Gradient Polyconvexity and Modeling of Shape Memory Alloys

Martin Horák, Martin Kružík, Petr Pelech, and Anja Schlömerkemper

1 Introduction

Gradient polyconvex functionals, introduced originally in [10], depend on the gradients of nonlinear minors of the deformation gradient, i.e., they involve not only the first but also the second spatial derivatives of the deformation field. Materials having such a broader energy dependence are generally called non-simple [54] and their idea can be traced back to 1901 when Korteweg [32] considered a gradient of the density in his model of fluid capillarity. Considering more than only the first deformation gradient in the description of elastic behavior of solids goes back to the 1960s and appeared in the work of Toupin [52, 53], and Green and Rivlin [29]. Such materials are usually called N-grade materials, where N refers to the highest deformation gradient appearing in the model. This approach has brought questions on thermodynamical consistency of such models, treated in [13, 22], for instance. Since then, it has been used and analyzed in many works; see, e.g.,

M. Horák
Faculty of Civil Engineering, Czech Technical University, Prague, Czechia
e-mail: martin.horak@fsv.cvut.cz

M. Kružík (✉)
Czech Academy of Sciences, Institute of Information Theory and Automation, Prague, Czechia

Faculty of Civil Engineering, Czech Technical University, Prague, Czechia
e-mail: kruzik@utia.cas.cz

P. Pelech
Weierstrass Institute, Berlin, Germany
e-mail: petr.pelech@wias-berlin.de

A. Schlömerkemper
Institute of Mathematics, University of Würzburg, Würzburg, Germany
e-mail: anja.schloemerkemper@mathematik.uni-wuerzburg.de

P. M. Mariano (ed.), *Variational Views in Mechanics*, Advances in Continuum Mechanics, https://doi.org/10.1007/978-3-030-90051-9_5

[7, 20, 21, 24, 26, 27, 34, 46, 50, 51]. From the material point of view, the more general energy functionals in higher grade continua lead to an additional force interaction in a form of an edge traction or the so-called couple-stress or double force acting on the boundary; see [38, 42, 48, 49].

Mathematically, the presence of higher-order gradients in the model brings additional compactness properties for the set of admissible functions and ensures the existence of minimizers. We refer to recent related results on the mathematical treatment of shape memory materials and solid-to-solid interfaces: [1, 4, 6, 18, 19]. We also refer to [9] for an overview of recent mathematical results in the calculus of variations. For computational results on NiMnGa see, e.g., [1].

The aim of this contribution (cf. [37]) is to apply a new class of non-simple material models introduced in [10] (called *gradient polyconvex materials*) to evolutionary problems of shape memory alloys and to consider a computational experiment. The novelty consists in considering only gradients on nonlinear minors in the stored energy density of the material. It is shown there, and also in Example 2 below, that corresponding deformations do not necessarily have integrable second weak derivatives. Nevertheless, it is possible to prove existence of an energetic solution.

The plan of the paper is as follows. We first introduce necessary notation and tools in Sect. 2. The notion of gradient polyconvexity is thoroughly discussed in Sect. 3 and the quasistatic evolution in Sect. 4. Finally, in Sect. 5 we consider a bar made of a specific shape memory material (NiMnGa) and provide first computational results on the evolution of a solid-to-solid phase transformation in a tension experiment.

2 Preliminaries

Hyperelasticity is a special area of Cauchy elasticity, where one assumes that the first Piola-Kirchhoff stress tensor P possesses a potential (called stored energy density) $W : (0, +\infty) \times \mathbb{R}^{3\times 3} \to (-\infty, \infty]$. In other words,

$$P(\theta, F) := \frac{\partial W(\theta, F)}{\partial F} \tag{1}$$

on its domain, where $F \in \mathbb{R}^{3\times 3}$ is such that $\det F > 0$ and θ stands for the absolute temperature. This concept emphasizes that all work done by external loads on the specimen is stored in it. The principle of frame-indifference requires that W satisfies, for all $F \in \mathbb{R}^{3\times 3}$ and all proper rotations $R \in \mathrm{SO}(3)$,

$$W(\theta, F) = W(\theta, RF) = \tilde{W}(\theta, F^\top F) = \tilde{W}(\theta, C),$$

where $C := F^\top F$ is the right Cauchy-Green strain tensor and $\tilde{W} : (0, +\infty) \times \mathbb{R}^{3\times 3} \to (-\infty, \infty]$.

Additionally, every elastic material is assumed to resist extreme compression, which is modeled by

$$W(\theta, F) \to +\infty, \text{ if } \det F \searrow 0. \tag{2}$$

Let the reference configuration be a bounded Lipshitz domain $\Omega \subset \mathbb{R}^3$. Deformation $y : \bar{\Omega} \to \mathbb{R}^3$ maps the points in the closure of the reference configuration $\bar{\Omega}$ to their positions in the deformation configuration. Solutions to the corresponding elasticity equations can then be formally found by minimizing the energy functional

$$I(\theta, y) := \int_\Omega W(\theta, \nabla y(x))\, \mathrm{d}x - \ell(y) \tag{3}$$

over the class of admissible deformations. Here, ℓ is a functional on the set of deformations, expressing (in a simplified way) the work of external loads on the specimen, and ∇y is the deformation gradient, which quantifies the strain. We only allow for deformations, which are orientation-preserving, i.e., if $a, b, c \in \mathbb{R}^3$ satisfy $(a \times b) \cdot c > 0$, then $(Fa \times Fb) \cdot Fc > 0$ for every $F := \nabla y(x)$ and $x \in \Omega$, which means that $\det F > 0$. This condition can be expressed by extending W by infinity on matrices with non-positive determinants, i.e.,

$$W(\theta, F) := +\infty, \text{ if } \det F \leq 0. \tag{4}$$

In view of (1), (2), and (4), we see that $W : (0, +\infty) \times \mathbb{R}^{3\times 3} \to (-\infty, +\infty]$, is continuous in the sense that if $F_k \to F$ in $\mathbb{R}^{3\times 3}$ for $k \to +\infty$, then $\lim_{k\to+\infty} W(\theta, F_k) = W(\theta, F)$. Furthermore, $W(\theta, \cdot)$ is differentiable on the set of matrices with positive determinants.

Relying on the direct method of the calculus of variations, the usual approach to prove the existence of minimizers is to study (weak) lower semicontinuity of the functional I on appropriate Banach spaces containing the admissible deformations. For definiteness, we assume that $y \mapsto -\ell(y)$ is weakly sequentially lower semicontinuous. Thus, the question reduces to a discussion of the assumptions on W. It is well-known that (2) prevents us from assuming convexity of W. See, e.g., [17] or the recent review for a detailed exposition of weak lower semicontinuity. Following earlier work by C.B. Morrey, Jr., [43], J.M. Ball [2] defined a polyconvex stored energy density W by assuming that there is a convex and lower semicontinuous function $\overline{W}(\theta, \cdot) : \mathbb{R}^{19} \to (-\infty, +\infty]$ such that

$$W(\theta, F) := \overline{W}(\theta, F, \operatorname{Cof} F, \det F) \quad \forall F \in \mathbb{R}^{3\times 3}.$$

Here, $\operatorname{Cof} F$ denotes the cofactor matrix of F, which, for F being invertible, satisfies Cramer's rule:

$$\operatorname{Cof} F = (\det F)(F^{-1})^\top.$$

Hence, $\det \operatorname{Cof} F = \det^2 F$ and because we assume that $\det F > 0$ we have that

$$F = \Big(\frac{\operatorname{Cof} F}{\sqrt{\det \operatorname{Cof} F}}\Big)^{-\top},$$

i.e., we can reconstruct F from $\operatorname{Cof} F$. It is well-known that polyconvexity is satisfied for a large class of constitutive functions and allows for the existence of minimizers of I under (2) and (4). On the other hand, there are still situations where polyconvexity cannot be adopted. A prominent example is shape memory alloys, where W has the so-called multi-well structure; see, e.g., [5, 11, 44]. Namely, there is a high-temperature phase, called austenite, which is usually of cubic symmetry, and a low-temperature phase, called martensite, which is less symmetric and exists in more variants, e.g., in three for the tetragonal structure (NiMnGa) or in twelve for the monoclinic one (NiTi). We can assume that

$$W(\theta, F) := \min_{0 \le i \le M} W_i(\theta, F), \tag{5}$$

where $W_i : (0, +\infty) \times \mathbb{R}^{3\times 3} \to (-\infty, +\infty]$ is the stored energy density of the i-th variant of martensite if $i > 0$, and W_0 is the stored energy density of the austenite. For every admissible i, we have $W_i(\theta, \cdot)$ is minimized if and only if $F = RF_i$ for a given matrix $F_i \in \mathbb{R}^{3\times 3}$ and an arbitrary proper rotation $R \in \mathrm{SO}(3)$. This means that each variant of the martensite and the austenite is modeled as a hyperelastic material with its own stored energy density W_i. We also assume that each $W_i(\theta, \cdot)$ is differentiable on the set of matrices with positive determinants. Thus the variants can be described independently of each other, i.e., the elastic constants can be chosen differently. The drawback is obviously the non-smoothness of W, however, physically realistic elastic strain values do not occur in the set where W is not differentiable. We refer, e.g., to [39] for other models of the stored energy density of shape memory alloys.

