A micromechanically consistent energy estimate for polycrystalline shape-memory alloys. I - General formulation

Michaël Peigney

A micromechanically consistent energy estimate for polycrystalline shape-memory alloys. I - General formulation.

Michaël Peigney

Lab Navier, Univ Gustave Eiffel, ENPC, CNRS, F-77447 Marne la Vallée, France

Abstract

We present an estimate for the effective energy of polycrystalline shape memory alloys that takes both intra-grain compatibility conditions and inter-grain constraints into account. That estimate is constructed using ideas and results from homogenization theory and from the translation method. The proposed estimate depends on the texture through 2-point statistics and is proved to satisfy known rigorous bounds that depend on the same microstructural information. Several examples are provided to give some insight on the proposed estimate and illustrate how it behaves relative to the bounds. The presented examples range from strain-driven uniaxial tension in the superelastic regime (high temperature) to the estimation of recoverable strains in the shape memory regime (low temperature).

Keywords: Shape Memory Alloys, Polycrystals, Micromechanics, Relaxation, Martensitic phase transformation

1. Introduction

The peculiar properties of Shape Memory Alloys (SMAs) result from a temperature-induced or stress-induced solid/solid phase transformation between austenite and martensite. Martensite has a crystallographic structure with less symmetry than the austenite, which implies that several martensitic variants need to be distinguished depending on the orientation of the martensitic lattice with respect to the austenitic lattice. Applying a mechanical or thermal loading gives rise to the spontaneous formation of austenite-martensite microstructures at a very fine scale. In a polycrystalline SMAs, three length scales typically coexist: the *microscopic* scale of the

Email address: michael.peigney@polytechnique.org (Michaël Peigney)

martensitic microstructures, the *mesoscopic* scale of the individual grains (usually much larger than the microscopic scale) and the *macroscopic* scale of a representative volume element of the polycrystal, consisting in numerous grains. The modeling of martensitic phase transformation has received a lot of attention and can be achieved using phenomenological and/or micromechanical considerations (see e.g. Cisse et al. (2016) for a review of constitutive models and modeling techniques). A possible approach, pioneered by Kohn (1991); Ball and James (1992), is to resort to elastic energy minimization across the scales. The guiding idea is to start from given constitutive relations at the microscopic level and to use upscaling strategies for obtaining the macroscopic energy. More specifically, each individual phase is modeled as a linear elastic material, resulting in a microscopic energy that is piecewise quadratic (and non convex). The mesoscopic energy of a single crystal is obtained by optimizing the arrangements of the phases in space so as to minimize the total energy. This so-called *relaxation* problem is difficult and exact solutions are known only in few special cases. Considering the macroscopic scale of a polycrystal adds a level of complexity as the interaction between the grains need to be taken into account. Viewing group of grains having the same orientation as homogeneous materials governed by mesoscopic energies, a polycrystalline SMAs can be seen as a composite material and its macroscopic (or effective) energy is obtained by *homogenization*. The exact expression of the effective energy remains elusive but rigorous bounds can be obtained. Making the assumption of constant strain or constant stress leads to convex bounds of the Taylor- and Sachs-type. That approach has notably been used by Hackl et al. (2008) for bounding the effective energy and by Bhattacharya and Kohn (1997) for the related problem of estimating the recoverable strains. We note that in general neither the effective energy nor the set of recoverable strains is convex. Several bounds which account for the possible nonconvexity of the effective energy – improving on the convex bounds in the process – have been proposed (Smyshlyaev and Willis, 1998; Hackl and Heinen, 2008; Peigney, 2009). Drawing ideas from the derivation of bounds, we present in this paper a nonconvex estimate for the effective energy of SMA polycrystals. The proposed estimate depends on the same information as the nonconvex bounds but has a simpler mathematical structure and is therefore easier to use. As will be discussed, that estimate avoids some shortcomings of available bounds and is tentatively closer to the true energy. The proposed estimate relies crucially on the lower bound formerly obtained in Peigney (2009) for the relaxed energy of single crystals. That lower bound is presented in Sect. 2 and illustrated on the two- and three-phase problems. Exact expressions of the relaxed energy are available in those cases (Pipkin, 1991; Kohn, 1991; Smyshlyaev and Willis, 1999) and recovered from the presented lower bound. In particular, use

of the proposed lower bound shows that (under mild conditions) the solution of the 2-phase problem actually admits a very simple expression, which does not seem to have been observed before. For the 3-phase problem corresponding to the tetragonal to orthorhombic transformation, we obtain similarly an explicit expression for the lower bound which coincides with the exact expression of Smyshlyaev and Willis (1999). Building on the presented lower bound for single crystals and drawing ideas from the construction of nonconvex bounds for polycrystals, a general estimate for the effective energy of polycrystalline SMAs is constructed in Sect. 3. The obtained estimate depends on the polycrystalline texture through the 2-point statistics. The validity of such an estimate can be assessed by comparison with rigorous bounds that depend on the same microstructural information. In that regard, we give a general proof that the proposed estimate satisfies the upper bound of Hackl and Heinen (2008) and the lower bound of Peigney (2009) on the effective energy. In Sects. 4 to Sect. 7 we consider several examples of increasing complexity corresponding to 2, 3, 7 and 13 phases per grain. Those example provide some insight on the proposed estimate and illustrate how it behaves compared to the bounds. In Sects. 4 and 2.2 we first consider relatively simple examples for which all (or most of) the calculations can be performed in closed form. In Sect. 6 and 7 we consider more complex examples related to the cubic to orthorhombic transformation and the cubic to monoclinic transformation. Concluding remarks follow in Sect. 8.

2. Relaxed energy of a single crystal

The microscopic energy in a single crystal is defined in terms of the energies of the individual phases. Adopting the geometrically linear setting, the free energy of the austenite is

$$\Psi_0(\boldsymbol{\varepsilon}) = \frac{1}{2}\boldsymbol{\varepsilon} : \mathbb{L}_0 : \boldsymbol{\varepsilon}$$

where $\boldsymbol{\varepsilon}$ is the linearized strain and $\mathbb{L}_0 \succ 0$ is the symmetric fourth-order elasticity tensor. Here and in the following, the notation $\mathbb{L} \succ 0$ is used to indicate that a fourth-order tensor $\mathbb{L}$ is positive definite, i.e. satisfies $\boldsymbol{u} : \mathbb{L} : \boldsymbol{u} > 0$ for all symmetric second-order tensors $\boldsymbol{u}$. For a given second-order tensor $\boldsymbol{a}$, we will use the similar notation $\boldsymbol{a} \succeq 0$ to indicate that $\boldsymbol{u} \cdot \boldsymbol{a} \cdot \boldsymbol{u} \geq 0$ for all vectors $\boldsymbol{u}$. The free energy of martensitic variant r ($r = 1, \cdots, n$) is

$$\Psi_i(\boldsymbol{\varepsilon}) = \frac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\tau}_i) : \mathbb{L}_i : (\boldsymbol{\varepsilon} - \boldsymbol{\tau}_i) + w_i \tag{1}$$

where $\boldsymbol{\tau}_i$ is the transformation strain and $\mathbb{L}_i \succ 0$ is the fourth-order elasticity tensor of the variant. The number n of variants and the values of the transformation strains

$\boldsymbol{\tau}_i$ are obtained from the crystallographic structure of the alloy (see Bhattacharya (2003) for details and examples). In (1), w_i is the minimum energy of martensite. A common expression for w_i is

$$w_i = \lambda_0 \frac{T - T_0}{T_0} \tag{2}$$

where λ_0 is the latent heat, T is the temperature and T_0 is the transformation temperature. It is convenient to set $\boldsymbol{\tau}_0 = 0$ and $w_0 = 0$ so that expression (1) remains valid for $i = 0$.

In the energy minimization approach of martensitic phase transformation (Kohn, 1991; Ball and James, 1992), the microscopic free energy is taken as

$$\Psi(\boldsymbol{\varepsilon}) = \min_{0 \le i \le n} \Psi_i(\boldsymbol{\varepsilon}) \tag{3}$$

and the mesoscopic free energy is the *relaxation* of Ψ, defined as

$$Q\Psi(\bar{\boldsymbol{\varepsilon}}) = \inf_{\boldsymbol{\varepsilon} \in \mathcal{K}(\bar{\boldsymbol{\varepsilon}})} \frac{1}{|\Omega|} \int_\Omega \Psi(\boldsymbol{\varepsilon})\, d\boldsymbol{x} \tag{4}$$

where

$$\mathcal{K}(\bar{\boldsymbol{\varepsilon}}) = \left\{ \boldsymbol{\varepsilon} | \exists \boldsymbol{u}(\boldsymbol{x}) \text{ such that } \boldsymbol{\varepsilon} = \frac{(\nabla \boldsymbol{u} + \nabla^T \boldsymbol{u})}{2} \text{ in } \Omega; \boldsymbol{u}(\boldsymbol{x}) = \bar{\boldsymbol{\varepsilon}} \cdot \boldsymbol{x} \text{ on } \partial\Omega \right\}. \tag{5}$$

The superscript ^{T} in (5) denotes the transpose operator. The result of (4) does not depend on the domain Ω (Dacorogna, 2007). In informal terms, Eq. (3) expresses the idea that the material locally transforms to the phase of minimum energy for a given microscopic strain $\boldsymbol{\varepsilon}$. The resulting energy function Ψ has a multiwell structure and therefore is not convex. As a consequence, minimizing sequences in (4) display oscillations at a finer and finer scale, preventing them to converge in the usual sense. This physically corresponds to the formation of austenite-martensite microstructures. Some information on those energy minimizing microstructures can obtained by introducing the relaxation at fixed martensitic volume fractions $\boldsymbol{\theta}$, defined by

$$Q\Psi(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta}) = \inf_{\{\chi_i\}} \inf_{\boldsymbol{\varepsilon} \in \mathcal{K}(\bar{\boldsymbol{\varepsilon}})} \frac{1}{|\Omega|} \int_\Omega \sum_{i=0}^{n} \chi_i(\boldsymbol{x}) \Psi_i(\boldsymbol{\varepsilon})\, d\boldsymbol{x} \tag{6}$$

for any $\boldsymbol{\theta} \in \mathcal{T}_n = \{(\theta_1 \cdots, \theta_n) \in \mathbb{R}^n_+ : \sum_{i=1}^n \theta_i \le 1\}$. For any given $\boldsymbol{\theta} \in \mathcal{T}_n$, the volume fraction of the austenite is denoted by θ_0 and given by $\theta_0 = 1 - \sum_{i=1}^n \theta_i$. The

first infimum in (6) is taken over characteristic functions $\chi_0, \cdots, \chi_n$ compatible with volume fractions $\boldsymbol{\theta}$. Such functions satisfy

$$\chi_i(\boldsymbol{x}) \in \{0,1\};\ 1 = \sum_{i=0}^{n} \chi_i(\boldsymbol{x});\ \theta_i = \frac{1}{|\Omega|}\int_\Omega \chi_i(\boldsymbol{x})\, d\boldsymbol{x}$$

for $i = 0, \cdots, n$. We have the relation (Kohn, 1991)

$$Q\Psi(\bar{\boldsymbol{\varepsilon}}) = \inf_{\boldsymbol{\theta}\in\mathcal{T}_n} Q\Psi(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta}). \tag{7}$$

Under mild conditions, the finite-dimensional minimization problem in (7) has minimizers $\boldsymbol{\theta}$. The latter correspond to the martensitic volume fractions in the microstructures that are expected to appear for a given mesoscopic strain $\bar{\boldsymbol{\varepsilon}}$. Finding the expression of the relaxation $Q\Psi(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta})$ is key to obtain the mesoscopic energy but this remains a largely unsolved problem in the general case. Some progress can be made if all the phases have the same elasticity tensor $\mathbb{L}$, which we assume from now on. In that case, $Q\Psi(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta})$ is known to have the structure

$$Q\Psi(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta}) = \sum_{i=0}^{n} \theta_i \Psi_i(\bar{\boldsymbol{\varepsilon}}) + \Psi_{mix}(\boldsymbol{\theta}) \tag{8}$$

where the mixing energy Ψ_{mix} does not depend on $\bar{\boldsymbol{\varepsilon}}$ (Govindjee et al., 2003). The exact expression of Ψ_{mix} is generally not available but rigorous bounds can be obtained. In that regard, a useful lower bound on Ψ_{mix} is the Reuss bound Ψ_{Reuss} (Govindjee et al., 2003), given by

$$\Psi_{Reuss}(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{i=1}^{n} \theta_i \boldsymbol{\tau}_i : \mathbb{L} : \boldsymbol{\tau}_i + \frac{1}{2}\sum_{i,j=1}^{n} \theta_i\theta_j \boldsymbol{\tau}_i : \mathbb{L} : \boldsymbol{\tau}_j. \tag{9}$$

Replacing Ψ_{mix} with Ψ_{Reuss} in (8) leads to a lower bound on $Q\Psi(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta})$, denoted by $\Psi^C(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta})$ and equal to

$$\Psi^C(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta}) = \frac{1}{2}(\bar{\boldsymbol{\varepsilon}} - \sum_{i=1}^{n} \theta_i \boldsymbol{\tau}_i) : \mathbb{L} : (\bar{\boldsymbol{\varepsilon}} - \sum_{i=1}^{n} \theta_i \boldsymbol{\tau}_i) + \sum_{i=1}^{n} w_i\theta_i. \tag{10}$$

The resulting lower bound on $Q\Psi$, obtained by replacing $Q\Psi(\bar{\boldsymbol{\varepsilon}}, \theta)$ with $\Psi^C(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta})$ in (7), is the convexification (or convex envelope) of Ψ, i.e. the largest convex function F verifying $F \leq \Psi$. The convex envelope of Ψ, henceforth denoted by Ψ^C, gives

the exact expression of the mesoscopic energy $Q\Psi$ if all the phases are pairwise compatible, i.e. for any $0 \leq i, j \leq n$ there exists some vectors $(\boldsymbol{a}, \boldsymbol{n})$ such that

$$\boldsymbol{\tau}_i - \boldsymbol{\tau}_j = \boldsymbol{a} \otimes \boldsymbol{n} + \boldsymbol{n} \otimes \boldsymbol{a}. \tag{11}$$

