Systematizing the Application of Shape Memory

Duerig, Melton

Shape Memory Alloy' 86
(eds.) C. Youri et al
pp. 397-404

1986

SYSTEMATIZING THE APPLICATION OF SHAPE MEMORY

T. W. Duerig and K. N. Melton
(Raychem Corporation, Menlo Park, CA 94025, USA)

Designing with shape memory is far more complicated than with a non-transforming material. The field requires new terminology - usual design parameters such as "ultimate tensile strength" become less important than "recovery strain" and "M_s". Shape memory events take place in three dimensions (stress, strain and temperature). Moreover, these variables are not state functions: e.g. stipulating a temperature and a strain are insufficient to determine stress. But probably the most complicating aspect of this new technology is the wide range of applications to which it can be applied. This makes it very difficult to screen applications (distinguish the viable from the nonsensical), and to complete the product development process.

An approach will be presented to define, categorize, and screen applications into four fundamental groups. These divisions are made according to the function of the shape memory alloy itself. Each represents a very different "event" in this stress-strain-temperature space and each requires a different set of design parameters and alloy properties. The intent of this paper is not to provide specific data, but to introduce the general screening philosophy developed by Raychem during nearly 20 years of memory product design.

FREE RECOVERY

Free recovery is the simplest of the shape memory events, consisting of a deformation of the martensite and then heating to recover the original shape (see Figure 1). Some key descriptors of the event shown in Figure 1a are the total deformation strain (ε_T), the plastic strain (ε_p), and the amnesia, or final strain (ε_f). Implicitly defined is the recovery strain ($\varepsilon_p - \varepsilon_f$). As the total strain increases, so do the plastic and amnesia strains; the recovery strain increases with until some maximum (8% in the case of NiTi alloys), and then decreases. Figure 1b is the strain-temperature projection of the same event, showing three more parameters: the austenite start and the austenite finish temperatures (A_s and A_f), and the deformation temperature, T_d.

There are few practical uses of the free recovery event, the one celebrated exception being space antenna[1]. The most common use of the event today is in toys and demonstrations.

