following:

$$\Delta H_{el} = \Delta G_{el} + T\Delta S_{el}$$

(45)

where $\Delta G_{el}, \Delta S_{el}$ are the free elastic energy and entropy changes.

3) $\Delta H_i$ arises mainly from the twin interfaces in the martensitic phase. It does not create a significant contribution to $\Delta H_{net}$, however, together with $\Delta H_{el}, \Delta H_i$ increases as the transformation proceeds. The magnitude of $\Delta H_i$ is about two orders of magnitude smaller than $\Delta H_{ch}$. Approximately $\Delta H_i$ can be expressed as:

$$\Delta H_i = \frac{\gamma V_m}{s_t}$$

(46)

where $\gamma, V_m$ and $s_t$ are the twin-interfacial energy, molar volume, and twin spacing, respectively.

The elastic enthalpy and internal interface enthalpy contributions are opposite in sign to the chemical enthalpy change. Therefore, they cause a reduction in the total enthalpy change according to Equation 44. Another difference between the elastic and chemical terms is that the chemical driving force is constant throughout the whole chemically homogeneous specimen, while the elastic strain energy will be inhomogeneous in a partially transformed specimen.
5.3. SMA work output

When sensible heat is added to the system, the wire temperature changes from $M_f^{\sigma_2}$ to $A_f^{\sigma_1}$. The heat from $\Delta H_{\text{net}}$ promotes the absorption of latent heat by the alloy and initiates the transformation from martensite at $\sigma_2, \varepsilon_M$ to austenite at $\sigma_1, \varepsilon_M$. The material recovers its shape under constraint from $\varepsilon_M$ to $\varepsilon_A$. Cooling the material from $A_f^{\sigma_1}$ to $M_f^{\sigma_2}$ reverts material to the martensite state at $\varepsilon_A$. Later, due to the pre-deformed state induced by initial force $F_{\text{init}}$, martensite strains to $\varepsilon_M$. This process is schematically shown in Figure 20.

![Figure 20. Schematic of SMA work output during a thermal cycle, following Ref. [3, 22]](image)

It is important to highlight that SMAs produce work only through the reverse transformation from martensite to austenite and do not produce substantial work upon cooling [22]. The unit work output that a SMA is able to produce in one cycle can be estimated as the area between austenite and martensite curves in the stress-strain graph as is illustrated by the cross hatched area shown in Figure 20 and can be expressed as:
\[ W_{SMA} = \int \sigma d\varepsilon = \int_{\varepsilon^M}^{\varepsilon^A} \sigma_1 d\varepsilon - \int_{\varepsilon^M}^{\varepsilon^A} \sigma_2 d\varepsilon \approx \Delta \sigma \cdot \Delta \varepsilon \] (47)

From the analyses of Figure 19 and Equation 47, maximum material work output under \( \sigma_1 \) is equal to the stress slightly above the twinning stress-plateau of martensite and \( \sigma_2 \) is equal to the stress below that for plastic deformation in austenite.

While SMA elements with the two-way shape memory effect (TWSME) are able to loop the performance of the engine without additional components in the system, SMA with one-way memory elements require biasing components to cycle the work of the engine. When a bias component is present, part of the work done by the SMA is used to overcome the biasing force \( W_{bias} \). As it was pointed out in Chapter 3, in the specific example of the Stirling engine, both pulleys are biasing mechanisms.

By recovering strain against the biasing force introduced by the wire bending around the hot pulley, the SMA forces the pulley to rotate. The biasing force is introduced by bending the wire over the cold pulley. Therefore, the magnitude of the useful power generated by the SMA engine is the work done by the material to change the shape of the wire under applied stress and strain minus the bias work:

\[ W_{SMA} - W_{bias} = W_{out} \] (48)

5.4. Hysteresis

As it is seen from Equation 37 – Equation 38 and Figure 18 – Figure 19, all three components \( Q_s, \Delta H_{net} \), and \( W_{SMA} \) depend on the alloy hysteresis.

Hysteresis is a manifestation of the energy losses between the forward and reverse martensitic transformation. Macroscopically it can be observed by Differential Scanning Calorimetry (DSC) obtained under no applied stress in accordance with ASTM
F2082/F2082M-16 (Standard Test Method for Determination of Transformation Temperature of Nickel-Titanium Shape Memory Alloys by Bend and Free Recovery). The SMAs thermal response for the forward and reverse transformation occurs in different paths as shown in Figure 18.

To analyze energy losses associated with the hysteresis, it is more convenient to express the required heat with elastic and inelastic components:

\[ Q_{in}(T, \sigma/\varepsilon) = E_{el} + E_{inet} \] (49)

where \( E_{el} \) and \( E_{inet} \) are the elastic strain energy and inelastic energy terms, respectively.

Elastic strain energy arises around the martensite plates formed during the transformation and from the material thermal expansion. This contribution is fully recoverable. Inelastic energy is wasted heat which does not produce any useful work. In addition, inelastic energy may also include contributions from a residual martensite phase that nominally would have been expected to transform.

Since the martensitic transformation is associated with a shape change, when a martensite plate forms in the parent phase (forward, A → M transformation), a large strain arises around the martensite particle. In the case of a thermoelastic transformation, when the transformational shape change is accommodated fully elastically, the build-up of elastic strain energy opposes further growth of a martensite plate, which was initially driven by the chemical free-energy change [23].

However, the shape change may not be fully accommodated elastically. If during the transformation appreciable plastic accommodation by slip takes place, the relaxation of the elastic strain field around the particle will destroy the ideally thermodynamic restoring force responsible for the small transformational hysteresis [5]. To the degree that plastic
accommodation takes place, the corresponding decrease in stored elastic energy will require higher energy required to reverse the transformation in SMA with reduced stored elastic energy. Moreover, the concomitant reduction in elastic energy will affect $\Delta H_{net}$. Regardless of the amount of stored elastic energy, the hysteresis always exists because the main contribution to the hysteresis comes from frictional resistance to the interfacial motion between phases and the compatibility at the austenite/martensite interface. In addition, plastic accommodation of transformation strain will cause a higher dislocation density in the matrix which in a non-cycled sample increases the interfacial friction stress and therefore contributes an additional irreversible component to the overall energetics.

Both, considerable frictional resistance during the forward and reverse transformations and amount of stored elastic energy will introduce a deviation from thermoeelastic equilibrium [5].
CHAPTER SIX: MATERIAL MICROSTRUCTURE

Since the SMA phase transformation in general is thermoelastic, internal and external stresses can significantly affect the behavior of these alloys and change the energetics. Therefore, grain size, nucleation mechanism, martensite texture, dislocations, training, and precipitates in austenite matrix may significantly affect $\Delta H_{\text{net}}$.

6.1. Parameters affecting the internal stress state of the SMA

According to Ref. [24], the activation energy for homogeneous nucleation of a single variant martensitic is very high ($\sim 10^5 k_B T$) to be overcome by a moderate amount of thermal energy. However, defects such as grain boundaries and dislocations, or precipitates increase the internal stress state of SMAs and allow the nucleation of martensite at relatively low thermal fluctuations or even introduce a multi-stage martensitic transformation. It affects overall equilibrium, hysteresis width, shape of DSC curve ($\Delta H_{\text{net}}$) and other parameters. In changes of the internal stress state of the material with dislocations or by inducing finer microstructure, only mechanical energy is involved, while stress changes associated with precipitates also includes chemical energy change.