Given a deformation gradient F, we need to decide if the corresponding deformation is in the well of the austenite, or in a martensitic variant. In order to do so, we define a volume fraction $\lambda(F)$ as follows: Let $\lambda : \mathbb{R}^{3\times 3} \to \mathbb{R}^{M+1}$. Set

$$\lambda^j(F) := \frac{1}{M}\left(1 - \frac{\operatorname{dist}(C, \mathcal{N}_j(C_j))}{\sum_{i=0}^{M} \operatorname{dist}(C, \mathcal{N}_i(C_i))}\right) \; \forall C = F^T F \in \mathbb{R}^{3\times 3},\; j = 0, \ldots, M, \tag{6}$$

where $\{\mathcal{N}_i(C_i)\}_i$ are pairwise disjoint neighborhoods of the right Cauchy-Green strain tensors $C_i = F_i^\top F_i$, for $i \in \{0, \ldots, M\}$. Notice that $\sum_{j=0}^{M} \lambda^j(F) = 1$ for every F, which, together with $\lambda^j \ge 0$, allows us to interpret λ as a volume fraction.

Moreover, note that λ is continuous and frame-indifferent in the sense that $\lambda(F) = \lambda(RF)$ for every proper rotation R. Volume fractions will play an important role in the definition of our evolutionary model in Sect. 4.

Remark 1 Note that this particular choice of λ allows for some elastic behavior close to the wells $SO(3)F_i$, $i = 0, \dots, M$, since the volume fraction remains constant on the neighborhoods $\mathcal{N}_i(C_i)$, $i = 0, \dots, M$.

Let us emphasize that (5) ruins even generalized notions of convexity as, e.g., rank-one convexity. (We recall that rank-one convex functions are convex on line segments with endpoints differing by a rank-one matrix and that rank-one convexity is a necessary condition for polyconvexity; cf. [17], for instance.) Namely, it is observed (see, e.g., [5, 11]) that there is a proper rotation R_{ij} such that $\operatorname{rank}(R_{ij}F_i - F_j) = 1$. if $0 < i \neq j > 0$. Hence, generically, $W(\theta, R_{ij}F_i) = W(\theta, F_j) = -w_i(\theta)$, but $W(\theta, F) > -w_i(\theta)$ if F is on the line segment between $R_{ij}F_i$ and F_j. Nevertheless, not having a convexity property at hand that implied existence of minimizers is in accordance with experimental observations for these alloys.

Indeed, nonexistence of a minimizer corresponds to the formation of microstructure of strain-states. This is mathematically manifested via a faster and faster oscillation of deformation gradients in minimizing sequences, driving the functional I to its infimum. One can then formulate a minimization problem for a lower semicontinuous envelope of I, the so-called relaxation, see, e.g., [17]. Such a relaxation yields information of the effective behavior of the material and on the set of possible microstructures. Thus relaxation is not only an important tool for mathematical analysis, but also for applications. For numerical considerations it is a challenging problem, because the relaxation formula is generically not obtained in a closed form. Further difficulties come from the fact that a sound mathematical relaxation theory is developed only if W has p-growth; that is, for some $c(\theta)$, $c > 1$, $p \in]1, +\infty[$ and all $F \in \mathbb{R}^{3\times 3}$, the inequality

$$\frac{1}{c}(|F| - c(\theta)) \leq W(\theta, F) \leq c(1 + |F|^p + c(\theta))$$

is satisfied. This in particular implies that $W < +\infty$. We refer, however, to [8, 16, 33] for results allowing for infinite energies. Nevertheless, these works include other assumptions that severely restrict their usage. Let us point out that the right Cauchy-Green strain tensor $F^\top F$ maps $\mathrm{SO}(3)F$ as well as $(\mathrm{O}(3)\setminus\mathrm{SO}(3))F$ to the same point. Here, O(3) are the orthogonal matrices with determinant ± 1. Thus, for example, $F \mapsto |F^\top F - \mathbb{I}|$ is minimized on two energy wells, on SO(3) and also on $\mathrm{O}(3)\setminus\mathrm{SO}(3)$. However, the latter set is not acceptable in elasticity, because the corresponding minimizing affine deformation is a mirror reflection. In order to distinguish between these two wells, it is necessary to incorporate $\det F$ in the model properly.

Besides relaxation, another approach guaranteeing existence of minimizers is to resort to non-simple materials, i.e., materials, whose stored energy density depends also on higher-order derivatives. Simple examples are functionals of the form

$$I(\theta, y) := \int_\Omega W(\theta, \nabla y(x)) + \varepsilon|\nabla^2 y(x)|^p \,\mathrm{d}x - \ell(y),$$

where $\varepsilon > 0$. Obviously, the second-gradient term brings additional compactness to the problem, which allows to require only strong lower semicontinuity of the term

$$\nabla y \mapsto \int_\Omega W(\theta, \nabla y(x))\,\mathrm{d}x$$

for existence of minimizers.

Here, we follow a different approach suggested in [10], which is a natural extension of polyconvexity exploiting weak continuity of minors in Sobolev spaces. Instead of the full second gradient, it is assumed that the stored energy density of the material depends on the deformation gradient ∇y and on gradients of nonlinear minors of ∇y, i.e., on $\nabla[\mathrm{Cof}\,\nabla y]$ and on $\nabla[\det \nabla y]$. The corresponding functionals are then called gradient polyconvex. While we assume convexity of the stored energy density in the two latter variables, this is not assumed in the ∇y variable. The advantage is that minimizers are elements of Sobolev spaces $W^{1,p}(\Omega, \mathbb{R}^3)$, and no higher regularity is required.

The following example is inspired by a similar one in [10]. It shows that there are maps with smooth nonlinear minors whose deformation gradient is *not* a Sobolev map. Hence, gradient polyconvex energies are more general than second-gradient ones.

Example

Let $\Omega = \,]0, 1[^3$. For functions $f, g :]0, 1[\to]0, +\infty[$ to be specified later, let us consider the deformation

$$y(x_1, x_2, x_3) := (x_1, x_2 f(x_1), x_3 g(x_1))\,.$$

Then,

$$\nabla y(x_1, x_2, x_3) = \begin{pmatrix} 1 & 0 & 0 \\ x_2 f'(x_1) & f(x_1) & 0 \\ x_3 g'(x_1) & 0 & g(x_1) \end{pmatrix},$$

$$\mathrm{Cof}\,\nabla y(x_1, x_2, x_3) = \begin{pmatrix} f(x_1)g(x_1) & -x_2 f'(x_1)g(x_1) & -x_3 f(x_1)g'(x_1) \\ 0 & g(x_1) & 0 \\ 0 & 0 & f(x_1) \end{pmatrix}$$

and

$$\det \nabla y(x_1, x_2, x_3) = f(x_1)g(x_1) > 0\,.$$

Finally, the nonzero entries of $\nabla^2 y(x_1, x_2, x_3)$ are

$$x_2 f''(x_1), \quad f'(x_1), \quad x_3 g''(x_1), \quad g'(x_1). \tag{7}$$

Note that we have in particular

$$|\nabla^2 y(x_1, x_2, x_3)| \geq |x_2||f''(x_1)|.$$

Any functions f, g such that $y \in W^{1,p}(\Omega; \mathbb{R}^3)$, $\operatorname{Cof} \nabla y \in W^{1,q}(\Omega; \mathbb{R}^{3\times 3})$, $0 < \det \nabla y \in W^{1,r}(\Omega)$, $(\det \nabla y)^{-s} \in L^1(\Omega)$ for some $p, q, r \geq 1$ and $s > 0$, but such that one of the quantities in (7) is not a function in $L^p(\Omega)$ yield a useful example since then $y \notin W^{2,p}(\Omega; \mathbb{R}^3)$. To be specific, we choose, for $1 > \varepsilon > 0$,

$$f(x_1) = x_1^{1-\varepsilon} \quad \text{and} \quad g(x_1) = x_1^{1+\varepsilon}.$$

Hence

$$f'(x_1) = (1-\varepsilon)x_1^{-\varepsilon}, \qquad g'(x_1) = (1+\varepsilon)x_1^{\varepsilon},$$
$$f''(x_1) = -\varepsilon(1-\varepsilon)x_1^{-1-\varepsilon}, \qquad g''(x_1) = \varepsilon(1+\varepsilon)x_1^{-1+\varepsilon}.$$

Since $x_2 f''(x_1)$ is not integrable, we have $\nabla^2 y \notin L^1(\Omega; \mathbb{R}^{3\times 3\times 3})$ and thus $y \notin W^{2,1}(\Omega; \mathbb{R}^3)$. We have only $y \in W^{1,p}(\Omega; \mathbb{R}^3) \cap L^\infty(\Omega; \mathbb{R}^3)$ for every $1 \leq p < 1/\varepsilon$. Moreover, direct computation shows that both $\operatorname{Cof} \nabla y$ and $\det \nabla y$ lie in $W^{1,\infty}$. Finally, $\det \nabla y = x_1^2 > 0$ and $(\det \nabla y)^{-s} \in L^1(\Omega)$ for all $0 < s < 1/2$.

