TECHNISCHE MECHANIK, Band 25, Heft 1, (2005), 39– 58
Manuskripteingang: 2. August 2004

Pseudoelastic Behavior of Shape Memory Alloys: Constitutive Theory and Identification of the Material Parameters Using Neural Networks*

D. Helm

In shape memory alloys, some exceptional effects, like the one- and two-way shape memory effect, are observable. In a certain temperature range, the so-called pseudoelastic effect due to stress induced martensitic phase transitions is additionally observed in NiTi and other shape memory alloys. To identify the pseudoelastic material behavior of NiTi shape memory alloys, tension tests with different strain-rates are discussed. The observed phenomena are modeled within the framework of continuum thermomechanics regarding a geometric linear theory. The proposed model consists of a free energy function in order to represent the occurring energy storage and release effects. Additionally, evolution equations for internal variables, like the inelastic strain tensor and the fraction of martensite, are introduced. The developed system of constitutive equations represents the observed history-dependent material behavior. For the identification of the material parameters, the theory of neural networks is applied. Finally, the numerical simulations show a good agreement between the experimental observations and the predictions of the phenomenological model.

1 Introduction

The pseudoelastic behavior of shape memory alloys enables exceptional products like for example stents, endoscopic devices for minimally invasive surgery, and eyeglass frames. In order to develop and improve such innovative products, the structural behavior must be simulated in the framework of thermomechanics. The pseudoelastic properties of shape memory alloys are well known from uniaxial tension and torsion as well as biaxial experiments: e.g. Funakubo (1987), Helm (2001), Huo and Müller (1993), Lim and McDowell (1999), Raniecki et al. (2001), Shaw and Kyriakides (1995), and Tobushi et al. (1998). Additionally, the microstructural origin of the observed macroscopic behavior is well understood (cf. Funakubo (1987), Miyazaki (1996), and Otsuka and Wayman (1998)): In a certain temperature range, stress induced phase transitions from the parent phase (austenite) into martensite occur if a critical stress is reached. Due to this stress state, the atoms move in a cooperative manner until the crystal structure of the produced phase (martensite) is reached. The occurring phase transitions take place without diffusion processes. In the case of NiTi, the austenite is body-centered cubic and the martensite consists of a monoclinic lattice. Consequently, the martensitic phase transition is accompanied by a remarkable shear deformation. In contrast to this, the change in volume is negligible. Figure 1 illustrates the discussed material behavior. Therein, the abbreviations A for austenite and $\vec{\mathrm{M}}$ for oriented martensite as well as M_{f} (martensite finish), M_{s} (martensite start), A_{s} (austenite start), and A_{f} (austenite finish) for the characteristic phase transition temperatures are used. Additionally, the temperature M_{d} represents the beginning of plastic deformations due to the production and motion of dislocations. Above M_{d}, the necessary stress to induce plastic deformations is lower than the stress necessary to produce martensitic phase transitions. Between A_{f} and M_{d}, the pseudoelastic behavior of shape memory alloys is observed. During loading, phase transitions from austenite into martensite take place. In the case of NiTi, a nearly horizontal phase transition plateau is often observed if the temperature is approximately constant. After a certain deformation, the phase transition is almost finished and the produced martensite is stretched elastically. During unloading the retransformation occurs at a lower stress level. After loading, the material reaches the initial configuration. Therefore, this material behavior is called pseudoelasticity.

In the last decades, material models of different types have been proposed in the literature: A number of models describe merely the uniaxial behavior like e.g. Müller (1982), Seelecke (1997), Falk (1983), and Frémond (1996). These models are able to represent the behavior of one-dimensional structures like wires. However, the simulation of complex structures requires three-dimensional constitutive models: e.g. Boyd and Lagoudas (1994), Helm and

*This article is based on a contribution to the 11th SPIE Annual International Symposium on Smart Structures and Materials: cf. Helm (2004).

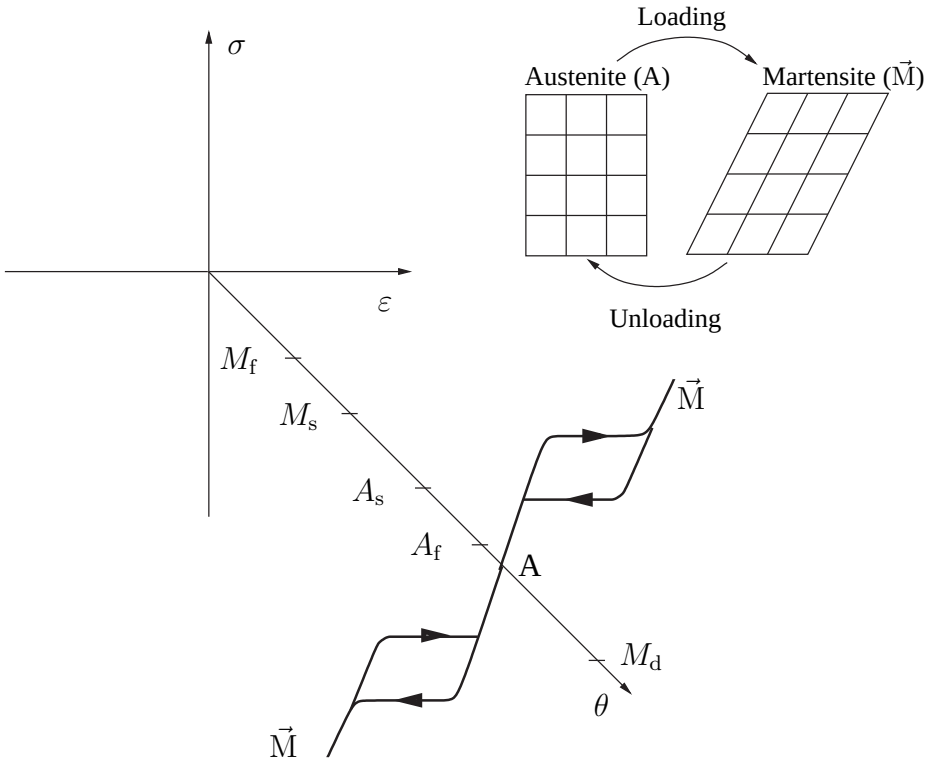

Figure 1: Pseudoelasticity due to stress induced phase transitions

Haupt (2003), Lubliner and Auricchio (1996), Raniecki et al. (1992), and Raniecki and Lexcellent (1994). Due to the strong thermomechanical coupling phenomena reported for example in Leo et al. (1993), Shaw and Kyriakides (1995), and recently in Helm and Haupt (2001), an adequate constitutive theory must be formulated in the context of thermomechanics.

The article is organized as follows: In Sect. 2, the results of tension experiments are discussed. For the representation of the pseudoelastic behavior of shape memory alloys, a fully three-dimensional material model is proposed in the framework of continuum thermomechanics (Sect. 3): The introduced free energy function is the result of a two-phase mixture between austenite and martensite. Additionally, evolution equations for internal variables are defined. Altogether, the material model contains a lot of material parameters. In order to identify these material parameters, the theory of neural networks is applied (Sect. 4). Here, the neural network is trained by results of appropriate simulations. Thereafter, the experimental data are used as input variables for the identification process. A part of the discussed tension test results are applied to identify the set of material parameters. The other data are useful to verify the proposed theory.

2 Experimental Investigations

In a previous experimental study, cf. Helm (2001) and Helm and Haupt (2001), the material behavior of thin-walled NiTi tubes was investigated. Some results from this series of experiments are summarized in this section. The experimental setup and the tested material is explained in Helm (2001) and Helm and Haupt (2001): To summarize, a hydraulic testing machine is used to apply the mechanical loads, the deformation is measured by means of an extensometer (initial length $l_0 = 12.5$mm, measurement range ± 2.5mm), the axial force is determined with a load cell (measurement ranges: ± 5kN), and the surface temperature field of the specimen is recorded by means of an infrared thermography system (minimal resolution ± 0.03K). Using this experimental setup, the thermomechanical properties of the investigated shape memory alloy are observed and not only the pure mechanical behavior. The investigated tubes consist of a Ni-content of 55.92 wt.-% and a complementary Ti-content of 44.08wt.-%. The tubes have an outer diameter of 4.674mm and the inner diameter amounts 3.978mm. The specimen has a full length of 80mm and a free length of 30mm. At room temperature the material is able to show pseudoelasticity due to an A_f-temperature of 11.4°C.

First, the pure tension behavior is studied using a relatively slow strain rate of $|\dot{\varepsilon}| = 0.0001\mathrm{s}^{-1}$: Figure 2 depicts

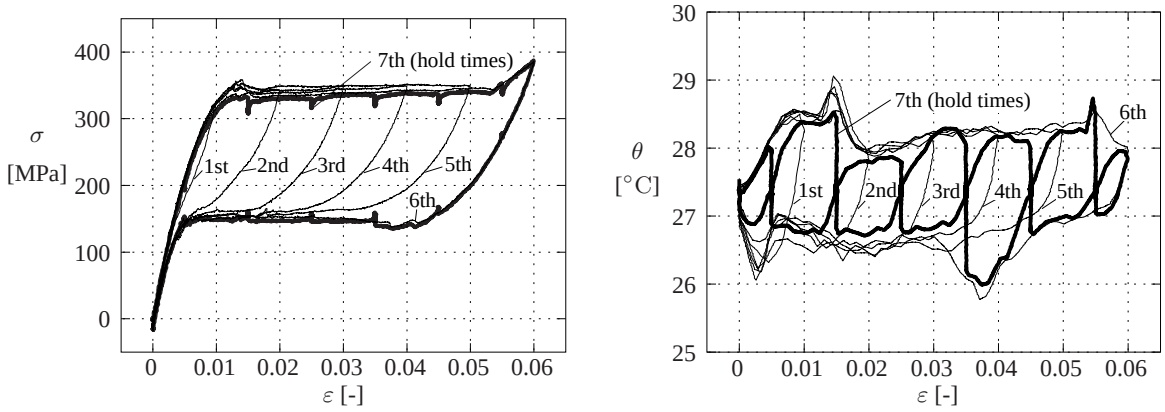

Figure 2: Step test: $|\dot{\varepsilon}| = 0.0001\text{s}^{-1}$, hold times 600s

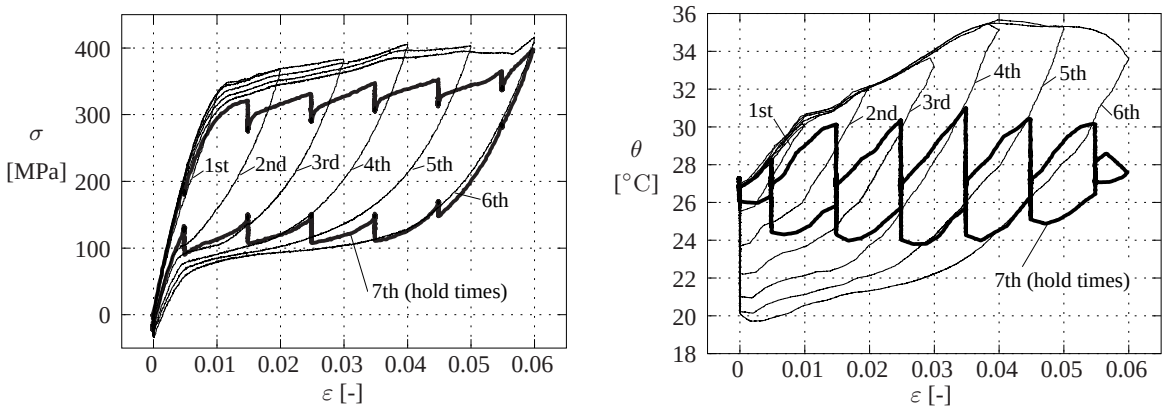

Figure 3: Step test: $|\dot{\varepsilon}| = 0.001\text{s}^{-1}$, hold times 600s

the stress-strain diagram as well as the accompanying temperature vs. strain curve. The shown temperature values are average temperatures arising from a small area between the blades of the extensometer. In the so-called step test, the strain amplitude increases stepwise from $\varepsilon_{\max} = 0.01$ in the first cycle to $\varepsilon_{\max} = 0.06$ in the last cycle. Thereafter, a full cycle up to $\varepsilon_{\max} = 0.06$ is executed and the deformation process is interrupted by 12 hold times of 600s at $\varepsilon = 0.005, 0.015, \ldots, 0.055, \ldots, 0.015, 0.005$. This strain controlled experiment is qualified to discuss the main properties of the investigated alloy: Up to a stress of approximately 75MPa, the material behavior is approximately elastic. But already at a strain far below $\varepsilon = 0.01$, partially martensitic phase transitions from austenite to martensite occur in grains with preferred orientations (cf. Brinson et al. (2004) and Li and Sun (2002)). Thereafter, the main phase transition initiates and progresses at nearly constant normal stress. Although the strain rate is relatively small, the temperature increases due to the exothermic nature of the austenite to martensite transformation. After the initiation of the phase transition in preferred regions it is well known that an inhomogeneous deformation takes place because phase transition fronts move through the specimen: cf. Brinson et al. (2002), Brinson et al. (2004), Li and Sun (2002), Shaw and Kyriakides (1995), and Shaw (2000). However, the investigations in Brinson et al. (2004) on polycrystalline NiTi verify that the phase transition occurs throughout the specimen. Only parts of a grain transform to martensite. Consequently, the macroscopically observed phase transition fronts represent regions, where a larger amount of austenite transforms to martensite. At the end of the phase transition plateau, a strong stress increase is observed. In this region, the microstructure consists of an austenite-martensite mixture (cf. Brinson et al. (2004), Helm (2001)), which is deformed elastically. However, further phase transitions likewise occur. Consequently, different loading and unloading tangents exist in this region. During unloading, the reverse transformation from martensite to austenite takes place after a small elastic region. All unloading processes lead to almost horizontal retransformation plateaus and, due to the endothermic nature of the reverse transformation, the temperature is below the initial temperature. After each loading/unloading path, the specimen returns approximately to its initial state. After the sixth step, a further loading/unloading process with hold times is carried out. During the hold times, both stress relaxation and temperature compensation processes take place.