6.1.1. Introduction of elastic constraints by grain boundaries

Ref. [5] compares transformations of Cu - 14Al - 2.5Ni with and without the elastic strain energy involved. In single crystal and single interface transformations which takes place against the atmosphere, there is no stored elastic strain. As it is schematically shown in Figure 21 (a), the transformation results in a sharp recovery profile occurring at temperatures: $M_s = M_f$ and $A_s = A_f$. $T_o$ is the temperature at which austenite and martensite have the same amount of chemical energy.
However, in the general case, $\Delta H_{\text{net}}$ is less negative than $\Delta H_{\text{ch}}$ in a thermoelastic martensitic transformation, mainly because of the elastic energy. Introduction of elastic constraints such as multiple interfaces and/or grains, tilts the shape of the hysteresis by unevenly lowering transformation temperatures relative to the temperatures of the transformation with no elastic strain energy involved.

In the single crystal and multiple interface case, as Figure 21 (b) schematically illustrates, $M_s$ and $A_f$ are still close to the transformation temperatures of single crystal and single interface case while significant changes in lowering $M_f$ and $A_s$ are observed. The reason for this is that the elastic strain energy is not generated when the first martensitic particle forms at $M_s$ because the nucleation takes place at a free edge of the crystal. The
similar geometric situation holds true at \( A_f \). The trend of lowering transformational temperatures with introduction of elastic constraints is reinforced even more clearly by the results which were obtained in Ref. [5] for polycrystalline materials of different grain diameters.

The higher amount of stored elastic energy is involved in transformation the smaller associated \( \Delta H_{net} \) in accordance with Equation 44. Therefore, less energy can be absorbed/released by the material and, consequently, smaller temperature suppression during the transformation can be achieved for one cycle (or one revolution of hot pulley for the engine of the subject framework). In general, a downward trend in the transformation temperature range with decreasing grain size, or which is equal to with increase in elastic energy, exists.

6.1.2. Nucleation mechanisms

In contrast to surface nucleation, when the transformation takes place in a bulk material, elastic strain energy is operative even at \( M_s \) and \( A_f \). In Ref. [5] to initiate bulk nucleation, a very thin plating of nickel 25 \( \mu \text{m} \) was deposited on the samples. Nickel plating inhibits surface nucleation and causes supercooling to lower temperatures for bulk nucleation due to increased stored elastic energy. The schematic change in shape of the heat-evolution curve for unplated single interface (zero elastic strain) and the same sample but Ni-plated (bulk nucleation, stored elastic energy) is shown in Figure 22. Figure 22 (a) illustrates a very narrow shape of DSC curve with a very sharp pick which are typical for transformation with no stored elastic strain energy. The broader shape of the DSC curve and blunt pick in Figure 22 (b) corresponds to transformation with stored elastic strain energy.
energy. It should be noted that the area under the curve in (a) is larger than area in (b) in accordance with Equation 44. Moreover, as previously discussed, $M_s^{(a)}$ is higher than $M_s^{(b)}$.

![Figure 22. Schematic of DSC curves for (a) transformation without stored elastic strain energy (b) transformation with stored elastic strain energy following Ref. [5]](image)

6.1.3. Martensite variants

The lattice correspondence between the austenite and martensite lattices is a unique relationship. In NiTi the lattice correspondence between B2 austenite (subscript “A” in Equation 50, Table 1 and Table 2) and B19’ martensite (subscript “M” in Equation 50 and Table 1) phases has been established to be of the type [25]:

$$
(001)_M // (011)_A, [\bar{1}10]_M // [\bar{1}11]_A.
$$

(50)

From symmetry Equation 50 can be accomplished by a total of twelve crystallographically different correspondent martensite variants as listed in Table 1. All these correspondence-variants yield an identical B19’ structure, but they are different in crystallographic orientation.
Table 1
Lattice correspondence between martensitic variants and austenitic phase, [25]

<table>
<thead>
<tr>
<th>Variant</th>
<th>$[100]_M$</th>
<th>$[010]_M$</th>
<th>$[001]_M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$[011]_A$</td>
<td>$[100]_A$</td>
<td>$[011]_A$</td>
</tr>
<tr>
<td>2</td>
<td>$[011]_A$</td>
<td>$[100]_A$</td>
<td>$[011]_A$</td>
</tr>
<tr>
<td>3</td>
<td>$[011]_A$</td>
<td>$[100]_A$</td>
<td>$[011]_A$</td>
</tr>
<tr>
<td>4</td>
<td>$[011]_A$</td>
<td>$[100]_A$</td>
<td>$[011]_A$</td>
</tr>
<tr>
<td>5</td>
<td>$[101]_A$</td>
<td>$[010]_A$</td>
<td>$[101]_A$</td>
</tr>
<tr>
<td>6</td>
<td>$[101]_A$</td>
<td>$[010]_A$</td>
<td>$[101]_A$</td>
</tr>
<tr>
<td>7</td>
<td>$[101]_A$</td>
<td>$[010]_A$</td>
<td>$[101]_A$</td>
</tr>
<tr>
<td>8</td>
<td>$[101]_A$</td>
<td>$[010]_A$</td>
<td>$[101]_A$</td>
</tr>
<tr>
<td>9</td>
<td>$[110]_A$</td>
<td>$[001]_A$</td>
<td>$[110]_A$</td>
</tr>
<tr>
<td>10</td>
<td>$[110]_A$</td>
<td>$[001]_A$</td>
<td>$[110]_A$</td>
</tr>
<tr>
<td>11</td>
<td>$[110]_A$</td>
<td>$[001]_A$</td>
<td>$[110]_A$</td>
</tr>
<tr>
<td>12</td>
<td>$[110]_A$</td>
<td>$[001]_A$</td>
<td>$[110]_A$</td>
</tr>
</tbody>
</table>

Since thermoelastic martensitic transformation is associated with a shape change, a large strain arises around the martensite plate in the parent matrix. Therefore, the formation energy of a single-variant martensite could be not favorable due to the high amount of elastic energy involved, and for that reason, self-accommodated multi-variant martensite with lower activation energy barriers is energetically favorable and usually formed in temperature-induced transformation [7]. Self-accommodation is achieved by combining four kinds of variant clusters: A, B, C, and D as schematically shown in Figure 23. Shape strains of the plates compensate each other providing a zero microscopic shape change.
The shape memory response of SMA under an applied load is based on reorientation of the martensite structure towards a more favorable configuration of correspondence-variants with respect to the stress. Under unloading, reverse reorientation of the martensite does not appear. However, on heating by mutual conversion among the correspondence-variants a reverse transformation from reoriented martensite to the austenite of the original structure takes place. The main mechanism of the variants’ conversion is deformation twinning. Two types of twinning in martensitic NiTi have been reported: Type I twinning, with a mirror reflection about the twinning plane \( K1 = (11\bar{1}) \) and Type II twinning with a twinning plane \( K1 = (0.7205 \ 1 \ \bar{1}) \), leading to a rotation by \( 180^\circ \) around the shear direction \( \eta = [0\bar{1}1] \) [26].