Therefore, for any $r, q \geq 1, s > 0$, requiring a deformation $y : \Omega \to \mathbb{R}^3$ to satisfy $\det \nabla y \in W^{1,r}(\Omega)$, $(\det \nabla y)^{-s} \in L^1(\Omega)$ and $\operatorname{Cof} \nabla y \in W^{1,q}(\Omega; \mathbb{R}^{3\times 3})$ is a weaker assumption than $y \in W^{2,1}(\Omega; \mathbb{R}^3)$.

3 Gradient Polyconvexity

We start with a definition of gradient polyconvexity.

Definition 1 (See [10, 36]) Let $\hat{W} : (0, +\infty) \times \mathbb{R}^{3\times 3} \times \mathbb{R}^{3\times 3\times 3} \times \mathbb{R}^3 \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function, and let $\Omega \subset \mathbb{R}^3$ be a bounded open domain. The functional

$$J(\theta, y) = \int_\Omega \hat{W}(\theta, \nabla y(x), \nabla[\operatorname{Cof} \nabla y(x)], \nabla[\det \nabla y(x)])\mathrm{d}x, \tag{8}$$

defined for any measurable function $y : \Omega \to \mathbb{R}^3$ for which the weak derivatives ∇y, $\nabla[\operatorname{Cof} \nabla y]$, $\nabla[\det \nabla y]$ exist and which are integrable, is called *gradient polyconvex* if the function $\hat{W}(F, \cdot, \cdot)$ is convex for every $F \in \mathbb{R}^{3\times 3}$.

With J defined as in (8) and a functional $y \mapsto -\ell(y)$ expressing the work of external loads, we set

$$I(\theta, y) := J(\theta, y) - \ell(y). \tag{9}$$

Besides convexity properties, the results of weak lower semicontinuity of $I(\theta, \cdot)$ on $W^{1,p}(\Omega; \mathbb{R}^3)$, in the case $1 \le p < +\infty$, rely on suitable coercivity properties. Here we assume that there are numbers $q, r > 1$ and $c, c(\theta), s > 0$ such that for every $F \in \mathbb{R}^{3\times 3}$, $\Delta_1 \in \mathbb{R}^{3\times 3\times 3}$, and every $\Delta_2 \in \mathbb{R}^3$

$$\hat{W}(\theta, F, \Delta_1, \Delta_2) \\ \ge \begin{cases} c\big(|F|^p + |\operatorname{Cof} F|^q + (\det F)^r + (\det F)^{-s} + |\Delta_1|^q + |\Delta_2|^r\big) - c(\theta), & \text{if } \det F > 0, \\ +\infty, & \text{otherwise.} \end{cases} \tag{10}$$

The following existence result is taken from [10] where it is stated without the explicit dependance on θ. For the reader's convenience, we provide a proof below.

Proposition 1 *Let $\theta > 0$ be fixed. Let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain, and let $\Gamma = \Gamma_0 \cup \Gamma_1$ be an $\mathcal{H}^2$-measurable partition of $\Gamma = \partial\Omega$ with the area of $\Gamma_0 > 0$. Let further $-\ell : W^{1,p}(\Omega; \mathbb{R}^3) \to \mathbb{R}$ be a weakly lower semicontinuous functional satisfying, for some $\tilde{C} > 0$ and $1 \le \bar{p} < p$,*

$$\ell(y) \le \tilde{C} \|y\|^{\bar{p}}_{W^{1,p}(\Omega;\mathbb{R}^3)}, \qquad \text{for all } y \in W^{1,p}(\Omega; \mathbb{R}^3). \tag{11}$$

Further, let J, as in (8), be gradient polyconvex on Ω and such that there is a $\hat{W}$ as in Definition 1 which in addition satisfies (10) for $p > 2$, $q \ge \frac{p}{p-1}$, $r > 1$, $s > 0$. Moreover, assume that, for some given measurable function $y_0 : \Gamma_0 \to \mathbb{R}^3$, the following set

$$\mathcal{A} := \big\{ y \in W^{1,p}(\Omega; \mathbb{R}^3) : \ \operatorname{Cof} \nabla y \in W^{1,q}(\Omega; \mathbb{R}^{3\times 3}), \ \det \nabla y \in W^{1,r}(\Omega), \\ (\det \nabla y)^{-s} \in L^1(\Omega), \ \det \nabla y > 0 \text{ a.e. in } \Omega, \ y = y_0 \text{ on } \Gamma_0 \big\}$$

is nonempty. If $\inf_{\mathcal{A}} I(\theta, \cdot) < \infty$ for I from (9), then the functional I has a minimizer on $\mathcal{A}$.

Proof Our proof closely follows the approach in [10]. Let $\{y_k\} \subset \mathcal{A}$ be a minimizing sequence of I. Due to coercivity assumption (10), the bound on the loading (11), the Poincaré inequality, and the Dirichlet boundary conditions on Γ_0, we obtain that

$$\sup_{k\in\mathbb{N}} \Big(\|y_k\|_{W^{1,p}(\Omega;\mathbb{R}^3)} + \|\operatorname{Cof}\nabla y_k\|_{W^{1,q}(\Omega;\mathbb{R}^{3\times 3})} + \|\det\nabla y_k\|_{W^{1,r}(\Omega)} + \|(\det\nabla y_k)^{-s}\|_{L^1(\Omega)} \Big) < \infty. \tag{12}$$

Hence, by standard results on weak convergence of minors, see, e.g., [14, Thm. 7.6-1], there are (not explicitly labeled) subsequences such that

$$y_k \rightharpoonup y \text{ in } W^{1,p}(\Omega;\mathbb{R}^3), \quad \operatorname{Cof}\nabla y_k \rightharpoonup \operatorname{Cof}\nabla y \text{ in } L^q(\Omega;\mathbb{R}^{3\times 3}),$$
$$\det\nabla y_k \rightharpoonup \det\nabla y \text{ in } L^r(\Omega)$$

for $k \to \infty$. Moreover, since bounded sets in uniformly convex Sobolev spaces are weakly sequentially compact,

$$\operatorname{Cof}\nabla y_k \rightharpoonup H \text{ in } W^{1,q}(\Omega;\mathbb{R}^{3\times 3}), \quad \det\nabla y_k \rightharpoonup D \text{ in } W^{1,r}(\Omega) \tag{13}$$

for some $H \in W^{1,q}(\Omega;\mathbb{R}^{3\times 3})$ and $D \in W^{1,r}(\Omega)$. Since the weak limit is unique, we have $H = \operatorname{Cof}\nabla y$ and $D = \det\nabla y$. By compact embedding, also $\operatorname{Cof}\nabla y_k \to H$ in $L^q(\Omega;\mathbb{R}^{3\times 3})$ and hence we obtain a (not explicitly labeled) subsequence such that, for $k \to \infty$,

$$\operatorname{Cof}\nabla y_k \to \operatorname{Cof}\nabla y \quad \text{a.e. in } \Omega. \tag{14}$$

Since, by Cramer's formula, $\det(\operatorname{Cof}\nabla y) = (\det\nabla y)^2$, we have, for $k \to \infty$, that

$$\det\nabla y_k \to \det\nabla y \quad \text{a.e. in } \Omega. \tag{15}$$

Next we show that y belongs to the set of admissible functions $\mathcal{A}$. Notice that $\det\nabla y \geq 0$ since $\det\nabla y_k > 0$ for any $k \in \mathbb{N}$. Further, the conditions (10), (11), (12), and the Fatou lemma imply that

$$+\infty > \liminf_{k\to\infty} I(\theta, y_k) + \ell(y_k) \geq \liminf_{k\to\infty} \int_\Omega \frac{1}{(\det\nabla y_k(x))^s}\,\mathrm{d}x$$
$$\geq \int_\Omega \frac{1}{(\det\nabla y(x))^s}\,\mathrm{d}x.$$

Hence, inevitably, $\det\nabla y > 0$ almost everywhere in Ω and $(\det\nabla y)^{-s} \in L^1(\Omega)$. Since the trace operator is continuous, we obtain that $y \in \mathcal{A}$.

By Cramer's rule, the inverse of the deformation gradient satisfies, for almost all $x \in \Omega$ and $k \to \infty$, that

$$(\nabla y_k(x))^{-1} = \frac{(\operatorname{Cof} \nabla y_k(x))^\top}{\det \nabla y_k(x)} \longrightarrow \frac{(\operatorname{Cof} \nabla y(x))^\top}{\det \nabla y(x)} = (\nabla y(x))^{-1}. \tag{16}$$

Notice that, for almost all $x \in \Omega$,

$$\begin{aligned}\sup_{k\in\mathbb{N}} |\nabla y_k(x)| &= \sup_{k\in\mathbb{N}} \det \nabla y_k(x)\ |((\operatorname{Cof}(\nabla y_k(x)))^{-1}))^\top| \\ &\le \sup_{k\in\mathbb{N}} \frac{3}{2} \det \nabla y_k(x)\ |(\nabla y_k(x))^{-1}|^2 < \infty\end{aligned}$$

because of the pointwise convergence of $\{\det \nabla y_k\}$ and (16).