Relation (11) is known as the Hadamard compatibility condition between strains $\boldsymbol{\tau}_i$ and $\boldsymbol{\tau}_j$. Denoting by $\lambda_1 \leq \lambda_2 \leq \lambda_3$ the ordered eigenvalues of $\boldsymbol{\tau}_i - \boldsymbol{\tau}_j$, it is known that $\boldsymbol{\tau}_i - \boldsymbol{\tau}_j$ can be written in the form (11) iff $\lambda_2 = 0$. In common shape memory alloys, condition (11) is not satisfied for all pair of phases so that the convex envelope is not equal to $Q\Psi$. A lower bound that improves on (10) in such cases has been proposed in Peigney (2009) using ideas from the translation method (Lurie and Cherkaev, 1984; Murat, 1985; Milton, 2002). Let $\boldsymbol{\varepsilon}^*$ be the adjugate of $\boldsymbol{\varepsilon}$, i.e the tensors with components

$$\varepsilon^*_{ii} = \varepsilon_{jj}\varepsilon_{kk} - \varepsilon^2_{jk} \ , \ \varepsilon^*_{jk} = \varepsilon_{ji}\varepsilon_{ki} - \varepsilon_{jk}\varepsilon_{ii}.$$

for any (i, j, k) permutation of $(1, 2, 3)$. For any given symmetric second-order tensor $\boldsymbol{a}$, $-\boldsymbol{a} : \boldsymbol{\varepsilon}^*$ is quadratic in $\boldsymbol{\varepsilon}$ and thus can be written as

$$\frac{1}{2}\boldsymbol{\varepsilon} : \mathbb{K}(\boldsymbol{a}) : \boldsymbol{\varepsilon} = -\boldsymbol{a} : \boldsymbol{\varepsilon}^* \tag{12}$$

for some symmetric fourth-order tensor $\mathbb{K}(\boldsymbol{a})$. In indicial notation, we have $\boldsymbol{\varepsilon} : \mathbb{K}(\boldsymbol{a}) : \boldsymbol{\varepsilon} = \sum_{i,j,k,l} \varepsilon_{ij} K_{ijkl} \varepsilon_{kl}$ with

$$K_{iijj} = -a_{kk}, K_{iijk} = K_{jkii} = \frac{a_{jk}}{2}, K_{ijij} = K_{jiij} = \frac{a_{kk}}{2},$$
$$K_{ijik} = K_{jiik} = K_{ijki} = K_{jiki} = -\frac{a_{jk}}{4}$$

for any (i, j, k) permutation of $(1, 2, 3)$. It is proved in Peigney (2009) that

$$Q\Psi(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta}) \geq \Psi^C(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta}) + f(\boldsymbol{\theta}) \tag{13}$$

where

$$f(\boldsymbol{\theta}) = \sup_{\boldsymbol{a} \in \mathcal{C}} \frac{1}{4} \sum_{i,j=0}^{n} \theta_i \theta_j (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) : \mathbb{M}(\boldsymbol{a}) : (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) \tag{14}$$

with

$$\mathbb{M}(\boldsymbol{a}) = \mathbb{L} - \mathbb{L} : (\mathbb{L} - \mathbb{K}(\boldsymbol{a}))^{-1} : \mathbb{L} \tag{15}$$

and

$$\mathcal{C} = \{\boldsymbol{a} \succeq 0 : \mathbb{L} - \mathbb{K}(\boldsymbol{a}) \succ 0\}. \tag{16}$$

The bound (13) relies crucially on the fact that the function $\frac{1}{2}\boldsymbol{\varepsilon} : \mathbb{K}(\boldsymbol{a}) : \boldsymbol{\varepsilon}$ is (symmetric) quasiconvex (see Boussaid et al. (2019) for recent advances on the characterization of symmetric quasiconvex functions). For later reference, we note that Eq. (13) can be rewritten equivalently as

$$\Psi_{mix}(\boldsymbol{\theta}) \geq \Psi_{Reuss}(\boldsymbol{\theta}) + f(\boldsymbol{\theta}) \tag{17}$$

showing that $\Psi_{Reuss} + f$ is a lower bound on Ψ_{mix}. Observing that $0 \in \mathcal{C}$ and $M(0) = 0$, we further note that $f(\boldsymbol{\theta}) \geq 0$ for all $\boldsymbol{\theta}$. The term $\Psi_{Reuss} + f$ is thus potentially a tighter lower bound on Ψ_{mix} than Ψ_{Reuss}.

A special role is played by the set $\mathcal{S}$ of energy-minimizing strains, defined as

$$\mathcal{S} = \{\bar{\boldsymbol{\varepsilon}} : Q\Psi(\bar{\boldsymbol{\varepsilon}}) = \min(0, w_1)\}. \tag{18}$$

Adopting expression (2) for w_i ($i \geq 1$), the set $\mathcal{S}$ at $T < T_0$ is indeed the set of stress-free strains that can be recovered by the shape memory effect (Bhattacharya and Kohn, 1997). Energy bounds translate as bounds (in the sense of inclusion) on $\mathcal{S}$. In particular, the convex bound (10) implies that

$$\mathcal{S} \subset \Big\{ \sum_{i=1}^{n} \theta_i \boldsymbol{\tau}_i : \theta_i \geq 0; \sum_{i=1}^{n} \theta_i = 1 \Big\}, \tag{19}$$

i.e. $\mathcal{S}$ is included in the convex hull of $\boldsymbol{\tau}_1, \cdots, \boldsymbol{\tau}_n$. The energy bound (13) gives the tighter bound $\mathcal{S} \subset \mathcal{S}^+$ with

$$\mathcal{S}^+ = \Big\{ \sum_{i=1}^{n} \theta_i \boldsymbol{\tau}_i : \theta_i \geq 0; \sum_{i=1}^{n} \theta_i = 1; f(\boldsymbol{\theta}) = 0 \Big\}. \tag{20}$$

The set $\mathcal{S}^+$ in (20) does not depend on the elasticity tensor $\mathbb{L}$. It is shown indeed in Peigney (2013) that

$$\mathcal{S}^+ = \Big\{ \sum_{i=1}^{n} \theta_i \boldsymbol{\tau}_i : \theta_i \geq 0; \sum_{i=1}^{n} \theta_i = 1; \sum_{i,j=1}^{n} \theta_i \theta_i (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j)^* \leq 0 \Big\}. \tag{21}$$

Eqs. (19), (20) and (21) hold for $T < T_0$. Corresponding expressions for $T = T_0$ are obtained by starting the sums over indices i and j from 0 instead of 1.

Emphasis has been put on lower bounds on the relaxed energy because they play a pivotal role in the polycrystalline estimate to be detailed later on. Upper bounds on the relaxed energy $Q\Psi$ and lower bounds on the set $\mathcal{S}$ can be constructed using

lamination techniques (Kohn, 1991; Govindjee et al., 2003, 2007). Let us briefly elaborate on those techniques for later reference, focusing on bounding the set $\mathcal{S}$ of energy minimizing strains at $T = T_0$. The fundamental argument is that $\mathcal{S}$ is (symmetrized) rank-1 convex, i.e. that $\theta\boldsymbol{\varepsilon} + (1-\theta)\boldsymbol{\varepsilon}' \in \mathcal{S}$ for any $\theta \in [0,1]$ and any strains $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{S}^2$ that are compatible in the sense of (11) (Kohn, 1991). Using the (symmetrized) rank-1 convexity of $\mathcal{S}$ and the fact that $\{\boldsymbol{\tau}_0, \cdots, \boldsymbol{\tau}_n\} \subset \mathcal{S}$, we can see that

$$\mathcal{R}_1 = \{\theta\boldsymbol{\tau}_i + (1-\theta)\boldsymbol{\tau}_j : \theta \in [0,1]; (\boldsymbol{\tau}_i, \boldsymbol{\tau}_j) \text{ satisfying (11)}\}$$

is a lower bound on $\mathcal{S}$. Repeating the argument in a iterative fashion leads to an increasing sequence $\{\mathcal{R}_k\}$ of lower bounds, defined by the recurrence relation

$$\mathcal{R}_{k+1} = \{\theta\boldsymbol{\varepsilon} + (1-\theta)\boldsymbol{\varepsilon}' : \theta \in [0,1]; (\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{R}_k^2 \text{ satisfying (11)}\}.$$

The best lower bound is achieved by taking k are large as allowed by the complexity of the calculations. The computational complexity of $\mathcal{R}_k$ increases exponentially with k so that it is actually difficult to go beyond $k = 2$. A similar (and even more accute) difficulty holds for lamination upper bounds on the energy.

In the rest of this Section, we illustrate the lower bound (13) on two examples for which the exact expression of $Q\Psi$ is known from other arguments than those used in the derivation of (13). For those two cases, we show that (13) agrees with the exact expression. Moreover, it turns out that the use of lower bound leads to simple explicit expressions of Ψ_{mix}, as will be detailed.

2.1. The two-well problem

The two-well problem corresponds to $n = 1$, i.e. to a single variant of martensite (in addition to the austenite). The left-hand of (13) is a lower bound on $Q\Psi(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta})$ and takes the form

$$Q\Psi^-(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta}) = \Psi^C(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta}) + \frac{1}{2}\theta_1(1-\theta_1)m_- \tag{22}$$

where

$$m_- = \sup_{\boldsymbol{a}\in\mathcal{C}} \boldsymbol{\tau}_1 : \mathbb{M}(\boldsymbol{a}) : \boldsymbol{\tau}_1. \tag{23}$$

It is known (Kohn, 1991) that the exact expression of $Q\Psi(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta})$ has a structure similar to (22), i.e.

$$Q\Psi(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta}) = \Psi^C(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta}) + \frac{1}{2}\theta_1(1-\theta_1)m \tag{24}$$

where m only depends on $\boldsymbol{\tau}_1$ and $\mathbb{L}$. Explicit but intricate expressions of m have been proposed if $\mathbb{L}$ is isotropic (Kohn, 1991). In the following we use (23) to show that m actually admits a very simple expression if the Poisson's ratio ν is positive and the ordered eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3$ of $\boldsymbol{\tau}_1$ satisfy

$$\lambda_1 \leq -\nu\lambda_2 \leq \lambda_3. \tag{25}$$

It can easily be verified that (25) implies that $\lambda_1 \leq 0 \leq \lambda_3$ and we can assume (without loss of generality[1]) that $\lambda_2 \leq 0$. Since the Poisson's ratio can not exceed $1/2$, the most severe restriction that can arise from (25) is

$$\lambda_1 \leq -\frac{1}{2}\lambda_2 \leq \lambda_3. \tag{26}$$

Conditions (26) are found to hold for common of shape-memory alloys. As an example, for NiTiCu (Appendix A) we have $(\lambda_1, \lambda_2, \lambda_3) = (-0.0633, -0.0300, 0.0987)$ so that (26) is satisfied.

The point of condition (25) is to allow the maximization problem in (23) to be solved in closed form. Let us first explain how a good candidate for the optimal value of $\boldsymbol{a}$ can be guessed. Consider a second-order tensor $\boldsymbol{a}_0 \succeq 0$ such that $\|\boldsymbol{a}_0\| = 1$. We have $t\boldsymbol{a}_0 \in \mathcal{C}$ for any t positive small enough and Eq. (15) shows that $\mathbb{M}(\boldsymbol{a}) = -t\,\mathbb{K}(\boldsymbol{a}_0)$ at the first order in t. Using (12) we obtain

$$\boldsymbol{\tau}_1 : \mathbb{M}(\boldsymbol{a}) : \boldsymbol{\tau}_1 = 2t\boldsymbol{a}_0 : \boldsymbol{\tau}_1^*$$

at the first order in t. Let $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$ be eigenvectors of $\boldsymbol{\tau}_1$, so that $\boldsymbol{\tau}_1 = \sum_i \lambda_i \boldsymbol{u}_i \otimes \boldsymbol{u}_i$ and $\boldsymbol{\tau}_1^* = \sum_i \mu_i \boldsymbol{u}_i \otimes \boldsymbol{u}_i$ with

$$\mu_1 = \lambda_2\lambda_3 \leq 0, \quad \mu_2 = \lambda_1\lambda_3 \leq 0, \quad \mu_3 = \lambda_1\lambda_2 \geq 0. \tag{27}$$

Writing $\boldsymbol{a}_0$ as $\boldsymbol{a}_0 = \sum_i \alpha_i \boldsymbol{v}_i \otimes \boldsymbol{v}_i$ where $\{\alpha_i\} \in \mathbb{R}^3_+$ and $\{\boldsymbol{v}_i\}$ are the eigenvalues and eigenvectors of $\boldsymbol{a}_0$, we obtain

$$\boldsymbol{a}_0 : \boldsymbol{\tau}_1^* = \sum_{i,j=1}^{3} \alpha_i \mu_j (\boldsymbol{u}_i \cdot \boldsymbol{v}_j)^2.$$

In view of (27), the maximum value taken by $\sum_{i,j=1}^3 \alpha_i\mu_j(\boldsymbol{u}_i \cdot \boldsymbol{v}_j)^2$ over tensors $\boldsymbol{a}_0$ verifying $\boldsymbol{a}_0 \succeq 0$ and $\|\boldsymbol{a}_0\| = 1$ can easily be seen[2] to be reached for $\boldsymbol{v}_3 = \boldsymbol{u}_3$,

[1] i.e. possibly replacing $\boldsymbol{\tau}_1$ par $-\boldsymbol{\tau}_1$. The value m_- in (23) is indeed invariant by the substitution $\boldsymbol{\tau}_1 \mapsto -\boldsymbol{\tau}_1$. Moreover, if $\boldsymbol{\tau}_1$ satisfies conditions (25), so does $-\boldsymbol{\tau}_1$.

[2] We have the inequalities $\alpha_i\mu_1 \leq 0$, $\alpha_i\mu_2 \leq 0$, $\alpha_i\mu_3 \geq 0$ and the relation $\|\boldsymbol{a}_0\| = 1$ implies that $\alpha_i \leq 1$ $(i = 1, 2, 3)$. It follows that $\sum_{i,j=1}^3 \alpha_i\mu_j(\boldsymbol{u}_i \cdot \boldsymbol{v}_j)^2 \leq \mu_3 \sum_i \alpha_i(\boldsymbol{u}_i \cdot \boldsymbol{v}_3)^2 \leq \mu_3$ with equality if $\alpha_3 = 1$, $\alpha_1 = \alpha_2 = 0$ and $\boldsymbol{v}_3 = \boldsymbol{u}_3$.