The transformation strains \( \varepsilon_i \) from austenite to a specific martensite variant for all twelve variants (i=1..12), can be mathematically described with transformation matrix \( U_1 \). The matrix brings each austenitic lattice vector \( e_i^a \) to martensitic lattice vector and is described in detail in Section 6.1.7. For all twelve variants along a \( (hkl) \) crystallographic directions in parent basis, \( \varepsilon_i(hkl) \) is expressed by a given Equation 51 and presented:
\[ \epsilon_i(hkl) = \frac{|U_1(e_i^A)|}{|e_i^A|} - 1 \]  

(51)

In SIM transformation the maximum strains under compression (-) and tension (+), \( \pm \epsilon_i(hkl) \), along [111]_A, [110]_A, and [100]_A directions were calculated in Ref. [25] and are listed in Table 2. The dependance of strain from loading direction in polycrystalline SMA bulk was obtained by averaging over all orientations which is in accordance with the crystallographic theory based on selective formation of different SIM variants under loading. Under assumption of random grain orientation in a polycrystalline sample without any preferred grain orientation, the tension strain is approximately 50% higher than compression strain.
Table 2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-3.7</td>
<td>-5.2</td>
<td>-4.2</td>
</tr>
<tr>
<td>2</td>
<td>-1.4</td>
<td>-5.2</td>
<td>-4.2</td>
</tr>
<tr>
<td>3</td>
<td>-3.7</td>
<td>+3.6</td>
<td>-4.2</td>
</tr>
<tr>
<td>4</td>
<td>+9.7</td>
<td>+3.6</td>
<td>-4.2</td>
</tr>
<tr>
<td>5</td>
<td>-3.7</td>
<td>+3.6</td>
<td>+2.6</td>
</tr>
<tr>
<td>6</td>
<td>+9.7</td>
<td>+3.6</td>
<td>+2.6</td>
</tr>
<tr>
<td>7</td>
<td>-3.7</td>
<td>-5.2</td>
<td>+2.6</td>
</tr>
<tr>
<td>8</td>
<td>-1.4</td>
<td>-5.2</td>
<td>+2.6</td>
</tr>
<tr>
<td>9</td>
<td>-3.7</td>
<td>-3.4</td>
<td>+2.6</td>
</tr>
<tr>
<td>10</td>
<td>-1.4</td>
<td>+8.4</td>
<td>+2.6</td>
</tr>
<tr>
<td>11</td>
<td>-3.7</td>
<td>-3.4</td>
<td>+2.6</td>
</tr>
<tr>
<td>12</td>
<td>+9.7</td>
<td>+8.4</td>
<td>+2.6</td>
</tr>
<tr>
<td></td>
<td>Max. compression</td>
<td>-3.7</td>
<td>-5.2</td>
</tr>
<tr>
<td></td>
<td>Max. tension</td>
<td>+9.7</td>
<td>+8.4</td>
</tr>
<tr>
<td></td>
<td>Strength differential, max.tension/max.compression</td>
<td>2.62</td>
<td>1.61</td>
</tr>
<tr>
<td></td>
<td>Avg. compression</td>
<td>-2.93</td>
<td>-4.6</td>
</tr>
<tr>
<td></td>
<td>Avg. tension</td>
<td>+9.7</td>
<td>+5.2</td>
</tr>
<tr>
<td></td>
<td>Strength differential, avg.tension/avg.compression</td>
<td>3.3</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>Avg. compression over all orientations</td>
<td>-3.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Avg. tension over all orientations</td>
<td>+5.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Strength differential over all orientations, avg.tension/avg.compression</td>
<td>1.49</td>
<td></td>
</tr>
</tbody>
</table>

These theoretical calculations were confirmed in Ref. [25] with experimental results.

Section 6.1.3.1. and Section 6.1.3.2 consider polycrystalline NiTi wire tested in tension and compression in martensitic and superelastic conditions.
6.1.3.1 Deformation of martensite

In the martensitic sample, self-accommodated structure with any of twelve correspondence-variants can be formed by cooling from the austenite phase. As previously discussed, to fulfill the conditions of lattice-invariant share, these variants are internally twinned. During loading, variant clusters, initially formed by cooling, are altered by transformation to a single variant which is able to accommodate the applied stress more effectively. In Figure 24 the plateau region corresponds to growth of one preferable variant at the expense of others. The further loading, as in case of average materials, causes elastic deformation and after reaching yield stress \( \sigma_y \), plastic deformation of the subject martensite variant. For some loading directions, symmetrically arranged groups of several variants can deliver the best strain accommodation. For example, in case of loading the sample along the [111]_A direction, transformation strain of 9.7% in tension can be achieved by clustering of 4, 6, and 12 variants it is seen from Table 2.

![Figure 24. Tensile and compression deformation of martensitic SMA, following Ref. [25]](image-url)
Similar results are in compression loading, when six variants with unique transformation strain-variants may be simultaneously formed: 1, 3, 5, 7, 9 and 11. However, the compression strain is limited to 3.7% only. Therefore, in the experiment, for a given stress level the tension part of the loading curve extends to much larger strains (9.7%) than the compression part.

6.1.3.2. Formation of stress induced martensite

The deformation of the SE sample is schematically presented in Figure 25. In load-free condition the sample is in austenite condition, but upon loading transforms to SIM. When the SE sample is loaded within the elastic limit of the B2 phase, there is no texture change. However, under further loading stress induced martensite variants corresponding to maximum strain accommodation are formed due to applied tension or compression. Since only a little transformation strain can be delivered by the variant formed during compression loading, part of the newly built SIM is deformed: first elastically and if sample further loaded plastically. This mechanism is responsible for a steeper rise in stress during creation of SIM in compression compared to tension as seen in Figure 25, where deformation is governed by creation of undeformed SIM maximum-strain-under-tension variant only. A notable difference in the elastic slopes of tension and compression is attributed to the presence of retained martensite even in the virgin austenite.
Lattice correspondence between martensite variants and austenite is responsible for a directional dependence of transformation strains in martensite or for development of the favorable microstructure or martensite texture. According to Table 2 and Figure 24 and Figure 25, in a NiTi polycrystalline sample the strength differential between tension and compression is about 2 times. Also, the maximum strength differential effect is observed along <111>_A direction. Therefore, the amount of strain recovery in SE and martensitic samples can be optimized by texturizing SMAs for a particular application. However, it also should be considered, that textured sample has lower stored elastic energy in comparison with a sample with randomly oriented grains.

6.1.4. Introduction of dislocations in the austenitic matrix

Ref. [27] studied superelastic NiTi samples with different level of induced plastic deformation (0%, 8%, and 11%). The schematic results are presented in Figure 26 and show that with an increase of plastic deformation level, there are decreased maximum
recoverable strain (residual strain is designated as $\varepsilon_{\text{resid}}$) and decreased hysteresis as well as increased linearity of stress-strain curve.

![Schematic of hysteresis in stress space for samples with and without induced plastic deformation](image)

Figure 26. Schematic of hysteresis in stress space for samples with and without induced plastic deformation, following Ref. [27]

As neutron diffraction spectra show, in unloaded state there is presence of martensite in samples with plastic deformation, while the sample with no prior plastic deformation was fully in austenitic condition. Therefore, induced plastic deformation in NiTi by generating crystallographic defects results in introducing an internal stress field which stabilizes the martensite even in unloaded condition.

