JOURNAL OF MATERIALS SCIENCE **39** (2004) 385–399

Review
Thermo-mechanical fatigue of shape memory alloys

E. HORNBOGEN
Ruhr-University, Bochum D-44801, Germany
E-mail: erhard.hornbogen@rz.ruhr-uni-bochum.de

Shape memory alloys (SMA) produce the largest anomalous reversible strains of all crystalline materials: $\varepsilon_{SM} < 10\%$ in comparison to $\varepsilon_{FM} < 1.0\%$ and $\varepsilon_{FE} < 0.1\%$ (FM magnetostrictive, FE electrostrictive). They are considered as active materials for applications in which large scale motions are required or because of their high damping capacity for mechanical vibrations.
Their use is however limited if a large number of cycles is applied. A survey is given on fatigue phenomena in SMA:

1. classical mechanical fatigue—accumulation of defects, formation and growth of cracks,
2. shape memory fatigue—change of transformation temperatures, reduction or loss of memory, loss of pseudo elasticity or damping capacity.

Finally, it is shown how a modification of microstructure by thermo-mechanical treatments (TMT) affects fatigue resistance. In this context primary defects (as introduced by TMT) and secondary defects (introduced by mechanical or thermal cycling) have to be distinguished. In case of training for the two-way effect secondary defects are introduced voluntarily.

Nomenclature

$A \equiv \beta$	High temperature phase, austenite
A_s, A_f, A_p (K)	Austenite start, finish, peak of dsc-signal
b (m)	Burgers vector of any crystal
b_α, b_β (m)	Burgers vector of α, β
b^*_α, b^*_β (m)	$b^* < b$, defect in boundary
b_{lid} (m)	Defect by lattice invariant deformation
b_{ghost} (m)	Ghost defects in austenite
D (m^2s^{-1})	Diffusion coefficient
ε (l)	Any strain
$\varepsilon_{\beta\alpha}$ (l)	Transformational strain
ε_t (l)	Training strain
$M \equiv \alpha$	Martensite, low T phase
M_s, M_f, M_p (K)	Martensite start, finish, peak of signal
M_d (K)	No strain induced martensite above this temperature
N (l)	Number of cycles
N_f (l)	Number of cycles to fracture
$S_{\beta\alpha}$ ($JK^{-1}m^{-3}$)	Transformational entropy
S_{th} ($JK^{-1}m^{-3}$)	Thermal entropy
S_{st} ($JK^{-1}m^{-3}$)	Structural entropy
σ (Nm^{-2})	Any external stress
$\sigma_{\beta\alpha}$(Nm^{-2})	Transformational stress
$\sigma_{y\beta}$ (Nm^{-2})	Yield stress of β
T (K)	Any temperature
T_1 (K)	Low T range, transformed
T_2 (K)	Medium T range, transforming
T_3 (K)	High T range, stable austenite
ΔT (K)	Stress induced raise of M_s
T_c (K)	$T > T_c = 1/3\ T_m$ considerable diffusion
T_m (K)	Melting temperature

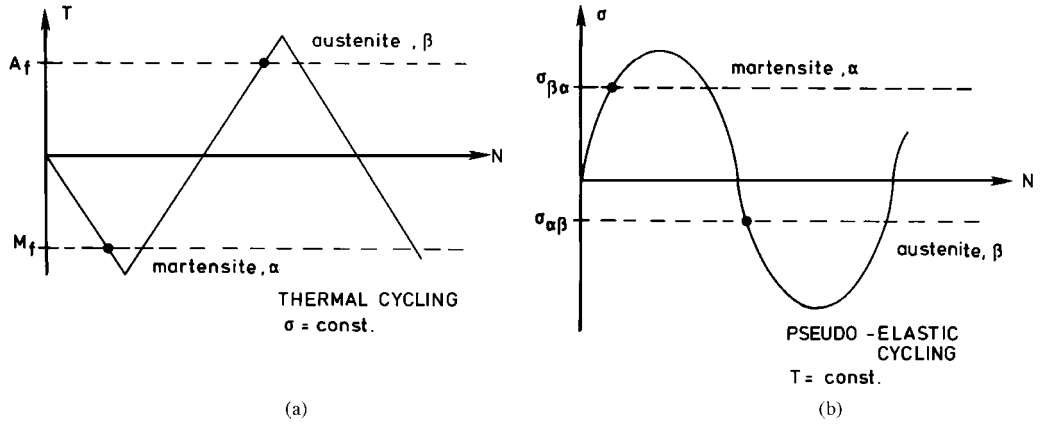
1. Definition of three temperature ranges: Stable austenite β, transforming $\beta \leftrightarrow \alpha$, fully transformed martensite α

Response of a material to cyclic loading can be classified by four combinations of structurally reversible friction and irreversible fatigue: Fatigue phenomena are due to irreversible localized formation and accumulation of structural changes. Internal friction implies a minimum or maximum dissipation of energy (heat). The phenomena are analogous to (external) friction and wear. Classical fatigue assumes a stable phase structure of the loaded material. Additional phenomena are to be expected, if phase transformations interfere. Of special concern is martensitic transformation. It can take place at low temperature, because it requires no diffusion ($D \approx 0$, Table I). High internal friction and stability against fatigue is a desirable combination which

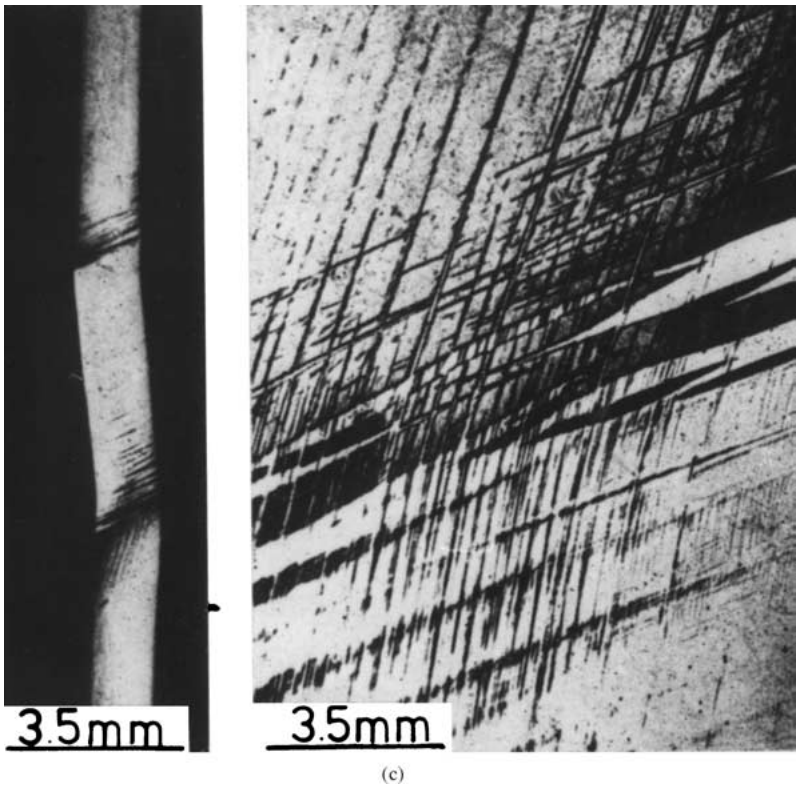
TABLE I Fatigue of SMA, temperature ranges (T/T_m ratio of any temperature T to melting temperature T_m: $0 < T/T_m < 1$)

$\frac{T}{T_m}$		T
	Creep-controlled fatigue (dislocations gliding and climbing) in austenite	
$3b$	Creep as well as precipitation or ordering controlled fatigue, stabilization of martensite	$T_3 \gg M_d$
$1/3$		T_c
$3a$	"Classical" fatigue of β (dislocation gliding, grain boundary embrittlement)	$T_3 > M_d$
	Fatigue affected by martensitic transformation following primary plastic deformation of β	$M_d > T_2 > M_s$
2	Fatigue by stress induced nucleation and motion of β/α_M-interfaces	$A_f > T_2 > M_s$
	Fatigue in ($\beta + \alpha_M$) mixture: motion of β/α_M-interfaces and reorientation of thermal martensite	$M_s > T_2 > M_f$
1	Fatigue in martensite: motion of $\alpha_M + /\alpha_M$-domain or twin boundaries: eventually: martensite $\Leftrightarrow$ martensite transformation	$M_f > T_1$
0		

$$T_m > T_{3b} > T_{3a} > M_d > T_2 > M_f > T_1$$



(a) (b)



(c)

Figure 1 Transformation cycles: (a/b) thermal and mechanical cycles, schematic and (c) CuZn, $T > M_s$, formation of an α (martensite) from a β (austenite) single crystal by external tensile stress.

could open shape memory alloys, and possibly steels, new fields of application. One-way martensitic transformation during cooling was utilized for more than 3000 years, as the phase transformation which is responsible for hardening of steel. Only in this century the complete reversibility of this diffusionless transformation was discovered [1]. Complete reversibility requires crystallographic reversibility of a cycle $\beta \leftrightarrow \alpha$, i.e., the high temperature phase β is restored faultlessly during heating of this martensitic structure (Fig. 1). Considerable shape changes have been reported in β-brass, if the thermal cycles are taking place under simultaneous mechanical stress (Fig. 2). Hardening by repeated stress-free thermal cycles has also reported [2] (Fig. 3).

Such transformation cycles can be induced by thermodynamical (variable temperature T) and/or mechanical energies (shear stress τ). Consequently three principle temperature ranges have to be defined for alloys which undergo martensitic transformation (or similar transformations which lead to dimensional changes: premartensitic R-phase, ferro-magnetic, ferro-electric [3]). The fatigue behavior differs depending on whether the high temperature phase ($\beta \equiv$ austenite) is stable: $T_3 > M_d$, transforming: $M_d > T_2 > M_f$; or fully transformed ($\alpha \equiv$ martensite): $T_1 < M_f$ (Table I). The transformation temperatures M_s depend on chemical composition and microstructure. The stress dependence of martensite start M_s is controlled by an equation analogous to Clausius Clapeiron equation:

$$M_s(\sigma) = M_s(0) + \Delta T(\sigma) \tag{1a}$$

$$\frac{dM_s}{d\sigma} = \frac{\varepsilon_{\beta\alpha}}{S_{\beta\alpha}} \tag{1b}$$

$$M_s(\sigma) = M_s(0) + \frac{\sigma_{\beta\alpha}\varepsilon_{\beta\alpha}}{S_{\beta\alpha}} \tag{1c}$$

$$\sigma_{\beta\alpha} \leq \sigma_{y\beta} \tag{1d}$$

The yield stress of β, $\sigma_{y\beta}$ provides the upper limit for stress-induced transformation by $\sigma_{\beta\alpha}$.

