A review of shape memory stability in NiTi alloys

N.B. Morgan and C.M. Friend

Department of Materials and Medical Sciences, Cranfield University, Shrivenham, Swindon, Wils SN6 8LA, U.K.

Stability and fatigue are terms often referred to when discussing the durability of the shape memory effect. This paper considers the meaning of these terms and how they relate to the macroscopic memory response and the microstructural accommodation response of NiTi alloys.

The results of previous workers are analysed and compared to establish how the three parameters that apparently control the macroscopic memory effect - temperature, stress and strain, are controlled themselves by the alloy’s microstructure. It will be shown that repeated phase changes will alter the microstructure and hysteresis of the transformation and in turn this will lead to changes in transformation temperatures, transformation stresses and transformation strains.

Both superelastic and memory effects are considered simultaneously so as to emphasise the similarities between the respective stress/strain and strain/temperature degradation effects. Although the mechanisms of both are often regarded as physically separate it is shown that in fact the macroscopic changes observed during cycling are a result of similar microstructural changes.

It is also shown how as a result of the inherently large transformation strains internal cracking may occur and that even apparently stable microstructures can fail catastrophically via conventional crack growth mechanisms.

1. INTRODUCTION

The lack of reliable data on shape memory fatigue effects is a major barrier to NiTi applications [1]. In particular, the acceptance of NiTi by the medical industry and its associated regulatory bodies depends heavily upon assessing the chemical and mechanical lifetime of NiTi products. Often shape memory and superelasticity are treated as entirely separate effects when investigating their fatigue behaviour and cyclic stability. However, due to the very close relationship of these effects many stability characteristics are common to both. This will be demonstrated within this paper.

2. CYCLIC STABILITY AND FATIGUE

'Fatigue', when used to describe the stability of shape memory alloys requires an extension of the usual definition. The following definition is adapted from Van Humbeeck [2] :-

Failure by fracture - Stress or strain cycling at constant temperature. Three situations are possible: the alloy is either in the martensite or parent phase during cycling, the martensite transformation is stress induced above \(A_t\) during each cycle.

Changes in physical, mechanical and memory properties - Due to pure thermal cycling through the transformation region.

Changes in physical, mechanical and memory properties - Due to cycling in or through the transformation region with a combination of applied stress and strain.

This definition is useful as it separates classical fatigue effects from transformation cycling effects. Failure by fracture encompasses conventional fatigue mechanisms and crack growth whilst changes in physical, mechanical and memory properties describes a process whereby these properties change due to repeated phase transformations. This paper will concentrate on the latter.
Following Van Humbeeck [2], the main parameters to consider when studying the memory stability of shape memory alloys are: temperature, stress and macroscopic shape strain. However, it should also be kept in mind that these are all macroscopic parameters that are controlled by the microstructural phase transformation within the alloy. Therefore, an understanding of the microstructural and transformation changes associated with the observed cycling effects is the key to defining how temperature, stress and strain relate to the stability of shape memory alloys. It has been shown that cycling results are often contradictory. If more emphasis was placed on relating microstructural effects with macroscopic parameters a better understanding of shape memory stability would likely result.

The following parameters are therefore very closely linked and all will affect the stability and fatigue behaviour of memory alloys:-

- Temperature
- Stress
- Strain
- Microstructure

Because microstructure is inherently linked to the thermodynamic behaviour of the alloy, the importance of these four factors and their inter-relationship is emphasised by the Claussius-Clapeyron equation in the form:

$$\frac{d\sigma}{dT} = -\frac{\Delta S}{\varepsilon}$$

Where $\sigma$ is a uniaxial stress, $T$ is transformation temperature, $\varepsilon$ is the strain associated with the transformation and $\Delta S$ is the entropy change per unit volume. It is the stability of these variables that defines how stable the alloy is with regards to changes in physical, mechanical and memory properties during repeated phase transformations.

The Claussius-Clapeyron equation emphasises the close relationship between superelastic cycling effects and one-way memory cycling effects. A study of these effects should not be polarised into one or other of them. For instance, a change of transformation temperatures during repeated memory cycles, may be caused by the same microstructural changes that result in a change of transformation stresses during superelastic cycles. Depending upon the type of application different parameters will define cyclic stability:-

**Superelastic applications** are likely to operate at an applied temperature (often ambient) over a particular stress envelope defined by the stress required for inducing the martensite transformation. It is likely that the stability of the transformation stresses and their associated strains will be of primary importance when considering the lifetime of the alloy.