Moreover, it was shown that residual or nonrecoverable strain $\varepsilon_{\text{resid}}$ for superelastic transformation in the presence of dislocations generated in austenite matrix can be significantly reduced by cycling. The cycling provides stabilization of additional martensite or texture evolution in the martensite [28]. Stress induced martensite (SIM) can accommodate the mismatch with martensite that is stabilized from stresses associated with the residual plastic strain. The mismatch is fully accommodated elastically following cycling. However, prior to this elastic accommodation, the initial
cycles involve inelastic accommodation, that is associated with phase fraction and texture evolution through twinning in martensite [27].

In thermoelastic martensite transformation classical thermodynamics associates the increase in the linearity of stress-strain curve with increased stored elastic strain energy (decrease in $\Delta H_{net}$). Plastic deformation defect density and accompanying stress field may cause such increase in $\Delta H_{el}$. It is also expected that the increased defect density and correspondent stress field create the conditions for favorable selection of martensite variants or result in their coalescence that reduces the overall energy dissipation. Reduced hysteresis of the cycled sample with induced dislocations is manifestation of decreased frictional resistance of interfacial motion and decreased stored elastic strain energy dissipation. In addition, it was established that the overall reversible strain scales with the volume fraction of transformable martensite, which decreases with residual plastic deformation.

6.1.5. SMA training and martensite texture evolution with cycling

In any actuating system along with increased work output or high speed and narrow hysteresis, good dimensional stability in reproducing an actuation event is of high interest. Typical actuator response of an SMA is shown in Figure 27, where instead of recovering in stress-free conditions, the SMA element remains in the loading condition.
Dimensional instability is evolution in part geometry during thermal cycling of the SMA part under load. If there is no high tolerance in deformation events, the actuating mechanisms become less useful in application. For efficient employment of SMAs for actuating applications, the dimensional instability should be reduced to a sufficient level providing no residual (irreversible) strain with cycling.

The phenomenon of instability from cycle to cycle has been attributed to martensite texture evolution as an equilibrium stress state is attained. Therefore, to properly stabilize or train SMA elements, the underlying microscopic mechanisms responsible for the instabilities in SMAs should be carefully defined and managed. Even though they are still not fully known, the underlying atomistic structure is a groundwork for the processes.

There are two types of material stabilization by training: isobaric training (conventional) and isothermal training. Conventional training uses thermal cycling of SMA elements under fixed load. The drawback of the method is reduced service life and time of
the process which can take weeks depending on facilities and size of the component. Isothermal training uses mechanical cycling at fixed temperatures. Even though the procedure still requires running a single isobaric condition on the component to pinpoint the amount of deformation that needs to be achieved isothermally, the process of stabilization takes fewer cycles and allows to achieve higher service life.

Material stress accommodation or compatibility is the driving force for the SMA to develop texture during cycling and/or training. Lattice defects are accumulated through the phase transformation with cycling. During thermomechanical cycling, favorable variants are formed with the increase in transformational strain at initial cycles. In later cycles, similar martensite variants are selected after each cycle corresponding to the maximum transformation strain [29]. When a dynamic equilibrium between defect generation and annihilation is achieved, the material is stabilized.

The earlier discussed tension-compression asymmetry in texture evolution is dependent on the strain and independent of stress and number of cycles. For the case when tension was applied before compression, preferred variants which were formed in tension partially remained following compression. In particular this material dependence on the strain and independence of stress and number cycles are the basis for achieving stabilization by means of isothermal cycling [29].

The reason why SMA transformation strain increases with cycling is because the microstructure evolves until there is little or no driving force for it to evolve further. However, through training, a state close to saturation is reached, at which the level of walking from cycle to cycle is reduced to a desired level (ideally approaching zero). The process causes an increase in stored elastic energy, as discussed in Chapter 5.
6.1.6. Precipitates in B2 matrix

The effects of the precipitates on the strength of the B2-matrix and the behavior of the subsequent transformations strongly depend on precipitate microstructures, their size and shape, interparticle spacing between, and stress and concentration fields around the precipitates.

**Introduction of intermediate R-phase in NiTi system:** In fully annealed near-equaiatomic NiTi, martensitic transformation occurs from the B2 austenitic phase directly to B19’ martensitic phase. However, in thermally cycled, thermo-mechanically treated, or alloyed with Fe, NiTi alloys transform to martensite through the formation of intermediate trigonal R-phase: B2 → R-phase → B19’. The R-phase transformation is associated with a narrow temperature hysteresis of 1.5 °C, and with the maximum recoverable strain of only about 1%, while a hysteresis associated with B2 → B19’ transformation exceeds 10 °C and possesses about 8% the recoverable strain in shape memory effect. The fatigue life for the R-phase is remarkably high compared to the full B19’ phase transformation [30].

As in case with B2 ↔ B19’ transformation, R-phase transition is associated with lattice distortion, therefore both SME and SE can be observed.

R-phase transition and full martensitic transformation are two competing processes. Therefore, R-phase can be achieved by suppressing the B2 → B19’ transformation. Due to thermoelastic nature of the martensitic transformations (including transformation through R-phase) in SMAs, they are sensitive to the internal stress state of the parent B2 phase. R-phase transition can be attributed to the presence of coherent/semicoherent Ni4Ti3
precipitates in the alloy B2 matrix[31]. The presence of the particles produces a strong resistance to large lattice variant transformation associated with the formation of B19’. While the transformation through the intermediate trigonal R-phase produces a significantly small lattice variant transformation and is much less affected by the particles. Due to low transformation hysteresis and increased service life, the R-phase is beneficial for thermal actuators applications where there is no need for large stroke [32]. The schematic quantitative comparison of hystereses widths and amounts of heat absorption and revealing by SMA through B2 ↔ B19’ and B2 ↔ R-phase is show in Figure 28.

![Diagram showing heat flow and temperature changes for B2 ↔ B19’ and B2 ↔ R-phase transformations.]

**Figure 28.** Schematic Differential scanning Calorimetry results for the B2 ↔ B19’ and B2 ↔ R-phase transformations

In accordance with the quantitative results obtained for NiTi in Ref. [33], the martensite to austenite phase transformation absorbs three times more energy (latent heat) and provides three times higher temperature suppression than the R-phase to austenite transition.
NiTi alloyed by Fe also manifests the transformation through the intermediate R-phase and is referred to low temperature SMA that at room temperature is in austenite phase. Therefore, NiTiFe is of particular interest due to its stability, reduced hysteresis, and improved fatigue life associated with R-phase transformation for cryogenic applications.

**Precipitates in high temperature alloys:**

*Precipitates of so-called P-phase and H-phase have been identified in high temperature NiTiPt and NiTiHf. Both phases are fully coherent precipitates embedded in the B2-matrix.*

(Ni,Pt)Ti system, P-phase precipitates

P-phase is introduced in the (Ni,Pt)Ti system by aging the material at a relatively high temperature (~600°C) for a relatively long time (100 hours). P-phase fine precipitates (Ti_{44}Ni_{36}Pt_{16}) have a monoclinic structure (B19′, a_p = 0.745 nm, b_p = 1.292 nm, c_p = 1.422 nm and β =100.45) and fully coherent with the B2 matrix phase (a_{B2} = 0.3087 nm). The martensitic phase is of B19 structure in the (Ni,Pt)Ti system. Each monoclinic unit cell of the P-phase includes 96 atoms in 16 layers that are parallel to the (111) plane of B2 cell. The precipitation shapes are relatively cuboidal, with edge length 200–400 nm. In total, there are twelve crystallographic equivalent variants possible for the P-phase [6].