Due to (16), we have, for almost all $x \in \Omega$ and $k \to \infty$, that

$$\begin{aligned}\nabla y_k(x) &= ((\operatorname{Cof}(\nabla y_k(x))^{-1})^\top \det \nabla y_k(x) \longrightarrow ((\operatorname{Cof}(\nabla y(x))^{-1})^\top \det \nabla y(x) \\ &= \nabla y(x),\end{aligned}$$

where we have used that the cofactor of some matrix is invertible whenever the matrix itself is invertible too. As the Lebesgue measure on Ω is finite, we get by the Egoroff theorem, c.f. [23, Thm. 2.22],

$$\nabla y_k \to \nabla y \ \text{ in measure.} \tag{17}$$

Since $\hat{W}(\theta, \cdot)$ is bounded from below and continuous on matrices with positive determinants and $\hat{W}(\theta, F, \cdot, \cdot)$ is convex, we may use [23, Cor. 7.9] to conclude, from (17) and (13), that

$$\begin{aligned}&\int_\Omega \hat{W}(\theta, \nabla y(x), \nabla \operatorname{Cof} \nabla y(x), \nabla \det \nabla y(x))\, \mathrm{d}x \\ &\le \liminf_{k\to\infty} \int_\Omega \hat{W}(\theta, \nabla y_k(x), \nabla \operatorname{Cof} \nabla y_k(x), \nabla \det \nabla y_k(x))\, \mathrm{d}x\ .\end{aligned}$$

To pass to the limit in the functional $-\ell$, we exploit its weak lower semicontinuity. Therefore, the whole functional I is weakly lower semicontinuous along $\{y_k\} \subset \mathcal{A}$ and hence $y \in \mathcal{A}$ is a minimizer of $I(\theta, \cdot)$.

Remark 2 Note that the pointwise convergence (15) of the determinant, necessary for obtaining the crucial convergence in (17), was not achieved by compact embedding, as it was done for $\operatorname{Cof} \nabla y$ in (14). Hence, the coercivity in $\nabla[\det \nabla y]$ is of minor importance and can be relaxed, provided the function $\hat{W}$ from (8) does not depend on its last argument, c.f. [10, Prop. 5.1]. On the other hand, although only $\nabla[\operatorname{Cof} \nabla y]$ is necessary for regularizing the whole problem, making the functional in (8) dependent also on $\nabla[\det \nabla y]$ may be interesting from the applications point of view.

Let $\mathcal{L}^3$ denote the Lebesgue measure in $\mathbb{R}^3$. If $p > 3$ and $y \in W^{1,p}(\Omega;\mathbb{R}^3)$ is such that $\det \nabla y > 0$ almost everywhere in Ω, then the so-called Ciarlet-Nečas condition

$$\int_\Omega \det \nabla y(x)\,\mathrm{d}x \le \mathcal{L}^3(y(\Omega)), \tag{18}$$

derived in [15], ensures almost everywhere injectivity of deformations. We also refer to [28, Sec. 6, Thm.2] and to [3] for other conditions ensuring injectivity of deformations, requiring, however, a prescribed Dirichlet boundary datum on the whole $\partial\Omega$, which is difficult to ensure in a physical lab. If

$$\frac{|\nabla y|^3}{\det \nabla y} \in L^\delta(\Omega) \tag{19}$$

for some $\delta > 2$ and (18) holds, then we even get invertibility everywhere in Ω due to [30, Theorem 3.4]. Namely, this then implies that y is an open map. Hence, we get the following corollary of Proposition 1.

Corollary 1 *Let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain, and let $\Gamma = \Gamma_0 \cup \Gamma_1$ be an $\mathcal{H}^2$-measurable partition of $\Gamma = \partial\Omega$ with the area of $\Gamma_0 > 0$. Let further $\ell : W^{1,p}(\Omega;\mathbb{R}^3) \to \mathbb{R}$ be a weakly upper semicontinuous functional and J as in (8) be gradient polyconvex on Ω such that $\hat{W}$ satisfies (10). Finally, let $p > 6$, $q \ge \frac{p}{p-1}$, $r > 1$, $s > 2p/(p-6)$, and assume that, for some given measurable function $y_0 : \Gamma_0 \to \mathbb{R}^3$, the following set*

$$\begin{aligned}\mathcal{A} := \{&y \in W^{1,p}(\Omega;\mathbb{R}^3) : \ \mathrm{Cof}\,\nabla y \in W^{1,q}(\Omega;\mathbb{R}^{3\times 3}),\ \det \nabla y \in W^{1,r}(\Omega),\\ &(\det \nabla y)^{-s} \in L^1(\Omega),\ \det \nabla y > 0 \textit{ a.e. in } \Omega,\ y = y_0 \textit{ on } \Gamma_0,\ (18) \textit{ holds}\}\end{aligned}$$

is nonempty. If $\inf_{\mathcal{A}} I < \infty$ for I from (9), then the functional I has a minimizer on $\mathcal{A}$ which is injective everywhere in Ω.

A simple example of an energy density which satisfies the assumptions of Proposition 1 and Corollary 1 is

$$\begin{aligned}&\hat{W}(\theta, F, \Delta_1, \Delta_2)\\ &= \begin{cases} W(\theta, F) + \varepsilon\big(|F|^p + |\mathrm{Cof}\,F|^q + (\det F)^r + (\det F)^{-s} + |\Delta_1|^q + |\Delta_2|^r\big), & \text{if } \det F > 0,\\ +\infty, & \text{otherwise,}\end{cases}\end{aligned}$$

for W defined in (5).

Remark 3 (Gradient Polyconvex Materials and Smoothness of Stress) Gradient polyconvex materials enable us to control regularity of the first Piola-Kirchhoff stress tensor by means of smoothness of the Cauchy stress. Assume that the Cauchy stress tensor $T^y : y(\Omega) \to \mathbb{R}^{3\times 3}$ is Lipschitz continuous, for instance. If $\operatorname{Cof} \nabla y : \Omega \to \mathbb{R}^{3\times 3}$ is Lipschitz continuous too, then the first Piola-Kirchhoff stress tensor P inherits the Lipschitz continuity from T^y because

$$P(x) := T^y(x^y) \operatorname{Cof} \nabla y(x),$$

where $x^y := y(x)$. In a similar fashion, one can transfer Hölder continuity of T^y to P via Hölder continuity of $x \mapsto \operatorname{Cof} \nabla y(x)$.

4 Evolution

If the loading changes in time or if the boundary condition becomes time-dependent, then the specimen evolves as well. We consider here the case, in which evolution is connected with energy dissipation. Experimental evidence shows that considering a rate-independent dissipation mechanism is a reasonable approximation in a wide range of rates of external loads. We hence need to define a suitable dissipation function.

Since we consider a rate-independent processes, this dissipation will be positively one-homogeneous. We associate the dissipation with the magnitude of the time derivative of the dissipative variable $z \in \mathbb{R}^{M+1}$, where $M \in \mathbb{N}$, i.e., with $|\dot{z}|_{M+1}$, where $|\cdot|_{M+1}$ denotes a norm on $\mathbb{R}^{M+1}$ (in our setting, the internal variable z can be seen as a vector of volume fractions of austenite and M variants of martensite). Therefore, the specific dissipated energy associated with a change from state z^1 to z^2 is postulated as

$$D(z^1, z^2) := |z^1 - z^2|_{M+1}. \tag{20}$$

Hence, for $z^i : \Omega \to \mathbb{R}^{M+1}$, $i = 1, 2$, the total dissipation reads

$$\mathcal{D}(z^1, z^2) := \int_\Omega D(z^1(x), z^2(x))\, dx,$$

and the total $\mathcal{D}$-dissipation of a time-dependent curve $z : t \in [0, T] \mapsto z(t)$, where $z(t) : \Omega \to \mathbb{R}^{M+1}$ is defined as

$$\mathrm{Diss}_{\mathcal{D}}(z, [s, t]) := \sup \left\{ \sum_{j=1}^{N} \mathcal{D}(z(t_{i-1}), z(t_i)) : N \in \mathbb{N}, s = t_0 \le \ldots \le t_N = t \right\}.$$

Let $\mathcal{Z}$ denote the set of all admissible states of internal variables $z : \Omega \to \mathbb{R}^{M+1}$ and $\mathcal{A}$ be the set of admissible deformations as before. For a given triple $(t, y, z) \in [0, T] \times \mathcal{A} \times \mathcal{Z}$, we define the total energy of the system by

$$\mathcal{E}(t, \theta, y, z) = \begin{cases} J(\theta, y) - L(t, y), & \text{if } z = \lambda(\nabla y) \text{ a.e. in } \Omega, \\ +\infty, & \text{otherwise,} \end{cases}$$

where $L(t, \cdot)$ is a functional on deformations expressing time-dependent loading of the specimen, and λ is defined in (6).

4.1 Energetic Solution

Suppose, that we look for the time evolution of $t \mapsto y(t) \in \mathcal{A}$ and $t \mapsto z(t) \in \mathcal{Z} := L^\infty(\Omega, \mathbb{R}^{M+1})$ during a process on a time interval $[0, T]$, where $T > 0$ is the time horizon. We use the following notion of solution from [25], see also [40, 41].