2. The evolution of lattice and surface defects in single- and poly-crystals

Table I summarizes different phenomena observed in the particular temperature ranges. More or less energy is dissipated as structural entropy S_{st} during a cycle N. If heat only is produced no fatigue will take place ($S_{th} > 0; S_{st} = 0$). Fatigue implies the formation of structural disorder, namely permanent defects

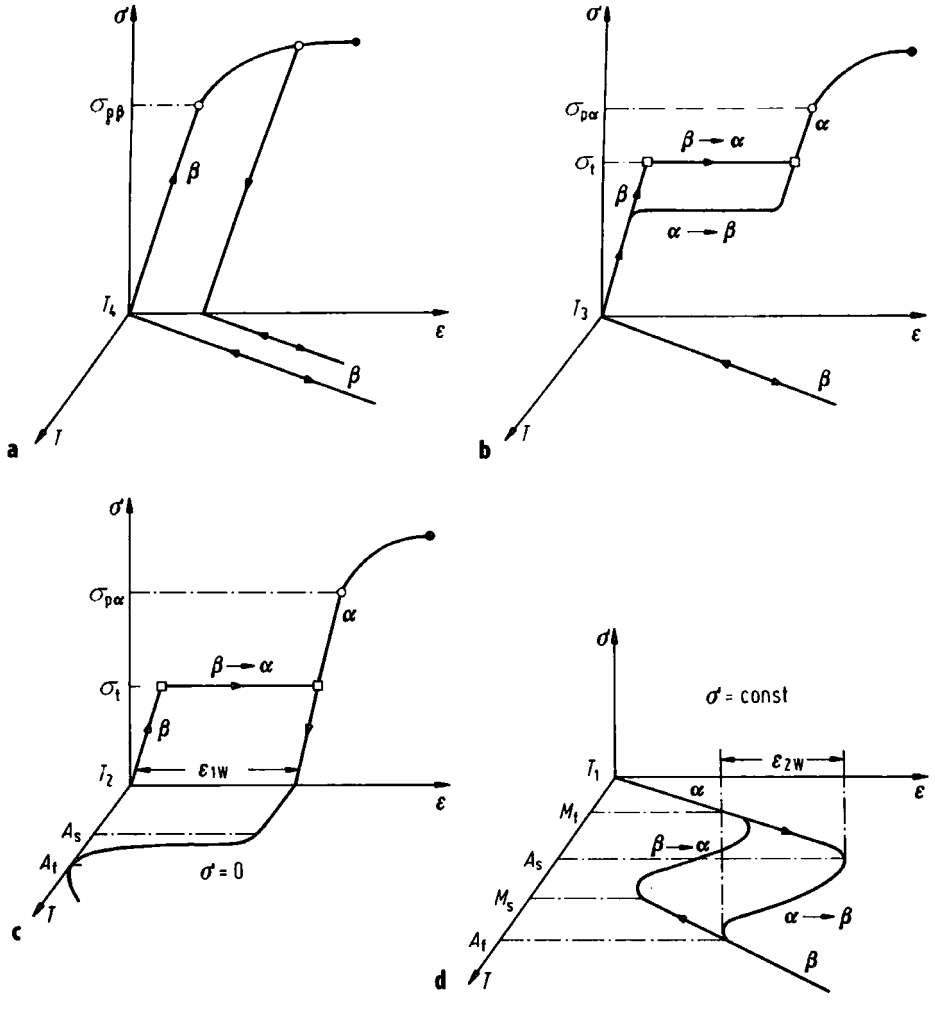
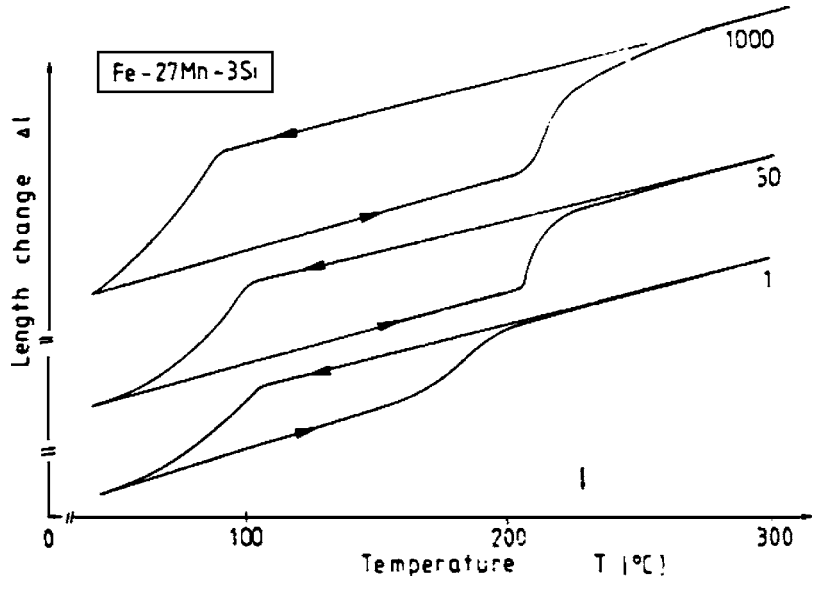
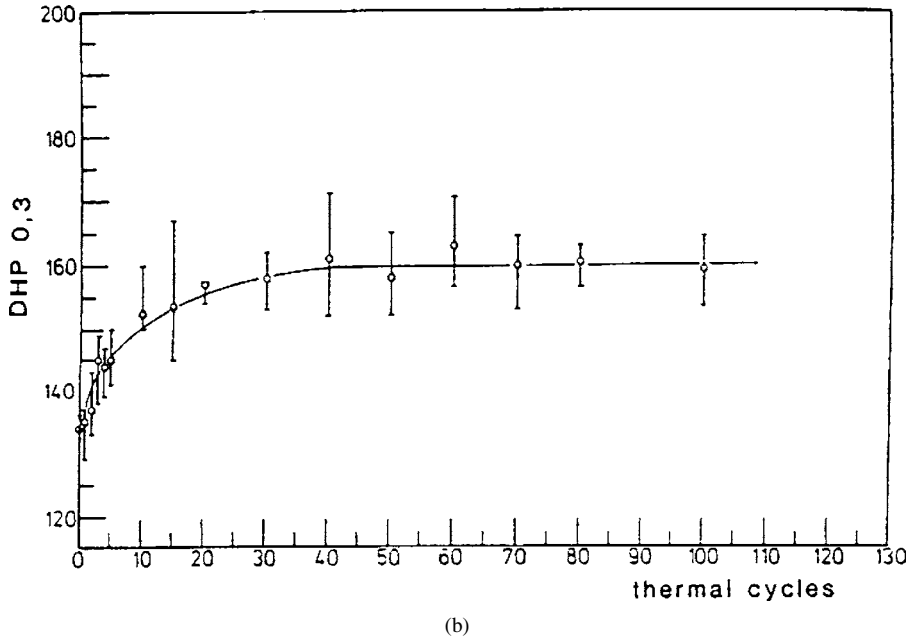


Figure 2 Shape changes due to martensitic transformation, schematic survey in σ-, ε-, T-space: (a) normal behaviour, no transformation, temperature range 3 a, Table I, (b) pseudo-elasticity, strain $\varepsilon_{\beta\alpha}$ is reversed by reduction of stress, and (c) pseudo-plasticity, one-way effect, stress induced strain $\varepsilon_{\beta\alpha}$ is reversed by heating to $T > A_f$ and (d) two-way effect, strain $\varepsilon_{\beta\alpha} \leftrightarrow \varepsilon_{\alpha\beta}$ is caused solely by thermal cycle: $\Delta T > T_3 - T, T_3 > A_f, T_1 < M_f$.



(a)



(b)

Figure 3 Fatigue effects are due to transformation and accumulation of defects (Fig. 4): (a) thermal cycling of through $\gamma \leftrightarrow \varepsilon$ transformation, increasing hysteresis $M_s(0) = 110°C$ and (b) hardening of thermally cycled CuZn39 (at.%) alloy, $M_s = -70°C$.

or disorder of atomic positions during repeated cycles (i.e., incomplete crystallographic reversibility):

$$0 < \frac{dS}{dN} = \frac{dS_{th}}{dN} + \frac{dS_{st}}{dN} \qquad (2)$$

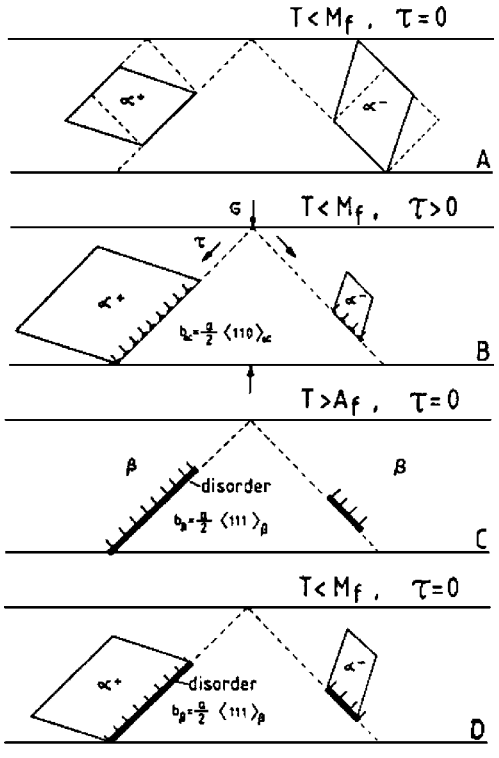
Fig. 3b shows the consequences of an accumulation of defects caused by purely thermal cycles: (Fig. 1b), and the change of a pseudo-elastic stress strain curve, due to $N > 1$ stress-induced cycles at T_2. The effects are caused by the localized formation of lattice defects, by structural changes in grain or domain boundaries or by stabilized residual martensite inside crystals. In addition, the morphology of the (polished) surface is modified in a characteristic way (Figs 3–5) [4–6]. A low value of the conventional yield stress favors the formation of such extrinsic defects for defined loading conditions. Defects, for example dislocations b_α, produced by lattice invariant shear in martensite α at $T_1 < M_f$ may transform into a lattice vector of austenite β during heating to $T_3 > A_f$:

$$b_\alpha \rightarrow b_\beta \qquad (3a)$$

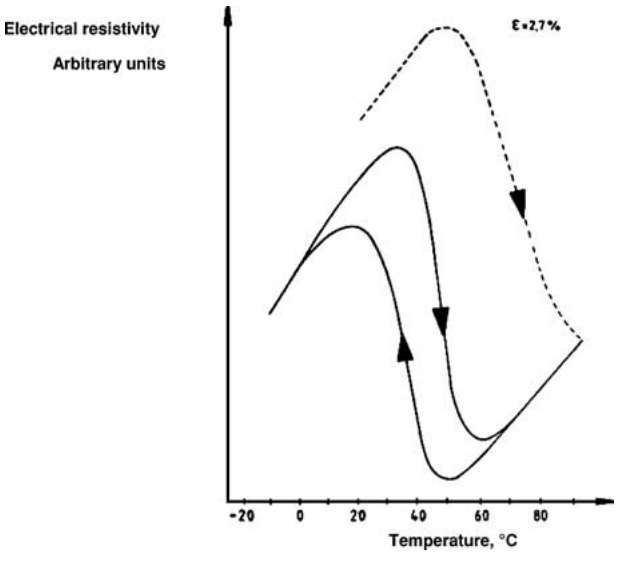
N-times repeated transformation cycles will cause in the interior an accumulation of tangles and persistent bands, or intrusions and extrusions in the surface [4]. Weaker defects $b^*_\alpha < b_\beta$, may form if martensite is exposed to one small strain amplitude: dislocations of b^*_α inside of martensitic domain boundaries or habit planes. Such defects can annihilate during the first reverse transformation by a reshuffling process (Fig. 4b):

$$b_\alpha > b^*_\alpha \rightarrow b_\beta = 0 \qquad (3b)$$

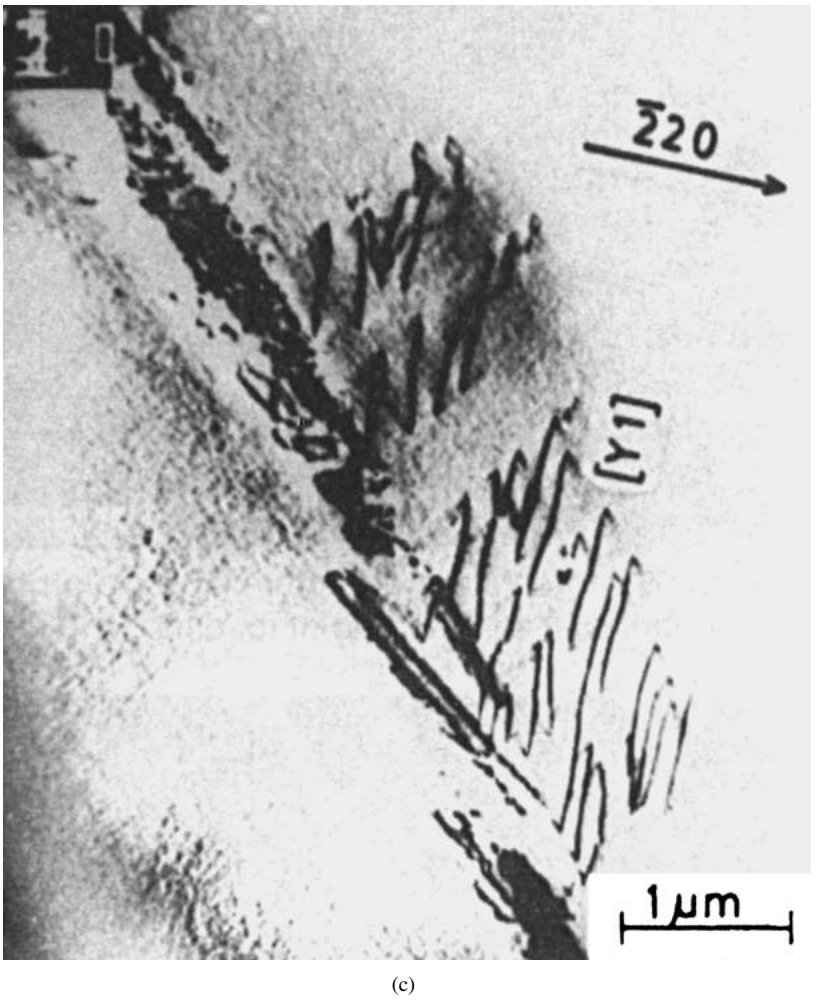
This leads to the effects of the first cycle (EFC), i.e., an anomalous raise of A_f for the first thermal fatigue cycle



(a)



(b)



(c)

Figure 4 Formation of lattice defects by transformation cycles: (a) Formation of defects in domain boundaries by one mechanical-thermal cycle: dislocation and disorder is created in domain boundaries of martensite which are not removed by heating into austenite $T > A_f$. They therefore induce two-way effect. (b) Effect of first cycle (EFC) in CuZnAl-alloy ($M_s = +50°C$). Martensite was deformed slightly, weak defects ($b^*_\alpha < b_\beta$, Equation 3b) are removed by first cycle ($N = 1$, dotted), subsequent cycles ($N > 1$) behave normally. (c) Microscopic analysis of transformation induced defects, and CuZnAl, dislocations in austenite after pseudo-elastic cycling (TEM) Cu16Zn, 16Al, at.%, $M_s = -15$.

[7, 8] (Fig. 4b and c). If they do not annihilate they stay in austenite β as ghost dislocations in an arrangement that resembles the interior structure of martensite:

$$b^*_\alpha \rightarrow b^*_\beta \tag{3c}$$

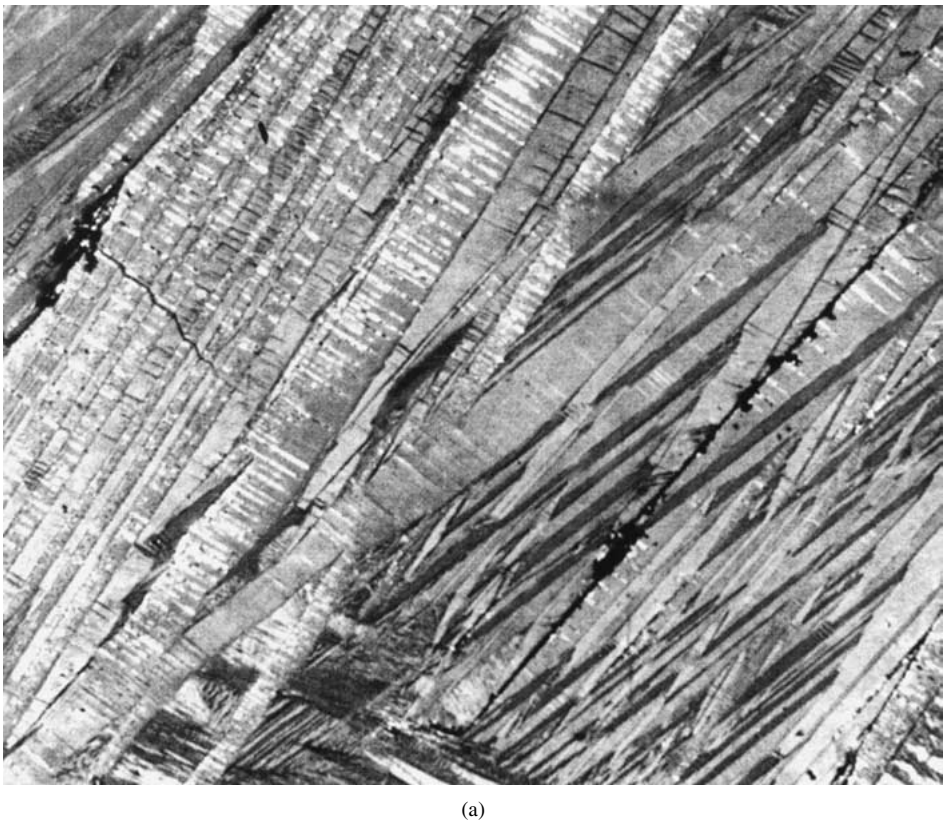
The consequence can be the induction of the two way effect caused by internal stresses (Figs 2d and 6)

3. Stages of mechanical fatigue: nucleation, propagation of cracks, transformation at crack tip, rupture

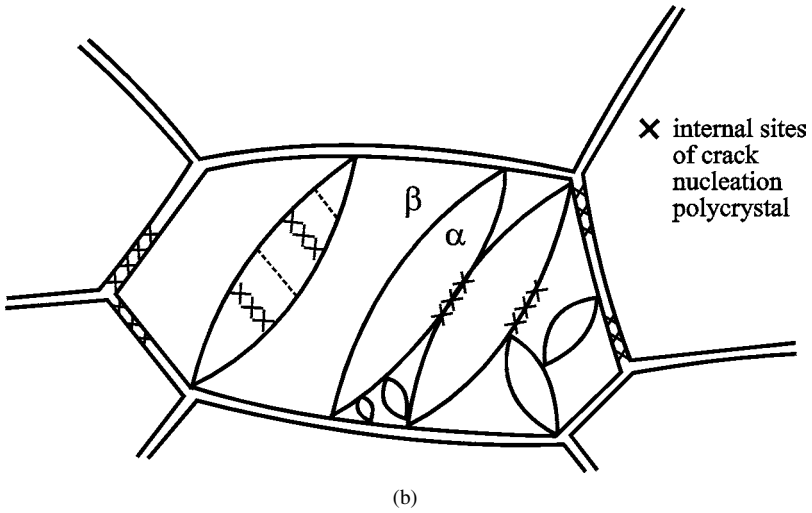
We have to distinguish shape memory fatigue from classical mechanical fatigue. The former comprises changes of transformation temperatures, the reduction of transforming volume and consequently a reduction of the amount of shape changes. The latter resembles phenomenologically 'classical' fatigue, i.e., local workhardening (Fig. 3) and the formation, subcritical and critical growth of cracks [9–13] (Fig. 5).

A remarkable microstructural feature is multiple crack formation. A consequence can be a relatively long fatigue life obtained by cycling exclusively in the martensitic state $T_1 < M_f$. This behavior is caused by preferred crack nucleation at variant boundaries which evidently act as weak spots under fatigue loading (Fig. 5).