P-phase precipitation causes changes in concentration in the B2-matrix due to depletions of Ni and enriching Pt. Both the chemical non-uniformity and stress field associated with the P-phase precipitates are in favor of the martensitic transformation
and affect the $M_s$. Therefore, concentration and stress fields surrounding a coherent precipitate should be determined by the interaction between chemical free energy and coherency elastic strain energy.

To investigate stress field associated with the P-phase precipitate effect, the elastic interaction energy $E_{el}^{int}$ between the P-phase and a nucleating martensitic particle can be calculated as follow:

$$E_{el}^{int} = -\sigma_{ij}\varepsilon_{ij}^{Mv}$$  \hspace{1cm} (52)

where $\sigma_{ij}$ and $\varepsilon_{ij}^{Mv}$ are the stress field generated by the P-phase precipitates and the stress-free transformation strains of a specific martensitic variant respectively. A negative value of the interaction energy corresponds to the favorable formation of martensite by the presence of the P-phase precipitates.

In accordance with the results obtained in [6], all three principal stress-free transformation strains of a P-phase particle are negative. Therefore, there are both compressive and tensile stresses around the particle in the B2-matrix and mostly tensile stresses inside the particle. The calculation of the interaction energy showed that locations with the most negative interaction energy are near the edges of a cube-like precipitate with an edge length of $\sim 116$ nm. Therefore, that edges are the most preferred sites for martensite to nucleate and different martensitic variants should be favored at different edges. However, since internally twinned martensite is usually formed in SMA, it is expected that internally twinned martensitic particles consisting of large fractions of the most favored variant will develop via an autocatalytic effect.
It was shown that the chemical driving force for the transformation, is about one-fifth of the maximum elastic interaction energy. Therefore, P-phase precipitates make a major contribution to the driving force for the B2 to B19 martensitic transformation. These changes in (Ni,Pt)Ti system stress state associated with precipitation of P-phase particles increases the $M_s$. Based on experimental measurements, it results in $M_s$ increasing by 100°C.

**Ni-(Ti,Hf) system, H-phase precipitates**

H-phase in Ni-(Ti,Hf) system (stoichiometry about Ni$_{50}$(Ti,Hf)$_{50}$) is a nano-size coherent particles which appear in B2 matrix ($a_{B2} = 0.309$ nm) either as thin plates when aging at a relatively high temperature (600°C) for a relatively long time (815 hours) or as diamond-like shaped particles when aging at a relatively low temperature (550°C) for a relatively short time (3 hours). H-phase precipitates are of B19 orthorhombic structure with lattice parameters $a_h = 1.264$ nm, $b_h = 0.882$ nm, $c_h = 2.608$ nm. Due to symmetry breaking, six correspondence variants of the H-phase are possible [7].

As in case of P-phase, H-phase precipitation process activates two phenomena which affect the strength of the B2 matrix and the behavior of the subsequent martensitic transformations. First, it creates a spatially inhomogeneous stress field around an H-phase particle. Second, because in the Ni-(Ti,Hf) system $M_s$ strongly correlates with the concentration of Ni and Hf in the B2-matrix, growth of H-phase particles will create a concentration gradient which will affect martensitic transformation. Moreover, at earlier stages of the aging, chemical effect dominates while at later stages mechanical effect is dominant.
The coherency stress field of H-phase precipitation in the B2 matrix was estimated in the same way as was proposed for P-phase by calculating the elastic interaction energy between the H-phase and a nucleating martensitic particle. The maximum elastic interaction energy between the H-phase and the twelve individual martensitic variants of B19’ which are nucleating in the stress field of the precipitates was found at the rims of the diamond-like shapes precipitates. Four martensitic variants (2, 3, 6, and 7) showed maximum value of the interaction energy at the rims (for diamond-like precipitant of 10.0 nm in the longest direction).

The maximum interaction energy between the precipitates and internally twinned self-accommodating variants came to maximum for variant pairs 1-4, 1-3, 2-4, and 2-3. However, this is approximately only 23.5% of the enthalpy of martensitic transformation for individual martensitic variants and 19.4% of the enthalpy of martensitic transformation for internally twinned martensite. Because the elastic interaction energy of a single variant martensite is more negative than that of twin pair, the average transformation strain of the twin pair is much smaller than that of single variant.

Chemical energy change for both, P-phase and H-phase was formulated based on existing thermodynamic databases and Landau theory and results are presented in Ref. [6, 7].

Nano-sized H-phase precipitates, increases $M_s$ by around 40-60°C. The thermal hysteresis is found to be increased for higher aging treatment (650°C, for 3 hours) but decreased for lower aging treatment (550°C, for 3 hours). Furthermore, functional fatigue resistance improves significantly. All these effects have been attributed to
different stress and concentration fields around H-phase precipitates in the B2 matrix [7].

Like any other stress inducing defects (dislocations and grain boundaries), the stress field associated with the formation of precipitates modifies the material internal stress ($\Delta H_{net}$) state and affects the martensite transformation start temperature, material work output, austenite strength, functional fatigue resistance, and dimensional stability. In addition to change in mechanical enthalpy changes, B2 matrix precipitates are associated with chemical enthalpy change.

6.1.7. Transformation matrix

The mathematical description of martensitic transformation can be made with transformation matrix $U_1$ introduced earlier in Section 6.1.3. The procedure is built on the fact that there is no volume change associated with thermoelastic martensitic transformation in SMA.

If the change in volume occurred, then a growing martensite phase in the surrounding austenite phase would generate stress and vice versa. Since distortion matrices deform unstressed austenite to unstressed martensite no volume change is involved in the process [34, 35]. The fact that martensitic transformation is displacive, allows to connect crystallography and microstructure of the alloy with macroscopic properties and required energy to activate the operation of an SMA heat engine in some precise and quantitative manner.

NiTi undergoes B2 to B19’ transformation. In the austenite state all the titanium atoms lie at the corners of the unit cell while atom of nickel lies at the center of the cubic
unit cell. Martensite is an almost face-centered monoclinic in which planes of atoms are slightly displaced from their natural position [34]. Transformation of the austenitic cell to the martensitic cell can be described as unequal stretching of three edges of tetragonal face-centered cell formed by four cubic unit cells with the following change the angle between two edges of length $a$ and $c$ as shown in Figure 29.

![Figure 29. The martensitic transformation in NiTi](image)

The austenite lattice vectors chosen along the edges of the tetragonal cell as shown in Figure 29:

$$e_1^a = a_0 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad e_2^a = a_0 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \quad e_3^a = a_0 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

where $a_0$ is the length of an austenitic cubic cell.