$\alpha_1 = \alpha_2 = 0$, $\alpha_3 = 1$. This shows that $\boldsymbol{u}_3 \otimes \boldsymbol{u}_3$ is the best search direction for maximizing $\boldsymbol{\tau}_1 : \mathbb{M}(\boldsymbol{a}) : \boldsymbol{\tau}_1$ with respect to $\boldsymbol{a}$. Let us now proceed to maximize $\boldsymbol{\tau}_1 : \mathbb{M}(\boldsymbol{a}) : \boldsymbol{\tau}_1$ over tensors $\boldsymbol{a} \in \mathcal{C}$ of the form $Ea\boldsymbol{u}_3 \otimes \boldsymbol{u}_3$ where E is the Young's modulus. Noting that the largest eigenvalue of $\mathbb{K}(\boldsymbol{a})$ in (12) is $E\|\boldsymbol{a}\|$ and that the smallest eigenvalue of $\mathbb{L}$ is $E/(1+\nu)$, we have $Ea\boldsymbol{u}_3 \otimes \boldsymbol{u}_3 \in \mathcal{C}$ for $0 \leq a < 1/(1+\nu)$. Now consider the (3-dimensional) space $\mathcal{D}$ of second-order tensors which are diagonal in the basis $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$. For $\boldsymbol{a} = Ea\boldsymbol{u}_3 \otimes \boldsymbol{u}_3$, both $\mathbb{L}$ and $\mathbb{K}(\boldsymbol{a})$ can be observed to map $\mathcal{D}$ to itself. Thus the calculation of $\boldsymbol{\tau}_1 : \mathbb{M}(\boldsymbol{a}) : \boldsymbol{\tau}_1$ only involves the restrictions of $\mathbb{L}$ and $\mathbb{K}(\boldsymbol{a})$ on $\mathcal{D}$, which are given by the matrix representations (in the basis $\{\boldsymbol{u}_i \otimes \boldsymbol{u}_i\}_{1 \leq i \leq 3}$ of $\mathcal{D}$)

$$\mathbb{L} = \frac{E}{1-\nu-2\nu^2} \begin{bmatrix} 1-\nu & \nu & \nu \\ \nu & 1-\nu & \nu \\ \nu & \nu & 1-\nu \end{bmatrix}, \; \mathbb{K}(\boldsymbol{a}) = E \begin{bmatrix} 0 & -a & 0 \\ -a & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

A straightforward calculation shows that restriction of $\mathbb{M}(\boldsymbol{a})$ on $\mathcal{D}$ is defined by the matrix

$$\mathbb{M}(\boldsymbol{a}) = \frac{Ea}{(1-\nu a)^2 - a^2} \begin{bmatrix} -a & 1-\nu a & 0 \\ 1-\nu & -a & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so that

$$\boldsymbol{\tau}_1 : \mathbb{M}(\boldsymbol{a}) : \boldsymbol{\tau}_1 = \frac{Ea}{(1-\nu a)^2 - a^2} \left(2\lambda_1\lambda_2 - a(\lambda_1^2 + 2\nu\lambda_1\lambda_2 + \lambda_2^2)\right). \tag{28}$$

Differentiating with respect to a shows that (28) has a stationary point at

$$a = \frac{\lambda_2}{\lambda_1 + \nu\lambda_2} \in [0, \frac{1}{1+\nu}]. \tag{29}$$

The corresponding value of $\boldsymbol{\tau}_1 : \mathbb{M}(\boldsymbol{a}) : \boldsymbol{\tau}_1$ is $E\lambda_2^2$. Hence we can state that

$$m_- \geq E\lambda_2^2. \tag{30}$$

Although it is not obvious at first sight, comparing with the expressions of Kohn (1991) shows that (30) actually holds as an equality. Eq. (4.11) in Kohn (1991) gives indeed

$$\frac{m}{E} = \frac{\nu}{1-\nu-2\nu^2}(\operatorname{tr}\boldsymbol{\tau}_1)^2 + \frac{\lambda_1^2+\lambda_2^2+\lambda_3^2}{1+\nu} - \frac{X}{2(1+\nu)} \tag{31}$$

where X is obtained by taking the maximum among several branches. Under Condition (25), Eqs (4.16) and (4.17) in Kohn (1991) show that $X \geq X_-$ with

$$X_- = \frac{2}{1-2\nu}\Big(\lambda_1^2 + \lambda_3^2 + 2\nu^2\lambda_2^2 + \nu(2\lambda_2(\lambda_1+\lambda_3) - (\lambda_1-\lambda_3)^2)\Big) \tag{32}$$

For $X = X_-$, the right-hand of (31) is found to be equal to $E\lambda_2^2$, hence $m \geq E\lambda_2^2$. Comparing with (30) and noting that $m_- \leq m$, we can conclude that

$$m_- = m = E\lambda_2^2 \tag{33}$$

and that

$$E\frac{\lambda_2}{\lambda_1 + \nu\lambda_2}\boldsymbol{u}_3 \otimes \boldsymbol{u}_3$$

reaches the maximum in (23). Expression (33) will be used later in the paper and will also play an important role in the companion paper devoted to localization in polycrystals.

Remark: As detailed by Kohn (1991), the double-well structure of the relaxed energy gives rise to different transition temperatures while heating (T_0^+) and cooling (T_0^-). It can be calculated from (24) that the thermal hysteresis $\Delta T = T_0^+ - T_0^-$ is equal to m/λ_0 where λ_0 is the latent heat parameter in (2). Assuming conditions (25) to be satisfied, Eq. (33) shows that ΔT is driven by the middle eigenvalue λ_2 of the transformation strain. In particular, ΔT vanishes when $\lambda_2 = 0$ i.e. when austenite and martensite are compatible in the sense of (11). There is a formal analogy with the experimental and theoretical results of Zhang et al. (2009), who showed that the thermal hysteresis in various CuAlNi alloys is governed by the middle eigenvalue of the transformation strain and minimized when austenite and martensite are compatible.

2.2. Tetragonal to orthorhombic transformation

We consider the three-well problem corresponding to the tetragonal to orthorhombic transformation, for which there are two martensitic variants with transformation strains

$$\boldsymbol{\tau}_1 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}, \quad \boldsymbol{\tau}_2 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \gamma \end{pmatrix} \tag{34}$$

in an orthonormal basis $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$. The martensitic variants can always be numbered in such a way that $\alpha < \beta$. $YBa_2Cu_3O_{6+x}$ is an example of material obeying

that transformation (Andersen et al., 1990). As can be seen on (34), the two martensitic variants are compatible with each other in the sense of (11) but they are not compatible with the austenite (unless very special values of the lattice parameters). It turns out that closed-form expressions of the lower bound f can obtained in some situations depending on the lattice parameters (α, β, γ). One such situation, that we detail in the following, is

$$\gamma < 0,\ \alpha < 0 < \beta \text{ with } |\alpha| < \beta. \tag{35}$$

Proceeding as for the two-well problem, we first use linearization to guess a good candidate for the optimal value of $\boldsymbol{a}$ in (14). For the transformation strains in (34), we have (at the first order in $\boldsymbol{a}$)

$$\sum_{i,j=0}^{n} \theta_i\theta_j(\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) : \mathbb{M}(\boldsymbol{a}) : (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) = \frac{1}{2}\sum_{i,j=0}^{n} \theta_i\theta_j(\boldsymbol{\tau}_i - \boldsymbol{\tau}_j)^* : \boldsymbol{a} = \sum_{i=1}^{3} \mu_i \boldsymbol{u}_i \otimes \boldsymbol{u}_i$$

with

$$\mu_1 = \theta_0\gamma(\theta_1\beta + \theta_2\alpha)\ \ \mu_2 = \theta_0\gamma(\theta_1\alpha + \theta_2\beta)\ \ \mu_3 = \theta_0(1-\theta_0)\alpha\beta - \theta_1\theta_2(\alpha-\beta)^2. \tag{36}$$

Let us first consider values of $\boldsymbol{\theta}$ such that $\alpha\theta_1 + \beta\theta_2 \leq 0$. In that case, it can easily be verified from (36) that $\mu_1 \leq 0, \mu_2 \geq 0, \mu_3 \leq 0$. This indicates that $\boldsymbol{u}_2 \otimes \boldsymbol{u}_2$ is a good search direction for solving (14), in a way similar to the two-well problem considered in Sect. 2.1. Choosing $\boldsymbol{a}$ in the form $Ea\boldsymbol{u}_2 \otimes \boldsymbol{u}_2$, the stationarity condition of (14) with respect to a yields a second degree polynomial equation with roots

$$a_\pm = \frac{\nu(X_+ - X_-) + (\sqrt{X_+} \pm \sqrt{X_-})^2}{(1+\nu)^2 X_+ - (1+\nu)^2 X_-}$$

where $X_\pm = \theta_0\theta_1(\alpha \pm \gamma)^2 + \theta_0\theta_2(\beta \pm \gamma)^2 + \theta_1\theta_2(\alpha - \beta)^2$. It can be verified that $a_+ > 1/(1+\nu) > a_- \geq 0$, which shows the tensor $\boldsymbol{a}_- = Ea_-\boldsymbol{u}_2 \otimes \boldsymbol{u}_2$ is in $\mathcal{C}$. Somewhat lengthy manipulations lead to

$$\frac{1}{4}\sum_{i,j=0}^{n} \theta_i\theta_j(\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) : \mathbb{M}(\boldsymbol{a}_-) : (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) = \frac{E}{8}\left(\frac{\nu(X_+ - X_-) + (\sqrt{X_+} - \sqrt{X_-})^2}{(1+\nu)\sqrt{X_+} - (1-\nu)\sqrt{X_-}}\right)^2. \tag{37}$$

Definition (14) implies that $f(\boldsymbol{\theta})$ is larger than or equal to the value in (37). A numerical optimization with respect to all tensors in $\mathcal{C}$ (not necessarily rank-1 or diagonal) suggests that $\boldsymbol{a}_-$ indeed reaches the optimum in (14) for $\alpha\theta_1 + \beta\theta_2 \leq 0$, so that

$$f(\boldsymbol{\theta}) = \frac{E}{8}\left(\frac{\nu(X_+ - X_-) + (\sqrt{X_+} - \sqrt{X_-})^2}{(1+\nu)\sqrt{X_+} - (1-\nu)\sqrt{X_-}}\right)^2 \text{ if } \alpha\theta_1 + \beta\theta_2 \leq 0. \tag{38}$$

Results for $\alpha\theta_2 + \beta\theta_1 \leq 0$ are obtained by swapping θ_1 and θ_2 in the previous developments, i.e.

$$f(\boldsymbol{\theta}) = \frac{E}{8}\left(\frac{\nu(X'_+ - X'_-) + (\sqrt{X'_+} - \sqrt{X'_-})^2}{(1+\nu)\sqrt{X'_+} - (1-\nu)\sqrt{X'_-}}\right)^2 \text{ if } \alpha\theta_2 + \beta\theta_1 \leq 0 \tag{39}$$

with $X'_\pm = \theta_0\theta_2(\alpha \pm \gamma)^2 + \theta_0\theta_1(\beta \pm \gamma)^2 + \theta_1\theta_2(\alpha - \beta)^2$.

When $\alpha\theta_1 + \beta\theta_2 \geq 0$ and $\alpha\theta_2 + \beta\theta_1 \geq 0$, all the three eigenvalues μ_i in (36) are negative so that, at the first order in $\boldsymbol{a}$, we have $\sum_{i,j}\theta_i\theta_j(\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) : \mathbb{M}(\boldsymbol{a}) : (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) \geq 0$ for all $\boldsymbol{a} \in \mathcal{C}$. In that case no direction leading to non negative value of $\sum_{i,j}\theta_i\theta_j(\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) : \mathbb{M}(\boldsymbol{a}) : (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j)$ can be found. As will be justified shortly, it can actually be shown that

$$f(\boldsymbol{\theta}) = \Psi_{mix}(\boldsymbol{\theta}) - \Psi_{Reuss}(\boldsymbol{\theta}) = 0 \text{ if } \alpha\theta_2 + \beta\theta_1 \geq 0 \text{ and } \alpha\theta_1 + \beta\theta_2 \geq 0. \tag{40}$$

Together Eqs (38), (39) and (40) define the values taken by the function f on the whole triangle

$$\mathcal{T}_2 = \{(\theta_1, \theta_2) : \theta_1 \geq 0, \theta_2 \geq 0, \theta_1 + \theta_2 \leq 1\} \tag{41}$$

of admissible values for the martensitic volume fractions (θ_1, θ_2). In Fig. 1 is shown the obtained function f in the case $\alpha = \gamma = -0.01$, $\beta = 0.02$, $\nu = 0.3$. As illustrated in Fig. 1, the volume fractions verifying $\theta_1\alpha + \theta_2\beta \geq 0$ and $\theta_2\alpha + \theta_1\beta \geq 0$ define a triangle $\boldsymbol{OAB}$ with vertices $\boldsymbol{O} = (0,0)$, $\boldsymbol{A} = (-\frac{\alpha}{\beta-\alpha}, \frac{\beta}{\beta-\alpha})$ and $\boldsymbol{B} = (\frac{\beta}{\beta-\alpha}, -\frac{\alpha}{\beta-\alpha})$. Let us prove that $\Psi_{mix} - \Psi_{Reuss} = 0$ on that triangle. We first observe that proving that $\Psi_{mix}(\boldsymbol{\theta}) - \Psi_{Reuss}(\boldsymbol{\theta}) = 0$ amounts to show that $\boldsymbol{\varepsilon}(\boldsymbol{\theta}) = \sum_i \theta_i\boldsymbol{\tau}_i$ is in the set of energy-minimizing strains $\mathcal{S}$ at $T = T_0$. Proving that $\boldsymbol{\varepsilon}(\boldsymbol{\theta}) \in \mathcal{S}$ can be achieved using sequential lamination techniques as discussed previously in Sect.2. Since $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ are compatible in the sense of (11), we have $\theta_1\boldsymbol{\tau}_1 + \theta_2\boldsymbol{\tau}_2 \in \mathcal{S}$ for any $\boldsymbol{\theta} = (\theta_1, \theta_2) \in \mathcal{T}_2$ such that $\theta_1 + \theta_2 = 1$. Taking in particular $\boldsymbol{\theta} = \boldsymbol{A}$ and $\boldsymbol{\theta} = \boldsymbol{B}$ shows that $\boldsymbol{\varepsilon}(\boldsymbol{A}) = \mathrm{diag}(\alpha + \beta, 0, \gamma)$ and $\boldsymbol{\varepsilon}(\boldsymbol{B}) = \mathrm{diag}(0, \alpha + \beta, \gamma)$ are in $\mathcal{S}$. Observing that $\boldsymbol{\varepsilon}(\boldsymbol{A})$, $\boldsymbol{\varepsilon}(\boldsymbol{B})$ and $\boldsymbol{\tau}_0 = 0$ are pairwise compatible, we obtain that $\boldsymbol{\varepsilon}(\boldsymbol{\theta}) \in \mathcal{S}$ and consequently that $\Psi_{mix}(\boldsymbol{\theta}) - \Psi_{Reuss}(\boldsymbol{\theta}) = 0$ for any $\boldsymbol{\theta}$ in the triangle $\boldsymbol{OAB}$. Since $0 \leq f \leq \Psi_{mix} - \Psi_{Reuss}$, we also have $f = 0$ in the triangle $\boldsymbol{OAB}$, justifying the statement (40).

