Shape-memory phenomena

The martensite transformation, first noticed in steels in the nineteenth century, gives materials properties that depend on their histories and that allow them to recover earlier shapes after apparently plastic deformation.

Ahmad A. Golestaneh

If you deform an ordinary material below its elastic limit and keep it below its annealing temperature, it will return to its original shape once the load is removed. If the stress is too large, however, or if the temperature is allowed to become too high, the deformation becomes permanent, and the material retains no memory of its original shape.

For the past four decades, metallurgists have been investigating an intriguing class of materials that can retain a memory of earlier shapes.[1] Such a material can, for example, retain an apparently plastic deformation strain if it is held below a critical temperature, T_c. If it is heated above that temperature it returns to an earlier shape, a shape given it by appropriate heat treatment. Such shape-memory alloys absorb heat as they return to their earlier shape, converting it into mechanical work. The engine shown in figure 1 is based on this principle.

In addition to their technological interest, the shape-memory materials also have a considerable theoretical interest for metallurgists, chemists and physicists. Two broad classes of materials exhibit shape-memory phenomena: metal alloys and organic polymers. In this article I will focus on the metallic alloys. The polymers in many cases show behavior that is similar in kind but different in detail from that of the alloys. For example, while the alloys are soft and pliable below their T_c and much more stiff above T_c, the polymers are stiff when cool and pliable when warm. Alloys are thus made to "recover" their memory by heating, polymers by cooling.

Ahmad A. Golestaneh is visiting scholar at the Northwestern University Technological Institute, Evanston, Illinois.

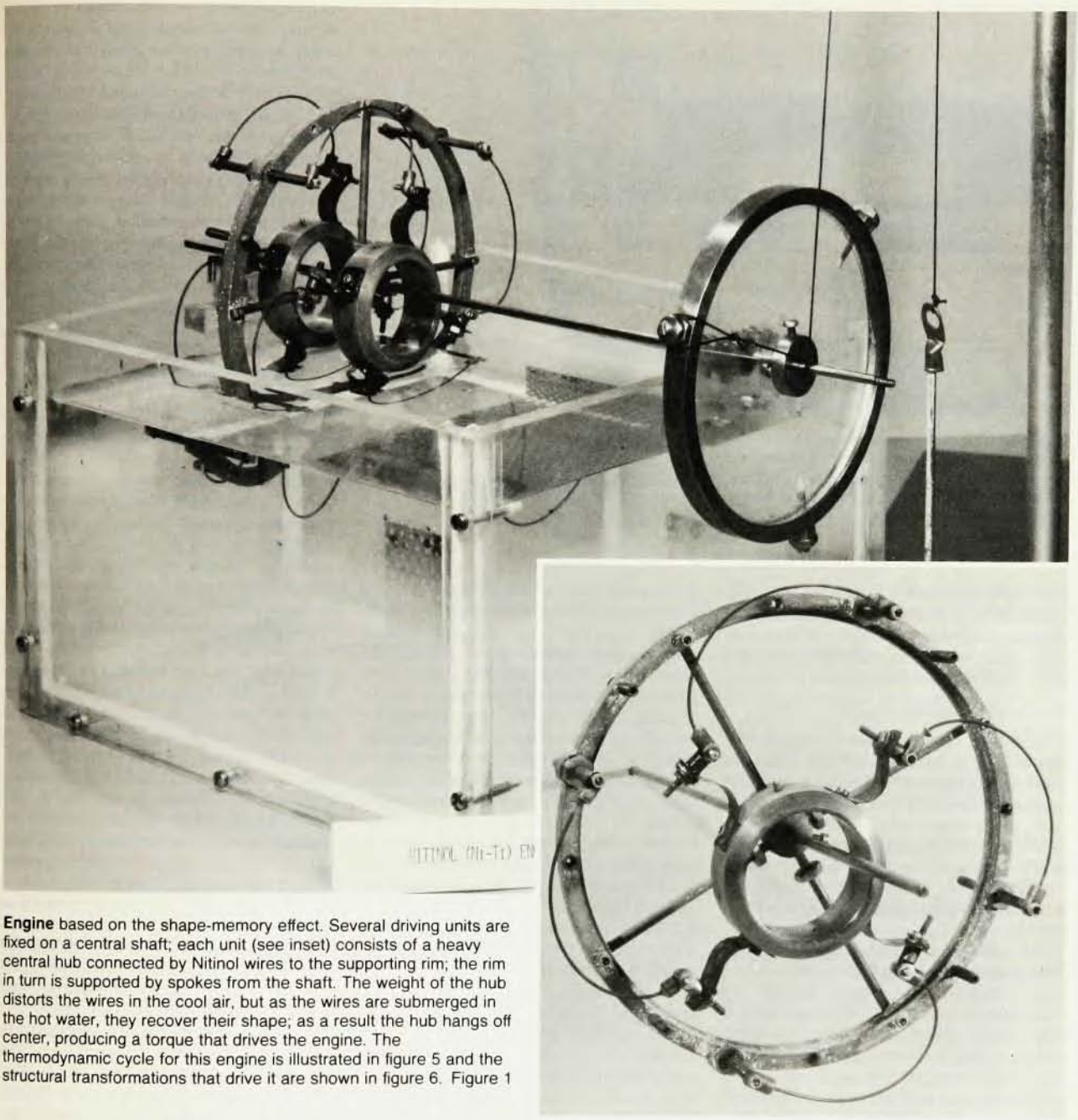
Alloys

At present there are about twenty elements that form shape-memory alloys, all from the central part of the periodic table, as shown in figure 2. These elements are from the same area from which most of the elements known to form superconductors are found. No ferromagnetic alloys apparently exhibit superconductivity or shape memory, but there are some superconducting alloys, such as Nb–Ti systems, that display a shape-memory effect, with T_c very close to their superconducting transition temperature.[2]