In the second tension test, the same strain path is carried out, but the strain rate is ten times larger than in the foregoing test: $|\dot{\varepsilon}| = 0.001\mathrm{s}^{-1}$. The experimental results are depicted in Fig. 3. The observed behavior is based on the same deformation mechanisms. In contrast to the experiment with the slower strain rate, the occurring thermomechanical coupling phenomena lead to a stronger increase (decrease) in temperature during loading (unloading): E.g. the maximum rise in temperature is approximately 10K during loading. Additionally, the required stress to stimulate the martensitic phase transition from austenite to martensite is larger than in an isothermal experiment. In the last cycle, twelve hold times with zero strain rate are carried out. Both stress relaxation as well as temperature compensation effects occur. The reason for the stress behavior can be explained as follows: The required stress to initiate and progress the martensitic phase transition is strongly temperature-dependent. In relation to the investigated NiTi alloy, the necessary phase transition stress in order to stimulate the austenite-martensite transformation increases with 5.5MPa/K. In contrast to this, the reverse transition stress changes with 8MPa/K. Consequently, the temperature compensation processes during the hold times lead to stress relaxation effects. The whole amount of relaxation cannot be explained by temperature change alone: E.g., during the fourth hold time a stress relaxation of 41MPa and a temperature decrease of 4K takes place. However, the temperature-dependence of the phase transition stress leads merely to an influence of approximately 22MPa (4K·5.5MPa/K). Consequently, there must be an additional rate-dependent phenomena due to the viscosity of the material (cf. Helm and Haupt (2001)). This material property can be separately observed in the case of pseudoplasticity (cf. Helm (2001)).

3 Modeling in the Framework of Continuum Thermomechanics

Based on the experimental observations (see Sect. 2 and Helm and Haupt (2001)), a suitable system of constitutive equations should represent the following phenomena: the mixture of austenite and martensite during phase transitions, the different loading and unloading slopes in the region of the second elasticity, the temperature-dependent phase transition stress and other temperature-dependencies, as well as the thermomechanical coupling effects. In previous studies (cf. Helm (2001), Helm and Haupt (2002), and Helm and Haupt (2003)) a basic phenomenological model was developed in the context of continuum thermomechanics, which is able to represent some of the experimentally observed phenomena. In the present article, an enhanced model is proposed, which incorporates among other details the different elastic properties of austenite and martensite as well as an improved phase transition function and phase transition rule.

3.1 Introducing the Basic Structure of the Material Model

In order to model the experimentally observed energy storage and release phenomena, the basic structure of the free energy function is introduced as follows:

$$\rho\psi(\theta, \mathbf{E}_{\mathrm{e}}^{\mathrm{M}}, \mathbf{E}_{\mathrm{e}}^{\mathrm{A}}, z, \{\mathbf{Y}_{\mathrm{e}_i}\}) = z\rho\psi_{\mathrm{e}}^{\mathrm{M}}(\theta, \mathbf{E}_{\mathrm{e}}^{\mathrm{M}}) + (1-z)\rho\psi_{\mathrm{e}}^{\mathrm{A}}(\theta, \mathbf{E}_{\mathrm{e}}^{\mathrm{A}}) + \rho\psi_{\mathrm{s}}(\theta, \{\mathbf{Y}_{\mathrm{e}_i}\}). \tag{1}$$

Therein, ρ is the mass density of the total mixture consisting of the partial densities of the austenite phase $\rho^{\mathrm{A}} = (1-z)\rho$ and martensite phase $\rho^{\mathrm{M}} = z\rho$:

$$\rho = \rho^{\mathrm{M}} + \rho^{\mathrm{A}} = z\rho + (1-z)\rho. \tag{2}$$

The internal variable $z \in [0;1]$ represents the fraction of martensite. It should be remarked that the partial densities ρ^{A} and ρ^{M} are not the material densities of austenite and martensite (cf. Bowen (1969); Haupt (2002)). The partial density ρ^{ω} (ω =A, M) is the mass of the component ω per volume of the mixture. In Eq. (1), θ is the absolute thermodynamic temperature; $\mathbf{E}_{\mathrm{e}}^{\mathrm{M}}$ and $\mathbf{E}_{\mathrm{e}}^{\mathrm{A}}$ represent the elastic strain state in the martensite and austenite phase, respectively. Additionally, a set $\{\mathbf{Y}_{\mathrm{e}_i}\}$ of n internal variables $\mathbf{Y}_{\mathrm{e}_i}$ is introduced to model the energy storage resulting from internal stress fields, which are also called back stresses.

As discussed in Helm and Haupt (2003), it is sufficient to apply the 2nd Law of thermodynamics in form of the Clausius-Duhem inequality,

$$\theta\gamma = -\dot{\psi} - \dot{\theta}\eta + \frac{1}{\rho}\mathbf{T}\cdot\dot{\mathbf{E}} - \frac{1}{\rho\theta}\boldsymbol{q}\cdot\operatorname{grad}\theta \geq 0, \tag{3}$$

instead of other irreversibility constrains (cf. Hutter (1977), Jou et al. (1996), and Müller and Ruggeri (1998)). In the Clausius-Duhem inequality, η represents the entropy, $\mathbf{T}$ the stress tensor, $\mathbf{E}$ the linearized strain tensor, and $\boldsymbol{q}$ the heat flux vector. The developed model structure allows the partition of the Clausius-Duhem inequality into the internal dissipation inequality,

$$\delta = -\dot{\psi} - \dot{\theta}\eta + \frac{1}{\rho}\mathbf{T}\cdot\dot{\mathbf{E}} \geq 0, \tag{4}$$

and the heat conduction inequality,

$$-\frac{1}{\rho\theta}\boldsymbol{q}\cdot\,\mathrm{grad}\,\theta\geq 0. \tag{5}$$

If the Fourier-model for heat conduction, $\boldsymbol{q}=-\lambda\,\mathrm{grad}\,\theta$, with the coefficient $\lambda>0$, is applied, the heat conduction inequality is always satisfied.

In order to develop a thermodynamically consistent material model, the internal dissipation inequality (4) is studied for the proposed structure of the free energy function according to Eq. (1). In this sense, the basic structure of the free energy function is introduced into the internal dissipation inequality to give

$$\begin{aligned}\rho\delta=&-\rho\frac{\partial\psi}{\partial z}\dot{z}-z\rho\frac{\partial\psi_{\mathrm{e}}^{\mathrm{M}}}{\partial\mathbf{E}_{\mathrm{e}}^{\mathrm{M}}}\cdot\dot{\mathbf{E}}_{\mathrm{e}}^{\mathrm{M}}-(1-z)\rho\frac{\partial\psi_{\mathrm{e}}^{\mathrm{A}}}{\partial\mathbf{E}_{\mathrm{e}}^{\mathrm{A}}}\cdot\dot{\mathbf{E}}_{\mathrm{e}}^{\mathrm{A}}-z\rho\frac{\partial\psi_{\mathrm{e}}^{\mathrm{M}}}{\partial\theta}\dot{\theta}-(1-z)\rho\frac{\partial\psi_{\mathrm{e}}^{\mathrm{A}}}{\partial\theta}\dot{\theta}-\rho\frac{\partial\psi_{\mathrm{s}}}{\partial\theta}\dot{\theta}-\rho\eta\dot{\theta}\\&-\sum_{i=1}^{n}\rho\frac{\partial\psi_{\mathrm{s}}}{\partial\mathbf{Y}_{\mathrm{e}_i}}\cdot\dot{\mathbf{Y}}_{\mathrm{e}_i}+\mathbf{T}\cdot\dot{\mathbf{E}}\geq 0.\end{aligned} \tag{6}$$

Due to the choice of the free energy function (Eq. (1)), the dissipation inequality contains different internal variables of strain type: In the context of a geometric linear theory, the linearized Green strain tensor is decomposed into an elastic and inelastic part:

$$\mathbf{E}=\mathbf{E}_{\mathrm{e}}+\mathbf{E}_{\mathrm{d}}. \tag{7}$$

Here, the elastic part $\mathbf{E}_{\mathrm{e}}$ can be understood as an average value of the elastic deformations occurring in the austenite and the martensite phase. In order to describe the elastic interaction between the austenite and martensite phases, the most simple approach is the assumption that the strains are equal in both phases:

$$\mathbf{E}_{\mathrm{e}}=\mathbf{E}_{\mathrm{e}}^{\mathrm{M}}=\mathbf{E}_{\mathrm{e}}^{\mathrm{A}}. \tag{8}$$

Additionally, the inelastic strain tensor $\mathbf{E}_{\mathrm{d}}$ is additively decomposed n-times:

$$\mathbf{E}_{\mathrm{d}}=\mathbf{Y}_{\mathrm{e}_i}+\mathbf{Y}_{\mathrm{d}_i}\quad\text{with}\quad i=1,\ldots,n. \tag{9}$$

Thus, the model contains n internal variables $\mathbf{Y}_{\mathrm{e}_i}$ and n internal variables $\mathbf{Y}_{\mathrm{d}_i}$, all being of strain type. The variables $\mathbf{Y}_{\mathrm{d}_i}$ are introduced to model dissipation effects during the generation of internal stress fields. Combining Eqs. (7)-(9) and the internal dissipation inequality (6), the resulting inequality,

$$\begin{aligned}\rho\delta=&\left[\mathbf{T}-z\rho\frac{\partial\psi_{\mathrm{e}}^{\mathrm{M}}}{\partial\mathbf{E}_{\mathrm{e}}^{\mathrm{M}}}-(1-z)\rho\frac{\partial\psi_{\mathrm{e}}^{\mathrm{A}}}{\partial\mathbf{E}_{\mathrm{e}}^{\mathrm{A}}}\right]\cdot\dot{\mathbf{E}}_{\mathrm{e}}-\rho\left[z\frac{\partial\psi_{\mathrm{e}}^{\mathrm{M}}}{\partial\theta}+(1-z)\frac{\partial\psi_{\mathrm{e}}^{\mathrm{A}}}{\partial\theta}+\frac{\partial\psi_{\mathrm{s}}}{\partial\theta}+\eta\right]\dot{\theta}\\&-\rho\frac{\partial\psi}{\partial z}\dot{z}+\sum_{i=1}^{n}\rho\frac{\partial\psi_{\mathrm{s}}}{\partial\mathbf{Y}_{\mathrm{e}_i}}\cdot\dot{\mathbf{Y}}_{\mathrm{d}_i}+\left[\mathbf{T}-\sum_{i=1}^{n}\rho\frac{\partial\psi_{\mathrm{s}}}{\partial\mathbf{Y}_{\mathrm{e}_i}}\right]\cdot\dot{\mathbf{E}}_{\mathrm{d}}\geq 0,\end{aligned} \tag{10}$$

leads to potential relations for the stress tensor

$$\mathbf{T}=z\rho\frac{\partial\psi_{\mathrm{e}}^{\mathrm{M}}}{\partial\mathbf{E}_{\mathrm{e}}^{\mathrm{M}}}+(1-z)\rho\frac{\partial\psi_{\mathrm{e}}^{\mathrm{A}}}{\partial\mathbf{E}_{\mathrm{e}}^{\mathrm{A}}} \tag{11}$$

and the entropy

$$\eta=-\frac{\partial\psi}{\partial\theta}=-z\frac{\partial\psi_{\mathrm{e}}^{\mathrm{M}}}{\partial\theta}-(1-z)\frac{\partial\psi_{\mathrm{e}}^{\mathrm{A}}}{\partial\theta}-\frac{\partial\psi_{\mathrm{s}}}{\partial\theta}, \tag{12}$$

if the standard argumentation is applied. The remaining internal dissipation inequality,

$$\rho\delta=-\rho\frac{\partial\psi}{\partial z}\dot{z}+\sum_{i=1}^{n}\rho\frac{\partial\psi_{\mathrm{s}}}{\partial\mathbf{Y}_{\mathrm{e}_i}}\cdot\dot{\mathbf{Y}}_{\mathrm{d}_i}+\left[\mathbf{T}-\sum_{i=1}^{n}\rho\frac{\partial\psi_{\mathrm{s}}}{\partial\mathbf{Y}_{\mathrm{e}_i}}\right]\cdot\dot{\mathbf{E}}_{\mathrm{d}}\geq 0, \tag{13}$$

motivates the definition of the internal stress tensors $\mathbf{X}_{\varepsilon_{\mathrm{i}}}$ as well as the resultant of all internal stresses:

$$\mathbf{X}_{\varepsilon_{\mathrm{i}}}=\rho\frac{\partial\psi_{\mathrm{s}}}{\partial\mathbf{Y}_{\mathrm{e}_i}}\qquad\text{and}\quad\mathbf{X}_{\varepsilon}=\sum_{i=1}^{n}\mathbf{X}_{\varepsilon_{\mathrm{i}}}. \tag{14}$$