Those arguments show in particular that f gives the exact value of $\Psi_{mix} - \Psi_{Reuss}$ on the triangle $\boldsymbol{OAB}$. Outside of that triangle, f *a priori* gives only a lower bound on $\Psi_{mix} - \Psi_{Reuss}$. For the three-well problem, Smyshlyaev and Willis (1999) reduced the evaluation of Ψ_{mix} to a finite-dimensional optimization problem. Solving the optimization problem formulated by Smyshlyaev and Willis (1999) seems difficult to

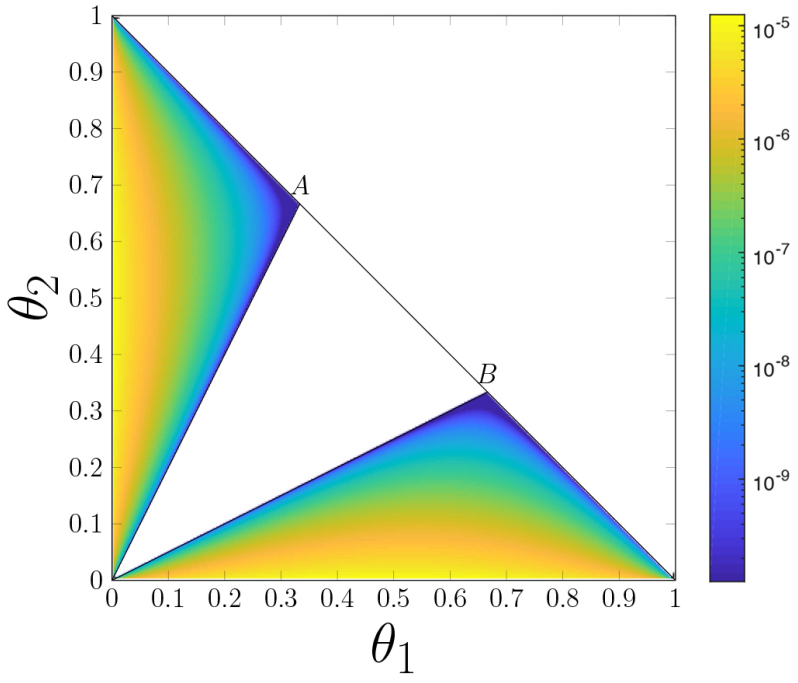

Figure 1: Values taken by f/E for the tetragonal to orthorhombic transformation. Case $\alpha = \gamma = -0.01$, $\beta = 0.02$, $\nu = 0.3$. The function f vanishes on the triangle $0AB$.

perform analytically but is tractable numerically. We have implemented the formula of Smyshlyaev and Willis (1999) to evaluate Ψ_{mix} for various choices of $(\nu, \alpha, \beta, \gamma)$ and compare the results with the function f given by the explicit expressions (38), (39) and (40). Up to the accuracy of the numerical calculations, Ψ_{mix} and $\Psi_{Reuss} - f$ were found to be equal. This suggests that the explicit expressions (38), (39) and (40) give the exact value of Ψ_{mix} – not just a lower bound.

Remark: In the analysis presented, the role of condition (35) is to ensure that $\sum \theta_i\theta_j(\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) : \mathbb{M}(\boldsymbol{a}) : (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j)$ has at most one non-negative eigenvalue. In that case, a simple search direction for $\boldsymbol{a}$ can easily be found and turns out to yield the optimal value for $\boldsymbol{a}$. Other conditions than (35) lead to a similar situation. Assuming without loss of generality that $\alpha < \beta$, it can be verified that if either

$$\gamma > 0, \alpha < 0 < \beta < |\alpha| \tag{42}$$

or

$$\alpha\beta > 0, \alpha\gamma < 0 \tag{43}$$

then problem (14) can be solved in closed-form and the optimal value of $\boldsymbol{a}$ is a rank-1 tensor of the form $a\boldsymbol{u}_i \otimes \boldsymbol{u}_i$. Unfortunately, conditions (35), (42) and (43) do not cover all possible values of the lattice parameters (α, β, γ). If for instance $\gamma < 0$, $\alpha < 0 < \beta$ but $|\alpha| > \beta$, then for some values of $\boldsymbol{\theta}$ the tensor $\sum \theta_i \theta_j (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) : \mathbb{M}(\boldsymbol{a}) : (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j)$ has two non-negative eigenvalues and the optimal value of $\boldsymbol{a}$ is a rank-2 tensor of the form $a\boldsymbol{u}_i \otimes \boldsymbol{u}_i + a'\boldsymbol{u}_j \otimes \boldsymbol{u}_j$ $(i \neq j)$. The calculations become much more intricate and do not seem tractable analytically.

3. Effective energy of a polycrystal

We now consider a polycrystal, consisting of multiple grains with distinct orientations. The domain Ω occupied by the polycrystal is divided into N subdomains $\Omega^1 \cdots, \Omega^N$ such that grains in Ω^r have the same orientation (defined by a rotation $\boldsymbol{R}^r$) relative to a reference crystal with microscopic energy Ψ. The microscopic free energy $\Psi^r(\boldsymbol{\varepsilon})$ and the mesoscopic free energy $Q\Psi^r(\boldsymbol{\varepsilon})$ in Ω^r are thus equal to $\Psi(\boldsymbol{R}^{r,T}\bar{\boldsymbol{\varepsilon}}\boldsymbol{R}^r)$ and $Q\Psi(\boldsymbol{R}^{r,T}\bar{\boldsymbol{\varepsilon}}\boldsymbol{R}^r)$, respectively. Set $\boldsymbol{\tau}_i^r = \boldsymbol{R}^r\boldsymbol{\tau}_i\boldsymbol{R}^{r,T}$ and let $\mathbb{L}^r$ be the rotated elasticity tensor, verifying $\boldsymbol{\varepsilon} : \mathbb{L}^r : \boldsymbol{\varepsilon} = (\boldsymbol{R}^{r,T}\boldsymbol{\varepsilon}\boldsymbol{R}^r) : \mathbb{L} : (\boldsymbol{R}^r\boldsymbol{\varepsilon}\boldsymbol{R}^{r,T})$ for all $\boldsymbol{\varepsilon}$. Eq. (3) shows that we can write Ψ^r as

$$\Psi^r(\boldsymbol{\varepsilon}) = \min_{0 \leq i \leq n} \Psi_i^r(\boldsymbol{\varepsilon})$$

where

$$\Psi_i^r(\boldsymbol{\varepsilon}) = \frac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\tau}_i^r) : \mathbb{L}^r : (\boldsymbol{\varepsilon} - \boldsymbol{\tau}_i^r) + w_i.$$

Eqs (7) and (8) give

$$Q\Psi^r(\boldsymbol{\varepsilon}) = \inf_{\boldsymbol{\theta} \in \mathcal{T}_n} Q\Psi^r(\boldsymbol{\varepsilon}, \boldsymbol{\theta}) \tag{44}$$

with

$$Q\Psi^r(\boldsymbol{\varepsilon}, \boldsymbol{\theta}) = \sum_{i=0}^{n} \theta_i \Psi_i^r(\boldsymbol{\varepsilon}) + \Psi_{mix}(\boldsymbol{\theta}). \tag{45}$$

We note that the function Ψ_{mix} is the same for all grains, i.e. is independent of the crystalline orientation. The effective energy Ψ_{eff} of the polycrystal is obtained by *homogenization*, i.e.

$$\Psi_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}) = \inf_{\boldsymbol{\varepsilon} \in \mathcal{K}(\bar{\boldsymbol{\varepsilon}})} \frac{1}{|\Omega|} \int_\Omega \sum_{r=1}^{N} \chi^r(\boldsymbol{x}) Q\Psi^r(\boldsymbol{\varepsilon})\, d\boldsymbol{x} \tag{46}$$

where χ^r is the characteristic function of Ω^r (i.e. $\chi^r(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in \Omega^r$ and $\chi^r(\boldsymbol{x}) = 0$ otherwise). We now derive an estimate for $\Psi_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}})$ in (46). For any $\boldsymbol{\theta}(\boldsymbol{x}) \in \mathcal{T}_n$, Eq. (44) implies that $Q\Psi^r(\boldsymbol{\varepsilon}) \leq Q\Psi^r(\boldsymbol{\varepsilon}, \boldsymbol{\theta}(\boldsymbol{x}))$. Considering a piecewise constant field

$$\boldsymbol{\theta}(\boldsymbol{x}) = \sum_{r=1}^{n} \chi^r(\boldsymbol{x}) \boldsymbol{\theta}^r \tag{47}$$

in which $\boldsymbol{\theta}(\boldsymbol{x})$ takes a uniform value $\boldsymbol{\theta}^r = (\theta_1^r, \cdots, \theta_n^r)$ in Ω^r, we find

$$\Psi_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}) \leq \inf_{\boldsymbol{\varepsilon} \in \mathcal{K}(\bar{\boldsymbol{\varepsilon}})} \frac{1}{|\Omega|} \int_\Omega \sum_{r=1}^{N} \chi^r(\boldsymbol{x}) Q\Psi^r(\boldsymbol{\varepsilon}, \boldsymbol{\theta}^r)\, d\boldsymbol{x}. \tag{48}$$

Using (45), Eq. (48) becomes

$$\Psi_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}) \leq W(\bar{\boldsymbol{\varepsilon}}) + \sum_{r=1}^{N} c_r \left(\Psi_{mix}(\boldsymbol{\theta}^r) + \sum_{i=0}^{n} \theta_i^r (\frac{1}{2} \boldsymbol{\tau}_i^r : \mathbb{L}^r : \boldsymbol{\tau}_i^r + w_i) \right) \tag{49}$$

with

$$W(\bar{\boldsymbol{\varepsilon}}) = \inf_{\boldsymbol{\varepsilon} \in \mathcal{K}(\bar{\boldsymbol{\varepsilon}})} \frac{1}{|\Omega|} \int_\Omega \sum_{r=1}^{N} \chi_r(\boldsymbol{x}) (\frac{1}{2} \boldsymbol{\varepsilon} : \mathbb{L}^r : \boldsymbol{\varepsilon} + \boldsymbol{\eta}^r : \boldsymbol{\varepsilon}) d\boldsymbol{x} \tag{50}$$

and $\boldsymbol{\eta}^r = -\mathbb{L}^r : \sum_i \theta_i^r \boldsymbol{\varepsilon}_i^r$. In (49), $c_r = |\Omega^r|/|\Omega|$ is the volume fraction of Ω^r. Solving the infimum problem in (50) amounts to solve the elasticity problem

$$\operatorname{div} \boldsymbol{\sigma} = 0, \quad \boldsymbol{\sigma}(\boldsymbol{x}) = \sum_r \chi_r(\boldsymbol{x}) (\mathbb{L}^r : \boldsymbol{\varepsilon}(\boldsymbol{x}) + \boldsymbol{\eta}^r), \quad \boldsymbol{\varepsilon} \in \mathcal{K}(\bar{\boldsymbol{\varepsilon}}).$$

The expression of $W(\bar{\boldsymbol{\varepsilon}})$ simplifies if $\mathbb{L}^1 = \cdots = \mathbb{L}^r$, which notably occurs if $\mathbb{L}$ is isotropic. In that case, the solution of (50) takes the form $\boldsymbol{\varepsilon} = \bar{\boldsymbol{\varepsilon}} - \Gamma \boldsymbol{\tau}$ where Γ is a singular integral operator related to Green's functions, and we have (Willis, 1981)

$$W(\bar{\boldsymbol{\varepsilon}}) = \frac{1}{2} \bar{\boldsymbol{\varepsilon}} : \mathbb{L} : \bar{\boldsymbol{\varepsilon}} + \sum_r c_r \boldsymbol{\eta}^r : \bar{\boldsymbol{\varepsilon}} - \frac{1}{2} \sum_{r,s} \boldsymbol{\eta}^r : \mathbb{A}_{rs} : \boldsymbol{\eta}^s \tag{51}$$

where

$$\mathbb{A}_{rs} = \frac{1}{|\Omega|} \int_\Omega \chi^r(\boldsymbol{x}) (\Gamma \chi^s)(\boldsymbol{x}) dx. \tag{52}$$

Using (51), Eq. (49) can be rewritten as

$$\begin{aligned} \Psi_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}) \leq \;& \Psi^C_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\Theta}) + \sum_r c_r \Psi_{mix}(\boldsymbol{\theta}^r) + \frac{1}{2} \sum_{r,i} c_r \theta_i^r \boldsymbol{\tau}_i^r : \mathbb{L} : \boldsymbol{\tau}_i^r \\ & - \frac{1}{2} \sum_{r,s} c_r c_s \boldsymbol{\eta}^r : \mathbb{L}^{-1} : \boldsymbol{\eta}^s - \frac{1}{2} \sum_{r,s} \boldsymbol{\eta}^r : \mathbb{A}_{rs} : \boldsymbol{\eta}^s \end{aligned} \tag{53}$$

with $\boldsymbol{\Theta} = (\boldsymbol{\theta}^1, \cdots, \boldsymbol{\theta}^r) \in \mathcal{T}_n^N$ and

$$\Psi_{\mathit{eff}}^C(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\Theta}) = \frac{1}{2}\left(\bar{\boldsymbol{\varepsilon}} - \sum_{r,i} c_r \theta_i^r \boldsymbol{\tau}_i^r\right) : \mathbb{L} : \left(\bar{\boldsymbol{\varepsilon}} - \sum_{r,i} c_r \theta_i^r \boldsymbol{\tau}_i^r\right) + \sum_{r,i} c_r \theta_i^r w_i. \tag{54}$$

The tensors $\mathbb{A}_{rs}$ in (52) take a simpler form if the texture has ellipsoidal symmetry, i.e. if the probability of finding orientation r at point $\boldsymbol{x}$ and orientation s at point $\boldsymbol{x}'$ is a function of $\|\mathbb{Z} \cdot (\boldsymbol{x}' - \boldsymbol{x})\|$ where $\mathbb{Z} \succ 0$ is a symmetric fourth-order tensor. In that case, we have (Willis, 1977, 1981)

$$\mathbb{A}_{rs} = c_r(\delta_{rs} - c_s)\mathbb{P} \tag{55}$$

where

$$\mathbb{P} = \frac{1}{4\pi \det \mathbb{Z}} \int_{\|\boldsymbol{\xi}\|=1} \mathbb{H}(\boldsymbol{\xi}) \|\mathbb{Z}^{-1} \cdot \boldsymbol{\xi}\|^{-3} d\xi. \tag{56}$$