Shape-memory effects had been noted as early as 1932 for gold–cadmium, indium–titanium and copper–zinc alloys. Thomas A. Read's group in the mechanical engineering department at the University of Illinois investigated[3] these materials beginning in the 1950s. The first demonstration that the effect can be used to convert heat to useful work came in 1958: David Lieberman and Read demonstrated[4] that a gold–cadmium alloy could exert a force during its phase transition. A few years later, metallurgists at the US Naval Ordnance Laboratory (now named the Naval Surface Weapons Center) developed Nitinol, a nickel titanium alloy that has the highest efficiency (ratio of mechancial energy to heat energy) of any shape-memory alloy known. By varying the Ni–Ti proportions slightly from 50% (atomic) or adding small amounts (½% to 2%) of another metal (such as Fe, Cu or Zr) one can vary both the alloy's efficiency in converting heat to work and its critical temperature for the phase transition that leads to the shape change. The Nitinol alloys have several other attractive features such as high strength and excellent corrosion resistance; without iron they are practically nonmagnetic. The range of available T_c is from $-50°$ to 150 °C, and the tensile strength of a Nitinol wire can be as high as 1500 MPa with an elongation of 10–15%. Because of these features Nitinols have been used, or considered, for several applications, such as control and sensor devices, fastening mechanisms and surgical devices. They have also been used in self-erecting structures for space applications. For example, a large antenna was built from Nitinol and folded into a compact package for launch; it unfolded in space when it was activated by solar radiation.[5] One of the most publicized uses of Nitinol has been the "solid-state engine." There are several patents for this type of engine, all using Nitinol alloys. Figure 1 shows one such engine, which I designed[6]; it is simple in concept and is, as far as I know, the most efficient of the current designs. I should note, however, that no solid-state engine with commercially acceptable specifications is yet available.

Some investigators have probably overestimated the usefulness of shape-memory alloys in technological applications, while others have taken a pessimistic attitude towards this mat-

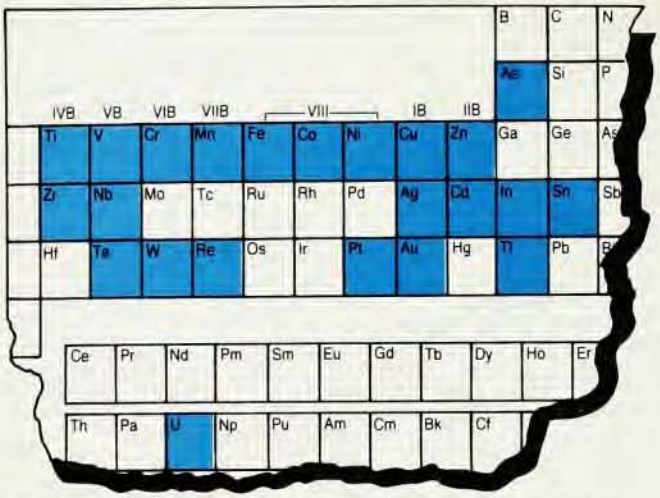
0031-9228 / 84 / 0400 62-08 / $01.00



Engine based on the shape-memory effect. Several driving units are fixed on a central shaft; each unit (see inset) consists of a heavy central hub connected by Nitinol wires to the supporting rim; the rim in turn is supported by spokes from the shaft. The weight of the hub distorts the wires in the cool air, but as the wires are submerged in the hot water, they recover their shape; as a result the hub hangs off center, producing a torque that drives the engine. The thermodynamic cycle for this engine is illustrated in figure 5 and the structural transformations that drive it are shown in figure 6. Figure 1

ter. The problem is related to a lack of the basic knowledge on the overall features of the materials. I will present a summary of the progress that has been made in recent years on describing the shape-recovery phenomenon on the basis of thermodynamics and continuum mechanics. However, there are other phenomena associated with shape recovery in these alloys that have not been well understood. For example, when one subjects a shape-memory alloy to a closed cycle of temperature changes about T_c, one observes a hysteresis loop with anomalous peaks in the strength of the specimen, its electrical resistivity, its elastic anisotropy, its internal friction, its magnetic properties and so on. The anomalous peak is always much more pronounced in the cooling curve than in the heating curve, as figure 3 shows. In addition, if one subjects a wire specimen of a shape-memory alloy to a temperature gradient, with one end above T_c and one end below, the wire exhibits a thermoelectric effect. The polarity of the voltage developed depends on the relative volume of the pieces of the wire above T_c and below T_c. The power produced also depends[7] strongly on the applied stress, that is, the material exhibits a piezoelectric behavior, although the effect is relatively weak compared to that exhibited by ferroelectric materials.

The source of these and the other phenomena mentioned above is known to be what metallurgists call the martensitic transformation, a phase transition that involves a change in crystal structure, and which occurs by nucleation and "diffusionless" growth. (The term diffusionless means that the atomic displacements are short on the scale of the interatomic distances and do not require thermal activation.) Such a phase change has also been called a "shear" or "displacive" transformation. It is beyond the scope of this article to discuss at length the martensite structure, a subject that has been discussed widely in the metallurgical literature.[8] A summary description of



Components of shape-memory alloys. This section of the periodic table shows (color) the elements that participate in alloys exhibiting shape memory. The table on page 70 shows the composition of such alloys. The heavy line indicates elements that are components of superconducting alloys. Figure 2

this topic is in order, however, not only because it is necessary for an understanding of shape-memory alloys, but also because it is important for the theory of phase changes in solids.

Martensite transformation

About a century ago, the German metallurgist Adolph Martens observed that steels hardened by rapid quenching had a needle-like appearance under the microscope. In his honor, the French metallurgist Floris Osmond in 1895 proposed to call this structure "martensite." The exact nature of martensite only became clear early in this century.

At ordinary temperatures, pure iron has a body-centered cubic lattice; the structure is called α-iron, or ferrite. At high temperatures (above about 900 °C) the stable structure is face-centered cubic, called γ-iron or austenite. Like most other useful forms of iron, steel contains some carbon, usually between 0.1% and 2%. Carbon is readily soluble in austenite, but it is not soluble in ferrite. When steel is cooled slowly, the solution separates into two components, ferrite and cementite (Fe_3C), with structures that depend on the history of cooling. When it is cooled quickly, however, the solution of carbon in γ-iron remains as a metastable structure. This hard and brittle material is martensite. It is, essentially, a supersaturated (or supercooled) solution of carbon in austenite.

Since 1929, investigators have reported similar structures in some nonferrous alloys, and even in pure metals, but without the famous hardness characteristic of martensite in steels.[8] For convenience, all these structures were also called martensites, despite the differences of properties found among these structures and the martensite in steels.