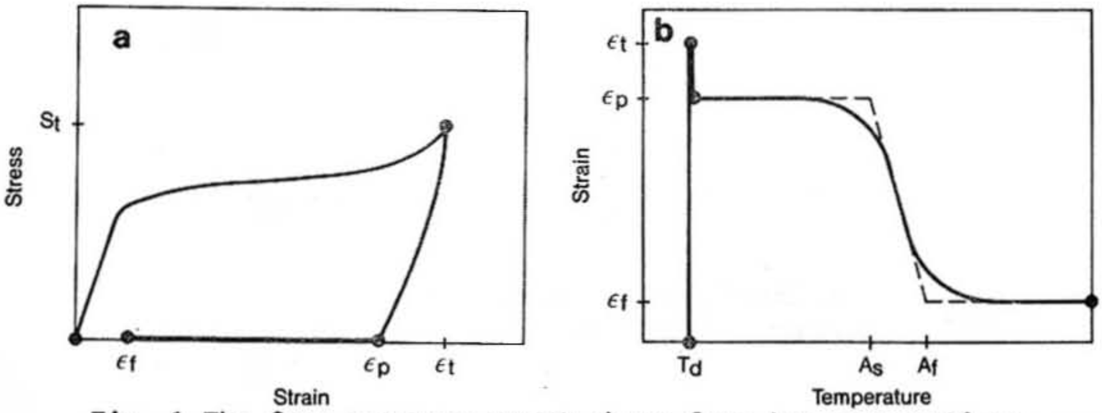


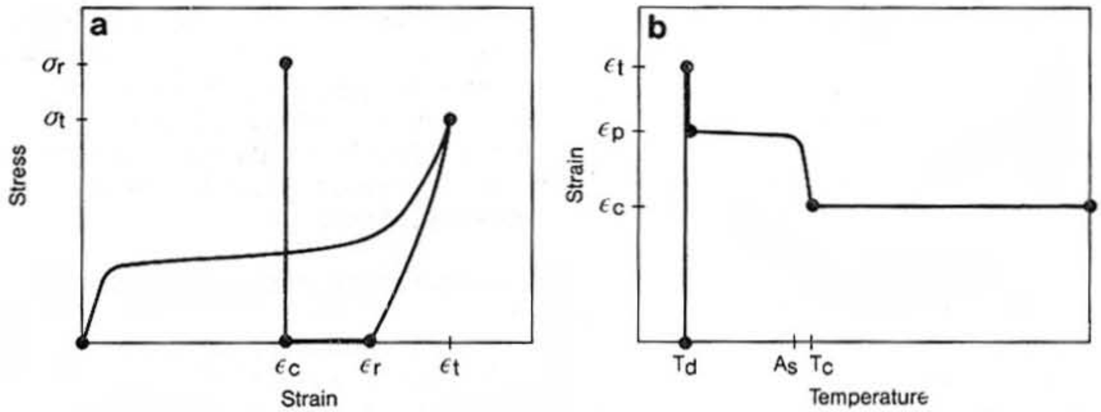
Fig. 1 The free recovery event shown from two perspectives.

CONSTRAINED RECOVERY

Constrained recovery is best visualized by imagining a solid, rigid pin, and a shape memory ring machined with an inside diameter slightly smaller than the pin. The ring is expanded in the martensitic state and placed over the pin. As it is heated, it freely recovers until it contacts the pin. The pin prevents full recovery and a stress is generated. Schematically, this is shown in Figure 2. As before, the event is multi-dimensional, and should be viewed in three planes: stress-strain, stress-temperature and strain-temperature. Two additional parameters appear in the stress-strain perspective (Figure 2a): the contact strain (ϵ_c) and the recovery stress (σ_r). Figure 2b shows the free recovery portion of the event and the contact temperature T_c. In Figure 2c, the stress rate (dS/dT) and M_d are shown. The stress rate is constant and can be derived theoretically from the Clausius-Clapeyron equation[2]. Typical values for this very important property (in MPa/C) are: 4. to 20. for the martensitic transformation in NiTi (30. to 70. for the R-phase) and 2. to 5. for brass. is defined as the temperature above which martensite cannot be stress induced, indicated by the leveling of the stress-temperature curve in Figure 2c.

Figure 3 shows that the recovery stress generally increases with contact strain, implicating that low contact strains are to be avoided. So although an alloy may have 8% "free" motion, only 6% is usable if a dependable stress level is expected. Note that the general envelope of Figure 3 resembles an austenitic stress-strain curve. In fact they do not overlap. The stresses developed during recovery fall typically 15% below the isothermal austenitic stress-strain curve.

The above description relates to constraint by an absolutely rigid substrate with the same coefficient of thermal expansion as the memory material. More often, the substrate deforms, either elastically or plastically. In the case of a pipe coupling, the pipe is crushed during recovery and the path is controlled by the pipe's tensile properties. Moreover, there is a differential thermal expansion so that the



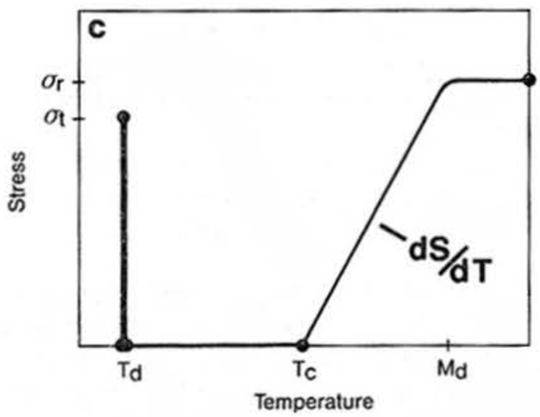


Fig. 2 Constained recovery schematically shown from 3 perspectives: stress-strain (a), strain-temperature (b) and stress-temperature (c).