Using these definitions, the remaining dissipation inequality can be written in the following form,

$$\rho\delta=-\rho\frac{\partial\psi}{\partial z}\dot{z}+\sum_{i=1}^{n}\mathbf{X}_{\varepsilon_{\mathrm{i}}}\cdot\dot{\mathbf{Y}}_{\mathrm{d}_i}+[\mathbf{T}-\mathbf{X}_{\varepsilon}]\cdot\dot{\mathbf{E}}_{\mathrm{d}}\geq 0, \tag{15}$$

which is necessary for the next considerations: Due to the physical origin of the pseudoelastic effect, the fraction of martensite depends on the occurring inelastic deformations. In the case of pseudoelasticity, the stress-induced phase transitions lead to inelastic deformations. This phenomenon is represented in the material model by a simple function (cf. Helm and Haupt (2003), Juhász et al. (2000), Levitas (1998), and Souza et al. (1998))

$$z = \frac{\|\mathbf{E}_{\mathrm{d}}\|}{\sqrt{\frac{3}{2}}\gamma_{\mathrm{d}}} \qquad \Longrightarrow \qquad \dot{z} = \frac{\mathbf{E}_{\mathrm{d}} \cdot \dot{\mathbf{E}}_{\mathrm{d}}}{\sqrt{\frac{3}{2}}\gamma_{\mathrm{d}}\|\mathbf{E}_{\mathrm{d}}\|}, \tag{16}$$

which is similar to the definition of a limit function as proposed in Bertram (1982). Therein, γ_{d} is a positive material parameter, which has an effect on the width of the hysteresis. Finally, Eq. (16) and a zero in form of $\mathbf{X}_\theta \cdot \dot{\mathbf{E}}_{\mathrm{d}} - \mathbf{X}_\theta \cdot \dot{\mathbf{E}}_{\mathrm{d}} = 0$ is introduced into the internal dissipation inequality (15), leading to the compact formulation

$$\rho\delta = \left[\mathbf{X}_\theta - \frac{\rho\frac{\partial\psi}{\partial z}\mathbf{E}_{\mathrm{d}}}{\sqrt{\frac{3}{2}}\gamma_{\mathrm{d}}\|\mathbf{E}_{\mathrm{d}}\|}\right] \cdot \dot{\mathbf{E}}_{\mathrm{d}} + \sum_{i=1}^{n} \mathbf{X}_{\varepsilon_i} \cdot \dot{\mathbf{Y}}_{\mathrm{d}_i} + [\mathbf{T} - \mathbf{X}] \cdot \dot{\mathbf{E}}_{\mathrm{d}} \geq 0 \tag{17}$$

with the definition

$$\mathbf{X} = \mathbf{X}_\varepsilon + \mathbf{X}_\theta. \tag{18}$$

In analogy to Helm and Haupt (2003), the internal variable $\mathbf{X}_\theta$ is introduced in order to model the temperature-dependent phase transition stress.

The remaining internal dissipation inequality (17) is always satisfied if appropriate evolution equations for $\mathbf{E}_{\mathrm{d}}$ and $\mathbf{Y}_{\mathrm{d}_i}$ are introduced and if a phase transition criterion is defined:

- First, evolution equations of the form

$$\dot{\mathbf{E}}_{\mathrm{d}} = \lambda_{\mathrm{d}} \frac{\mathbf{T}^{\mathrm{D}} - \mathbf{X}^{\mathrm{D}}}{\|\mathbf{T}^{\mathrm{D}} - \mathbf{X}^{\mathrm{D}}\|} \qquad \text{with} \qquad \lambda_{\mathrm{d}} \geq 0 \tag{19}$$

  for the inelastic strain tensor and

$$\dot{\mathbf{Y}}_{\mathrm{d}_i} = \xi_i \mathbf{X}_{\varepsilon_i} \qquad \text{with} \qquad \xi_i \geq 0 \tag{20}$$

  for the internal variables $\mathbf{Y}_{\mathrm{d}_i}$ are suggested. Therein, the scalar-valued proportionality factors λ_{d} and ξ_i must be non-negative. These proportionality factors are specified in Sect. 3.3.1 and 3.3.2.

- In order to guarantee the positiveness of the internal dissipation, an appropriate phase transition criterion must be defined. Inserting the evolution equations according to Eq. (19) and (20) into the remaining dissipation inequality,

$$\rho\delta = \lambda_{\mathrm{d}} \underbrace{\left[\left[\mathbf{X}_\theta - \frac{\rho\frac{\partial\psi}{\partial z}\mathbf{E}_{\mathrm{d}}}{\sqrt{\frac{3}{2}}\gamma_{\mathrm{d}}\|\mathbf{E}_{\mathrm{d}}\|}\right] \cdot \frac{\left[\mathbf{T}^{\mathrm{D}} - \mathbf{X}^{\mathrm{D}}\right]}{\|\mathbf{T}^{\mathrm{D}} - \mathbf{X}^{\mathrm{D}}\|} + \|\mathbf{T}^{\mathrm{D}} - \mathbf{X}^{\mathrm{D}}\|\right]}_{= f + \sqrt{\frac{2}{3}}k(\theta) \geq \sqrt{\frac{2}{3}}k(\theta)} + \sum_{i=1}^{n} \xi_i \|\mathbf{X}_{\varepsilon_i}\|^2 \geq 0, \tag{21}$$

  the dissipation is always non-negative if the following phase transition criterion or yield criterion is applied:

$$f = \left[\mathbf{X}_\theta - \frac{\rho\frac{\partial\psi}{\partial z}\mathbf{E}_{\mathrm{d}}}{\sqrt{\frac{3}{2}}\gamma_{\mathrm{d}}\|\mathbf{E}_{\mathrm{d}}\|}\right] \cdot \frac{\left[\mathbf{T}^{\mathrm{D}} - \mathbf{X}^{\mathrm{D}}\right]}{\|\mathbf{T}^{\mathrm{D}} - \mathbf{X}^{\mathrm{D}}\|} + \underbrace{\|\mathbf{T}^{\mathrm{D}} - \mathbf{X}^{\mathrm{D}}\| - \sqrt{\frac{2}{3}}k(\theta)}_{f_{\mathrm{Mises}}} \geq 0. \tag{22}$$

  Therein, the temperature-dependent material parameter $k(\theta)$ has an effect on the height of the hysteresis. In general, the yield radius $k(\theta, \ldots)$ depends on temperature and additional internal variables due to the representation of effects similar to isotropic hardening in metal plasticity. The applied phase transition criterion represents a modified v. Mises yield function f_{Mises}. In the case of multiaxial deformations, the proposed theory is a non-associative yield theory, because the evolution equation for the inelastic strains according to Eq. (19) cannot be derived by a normality rule from the phase transition function f. It should be mentioned that e.g. in the case of non-associative plasticity the direction of the inelastic deformations can be derived from a pseudopotential (cf. Lemaitre and Chaboche (1990)).

3.2 Free Energy Function

In the foregoing section, the basic structure of the material model was developed on the basis that the free energy function as introduced in Eq. (1) generally depends on temperature and a set of internal variables. In this section, the free energy function is specified in detail in order to model the occurring energy storage effects.

3.2.1 Elastic Part of the Free Energy Function

For a single phase material, consisting of austenite, $\omega = \mathrm{A}$, or martensite, $\omega = \mathrm{M}$, the free energy function

$$\begin{aligned}\psi_{\mathrm{e}}^{\omega} = \hat{\psi}_{\mathrm{e}}^{\omega}(\mathbf{E}_{\mathrm{e}}^{\omega}, \theta) = \frac{\mu^{\omega}(\theta)}{\rho}\mathbf{E}_{\mathrm{e}}^{\omega^{\mathrm{D}}} \cdot \mathbf{E}_{\mathrm{e}}^{\omega^{\mathrm{D}}} + \frac{\kappa^{\omega}(\theta)}{2\rho}\left(\mathrm{tr}\mathbf{E}_{\mathrm{e}}^{\omega}\right)^2 - \frac{3\alpha^{\omega}(\theta)\kappa^{\omega}(\theta)}{\rho}\left(\mathrm{tr}\mathbf{E}_{\mathrm{e}}^{\omega}\right)(\theta - \theta_0) + \\ + \int_{\theta_0}^{\theta} c_{\mathrm{d}_0}^{\omega}(\bar{\theta})\mathrm{d}\bar{\theta} + u_0^{\omega} - \theta\left[\int_{\theta_0}^{\theta}\frac{c_{\mathrm{d}_0}^{\omega}(\bar{\theta})}{\bar{\theta}}\mathrm{d}\bar{\theta} + \eta_0^{\omega}\right]\end{aligned} \tag{23}$$

represents the energy storage due to elastic deformations and thermomechanical effects. In general, all material parameters are temperature-dependent.

The simple mixture of the austenite and martensite phase,

$$\rho\psi_{\mathrm{e}}(\mathbf{E}_{\mathrm{e}}^{\mathrm{M}}, \mathbf{E}_{\mathrm{e}}^{\mathrm{A}}, \theta, z) = z\rho\psi_{\mathrm{e}}^{\mathrm{M}}(\mathbf{E}_{\mathrm{e}}^{\mathrm{M}}, \theta) + (1-z)\rho\psi_{\mathrm{e}}^{\mathrm{A}}(\mathbf{E}_{\mathrm{e}}^{\mathrm{A}}, \theta), \tag{24}$$

leads to the following elastic part of the free energy function for the two-phase material,

$$\begin{aligned}\rho\psi_{\mathrm{e}}(\mathbf{E}_{\mathrm{e}}, \theta, z) = \left[z\mu^{\mathrm{M}}(\theta) + (1-z)\mu^{\mathrm{A}}(\theta)\right]\mathbf{E}_{\mathrm{e}}^{\mathrm{D}} \cdot \mathbf{E}_{\mathrm{e}}^{\mathrm{D}} - 3\left[z\alpha^{\mathrm{M}}(\theta)\kappa^{\mathrm{M}}(\theta) + (1-z)\alpha^{\mathrm{A}}(\theta)\kappa^{\mathrm{A}}(\theta)\right](\mathrm{tr}\mathbf{E}_{\mathrm{e}})(\theta - \theta_0) \\ + \frac{z\kappa^{\mathrm{M}}(\theta) + (1-z)\kappa^{\mathrm{A}}(\theta)}{2}(\mathrm{tr}\mathbf{E}_{\mathrm{e}})^2 + z\rho\int_{\theta_0}^{\theta} c_{\mathrm{d}_0}^{\mathrm{M}}(\bar{\theta})\mathrm{d}\bar{\theta} + (1-z)\rho\int_{\theta_0}^{\theta} c_{\mathrm{d}_0}^{\mathrm{A}}(\bar{\theta})\mathrm{d}\bar{\theta} \\ + \rho\left[u_0^{\mathrm{A}} + z\Delta u_0\right] - \theta\left[z\rho\int_{\theta_0}^{\theta}\frac{c_{\mathrm{d}_0}^{\mathrm{M}}(\bar{\theta})}{\bar{\theta}}\mathrm{d}\bar{\theta} + (1-z)\rho\int_{\theta_0}^{\theta}\frac{c_{\mathrm{d}_0}^{\mathrm{A}}(\bar{\theta})}{\bar{\theta}}\mathrm{d}\bar{\theta} + \rho\left[\eta_0^{\mathrm{A}} + z\Delta\eta_0\right]\right],\end{aligned} \tag{25}$$

if the assumption of equal elastic strains is applied (cf. Eq. (8)). Therein, the abbreviations

$$\Delta u_0 = u_0^{\mathrm{M}} - u_0^{\mathrm{A}} \qquad \text{and} \qquad \Delta\eta_0 = \eta_0^{\mathrm{M}} - \eta_0^{\mathrm{A}} \tag{26}$$

are introduced: Both Δu_0 and $\Delta\eta_0$ play an important role in the description of the heat production during martensitic phase transitions (cf. Helm and Haupt (2003), Huo and Müller (1993)).