In this last expression, $\mathbb{H}(\boldsymbol{\xi})$ is the fourth-order tensor with components $\mathbb{H}(\boldsymbol{\xi})_{ijpq} = \xi_j [\mathbb{L}(\boldsymbol{\xi})]_{ip}^{-1} \xi_q$ and $\mathbb{L}(\boldsymbol{\xi})$ is the second-order tensor with components $\mathbb{L}(\boldsymbol{\xi})_{ip} = \sum_{jq} \mathbb{L}_{ijpq} \xi_j \xi_q$. Explicit expression of the $\mathbb{P}$ tensor are available for some useful cases (Castaneda and Suquet, 1997). Replacing $\mathbb{A}_{rs}$ with $c_r(\delta_{rs} - c_s)\mathbb{P}$ in (53) and taking the infimum over $\boldsymbol{\Theta} \in \mathcal{T}_n^N$, we obtain

$$\begin{aligned} \Psi_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}) \leq \inf_{\boldsymbol{\Theta}} \quad & \Psi_{\mathit{eff}}^C(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\theta}) + \sum_r c_r \Psi_{mix}(\boldsymbol{\theta}^r) + \frac{1}{2} \sum_{r,i} c_r \theta_i \boldsymbol{\tau}_i^r : \mathbb{L} : \boldsymbol{\tau}_i^r \\ & - \frac{1}{2} \sum_r c_r \boldsymbol{\eta}^r : \mathbb{P} : \boldsymbol{\eta}^r - \frac{1}{2} \sum_{r,s} c_r c_s \boldsymbol{\eta}^r : (\mathbb{L}^{-1} - \mathbb{P}) : \boldsymbol{\eta}^s. \end{aligned} \tag{57}$$

The choice (47) of a piecewise constant field $\boldsymbol{\theta}(\boldsymbol{x})$ is the only assumption preventing (57) to hold as an equality. A difficulty for progressing any further with (57) lies in the fact that Ψ_{mix} is unknown. A possible approach, explored by Smyshlyaev and Willis (1998); Hackl and Heinen (2008) is to replace Ψ_{mix} in (57) with an upper bound obtained from lamination. This results in an *upper bound* Ψ_{eff}^+ on Ψ_{eff} of the form

$$\Psi_{\mathit{eff}}^+(\bar{\boldsymbol{\varepsilon}}) = \inf_{\boldsymbol{\Theta}} \Psi_{\mathit{eff}}^C(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\Theta}) + l(\boldsymbol{\Theta}). \tag{58}$$

The function l considered by Hackl and Heinen (2008) accounts for energy minimization over twin-compatible second-rank laminates (Govindjee et al., 2007). An other possible approach, that we explore in the following, consists in replacing Ψ_{mix} with

the lower bound $\Psi_{Reuss} + f$ in (17). In that case we get an *estimate* of Ψ_{eff}, given by

$$\Psi^*_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}) = \inf_{\boldsymbol{\Theta}} \Psi^C_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\Theta}) + h(\boldsymbol{\Theta}) \tag{59}$$

with

$$h(\boldsymbol{\theta}) = \sum_r c_r f(\boldsymbol{\theta}^r) + \frac{1}{2}\sum_r c_r \boldsymbol{\eta}^r : (\mathbb{L}^{-1} - \mathbb{P}) : \boldsymbol{\eta}^r - \frac{1}{2}\bar{\boldsymbol{\eta}} : (\mathbb{L}^{-1} - \mathbb{P}) : \bar{\boldsymbol{\eta}} \tag{60}$$

and $\bar{\boldsymbol{\eta}} = \sum_r c_r \boldsymbol{\eta}^r$. Expression (60) of h can be rewritten in the more compact fashion

$$h(\boldsymbol{\theta}) = \sum_r c_r f(\boldsymbol{\theta}^r) + \frac{1}{2}\sum_r c_r \boldsymbol{h}_r : (\mathbb{L}^{-1} - \mathbb{P}) : \boldsymbol{h}_r \tag{61}$$

by setting

$$\boldsymbol{h}^r = \boldsymbol{\eta}^r - \bar{\boldsymbol{\eta}}. \tag{62}$$

The overall structure (59) of the estimate Ψ^*_{eff} is similar to that of the upper bound Ψ^+_{eff} in (58). However, the proposed estimate does not rely on any assumption on the type of austenite-martensite microstructures that develop in each grain. There are also differences in the computational complexity of the mixing energies l and h appearing in (58) and (59), respectively. The computational complexity of l grows rapidly with the number of variants n because l is defined by a combinatorial optimization problem that requires to keep track of all possible twin laminates. By contrast, the dimensionality of the optimization problem (14) defining f is equal to 6, independently of the number of variants.

The validity of the estimate Ψ^*_{eff} in (59) can be assessed by checking if it satisfies known bounds on the effective energy Ψ_{eff}. By construction, Ψ^*_{eff} satisfies the upper bounds of Smyshlyaev and Willis (1998); Hackl and Heinen (2008). The comparison with lower bounds is not as direct. It is known that the convex function

$$\Psi^C_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}) = \inf_{\boldsymbol{\Theta}} \Psi^C_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\Theta}) \tag{63}$$

is a lower bound on Ψ_{eff} (Hackl and Heinen, 2008). Definition (14) shows that $f \geq 0$ and it can be verified that $\mathbb{L}^{-1} - \mathbb{P} \succ 0$ (Appendix B). The function h in (61) is thus positive, which ensures that $\Psi^*_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}) \geq \Psi^C_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}})$ for all $\bar{\boldsymbol{\varepsilon}}$. A lower bound that potentially improves on Ψ^C_{eff} has been proposed by Peigney (2009). That bound, henceforth denoted by Ψ^-_{eff}, is defined as

$$\Psi^-_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}) = \inf_{\boldsymbol{\Theta}} \Psi^C_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}, \boldsymbol{\Theta}) + g(\boldsymbol{\Theta}) \tag{64}$$

with

$$g(\boldsymbol{\Theta}) = \sup_{(\boldsymbol{a},\tilde{\mathbb{L}})} \begin{array}{l} \dfrac{1}{4}\displaystyle\sum_{r,s,i,j} c_r c_s \theta_i^r \theta_j^s (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^s) : (\mathbb{L} - \mathbb{L} : (\mathbb{L} - \tilde{\mathbb{K}})^{-1} : \mathbb{L}) : (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^s) \\ + \dfrac{1}{2}\displaystyle\sum_r c_r \boldsymbol{h}^r : \Big(\mathbb{L} - \tilde{\mathbb{K}} + (\mathbb{L} - \tilde{\mathbb{K}}) : \tilde{\mathbb{P}} : (\mathbb{L} - \tilde{\mathbb{K}})\Big)^{-1} : \boldsymbol{h}^r \end{array} \tag{65}$$

where $\tilde{\mathbb{K}} = \mathbb{K}(\boldsymbol{a}) + \tilde{\mathbb{L}}$ and $\tilde{\mathbb{P}}$ is the polarization tensor associated to $\tilde{\mathbb{L}}$, defined as in (56) but replacing $\mathbb{L}$ with $\tilde{\mathbb{L}}$. In (65), the supremum is taken over symmetric second-order tensors $\boldsymbol{a}$ and symmetric fourth-order tensors $\tilde{\mathbb{L}}$ satisfying

$$\boldsymbol{a} \succeq 0, \quad \tilde{\mathbb{L}} \succ 0, \quad \mathbb{L} \succ \tilde{\mathbb{L}} + \mathbb{K}(\boldsymbol{a}). \tag{66}$$

The calculation of g is relatively difficult and it is not obvious from the definitions above that $\Psi^*_{\mathit{eff}} \geq \Psi^-_{\mathit{eff}}$. It can actually be shown that

$$h(\boldsymbol{\Theta}) \geq g(\boldsymbol{\Theta}) \tag{67}$$

for all $\boldsymbol{\Theta} \in \mathcal{T}_n^N$, which from (59) and (64) ensures that Ψ^*_{eff} respects the lower bound Ψ^-_{eff}. The proof of the inequality (67) is relatively technical and therefore reported in Appendix C not to obscure the presentation. We only mention here that (67) crucially relies on some properties of the $\mathbb{P}$ tensor in relation with tensors $(\boldsymbol{a}, \mathbb{L})$ satisfying (66). Those properties are proved as preliminaries in Appendix B. In summary we have

$$\Psi^-_{\mathit{eff}} \leq \Psi^*_{\mathit{eff}} \leq \Psi^+_{\mathit{eff}} \tag{68}$$

i.e the proposed estimate satisfies known lower and upper bounds that depend on the same microstructural information. Inequality (67) plays an essential role in (68). In the following, properties (67) and (68) are illustrated on several examples of increasing complexity. Those examples also provide some insight on the behavior of the energy estimate Ψ^*_{eff} and its mechanicals implications.

Remark: The obtained energy estimate (59) only depends on the texture through the volume fractions of the crystalline orientations and can thus be expressed in terms of the Orientation Distribution Function (ODF) p of the texture as

$$\Psi^*_{\mathit{eff}}(\boldsymbol{\varepsilon}) = \min_{\boldsymbol{\theta}:\mathrm{SO}(3)\mapsto\mathcal{T}_n} \Psi^C_{\mathit{eff}}(\boldsymbol{\varepsilon}, \boldsymbol{\theta}) + h(\boldsymbol{\theta}) \tag{69}$$

where

$$\Psi^C_{\mathit{eff}}(\boldsymbol{\varepsilon},\boldsymbol{\theta}) = \frac{1}{2}(\boldsymbol{\varepsilon}-\bar{\boldsymbol{\tau}}(\boldsymbol{\theta})) : \mathbb{L} : (\boldsymbol{\varepsilon}-\bar{\boldsymbol{\tau}}(\boldsymbol{\theta})) + \int_{\boldsymbol{r}\in \mathrm{SO}(3)} \sum_{i=1}^{n} \theta_i(\boldsymbol{r}) w_i \, p(\boldsymbol{r}) d\mu \tag{70}$$

and

$$\begin{aligned} h(\boldsymbol{\theta}) = \; & \int_{\boldsymbol{r}\in \mathrm{SO}(3)} f(\boldsymbol{\theta}(\boldsymbol{r})) p(\boldsymbol{r}) d\mu \\ & + \frac{1}{2} \int_{\boldsymbol{r}\in \mathrm{SO}(3)} \boldsymbol{\tau}(\boldsymbol{r}) : \mathbb{L} : (\mathbb{L}^{-1} - \mathbb{P}) : \mathbb{L} : \boldsymbol{\tau}(\boldsymbol{r}) \, p(\boldsymbol{r}) d\mu \\ & - \frac{1}{2} \bar{\boldsymbol{\tau}}(\boldsymbol{\theta}) : \mathbb{L} : (\mathbb{L}^{-1} - \mathbb{P}) : \mathbb{L} : \bar{\boldsymbol{\tau}}(\boldsymbol{\theta}). \end{aligned} \tag{71}$$

The term $d\mu$ in (70) and (71) is the Haar measure (Bunge, 2013). The strains $\bar{\boldsymbol{\tau}}(\boldsymbol{\theta})$ in (70) and $\boldsymbol{\tau}(\boldsymbol{r})$ in (71) are defined as $\boldsymbol{\tau}(\boldsymbol{r}) = \sum_{i=1}^{n} \theta_i(\boldsymbol{r}) \boldsymbol{r}\boldsymbol{\tau}_i\boldsymbol{r}^T$ and $\bar{\boldsymbol{\tau}}(\boldsymbol{\theta}) = \int_{\boldsymbol{r}\in \mathrm{SO}(3)} \boldsymbol{\tau}(\boldsymbol{r}) p(\boldsymbol{r}) d\mu$. Expression (69) is the starting point of the companion paper devoted to strain localization in SMAs.

4. Analytical example

We first consider a polycrystal with 2 orientations. We assume that only one martensitic variant (variant 1) is active in orientation 1 and that orientation 2 remains purely austenitic. The rotation $\boldsymbol{R}^1$ defining orientation 1 is equal to the identity. Let θ be the volume fraction of the active martensitic variant in orientation 1. The energy estimate (59) specializes as

$$\Psi^*_{\mathit{eff}}(\boldsymbol{\varepsilon}) = \inf_{0\le\theta\le 1} \frac{1}{2}(\boldsymbol{\varepsilon} - c_1\theta\boldsymbol{\tau}_1) : \mathbb{L} : (\boldsymbol{\varepsilon} - c_1\theta\boldsymbol{\tau}_1) + c_1\theta w_1 + h(\theta) \tag{72}$$

where

$$h(\theta) = \frac{1}{2} c_1\theta(1-\theta) m_- + \frac{1}{2} c_1(1-c_1)\theta^2(\boldsymbol{\tau}_1 : \mathbb{L} : \boldsymbol{\tau}_1 - \kappa) \tag{73}$$

with $\kappa = \boldsymbol{\tau}_1 : \mathbb{L} : \mathbb{P} : \mathbb{L} : \boldsymbol{\tau}_1$ and m_- defined as in (23). The function h in (73) being quadratic in θ, there is no difficulty in solving the minimization problem defining $\Psi^*_{\mathit{eff}}(\boldsymbol{\varepsilon})$ in (72). For values of $\boldsymbol{\varepsilon}$ such that $\boldsymbol{\varepsilon} : \mathbb{L} : \boldsymbol{\tau}_1$ falls between $w_1 + \frac{1}{2}m_-$ and $w_1 + \boldsymbol{\tau}_1 : \mathbb{L} : \boldsymbol{\tau}_1 - \frac{1}{2}m_- - (1-c_1)\kappa$, we obtain

$$\Psi^*_{\mathit{eff}}(\boldsymbol{\varepsilon}) = \frac{1}{2}\boldsymbol{\varepsilon} : \mathbb{L} : \boldsymbol{\varepsilon} - \frac{1}{2} c_1 \frac{(\boldsymbol{\varepsilon} : \mathbb{L} : \boldsymbol{\tau}_1 - \frac{m_-}{2} - w_1)^2}{\boldsymbol{\tau}_1 : \mathbb{L} : \boldsymbol{\tau}_1 - m_- - (1-c_1)\kappa}. \tag{74}$$

Expression (74) corresponds to the situation where the minimizing volume fraction θ in (72) is strictly between 0 and 1, i.e. phase transformation in orientation 1 is not

complete. To simplify the discussion, we assume from now on that $\boldsymbol{\tau}_1$ is deviatoric and the texture is isotropic, in which case

$$\frac{m_-}{E} = \lambda_2^2, \;\; \frac{\kappa}{E} = \frac{2(5\nu-4)}{15(\nu^2-1)}\|\boldsymbol{\tau}_1\|^2. \tag{75}$$

Deviatoric tensors indeed verify condition (25), so that expression (33) of m_- applies. Denoting the second-order and the fourth-order identity tensors by $\boldsymbol{I}$ and $\mathbb{I}$ respectively, the $\mathbb{P}$ tensor for an isotropic texture is given by (Castaneda and Suquet, 1997)

$$\mathbb{P} = \frac{1}{15E}\frac{\nu+1}{\nu-1}(\boldsymbol{I}\otimes\boldsymbol{I} + 2(5\nu-4)\mathbb{I})$$

from which the expression of κ in (75) follows.