In current usage, then, martensite refers to a family of structures that have the following features:

Physical appearance: Under the microscope one generally sees needle-like edges, in parallel bands and sometimes interpenetrating each other at nearly 60°; figure 2 shows some typical structures.

Crystal configuration: Usually the structure is distorted body-centered, hexagonal close-packed, or face-centered cubic; sometimes, however, one sees distorted cubic, orthorhombic, or even monoclinic crystals.

Shape deformation: On a macroscopic scale, the transition to a martensite phase creates some surface upheaval or "tilting," often called "surface relief," which can be observed on the flat polished surface of the specimen, as seen in figure 3.

Metastability: The martensite structures are metastable—they are generally formed by quenching—and this is particularly noticeable in interstitial alloys such as carbon steels. In fact, martensite structures decompose to more stable structures over times ranging from seconds to several thousand years, depending on the nature of the material, and the temperature.

Hardness: In general, martensite materials are not hard; the outstanding hardness of martensite in steels arises because the interstitial C or N atoms not only contribute to lattice distortions but also block the motions of dislocations.

The transformation from parent to martensite occurs by nucleation (mostly heterogeneous) growth either athermally or isothermally or by a stress applied to the specimen. As one cools a specimen, the athermal parent–martensite reaction occurs spontaneously over a temperature range starting at a temperature denoted by M_s and completed at the lower temperature M_f. When a specimen of martensite material is heated it begins to revert to the parent form at a temperature A_s; the transition is completed by an upper temperature A_f. The isothermal parent–martensite transition occurs at a temperature M_t that is very nearly the average of the athermal temperatures, that is, very nearly $\frac{1}{4}(M_s + M_f + A_s + A_f)$. In the case of the shape-memory alloys, the critical temperature T_c that I mentioned earlier is roughly the same as M_t. All these temperatures depend on the applied stress.

The nucleation occurs at some "preferred" sites and not at random. In repeated cyclings between parent and martensite, martensite plates are frequently nucleated from the same sites—provided the specimen temperature does not greatly exceed A_f. No satisfactory knowledge of the structure of these sites is available yet.

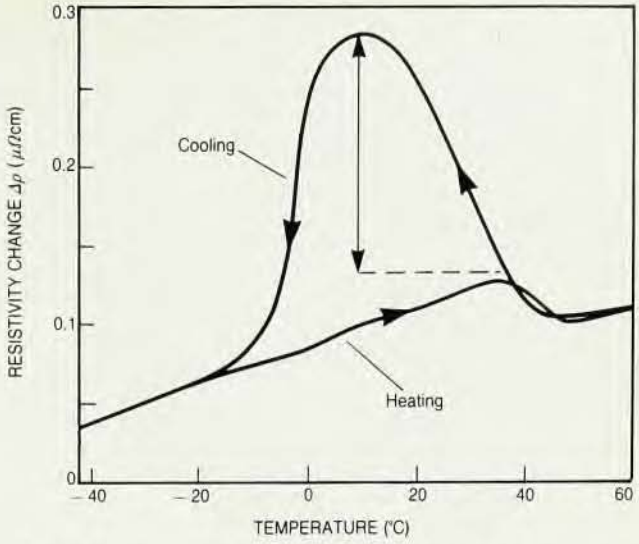
During the transition, the interface between the martensite and parent phases remains invariant in orientation while following a definite growth direction. This type of growth is realized by a combination of homogeneous and inhomogeneous lattice shears and some changes in lattice dimensions. Typically, the martensite crystal contains various kinds of fault layers, which lead to the "shape" deformation mentioned above plus small volume change.

Because the growth is diffusionless, the composition and the state of ordering (or disordering) of the parent and martensite phases are the same.

The martensite plates have an internal structure that depends on the kinetics of nucleation. What is called "burst" martensite grows with a speed approaching the speed of sound in the specimen—usually with some audible click—for about 10^{-7} sec, then thickening or growing edgewise; its structure is dominantly twin bands. (Such twin bands form when two parts of a crystal undergo different degrees of deformation resulting in one part being the mirror image of the other part.) When the burst martensite stops growing, another form, lenticular martensite, is formed; this grows at a much lower rate, about 0.01–0.1 m/sec. The internal structure of this martensite is dominated by slip and stacking fault layers. It should be understood that like most phase changes in solids, the martensite transition ends asymptotically with time, becoming more and more sluggish as it nears completion.

There are two principal means for inducing a martensite transition: quenching or stressing the sample. They differ somewhat in structure and properties. (See figure 4.) Stress-induced martensite, for example, has fine, sharp slip layers oriented along the applied shear stress; quench-induced martensite is isotropic, with faults oriented at random. Because the grains in stress-induced martensite are rotated into a common orientation, one can sometimes make use of repeated parent–martensite transitions to convert a polycrystalline material into a single crystal.[9]

An important difference between the stress- and quench-induced martensites is that they release different amounts of heat in the transition to the parent form. We will see later that it is the difference in these two heats of transition that plays a role in the shape-memory heat engines.



Resistivity change in Nitinol wire under zero stress. The wire was heated and cooled at a rate of about 1 °C/min in a water bath. Note that the resistivity peaks at a different temperature in the two curves. Figure 3

Shape-memory behavior

As I have mentioned, there are about 20 elements that form the known shape-memory alloys. In most cases, the structure of the parent phase is a body-centered cubic lattice. The lattice may be in a disordered state, but more commonly it is in one of the ordered states denoted as β_1, β_2 or β_3—that is, of the Fe_3 type, the CsCl type or the $MnCu_3Al$ type of crystal. In Nitinols, for example, the parent lattice is generally of the β_2, or CsCl, type. The martensite phases, on the other hand, usually have a hexagonal close-packed lattice.

The rearrangement of atoms in the martensite transition gives rise to two classes of phenomena: mechanical and electronic. The mechanical phenomena, such as the changes in material strength and elasticity and the shape-recovery effect, are mainly due to the changes of cohesive forces among the atoms. The electronic phenomena, such as the ferroelectric effect, anomalies in the electrical resistivity and variations in the internal friction coefficient, are due to changes in the states of the outer electrons as well as concomitant changes in the electron–phonon interaction.