3.2.2 Inelastic Part of the Free Energy Function

As discussed in Sect. 2, different slopes of the loading and unloading tangents are observed in the region of the second elasticity. To represent this phenomenon and also the stress evolution at the beginning of the phase transition plateaus, the inelastic part of the free energy function

$$\rho\psi_{\mathrm{s}} = \rho\hat{\psi}_{\mathrm{s}}(\theta, \{\mathbf{Y}_{\mathrm{e}_i}\}) = \sum_{i=1}^{n}\frac{c_i(\theta)}{2}\mathbf{Y}_{\mathrm{e}_i} \cdot \mathbf{Y}_{\mathrm{e}_i} \tag{27}$$

is proposed. In general, the coefficients $c_i(\theta)$ are non-negative temperature-dependent material parameters. If required, the free energy function may depend on further internal variables. For example, the experiments depicted in Fu et al. (1992) show that the number of interfaces between austenite and martensite is proportional to the fraction of martensite. Consequently, the modeling of such an effect in the present theoretical framework would require an inelastic part of the free energy, which depends on the fraction of martensite z.

3.2.3 Implications of the Free Energy Function

Due to the potential relation Eq. (11), the proposed free energy function according to Eq. (1) and (24)–(27) leads to the stress relation

$$\begin{aligned}\mathbf{T} = {} & 2\left[z\mu^{\mathrm{M}}(\theta) + (1-z)\mu^{\mathrm{A}}(\theta)\right]\mathbf{E}_{\mathrm{e}}^{\mathrm{D}} + \left[z\kappa^{\mathrm{M}}(\theta) + (1-z)\kappa^{\mathrm{A}}(\theta)\right](\mathrm{tr}\mathbf{E}_{\mathrm{e}})\mathbf{1} \\ & - 3\left[z\alpha^{\mathrm{M}}(\theta)\kappa^{\mathrm{M}}(\theta) + (1-z)\alpha^{\mathrm{A}}(\theta)\kappa^{\mathrm{A}}(\theta)\right](\theta-\theta_0)\,\mathbf{1}.\end{aligned} \tag{28}$$

Furthermore, the internal stress tensors $\mathbf{X}_{\varepsilon_{\mathrm{i}}}$ are defined according to Eq. (14):

$$\mathbf{X}_{\varepsilon_{\mathrm{i}}} = c_i(\theta)\mathbf{Y}_{\mathrm{e}_i}. \tag{29}$$

As a consequence of the proposed material model, the partial derivative

$$\begin{aligned}\rho\frac{\partial\psi}{\partial z} = {} & \left[\mu^{\mathrm{M}}(\theta) - \mu^{\mathrm{A}}(\theta)\right]\mathbf{E}_{\mathrm{e}}^{\mathrm{D}}\cdot\mathbf{E}_{\mathrm{e}}^{\mathrm{D}} - 3\left[\alpha^{\mathrm{M}}(\theta)\kappa^{\mathrm{M}}(\theta) - \alpha^{\mathrm{A}}(\theta)\kappa^{\mathrm{A}}(\theta)\right](\mathrm{tr}\mathbf{E}_{\mathrm{e}})(\theta-\theta_0) \\ & + \frac{\kappa^{\mathrm{M}}(\theta) - \kappa^{\mathrm{A}}(\theta)}{2}(\mathrm{tr}\mathbf{E}_{\mathrm{e}})^2 + \rho\left[\int_{\theta_0}^{\theta}\left[c_{\mathrm{d}_0}^{\mathrm{M}}(\bar\theta) - c_{\mathrm{d}_0}^{\mathrm{A}}(\bar\theta)\right]\mathrm{d}\bar\theta + \Delta u_0\right] \\ & - \rho\theta\left[\int_{\theta_0}^{\theta}\frac{c_{\mathrm{d}_0}^{\mathrm{M}}(\bar\theta) - c_{\mathrm{d}_0}^{\mathrm{A}}(\bar\theta)}{\bar\theta}\mathrm{d}\bar\theta + \Delta\eta_0\right]\end{aligned} \tag{30}$$

is of particular importance. In contrast to the previous study (Helm and Haupt (2003)), the consideration of different elastic properties of the austenite and the martensite phase leads to additional terms in the partial derivative $\frac{\partial\psi}{\partial z}$. As will be shown in the next section, these additional terms influence the phase transition range.

3.3 Evolution Equations for Internal Variables

To complete the evolution equations for the internal variables, the proportionality factors λ_{d} and ξ_i in the evolution equations for $\mathbf{E}_{\mathrm{d}}$ and $\mathbf{Y}_{\mathrm{d}_i}$ are defined in this section.

3.3.1 Evolution Equation for $\mathbf{E}_{\mathrm{d}}$

The final structure of the remaining dissipation inequality (17) motivates the evolution equation (19):

$$\dot{\mathbf{E}}_{\mathrm{d}} = \lambda_{\mathrm{d}}\,\mathbf{N} \qquad \text{with} \qquad \mathbf{N} = \frac{\mathbf{T}^{\mathrm{D}} - \mathbf{X}^{\mathrm{D}}}{\|\mathbf{T}^{\mathrm{D}} - \mathbf{X}^{\mathrm{D}}\|} = \frac{\mathbf{T}^{\mathrm{D}} - \left[\mathbf{X}_{\varepsilon}^{\mathrm{D}} + \mathbf{X}_{\theta}^{\mathrm{D}}\right]}{\|\mathbf{T}^{\mathrm{D}} - \left[\mathbf{X}_{\varepsilon}^{\mathrm{D}} + \mathbf{X}_{\theta}^{\mathrm{D}}\right]\|}. \tag{31}$$

Here, the abbreviation $\mathbf{N}$ for the direction of the inelastic strain rate is used. Obviously, the tensor $\mathbf{N}$ has the property $\|\mathbf{N}\| = 1$. The deviatoric evolution of $\mathbf{E}_{\mathrm{d}}$ is assumed, because the phase transition occurs approximately at constant volume.

Until now, the internal variable $\mathbf{X}_\theta$ of stress type was introduced in view of the basic structure of the theory, but a constitutive relation for $\mathbf{X}_\theta$ is still missing: This variable is an essential element of the model: it represents the temperature-dependent phase transition stress. Regarding the remaining dissipation inequality (21), the definition

$$\mathbf{X}_\theta = \frac{\rho\frac{\partial\psi}{\partial z}\mathbf{E}_{\mathrm{d}}}{\sqrt{\frac{3}{2}}\gamma_{\mathrm{d}}\|\mathbf{E}_{\mathrm{d}}\|} \tag{32}$$

could be reasonable to be introduced (see the discussion in Helm and Haupt (2003) and cf. Juhász et al. (2000) as well as Souza et al. (1998)). However, this equation contains the term $\|\mathbf{E}_{\mathrm{d}}\|$ in the denominator, which is sometimes zero. This leads to singularities in the tensor $\mathbf{N}$ and the phase transition function, but these singularities can be removed, if a v. Mises yield function is applied (see Helm and Haupt (2003) for details). In order to avoid the singularity in the tensor $\mathbf{N}$, the constitutive relation

$$\mathbf{X}_\theta = \frac{\rho\left\langle\frac{\partial\psi}{\partial z}\right\rangle\mathbf{E}^{\mathrm{D}}}{\sqrt{\frac{3}{2}}\gamma_{\mathrm{d}}\left(\|\mathbf{E}^{\mathrm{D}}\| + a\right)} \tag{33}$$

is suggested in the present model. Instead of the inelastic strain tensor $\mathbf{E}_\mathrm{d}$, the internal variable depends on the deviatoric part of the linearized strain tensor $\mathbf{E}$. In the additional case of pseudoplasticity due to stress induced phase transitions, $\|\mathbf{E}^\mathrm{D}\|$ can be zero: Therefore, the regularization parameter $a > 0$ is considered in avoidance of the singularity.

Finally, the inelastic multiplier λ_d is defined: In order to model the experimentally observed rate-dependent effect due to the viscosity of the material (cf. Sect. 2), an inelastic multiplier of Perzyna-type (cf. Perzyna (1963) and Hohenemser and Prager (1932)) is applied:

$$\lambda_\mathrm{d} = \begin{cases} \dfrac{1}{\eta_\mathrm{d}(\theta)} \left\langle \dfrac{f}{r_\mathrm{d}} \right\rangle^{m_\mathrm{d}(\theta)} & \mathrm{A} \rightarrow \vec{\mathrm{M}} \text{ if: } z < 1,\ \frac{\partial \psi}{\partial z} > 0, \text{ and } \mathbf{E}_\mathrm{d} \cdot \mathbf{N} \geq 0 \\ & \vec{\mathrm{M}} \rightarrow \mathrm{A} \text{ if: } z > 0,\ \frac{\partial \psi}{\partial z} > 0, \text{ and } \mathbf{E}_\mathrm{d} \cdot \mathbf{N} < 0 \\ 0 & \text{in all other cases} \end{cases} . \tag{34}$$

In general, the introduced material parameters η_d and m_d are temperature-dependent. In contrast to a general theory, r_d is merely used to obtain a dimensionless argument in the McCauley bracket. The McCauley bracket is defined by

$$\langle x \rangle = \frac{|x| + x}{2}. \tag{35}$$

In order to distinguish between the austenite $\rightarrow$ martensite and the reverse transformation, a few case distinctions are added: The $\mathrm{A} \rightarrow \vec{\mathrm{M}}$ phase transition occurs if austenite is available ($z < 1$), the partial derivative $\frac{\partial \psi}{\partial z}$ is positive, and $\mathbf{E}_\mathrm{d} \cdot \mathbf{N} \geq 0$. The last condition guarantees that the fraction of martensite increases (cf. Eq. (16)). In contrast to this, the reverse transformation $\vec{\mathrm{M}} \rightarrow \mathrm{A}$ takes place if orientated martensite $\vec{\mathrm{M}}$ is available ($z > 0$), the partial derivative $\frac{\partial \psi}{\partial z}$ is positive, and the fraction of martensite decreases $\mathbf{E}_\mathrm{d} \cdot \mathbf{N} < 0$.