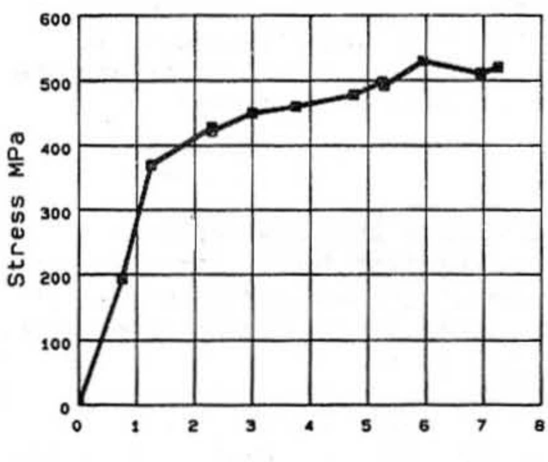


Fig. 3 Stresses developed in an annealed NiTi-based alloy by rigidly constraining at various contact strains.

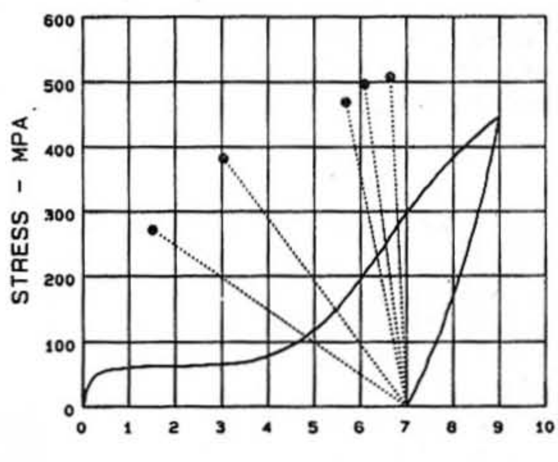


Fig. 4 Stresses developed in an annealed NiTi-based alloy alloy by elastically con- staining at five compli

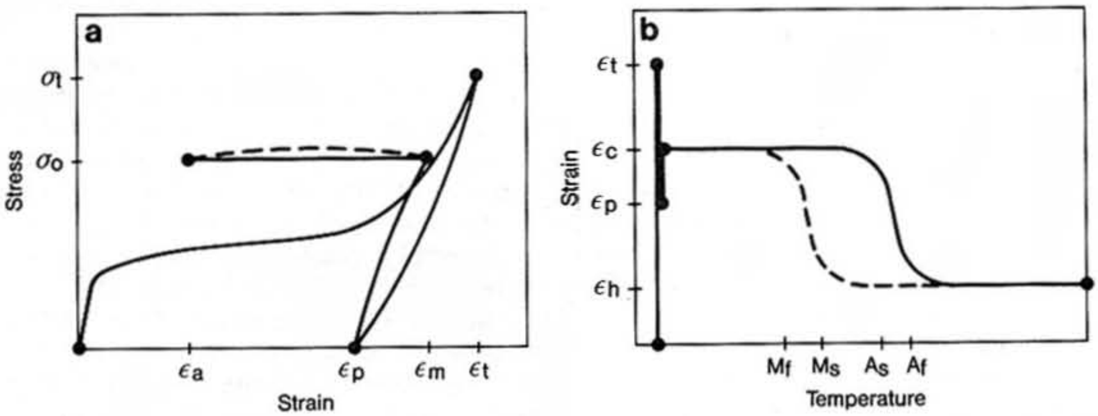
interference stress at any given time is given by: $\Delta\sigma = E\Delta\alpha\Delta T$. Predicting recovery stresses is therefore more difficult than simply constraining a tensile specimen and heating; measurements must be made dynamically, by enforcing a particular stress-strain path (to represent substrate deformation) and a particular strain-temperature path (to represent differential thermal expansion). As an example, Figure 4 shows the stresses developed during linear-elastic recovery experiments with no differental thermal expansion. As one would expect, greater substrate compliance leads to lower recovery stresses.