A peculiar fracture-mechanical situation arises, if stress-induced transformation is localized to the crack

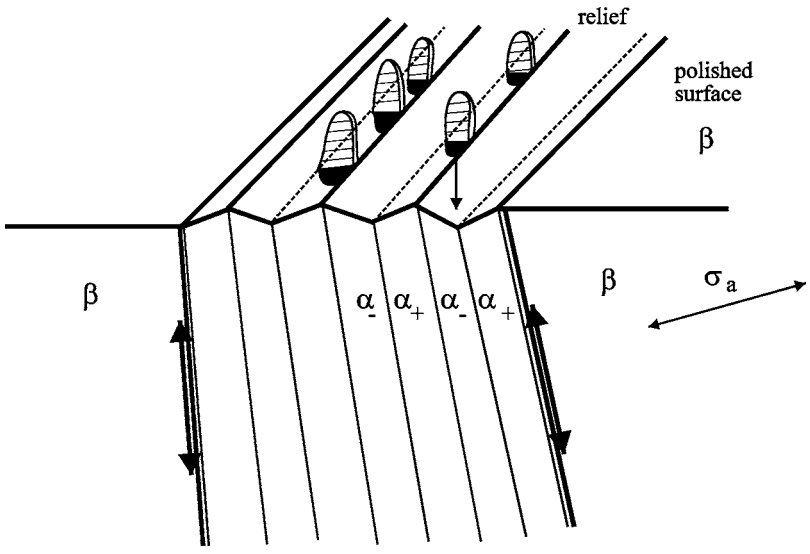


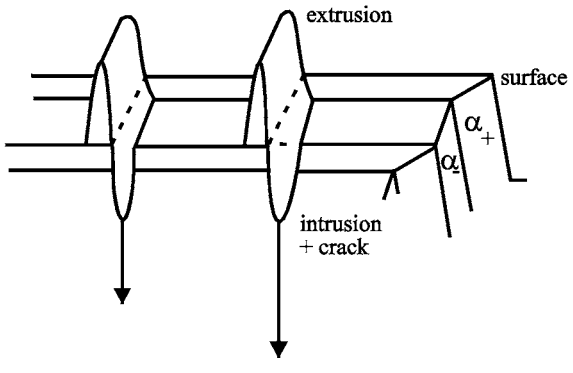
(a)



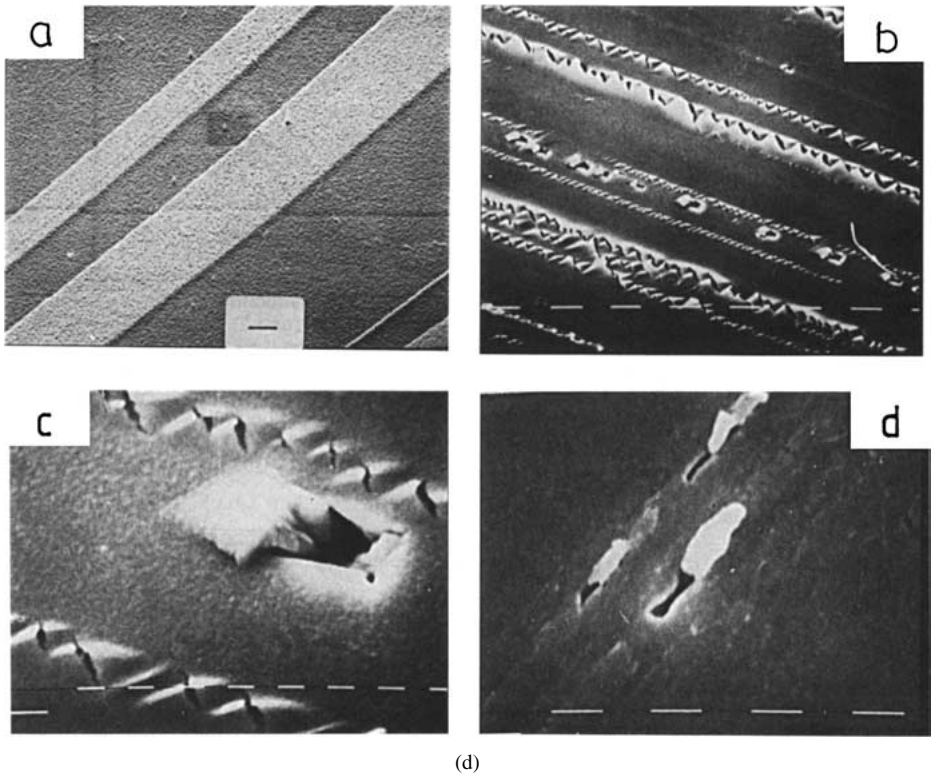
(b)

Figure 5 Initiation of cracks by repeated phase transformation: (a) multiple crack nucleation at domain boundaries: β-CuZnAl, $\sigma_a = 240$ MPa, $N = 3 \times 10^3$, $M_s = +120°C$, (b) nucleation sites in the interior of the alloy, schematic, (c) tongue-shaped extrusions/intrusions in polished surfaces: (d) Cu15 · 5Zn16.4Al, $N = 3 \times 10^3$, SEM, and formation of tongue shaped extrusions and cracks at domain boundaries by repeated transformation, schematic (SEM). *(Continued)*



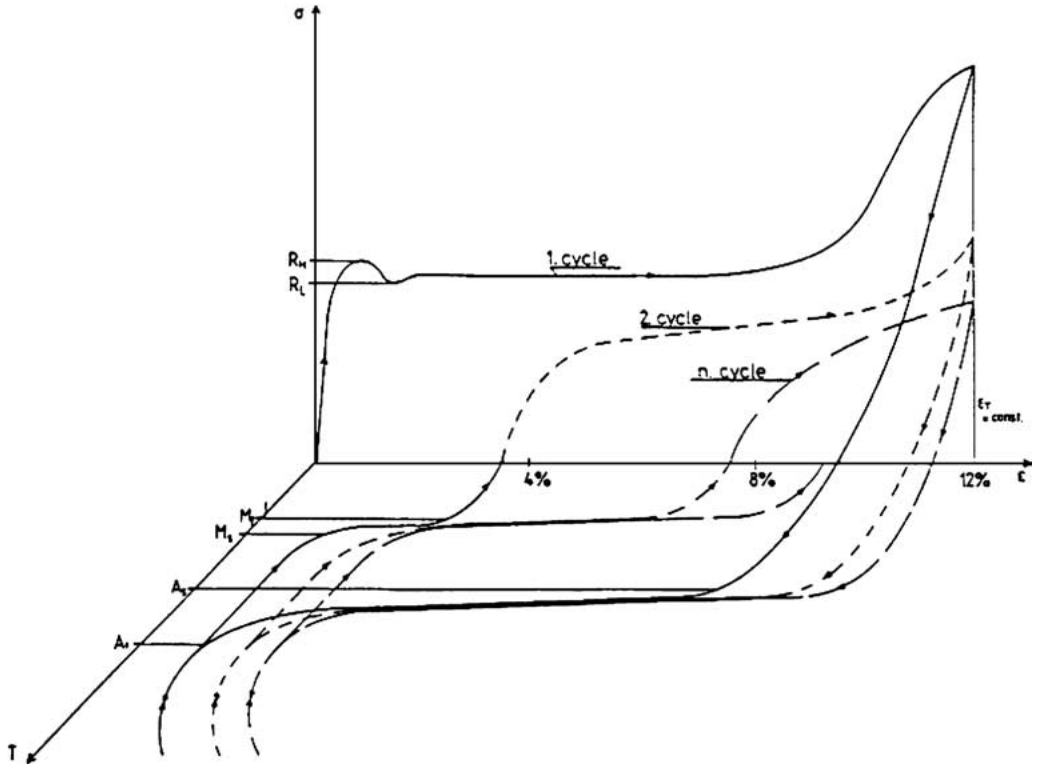


(c)

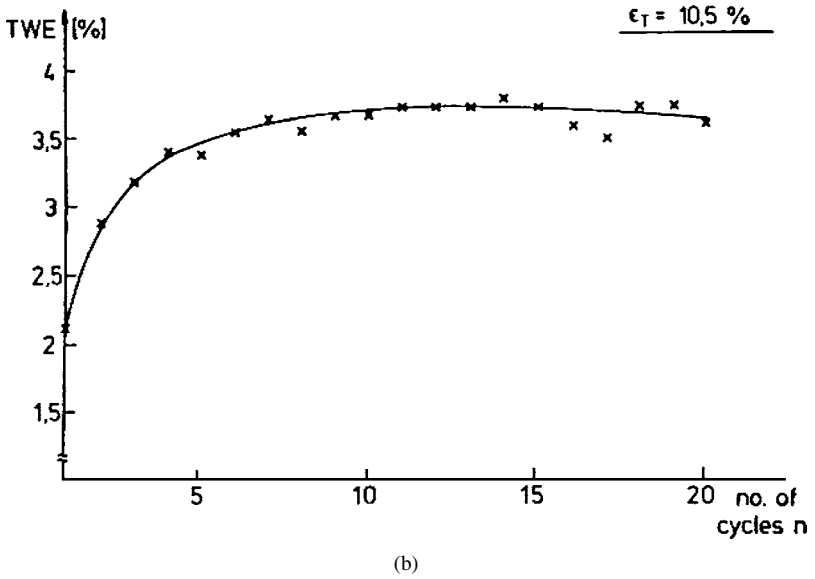


(d)

Figure 5 (Continued).



(a)



(b)

Figure 6 Training for the two-way effect of binary NiTi (50.3 at.%Ni) by thermo-mechanical cycling, i.e., intentional formation of defects, without cracking: (a) mechano-thermal training cycles, $\varepsilon_T = 12\%$, semi-schematic, compare Fig. 2c and (b) amount of two-way effect $\varepsilon_{TWE}\%$) vs. number of cycles N for $\varepsilon = 10.5\%$, $N_T \approx 10$, Fig. 2d.

tip. There are two reasons for special effects on crack growth: the transformation is reversible $\alpha \leftrightarrow \beta$, or it is connected with an increase of specific volume $V_{\beta\alpha} > 0$. Retardation or stoppage of crack growth is the consequence (Fig. 7a and 8c, d) [14]. The former is observed in shape memory alloys, the latter in austenitic steels (or ZrO_2 containing ceramics). In steels the martensitic transformation is also associated with the formation of a ferromagnetic phase. This in turn may be used as a sensor, if a certain stage of crack growth has been reached.