The advantage of the new definition (33) for $\mathbf{X}_\theta$ is that the tensor $\mathbf{N}$ is well-defined and that $\mathbf{X}_\theta$ is a smooth function of $\mathbf{E}$. Through this, the movement of the phase transition surface in the stress space due to $\mathbf{X}_\theta$ is continuous. Unfortunately, the definition of $\mathbf{X}_\theta$ according to Eq. (33) leads also to a phase transition function (Eq. (22)), i.e.

$$f = \|\mathbf{T}^\mathrm{D} - \mathbf{X}^\mathrm{D}\| - \sqrt{\frac{2}{3}} \left[k - \frac{\rho}{\gamma_\mathrm{d}} \left[\frac{\left\langle \frac{\partial \psi}{\partial z} \right\rangle \mathbf{E}^\mathrm{D}}{\|\mathbf{E}^\mathrm{D}\| + a} - \frac{\frac{\partial \psi}{\partial z} \mathbf{E}_\mathrm{d}}{\|\mathbf{E}_\mathrm{d}\|} \right] \cdot \mathbf{N} \right], \tag{36}$$

wherein a singularity at $\|\mathbf{E}_\mathrm{d}\| = 0$ exists. However, the singularity in f is simply removable: In the evolution equation for $\mathbf{E}_\mathrm{d}$ (Eq. (31)), the time derivative is approximated by the finite difference:

$$\frac{\mathbf{E}_\mathrm{d}(t + \Delta t) - \mathbf{E}_\mathrm{d}(t)}{\Delta t} = \lambda_\mathrm{d}\, \mathbf{N}. \tag{37}$$

The case $\|\mathbf{E}_\mathrm{d}(T)\| = 0$ is valid at the beginning of the phase transition. The associated time is denoted by T and Eq. (37) merge into

$$\frac{\mathbf{E}_\mathrm{d}(T + \Delta t)}{\Delta t} = \lambda_\mathrm{d}\, \mathbf{N} \tag{38}$$

with the property

$$\|\mathbf{E}_\mathrm{d}(T + \Delta t)\| = \lambda_\mathrm{d} \Delta t \tag{39}$$

due to $\|\mathbf{N}\| = 1$. Consequently, the ratio

$$\frac{\mathbf{E}_\mathrm{d}}{\|\mathbf{E}_\mathrm{d}\|} = \mathbf{N} = \frac{\mathbf{T}^\mathrm{D} - \mathbf{X}^\mathrm{D}}{\|\mathbf{T}^\mathrm{D} - \mathbf{X}^\mathrm{D}\|}. \tag{40}$$

is determined for $\|\mathbf{E}_\mathrm{d}\| = 0$. Only in this case ($\|\mathbf{E}_\mathrm{d}\| = 0$), the phase transition function

$$f = \|\mathbf{T}^\mathrm{D} - \mathbf{X}^\mathrm{D}\| - \sqrt{\frac{2}{3}} \left[k - \frac{\rho}{\gamma_\mathrm{d}} \left[\frac{\left\langle \frac{\partial \psi}{\partial z} \right\rangle \mathbf{E}^\mathrm{D} \cdot \mathbf{N}}{\|\mathbf{E}^\mathrm{D}\| + a} - \frac{\partial \psi}{\partial z} \right] \right] \tag{41}$$

is valid.

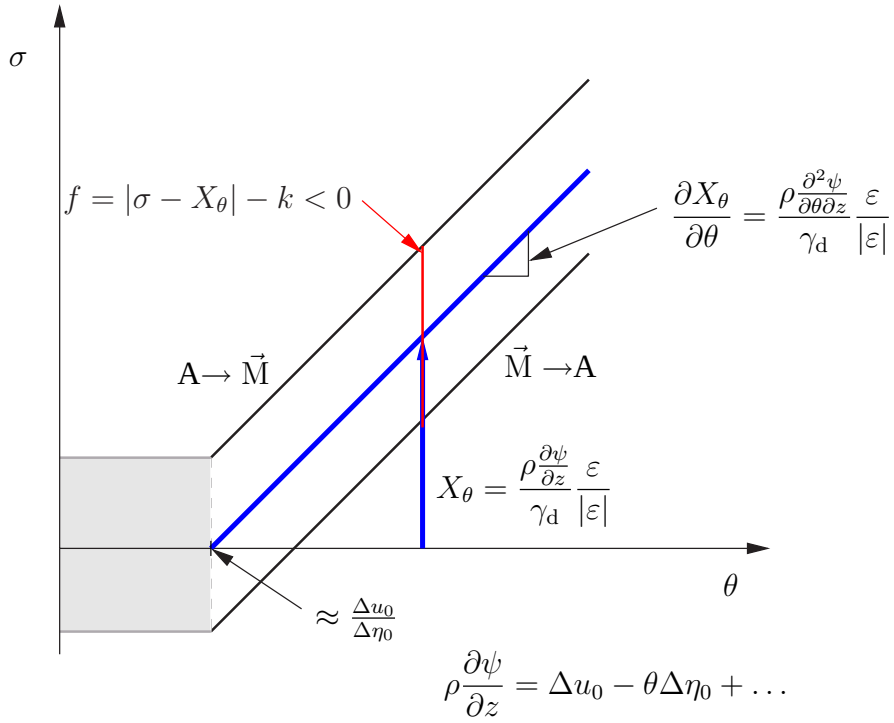

Figure 4: Basic structure of the constitutive theory in a stress-temperature diagram

3.3.2 Evolution Equation for $\mathbf{Y}_{\mathrm{d}_i}$

The non-negative entropy production is guaranteed if the internal variables $\mathbf{Y}_{\mathrm{d}_i}$ are proportional to $\mathbf{X}_{\varepsilon_\mathrm{i}}$ with non-negative proportionality factors ξ_i (cf. Eq. (20)). For the modeling of the different slopes in the loading and unloading path in a stress-strain diagram (see Fig. 2), a proportionality factor ξ_i depending on the temperature θ, the rate of the accumulated inelastic strain

$$\dot{s}_\mathrm{d} = \sqrt{\frac{2}{3}} \|\dot{\mathbf{E}}_\mathrm{d}\|, \tag{42}$$

and the fraction of martensite z is specified:

$$\xi_i = \hat{\xi}_i(\theta, \dot{s}_\mathrm{d}, z) = \begin{cases} \frac{1}{2}\left(\tanh\left[-\beta_i(\theta)(z - \gamma_i(\theta))\right] + 1\right) \frac{b_i(\theta)}{c_i(\theta)} \dot{s}_\mathrm{d} & \text{if: } \dot{z} > 0 \\ \frac{b_i(\theta)}{c_i(\theta)} \dot{s}_\mathrm{d} & \text{if: } \dot{z} < 0 \end{cases} . \tag{43}$$

In general, the material parameters b_i, β_i, and γ_i depend on temperature.

3.4 Effect of the Internal Variable $\mathbf{X}_\theta$ on the Phase Transition Stress

Figure 4 outlines the effect of the introduced internal variable $\mathbf{X}_\theta$ on the stress necessary to initiate and progress martensitic phase transitions. In this Figure, the uniaxial case is considered: $\mathbf{T} \rightarrow \sigma$, $\mathbf{X}_\varepsilon = \mathbf{0}$, $\mathbf{X}_\theta \rightarrow X_\theta$, $\mathbf{E}_\mathrm{d} \rightarrow \varepsilon_\mathrm{d}$, and $\mathbf{E} \rightarrow \varepsilon$. For simplicity, $a = 0$ is assumed. The phase transition from austenite to martensite starts in such a diagram if the stresses cross the $\mathrm{A} \rightarrow \vec{\mathrm{M}}$-start line. In contrast to this, the retransformation starts if the stress is lower than the $\vec{\mathrm{M}} \rightarrow \mathrm{A}$-start line. The introduced internal variable X_θ represents the center line between the martensite and austenite start lines. Due to the partial derivative $\frac{\partial \psi}{\partial z} = \Delta u_0 - \theta \Delta \eta_0 + \ldots$, the internal variable X_θ depends approximately linear on temperature, because the ratio $\varepsilon / |\varepsilon|$ is plus or minus one. The root of X_θ is nearly given by $\Delta u_0 / \Delta \eta_0$ and the slope of X_θ is mainly influenced by $\Delta \eta_0$. Furthermore, the elastic range is described by $f < 0$. Due to the regarded uniaxial case, the additional non-v. Mises terms in the phase transition function (Eq. (22)) vanishes. Consequently, the uniaxial phase transition function is given by $f = |\sigma - X_\varepsilon - X_\theta| - k$.

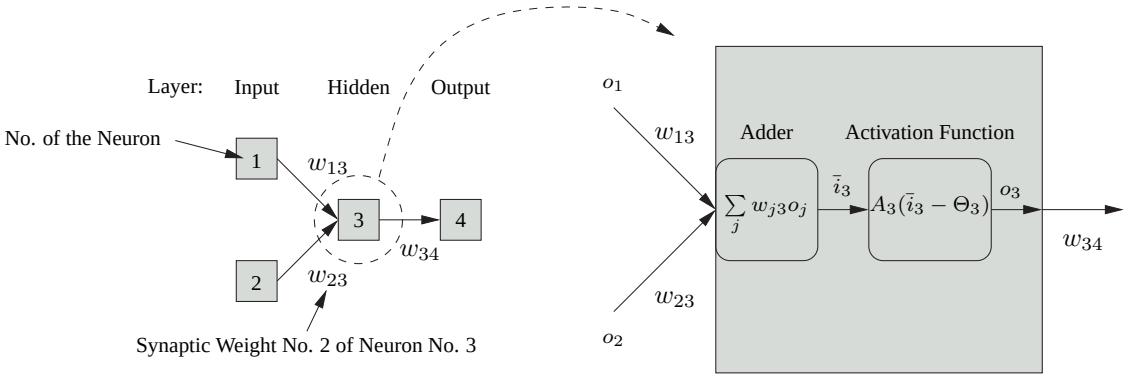

Figure 5: Simple example of a feedforward network

4 Identification of the Material Parameters using the Theory of Neural Networks

In the present study, the theory of neural networks is applied in order to identify the material parameters of the proposed constitutive model. The theory of neural networks is described in many textbooks: e.g. Haykin (1994), Rojas (1996), and Zell (1994). A short description of neural networks in Sect. 4.1 and the learning strategy in Sect. 4.2 is explained on the basis of these textbooks. In order to apply the method of neural networks in Sect. 4.3, the software package SNNS (Stuttgart Neural Network Simulator; SNNS (2000) and Zell (1994)) is used. The application of neural networks in order to identify the material parameters is a well-known strategy: cf. Huber (2000) and Huber and Tsakmakis (2001) and the detailed citations therein. The neural network method has the advantage that a trained network is able to identify the temperature-dependent material parameters of different alloys.

4.1 Basic Structure of Multilayered Neural Networks

Basically, an artifical neural network consists of neurons and connecting links, which are also called synapses. In the theory of neural networks, different kinds of networks are introduced. In the present context, so-called feedforward networks, as exemplarily shown in Fig. 5, are applied: In a feedforward network with at least one hidden layer (so-called multilayer networks), the neurons are connected from an input layer over the hidden layer(s) to an output layer. Consequently, connections in the opposite direction (feedback) and connections between neurons of the same layer do not exist in a feedforward network. If the feedforward network is fully connected, a neuron can only be chained with neurons in the next layer. In contrast to this, if shortcut connections are considered, a neuron can be aligned with neurons in all next layers. A synaptic weight w_{jk} is assigned to a connecting link. Here, the synaptic weight w_{jk} represents a connection from the neuron j to neuron k.

In a natural classification, the neurons are assigned to different layers. In the present context, the layers are counted from zero (input layer) to L (output layer). If the neuron k is an element of the layer number m, the neuron numerator k is a component of the set

$$\mathcal{L}_m = \{k \in \mathbb{N} | \text{No. of the neurons in the layer number } m\}. \tag{44}$$

The task of the neurons in the input layer $\mathcal{I} = \mathcal{L}_0$ is simple, because the input value i_k is equal to the output o_k of the neuron:

$$o_k = i_k \qquad \forall \quad k \in \mathcal{I}. \tag{45}$$

A simple mathematical model of a neuron in the other layers can be introduced as follows (cf. Fig. 5): First, the neuron k consists of an adder for summing up the input signals. In a linear combiner, a single input signal of the neuron k is equal to the output of the neuron j multiplied by the synaptic weight w_{jk}:

$$\bar{i}_k = \sum_j w_{jk} o_j. \tag{46}$$

As a result of this model, a positive weight $w_{jk} > 0$ amplifies and a negative value $w_{jk} < 0$ reduces the influence of the output o_j.

In order to represent the degree of stimulus, a component of the neuron is an activation function A_k:

$$a_k = A_k(\bar{i}_k - \Theta_k). \tag{47}$$

Therein, Θ_k is the threshold value, which models the stimulus threshold of a neuron. Different kinds of activation functions can be used in the theory of neural networks: e.g. the identity function

$$A(x) = x \tag{48}$$

or the logistic function

$$A(x) = \frac{1}{1 + \mathrm{e}^{-x}}, \tag{49}$$

which is a sigmoid function. In general, the activation function must not be differentiable, but some learning strategies require an appropriate differentiability (see Sect. 4.2).

In a natural partition of a multilayer network, the input layer $\mathcal{I} = \mathcal{L}_0$ is followed by layers of hidden neurons. The neurons in a hidden layer with the number $0 < m < L$ are specified by the set $\mathcal{H}_m = \mathcal{L}_m$. Additionally, the hidden layers are followed by an output layer: $\mathcal{O} = \mathcal{L}_L$. The set of all neurons is given by $\mathcal{L}$, the set of all neurons of the first f-layers is called $\mathcal{P}_f$, and the set of all hidden layers is denoted by $\mathcal{H}$:

$$\mathcal{L} = \bigcup_{m=0}^{L} \mathcal{L}_m, \qquad \mathcal{P}_f = \bigcup_{m=0}^{f} \mathcal{L}_m, \qquad \text{and} \qquad \mathcal{H} = \bigcup_{m=1}^{L-1} \mathcal{H}_m. \tag{50}$$

Due to simple algebraic operations, the output of a neural network can be calculated:

$$o_k = A_k\Big(\sum_{j \in \mathcal{P}_{L-1}} w_{jk} o_j - \Theta_k\Big) \qquad \forall \quad k \in \mathcal{O}. \tag{51}$$

However, the synaptic weights w_{jk} and the threshold values are unknown. The determination of these variables is one of the major topic in neural network technology. A well-known learning method is described in the next section.