To date, applications described by the constrained recovery event have been the most successful. Some examples:
1. Tube and pipe couplings: Sleeves are fit around tube or pipe and shrunk to make extremely easy, fast and reliable joints[3,4].
2. Fasteners: Memory rings are used to fasten braided shielding to electrical connectors, to fix bearings on shafts, to fasten two dissimilar materials (such as a ceramic to a metal), etc.[5,6].
3. Electrical connectors: These contacts have zero insertion forces, high retention forces, and are very compact, making them well suited to high pin density connection problems[7].
4. Repair sleeves: Sleeves can be inserted into, or be shrunk onto, a pipe for repair or surface protection. Generally the memory metal presses a polymeric sleeve against the surface to be sealed.

Stable NiTi based alloys are generally preferred in these applications for several reasons: greater motion (critical to maintaining sufficient unresolved recoveries on normal substrate tolerances), high corrosion resistance, very low stress relaxation rates up to 350 C, high recovery stresses, and high ductilities. Until recently, constrained recovery applications required cryogenic alloys so that components would remain austenitic over a sufficiently large temperature range. Recently, however, it has become possible to introduce a large hysteresis in certain NiTi alloys, making it possible to store in martensite at room temperature, but remain austenitic below -55 after recovery[8].

WORK PRODUCTION

The third branch of application includes those that move against a resisting force - in the purest sense, a constant force. Visually, one can imagine a wire of memory metal, fixed at one end and a dead weight hanging from the other; the wire is stretched in the martensitic state, and will lift the weight when heated. The stretching can be done by applying an additional load, or by the weight itself if it is of sufficiently large mass. In the latter case, the system is self-resetting and the event can be performed repeatedly. Most applications of this nature are, in fact, cyclic, (self-resetting). Schematically, the event is modeled in Figure 5. Again we deform and unload just as in the first two cases. We then apply a constant load (σ_c) and begin heating. Key descriptors are the martensitic strain (ε_m), the austenitic strain (ε_a) and the recovery strain ($\varepsilon_r = \varepsilon_m - \varepsilon_a$).



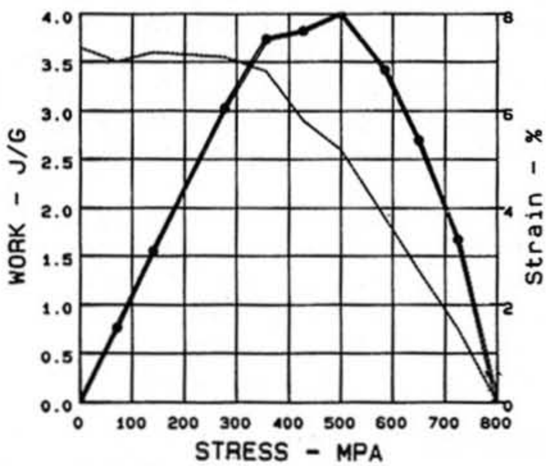


Fig. (above) 5 Schematic representation of the work generation event.

Fig. (left) 6 Work generation data from a hardened NiTi-based alloy. The work output is shown by the double line, the recovery strain by the dotted.

As pointed out above, subsequent cooling may again deform the memory metal, though not necessarily to e . The motion on cooling occurs at the M_s and M_f temperatures. A_s, A_f, M_s, and M_f all increase linearly with the applied stress at a rate given by the stress rate (dS/dT) of the material.

The shape memory event shown in Figure 5 does work since there is a displacement against a force. Figure 6 shows the work output of a NiTi shape memory alloy, low at small stresses since the force term is small, and low at high stresses because the alloy is unable to displace the high resisting force. Such curves are important in optimizing the use of memory material. Fatigue is also a key issue in most applications of this sort, making cyclic work outputs significantly lower than those shown in Figure 6. A fatigue failure in a common metal is defined as fracture. In a shape memory alloy, however, there are many modes of failure: a shift in the transformation temperatures, a reduction in the stroke or recovery strain, or ratcheting. Ratcheting is defined as a migration of the austenitic strain and progresses logrithmically with the number of cycles. Fatigue life can be influenced by heat treatment, alloy selection, and cycle design. Still a large alloy improvement is necessary before products such as a practical heat engine could be considered realistic.