Crack initiation may take place by the accumulation of dislocations in austenite β, which cause highly localized internal stresses by pile-ups as in classical fatigue. In addition, localized sliding inside α/β-interfaces and α^+/α^--domain boundaries in martensite is often observed as a cause of crack formation. Tongue-shaped intrusions/extrusions form associated with slip inside the martensitic structure (Fig. 5). There is a large number of special investigations, most of them are concerned with Cu-base alloys single- or polycrystalline [5–10, 13–17], some with NiTi [11, 12], and few with Fe-base alloys [18–21]. These behave differently from the SMA because of the volume change and lattice invariant shear by dislocation in martensite.

Especially in CuZn-based alloys intercrystalline crack paths along austenite grain boundaries are

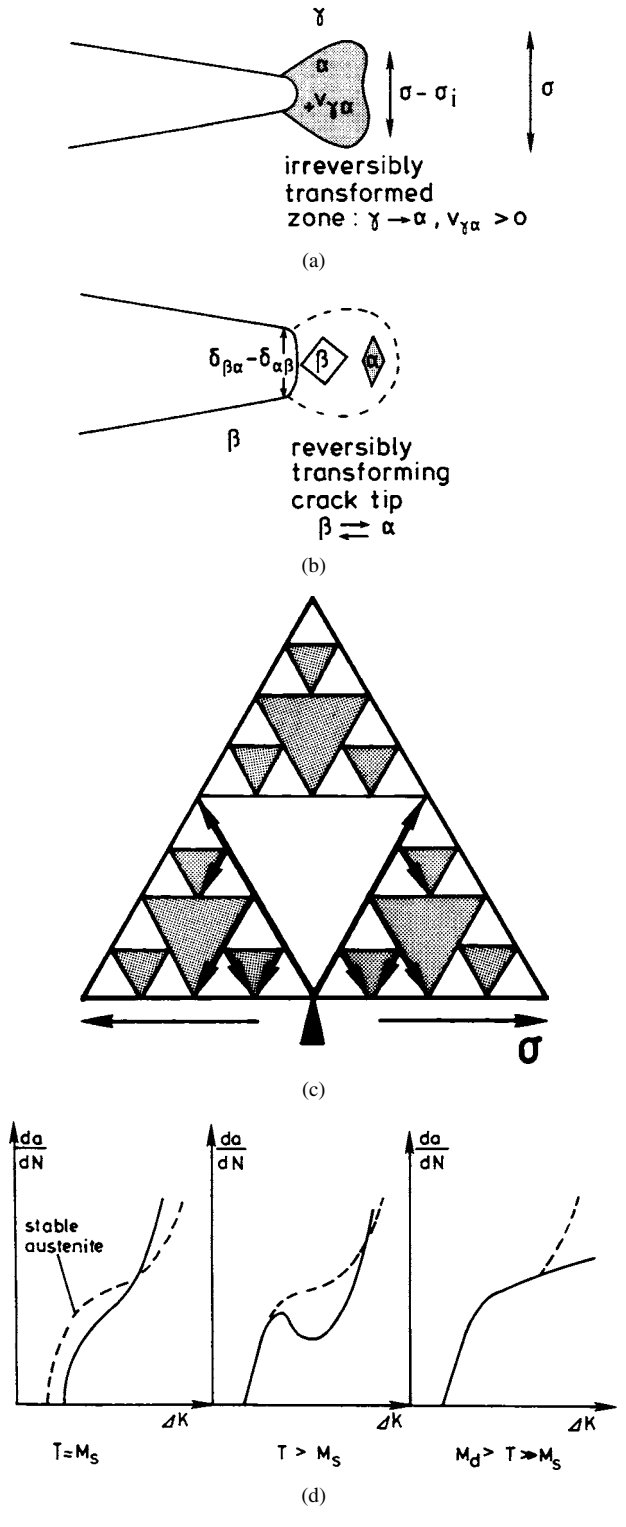


Figure 7 Crack propagation and martensitic transformation in austenitic steels and SMA, schematic. The test temperature T is related to the martensite temperature M_s. A higher amplitude of stress intensity ΔK is required the larger is $\Delta T = T - M_s$. (a) Localised transformation associated with increased volume ($+V_{\gamma\alpha}$, steels), pseudo-elastic shear and/or multiple crack formation at crack tip, (b) fractal crack propagation at weak phase- or domain-boundaries, see Fig. 5a and (c) typical *da/dN*-functions of metastable austenite retarded crack growth if an expanded and transformed zone is surrounded by non-transformed material (stable austenite: dotted).

frequently observed. In such cases intra-martensitic mechanisms will become less relevant. Changing the grain boundary morphology by hot rolling treatments often provides a successful therapy against such an efficient crack growth mechanism [16]. If the cracks are initiated and propagated in a martensitic structure. The fracture surface is rough and reflects the fractal structure of martensite. Experimental results on crack propagation in transforming SMA show no clear advantage of the transformation. This is probably due to the weakening effect of the domain boundaries. The positive volume change in iron alloys seems to retard crack growth always in a certain range of stress intensities (Fig. 7).

4. Thermal cycling, effects of first cycle, martensitic ghosts, and other structural irreversibilities

Full crystallographic reversibility, i.e., the faultless restoration of austenite has to be aspired for good fatigue resistance. This is difficult to achieve even for stress-free transformation cycles (Figs 1–3) because the martensitic transformation is always connected with large amounts of shear strain. The formation of internal stresses is unavoidable. They are reduced by intrinsic defects due to lattice invariant deformation: domain boundaries, stacking faults, twins, and in steels by dislocations with burgers vectors b_{lid} of the bcc lattice. All of these defects should annihilate during the reverse transformation for a fully reversible transformation and a consequent high fatigue resistance (Figs 4 and 5):

$$b_{\alpha\mathrm{lid}} \rightarrow b_{\beta} = 0 \tag{4}$$

If they cannot annihilate, as for example $^{a}/_{2}\,\langle 111\rangle$- or a $\langle 100\rangle$-dislocations in martensitic FeNiX-alloys, then they become transferred into austenite β, where they accumulate during subsequent cycles [5, 20, 21]. The structures can be revealed in the austenite even by light microscopy. The etching effects resemble martensite and are designated as 'martensite ghost.' They can serve as an indication for incipient thermal fatigue, even if no external stress is acting (Fig. 4).

$$b_{\alpha\mathrm{lid}} \rightarrow b_{\beta\mathrm{ghost}} \tag{5}$$

Other defects, mainly dislocations are formed in austenite in the environment of growing martensite. This is especially true in steels because its austenite is soft. Strengthening of austenite by ausforming or coherent precipitation will reduce this effect and consequently sensitivity to fatigue. Some defects are in fact annihilated by the reverse $\alpha \rightarrow \beta$ transformation as implied by Equations 3b and 4. These are defects inside domain boundaries with small Burgers vectors ($b^{*}_{\alpha} \ll b_{\mathrm{a}}$). Their annihilation requires an additional driving force (overheating of A_s) which leads to the "effect of the first cycle" (Fig. 4b) in thermal cycling experiments.

5. Training and fatigue of one-way and two-way effects

The training of SMA is related to fatigue, because defects are introduced, intentionally to affect the martensitic domain structure (Fig. 6). Training often consists of a cyclic mechano-thermal treatment: pseudo-plastic deformation of martensite α at $T_1 < M_f$, slightly above the yield stress, plus heating to $T_3 > A_f$. There exists

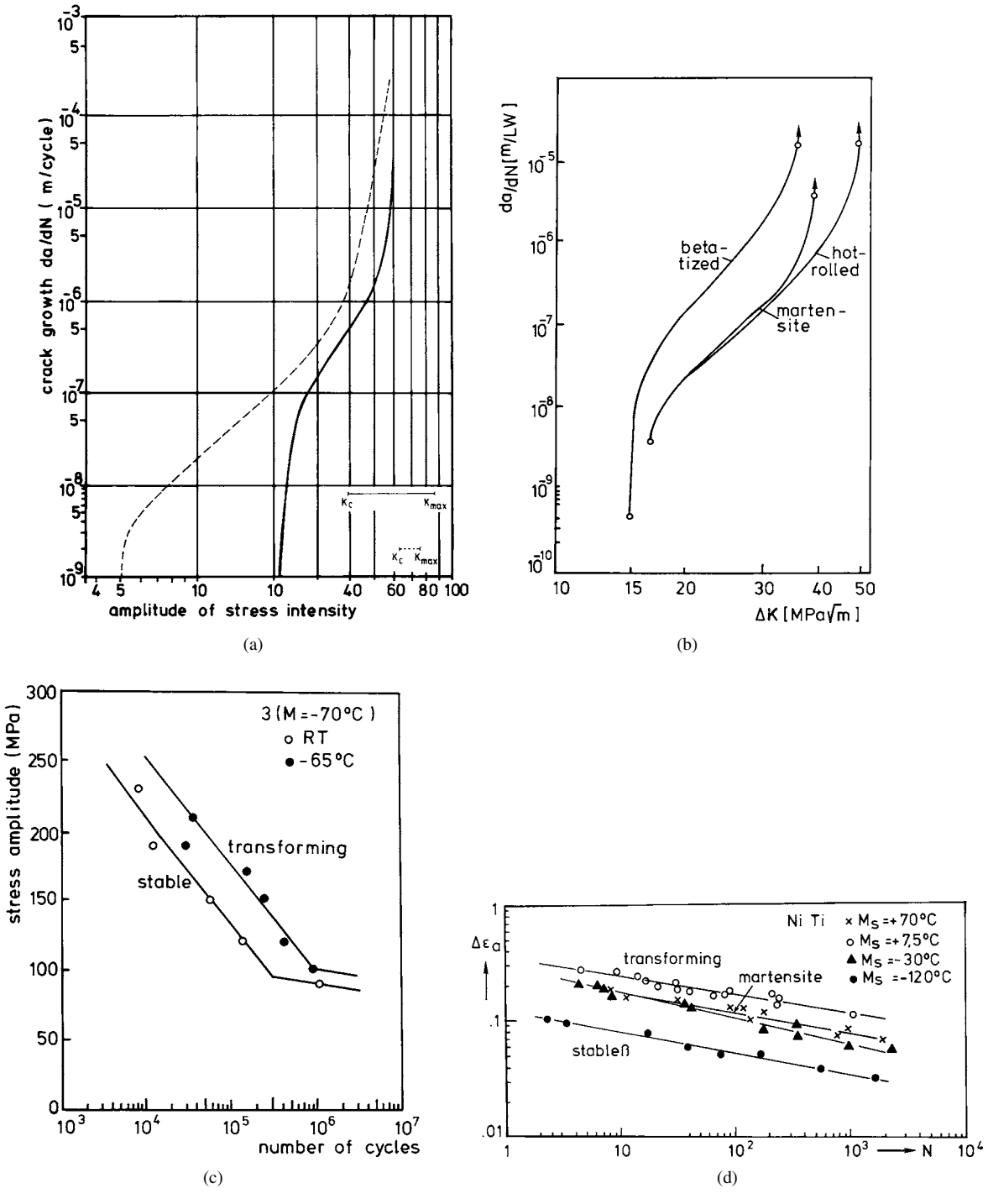


Figure 8 Crack growth and martensitic transformations. (a) Austenitic alloy (FeNi29.0Al7.1 at.%, $M_s = -32°C$) cooled to $T < M_f$ (dotted line) compared with transforming state $T > M_s$), Slower crack growth of metastable austenite (full line) in spite of lower hardness as compared to martensite (dotted) (see Fig. 7d). (b) da/dN-curves of CuZn-based alloys measured at 20°C with different transformation temperatures. The unfavourable behaviour of "betatized" can be due to grain boundary embrittlement. (c) Wöhler curves of stable (β) and transforming ($\beta \leftrightarrow \alpha$) CuZn39.5 at.% alloys. (d) Wöhler curves of NiTi-alloy with different transformation behaviour: higher strain amplitudes of transforming states.