Because we are here concerned chiefly with shape memory, we shall concentrate on the mechanical aspects of martensite. As I have noted, the structure and properties of martensite at low temperatures depend on how the material was formed: quenched from high temperatures in the absence of stress or formed under some stress. Figure 5 shows a phase diagram of a typical Nitinol alloy. There are four important regions in the diagram.

▶ At high temperatures and most values of stress, the alloy retains the structure of the parent material, albeit altered by the supersaturation of the lattice. I shall denote this structure as β.

▶ At high values of stress, the alloy assumes the lamellar, anisotropic structure of stress-induced martensite. I shall denote this structure as β'.

▶ At very low values of stress and at temperatures below T_c, the structure is that of quench-induced martensite. I shall denote this structure as γ'.

▶ At intermediate values of stress and temperatures below T_c, the alloy contains a mixture of the isotropic, uniform γ' configuration and the anisotropic β' configuration. I shall denote the mixed state as β''.

Because of its structure, pure γ' is the softest of the four, and, because of its stress-induced dislocations misfits, β' is the hardest. Note that while γ' can exist only below T_c, β' exists both below and above T_c. It is this difference in thermal properties that makes possible an engine based on shape memory.

Consider a sample of wire with a pure γ' structure. Because it is soft, it is easy to deform it; the deformation, of course, stores strain energy in material as dislocations, misfits, and other defects around and within each crystal grain. Because of the applied stress the structure becomes a mixed, β'' structure; at a temperature below T_c the fraction of β' and γ' in the resulting material is determined by the stress that was applied. The alignment of the β' crystals is, of course, along the direction of the applied shear.

Now suppose the deformed specimen is brought to a temperature above T_c at zero applied stress. Because of the change in the free energies of the various constituents, both the γ' and β' components of the mixed structure transform to the parent β structure. Furthermore, the increase in temperature results in an increased dislocation mobility and dissociation, so that the material can release the stored strain energy. Essentially, the material rebounds as if it were an ordinary deformed elastic material whose applied stress has just been removed. The strain release is more or less along the direction of the original strain, because the β' component of the deformed martensite is oriented in that direction. Note that the deformation of the specimen and the shape recovery are both associated with the stress-induced β' structure. Because its crystals are oriented randomly, the quench-induced γ' structure contributes to the shape memory not directly by conversion to β, but only indirectly via conversion to the β' structure by stress above T_c.

If, instead of releasing the stress on the deformed specimen, we heat it to a temperature above T_c while it is still under stress, the response is similar, but differs in details. Because at nonzero stress the fraction of β' and γ' components of the β'' structure varies with temperature, some of the γ' changes to β'; this releases an amount of heat proportional to the difference ΔH in the heats of formation of the two structures, the total quantity of heat released being determined by the amount of material changing from γ' and β'. Only subsequently does the material change to the parent β structure if the temperature is sufficiently high. Associated with this change in structure is, as before, a change in the elastic modulus and the strength of the

system. If the energy released by the change in structure is large enough to release the strain, the sample recovers its original shape against the outside force. We can see this quantitatively as follows. Denote the fraction of γ' that is changed to β' by B and the difference in the Young's modulus between the β' and β structures by ΔY. The amount of energy available from the transition $\beta'' \rightarrow \beta$ is

$$B\rho\Delta H$$

where ρ is the density of the material. The energy to deform the specimen by an average strain ϵ is

$$\tfrac{1}{2}\epsilon^2\Delta Y$$

This is, essentially, the work that can be done by the system. Its maximum value is just $B\rho\Delta H$.

One can also calculate[10] the speed with which the shape recovery takes place, based on the theory I have sketched here. However, in practice the speed depends sensitively on the material composition, the defects present, and how the external energy is supplied to the specimen. I have observed the shape recovery of Nitinol wire by flame heating, electrical heating and laser heating; none of these methods produce a speed equal to that achieved by immersing the wire in hot water: Apparently immersion provides the most effective means of heat transfer and heat distribution in the specimen.

The solid-state engine

The effects I have just described form the basis of a solid-state engine; we need only a closed loop in the stress–temperature diagram to complete the analysis. Such a loop is shown in figure 5 and diagrammed in figure 6. We start, say, with an unstressed sample at a low temperature T_L and apply a stress; some of the γ' structure converts to β', absorbing heat. We transfer the sample to a high-temperature resevoir, T_H, where it reverts to the parent structure β; it absorbs heat and converts some of the heat released by the transformation to work—this is the amount $B\Delta H\rho$ mentioned above. The associated shape recovery results in removing the stress from the sample, and the unstressed sample is returned to the low temperature, where its structure is quenched again to the γ' configuration, releasing the heat H to the low-temperature resevoir.

The coefficient of heat-to-energy conversion is defined as the ratio of work, $\frac{1}{2}\epsilon^2\Delta Y$, to the heat released by quench-induced martensite formation, ρH:

$$\alpha = \tfrac{1}{2}\epsilon^2\Delta Y/\rho H$$
$$= B\Delta H/H$$

One can use an argument analogous to the sort of argument used to derive the Clausius–Clapeyron equation in classical thermodynamics (that is, consider a cyclic process in which the system is carried through a phase transition at two different temperatures) to show that α has an upper limiting value given by

$$\alpha_{max} = \ln(T_H/T_c)$$

for values of ϵ below about 3.5%. The thermodynamic efficiency is given in terms of a parameter called the "characteristic temperature" of the alloy

$$h = H/C$$

where C is the average specific heat of the alloy. I will not present the derivation[6,11] here, but will only mention the results. The peak efficiency is approximately

$$\eta = \alpha h/(T_c + h)$$

and the best engine performance takes place when the hot and cold temperatures are given by

$$T_L T_H = T_c^2$$

the efficiency is, as it must be, less than the Carnot efficiency

$$\eta < 1 - T_L/T_H$$

The speed of the engine, it turns out, is determined by the coefficient α mentioned above.