Thermal and electrical actuators are the most common example of this shape memory event. Thermally actuated devices would include fire protection devices[9], anti-scald shower heads, window opening devices[10], and air conditioning vent controls. In these devices, the memory metal acts both as a temperature sensor and as an actuator. The choice of an alloy is not clear. The brasses and bronzes remain the only practical alloy for high actuation temperature (over 100 C) applications, but can suffer from stress corrosion, instability and ductility problems. In most of these applications, memory is directly competing with bimetals making cost another important alloy selection issue. NiTi has the advantage of the R-phase transformation, which is hysteresis-free (often called ARSME). This can be a big advantage in temperature control devices, but there are limitations: small recovery strains (roughly 1.%), narrow temperature range over which the effect is observed (-60 to 50), and succeptability to damage from excessive heating or cooling. Both memory alloy systems have significantly greater motions and work outputs than do bimetals, and can deliver their motion in any form (tension, bending, torsion, etc). Bimetals, on the other hand are hysteresis-free and exhibit linear motions over a wide temperature range.

Electrical actuation is the second branch of actuator application. Examples include remote louvre opening devices,[11] head lifters for disk drive units,[12] circuit breakers, door opening devices, etc. Direct electrical heating is used in these devices, so the purpose of the shape memory element is to replace a solonoid or servo-motor. The advantages of shape memory are that it is quiet, smooth, light, and compact; cost is potentially an advantage in some applications. NiTi based alloys have a large advantage over the Cu-based alloys due to their high resistivity and concomitant lower current requirements.

Motors and pumps have become popular discussion pieces, but are not close to becoming a commercial reality. Fatigue and inefficiency are the key factors. To maximize work output and keep costs low, high stresses must be designed into the device, causing unacceptable fatigue lifetimes. Moreover, the hysteresis and the high heat of transformation associated with the martensite-austenite transformations lead to low theoretical efficiencies (below 5%). There is at least one commercially available demonstration engine that uses the R-phase to austenite transformation instead of the martensite to austenite; this dramatically decreases losses from the hysteresis and the heat of transformation, but also decreases the memory strain and mechanical work output. Of all applications, the design of a useful heat engine appears most remote.

Finally, some applications require that the event is resetting with no load. This requires that the material has a two-way effect. Two-way effects are more easily obtained in the NiTi alloys but fatigue and stability issues are incompletely understood. Moreover, alloys and conditions exhibiting the greatest two-way effects are not the alloys with the highest capacity to do work. With our current knowledge base it is advisable to avoid relying on the material to self-reset.

PSUEDOELASTIC APPLICATIONS

The fourth type of shape memory event is isothermal and can be completely viewed in the stress-strain plane. When a shape memory alloy is deformed above A and below M , martensite is stress-induced; during unloading, the martensite again becomes unstable and the material reverts to its original shape. This is shown in Figure 7. S_l is the transformational stress upon loading (the stress needed to start the transformation to martensite) and S_u is the transformatinal stress upon unloading (corresponding to reversion back to austenite). In practice it is more useful to define these as the inflection points in the loading and unloading plateaus instead of using the standard offset methods. The hysteresis (S_l - S_u) can be related to the temperature hysteresis during loaded recovery (part III) through the Clausius-Clapeyron equation.

The benefit derived from the pseudoelastic event is energy storage. Listed below are the springback strains and stored elastic energies (as defined by $S\sigma d\varepsilon$) for high performance spring materials:

Material	Max. Springback Strain	Stored Energy
Steel	0.8%	8. Joules/cc
Beta Titanium	1.7%	14. Joules/cc
NiTi	8.3%	42. Joules/cc

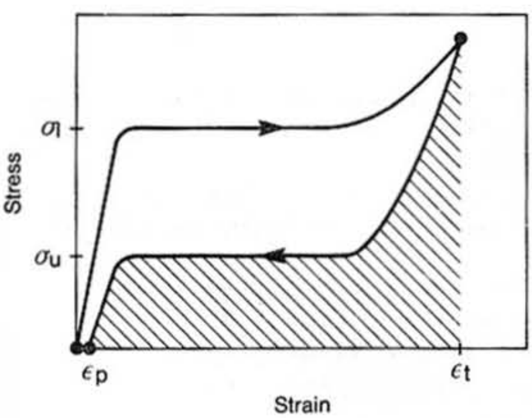


Fig. 7 Schematic representation of pseudoelasticity. The shaded region illustrates the stored elastic energy.