The martensite lattice vectors chosen along the edges of the monoclinic cell:

$$e_1^m = \frac{c}{\sqrt{2}} \begin{pmatrix} \sqrt{2} \cos \beta \\ \sin \beta \\ \sin \beta \end{pmatrix}, \quad e_2^m = \frac{b}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \quad e_3^m = a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

where $a, b, c$ and $\beta$ are lengths of martensitic monoclinic cell and the angle between edges $a$ and $c$ respectively.
Then, in orthonormal basis parallel to the cubic unit cell:

\[ e_i^m = T_1 e_i^a \]  \hspace{1cm} (52)

where

\[
T_1 = \begin{pmatrix}
a & \frac{1}{\sqrt{2}} \frac{c}{\sqrt{2} a_0} \cos \beta & \frac{1}{\sqrt{2}} \frac{c}{\sqrt{2} a_0} \cos \beta \\
\frac{a_0}{a} & 1 & \frac{1}{2} \left( \frac{c}{\sqrt{2} a_0} \sin \beta + \frac{b}{\sqrt{2} a_0} \right) \\
0 & \frac{1}{2} \left( \frac{c}{\sqrt{2} a_0} \sin \beta - \frac{b}{\sqrt{2} a_0} \right) & 1 \\
0 & \frac{1}{2} \left( \frac{c}{\sqrt{2} a_0} \sin \beta - \frac{b}{\sqrt{2} a_0} \right) & 1 \\
\end{pmatrix}
\]

and \( e_i^m \) and \( e_i^a \) are martensite and austenite cell vectors along axis 1, 2, and 3 respectively.

In the subject transformation distortion matrix \( T_1 \) which transforms austenite to martensite is not symmetric, however, can be presented as:

\[ T_1 = QU_1 \]  \hspace{1cm} (53)

where \( Q \) is a rotation and \( U_1 \) is a symmetric and positive-definite matrix.

Because \( T_1 \) and \( U_1 \) are related through a rotation which does not change the state of a lattice, it is conventional to use \( U_1 \) to describe the transformation, and not \( T_1 \). The relationship between \( T_1 \) and \( U_1 \) can be described as: \( T_1^T T_1 = U_1^2 \).

In general case if the lattice vectors of the austenite phase are \{\( e_1^a, e_2^a, e_3^a \)\} and lattice vectors of the martensite phase are \{\( e_1^m, e_2^m, e_3^m \)\} then the diffusionless homogeneous transformation can be expressed as:

\[ e_i^m = U_1 e_i^a \]  \hspace{1cm} (54)

Therefore, for B2 to B19’ transformation \( U_1 \) is [34]:

\[
U_1 = \begin{pmatrix}
\gamma & e & e \\
e & \alpha & \delta \\
e & \delta & \alpha
\end{pmatrix}
\]  \hspace{1cm} (55)

where
\begin{align*}
\gamma &= \frac{a(\sqrt{2}a + c \sin \beta)}{a_0 \sqrt{2a^2 + c^2 + 2\sqrt{2}ac \sin \beta}} \\
\epsilon &= \frac{ac \cos \beta}{\sqrt{2}a_0 \sqrt{2a^2 + c^2 + 2\sqrt{2}ac \sin \beta}} \\
\alpha &= \frac{1}{2\sqrt{2}a_0} \left( \frac{c(c + \sqrt{2}a \sin \beta)}{\sqrt{2a^2 + c^2 + 2\sqrt{2}ac \sin \beta}} + b \right) \\
\delta &= \frac{1}{2\sqrt{2}a_0} \left( \frac{c(c + \sqrt{2}a \sin \beta)}{\sqrt{2a^2 + c^2 + 2\sqrt{2}ac \sin \beta}} - b \right)
\end{align*}

For example, transformation matrix for Ni_{50}Ti_{50} [35]:

\[ U_1 = \begin{pmatrix}
1.0243 & 0.05803 & -0.04266 \\
0.05803 & 1.0243 & -0.04266 \\
-0.04266 & -0.004266 & 0.9563
\end{pmatrix} \]

with tailoring coefficients between austenite-martensite crystals (eigenvalues of the matrix \( U_1 \)): \( \lambda_1 = 1.1066, \lambda_2 = 0.9663, \lambda_3 = 0.9321 \).

Per definition [23], the determinant of the transformational matrix measures the volume ratio of martensite to austenite. Therefore, \( \text{det}|U_1| = 1 \) means no volume change and is a necessary condition for SME and SE phenomena [35, 36]. In the example of Ni_{50}Ti_{50} transformation matrix, \( \text{det} U_1 = 0.9966 \).

In general case the interface between two phases is a transition layer containing elastic and interfacial friction energies. It appears as a laminated mixture of martensite correspondence variants. Neither correspondence variant itself is compatible with austenite. However, a fine laminated mixture of variants can be approximately compatible
with austenite. Therefore, alternating layers of martensitic variants meeting a homogeneous region of austenite, is a representation of the typical interface between martensite and austenite. The fineness of the microstructure is determined by a competition between the elastic energy stored in the transition layer and the interfacial energy on the twin boundaries between the variants [37].

The condition $\lambda_2 = 1$ describes the exact compatible austenite with the martensite. It means that no elastic and interfacial energy (except for the single interface separating austenite and martensite) are involved in the transformation. As it was previously discussed, in Chapter 5 the amount of elastic and interfacial energy is directly linked to $\Delta H_{net}$ and hysteresis width. The closer $\lambda_2$ to 1, the more compatible the austenite and the martensite are, and therefore the smaller the hysteresis is.

6.2. Influence of microstructure on $\Delta H_{net}$

Here the various aforementioned descriptions of SMA microstructure to the enthalpy of the transformation that includes chemical, elastic and interface (i.e., mainly friction) terms are extending. Since the martensitic transformation in SMAs is thermoelastic, internal and external stresses significantly affect the behavior of these alloys. The internal stresses in the SMAs may occur due to mismatch between grains, dislocations, precipitates or second phases.

Both mismatch among grains, variants and dislocations cause changes in $\Delta H_{el}$ and $\Delta H_i$. Mathematically, a decrease in $\Delta H_i$ or approaching a coherent interface between austenite and martensite during the transformation points to the middle eigenvalue $\lambda_2$ of the transformation matrix approaching 1, that gives rise to the reduction of the hysteresis. B2 matrix precipitates in addition to $\Delta H_{el}$ and $\Delta H_i$ also affect $\Delta H_{ch}$ due to the existence of
concentration gradients. In accordance with Equation 44, $\Delta H_{el}$, and $\Delta H_t$, enthalpy changes affect the hysteresis width. Therefore, according to Equation 29 it affects the maximum angular speed.

As a response to applied stress, SMAs develop texture. Martensite is able to reorient its variant structure or grow favorable ones at the expense of ones not favorable or coalesce towards overall a more favorable configuration with respect to stress. During the transformation material generates defects. With cycling, the number of defects generated and annihilated reaches a dynamic equilibrium. The approach of the equilibrium is associated with increased elastic stored energy compared to the virgin material. The ability of martensite to reorient the correspondence-variants or variant coalescence with cycling lowers the overall energy dissipation.

The TWSME removes the need for biasing mechanisms and causes the desired increase in stored elastic energy for maximum speed. However, the previously discussed issue with small force for the shape change in the forward transformation compared to that in the reverse transformation should be carefully considered in the SMA heat engine design. The TWSME influences the thermodynamics through texture and re-distributing internal stresses more than resulting in a strong biasing force.
CHAPTER SEVEN: PERFORMANCE OPTIMIZATION AND FIGURE OF MERIT

The intent of this chapter is to extend the previous mechanical, thermal, and materials aspects presented in Chapter 3 through Chapter 6 and to generalize the approach for achieving desired outputs for the particular applications of the SMA Stirling engine.