4.2 Learning due to Supervised Training Processes using the Back-Propagation Algorithm

Before the so-called back-propagation algorithm is described, the structure of the neural network is slightly reorganized: In order to regard the threshold values Θ_k effectively in the mathematical framework of training algorithms, an input neuron $i_0 = -1$ is introduced in the network. Consequently, additional weights w_{0k} with the property

$$w_{0k} = \Theta_k \tag{52}$$

are considered for each neuron k. Through this, the threshold value can be eliminated from the activation function:

$$o_k = A_k(\bar{i}_k). \tag{53}$$

The determination of the synaptic weights w_{jk} can be done in different ways. However, to determine the weights a set of training data is necessary. In a supervised training process, the input of the neurons in the input layer and output of the neurons in the output layer must be known: e.g. as the result of numerical simulations. Through this, the synaptic weights can be determined with different methods: E.g. if an appropriate error measure is additionally defined, classical optimization strategies can be applied. In the classical back-propagation algorithm, the sum $E(\mathbf{W})$ of the squared error $e(p, \mathbf{W})$ over all training patterns p is regarded as error function:

$$E(\mathbf{W}) = \sum_p e(p, \mathbf{W}) \to \min \quad \text{with} \quad e(p, \mathbf{W}) = \frac{1}{2} \sum_{k \in \mathcal{O}} \left[t_k(p) - o_k(p, \mathbf{W})\right]^2. \tag{54}$$

Therein, $t_k(p)$ represent the output value of the neuron k in the training pattern p and $\mathbf{W}$ is the weight matrix consisting of the elements w_{jk}. The weights of non-existing links are zero. To find the minimum of E, the back-propagation algorithm uses the method of gradient descent, which is given by the series

$${}^{i+1}w_{mn} = {}^{i}w_{mn} + \Delta w_{mn} \tag{55}$$

with the weight correction terms

$$\Delta w_{mn} = -\eta_{\mathrm{L}} \frac{\partial E(\mathbf{W})}{\partial w_{mn}} = -\eta_{\mathrm{L}} \sum_p \frac{\partial e(p, \mathbf{W})}{\partial w_{mn}}. \tag{56}$$

Therein, η_{L} is the so-called learning-rate, which is very important in a gradient descent method in order to obtain a convergent series ${}^{i+1}w_{mn}$. In Eq. (56), the gradient $\frac{\partial e(p,\mathbf{W})}{\partial w_{mn}}$ must be determined. Therefore, the chain-rule is applied,

$$\frac{\partial e(p, \mathbf{W})}{\partial w_{mn}} = \underbrace{\frac{\partial e(p, \mathbf{W})}{\partial o_n} \frac{\partial o_n(p)}{\partial i_n}}_{-\delta_n(p)} \frac{\partial i_n(p)}{\partial w_{mn}}, \tag{57}$$

wherein the abbreviation

$$\delta_n(p) = -\frac{\partial e(p, \mathbf{W})}{\partial o_n} \frac{\partial o_n(p)}{\partial i_n} \tag{58}$$

is introduced for simplification: $\delta_n(p)$ is called the local gradient. The first partial derivative $\frac{\partial e(p,\mathbf{W})}{\partial o_n}$ will be discussed at the end, because merely the difference between the training data t_k and the output neurons o_k of the output layer $\mathcal{O}$ are calculated in the error function. The second derivative is nothing else as the derivative of the applied activation function:

$$\frac{\partial o_n(p)}{\partial i_n} = \frac{\partial A_n(i_n(p))}{\partial i_n} = A_n'(i_n(p)). \tag{59}$$

In the back-propagation algorithm, the differentiability of the activation function is necessary. The last partial derivative is equal to the output value of the neuron m:

$$\frac{\partial i_n(p)}{\partial w_{mn}} = \frac{\partial}{\partial w_{mn}} \sum_m w_{mn} o_m(p) = o_m(p). \tag{60}$$

Therefore, the weight correction terms can be written in the following form :

$$\Delta w_{mn} = \eta_{\mathrm{L}} \sum_p \delta_n(p) o_m(p). \tag{61}$$

However, for the determination of the first partial derivative $\frac{\partial e(p,\mathbf{W})}{\partial o_n}$ in Eq. (58), two cases must be considered:

1. If the output o_n belong to the neurons in the output layer (i.e. $n \in \mathcal{O}$), the partial derivative

$$\frac{\partial e(p, \mathbf{W})}{\partial o_n} = -\left[t_n(p) - o_n(p, \mathbf{W})\right] \tag{62}$$

   is simply calculable and the belonging local gradient is given by

$$\delta_n(p) = -\frac{\partial e(p, \mathbf{W})}{\partial i_n} = A_n'(i_n(p)) \left[t_n(p) - o_n(p, \mathbf{W})\right]. \tag{63}$$

2. The second case regards output neurons in the set of hidden neurons. These neurons are hidden due to the neurons in the output layer. Consequently, the chain-rule must be carefully applied. Due to the considerations in Appendix A, the local gradient δ_n of a hidden neuron in the hidden layer number m is given by

$$\delta_n = A_n'(i_n(p)) \sum_{l \in \mathcal{L}_{m+1}} \delta_l w_{nl} \qquad \forall \quad n \in \mathcal{H}_m = \mathcal{L}_m. \tag{64}$$

   Here, the local gradient δ_n of a neuron in the hidden layer $\mathcal{H}_m$ is influenced by all local gradients of connected neurons, which lie in the next layer $\mathcal{L}_{m+1}$ (for hidden layers: $0 < m < (L-1)$).

On the basis of the equations above, the weight correction terms Δw_{mn} can be calculated:

$$\Delta w_{mn} = \eta_{\mathrm{L}} \sum_p \delta_n(p) o_m(p) = \begin{cases} \eta_{\mathrm{L}} \sum_p A_n'(i_n(p)) \left(t_n(p) - o_n(p, \mathbf{W})\right) o_m(p) & \text{if: } n \in \mathcal{O} \\ \eta_{\mathrm{L}} \sum_p A_n'(i_n(p)) \left(\sum_{l \in \mathcal{L}_{m+1}} \delta_l w_{nl} \right) o_m(p) & \text{if: } n \in \mathcal{H} \end{cases}. \tag{65}$$

This equation clarifies the used terminology: back-propagation. The determination of the weights starts in the output layer. Thereafter, the δ_n's in the hidden layers can be determined layer-by-layer, because the values δ_l's in the summation over l are known from the foregoing regarded layer.

In the exact method of gradient descent (cf. Eq. (65)), the weight correction terms of each pattern p must be summed up. This back-propagation algorithm is often called batch back-propagation and requires more local storage for the synaptic weights as the so-called pattern-by-pattern mode: This modification changes the synaptic weights after each presentation of a training pattern, i.e. $\Delta w_{mn} = \eta_{\mathrm{L}} \delta_n(p) o_m(p)$. To avoid that informations about the sequence of training patterns are learned by the network in a pattern-by-pattern mode, the training data are shuffled after each training epoch. An epoch is finished if the network was trained with all patterns. In addition to this modification (pattern-by-pattern mode), a lot of additional modifications of the back-propagation algorithm are developed in order to improve the properties of the neural network as well as the training algorithm. An useful modification, which is used in the next section, is the so-called weight decay: The minimization of the mean square error according to Eq. (54) leads in some cases to large synaptic weights. Consequently, the participated neurons show strong activations and the according derivative of the activation function is in the worst-case approximately zero, which is disadvantageous in a gradient decent method. Applying the modified error function

$$E(\mathbf{W}) = \sum_p e(p, \mathbf{W}) \rightarrow \min \qquad \text{with} \qquad e(p, \mathbf{W}) = \frac{1}{2} \sum_{k \in \mathcal{O}} \left[t_k(p) - o_k(p, \mathbf{W}) \right]^2 + \frac{\beta}{2} \sum_{m,n \in \mathcal{L}} {}^i w_{mn}, \tag{66}$$

large synaptic weights are penalized due to the penalty factor β. Through this modification, the method of gradient descent leads to

$$\Delta w_{mn} = -\eta_{\mathrm{L}} \frac{\partial E(\mathbf{W})}{\partial w_{mn}} = \sum_p \left[-\eta_{\mathrm{L}} \frac{\partial e(p, \mathbf{W})}{\partial w_{mn}} - \beta \, {}^i w_{mn} \right] = \sum_p \left[\eta_{\mathrm{L}} \delta_n(p) o_m(p) - \beta \, {}^i w_{mn} \right]. \tag{67}$$

4.3 Application of Neural Network Technology to Identify the Material Parameters

The proposed constitutive theory contains a series of material parameters, some of them are in general temperature-dependent: The representation of the thermoelastic part of the model requires the identification of eleven material parameters (μ^{A}, μ^{M}, κ^{A}, κ^{M}, α^{A}, α^{M}, $c^{\mathrm{A}}_{\mathrm{d}_0}$, $c^{\mathrm{M}}_{\mathrm{d}_0}$, ρ, Δu_0, and $\Delta \eta_0$). In contrast to these parameters, the values of u_0^{A} and η_0^{A} are not of interest. For each internal stress state $\mathbf{X}_{\varepsilon_{\mathrm{i}}}$, the inelastic part of the free energy function (Eq. (27)) contains one material parameter c_i. In the evolution equation of the inelastic strain tensor $\mathbf{E}_{\mathrm{d}}$, two material parameters must be identified: η_{d} and m_{d}. The parameter r_{d} is used to obtain a dimensionless argument in the McCauley bracket. In addition, the yield radius k in the yield function and the parameter γ_{d} in $z = \hat{z}(\mathbf{E}_{\mathrm{d}})$ must be identified. Finally, the evolution equation for $\mathbf{Y}_{\mathrm{d}_{\mathrm{i}}}$ contains for each internal stress state $\mathbf{X}_{\varepsilon_{\mathrm{i}}}$ three material parameters: b_i, β_i, and γ_i. In all, the model contains $(15 + 4 \cdot n)$ material parameters. Due to the large number of material parameters, the identification of these parameters is a challenge.

In a first study, the problem is reduced as follows: An experiment with an appropriate slow loading rate is regarded as nearly isothermal. In such a case, the specific heat capacities and the thermal expansion coefficients are not of interest. Additionally, for appropriate slow loading rates the viscous phenomena of the material can be neglected. Therefore, parameters η_{d} and m_{d} play a minor rule: To solve the system of constitutive equations, the values $\eta_{\mathrm{d}} = 2 \cdot 10^7 \frac{\mathrm{MPa}}{\mathrm{s}}$ and $m_{\mathrm{d}} = 3$ are used. Indeed, the temperature-dependent phase transition stress cannot be identified in an isothermal experiment. Therefore, the temperature-dependence is considered as follows: In the material model the internal variable $\mathbf{X}_\theta$ is responsible for the representation of the temperature-dependent phase transition stress. At $\chi_0 = \{\theta = \theta_0, a = 0, \mathbf{E}_{\mathrm{e}} = \mathbf{0}\}$, the partial derivative

$$\zeta_0 = \left\| \left. \frac{\partial \mathbf{X}_\theta}{\partial \theta} \right|_{\chi_0} \right\| = -\frac{\rho \Delta \eta_0}{\sqrt{\frac{3}{2}} \gamma_{\mathrm{d}}} \qquad \Longleftrightarrow \qquad \Delta \eta_0 = -\frac{\sqrt{\frac{3}{2}} \gamma_{\mathrm{d}} \zeta_0}{\rho} \tag{68}$$

represents in a first approximation the slope of the temperature-dependent phase transition stress. Note, that due to thermodynamical restrictions, $\Delta \eta_0$ and Δu_0 are non-positive material parameters if equal thermoelastic material parameters are regarded (cf. Helm and Haupt (2003)). In appropriate experiments, an average slope of 6.75MPa/K was measured for the investigated alloy (cf. Sect. 2). Consequently, for $\zeta_0 = 6.75$MPa/K the material parameter $\Delta \eta_0$ can be determined for given γ_{d} and ρ. In the present study, the material density is $\rho = 6400\mathrm{kg/m}^3$. The important parameter Δu_0 is not determined directly: Instead of Δu_0, the auxiliary parameter ϑ_0

$$\vartheta_0 = \frac{\Delta u_0}{\Delta \eta_0} \qquad \Longleftrightarrow \qquad \Delta u_0 = \vartheta_0 \, \Delta \eta_0 \tag{69}$$

is identified. Finally, the Poisson number $\nu^{\mathrm{A}} = \nu = 0.36$ is predetermined in a torsion test and the Poisson number of the martensite phase is assumed to be equal to the Poisson number of the austenite phase: $\nu^{\mathrm{M}} = \nu^{\mathrm{A}}$. Therefore, the bulk modulus is given by $\kappa^{\omega} = 2\mu^{\omega}(1+\nu^{\omega})/(3(1-2\nu^{\omega}))$ for $\omega = \mathrm{A}, \mathrm{M}$. For simplicity, only one internal variable $\mathbf{Y}_{\mathrm{d}_i}$ is used in the present framework ($n = 1$ in Eq. (27)). Therefore, nine material parameters (cf. Tab. 1) must be determined.