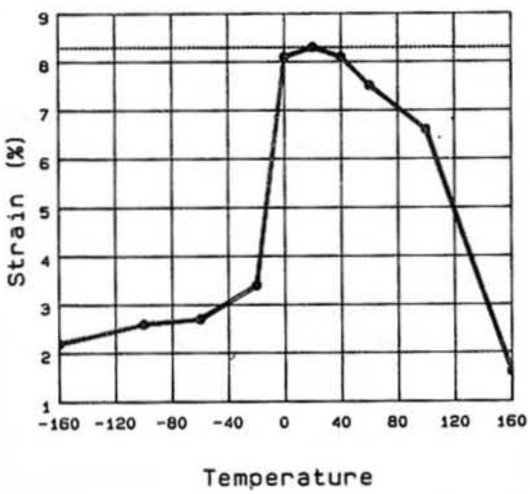


Fig. 8 Strain during unloading (e_t - e_f) of a pseudoelastic wire deformed 8.3% at various temperatures.

This, however, is in a single cycle, isothermal event - an ideal scenerio for shape memory and one seldom found outside the laboratory. Consider first the isothermal condition. Figure 8 shows the response of a pseudoelastic material to changes in temperature: stress continuously changes according to the material stress rate and the permanent set is low over a rather limited temperature range (80 degrees, in the case of NiTi). Though the cyclic properties of pseudoelastic materials have not yet been fully defined, there can be substantial "walking" or ratcheting as it goes through the transformation cycle.

Applications using the pseudoelastic event are therefore limited to low cycle, isothermal situations in which geometric or weight requirements require high reversible motions or elastic energy storage. The first such application was as orthodontic arch wire, where NiTi pseudoelastic wire reduces the need for adjustments and is more comfortable than conventional materials.[13] The body appears to be an ideal environment for pseudoelasticity since the temperature is well controlled and requires a biocompatible material. NiTi alloys are generally preferred to the Cu-based alloys for these applications due to their greater corrosion resistance, larger strains and longer fatigue lifetimes.

CONCLUSIONS

Four basic types of shape memory event have been discussed. It is clear that each is individual in terms of function and application potential. Together they encompass applications as different as springs for dentistry, large pipe couplings, circuit breakers, and wire for bras. Of all these applications only one, the pipe couplings, can be considered a proven "high volume" application. It remains to be seen which other areas will be successful.

REFERENCES

1. C.M. Wayman and K. Shimizu, Metal Science Journal, 6(1972), 175.
2. J. Perkins et al, in Shape Memory Effects in Alloys, Plenum Press (1975),273.
3. R.W. Benson et al, 15th National SAMPLE Conf, (October 1983).
4. J.D. Harrison, D.P. Hodgson, in Shape Memory Effects in Alloys, Plenum Press (1965), 517.
5. C. Dietemann, Machine Design, to be published, August 1986.
6. W. Schwenk and J. Huber, SAMPLE Quarterly 1 (1974),7.
7. R. Flot et al, "Shape-Memory Effect Alloys as an Interconnection Technology for High Density IC Packages", Elec. Conf (1984).
8. K. Melton and T. Duerig, ICOMAT 1986.
9. W. Vermeersch, "Proteus Shape Memory Alloy Applications" TMS-AIME 112th Annual Meeting (March 1983).
10. C.M. Wayman, J. of Metals, 6(1980), 129.11. D. Yaeger, Mechanical Engineering, 106(1984).
12. D. Yaeger, Design News, to be published.
13. G. Andreasen et al, Quintessence International, 9(1985),623.