7.1. Engine outputs

The total power generated by an SMA wire in the Stirling engine with the configuration shown in Figure 11 is induced by the recovery stress of the bending wire around the hot pulley. The total output power $P_{out}$ of the SMA Stirling engine can be estimated by the output torque (torque on the cold pulley) multiplied by the speed of the cold pulley:

$$P_{out} = T_{out} \omega_c$$

Separately consider cases of increasing $T_{out}$ and $\omega_c$ for maximizing $P_{out}$.

7.1.1. Increasing torque

Recall output torque for the engine in Figure 11 from Equation 20:

$$T_{out} = \frac{\pi d^3}{32} \left( \frac{R_c}{R_h} \left( \sigma_{exit} - \sigma_{entr} \right) - \left( \sigma_{c,h} - \sigma_{c,c} \right) \right)$$

Analyzing Equation 20, it is clear that an increase in $T_{out}$ can be achieved by increasing the wire diameter (or, in the general case, an increase in the section modulus W), radius of the cold pulley $R_c$, bending stresses (moments) acting over the cross section of the SMA element after exiting the heat source and after last contact of the wire with the
cold pulley, $\sigma_{exit}, \sigma_{c,c},$ respectively or by decreasing the bending stresses (moments) acting over the cross section of the SMA element prior to the heat source entrance and first contact of the wire with the cold pulley: $\sigma_{entr}, \sigma_{c,h},$ respectively, and the radius of the hot pulley $R_h$. The variables are shown in Figure 17, Figure 18, and Figure 30.

Figure 30 illustrates the stress distribution in the SMA wire at different temperatures for each position.

All temperatures $T_{exit}, T_{entr}, T_{c,h}$ and $T_{c,c}$ increase with an increase in the heat source temperature. However, due to the non-linear stress-temperature curves, as presented in Figure 30 (a), the moment when the wire transforms to the austenite state, stress is constant at the pulley. Therefore, further increase in the wire temperature due to an increase in temperature of heat source will decrease the stress differences which drive the engine. That is the continuous increase in temperature of heat source will decrease the output torque $T_{exit}$.
An important statement made in Chapter 5, is that bulk nucleation in SMA results in an increase in the stored elastic energy and increases the hysteresis ($A_s$-$A_f$ and $M_s$-$M_f$) in accordance with Figure 21. Therefore, the requirements of a wider hysteresis per Equation 18 and Figure 20 - Figure 21, Figure 30 indicate that bulk nucleation is a desirable method for martensitic nucleation for achieving larger $T_{exit}$ for an SMA Stirling engine.

### 7.1.2. Maximum angular speed

Multiplying $\omega_{h,\text{max}}$ from Equation 29 by the transmission ratio $i$, the angular speed of the cold pulley can be obtained (note: maximum speed):

$$\omega_{c,\text{max}} = \omega_{h,\text{max}} \cdot i = -\frac{4 \cdot h_1 \cdot \rho \cdot c_p \cdot d \cdot \ln \left( \frac{T_{source} - T_{exit}}{T_{source} - M_f} \right)}{R_h} \frac{R_h}{R_c}$$

(57)

It can be recalled from Chapter 3 that $\omega_{h,\text{max}}$ was obtained under the assumption that the transformation time $\tau$ is equal to the time $t$ required to overcome the distance corresponding to angle $\varphi$.

Analyzing Equation 57, it is clear that an increase in $\omega_{c,\text{max}}$ can be achieved by decreasing the wire diameter (or, in the general case, decreasing the section modulus $W$), radius of the engine cold pulley $R_c$, decrease in hysteresis $A_f - M_f$ or by increasing the radius of the hot pulley $R_h$ and increasing the temperature of the source. In contrast to the requirement for increased $T_{exit}$, surface nucleation is necessary to achieve $\omega_{c,\text{max}}$ for the SMA Stirling engine.
These derived mathematical descriptions of the SMA Stirling engine outputs are in agreement with the established theory of traditional engine output. In which there is an interdependence between decreasing $T_{out}$ and proportional increasing in $\omega_c^{max}$ and vice versa.

For engines operating in a “dense” medium such as liquid, friction torque loss between rotating pulley and liquid should be considered. The losses will increase with increasing in speed.

7.2. Figure of Merit

To quantitatively estimate the performance of SMA elements in heat engines regardless of the working temperature range, component size and mass, but considering the need for elevated torque and energy absorption or speed, a relative measure of the material performance with the proposed Figure of Merit (FOM) can be employed.

A previously derived FOM [33] for phase change materials does not consider the effect of different energetic components in $\Delta H_{net}$ which is crucial for speed application. It was shown in Chapter 5, hysteresis is a measure of the dissipated work which is corelated with $\Delta H_{net}$ and directly relates to the engine output. The FOM is a convenient tool which may predict the relative output of the SMA element on the engine development stage depending on the customer’s requirement.

In accordance with all phenomena associated with martensitic transformations in SMA studied in Chapter 3 through Chapter 6, and the considering previously proposed FOM [33], the material efficiency for both heat engine applications requiring either high torque and maximum latent heat or high speed, can be estimated with the FOM presented
in Equation 58. However, for increased $T_{out}$ of Stirling engine and maximum energy storage applications, the FOM should be as high as possible, while for maximizing $\omega_c^{max}$, it should be as low as possible.

$$FOM = \frac{\Delta H_{net} \cdot c_p}{\rho \cdot h \cdot k_A}$$

(58)

where $\rho$, $\Delta H_{net}$, $k_A$, $c_p$, and $h$ are the density, the latent heat of endothermic transformation, the thermal conductivity of austenite, specific heat capacity, and the heat transfer coefficient between the element and the heat source, respectively.

The FOM from Equation 58 is an inverse measure of the volumetric capacity of SMAs to readily recover mechanical energy stored in a system by activating the work of a Stirling engine with excessive heat.

As was previously shown in Figure 21, elastic constraints in the form of “introduced” grains, change the amount of stored elastic energy. A downward trend in the transformation-temperature range develops with decreasing grain size or with an increase in the amount of stored elastic energy thereby affecting $\Delta H_{net}$.