For the example at hand, the energy estimate Ψ^*_{eff} coincide with the upper bound Ψ^+_{eff} in (58). It is known indeed that the relaxed energy for the two-well problem is attained by rank-1 laminates (Kohn, 1991). The lower bound (64) takes the form

$$\Psi^-_{\mathit{eff}}(\bar{\boldsymbol{\varepsilon}}) = \inf_{0\le\theta\le 1} \frac{1}{2}(\bar{\boldsymbol{\varepsilon}} - c_1\theta\boldsymbol{\tau}_1) : \mathbb{L} : (\bar{\boldsymbol{\varepsilon}} - c_1\theta\boldsymbol{\tau}_1) + c_1\theta w_1 + g(\theta) \tag{76}$$

with

$$\begin{aligned} g(\theta) = \sup_{(\boldsymbol{a},\tilde{\mathbb{L}})} \;& \frac{1}{2}c_1\theta(1-c_1\theta)\boldsymbol{\tau}_1 : (\mathbb{L} - \mathbb{L} : (\mathbb{L}-\tilde{\mathbb{K}})^{-1} : \mathbb{L}) : \boldsymbol{\tau}_1 \\ & + \tfrac{1}{2}c_1(1-c_1)\theta^2 : \boldsymbol{\tau}_1 : \mathbb{L} : (\mathbb{L} - \tilde{\mathbb{K}} + (\mathbb{L}-\tilde{\mathbb{K}}) : \tilde{\mathbb{P}} : (\mathbb{L}-\tilde{\mathbb{K}}))^{-1} : \mathbb{L} : \boldsymbol{\tau}_1. \end{aligned} \tag{77}$$

The function g in (77) is clearly more difficult to evaluate than the function h. Closed-form expression can be obtained by restricting the minimization in (77) to isotropic tensors $\tilde{\mathbb{L}}$ and tensors $\boldsymbol{a}$ of the form $a\boldsymbol{u}_3\otimes\boldsymbol{u}_3$ where $\boldsymbol{u}_3$ is an eigenvector of $\boldsymbol{\tau}_1$ for its maximum eigenvalue λ_3. This special choice of $\boldsymbol{a}$ is motivated by the fact that it corresponds to the solution of the two-well problem recovered by setting $c_1 = 1$ in (77), as discussed in Sect. 2.1. To further simplify the calculations, we consider the case of incompressible elasticity ($\nu = 1/2$) and choose accordingly incompressible tensors $\tilde{\mathbb{L}}$. The corresponding tensors $(\boldsymbol{a},\tilde{\mathbb{L}})$ are thus parameterized by the 2 scalars $(\tilde{E}, a)$ where $\tilde{E}$ is the Young's modulus of $\tilde{\mathbb{L}}$. The restrictions (66) can be verified to reduce to

$$0 < \tilde{E} < E, \quad 0 \le a < \frac{2}{3}(E-\tilde{E}). \tag{78}$$

Setting

$$r = \frac{\tilde{E}}{E}, \quad s = 3\frac{a}{E},$$

we obtain

$$g(\theta) = \sup_{(r,s)} \frac{E}{3} c_1 \theta \left((1 - c_1\theta)(\|\boldsymbol{\tau}_1\|^2 - p(r,s)) + c_1(1-c_1)\theta q(r,s) \right) \tag{79}$$

where

$$p(r,s) = 4\frac{3(r-1)\|\boldsymbol{\tau}_1\|^2 + s(\|\boldsymbol{\tau}_1\|^2 + 3\lambda_1\lambda_2)}{(s+2r-2)(s+6-6r)},$$

$$q(r,s) = 20r\frac{(9(1-r)(2+3r) - 3s(4+r) + 7/2s^2)\|\boldsymbol{\tau}_1\|^2 + 3s\lambda_1\lambda_2(2s - 3(4+r))}{(s-3r-2)(s+2r-2)(18(1-r)(2+3r) + 3s(r+4) + s^2)}.$$

The supremum in (79) is taken over values (r,s) satisfying $0 < r < 1$ and $0 \leq s < 2(r-1)$. Observe that the function g in (79) is an isotropic function of $\boldsymbol{\tau}_1$ as it only depends on $\boldsymbol{\tau}_1$ through its eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$. Moreover, the function g is homogeneous of degree 2 in $\boldsymbol{\tau}_1$, i.e. changing $\boldsymbol{\tau}_1$ to $x\boldsymbol{\tau}_1$ has the effect of changing $g(\theta)$ to $x^2 g(\theta)$. Similar remarks apply to the function h in (73). Due to those invariance properties, we can assume without loss of generality that $\|\boldsymbol{\tau}_1\|^2 = 3/2$. This allows one to write the eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ of the deviatoric tensor $\boldsymbol{\tau}_1$ as

$$(\lambda_1, \lambda_2, \lambda_3) = (\cos(\alpha + \frac{2\pi}{3}), \cos(\alpha + \frac{4\pi}{3}), \cos\alpha) \tag{80}$$

where $\alpha \in [0, \pi/6]$ is the angle introduced by Kachanov (2004). In the following we examine the two limiting situations $\alpha = \pi/6$ and $\alpha = 0$. The case $\alpha = \pi/6$ corresponds to $\lambda_2 = 0$, i.e. to the situation where austenite and martensite are geometrically compatible. By contrast, the case $\alpha = 0$ corresponds to the situation where λ_2^2 is maximized, i.e the incompatibility between austenite and martensite is maximized.

4.1. Case $\alpha = \pi/6$

The case $\alpha = \pi/6$ corresponds to $(\lambda_1, \lambda_2, \lambda_3) = (-\frac{\sqrt{3}}{2}, 0, \frac{\sqrt{3}}{2})$. Carrying out the optimization with respect to (r,s) in (79), the final expression for g is

$$\frac{g(\theta)}{E} = \begin{cases} 0 & \text{for } \theta \leq \dfrac{2}{5-3c_1}, \\ c_1\dfrac{\theta}{60}\dfrac{(5\Delta - 6(1-c_1\theta))^2}{3 - 2c_1\theta - \theta - 2\Delta} & \text{for } \theta \geq \dfrac{2}{5-3c_1}. \end{cases} \tag{81}$$

where $\Delta = \sqrt{6c_2(1-\theta)\theta}$. The functions g in (81) and h in (73) are plotted in Fig. 2 for the case $c_1 = 0.6$. We can observe that $g(\theta) \leq h(\theta)$ for all θ, which implies that

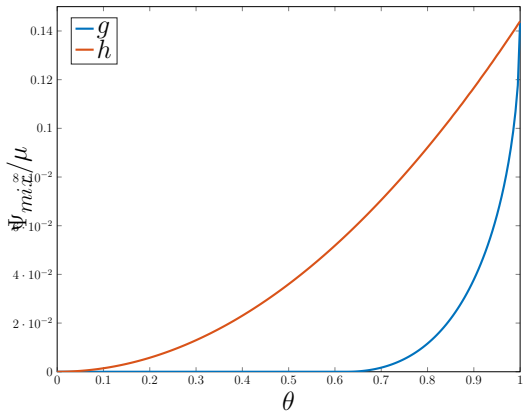

Figure 2: Functions h and g in the case $\alpha = \pi/6$, $c_1 = 0.6$.

the estimate (72) of the free energy satisfies the lower bound (76) for all values of the strain $\boldsymbol{\varepsilon}$.

In Fig. 3(left, red curve) are shown the values taken the estimate $\Psi^*_{\mathit{eff}}(\boldsymbol{\varepsilon})$ in the direction $\boldsymbol{\tau}_1$, i.e. for $\boldsymbol{\varepsilon}$ of the form $t\boldsymbol{\tau}_1$. We have set $w_1 = 0$ and $c_1 = 0.6$. The values taken by the convex bound Ψ^C_{eff} and the lower bound Ψ^-_{eff} are shown in green and blue, respectively. In Fig. 3(right, red curve) is shown the volume fraction θ^* reaching the minimum in (72) for $\boldsymbol{\varepsilon} = t\boldsymbol{\tau}_1$. The minimizing volume fractions for the optimization problems (63) and (76) defining $\Psi^C_{\mathit{eff}}(t\boldsymbol{\tau}_1)$ and $\Psi^-_{\mathit{eff}}(t\boldsymbol{\tau}_1)$ are denoted by θ^C and θ^-, respectively. They are shown as green and blue curves in Fig. 3(right). The convex bound Ψ^C_{eff} is piecewise quadratic and the corresponding volume fraction θ^C grows with t in a piecewise linear fashion as can be observed in Fig. 3. For $t \geq 2c_1/(5 - 3c_1)$, we have $\theta^- \leq \theta^C$ i.e. Ψ^-_{eff} requires a larger strain than Ψ^C_{eff} for reaching the same level of phase transformation. This hardening effect results from the elastic interaction between grains in orientation 1 and orientation 2 which is captured in Ψ^-_{eff} (and Ψ^*_{eff}) but not taken into account in Ψ^C_{eff}. As can be observed in Fig. 3(right), the hardening effect in Ψ^-_{eff} is such that full transformation into martensite is achieved only in the limit $t \to +\infty$. We can also observe that the hardening vanishes (i.e. $\theta^- = \theta^C$) for $0 < t \leq 2c_1/(5 - 3c_1)$. In that range of values for t, the lower bound Ψ^-_{eff} surely underestimates the true effective energy Ψ_{eff} as the elastic interaction between orientation 1 and 2 is still present and contributes to increase the energy compared to the convex envelope Ψ^C_{eff}. The proposed energy

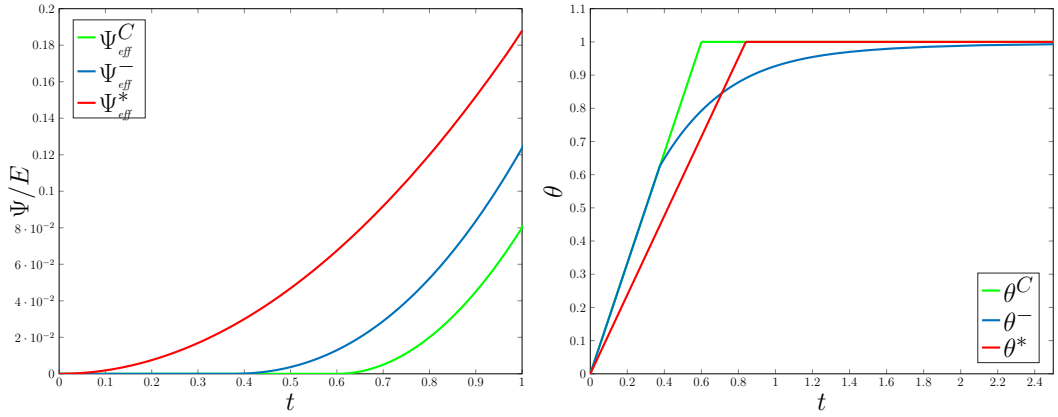

Figure 3: (left) Values of the energy functions for a strain of the form $t\boldsymbol{\tau}_1$. Case $\alpha = \pi/6$, $c_1 = 0.6$

estimate Ψ^*_{eff} is free from such shortcoming: We have indeed $\Psi^*_{eff} > \Psi^C_{eff}$ and $\theta^* < \theta^C$ even for low values of t, as can be observed in Fig. 3. The behaviors of Ψ^-_{eff} and Ψ^*_{eff} also differ for high values of t: In contrast with the lower bound Ψ^-_{eff}, the estimate Ψ^*_{eff} allows for full transformation into martensite to be achieved for a finite value of the strain, see Fig. 3(right).

4.2. Case $\alpha = 0$

The case $\alpha = 0$ corresponds to $(\lambda_1, \lambda_2, \lambda_3) = (-\frac{1}{2}, -\frac{1}{2}, 1)$. Carrying the optimization with respect to (r, s) in (79) gives

$$\frac{g(\theta)}{E} = \begin{cases} \dfrac{1}{8}(1 - c_1\theta)c_1\theta & \text{for } \theta \leq \dfrac{8}{15 - 7c_1} \\ c_1\theta \dfrac{96c_2^2\theta^2 + 412c_2\theta(1-\theta) + 91(1-\theta)^2 - 80D(1 - c_1\theta)}{40(7 - 4D + (1 - 8c_1)\theta)} & \text{for } \theta \geq \dfrac{8}{15 - 7c_1} \end{cases} \tag{82}$$

with $D = \sqrt{14c_2(1-\theta)\theta}$. The function g given by (82) is plotted in Fig. 4 (blue curve) for the cases $c_1 = 0.25$ and $c_1 = 0.75$. The function h given by (73) is shown in red. We can observe that $g \leq h$ in accordance with (67). As illustrated in Fig. 4 , the function h is convex in θ for low values of c_1 and concave for high values of c_1 [3]. Whereas h is always monotonically increasing, the function g corresponding to

[3] it can be calculated that the transition from convexity to concavity occurs at $c_1 = 7/12 \simeq 0.58$

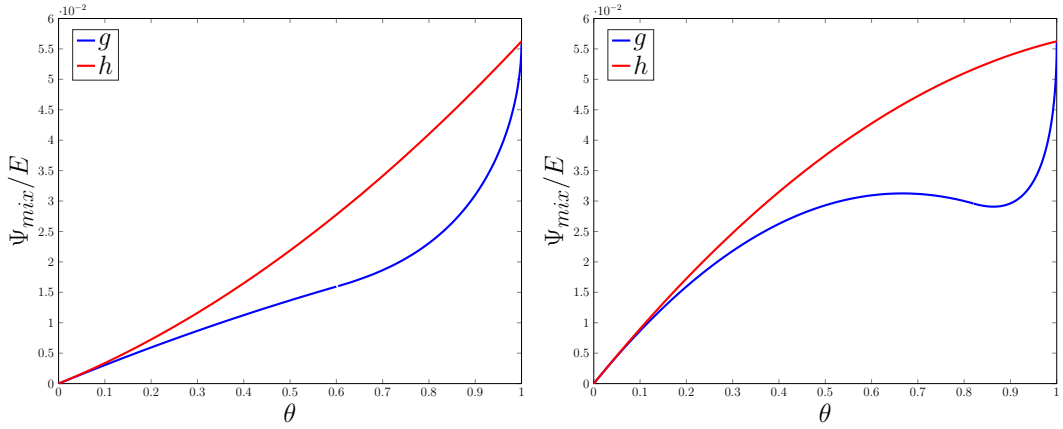

Figure 4: Functions h and g in the case $\alpha = 0$ with $c_1 = 0.25$ (left), $c_1 = 0.75$ (right).

the lower bound has a up-down-up behavior for $c_1 \geq 15/23 \simeq 0.65$, as illustrated in Fig. 4 (right). It can further be noted that h and g are equal at $\theta = 1$ (their common value is $3Ec_1c_2/10$) and coincide at the first order at $\theta = 0$ (their common first-order expansion is $Ec_1\theta/8$). That last property implies that h actually gives the exact value of $\Psi_{mix} - \Psi_{Reuss}$ at the first order in θ. We have indeed the general property $g \leq \Psi_{mix} - \Psi_{Reuss} \leq l$ with $h = l$ for the example at hand. Consequently we can state that $\Psi_{mix} - \Psi_{Reuss} = 3Ec_1c_2\theta/10$ at the first order in θ. A similar remark actually applies in the case $\alpha = 0$, giving $\Psi_{mix} - \Psi_{Reuss} = 0$ at the first order in θ.