Number j	Parameter p_j	$p_j^{\min}$	$p_j^{\max}$	Unit
1	μ_{A}	13000	20000	MPa
2	μ_{M}	6000	12000	MPa
3	k	10	50	MPa
4	ϑ_0	255	280	K
5	γ_{d}	0.0425	0.06	–
6	c_1	15000	40000	MPa
7	c_1/b_1	20	90	MPa
8	β_1	5	15	–
9	γ_1	0.7	1	–

Parameter	Value	Unit
$\nu = \nu^{\mathrm{A}} = \nu^{\mathrm{M}}$	0.36	–
η_{d}	$2 \cdot 10^7$	$\frac{\mathrm{MPa}}{\mathrm{s}}$
m_{d}	3	–
ζ_0	6.75	$\frac{\mathrm{MPa}}{\mathrm{K}}$
ρ	6400	$\frac{\mathrm{kg}}{\mathrm{m}^3}$

Table 1: Left table: training range of the material parameters; right table: predetermined material parameters

For the identification process, the experimental data of the step test with the amplitudes $\varepsilon_{\max} = 0.03$ and $\varepsilon_{\max} = 0.06$ are used. Due to the strain rate of $|\dot{\varepsilon}| = 0.0001\mathrm{s}^{-1}$ the loading process is in a first approximation isothermal. The other parts of the step test are applied to verify the identified material parameters. These different strain amplitudes are necessary in order to capture the different stiffness of the material during the phase transitions. The training of the neural network requires data of an appropriate simulation. Concerning this, the initial-value problem is solved for a homogeneously deformed rod under simple tension (cf. the procedure in Helm and Haupt (2003)). In the simulation, the rod is loaded to six percent strain and then unloaded in the strain free state within 1200s. This pre-process leads to an evolution of the internal stress tensor $\mathbf{X}_{\varepsilon_{\mathrm{i}}}$, which strongly influences the stress-strain behavior. Thereafter, the training data are calculated: loading/unloading cycles to a maximum amplitude of $\varepsilon_{\max} = 0.03$ and $\varepsilon_{\max} = 0.06$ with $|\dot{\varepsilon}| = 0.0001\mathrm{s}^{-1}$. The training of the neural network takes place with 400 sets of material parameters, which are randomly distributed in the selected ranges given in Tab. 1. Additionally, 20 sets of material parameters are applied to verify the training process. The used feedforward network consists of 102 input units, 36 hidden units in the first hidden layer as well as 18 hidden units in the second hidden layer, and also nine output units (cf. Fig. 6). The output units representing the nine material parameters. In order to consider most of the observed experimental phenomena in the input data of the neural network, the input units consists of 102 stress values in chronological order. The number of hidden layers and also the number of units in each hidden layer are the result of an iterative process based on experience. In the applied neural network, all units are only connected with units of the next layer (full connection). The logistic function (cf. Eq. (49)) is used as activation function in each unit. In order to improve the efficiency of the network, the input data σ_k and output data p_k are linearly scaled:

$$i_k = -0.25 + \frac{1}{2}\frac{\sigma_k - \sigma_k^{\min}}{\sigma_k^{\max} - \sigma_k^{\min}} \qquad \forall\, k \in \mathcal{I} \qquad \text{and} \qquad o_k = 0.25 + \frac{1}{2}\frac{p_k - p_k^{\min}}{p_k^{\max} - p_k^{\min}} \qquad \forall\, k \in \mathcal{O}. \tag{70}$$

As training algorithm, the pattern-by-pattern back-propagation algorithm with weight decay as described in Sect. 4.2 ($\eta_{\mathrm{L}} = 0.2$ and $\beta = 2 \cdot 10^{-6}$) were applied. In contrast to this strategy, Helm (2004) uses the resilient propagation method (RProp). After 2000 training cycles (epochs), the mean square error (MSE) according to the definition (P: number of training patterns; O: number of output neurons)

$$e_{\mathrm{MSE}} = \frac{E(\mathbf{W})}{PO} = \frac{1}{PO}\sum_{p}\sum_{k\in\mathcal{O}}\frac{1}{2}\left[t_k(p) - o_k(p,\mathbf{W})\right]^2 \tag{71}$$

amounts 0.00599 for the set of training data and 0.00595 for the set of validation data.

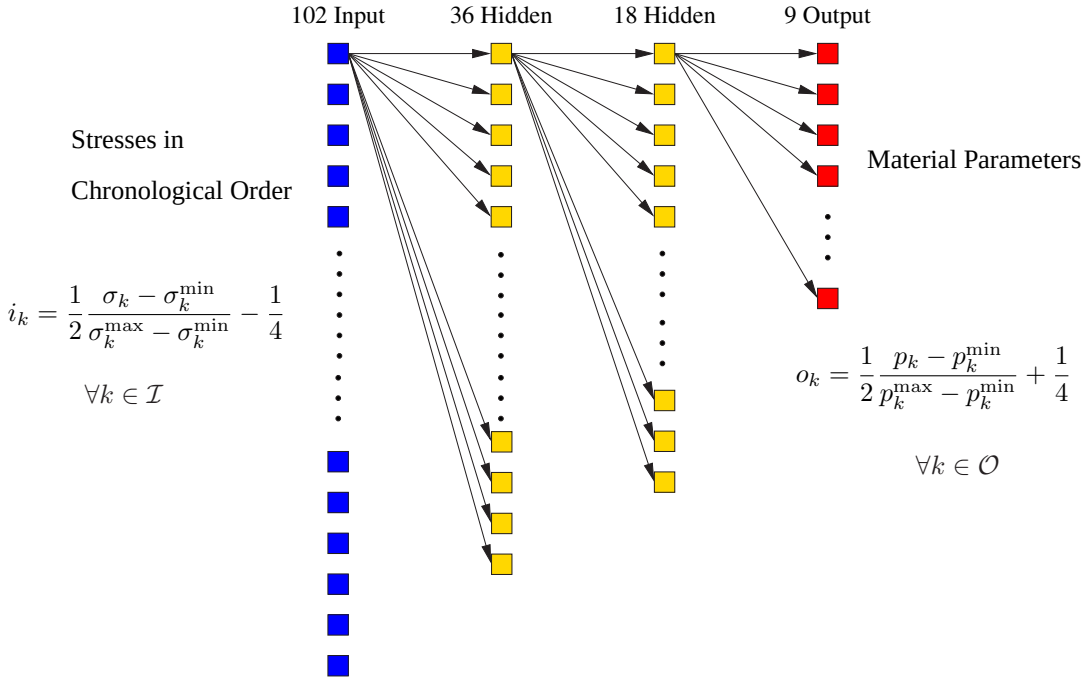

Figure 6: Outline of the applied neural network

Thereafter, the trained neural network is applied to identify the material parameters: 102 measured stress values, depicted in Fig. 7, are applied as the input values for the neural network. The results of the neural network are the identified material parameters, which are summarized in Tab. 2.

μ_A	μ_M	k	Δu_0	$\Delta\eta_0$	γ_d	c_1	b_1	β_1	γ_1
MPa	MPa	MPa	J/kg	J/(kg K)	–	MPa	–	–	–
16377.1	9139.4	13.6	-17877.91	-66.42	0.05142	21790.5	357.5	8.6	0.8471

Table 2: Identified set of material parameters (note: $\Delta\eta_0$ results from Eq. (68) and Δu_0 from Eq. (69))

The calculated model response is also depicted in Fig. 7 (left figure). Concerning the depicted tension behavior, the proposed system of constitutive equations in combination with the set of identified material parameters represents most of the experimentally observed effects: In particular, the model represents the stiffness of the austenite-martensite mixture, the initiation of the phase transition at low stresses, the approximately horizontal phase transition plateaus, the stress increase at a larger amount of martensite, as well as the very small elastic ranges during unloading. The whole step test is likewise depicted in Fig. 7 (right figure): As expected, this simple verification of the model leads to a good agreement between model response and the experimental data. Taking into account the capability of the model, the identification process seems to be successful.

5 Conclusions

The material behavior of shape memory alloys strongly depends on temperature and loading path history. The basic material characteristic can be identified in uniaxial tension tests: e.g. different elastic properties of austenite and martensite phase, temperature-dependent phase transition stress, different loading and unloading slopes in the region of the second elasticity, as well as impressing thermomechanical coupling phenomena. An appropriate thermomechanical material model, which represents these observations, is proposed in the article. The model consists of a free energy function and evolution equations for internal variables. In particular, the introduced internal variable $\mathbf{X}_\theta$ is an important part of the model: This variable moves the phase transition surface in the stress space if the temperature is changed.

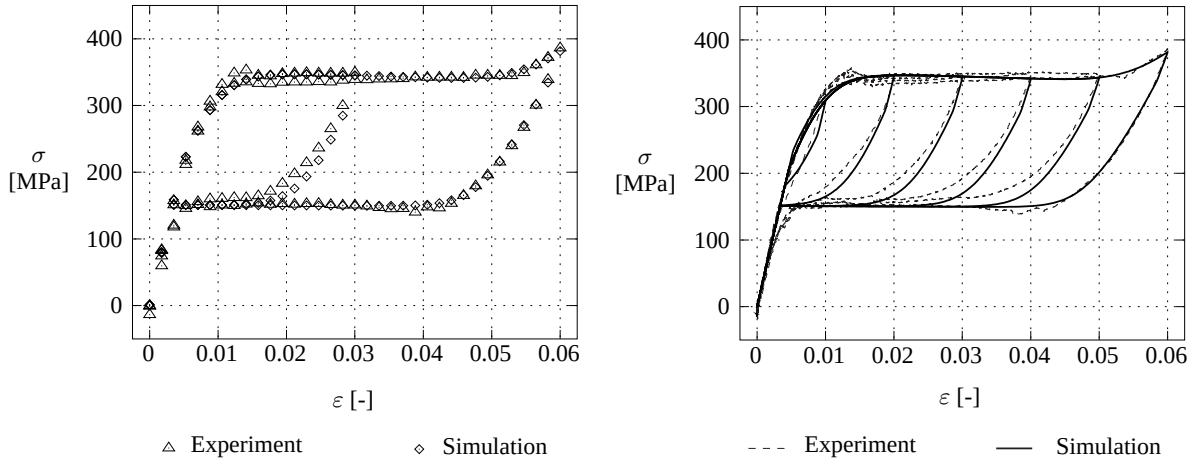

Figure 7: Experiment vs. model response ($|\dot{\varepsilon}| = 0.0001\text{s}^{-1}$): used training process (left figure) and the total step test (right figure)

Due to the complexity of the material properties a constitutive theory contains many material parameters, which are generally functions on temperature and maybe other variables. In spite of the expensive and precise production process, the material properties change in each charge. These circumstances motivate the usage of neural network technology for the identification process, because a trained network is able to identify the temperature-dependent material parameters of different alloys if the applied constitutive theory has the capability to represent the experimental observation. However, the generation of an appropriate neural network needs a lot of experience. In the present context, a fully connected feedforward network is applied whose synaptic weights are determined by an pattern-by-pattern back-propagation algorithm with weight decay. The depicted results show that the neural network is able to identify the material parameters and that the proposed constitutive theory represents the experimentally observed material behavior.