Definition 2 (Energetic Solution) Let an energy $\mathcal{E} : [0, T] \times (0, +\infty) \times \mathcal{A} \times \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$ and a dissipation distance $\mathcal{D} : \mathcal{Z} \times \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$ be given. The set of admissible configurations is defined as

$$\mathcal{Q} := \{(y, z) \in \mathcal{A} \times \mathcal{Z} : \lambda(\nabla y) = z \text{ a.e. in } \Omega\}.$$

We say that $(y, z) : [0, T] \to \mathcal{Q}$ is an energetic solution to $(\mathcal{Q}, \mathcal{E}, \mathcal{D})$, if the mapping $t \mapsto \partial_t \mathcal{E}(t, \theta, y(t), z(t))$ is in $L^1(0, T)$ and if, for all $t \in [0, T]$, the stability condition

$$\mathcal{E}(t, \theta, y(t), z(t)) \le \mathcal{E}(t, \theta, \tilde{y}, \tilde{z}) + \mathcal{D}(z(t), \tilde{z}) \qquad \forall (\tilde{y}, \tilde{z}) \in \mathcal{Q} \tag{S}$$

and the energy balance

$$\begin{aligned} \mathcal{E}(t, \theta, y(t), z(t)) + \mathrm{Diss}_{\mathcal{D}}(z; [s, t]) &= \mathcal{E}(s, \theta, y(s), z(s)) \\ &+ \int_s^t \partial_t \mathcal{E}(a, \theta, y(\theta), z(\theta))\, \mathrm{d}a \end{aligned} \tag{E}$$

are satisfied for any $0 \le s < t \le T$.

An important role is played by the set of so-called stable states, defined for each $t \in [0, T]$ as

$$\begin{aligned} \mathbb{S}(t) := \{(y, z) \in \mathcal{Q} :\ & \mathcal{E}(t, \theta, y, z) < +\infty \text{ and } \mathcal{E}(t, \theta, y, z) \le \mathcal{E}(t, \theta, \tilde{y}, \tilde{z}) \\ & + \mathcal{D}(z, \tilde{z})\ \ \forall (\tilde{y}, \tilde{z}) \in \mathcal{Q}\}. \end{aligned}$$

4.2 Existence of an Energetic Solution

A standard way how to prove the existence of an energetic solution is to construct time-discrete minimization problems and then to pass to the limit. Before we give the existence proof we need some auxiliary results. For given $N \in \mathbb{N}$ and for $0 \le k \le N$, we define the time increments $t_k := kT/N$. Furthermore, we use the abbreviation $q := (y, z) \in \mathcal{Q}$. We assume that there exists an admissible deformation y^0 being compatible with the initial volume fraction z^0, i.e., $q^0 := (y^0, z^0) \in \mathbb{S}(0)$. For $k = 1, \dots, N$, we define a sequence of minimization problems

$$\text{minimize } \mathcal{I}_k(\theta, y, z) := \mathcal{E}(t_k, \theta, y, z) + \mathcal{D}(z, z^{k-1}), \quad (y, z) \in \mathcal{Q}. \tag{21}$$

We denote a minimizer of (21), for a given k, as $q_k^N := (y^k, z^k) \in \mathcal{Q}$ for $1 \le k \le N$. The following lemma shows that a minimizer always exists if the elastic energy is not identically infinite on $\mathcal{Q}$:

Lemma 1 *Let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain, and let $\Gamma = \Gamma_0 \cup \Gamma_1$ be an $\mathcal{H}^2$-measurable partition of $\Gamma = \partial\Omega$ with the area of $\Gamma_0 > 0$. Let J, of the from (8), be gradient polyconvex on Ω and such that the stored energy density $\hat{W}$ satisfies (10). Moreover, let $L \in C^1[0, T] \times W^{1,p}(\Omega; \mathbb{R}^3)$ be such that, for some $C > 0$ and $1 \le \alpha < p$,*

$$L(t, y) \le C\|y\|^{\alpha}_{W^{1,p}}, \quad \textit{for all } t \in [0, T]$$

and $y \mapsto -L(t, y)$ is weakly lower semicontinuous on $W^{1,p}(\Omega; \mathbb{R}^3)$ for all $t \in [0, T]$. Finally, let $p > 6$, $q \ge \frac{p}{p-1}$, $r > 1$, $s > 2p/(p-6)$.

If there is $(y, z) \in \mathcal{Q}$ such that $\mathcal{I}_k(y, z) < \infty$ for $\mathcal{I}_k$ from (21), then the functional $\mathcal{I}_k$ has a minimizer $q_k^N = (y^k, z^k) \in \mathcal{Q}$ such that y_k is injective everywhere in Ω. Moreover, $q_k^N \in \mathbb{S}(t_k)$ for all $1 \le k \le N$.

Proof Since the discretized problem (21) has a purely static character, we can follow the proof of Proposition 1. Let $\{(y_j^k, z_j^k)\}_{j\in\mathbb{N}} \subset \mathcal{Q}$ be a minimizing sequence. As

$$\nabla y_j^k \longrightarrow \nabla y^k \quad \text{strongly in } L^{\tilde{p}}(\Omega, \mathbb{R}^{3\times 3}) \text{ as } j \to \infty$$

for every $1 \le \tilde{p} < p$ and $\lambda \in C(\mathbb{R}^{3\times 3}, \mathbb{R}^{M+1})$ is bounded, we obtain that

$$z_j^k = \lambda(\nabla y_j^k) \longrightarrow \lambda(\nabla y^k) \quad \text{strongly in } L^{\tilde{p}}(\Omega, \mathbb{R}^{M+1}) \text{ as } j \to \infty.$$

Since $\|z_j^k\|_{L^1(\Omega,\mathbb{R}^{M+1})}$ is uniformly bounded in j, there is a subsequence (not explicitly relabeled) such that $z_j^k \overset{*}{\rightharpoonup} \mu^k$ in Radon measures on Ω. This shows that $z^k := \mu^k = \lambda(\nabla y^k)$ and hence $q_k^N = (y^k, z^k) \in \mathcal{Q}$. Since $\mathcal{D}(\cdot, z^{k-1})$ is convex, we obtain that q_k^N is indeed a minimizer of $\mathcal{I}_k$. Moreover, y_k is injective everywhere by the reasoning used for proving Corollary 1. The stability $q_k^N \in \mathbb{S}(t_k)$ follows by standard arguments; see, e.g., [25].

Denoting by $B\,([0,T];\mathcal{A})$ the set of bounded maps $t \in [0,T] \mapsto y(t) \in \mathcal{A}$, we have the following result showing the existence of an energetic solution to the problem $(\mathcal{Q},\mathcal{E},\mathcal{D})$:

Theorem 1 *Let $\theta > 0$ be fixed. Let $T > 0$ and let the assumptions in Lemma 1 be satisfied. Moreover, let the initial condition be stable, i.e., $q^0 := (y^0, z^0) \in \mathbb{S}(0)$. Then there is an energetic solution to $(\mathcal{Q},\mathcal{E},\mathcal{D})$ satisfying $q(0) = q^0$ and such that $y \in B\,([0,T];\mathcal{A})$, $z \in \mathrm{BV}\big([0,T]; L^1(\Omega;\mathbb{R}^{M+1})\big) \cap L^\infty(0,T;\mathcal{Z})$, and such that for all $t \in [0,T]$ the identity $\lambda(\nabla y(t,\cdot)) = z(t,\cdot)$ holds a.e. in Ω. Moreover, for all $t \in [0,T]$, the deformation $y(t)$ is injective everywhere in Ω.*

Proof Let $q_k^N := (y^k, z^k)$ be the solution of (21), which exists by Lemma 1, and let $q^N : [0,T] \to \mathcal{Q}$ be given by

$$q^N(t) := \begin{cases} q_k^N, & \text{if } t \in [t_k, t_{k+1}[\text{ if } k = 0,\dots,N-1, \\ q_N^N, & \text{if } t = T. \end{cases}$$

Following [25], we get, for some $C > 0$ and for all $N \in \mathbb{N}$, the estimates

$$\|z^N\|_{BV(0,T;L^1(\Omega;\mathbb{R}^{M+1}))} \le C, \qquad \|z^N\|_{L^\infty(0,T;BV(\Omega;\mathbb{R}^{M+1}))} \le C, \tag{22a}$$

$$\|y^N\|_{L^\infty(0,T;W^{1,p}(\Omega;\mathbb{R}^3))} \le C, \tag{22b}$$

as well as the following two-sided energy inequality

$$\begin{aligned}\int_{t_{k-1}}^{t_k} \partial_t \mathcal{E}(a,\theta,q_k^N)\,\mathrm{d}a &\le \mathcal{E}(t_k,\theta,q_k^N) + \mathcal{D}(z^k, z^{k-1}) - \mathcal{E}(t_{k-1},\theta,q_{k-1}^N) \\ &\le \int_{t_{k-1}}^{t_k} \partial_t \mathcal{E}(a,\theta,q_{k-1}^N)\,\mathrm{d}a. \end{aligned}\tag{23}$$

The second inequality in (23) follows since q_k^N is a minimizer of (21) and by comparison of its energy with $q := q_{k-1}^N$. The lower estimate is implied by the stability of $q_{k-1}^N \in \mathbb{S}(t_{k-1})$, see Lemma 1, when compared with $\tilde{q} := q_k^N$. By this inequality, the a-priori estimates and a generalized Helly's selection principle [41, Cor. 2.8], we get that there is indeed an energetic solution obtained as a limit for $N \to \infty$.