an optimum strain ε_p and number of cycles N. If this optimum is surpassed, fatigue starts (Fig. 6). Training is identical with the introduction of defined internal stresses caused by dislocation groupings, or of residual martensite which aids further martensitic growth. Some alloys provide a third way, namely stress fields of coherent particles. The latter requires an aging treatment of suitable alloys as part of the training. Trained-in two-way behavior should stay stable for $N \rightarrow$ max cycles. Its loss due to thermal cycles is designated as of shape memory fatigue. Introduction of large plastic strains can cause partial or complete loss of transformability and therefore memory. Such alloys can only be revitalized by betatisation at temperatures at which recovery, recrystallisation and reordering in the β-phase can take place (>300°C).

6. Aging effects and fatigue

Up to now fatigue effects have been explained as being due to mechanically induced structural irreversibilities. This must be true for a temperature range $T < T_c$ in

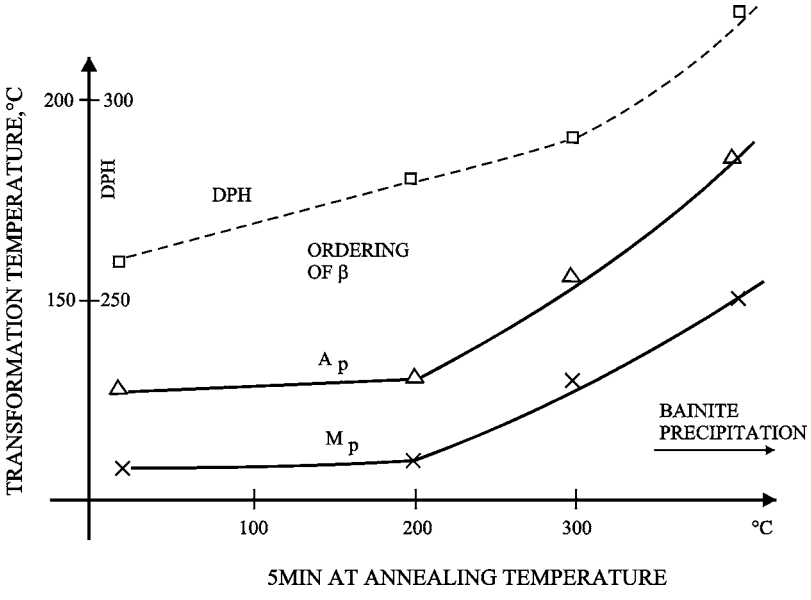


Figure 9 Aging and cycling loading: CuAlNi-alloy as-quenched. $\alpha \rightarrow \beta$ reverse transformation takes place at increasing temperatures because A_f is in the temperature range where diffusion is not negligible: increased ordering of the β-phase, CuAl12Ni5Mn2Ti, first cycle: $M_s = +126$, $M_f = +105$; $A_s = +124$, $A_f = +139$.

which negligible diffusivity can be expected ($D \approx 0$), namely between $0K < 1/3T_m = T_c < A_f$ (Table I). The temperature T_c is related to the melting temperature T_m. This relationship is based on substitutional diffusion and it can be defined by an extended period of time for example $t_c \approx 1$ year required for one atomic hop:

$$1a < t_c = \frac{b^2}{D} \tag{6}$$

The $D \approx 0$-temperature T_c is, 100° for β-CuZn (Fig. 9), and 300° for β-NiTi. It should be still higher for γ-Fe-based alloys (steels $T_c = 400$°C), if interstitial diffusivity by traces of Carbon or Nitrogen does not play a role.

If aging effects interfere with martensitic transformation in a static way its effect is known as stabilization of martensite [22]. The prerequisite is that $T_c < A_f$. In this case diffusional processes can take place inside of martensite. For example reordering or hops of atoms into structural defects will take place, the amount of which depends on temperature $T > T_c > {}^1/_3 T_m$ and time. The effect is a sometimes considerable time dependent raise in the temperatures of the reverse transformation A_s and A_f (Fig. 9). This is relevant and has been well investigated in brass-type alloys [23–26]. In our context the aging behavior during repeated thermal cycles above $A_f > T_c$ is of interest. The same is true for isothermal mechanical cycling in this temperature range of limited diffusivity. Ordering effects have to be considered, especially in alloys which have to be quenched from a disordered β-phase to obtain homogeneity. Thus alternating stabilization of martensite and ordering of austenite may take place during thermal or isothermal mechanical cycling in the appropriate temperature range causing a complex fatigue behaviour [23–26].

The superposition of pseudo-elastic fatigue and aging effects explains the difference in shape of the stress strain curves for cycling below and above T_c. Defects which accumulate at low temperature evidently annihilate at higher temperature by diffusion-aided process. This complex behavior is the disadvantage of CuZn-based alloys as compared to more stable NiTi because time plays a role in all fatigue conditions in copper base alloys.

For consideration of aging effects not only environmental temperatures are relevant. It has been shown for CuZn- as well as for NiTi-alloys that considerable adiabatic heating can take place during pseudo- elastic straining. The amount and localization of heating depends on the strain rate. The local raise in temperature (+50°) stabilizes the austenite, retards or stops the $\beta \rightarrow \alpha$-transformation front. It therefore favors new nucleation of martensite elsewhere in the specimen in case of high strain rates. Therefore a strain rate dependence of the martensitic microstructure (a shift from discontinuous to continuous) in addition to aging effects can be expected for high frequencies [27, 28]. The discontinuous mechanism implies a Lüdersband-like pseudo-elastic deformation. A transformation front moves forward and backward through a single or polycrystalline material in the discontinuous mode (Fig. 1c).

7. Intentional introduction of defects by thermo-mechanical treatments

For common NiTi-based alloys aging below A_f does not interfere. Only mechanical creation of defects is relevant for the fatigue behavior. Then the simple requirement for achievement of a high number of stable cycles is: a low ratio of pseudo-elastic stress $\sigma_{\beta\alpha}$ to true yield stress $\sigma_{y\beta}$, (Equation 1d):