Appendix

A Calculating the weight correction terms in the hidden layers during back-propagation

In Sect. 4.2, the back-propagation algorithm is introduced to determine the synaptic weights. In this context, the partial derivatives $\frac{\partial e(p,\mathbf{W})}{\partial o_n}$ are needed to specify the local gradient δ_n. However, the outputs of the hidden neurons are functions of the neurons in the output layer. Consequently, the chain rule must be applied: For example, the output values o_n with $n \in \mathcal{H}_{L-1}$ in the last hidden layer influences the mean square error due to

$$e(p,\mathbf{W}) = \frac{1}{2}\sum_{k\in\mathcal{O}}\left[t_k(p) - A_k(i_k)\right]^2 = \frac{1}{2}\sum_{k\in\mathcal{O}}\left[t_k(p) - A_k(\sum_{n\in\mathcal{P}_{L-1}} w_{nk}o_n)\right]^2. \tag{72}$$

This results in the derivative

$$\frac{\partial e(p,\mathbf{W})}{\partial o_n} = -\sum_{k\in\mathcal{O}} \underbrace{\left[t_k(p) - o_k(p,\mathbf{W})\right] A_k'(i_k(p))}_{\delta_k} w_{nk} = -\sum_{k\in\mathcal{O}} \delta_k w_{nk} \tag{73}$$

and the local gradient

$$\delta_n = A_n'(i_n(p)) \sum_{k\in\mathcal{O}} \delta_k w_{nk}. \tag{74}$$

Consequently, the local gradients δ_n of neurons in the last hidden layer are influenced by all local gradients δ_k of neurons in the next layer, which is in this case the output layer.

In contrast to this, the neurons o_n in the hidden layer $\mathcal{H}_{L-2}$ have the local gradient

$$\delta_n = A_n'(i_n(p)) \sum_{l \in \mathcal{H}_{L-1}} \delta_l w_{nl}. \tag{75}$$

because of

$$e(p, \mathbf{W}) = \frac{1}{2} \sum_{k \in \mathcal{O}} \left[t_k(p) - A_k(\sum_{l \in \mathcal{P}_{L-1}} w_{lk} A_l(\sum_{n \in \mathcal{P}_{L-2}} w_{nl} o_n)) \right]^2. \tag{76}$$

and

$$\frac{\partial e(p, \mathbf{W})}{\partial o_n} = \underbrace{\sum_{k \in \mathcal{O}} - \left[t_k(p) - o_k(p, \mathbf{W}) \right] A_k'(i_k(p)) \sum_{l \in \mathcal{P}_{L-1}} w_{lk} A_l'(\bar{i}_l) w_{nl}}_{\sum_{l \in \mathcal{P}_{L-1}} \delta_l w_{nl}} = \sum_{l \in \mathcal{H}_{L-1}} \delta_l w_{nl}. \tag{77}$$

This equation has the same interpretation as the result discussed before: The local gradient δ_n of neurons in the hidden layer $\mathcal{H}_{L-2}$ is influenced by all local gradients of connected neurons, which lie in the next layer $\mathcal{H}_{L-1}$.

As a result of this, the local gradients for neurons in the hidden layers

$$\delta_n = A_n'(i_n(p)) \sum_{l \in \mathcal{L}_{m+1}} \delta_l w_{nl} \qquad \forall \quad n \in \mathcal{H}_m = \mathcal{L}_m, \tag{78}$$

are influenced by the local gradients of the next layer multiplied by the synaptic weights.

Acknowledgments

The author gratefully acknowledges the support of this work by German Research Foundation (DFG).

References

Bertram, A.: Thermo–Mechanical Constitutive Equations for the Description of Shape Memory Effects in Alloys. *Nuclear Engineering and Design*, 74, (1982), 173–182.

Bowen, R.: Diffusion Models Implied by the Theory of Mixtures. In: C. Truesdell, ed., *Rational Thermodynamics*, pages 237–263, Springer–Verlag (1969).

Boyd, J.; Lagoudas, D.: *A Constitutive Model for Simultaneous Transformation and Reorientation in Shape Memory Materials*. in L.C. Brinson und B. Moran (Eds.), Mechanics of Phase Transformations and Shape Memory Alloys, The American Society of Mechanical Engineers (1994).

Brinson, L. C.; Schmidt, I.; Lammering, R.: Micro and Macroscopical Investigations of CuAlNi Single Crystal and CuAlMnZn Polycrystalline Shape Memory Alloy. *Journal of Intelligent Material Systems and Structures*, 13, (2002), 751–772.

Brinson, L. C.; Schmidt, I.; Lammering, R.: Stress-induced transformation behavior of a polycrystalline NiTi shape memory alloy: micro and macromechanical investigations via in situ optical microscopy. *Journal of the Mechanics and Physics of Solids*, 52, (2004), 1549–1571.

Falk, F.: One–dimensional model of shape memory. *Archives of Mechanics*, 35, (1983), 63–84.

Frémond, M.: Shape Memory Alloys: A Thermomechanical Macroscopic Therory. In: M. Frémond; S. Miyazaki, eds., *Shape Memory Alloys*, pages 3–68, Springer (1996).

Fu, S.; Müller, I.; Xu, H.: The Interior of the Pseudoelastic Hysteresis. In: C. Liu; H. Kunsmann; K. Otsuka; M. Wuttig, eds., *Shape–Memory Materials and Phenomena – Fundamental Aspects and Applications*, pages 39–42, Materials Research Society Symposium Proceedings Vol. 246, Pittsburg, Pennsylvania (1992).

Funakubo, H.: *Shape Memory Alloys*. Gordon and Breach Science Publishers, New York et al. (1987).

Haupt, P.: *Continuum Mechanics and Theory of Materials*. Springer Verlag, Berlin (2002).

Haykin, S.: *Neural Networks: A Comprehensive Foundation*. Macmillan College Publishing Company (1994).

Helm, D.: *Shape Memory Alloys: Experimental Investigation, Phenomenological Modelling and Numerical Simulation of the Thermomechanical Material Behaviour (in german)*. Ph.D. thesis, Institute of Mechanics, Department of Mechanical Engineering, University of Kassel (2001).

Helm, D.: Pseudoelasticity: experimental observations, thermomechanical modeling, and identication of the material parameters. In: D. Lagoudas, ed., *Proceedings of Smart Structures and Materials 2004: Active Materials: Behavior and Mechanics, Proceedings of SPIE*, vol. 5387, pages 198–209 (2004).

Helm, D.; Haupt, P.: Thermomechanical Behaviour of Shape Memory Alloys. In: S. Lynch, ed., *Proceedings of SPIEs Smart Structures and Materials 2001; Active Materials: Behavior and Mechanics, Proceedings of SPIE*, vol. 4333, pages 302–313 (2001).

Helm, D.; Haupt, P.: Thermomechanical Representation of the Multiaxial Behavior of Shape Memory Alloys. In: S. Lynch, ed., *Proceedings of SPIEs Smart Structures and Materials 1999; Mathematics and Control in Smart Structures*, vol. 4699, pages 343–354, SPIE (2002).

Helm, D.; Haupt, P.: Shape memory behaviour: modelling within continuum thermomechanics. *International Journal of Solids and Structures*, 40, (2003), 825–849.

Hohenemser, K.; Prager, W.: Über die Ansätze der Mechanik isotroper Kontinua. *ZAMM 1932*, 12, (1932), 216–226.

Huber, N.: *Anwendung Neuronaler Netze bei nichtlinearen Problemen der Mechanik*. Habilitationsschrift, Forschungszentrum Karlsruhe (2000).

Huber, N.; Tsakmakis, C.: A neural network tool for identifying the material parameters of a finite deformation viscoplasticity model with static recovery. *Computer Methods in Applied Mechanics and Engineering*, 191, (2001), 353–384.

Huo, Y.; Müller, I.: Nonequilibrium Thermodynamics of Pseudoelasticity. *Continuum Mechanics and Thermodynamics*, 5, (1993), 163–204.

Hutter, K.: The Foundations of Thermodynamics, Its Basic Postulate and Implications. A Review Article. *Acta Mechanica*, 27, (1977), 1–54.

Jou, D.; Casas-Vázquez, J.; Lebon, G.: *Extended Irreversible Thermodynamics*. Springer–Verlag, Berlin Heidelberg New York (1996).

Juhász, L.; Andrä, H.; Hesebeck, O.: Simulation of the thermomechanical behaviour of shape memory alloys under multi–axial non–proportional loading. In: S. Lynch, ed., *Proceedings of SPIEs Smart Structures and Materials 2000; Active Materials: Behavior and Mechanics*, pages 484–495 (2000).

Lemaitre, J.; Chaboche, J.: *Mechanics of Solid Materials*. Cambridge Press, Cambridge, New–York (1990).

Leo, P.; Shield, T.; Bruno, P.: Transient Heat Transfer Effects on the Pseudoelastic Behavior of Shape Memory Wires. *Acta Metallurgica et Materialia*, 41, 8, (1993), 2477–2485.

Levitas, V. I.: Thermomechanical Theory of Martensitic Phase Transformations in Inelastic Materials. *Int. J. of Solids and Structures*, 35, (1998), 889–940.

Li, Z. Q.; Sun, Q. P.: The initiation and growth of macroscopic band in nano-grained NiTi microtube under tension. *International Journal of Plasticity*, 18, (2002), 1481–1498.

Lim, T.; McDowell, D.: Mechanical Behavior of an Ni–Ti Shape Memory Alloy Under Axial–Torsional Proportional and Nonproportional Loading. *Journal of Engineering Materials and Technology*, 121, (1999), 9–19.

Lubliner, J.; Auricchio, F.: Generalized Plasticity and Shape Memory Alloys. *International Journal of Solids and Structures*, 33, (1996), 991–1003.

Miyazaki, S.: Shape Memory Alloys: Development and Characterization of Shape Memory Alloys. In: M. Frémond; S. Miyazaki, eds., *Shape Memory Alloys*, pages 69–147, Springer (1996).

Müller, I.: Stress–Strain–Temperature Curves in Pseudoelastic Bodies. In: D. E. Carlson; R. T. Shield, eds., *Finite Elasticity*, IUTAM Symposium Lehigh University, PA,USA (1982).

Müller, I.; Ruggeri: *Rational Extended Thermodynamics*. Springer–Verlag, Berlin Heidelberg New York (1998).

Otsuka, K.; Wayman, C.: *Shape Memory Materials*. Cambridge University Press (1998).

Perzyna, P.: The Constitutive Equations for Rate Sensitive Plastic Materials. *Quartely of Applied Mathematics*, 20, (1963), 321–332.

Raniecki, B.; Lexcellent, C.: R_L–models of pseudoelasticity and their specification for some shape memory solids. *European Journal of Mechanics A–Solids*, 13, (1994), 21–50.

Raniecki, B.; Lexcellent, C.; Tanaka, K.: Thermodynamic models of pseudoelastic behavior of shape memory alloys. *Archives of Mechanics*, 44, (1992), 261–284.

Raniecki, B.; Tanaka, K.; Ziółkowski, A.: Testing and Modeling of NiTi SMA at Complex Stress State – Selected Results of Polish–Japanese Research Coorperation. *Materials Science Research International, Special Technical Publication*, 2, (2001), 327–334.

Rojas, S.: *Neural Networks: A Systematic Introduction*. Springer-Verlag Berlin Heidelberg New York (1996).

Seelecke, S.: Torsional vibration of a shape memory wire. *Continuum Mechanics and Thermodynamics*, 9, (1997), 165–173.

Shaw, J.: Simulation of localized thermo-mechanical behavior in a NiTi shape memory alloy. *International Journal of Plasticity*, 16, (2000), 541–562.

Shaw, J.; Kyriakides, S.: Thermomechanical Aspects of NiTi. *Journal of the Mechanics and Physics of Solids*, 8, (1995), 1243–1281.

SNNS (2000): *SNNS, Stuttgart Neural Network Simulator, User Manual, Version 4.2*. Institute for Parallel and Distributed High Performance Systems, University of Stuttgart; Wilhelm-Schickard-Institute for Computer Science, University Tübingen (2000).

Souza, A.; Mamiya, E.; Zouain, N.: Three–dimensional model for solids undergoing stress–induced phase transitions. *Eur. J. Mech. A/Solids*, 17, (1998), 789–806.

Tobushi, H.; Shimeno, Y.; Hachisuka, T.; Tanaka, K.: Influence of strain rate on superelastic properties of TiNi shape memory alloy. *Mechanics of Material*, 30, (1998), 141–150.

Zell, A.: *Simulation Neuronaler Netze*. Addison–Wesley (1994).

Address: Dr.-Ing. Dirk Helm, Institute of Mechanics, Department of Mechanical Engineering, University of Kassel, Mönchebergstr. 7, 34109 Kassel, Germany; email: `helm@ifm.maschinenbau.uni-kassel.de`