The solid-state engine shown in figure 1 represents an application of these principles.[6] It contains several driving units fixed coaxially on a shaft. Each unit consists of a massive hub that is suspended by a number of nitinol wires connected to a rim; the rim in turn is supported by spokes fixed on the shaft. The engine in figure 1 has hubs suspended by four 1-mm-diameter Nitinol wires with a T_c of 33 ± 7 °C. Each wire is trained to have a curved shape. The shaft is installed horizontally on the top of container in which water is maintained at T_H between 50 and 75 °C. The set-up is such that about half the wires of the units are in the air at room temperature (T_L is 20–24 °C), while remaining wires are in the hot water. The wires in the air are deformed easily by the weight of the hub, as they have β'' structure. However, because T_H is above T_c, the wires in the hot water recover their trained shape as the β'' structure reverts to the parent β structure. The hub is thus pushed off center, producing a torque proportional to the hub weight and its off-center distance to the shaft. As the shaft turns, the cold wires that are deformed by the off-center hubs enter the hot water, straighten out, and the process repeats. The engine revolution rate depends on the coefficient α, as well as on power losses and the load applied to the shaft. The engine shown in figure 1 has delivered a few miliwatts and reached 52 revolutions per minute under zero load, while operating between 75 °C and 20 °C.

Nitinol alloys have a characteristic temperature h in the range 17 °C to 24 °C, the largest values known. The heat of martensite formation, H, for nitinols is about 1.5 to 2.2 cal/g. For the alloy used in the engine shown, h is about 24 °C; the maximum temperature at which martensite forms, T_m, is about 80 °C. The theoretical maximum efficiency given above is about 5 to 7%; the maximum measured efficiency of 4.7% compares quite favorably with the calculation.

The calculation shows that an increase in efficiency would require an increase in h or a decrease in T_c. Materials with such properties may be found among ternary alloys of nickel, titanium and zirconium, which I have found to show a higher shape-recovery strength than the binary Nitinols. Exploiting these will, however, require some research, as h and T_c are related to each other through material and structural parameters, and at present we have no formalism for these relationships.

A shortcoming of the shape-memory elements in the solid-state engine, or any other repetitive device, is the gradual loss of the shape memory as a result of overloading or overheating. If the engine is overloaded, the shape recovery cannot occur completely—or at all, depending on the load. The wires thus gradually adopt whatever shape the load imposes on them, and the engine output is reduced. To avoid this, the wires must be occasionally "retrained" to the desired shape.

Theories of martensite

Theoretical work on martensite and the martensite transformation has progressed rather slowly because the phenomena have been uncovered only gradually. The experiments are delicate and often difficult, and many materials must be investigated to identify features characteristic of martensite; the data are often difficult to interpret because of hidden factors or phenomena unrelated to martensite.

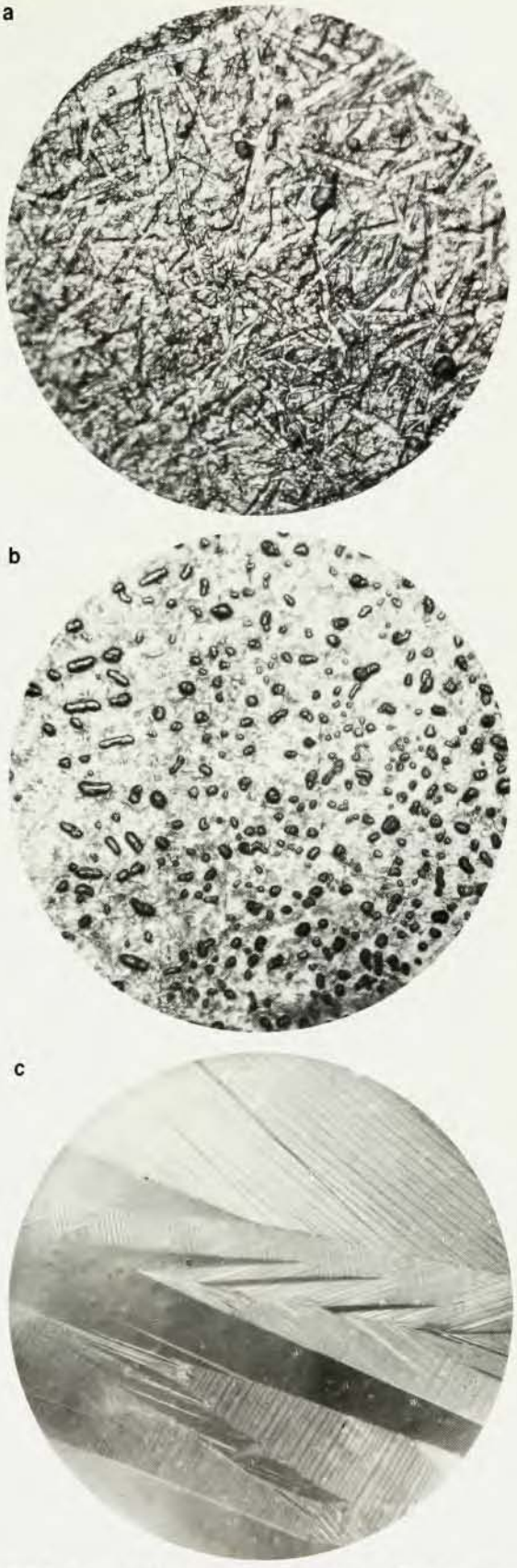
The descriptive crystallography of martensite structures is fairly well developed, although more complicated, and precise experiments are still needed. Kinetic theories of the martensite transition, however, are not yet fully satisfactory, even on a phenomenological basis. While some aspects can be adequately described in terms of temperature, stress, strain and so forth, there are many problems that still need basic study. Most features of martensite and the martensite transformation are now understood qualitatively, but not yet in quantitative terms. The remaining problems can be grouped under three headings: nucleation and growth, martensite formation, and

electron, lattice and phonon effects.

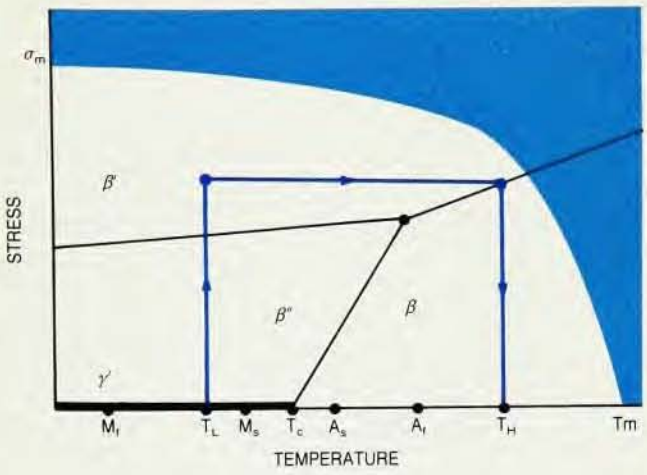
Nucleation and growth. How is the nucleus for a martensite crystal formed and what is its structure? What are the preferred sites referred to above? How does the martensite nucleation rate depend on temperature, stress and time?