$$\frac{\sigma_{\beta\alpha}}{\sigma_{y\beta}} = \min \tag{7}$$

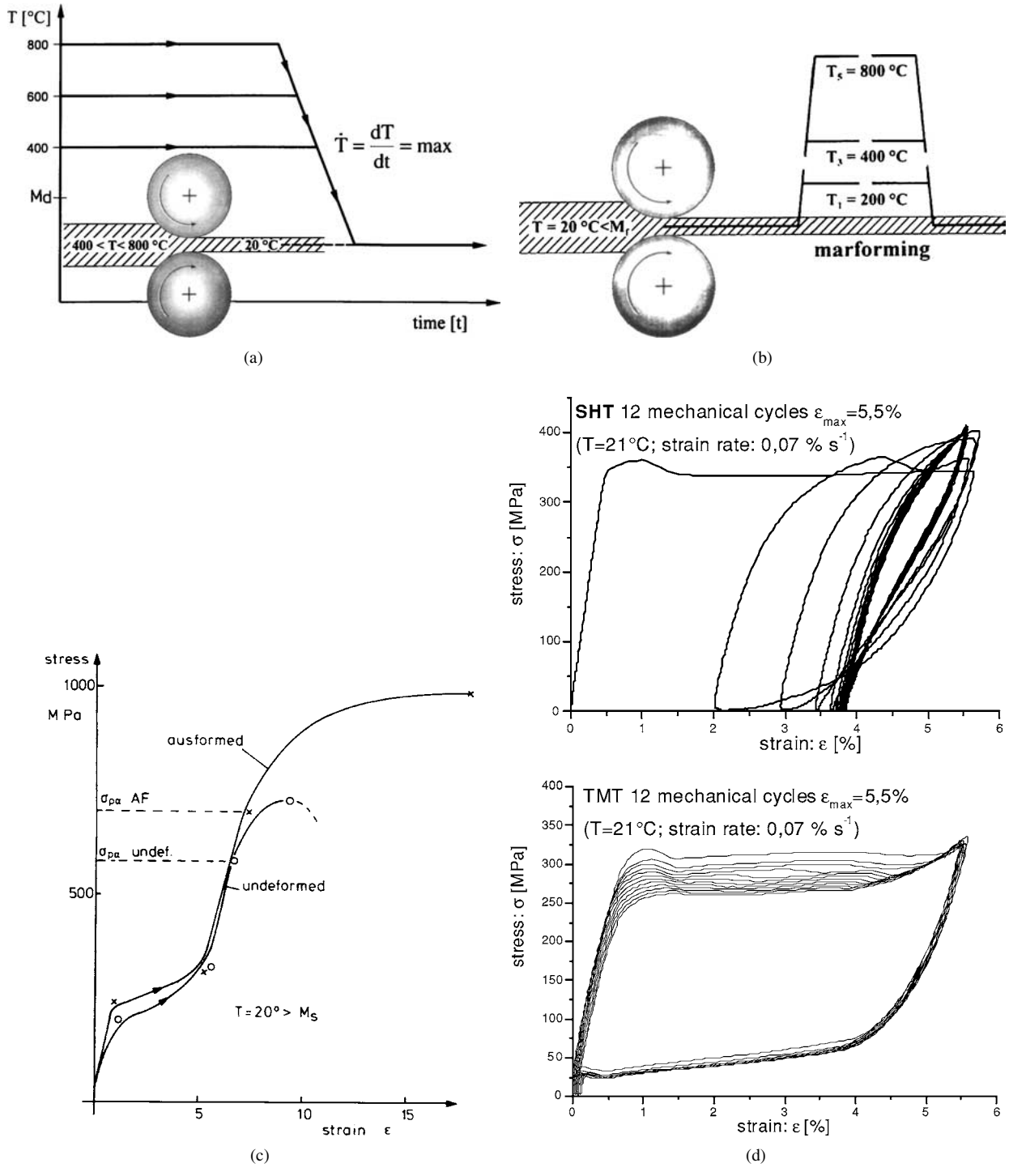


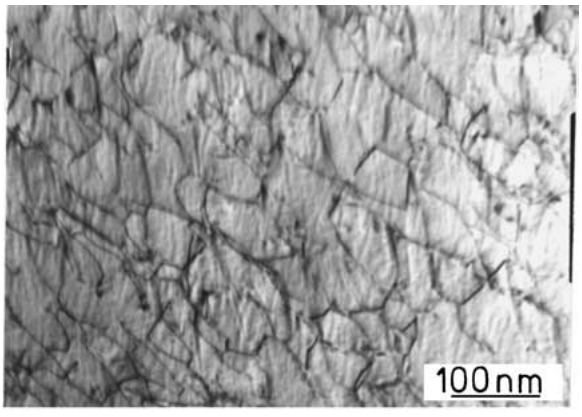
Figure 10 Introduction of primary defects and particles by thermo-mechanical treatments (TMT). (a) temperature-time scheme of ausforming, (b) temperature-time scheme of marforming, (c) stress strain curves of NiTi (50.3 at.%Ni) as betatised, ausformed, and (d) effect of TMT (marforming and aging), SHT (solution heat treated) on high amplitude low cycle fatigue of NiTi (Ti 50.7 at.% Ni, $M_s = -30°C$).

and not to exhaust the total plateau strain $\varepsilon_{\beta\alpha}$, because toward its end true plasticity is likely to occur (Fig. 10c).

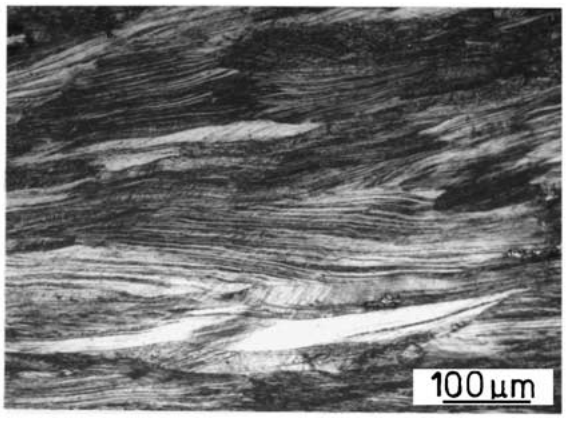
$$\frac{\varepsilon}{\varepsilon_{\beta\alpha}} \leftrightarrow \min \tag{8}$$

Therefore NiTi-alloys hardened by thermo-mechanical processes are the most promising candidates for safe use of SMA under cycle loading conditions [29, 30]. Ausforming and marforming plus tempering have been developed for production of ultra high strength steels. These methods are also useful for the production of SMA-semifinished products with high strength.

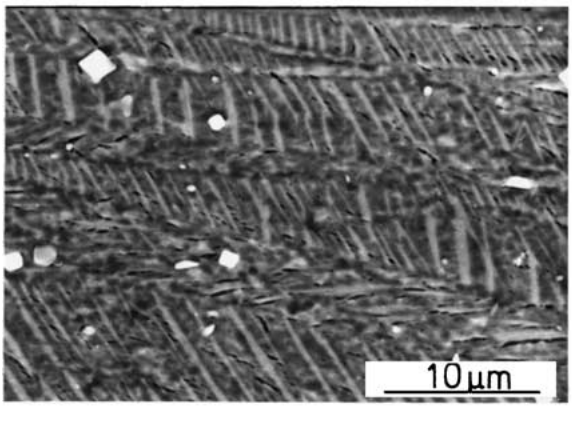
Fatigue resistance of SMA is related to the difficulty to create and accumulate defects during temperature- or stress-induced martensitic cycles. Lattice invariant plastic deformation of martensite by dislocations is to be avoided. Secondly, the yield strength of austenite should be high. This is of special concern for the development of shape memory steels. A disadvantage is the softness of fcc austenite for alloys of iron. (Fig. 3) But also for bcc SMA strengthening of the β-lattice has to be aspired. This is achieved by an increase



(a)



(b)



(c)

Figure 11 Microstructures of SMA after various TMT: (a) CuZnAl, TEM, 800°C, $\varepsilon = 1.66$, dislocations in β, TEM, (b) NiTi, (50.3 at.%), LM, 500°C, $\varepsilon = 0.71$, lamellar structure, and (c) NiTi, (50.3 at.%), SEM, 400°C, $\varepsilon = 0.30$, cooled to $T < M_f$, ultra-fine martensite between lamellae.

in crystallographic order (aging treatments in some Cu-base alloys, Fig. 9) [23] or by the introduction of three-dimensional dislocation networks into austenite β. These extrinsic dislocations are transferred into martensite α and back to β during transformation cycles (Figs 5c, 10c, and 11):

$$b_\beta \leftrightarrow b_\alpha \tag{9}$$

Less fatigue is caused, the higher the yield stress of the austenite $\sigma_{y\beta}$ is raised (Equation 6). The superior fatigue resistance of NiTi-based alloys as compared to Cu-alloys is related to the intrinsic strength due to the high ordering energy of NiTi and their higher stability against aging during cyclic loading. Thermo-mechanical treatments such as ausforming or marforming plus tempering provide a very effective means of additional strengthening for all types of SMA (Figs 10 and 11) [29]. According to recent experimental results, marforming plus tempering above 350° leads to a β-subgrain structure with the best fatigue properties. This treatment is however limited by the (rather modest, $\varepsilon_p < 20\%$) plastic deformability of NiTi-martensite.

$\sigma_{\beta\alpha}$ is the stress which is required to induce $\beta \rightarrow \alpha$-transformation at $T_2 > M_s$ in Equation 6. It is highly temperature dependent and controlled by the Clausius-Clapyron equation (Equation 1b) [31]. The yield stress of austenite $\sigma_{y\beta}$ is not strongly temperature dependent, but can be raised considerably by thermo-mechanical treatments [29, 30] (Figs 10 and 11).

8. Improvements of fatigue properties, prerequisites and limits for good fatigue behavior

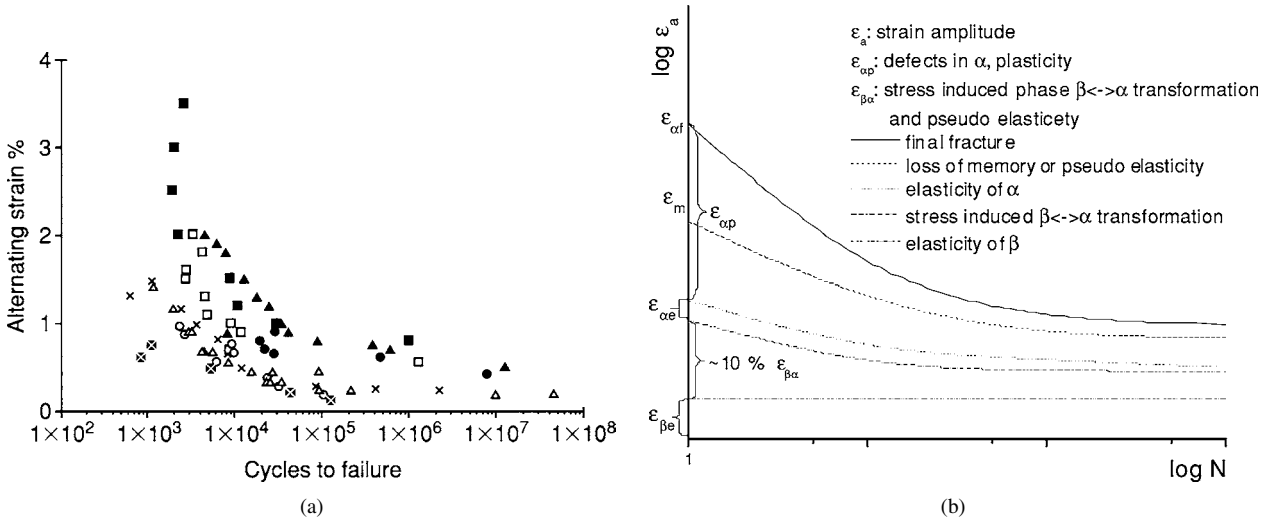
Iron- and copper-based alloys must be excluded as favorable candidates for good fatigue behavior for different reasons. The fcc austenite of steels in usually soft, which favors accumulation of defects in austenite. Intramartensitic lattice invariant deformation takes place by dislocations in most steels. The $^a/_2\langle 111\rangle$ dislocations in the bcc lattice interact to form nodes which can not annihilate during the reverse transformation. The creation of dislocations in the fcc lattice is also favored by the positive volume change which is associated with the martensitic transformation in most alloys of iron. It is related to the simultaneous magnetic transformation. This explains why iron-based alloys with the fcc $\Leftrightarrow$ bcc or bct-transformation show the worst fatigue behavior. The fcc $\Leftrightarrow$ hcp-transformation in FeMnX-alloys is connected with a smaller and negative volume-change. The fatigue effects are less than for the bcc and bct martensites. Because of the softness of both phases, the number of cycles obtainable by thermal or mechanical cycles is still much too small for applications for which $N > 10^3$ is required (Table II).