In Fig. 5(left) are shown the values taken by the energy functions in the direction $\boldsymbol{\tau}_1$. The minimizing volume fractions are shown in Fig. 5(right). The results in Fig. 5 have been obtained using the values $c_1 = 0.75$ and $w_1 = 0$. Compared to the case $\alpha = \pi/6$ studied previously, a first observation is that $\Psi^-_{\mathit{eff}}(t\boldsymbol{\tau}_1) > 0$ for low values of t. This results from the incompatibility between austenite and martensite: In the case $\alpha = 0$, austenite and martensite do not satisfy the Hadamard conditions (11) so that the formation of a microstructures entails some elastic energy cost. The impact of compatibility is also reflected in Fig. 5(right): For θ^* to be strictly positive, t needs to be larger than some threshold value t^*, equal to $1/8$ and independent of c_1. The corresponding energy is the energy barrier that needs to be overcome for martensite to appear. For large values of t, the behavior of θ^- and θ^* is similar to the case $\alpha = \pi/6$ depicted in Fig. 3: Whereas θ^- only reaches the value 1 in the limit $t \to \infty$, θ^* is equal to 1 as soon as t reaches a finite threshold.

We close the study of this example with some comments on the tangent stifness

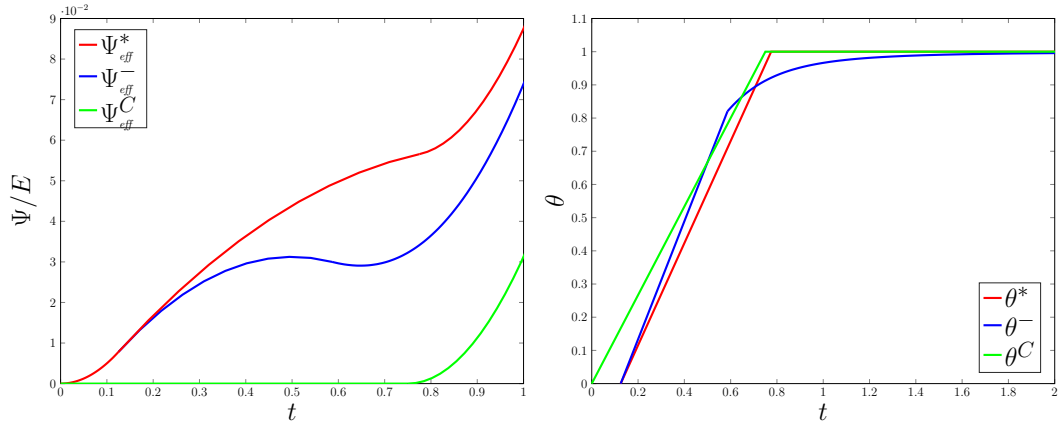

Figure 5: (left) Values of energy functions in the direction $\boldsymbol{\tau}_1$. Case $\alpha = 0$, $c_1 = 0.75$.

$\mathbb{L}^* = d^2\Psi^*_{eff}/d\boldsymbol{\varepsilon}^2$ of the energy estimate Ψ^*_{eff}. Expression (74) yields

$$\mathbb{L}^* = \mathbb{L} - c_1 \frac{(\mathbb{L} : \boldsymbol{\tau}_1) \otimes (\mathbb{L} : \boldsymbol{\tau}_1)}{\boldsymbol{\tau}_1 : \mathbb{L} : \boldsymbol{\tau}_1 - m_- - (1 - c_1)\kappa}. \tag{83}$$

It can be verified from (83) that the minimum value taken by $\boldsymbol{n} : \mathbb{L}^* : \boldsymbol{n}$ over unit second-order tensors $\boldsymbol{n}$ is attained for $\boldsymbol{n}$ parallel to $\boldsymbol{\tau}_1$. We have

$$\boldsymbol{\tau}_1 : \mathbb{L}^* : \boldsymbol{\tau}_1 = \frac{-m_- + (1 - c_1)(\boldsymbol{\tau}_1 : \mathbb{L} : \boldsymbol{\tau}_1 - n)}{\boldsymbol{\tau}_1 : \mathbb{L} : \boldsymbol{\tau}_1 - m_- - (1 - c_1)\kappa} \boldsymbol{\tau}_1 : \mathbb{L} : \boldsymbol{\tau}_1. \tag{84}$$

For the case $\alpha = \pi/6$ (i.e. when austenite and martensite are compatible), Eq. (84) gives $\boldsymbol{\tau}_1 : \mathbb{L}^* : \boldsymbol{\tau}_1 > 0$ and therefore $\mathbb{L}^*$ is positive definite. Now for $\alpha = 0$, Eq (84) yields

$$\boldsymbol{\tau}_1 : \mathbb{L}^* : \boldsymbol{\tau}_1 = 3E \frac{9 - 10\nu + 5\nu^2 - 2c_1(7 - 5\nu)}{2(1 + \nu)(9 - 10\nu + 5\nu^2 + 4c_1(4 - 5\nu))}.$$

The denominator of the expression above is positive, hence $\boldsymbol{\tau}_1 : \mathbb{L}^* : \boldsymbol{\tau}_1 < 0$ for $c_1 > (9 - 10\nu + 5\nu^2)/2(7 - 5\nu)$ ($\simeq 0.58$ for $\nu = 0.3$). It follows that the tangent stiffness $\mathbb{L}^*$ is *not* positive definite for high values of c_1. In that case, the energy estimate may give rise to material instabilities, which are observed indeed in SMAs. The formation of Lüders-type bands in NiTi specimen under tension is a well documented example of such instabilities (Shaw and Kyriakides, 1997; Daly et al., 2007; Churchill et al., 2009). In the framework of the present estimate, material instability is strongly conditioned to the austenite-martensite compatibility.

5. Tetragonal to orthorombic transformation in a polycrystal

We consider the tetragonal to orthorhombic transformation in a polycrystal with two isotropically distributed orientations. The rotation $\boldsymbol{R}^1$ in orientation 1 is set to identity. As in Sect. 4 we assume that grains in orientation 2 remain purely austenitic. In that case, the function h in (60) only depends of the martensitic volume fractions $\boldsymbol{\theta} = (\theta_1, \theta_2)$ in orientation 1 and we have

$$h = c_1 f_{\text{T-O}}(\boldsymbol{\theta}) + c_1 c_2 (\theta_1 \boldsymbol{\tau}_1 + \theta_2 \boldsymbol{\tau}_2) : (\mathbb{L} - \mathbb{L} : \mathbb{P} : \mathbb{L}) : (\theta_1 \boldsymbol{\tau}_1 + \theta_2 \boldsymbol{\tau}_2) \tag{85}$$

where $f_{\text{T-O}}(\boldsymbol{\theta})$ corresponds to the solution of the three-well relaxation problem for the tetragonal to orthorhombic transformation, as studied in Sect. 2.2. To fix ideas, we consider the case

$$\nu = \frac{1}{2}, \boldsymbol{\tau}_1 = \eta \, \text{diag}(-1, 2, -1), \ \boldsymbol{\tau}_2 = \eta \, \text{diag}(2, -1, -1)$$

for some η, so that

$$h = c_1 f_{\text{T-O}}(\boldsymbol{\theta}) + \frac{6}{5} E \eta^2 c_1 c_2 (\theta_1^2 + \theta_2^2 - \theta_1 \theta_2). \tag{86}$$

The term $f_{\text{T-O}}(\boldsymbol{\theta})$ in (86) is known explicitly as a special case of the expressions obtained in Sect. 2.2. The function h is plotted in Fig. 6 for $\eta = 1$ along with the functions g and l defining the lower and the upper bounds on the effective energy in (58) and (64). The function l can be calculated explicitly from the expressions given by Hackl and Heinen (2008) and is given by

$$\begin{aligned} \frac{l}{E\eta^2} = \ & c_1 \left(2(\theta_1(1-\theta_1) + \theta_2(1-\theta_2) + 2\theta_1\theta_2) - 6\tfrac{\theta_1\theta_2}{\theta_1+\theta_2} - \tfrac{3}{2}\theta_0(1-\theta_0) \left(|\tfrac{\theta_1}{\theta_1+\theta_2} - \tfrac{1}{2}| + \tfrac{1}{2} \right)^2 \right) \\ & + \frac{6}{5} c_1 c_2 (\theta_1^2 + \theta_2^2 - \theta_1\theta_2). \end{aligned} \tag{87}$$

The function g is obtained by solving the optimization problem (65) numerically. Each subfigure in Fig. 6 shows the values taken by the energy functions as a function of θ_1 for prescribed values of the austenitic volume fraction θ_0 and volume fraction c_1 of orientation 1. The subfigures on the left correspond to $c_1 = 0.2$ and those on the right correspond to $c_1 = 0.8$. The value of θ_0 decreases from 0.9 to 0.3 from top to bottom. For a given value of c_1, the plots in Fig. 6 show that h behaves as g for high values of θ_0 and as l for low values of θ_0. For a given value of θ_0, h becomes increasingly closer to g as c_1 increases. In the limit $c_1 \to 1$, it can observed from (86) that h recovers the exact expression $f_{\text{T-O}}$ of the relaxed energy.

This not the case for the function l corresponding to the upper bound. We have for instance $l(\theta,\theta) \to \theta(1-2\theta)/4$ as $c_1 \to 1$, whereas we know from Sect. 2.2 that $(\Psi_{mix} - \Psi_{Reuss})(\theta,\theta) = 0$. The reason is that second-rank lamination – as considered implicitly in the function l – is not sufficient to achieve the relaxed energy.

In Fig. 7(left) are shown the values taken by $\Psi^*_{\mathit{eff}}(\boldsymbol{\varepsilon})$ (red curve) and $\Psi^-_{\mathit{eff}}(\boldsymbol{\varepsilon})$ (blue curve) in the direction $\boldsymbol{\varepsilon} = t(0.6\boldsymbol{\tau}_1 + 0.4\boldsymbol{\tau}_2), t \geq 0$. The convex bound Ψ^C_{eff} is shown in green. The volume fraction c_1 and the ratio w_i/E $(i \geq 1)$ have been set to 0.2 and $4/45000$ respectively. Some insight on the behavior of the energy estimate Ψ^*_{eff} can be gained by looking at the minimizing volume fractions of martensite, see Fig. 7(right). Since a strictly positive value of w_1 has been used, martensite becomes energetically favorable only when t exceeds a certain transformation threshold t_M. For the convex envelope, that threshold is equal to $w_1/(2\mu\boldsymbol{\tau}_1 : (0.8\boldsymbol{\tau}_1 + 0.2\boldsymbol{\tau}_2)) \simeq 0.55$. Beyond that transformation threshold, the minimizing volume fractions $\boldsymbol{\theta}^C$ for the convex bound only involve variant 1 which is the most favorably oriented with respect to the applied strain $\bar{\boldsymbol{\varepsilon}}$. By contrast, the minimizing volume fractions $\boldsymbol{\theta}^* = (\theta^*_1, \theta^*_2)$ corresponding to the energy estimate Ψ^*_{eff} involve both variants. In a first stage, θ^*_1 and θ^*_2 grow linearly with t until $\theta^*_1 + \theta^*_2$ reaches the limit value 1, indicating that austenite is fully transformed into martensite. For larger values of t, θ^*_1 continues to grow while θ^*_2 decreases, the sum $\theta^*_1 + \theta^*_2$ being equal to 1. This corresponds to the reorientation of variant 2 into variant 1. The minimizing volume fractions $\boldsymbol{\theta}^-$ corresponding to the lower bound Ψ^-_{eff} show a similar behavior, with the difference that full transformation into martensite is reached sooner. The behavior of $\boldsymbol{\theta}^*$ for small values of t above t_M is explained by the fact that austenite and martensite are not compatible in the example considered. Thus, microstructures mixing austenite and a single variant of martensite – as predicted by the convex bound – have a high energy cost $f(\boldsymbol{\theta})$. Since the two martensitic variants are compatible in the example at hand, low values of $f(\boldsymbol{\theta})$ can be achieved by mixing the three phases as discussed in Sect. 2.2. For low values of t above t_M, the minimizing volume fractions $\boldsymbol{\theta}^*$ result from a trade off between the microstructural energy $f(\boldsymbol{\theta})$ – which favors the presence of both variants – and the convex term $\Psi^C_{\mathit{eff}}(\boldsymbol{\varepsilon}, \boldsymbol{\theta})$ – which favors variant 1. The situation changes when $\theta^*_1 + \theta^*_2$ reaches the value 1, which occurs for large values of t. In that case, the compatibility of the variants implies that $f(\boldsymbol{\theta}) = 0$ for any value of $\boldsymbol{\theta} = (\theta_1, \theta_2)$ such that $\theta_1 + \theta_2 = 1$. Hence the energy term f does not play a role anymore. For high values of t, the value of $\boldsymbol{\theta}^*$ strikes a balance between the inter-grain interaction energy and the convex term $\Psi^C_{\mathit{eff}}(\boldsymbol{\varepsilon}, \boldsymbol{\theta})$, in a way similar to the example considered in Sect. 4.1. As in Sect. 4.1, θ^*_1 reaches the limit value 1 for a finite value of t whereas θ^-_1 only reaches 1 in the limit $t \to +\infty$. For the case