It is known that thermal energy alone is insufficient to produce a transition as spontaneous and fast as seen in the martensite transformation. However, it is not known how—or how much—strain energy is accumulated at a given temperature in the nucleus to trigger the martensite formation—either "burst" or "slow"—in a given specimen. Another question is why the burst and slow martensites have, respectively, the lath and lenticular shapes rather than, say, a spherical shape. The interface between the parent and martensite is observed to be irrational in the sense that it has a high crystallographic index. This and other geometrical features such as the volume and shape changes are all consequences of the same rules (or criteria) that dictate that the transformation take place along a definite crystal growth direction as well as dictating the macroscopic invariance of the parent–martensite interface. Finally, we need to understand the conditions for an alloy to have shape memory and how the critical and characteristic temperatures, T_c and h, are related to the material parameters. We would also like to understand why the martensite transition in shape memory alloys is accompanied by an appreciably larger change in Young's modulus but a smaller change in volume than for other materials.

Models for martensite formation. Clearly the formation of a martensite crystal with some internal structure and its accomodation in its surrounding material require various kind of dislocation. Moreover, during the transition, the parent and martensite crystals meet at their interface, the "habit plane," through many glissile dislocations, all within an interfacial zone approximately 20 Å thick. It is known that screw dislocations are very much involved in the mechanism of the transition. Since 1951, several authors have proposed specific dislocation models for the interfacial zone and its advance along a definite growth direction.[12] Another model for the transformation is based[8] on the fine stack of twin bands mentioned above. Some authors have used some combination of translation, rotation and deformation of the lattice to show graphically or mathematically how the martensite unit cells can be obtained from the parent unit cells. These descriptive models are interesting, but they are associated with some specific modes of the crystal formation



Martensite structures: (**a**) Quench-induced γ' martensite in a Nitinol wire. (**b**) Stress-induced β' martensite in the same specimen (viewed perpendicular to the β' orientation). Both views are of sections. (**c**) Surface view the β' structure in a Cu–Al–Ni alloy; note the v-shaped plates superposed by the surface relief lines. The micrographs are about 400X magnification. (Photo **c** courtesy of C. Marvin Wayman, University of Illinois.) Figure 4



Stress-temperature diagram for a typical Nitinol shape-memory alloy. Martensite structures are only stable (or, rather, metastable) within the area bordered by the colored region. There are three fundamental structures involved: the parent structure (stable also in the colored area) β; stress-induced martensite, β'; and quench-induced martensite γ'. In the region marked β'' the stable structure is a mixture of β' and γ'. The colored box represents the thermodynamic path of the solid-state engine of figure 6. Figure 5

that disregard atomic forces. The fact is that as yet we cannot know exactly what occurs in the transformation zone and how this zone advances in the specimen matrix; all we know for sure are the features and rules described above. No dynamical theory has been developed yet to support these models. A theory based either on the crystal lattice changes or on a dislocation mechanism would require extensive work, which would, among other things, involve the evaluation of the interatomic cohesive energy and a good dislocation theory. Such micropic theories themselves need fresh development beyond what we know from continuum elasticity and classical thermodynamics. One related subject is the evaluation of what is called the "screw dislocation core." The instability or stability of this zone, which depends on the crystal composition and configuration, affects many properties of the specimen, thereby affecting the martensite transition in many ways. So far, however, the screw dislocation core has been either neglected or treated poorly by continuum elasticity theories rather than being evaluated from correct expression for the interatomic cohesive energy.

To avoid some of these difficulties, I have recently started to develop a phenomenological theory of martensite growth, using a formalism akin to the one used for the strong interactions in nuclear or particle physics. Briefly the method may be outlined as follows. We note that the diffusionless growth takes place within a thin transformation zone mentioned above. Within this zone, martensite nuclei appear simultaneously as soon as the specimen temperature is brought below the critical temperature. This zone can be considered as a "black box," as atoms there are in a highly disordered state, and their motions cannot be followed. The one thing we know is that at a temperature T_c the parent structure is unstable; to reduce the free energy, a number N_p of the parent cells disappears while N_m martensite cells appear. Usually there is a small volume reduction involved in the transition, so N_m is greater than N_p. The initial and final stages of the reaction are characterized by the matrices ψ^p and ψ^m, whose components are proportional to the lattice vectors of the parent and martensite crystal cells. There is a nonlocal correspondence between the parent and martensite states, which can be expressed by an S-matrix:

$$\psi^m = S\psi^p$$

where S must satisfy all features of the reaction. For instance, at T_c both the martensite-to-parent and the reverse transitions must have equal probability; this and the reversibility of the transition require S to be a hermitian and unitary operator at T_c. Also S must satisfy conditions imposed by the conservation of energy and momentum, as well as other crystallographic conditions. The S matrix should also take account of the symmetry of the lattices. We can represent the essential differences between the ψ^p and ψ^m states by the elastic anisotropy A_e and the generalized Cauchy parameter D; these are defined in terms of the crystal elastic constants C_{ij}:

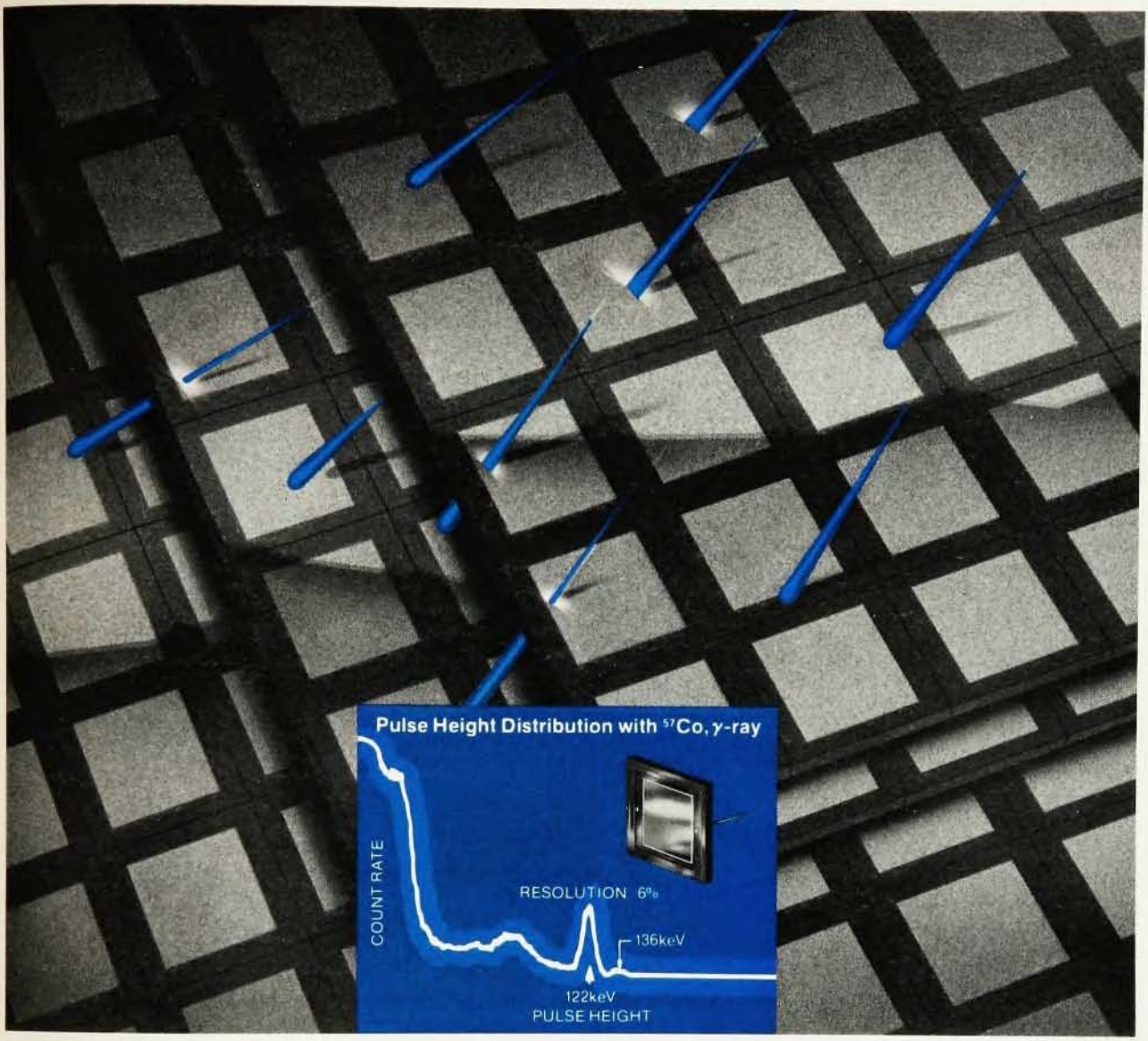
$$A_e = {}^1\!/_2 C_{44}/(C_{11} - C_{12})$$
$$D = C_{12} - C_{44}$$

($D = 0$ corresponds to a perfect cubic crystal.) The change in D during the transition is a measure of the crystal distortions in the martensite crystal.

The S matrix can be separated into two terms: one that operates during the very short time to produce burst martensite; the other creates martensite with a relatively slow and diminishing rate. The time dependence of the S matrix is given by an equation of motion derived from a Lagrangian that contains the cohesive energy and the restrictive energy of dislocation pile ups.

Electron and lattice effects. I have mentioned that the states of outer electrons and electron–phonon interactions are primarily responsible for the anomalies and hysteresis loops one sees in specimen properties during the martensite transition. An understanding of these phenomena is essential in the development of a general theory of the diffusionless phase change in solids. Yet they have received relatively little attention. To illustrate the problems, I will summarize some limited work I carried out[7] several years ago, using a thin Nitinol wire; the critical temperature of the specimen was 33 ± 7 °C. During the transition from the γ' structure to the β structure, its Young's modulus increased by a factor of 3 to 4. This is the sample whose resistivity is plotted in figure 2. An interesting feature is that the anomaly in the resistivity depends on the degree of anisotropy and perturbation of the lattice distortions during the martensite transition. That is, the resistivity of the sample measured at a temperature T is larger when the sample is quenched directly to T from the parent state than when it is quenched to a temperature far below T and then heated to reach T. An explanation for this difference is that the newly formed martensite produces many distortions and anisotropies in the electron–phonon scattering amplitudes. Subsequent heating removes these anisotropies and distortions. The anomaly is eliminated by heating and cooling and the specimen repeatedly, each time reducing the spread in temperatures. The electron–phonon scattering is reduced to its natural level when the upper temperature has reached T_c.

Other phenomena arising from the martensite transformation are the anomalous changes in the internal



Now, for the first time, high energy resolution from PIN silicon photodiodes used as nuclear counters

The new S1723 silicon photodiode provides the low junction capacitance and high shunt resistance needed for high speed response and low noise. The UV response is particularly suitable for use with BGO and other scintillation crystals. A sensitive area greater than 100mm^2 is provided in a very compact package.

This new detector is less than 3mm thick compared with 60mm or more for PMT's. Thus, the S1723 occupies a small fraction of the space, and costs about half as much when used with today's solid state amplifiers.

Applications include scintillation detection in the fields of high energy physics, medical diagnostics and industrial instrumentation. Hamamatsu engineers will modify the S1723 or other detectors to meet your needs.

Call or write for product bulletin and prices.

HAMAMATSU

HAMAMATSU CORPORATION • 420 SOUTH AVENUE • MIDDLESEX, NEW JERSEY 08846 • PHONE (201) 469-6640

International Offices in Major Countries of Europe and Asia.

Circle number 21 on Reader Service Card

friction, also known as the "quality factor"; a change in the velocity of the acoustic or elastic waves in the specimen; changes in specific heat, crystal elastic factors C_{ij} and so forth. All these changes are related to each other and to the interactions among phonons, the lattice and the electron states in the specimen.