Let us comment more on the two main properties of the minimizer, namely that it is orientation-preserving and injective everywhere in Ω. The condition $\det \nabla y > 0$ a.e. in Ω follows from the fact that if $t_j \to t$, $(y_{(j)}, z_{(j)}) \in \mathbb{S}(t_j)$ and $(y_{(j)}, z_{(j)}) \rightharpoonup$

(y,z) in $W^{1,p}(\Omega;\mathbb{R}^3)\times BV(\Omega;\mathbb{R}^{M+1})$, then $(y,z)\in\mathbb{S}(t)$. Indeed, we have $z_{(j)}\to z$ in $L^1(\Omega;\mathbb{R}^{M+1})$ in our setting and hence for all $(\tilde{y},\tilde{z})\in\mathcal{Q}$, we get

$$\begin{aligned}\mathcal{E}(t,\theta,y,z) &\le \liminf_{j\to\infty}\mathcal{E}(t_j,\theta,y_{(j)},z_{(j)}) \le \liminf_{j\to\infty}(\mathcal{E}(t_j,\theta,\tilde{y},\tilde{z})+\mathcal{D}(z_{(j)},\tilde{z}))\\ &= \mathcal{E}(t,\theta,\tilde{y},\tilde{z})+\mathcal{D}(z,\tilde{z}).\end{aligned}$$

In particular, as $\mathcal{E}(t_j,\theta,\tilde{y},\tilde{z})$ is finite for some $(\tilde{y},\tilde{z})\in\mathcal{Q}$, we get $\mathcal{E}(t,\theta,y,z)<+\infty$ and thus $\det\nabla y>0$ a.e. in Ω in view of (10).

To prove injectivity, we profit again from the fact that quasistatic evolution of energetic solutions is very close to a purely static problem. In view of (22b), we obtain, for each $t\in[0,T]$, all necessary convergences that were used in the proof of Corollary 1 to pass to the limit in the conditions (18) and (19).

5 Computational Experiments

In this section, we demonstrate computational performance of the above model on a numerical experiment. We will use a *St.Venant-Kirchhoff*-like form of the stored energy of each particular phase variant, which allows for an explicit reference to measured data and can easily be applied to various materials. We consider that the material can occur in $M+1$ stress-free configurations that are determined by *distortion matrices* F_i, $i=0,\ldots,M$, which are independent of θ, i.e., thermal expansion is neglected. The austenite well is defined by $F_0=\mathbb{I}$.

The frame-indifferent free energy of particular phase (variant) is considered as a function of *Green strain* tensor ε^ℓ related to the distortion of this phase(variant). In the simplest case (cf. [47, Sect. 6.6], or [35], e.g.), one can consider a function quadratic in terms of ε^ℓ of the form (if $\det F>0$)

$$\begin{aligned}W_\ell(F,\theta) &= \sum_{i,j,k,l=1}^{d}\varepsilon^\ell_{ij}\mathcal{C}^\ell_{ijkl}\varepsilon^\ell_{kl}+d_\ell(\theta)+\alpha((\det F)^{-2}+|\nabla[\operatorname{Cof}F]|^2),\\ \varepsilon^\ell &= \frac{(F_\ell^\top)^{-1}F^\top F F_\ell^{-1}-\mathbb{I}}{2},\end{aligned}\tag{24}$$

where $\mathcal{C}^\ell=\{\mathcal{C}^\ell_{ijkl}\}$ is the fourth-order tensor of elastic moduli satisfying the usual symmetry relations depending also on symmetry of the specific phase(variant) ℓ and d_ℓ is some offset. The overall stored energy is assembled as in (5).

The data required for the potential are available for many alloys, except perhaps the measurements of the elastic tensor $\mathcal{C}^\ell$, which are standardly done (with few exceptions) only for the austenite so that elastic response of the martensitic variants has to be extrapolated. The heat capacities c_ℓ are usually obtained experimentally, while the offsets d_ℓ are then to be fitted to get the agreement with energetical equilibrium between martensite and austenite at a specific temperature. Typically, heat capacity of austenite is larger than that of martensite, which is just what causes the shape memory effect.

We performed our computation on a prismatic single crystal of Ni_2MnGa in a specific orientation, mostly (1,0,0). This alloy (or, more precisely, intermetallic) undergoes a cubic/tetragonal transformation, which is relatively easy to model because the martensite forms only 3 variants, i.e., $M = 3$.

Following [12] we describe the variants of martensite by $F_1 = \mathrm{diag}(\eta_2, \eta_1, \eta_1)$, $F_2 = \mathrm{diag}(\eta_1, \eta_2, \eta_1)$ and $F_3 = \mathrm{diag}(\eta_1, \eta_1, \eta_2)$ where $\eta_1 = 0.9512$ and $\eta_2 = 1.130$. The stretch tensor of the austenite is the identity, i.e., $F_0 = \mathrm{diag}(1, 1, 1)$. The Euclidean distance between any two variants of the martensite is about 0.253 while the distance between the austenite and any variant of the martensite is 0.147. The distances here are calculated as the Frobenius norms of the corresponding right Cauchy-Green strains. Hence, we can define $\mathcal{N}_i(C_i) = \{C \in \mathbb{R}^{3\times 3} : |C - C_i| < \epsilon_i\}$ for some $\epsilon_i > 0$. Then

$$\mathrm{dist}(C, \mathcal{N}_i(C_i)) = \begin{cases} 0 & \text{if } |C - C_i| < \epsilon_i, \\ |C - C_i| - \epsilon_i & \text{otherwise.} \end{cases}$$

We can take $\epsilon_i = 0.07$ for every $0 \le i \le 3$. This formula is then used in (6). As the elastic moduli are much bigger than the transformation strains, the volume fraction λ will have one dominant component because ∇y must be pointwise in a small vicinity of some energy well. Using [5] we can see that the martensitic variants are rank-one connected with each other while none of them is rank-one connected with the austenite. Rank-one connection allows for the formation of a planar interface between two martensitic variants.

We prescribe the dissipation energy density as 0.35 MPa for transformations between the austenite and any martensitic variant [1] and almost no dissipation is assumed for transformations among martensitic variants. This can be done by setting $|z|_4 := \sum_{i=0}^{3} \gamma_i |z_i|$ in (20) and taking $\gamma_0 = 35 \times 10^4$ Pa and $\gamma_i = 1$ Pa if $i \neq 0$. The equilibrium temperature θ_0 of the austenite and the martensite is about 288 K. The Clausius-Clapeyron constant describing the rate of the increase of the bottoms of the martensitic wells with respect to the austenite is about 0.2 MPa/K. Therefore, we can take $d_\ell(\theta) = 0.2\,\mathrm{MPa}\,(\theta - 288\,\mathrm{K})$ for $\ell > 0$ and $d_0(\theta) = 0$.

Elastic moduli of the austenite are taken zero but $\mathcal{C}^0_{1111} = 136$ GPa, $\mathcal{C}^0_{1122} = \mathcal{C}^0_{2211} = 92$ GPa, $\mathcal{C}^0_{2323} = \mathcal{C}^0_{2332} = \mathcal{C}^0_{3223} = \mathcal{C}^0_{3232} = 102$ GPa.

We consider a simple problem of uniaxial tension of a three-dimensional bar, i.e., the horizontal displacements are fixed at the left end and all the nodes at the right end are loaded by increasing horizontal displacements, while the vertical displacements at the both ends are prescribed such as the rigid body modes are removed but the bar is free to deform laterally. In the case the bar is considered as perfectly uniform, the onset of phase transition from austenite to martensite is reached for all the points at the same time. This situation can be studied analytically, assuming zero dissipation for simplicity. First, we know that the only nonzero component of the second Piola-

Kirchhoff stress tensor S^ℓ is S^ℓ_{33} calculated as

$$S^\ell_{33} = C_{33}\varepsilon^\ell_{33} + C_{23}\varepsilon^\ell_{22} + C_{13}\varepsilon^\ell_{11}. \tag{25}$$

The condition of zero stress components S^ℓ_{11} and S^ℓ_{22} can be written as

$$S^\ell_{11} = C_{11}\varepsilon^\ell_{11} + C_{12}\varepsilon^\ell_{22} + C_{13}\varepsilon^\ell_{33} = 0 \tag{26}$$

$$S^\ell_{22} = C_{12}\varepsilon^\ell_{11} + C_{22}\varepsilon^\ell_{22} + C_{23}\varepsilon^\ell_{33} = 0, \tag{27}$$

where C_{ij} are components of the stiffness tensor in Voigt notation, i.e., $C_{12} = C_{21} = C_{2211} = C_{1122}$, $C_{22} = C_{2222}$, $C_{23} = C_{23} = C_{2233} = C_{3322}$, etc. Solution of the above system of two equations is given as

$$\varepsilon^\ell_{11} = \varepsilon^\ell_{22} \tag{28}$$

$$\varepsilon^\ell_{22} = -\frac{C_{23}}{C_{22} + C_{12}}\varepsilon^\ell_{33}. \tag{29}$$