**Actuator applications** are likely to operate against an applied stress over a particular temperature envelope defined by the $M_f - A_f$ range required for martensite transformation. It is likely that the stability of the transformation temperatures and the associated strains at various temperatures will be of primary importance when considering the lifetime of the alloy.

Measuring changes in transformation temperature, transformation stress and transformation strain can therefore yield important information on cycling mechanisms and how the transforming microstructure changes. The following sections will consider the stability of NiTi alloys during thermal actuation and superelastic cycling with respect to these parameters. In addition it will be shown that damage and catastrophic failure by fast fracture may also occur via conventional fatigue mechanisms in alloys that are cycled over the memory transformation strain against an applied stress.
3. START AND FINISH OF TRANSFORMATION

3.1 Thermal Cycling

The drifting of transformation temperatures during cycling is perhaps one of the most interesting cycling effects to consider first. Transformation temperatures are very dependent upon the microstructure and internal stress of the alloy. For this reason ageing effects and dislocations can significantly change the transformation temperatures and hysteresis profile. Authors are in agreement that it is chiefly dislocation build up that is responsible for shifts in transformation temperatures during repeated phase transformation cycling. The most stable alloys when actuating against constant stress are those that are nickel rich and are heat treated at temperatures below 500°C after substantial prior cold work [3]. This is attributed to the fact that these microstructures are most likely to resist further dislocation generation and that cold work stabilises the microstructure.

The stabilisation of transformation temperatures by prior cold work is demonstrated by a previous finding of Miyazaki et al. [4] [5]. This involved pre-straining a 49.8at.%Ni alloy with no previous cold work so that a residual plastic strain could be measured after unloading. They found that pre-straining the alloys resulted in lower martensite transformation temperatures. However, by repeated thermal cycling of the alloys against zero stress, they found that $M_s$ increased with pre-strains of 4.6% and 11.8% and was almost constant when the pre-strain was as high as 22.8%. When no pre-strain was induced (0%) a decrease in $M_s$ was observed. These results are shown in Figure 1 where the change in $M_s$ is described as a function of the number of thermal cycles. It is found that the degree of change in $M_s$ depends upon the amount of pre-strain and that therefore the effects of thermal cycling against zero stress are suppressed by work hardening. If the pre-strain is small, then new dislocations may be generated or existing dislocations may be rearranged such that the transformation gradually becomes easier. This will lead to a change of the internal stress field resulting in a shift of $M_s$. In a strongly work hardened condition, dislocations will find it very difficult to move during thermal cycling and thus $M_s$ will change very little. If there is no pre-strain prior to cycling then the $M_s$ will gradually decrease as the dislocations generated during cycling interfere with the transformation.

![Figure 1](image-url)

Figure 1 Temperature shift in $M_s$ due to thermal cycling for specimens with various pre-strains[5]

Figure 1 highlights the type of contradictions that can make shape memory degradation processes difficult to understand. However, careful analysis of previous papers helps explain why anomalies are found when analysing thermal cycling effects. Where authors find decreasing $M_s$ temperatures [4][6] it tends to be associated with pure thermal cycling or very low stress levels and implies that dislocations within the microstructure interfere with the transformation and increase the degree of undercooling required for the transformation to proceed. Similarly, where authors find increasing $M_s$ temperatures [6][7][8] it tends to be associated with an applied stress either prior to or during thermal cycling. In this case the changing microstructure appears to assist the transformation resulting in a decrease of the required undercooling.
Although cycling effects are attributed to dislocation generation [9][10] defect redistribution [11], or both [11], very little confirming evidence is ever provided although it has been shown by transmission electron microscopy that dislocations are generated during stress free transformation cycling [4]. It is obvious how complicated dislocation arrays generated during stress free cycling may interfere with the transformation accommodation mechanism thus reducing the $M_s$ temperature. It is perhaps less clear how the generation of dislocations may actually assist a displasive transformation. Otsuka states that an increase in stacking faults may actually act as nucleation sites for martensite plates [12] and thereby assist the transformation. It is likely that the observed increase of $M_s$ temperatures must occur through a similar mechanism as that which produces two-way memory, i.e. distributed stress fields assisting the transformation. A build up of internal stress has been confirmed using differential scanning calorimetry [3].