Future developments

Further development of shape-memory alloys is strongly dependent on the understanding of martensite transition and the structural properties of martensite. This subject, now about 100 years old, still persents many unsolved problems that need fresh physical insight and input. The kind of questions raised here about nucleation and growth, formation and stability of the screw dislocation core, anomalies in the specimen properties and so forth, all depend on what occurs over distances of a few atomic diameters. Thus they require in principle a quantum mechanical treatment, which must, however, be consistent with thermodynamics and continuum mechanics.

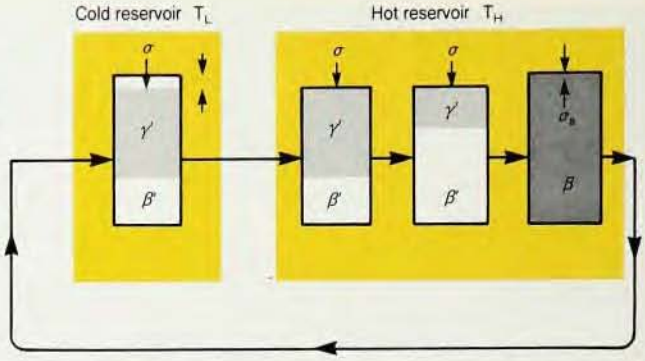
Before dealing with a complete atomic theory, however, some preparatory work is needed. For instance, I suspect that the screw dislocation core is an essential part of the preferred sites for martensite nucleation.[12] To calculate the internal and surface energies of the screw dislocation core we need to know[3] the cohesive energy E_c in a given crystal matrix as a function of temperature, stress, and so forth. The growth theory based on the S-matrix formalism also depends on E_c, as well as other parameters. The strain energy due to the change Young's modulus during the transition depends on the change of E_c as the reaction occurs. Thus a microscopic theory of the martensite transition requires an evaluation of E_c for perfect and distorted crystals in pure metals and alloys: an extremely difficult task.

Furthermore, a satisfactory theory of phase change in solids must take account of the particulars of the specimen structure as well as the statistical nature of the process. Undoubtedly the rules involved in the nucleation and growth dynamics include conservation of energy and momentum, as well as minimization of the free energy, which contains the cohesive-energy terms. (The cohesive-energy terms have not been taken into account analytically in the existing nucleation formalism.) Other criteria may also be involved in martensite nucleation.

As I mentioned, martensite transition are observed in materials other than metals and alloys: $BaTiO_3$, $SiTiO_3$, and organic polymers, for example. The details of the behavior of these systems are as yet less well understood that the alloys I have described.

Perhaps the situation can be described in the words of the Persian poet Hafez, writing in 1350:

> The search for truth seemed easy at first, but ran into difficulty. It is dark, and a deep sea, with turbulent undercurrents and surmounting waves, stands between us. Those in a calm harbor, how can they know what we must go through?



Solid-state engine. A shape-memory alloy is quenched to temperature T_L, causing it to have a β'' (mixture of β' and γ') structure; it is deformed by an external stress σ. Transferring it to the hot reservoir T_H changes the structure in the manner indicated to the parent β. The structural change is accompanied by an expansion and change in shape that does work against the external stress. Figure 6

Shape-memory alloys

Composition (atomic %)	Structure: Parent	Structure: Martensite
Au, (49–51) Cd	B2	Trigonal
Au, (95) Cd	CsCl, β_1	γ_1'
Ag, (44–49) Cd	B_2	2H, B19
Cu, (28) Al, (3.9) Ni	DO_3, β_1 (Fe_3Al) type	(2H) γ' Cu_3Ti type
Cu, (23–28) Au, (4.5–4.7) Zn	Hensler type	M3R
Cu, (15) Sn	DO_3 (Fe_3Al) type	(2H), (6R) Cu_3Ti type
Cu, (38.5–40) Zn	B2	M3R Ll_0 (3R)
Cu, Zn, (Si, Sn, or Al)	B2	3R and M3R
Ni, (49–51) Ti	B2	γ_1' and deformed β_1'
Ni, Ti, (1–2) Z*	B2	(B19)
Ni, 37 Al	B2	Ll_0 (IR)
In, (18–23) Al	fcc	TCT
Fe, Pt	(Fe_3Pt), γ	α'
Nb, Ti†	bcc	
Nb, (82) U†	bcc	
Nb_3Sn†	A15	
V_3Si†		
U, Mo	bcc	

For crystallographic notation see *Progress in Materials Science*, J. W. Christian, T. B. Laski, eds. volume 18 (1974).
*Z can be Co, Fe, Mn, Au, Zr and so forth
†Superconducting alloys are believed to exhibit shape recovery (see reference 6).

I would like to thank William J. Shack and Fred A. Nichols of Argonne National Laboratory and Jerome B. Cohen of Northwestern University for their comments and helpful discussions. I dedicate the article to the memory of A. Martens.

References

1. *Shape Memory Effects in Alloys*, J. Perkins, ed. Plenum, New York (1975).
2. D. S. Easton, C. C. Koch, in reference 1, page 431.
3. A. Olander, J. Am. Chem. Soc. **54**, 3819 (1932); L. C. Chang, T. A. Read, Trans. AIME **189**, 47 (1951).
4. D. S. Lieberman, in reference 1, page 208.
5. C. M. Jackson, H. J. Wagner, R. J. Wasilewski, *55-Nitinol—The Alloy with a Memory*, NASA-SP-5110, National Aeronautics and Space Administration (1972).
6. A. A. Golestaneh, Acta Metal. **28**, 1427 (1980). The engine in this paper, and in figure 1 here, is US Patent number 4302939 DOE/TIC/WG-82/181.
7. A. A. Golestaneh, Proc. Int. Conf. Martensitic Transformation, ICOMAT 1979, page 715.
8. See, for instance, J. W. Christian, *The Theory of Transformation in Metals and Alloys*, Peragmon, New York (1965); also *Phase Transformation*, Am. Soc. Metals, Metals Park, Ohio (1968).
9. T. Saburi, C. M. Wayman, K. Takata, S. Nemo, Acta. Metal. **28**, 15 (1980).
10. A. A. Golestaneh, Res Mechanica Lett. **1**, 397 (1981).
11. A. A. Golestaneh, J. Appl. Phys. **49**, 1241 (1978).
12. F. C. Frank, Acta. Metall. **1**, 15 (1953). For crystallographic models of the martensite, see also references 8 and 9 and G. M. Michel, R. Sinclair, Acta Cryst. **B37**, 1803 (1981).