Substituting back to (25) we arrive at

$$S^\ell_{33} = \underbrace{\left(C_{33} - 2\frac{C^2_{23}}{C_{22} + C_{23}}\right)}_{K}\varepsilon^\ell_{33}. \tag{30}$$

The transformation from austenite to the first variant of martensite happens when the energy of both phases is the same

$$W_0(F) = W_3(F) \tag{31}$$

which can be written in terms of strain as

$$K(\varepsilon^0_{33})^2 = K(\varepsilon^3_{33})^2, \tag{32}$$

where the strains are calculated as

$$\varepsilon^0_{33} = \frac{1}{2}\left(F^2_{33} - 1\right) \tag{33}$$

$$\varepsilon^3_{33} = \frac{1}{2}\left(\frac{F^2_{33}}{\eta^2_2} - 1\right). \tag{34}$$

Therefore, the critical stretch F_c of the bar at the onset of transformation from austenite to martensite can be determined as

$$F_c = \sqrt{\frac{2\eta_2^2}{\eta_2^2 + 1}} \tag{35}$$

for the given value of $\eta_2 = 1.13$, the stretch is $F_c = 1.059$, and the strains are

$$\varepsilon_{33}^0 = \frac{1}{2}\left(F_c^2 - 1\right) = 0.0608 \tag{36}$$

$$\varepsilon_{33}^3 = \frac{1}{2}\left(\frac{F_c^2}{\eta_2^2} - 1\right) = -0.0608. \tag{37}$$

The solution is represented graphically in Fig. 1.

Moreover, also remaining nonzero components of the strain tensor before and after transformation can be calculated as

$$\varepsilon_{22}^0 = -\frac{C_{23}}{C_{33} + C_{23}}\varepsilon_{33}^0 = -0.608\frac{92}{136 + 92} = -0.0245 \tag{38}$$

$$\varepsilon_{22}^3 = -\frac{C_{23}}{C_{33} + C_{23}}\varepsilon_{11}^1 = 0.608\frac{92}{136 + 92} = 0.0245 \tag{39}$$

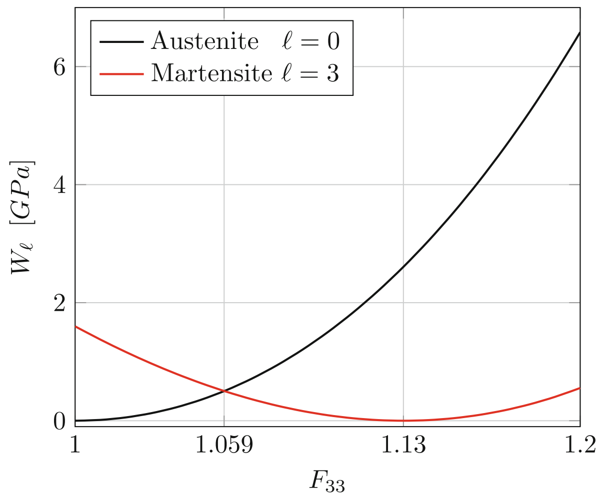

Fig. 1 Uniaxial tension: free energy of particular phase(variant)s, namely W_0 and W_3 in terms of F_{33}

and the stretches in the lateral direction before and after deformation are therefore given as

$$F_{22}^0 = \sqrt{2\varepsilon_{22}^0 + 1} = 0.9752 \tag{40}$$

$$F_{22}^3 = \sqrt{2\varepsilon_{22}^3 + 1} = 1.0242. \tag{41}$$

Let us now calculate also the stress at the point of transition from austenite to martensite. The first Piola-Kirchhoff stresses right before and after the transformation, i.e., P_{33}^0 and P_{33}^1 are calculated as

$$P_{33}^0 = F_c S_{33}^0 = F_c \left(C_{33} - 2\frac{C_{23}^2}{C_{22} + C_{23}} \right) \varepsilon_{33}^0 \tag{42}$$

$$= 1.059 \left(136 - 2\frac{92}{136 + 92} \right) 0.0608 = 8.705 \text{ GPa}. \tag{43}$$

$$P_{33}^3 = F_c S_{33}^3 = F_c \left(C_{33} - 2\frac{C_{23}^2}{C_{22} + C_{23}} \right) \varepsilon_{33}^1 \tag{44}$$

$$= 1.059 \left(136 - 2\frac{92}{136 + 92} \right) (-0.0608) = -8.705 \text{ GPa}. \tag{45}$$

Interestingly, jump from tension to compression occurs during the transformation, see Fig. 2 for the dependence of the first Piola-Kirchhoff stress on the stretch.

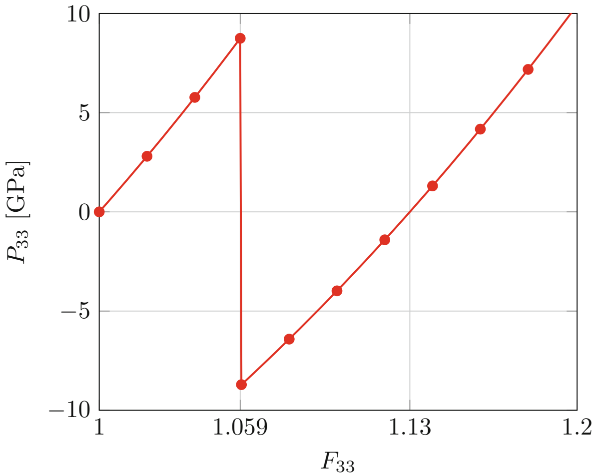

Fig. 2 Uniaxial tension: first Piola-Kirchoff stress—stretch graph

However, in reality, the material is never homogeneous and uniform but shows certain variation in material properties. Such a variation can trigger the transformation from austenite to martensite only in a small part of the bar. Nonetheless, such a uniaxial state would violate the equilibrium condition as well as the compatibility condition since the distortion matrices F_0 and F_3 are not rank-1 connected. Therefore, the bar must deform in a more complex way that is in general not possible to study analytically. Therefore, we simulate this case by the finite element method.

The proposed material model enhanced by gradient polyconvexity has been implemented into a finite element code OOFEM [45]. The implementation of gradient polyconvexity was based on the so-called micromorphic approach, see [31] for more details. Thus, in the present example we perform a uniaxial tension test of a bar with η_2 considered as a random variable with a Gaussian distribution, specified by mean $\mu = 1.13$ and standard deviation parameter $\sigma = 0.01$. As expected, the martensite transformation starts in several separated parts of the bar leading to violation of uniaxial stress state resulting into bending of the bar. Moreover, since the variants $\ell = 0$ and $\ell = 3$ are not rank-1 connected, an interface consisting of the other two variants of martensite is created. The transformation process is depicted in Fig. 3 where gradual change from the initial austenite state to the final state of martensite variant $\ell = 3$ is shown.

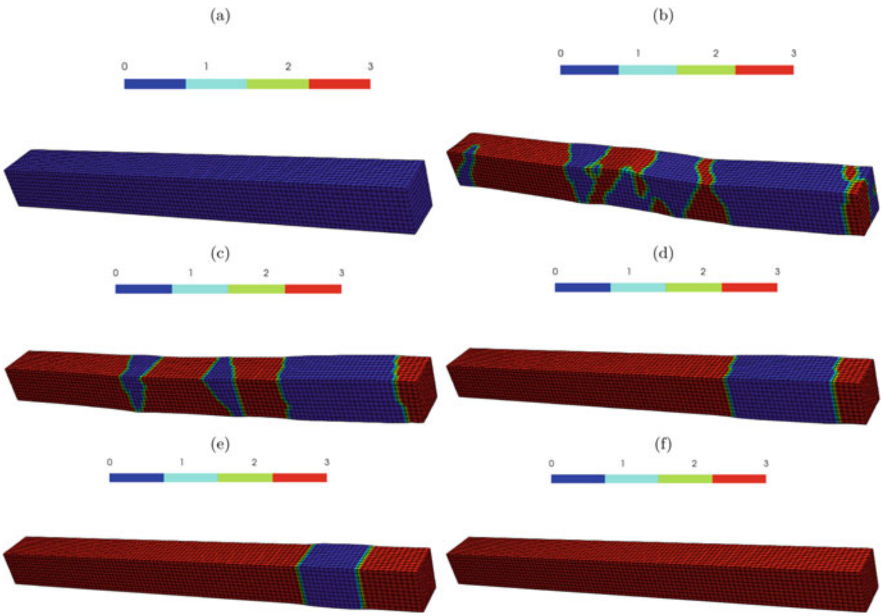

Fig. 3 Uniaxial tension test: evolution of a austenite-martensite transformation form (**a**) to (**f**). Blue color represents the austenite variant, while the remaining colors represent different variants of martensite according to the color bar

Note that the solution was obtained by the Newton-Raphson procedure which generally leads to a critical point rather than the global minima. Since the present problem involves several local minima, a more robust technique will be further implemented into OOFEM to allow development of austenite-martensite laminates without perturbing material parameters.

Acknowledgment MH and MK acknowledge the support of the Czech Science Foundation (project No. 19-26143X). P. P. was supported by Deutsche Forschungsgemeinschaft via SPP 2256 (project ID 441470105), project Analysis for thermo-mechanical models with internal variables.