3.2 Superelastic Cycling

This is perhaps one of the few areas where previous work is in good agreement. The stress required for the forward martensite transformation, $\sigma_{M_s}$, generally decreases with the number of cycles [13][14][15][16][17][18][19]. Associated with this decrease in stress, tends to be a decrease of the stress hysteresis and a change in the gradient of the stress induced martensite plateau, (i.e. the gradient gradually becomes steeper). These effects may be seen in the hysteresis profiles shown in Figure 2 [17].

One way of analysing the change in hysteresis gradients is to consider the origin of the hysteresis profile associated with shape memory. The transformation hysteresis is not solely controlled by chemical free energy resulting from the change in temperature there are also microstructural dependent factors that will influence the width and gradient of the hysteresis [20][21]. The inherent link between superelasticity (stress/strain) and thermal memory (temperature/strain) effects is emphasised by the observed changes to the hysteresis.

Elastic energy stored in the alloy during the forward martensite transformation and frictional energy lost due to the movement of interfaces are both controlling factors in the profile of the observed hysteresis. The gradually increasing stress plateau gradient that occurs during superelastic cycling corresponds with increasing elastic strain energy whilst the concurrent decreasing stress hysteresis corresponds with decreasing internal friction [3]. Miyazaki et al. [17] explains the gradually increasing stress plateau gradient in terms of an internal stress field that will have a 'gradient in its strength'. Therefore, the applied stress required to transform the parent phase to martensite does not stay steady and gradually increases with the volume fraction of transformed martensite. Again the origin of the internal stress field is attributed to permanent slip deformation due to repeated cycling and is essentially describing an increase of elastic energy such as that described in [3]. These changes can be directly compared to the gradually increasing temperature interval between the start and finish of transformation during thermal cycling of the memory effect against applied stress and the decrease of thermal hysteresis [3].

![Figure 2 The effect of cyclic deformation on the critical stress for inducing martensite](image-url)
Miyazaki et al. [17] also explains the decrease of transformation stress shown by Figure 2 in terms of an internal stress created by permanent slip deformation. The hypothesis being that an internal stress created by slip actually assists the formation of martensite and therefore reduces the critical applied stress for transformation. This supports the hypothesis of Morgan [3] that internal stresses caused by dislocations can actually assist the transformation of martensite. The stress results in an increase of the $M_s$ temperature during thermal cycling or a decrease in the critical stress required for transformation during superelastic cycling.

4. PERMANENT STRAINS

Another feature of thermal memory cycling against applied stress and superelastic transformation cycling are the permanent strains produced [3][8][22]. Permanent strain may be defined as the deviation of the starting strain after cycle number $N = n$ from the original starting strain at cycle $N = 1$ and can apply to both the hot and cold shape.

4.1 Thermal Cycling

Discrepancies can arise when considering the generation of permanent strains in alloys thermally cycled against applied stress. Some alloys display permanent shape deformation [8], some show no permanent deformation and some show greater deformation at low stresses than high stresses [6].

A good example of the confusion that can arise may be found in Todoroki et al. [6]. Here, alloys cycled against low stress levels displayed greater permanent strains than those cycled at high stress levels. The large increase of permanent strain in the alloys cycled at low stresses can be attributed to a gradually increasing volume fraction of transformed martensite. Both increasing $M_s$ temperatures and internal residual stress fields can cause a gradual increase in the volume fraction of transforming martensite at low stresses [3].

All references where thermal cycling against appreciable applied loads is carried out show some signs of permanent deformation. Often the greatest degree of permanent deformation appears in the parent phase shape. This is demonstrated in Figure 3 [8]. The amount of permanent strain in the high temperature shape is more significant than that in the low temperature shape. Because there is a greater degree of permanent strain in the parent phase the overall recoverable strain (stroke) decreases with cycling.