In accordance with Ref. [23], the critical chemical force at $M_s$ scales directly with the elastic modulus. The smaller the magnitude of the chemical driving force, the smaller will be the elastic strain energy that must be stored in the matrix to maintain the thermoelastic balance or vice versa. Therefore, the elastic modulus is another parameter already discussed in Chapter 5, which affects the hysteresis width and in turn $\Delta H_{net}$. 
In accordance with Hooke’s law, a low elastic modulus will cause a proportional reduction in elastic stresses, thus favoring elastic accommodation without exceeding the level of stress for slip.
CHAPTER EIGHT: CASE STUDY

8.1. Total amount of rejected heat

Using Equation 29 for maximum speed of the hot pulley, two modes of the SMA heat engine operating through $A \leftrightarrow M$ and $A \leftrightarrow R$ transformations for rejecting parasitic heat can be estimated for the entire period of the engine operation:

Consider the speed of a hot pulley associated with an $A \leftrightarrow M$ transformation:

$$
\omega_{h_{\text{max}}} = \frac{-4 \cdot h \cdot \varphi}{\rho \cdot c_p \cdot d \cdot \ln \left( \frac{T_{\text{source}} - A_f}{T_{\text{source}} - M_f} \right)}
$$

(29)

Similarly for maximum speed for the transformation through R phase $\omega_{h_{\text{max}, R}}$:

$$
\omega_{h_{\text{max}, R}} = \frac{-4 \cdot h \cdot \varphi}{\rho \cdot c_p \cdot d \cdot \ln \left( \frac{T_{\text{source}} - A_{f_R}}{T_{\text{source}} - R_f} \right)}
$$

(59)

where $R_f$ and $A_{f_R}$ are temperatures of the R-phase finish and austenite finish for the $A \leftrightarrow R$ transformation, respectively.

Keeping all the parameters fixed except for the temperatures, the ratio of the angular speeds is given:

$$
\left( \frac{\omega_{h_{\text{max}, R}}}{\omega_{h_{\text{max}}}} \right) = \frac{\ln \left( \frac{T_{\text{source}} - A_f}{T_{\text{source}} - M_f} \right)}{\ln \left( \frac{T_{\text{source}} - A_{f_R}}{T_{\text{source}} - R_f} \right)}
$$

(60)

The amount of heat rejected for the specific period of engine operation $Q_{\text{ref}}$:

$$
Q_{\text{ref}} = \omega_h \cdot \Delta H_{\text{net}}
$$

(61)
It should be kept in mind that the speed ratio from Equation 60 is highly sensitive to even a small relative change between $T_{source}$ and $A_f, M_f, A_{fR}$, or $R_f$. Therefore, the amount of heat rejected by the engine operating through the R-phase and through the B19’ phase cannot be easily compared.

However, it can be qualitatively estimated by considering the amount of heat rejected for the specific period of engine operation $Q_{rej}$ from Equation 61 with the FOM derived in Equation 58 for both maximum speed (FOM $\rightarrow$ min) and for increased torque (FOM $\rightarrow$ max).

In accordance with the FOM $\rightarrow$ min, for maximizing speed in addition to $\Delta H_{net} \rightarrow$ minimum, the requirement of $\Delta H_{et} \rightarrow$ minimum should be considered for achieving minimum hysteresis. However, to reject maximum heat following Equation 61, $\Delta H_{net}$ is required to be as large as possible. Therefore, maximizing heat rejection under maximum angular speed is not straightforward and careful experiment with real numbers is required.
CHAPTER NINE: CONCLUSIONS AND FUTURE WORK

9.1. Conclusions

This work brings together all current research results on SMA heat engines for thermal management and energy recovery. Prior work considered increased torque output as a measure of the SMA heat engine power and did not consider the angular speed. It was not surprising that, in accordance with derived equations, the parameters that provide increased torque lower the speed.

SMA Stirling engine work production with speed, as well as thermal efficiency of the SMA element operating in tension (bending) mode was evaluated in the work. The amount of heat energy involved in the process, its transformation into mechanical (work output) and internal work were discussed in detail from mechanical, thermal, and materials aspects. For the first time the work combined these aspects together to understand the behavior of SMAs in heat engines. This facilitated the development of a tool set for understanding engine operation based on SMA elements, and identified parameters that affect work done by the engine.

One of the emphases of this framework was to show that SMA heat engine can work utilizing ambient temperature or parasitic heat, which are free. Therefore, comparison of absolute values of Carnot efficiency and actual efficiency cannot be considered as an estimation of the SMA heat engine performance. However, a general approach should be discussed.
Maximum theoretical efficiency $\eta_{\text{Carnot}}^{\text{max}}$ [12]:

$$\eta_{\text{Carnot}}^{\text{max}} = \frac{T_H - T_L}{T_H} = 35.6\%$$

(62)

For the external combustion engines, the efficiency is calculated as follows:

$$\eta_{\text{SMA}} = \frac{W_{\text{out}}}{Q_{\text{in}}}$$

(63)

One of the highest thermal efficiencies of heat engines with SMA element $\eta_{\text{SMA}}$ obtained in practice in accordance with Equation 63 and reported in the literature is slightly less 12% [22].

However, even though thermal efficiency of the SMA Stirling engine is much lower than its maximum Carnot efficiency, the ability of material to perform work can be activated by ambient temperature (depending on application using hot ambient air of $T_H$ as heat source or low ambient air of $T_i$ as heat sink in accordance with Figure 10) or even waste and parasitic heat, which means that engine can perform work by utilizing free and accessible energy. Therefore, comparison of absolute values of Carnot and actual efficiency $\eta_{\text{SMA}}$ cannot be considered as an estimation of the SMA heat engine performance. Moreover, efficiency can be increased by modifying parameters affecting the martensitic transformation. Instead, the requirements for a particular application of SMA Stirling engine should dictate the choice of right parameters.

The influence of precipitates in the austenite phase, TWSME, SMA microstructure and its evolution with cycling, martensite twinning, variants reorientations and coalescence were considered as contributors to the whole energetics
of the martensitic transformation in SMAs through materials analyses. Special attention was paid to the difference between chemical enthalpy change and net enthalpy change involved in the transformation. For the first time in the literature, the connection between the nucleation mechanism (bulk or surface) in martensitic transformation to the engine ability of SMA heat engine to perform the work was discussed. Temperature distribution along the wire as a function of time and wire position and the heat transfer rate between the SMA element and the heat source or the heat sink were addressed in the analysis of thermal aspects. In addition, influence of the temperature of heat source and sink, the SMA element cross-section were outlined as parameters affecting engine outputs during the mechanical analysis of the SMA Stirling engine.

The work establishes the stage for further work in the area of thermal control and energy recovery with the SMA heat engines for particular applications, i.e., specific temperature ranges of operation and achieving the desired outputs (see e.g., Ref. [40]).

9.2. Future Work

The equations derived in the work are in agreement with the experimental data previously reported [38]. However, an experiment that captures all theoretical conclusions made in the subject work for the particular engine design in different applications is needed to fully develop the work for engineering use. The theoretical comparison of the heat rejection through R-phase and through B19’ phase considered in Chapter 8 is not straight forward. However, an experiment also would prove valuable.

Also, it was outlined in the work that the materials with wide hysteresis have the inherent capacity to harvest and store energy. The work on charging an external
capacitor with the SMA engine and releasing that stored energy on demand is very promising and should be developed.

The activating of engine work without external torque was achieved on Stirling engine of unsynchronized configuration operating between hot water and air of ambient temperature using the heat source of high temperature by heterogeneous transformation [39]. However, an initial torque generator for the applications with heat source of moderate temperature should be developed to optimize the engine design and improve its applicability.

Estimation of fatigue life of an SMA element especially for engines operating at high speed should be analyzed in future works. Moreover, joining NiTi-systems is a critical problem right now and a reliable solution for increasing engine life of operation is necessary.

Potential presence of stress induced martensite (even though presence can be controlled by pulleys size) and its effect on global energy balance may also discover further parameters affecting $\Delta H_{net}$ and therefore should also be taken into account.
REFERENCES