References

1. Arndt, M., Griebel, M., Novák, V., Roubíček, T., Šittner, P.: Martensitic transformation in NiMnGa single crystals: numerical simulation and experiments. Int. J. Plasticity **22**, 1943–1961 (2006)
2. Ball, J.M.: Convexity conditions and existence theorems in nonlinear elasticity. Arch. Ration. Mech. Anal. **63**, 337–403 (1977)
3. Ball, J.M.: Global invertibility of Sobolev functions and the interpenetration of matter. Proc. Roy. Soc. Edinb. **88A**, 315–328 (1981)
4. Ball, J.M., Crooks, E.C.M.: Local minimizers and planar interfaces in a phase-transition model with interfacial energy. Calc. Var. **40**, 501–538 (2011)
5. Ball, J.M., James, R.D.: Fine phase mixtures as minimizers of energy. Arch. Ration. Mech. Anal. **100**, 13–52 (1988)
6. Ball, J.M., Mora-Corral, C.: A variational model allowing both smooth and sharp phase boundaries in solids. Commun. Pure Appl. Anal. **8**, 55–81 (2009)
7. Ball, J.M., Currie, J.C., Olver, P.L.: Null Lagrangians, weak continuity, and variational problems of arbitrary order. J. Funct. Anal. **41**, 135–174 (1981)
8. Benešová, B., Kružík, M.: Characterization of gradient Young measures generated by homeomorphisms in the plane. ESAIM Control Optim. Calc. Var. **22**, 267–288 (2016)
9. Benešová, B., Kružík, M.: Weak lower semicontinuity of integral functionals and applications. SIAM Rev. **59**, 703–766 (2017)
10. Benešová, B., Kružík, M., Schlömerkemper, A.: A note on locking materials and gradient polyconvexity. Math. Mod. Meth. Appl. Sci. **28**, 2367–2401 (2018)
11. Bhattacharya, K.: Microstructure of Martensite. Why It Forms and How It Gives Rise to the Shape-Memory Effect. Oxford Univ. Press, New York (2003)
12. Bhattacharya, K., James, R.D.: A theory of thin films of martensitic materials with applications to microactuators. J. Mech. Phys. Solids **47**, 531–576 (1999)
13. Capriz, G.: Continua with latent microstructure. Arch. Ration. Mech. Anal. **90**, 43–56 (1985).
14. Ciarlet, P.G.: Mathematical Elasticity Vol. I: Three-dimensional Elasticity. North-Holland, Amsterdam (1988)
15. Ciarlet, P.G., Nečas, J.: Injectivity and self-contact in nonlinear elasticity. Arch. Ration. Mech. Anal. **97**, 171–188 (1987)
16. Conti, S., Dolzmann, G.: On the theory of relaxation in nonlinear elasticity with constraints on the determinant. Arch. Ration. Mech. Anal. **217**, 413–437 (2015)
17. Dacorogna, B.: Direct Methods in the Calculus of Variations, 2nd edn. Springer, Berlin (2008)
18. Davoli, E., Friedrich, M.: Two-well rigidity and multidimensional sharp-interface limits for solid–solid phase transitions. Calc. Var. PDE **59**, 44 (2020)
19. Davoli, E., Friedrich, M.: Two-well linearization for solid-solid phase transitions. Preprint, arXiv:2005.03892 (2020)

20. Dell'Isola, F., Steigmann, D.: A two-dimensional gradient-elasticity theory for woven fabrics. J. Elast. **118**, 113–125 (2014)
21. Dell'Isola, F., Sciarra, G., Vidoli, S.: Generalized Hooke's law for isotropic second gradient materials. Proc. R. Soc. Lond. A **465**, 2177–2196 (2009)
22. Dunn, J.E., Serrin, J.: On the thermomechanics of interstitial working. Arch Ration. Mech. Anal. **88**, 95–133 (1985)
23. Fonseca, I., Leoni, G.: Modern Methods in the Calculus of Variations: L^p Spaces. Springer, New York (2007)
24. Forest, S.: Micromorphic approach for gradient elasticity, viscoplasticity, and damage. J. Eng. Mech. **135**, 117 (2009)
25. Francfort, G., Mielke, A.: Existence results for a class of rate-independent material models with nonconvex elastic energies. J. Reine Angew. Math. **595**, 55–91 (2006)
26. Friedrich, M., Kružík, M.: On the passage from nonlinear to linearized viscoelasticity. SIAM J. Math. Anal. **50**, 4426–4456 (2018)
27. Friedrich, M., Kružík, M.: Derivation of von Kármán plate theory in the framework of three-dimensional viscoelasticity. Arch. Ration. Mech. Anal. **238**, 489–540 (2020)
28. Giaquinta, M., Modica, M., Souček, J.: Cartesian currents, weak diffeomorphisms and existence theorems in nonlinear elasticity. Arch. Ration. Mech. Anal. **105**, 97–159 (1990). Erratum and Addendum **109**, 385–392 (1990)
29. Green, A.E., Rivlin, R.S.: Multipolar continuum mechanics. Arch. Ration. Mech. Anal. **17**, 113–147 (1964)
30. Hencl, S., Koskela, P.: Lectures on Mappings of Finite Distortion. LNM 2096, Springer, Cham (2014)
31. Horák, M., Kružík, M.: Gradient polyconvex material models and their numerical treatment. Int. J. Solid Struct. **195**, 57–65 (2020)
32. Korteweg, D.J.: Sur la forme que prennent les équations du mouvement des fuides si lón tient compte des forces capillaires causées par des variations de densité considérables mais continues et sur la théorie de la capillarité dans l'hypothèse d'une variation continue de la densité. Arch. Néerl. Sci. Exactes Nat. **6**, 1–24 (1901)
33. Koumatos, K., Rindler, F., Wiedemann, E.: Orientation-preserving Young measures. Q. J. Math. **67**, 439–466 (2016)
34. Kouranbaeva, S., Shkoller, S.: A variational approach to second-order multisymplectic field theory. J. Geom. Phys. **35**, 333–366 (2000)
35. Kružík, M., Roubíček, T.: Mesoscopic model of microstructure evolution in shape memory alloys with applications to NiMnGa. Preprint IMA No. 2003, Univ. of Minnesota, Minneapolis (2004)
36. Kružík, M., Roubíček, T.: Mathematical Methods in Continuum Mechanics of Solids. Springer Nature, Cham (2019)
37. Kružík, M., Pelech, P., Schlömerkemper, A.: Gradient polyconvexity in evolutionary models of shape-memory alloys. J. Optim. Theory Appl. **184**, 5–20 (2020)
38. Mariano, P.M.: Geometry and balance of hyperstresses. Rendiconti Lincei Mat. Appl. **18**, 311–331 (2007)
39. Mielke, A., Roubíček, T.: Rate-Independent Systems: Theory and Applications. Springer, New York (2015)
40. Mielke, A., Theil, F.: On rate-independent hysteresis models. NoDEA Nonlinear Differ. Equ. Appl. **11**, 151–189 (2004)
41. Mielke, A., Theil, F., Levitas, V.I.: A variational formulation of rate-independent phase transformations using an extremum principle. Arch. Ration. Mech. Anal. **162**, 137–177 (2002)
42. Mindlin, R.D.: Micro-structure in linear elasticity. Arch. Ration. Mech. Anal. **16**, 51–78 (1964)
43. Morrey, C.B.: Quasi-convexity and the lower semicontinuity of multiple integrals. Pac. J. Math. **2**, 25–53 (1952)
44. Müller, S.: Variational Models for Microstructure and Phase Transitions. Lecture Notes in Mathematics 1713, pp. 85–210. Springer, Berlin (1999)

45. Patzák, B.: OOFEM—an object-oriented simulation tool for advanced modeling of materials and structures. Acta Polytech. **52**, 6 (2012)
46. Pideri, C., Seppecher, P.: A second gradient material resulting from the homogenization of an heterogeneous linear elastic medium. Contin. Mech. Thermodyn. **9**, 241–257 (1997)
47. Pitteri, M., Zanzotto, G.: Continuum Models for Phase Transitions and Twinning in Crystals. Chapman & Hall, Boca Raton (2003)
48. Podio-Guidugli, P., Caffarelli, G.V.: Surface interaction potentials in elasticity. Arch. Ration. Mech. Anal. **109**, 343–383 (1990)
49. Segev, R.: Geometric analysis of hyper-stresses. Int. J. Eng. Sci. **120**, 100–118 (2017)
50. Seppecher, P., Alibert, J.-J., dell'Isola, F.: Linear elastic trusses leading to continua with exotic mechanical interactions. J. Phys. Conf. Ser. **319**, 13 pp. (2011)
51. Šilhavý, M.: Phase transitions in non-simple bodies. Arch. Ration. Mech. Anal. **88**, 135–161 (1985)
52. Toupin, R.A.: Elastic materials with couple stresses. Arch. Ration Mech. Anal. **11**, 385–414 (1962)
53. Toupin, R.A.: Theory of elasticity with couple stress. Arch. Ration. Mech. Anal. **17**, 85–112 (1964)
54. Truesdell, C., Noll, W.: The Non-Linear Field Theories of Mechanics. Springer, Berlin (2004)