![Figure 3 Influence of repeated thermal cycling on the transformation behaviour of a Ti-50.2at% Ni wire annealed at 500°C for 15 minutes after cold drawing. Applied Stress = 200MPa [8]](image)

Higher stress levels also result in slip and plastic deformation. Dimensional stability at high stress levels tends to increase with decreasing heat treatment temperatures and increased prior cold work [22]. Many authors support this hypothesis including Lin and Wu [23] who showed that cold work and dislocations resist permanent strain and degradation of recovery strain. They conclude that the process of cold work
authors support this hypothesis including Lin and Wu [23] who showed that cold work and dislocations resist permanent strain and degradation of recovery strain. They conclude that the process of cold work and thermal cycling strengthens the NiTi alloys because of dislocation generation and a concurrent increase in critical stress for slip.

4.2 Superalelastic Cycling

The cause of residual strain in superelastic alloys is also probably due to the occurrence of slip during the transformation as demonstrated by Melton and Mercier [7]. The internal stress caused by slip can also assist the stress-induced martensite transformation by reducing the critical applied stress - $\sigma_M$ [3].

Permanent strain is linked to the starting microstructure of the alloy. Just as in the previous section – 4.1, it has been found that higher heat treatment temperatures result in greater permanent strains [13]. Again this confirms the importance of prior cold work and heat treatment temperature on the cyclic stability of all shape memory effects. The same slip that causes changes of internal friction and elastic energy (section 3.2) also results in permanent macroscopic shape strain.

Microstructural observations of the alloy surface before and during superelastic cycling [17] reveals residual martensite plates that do not revert back to the parent phase even when the applied stress is completely removed [3]. These retained plates are stabilised by the stress field formed by slip deformation. The plates tend to be located at grain boundaries where the dislocation density and associated stress fields are likely to be highest. Miyazaki et al. [17] consider slip deformation and retained martensite plates to be the cause of permanent strains after cycling. Interestingly, the variants of martensite which are retained are the same as those induced by the externally applied stress indicating that the stress fields formed around them are of the same sense as the applied stress. This confirms that the stress field actually assists externally applied stress [3]. The stabilised variants result in a reduction of the superelastic strain hysteresis and a greater degree of permanent strain in the parent phase as shown in Figure 4 (a) and (b) [17]. This is again directly analogous to the thermal cycling effects shown previously in Figure 3.

![Figure 4](image)

Figure 4 The effect of cyclic deformation on (a) strain hysteresis and (b) permanent strain [17]

5. INTERNAL CRACKING AND CATASTROPHIC FAILURE

It has been found that thermal cycling against applied loads can result in a large reduction of ultimate tensile strength (UTS) and elongation to failure [3]. Metallography of wires that show this reduction in strength has revealed a high degree of internal cracking Figure 5 (a). Cracks were not found to be present prior to cycling and it is possible that they form due to the high recovery strains associated with full memory cycles and the presence of applied stress. Miyazaki [5] has also found similar internal cracking in alloys that have been subjected to repeated superelastic transformations, Figure 5 (b). In this case they are ascribed to stress concentrations at grain boundaries. In both cases an applied stress and shape strain is required to initiate the crack.
Figure 5 (a) - Gross internal cracking in a 50.12at% Ti-Ni alloy thermally cycled against an applied stress of 185MPa [Errore. Il segnalibro non è definito.] (b) - Internal Cracking generated by superelastic cycling in a Ni rich alloy [5]

Analysis of a fracture surface in an alloy that failed after 83 000 thermal cycles operating against a constant applied stress reveals fatigue striations such as those found in conventional fatigue failures, Figure 6 [3]. It is not clear whether this is the result of the coalescence of internal cracks or if it is due to crack growth initiated at the site of a defect.

Figure 6 Fatigue striations on the fracture surface of a high cycle failure

6. CONCLUSIONS ON CYCLIC STABILITY

- Cycling effects displayed by repeated phase transformations in actuating thermal memory alloys and superelastic alloys originate from the same changes in microstructure.

- The most stable superelastic and thermal actuator alloys are those that have the highest resistance to slip, i.e. those that are heavily cold worked (30-40%) prior to shape setting at low temperatures (<500°C). There is also some evidence that Ni rich alloys are also more resistant to slip possibly due to the presence of Ni rich precipitates [3] [17].

- Internal cracking may occur at areas of stress concentration even in those alloys optimised to resist slip. This can badly degrade the mechanical integrity of the alloy.
Conventional fatigue crack growth can occur in alloys that are repeatedly cycled through the transformation strain when operating against applied stress.

References