The Cu-based alloys show a negligible volume change. Their major problem is instability of transformation temperature and pseudo-elastic stress strain curves due to aging effects. As they have to be quenched from the disordered β-phase to obtain homogeneity, they therefore usually show an undefined, incomplete

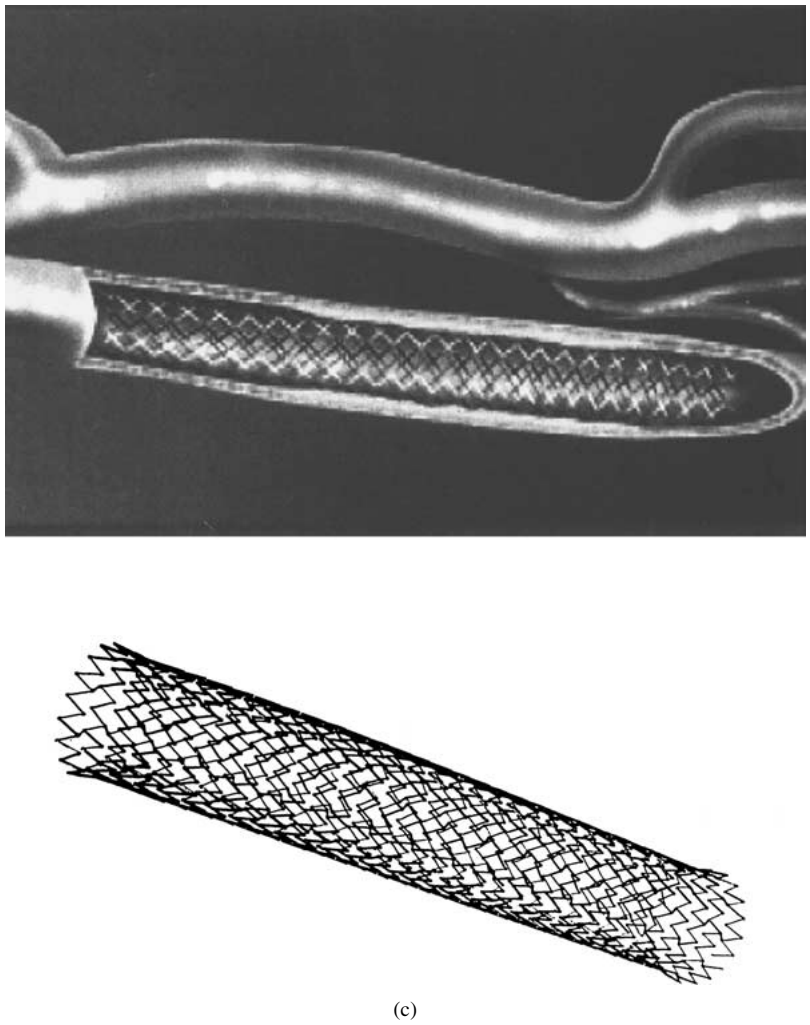
TABLE II Fields of application for which fatigue resistance is relevant (see Fig. 2)

Application	Type effect	N_{max}
Thermal valve control	2W	10^4
Positioning	2W	10^5
Grippers, robot fingers	2W	10^6
Orthodontic wires	PE	10^4
Stent in veins, arteries	PE	10^7
Damping, internal friction	PE	10^8
Tube coupling	IW	10^1
El. connectors	IW	10^2

PE: pseudo elasticity; 1W: one way effectet or pseudo plasticity; 2W: two way effect.



(a) (b)



(c)

Figure 12 Applications: (a) Wöhler-curve data of different NiTi alloys for stents [39, 40], (b) typical ε-N-curve of SMA indicating the additional pseudo-elastic strain $\varepsilon_{\beta\alpha}$, schematic, and (c) stent for opening up arteries (a complex strain amplitude $\varepsilon_{\beta\alpha}$ 5% and $N \approx 10^5$ cycles per day are usual loading conditions).

state of ordering. Diffusivity is not negligible at temperatures only slightly above ambient temperature, thus in the range of transformation temperatures.

At present NiTi-based alloys show a considerably better fatigue behavior as compared to the brass- or bronce-like alloys [29–33]. The major field of research is directed toward improving their microstructures for still better fatigue behavior. This is usually done by thermo-mechanical treatments. They lead to the intentional introduction of defects to increase the conventional yield stress (without spoiling transformability), or by the introduction of favorable textures to maximize the strain $\varepsilon_{\beta\alpha}$ (Equation 7). The best method to obtain a NiTi-alloy with optimum fatigue resistance consists at present of marforming plus tempering, so that a fine subgrain structure of the β-phase is formed.

Concomitant formation of coherent nano-meter size particles can lead to further improvement. Such a treatment is suitable to simultaneously producing a semi-finished product (sheet, wire) of shape memory alloys with the desired secondary properties: strength and fatigue resistance (controlled rolling).

Inspite of a considerable number of specialized papers on fatigue of shape memory alloys, a systematic, comprehensive approach to this field is still rare because the field has not received the attention which it deserves. There are only a few earlier incomplete reviews [34–38]. More work especially on NiTi-based alloys with well-defined pretreatment is required for their safe use as engineering on medical materials (Table II).

The achievement of a high precision and reproducibility of transformation temperatures and shape changes is a field in which scientists and engineers should cooperate. There is still room for improvement (Fig. 12).

In conclusion, it can be stated that fatigue resistance of shape memory alloys is affected by a large number of physical properties. High fatigue resistance requires multible conditions. Some of them are summarized as follows:

1. strength of austenite → max
2. degree of order → max
3. small coherent particles (in austenite) → max
4. volume change $V_{\beta\alpha} \rightarrow$ min
5. exploitation of shape strain $\varepsilon/\varepsilon_{\beta\alpha} = \min$, $\varepsilon_{\beta\alpha} = \max$
6. lattice invariant shear in martensite → twins, stacking faults, no dislocations
7. interference of transformation by diffusion → min, melting temperature → max
8. embrittlement of grain boundaries → min
9. structural incompatibilities (large non-transforming inclusions) → min

Acknowledgement

The work was part of DFG project SFB 459 Technology of shape memory alloys.

References

1. G. V. KURDYUMOV and E. KAMINSKY, *Metallwirtschaft* **15** (1936) 370; *Discussion in Trans. AIME* **128** (1938) 361.
2. E. HORNBOGEN and G. WASSERMANN, *Z. Metallkde* **47** (1956) 427.
3. E. HORNBOGEN and M. MERTMANN, *Metall* **50** (1997) 809.
4. M. SADE and E. HORNBOGEN, *Z. Metallkde*. **79** (1988) 782.
5. S. KAJIWARA and T. KIKUCHI, *Acta Met*. **30** (1982) 589.
6. M. SADE, R. RAPACIOLI and M. AHLERS, *ibid*. **33** (1985) 487.
7. Y. ZHANG and E. HORNBOGEN, *Z. Metallkde*. **79** (1988) 13.
8. E. HORNBOGEN, V. MERTINGER and J. SPIELFELD, *Scripta Mater*. **40** (1999) 1371.
9. M. SADE, J. KUMPFERT and E. HORNBOGEN, *Z. Metallkde*. **79** (1988) 119, 678.
10a. K. N. MELTON and O. MERCIER, *Scripta Met*. **13** (1979) 37.
10b. *Idem*., *ibid*. **13** (1979) 73.
10c. *Idem*., *Mater. Sci. Eng*. **40** (1979) 81.
11a. R. O. RITCHIE, MRS, *Adv. Mats*. **9** (1987) 243.
11b. A. L. KELVEY and R. O. RITCHIE, *Met. Mater. Trans* A **32** (2001) 731.
12. H. SEHITOGLU *et al., Mater. Sci. Eng*. A **314** (2001) 67.
13. E. HORNBOGEN and M THUMANN, *Z. Metallkde* **89** (1988) 119.
14. E. HORNBOGEN, *Acta Met*. **26** (1978) 147.
15. E. HORNBOGEN and M. SADE, *Z. Metallkde* **79** (1988) 782.
16. M. THUMANN and E. HORNBOGEN, "SMA, DGM Informationsgesellschaft Frankfurt" (1992) p. 89.
17. C. DAMIANI and M. SADE, *Mater. Sci. Eng*. A **273–275** (1999) 616.
18. E. HORNBOGEN and M. SADE, "The Martensite Transformation in Science and Technology, DGM Informationsgesellschaft Frankfurt" (1989) p. 249.
19. M. SADE, K. HALTER and E. HORNBOGEN, "SMA, DGM Informationsgesellschaft Frankfurt" (1992) p. 201.
20. E. HORNBOGEN and N. JOST, *Journ. De Physique*, Vol. 1, Coll. 4, **11** (1991) 199.
21. E. HORNBOGEN, *Phys. Stat. Sol*. (b) **172** (1992) 161.
22. J. VAN HUMBECK, *et al*., in Proceedings of ICOMAT-86, Nara, Japan (1986) p. 862.
23. E. HORNBOGEN, V. MERTINGER and J. SPIELFELD, *Z. Metallkde* **90** (1998) 318.
24. K. SHIMIZU, *Electron Microsc*. **34** (1985) 277.
25a. J. MALARRIA and M. SADE, *Scripta Met*. **30** (1994) 241.
25b. A. YAWNY, F. C. LOVEY and M. SADE, *Mater. Sci. Eng*. A **290** (2000) 108.
26. L. LU, M. O. LAI and A. S. LIM, *Scripta. Mater*. **34** (1996) 157.
27. J. A. SHAW and ST. KYRIAKIDES, *J. Mech. Phys. Solids* **43** (1995) 1243.
28. *Idem*., *Acta Mater*. **45** (1997) 683.
29. E. HORNBOGEN, "Displacive Phase Transformations and their Applications in Materials Engineering" (TMS, Warrendale, PA, 1997) p. 27.
30. D. TREPPMANN, D. WURZEL and E. HORNBOGEN, *Z. Metallkunde* **89** (1998) 126.
31. E. HORNBOGEN, *ibid*. **33** (1985) 595.
32. A. HECKMANN and E. HORNBOGEN, in Proc. ICOMAT-02, Helsinki 2002, to be published.
33. *Idem*., *Mater. Sci. Forum* **394** (2002) 325.
34. E. HORNBOGEN, in Proc. Euromech. Colloquium 255, edited by K. D. Herman, *et al*. (Paderborn, 1989) p. 156.
35. *Idem*., *Proc. Mater. Sci. Council (Taiwan)* **14** (1990) 350.
36. *Idem*., in "Technologie des Alliages à mémoire de forme" (Hermes, Paris, 1994) p. 133.
37. *Idem*., in Proceedings of the Conference on Fatigue, Wien 2001, edited by S. Stanzl-Tschegg (BoKu Wien, 2001) p. 157.
38. *Idem*., *Fatigue Fract. Eng. Mater. Struct*. **25** (2002) 785.
39. Y. S. KIM and S. MIYAZAKI, in Proc. 2nd Int. Conf. SMST—97, (Pacific Grove, CA, 1997) p. 473.
40. T. W. DUERIG, A. PELTON and D. STöCKEL, *Mater. Sci. Engg*. A **273–275** (1999) 149.

Received 28 November 2002
and accepted 2 July 2003