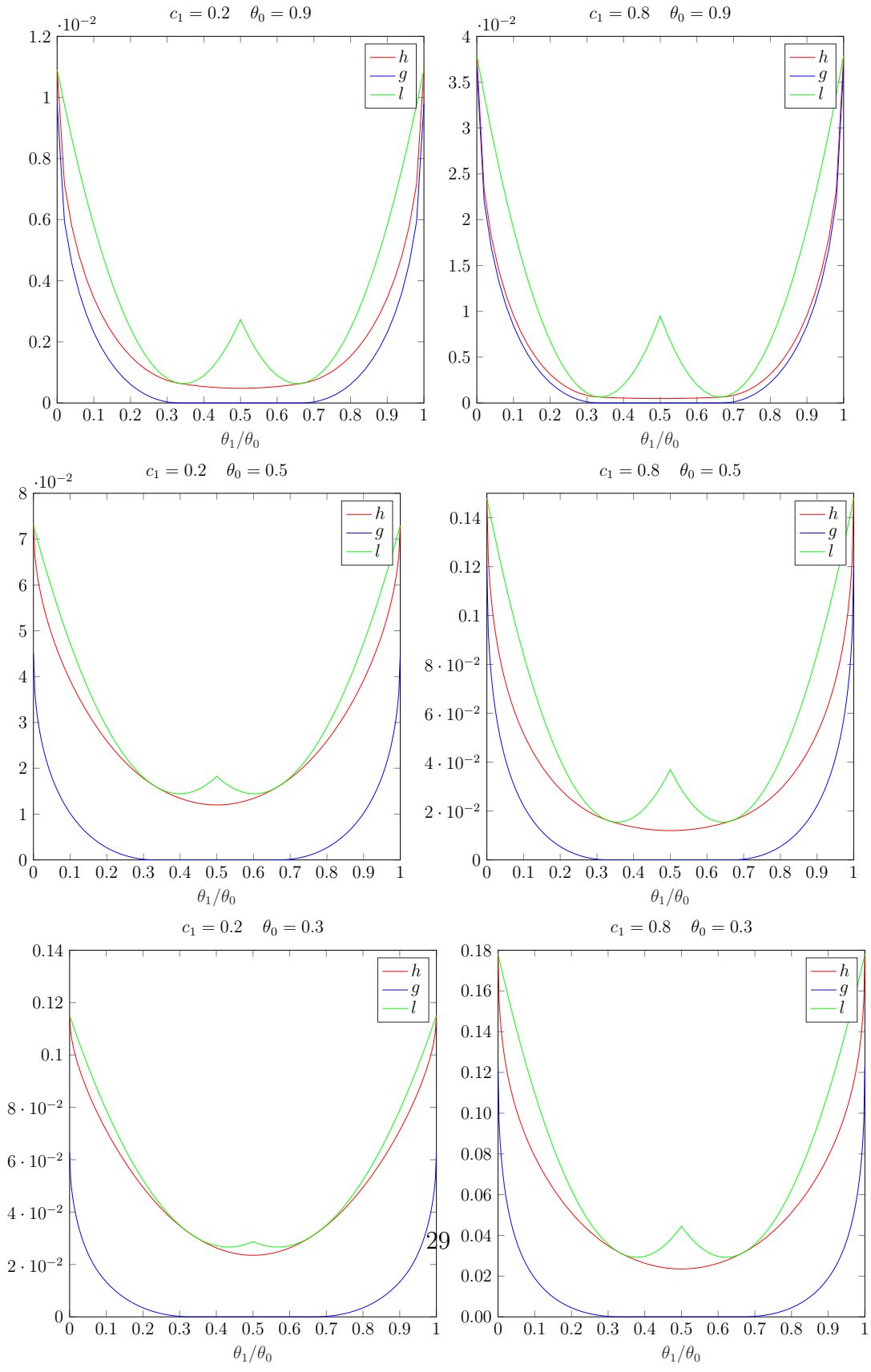

Figure 6: Bounds and estimate on the mixing energy for a three-well/two-orientation problem. Values of the energy are normalized with respect to the Young's modulus E.

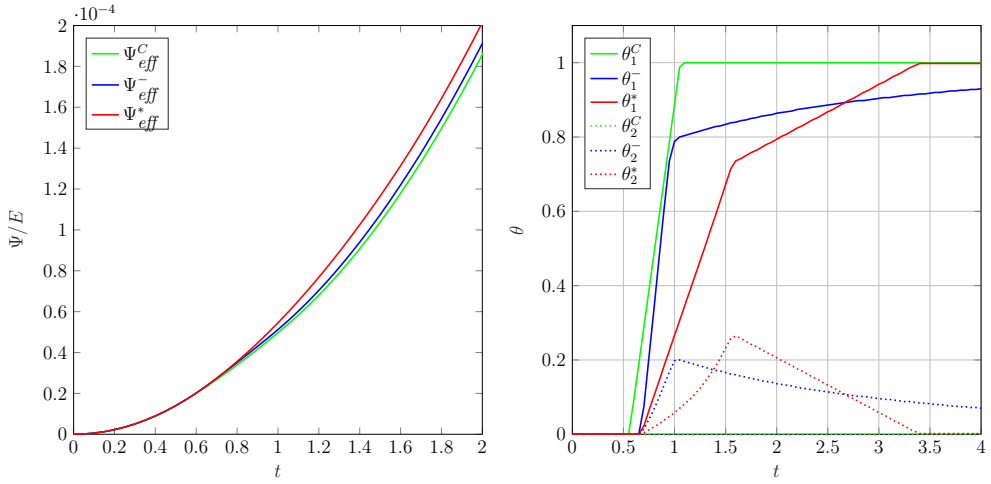

Figure 7: (left) Energy functions in the direction $0.6\boldsymbol{\tau}^1+0.4\boldsymbol{\tau}^2$, (right) Minimizing volume fractions. Case $w_i/E = 4/45000$.

depicted in Fig. 7, the energy estimate Ψ^*_{eff} is found to be really close to the upper bound Ψ^+_{eff}. In Fig. 8 is shown a situation where differences between Ψ^*_{eff} and Ψ^+_{eff} are noticeable. The plots in Fig. 8 correspond to $c_1 = 0.8$ and $w_1 = w_2 = 0$, i.e. to a temperature T close to the transformation temperature T_0.

We close this Section by comparing the stress component $\sigma_{11} = \boldsymbol{u}_1 \cdot \boldsymbol{\sigma} \cdot \boldsymbol{u}_1$ calculated from the different energy functions considered (Fig. 9). The stress calculated from the energy estimate Ψ^*_{eff} is

$$\frac{d\Psi^*_{eff}}{d\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}) = \mathbb{L} : (\boldsymbol{\varepsilon} - c_1\theta^*_1\boldsymbol{\tau}_1 - c_1\theta^*_2\boldsymbol{\tau}_2). \tag{88}$$

The stress $d\Psi^-_{eff}/d\boldsymbol{\varepsilon}$ calculated from the lower bound Ψ^-_{eff} is obtained by replacing (θ^*_1, θ^*_2) with (θ^-_1, θ^-_2) in (88). The stress calculated from the convex bound Ψ^C_{eff} is obtained similarly by replacing (θ^*_1, θ^*_2) with (θ^C_1, θ^C_2). Even though the values of the energy functions shown in Fig. 7(left) are close to one another, the calculated stresses show significant differences as illustrated in Fig. 9(left). This boils down to the fact that a small distance between two functions F and G does not warrant a small distance between their derivatives F' and G'. In the same spirit, the inequality $F \leq G$ does not imply that $F' \leq G'$ and the stress calculated from the convex envelope is not a lower bound on the stress calculated from Ψ^*_{eff} as can be seen in Fig. 9. Similarly, for the case depicted in Fig. 8, there is a large gap between the

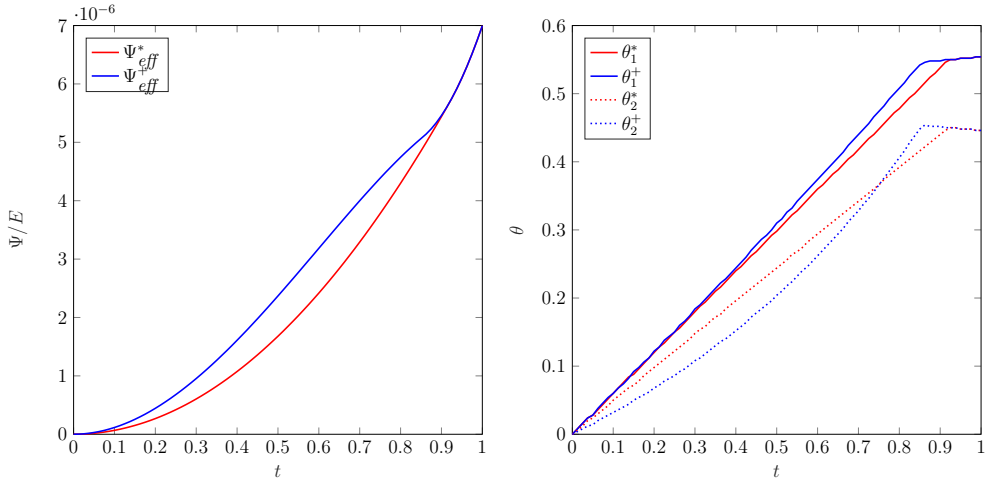

Figure 8: (left) Energy functions in the direction $0.55\boldsymbol{\tau}^1 + 0.45\boldsymbol{\tau}^2$, (right) Minimizing volume fractions. Case $w_i = 0$.

stresses calculated from the energy estimate and the upper bound as shown in Fig. 9(right).

6. Uniaxial tension in γ_1'CuAlNi

We now consider an example related to γ_1'CuAlNi, for which there are 6 martensitic variants with transformations strains listed in Table A.2. That example is used to illustrate the selection of variants operated by the energy estimate. We consider a polycrystal with 2 orientations defined by

$$\boldsymbol{R}^1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{R}^2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}. \tag{89}$$

The orientations in (89) correspond to the Eucken-Hirsh texture observed in rolled SMA ribbons (Eucken and Hirsch, 1990): the normal $\boldsymbol{u}_3$ to the ribbon is along the $\langle 001 \rangle$ crystallographic direction of the austenitic cubic lattice and the rolling direction $\boldsymbol{u}_1$ is parallel either to $\langle 100 \rangle$ or $\langle 110 \rangle$. We consider strain-driven uniaxial tension in the direction $\boldsymbol{u}_1$. For such a loading, the strain tensor is obtained by solving

$$\inf_{\boldsymbol{\varepsilon}:\boldsymbol{u}_1\cdot\boldsymbol{\varepsilon}\cdot\boldsymbol{u}_1=\epsilon} \Psi_{eff}(\boldsymbol{\varepsilon}) \tag{90}$$

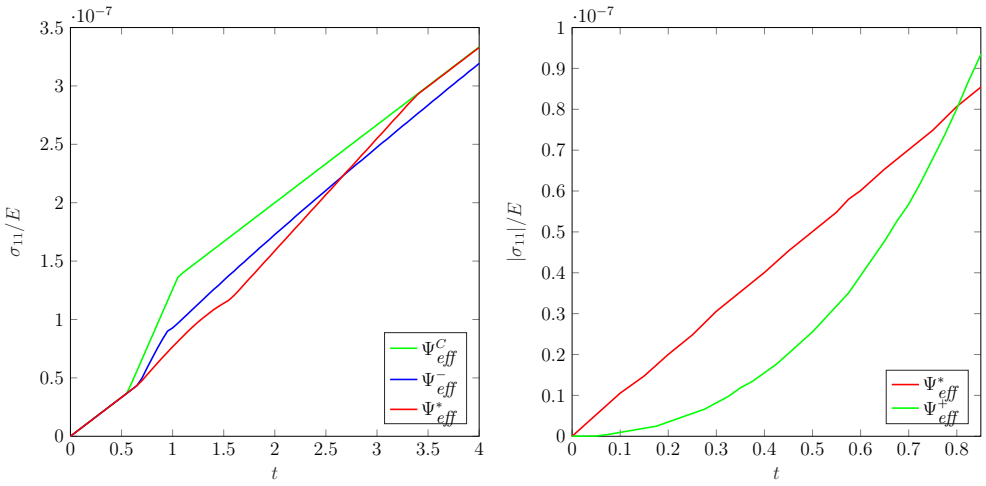

Figure 9: Stress for an applied strain $t(0.6\boldsymbol{\tau}^1+0.4\boldsymbol{\tau}^2)$ with $w_i/E = 4/45000$ (left) and for an applied strain $t(0.55\boldsymbol{\tau}^1 + 0.45\boldsymbol{\tau}^2)$ with $w_i/E = 0$ (right) as calculated from several energy functions.

where ϵ is the applied uniaxial strain. In Fig. 10(left) are shown the curves $\epsilon \mapsto \tilde{\Psi}^*_{eff}(\epsilon)$ and $\epsilon \mapsto \tilde{\Psi}^-_{eff}(\epsilon)$ where

$$\tilde{\Psi}^*_{eff}(\epsilon) = \inf_{\boldsymbol{\varepsilon}:\boldsymbol{u}_1\cdot\boldsymbol{\varepsilon}\cdot\boldsymbol{u}_1=\epsilon} \Psi^*_{eff}(\boldsymbol{\varepsilon}),\ \ \tilde{\Psi}^-_{eff}(\epsilon) = \inf_{\boldsymbol{\varepsilon}:\boldsymbol{u}_1\cdot\boldsymbol{\varepsilon}\cdot\boldsymbol{u}_1=\epsilon} \Psi^-_{eff}(\boldsymbol{\varepsilon}) \tag{91}$$

The volume fraction c_1 of orientation 1 has been set to 0.6 and the values $E = 45$ GPa, $\nu = 0.3$, $w_1 = 4$ MPa, $\alpha = 0.0425$, $\beta = -0.0822$, $\delta = 0.0194$ have been used (Otsuka and Shimizu, 1974). The property $\Psi_{eff} \geq \Psi^-_{eff}$ implies that $\tilde{\Psi}_{eff} \geq \tilde{\Psi}^-_{eff}$ i.e. $\tilde{\Psi}^-_{eff}$ is a lower bound on $\tilde{\Psi}_{eff}$. We can observe in Fig.10 that $\tilde{\Psi}^*_{eff}$ satisfies that lower bound. In Fig. 10(right) are shown the minimizing volume fractions $\boldsymbol{\Theta}$ obtained from the energy estimate Ψ_{eff}, i.e. such that $\tilde{\Psi}^*_{eff}(\epsilon) = \Psi^*_{eff}(\tilde{\boldsymbol{\varepsilon}}, \boldsymbol{\Theta})$ where $\tilde{\boldsymbol{\varepsilon}}$ reaches the minimum in (91). Among the 6 possible variants in each orientation, the calculations shows that only three variants appear, namely variants 1,2,4 in orientation 1 and variants 4,5,6 in orientation 2. The three variants that appear in orientation 1 are the most favorable with respect to the loading direction as can be observed from Table 1 showing the values taken by $\boldsymbol{u}_1 \cdot \boldsymbol{\tau}^j_i \cdot \boldsymbol{u}_1$. Regarding orientation 2, it is interesting to observe that the second best favorable variant (variant 3) is not part of the three variants that appear. Let us compare those predictions with those resulting from the consideration of twin laminates. The 6 martensitic variants in γ'_1CuAlNi are pairwise compatible and it is known that there exists 24 martensitic

Table 1: Values of $\boldsymbol{u}_1 \cdot \boldsymbol{\tau}_i^j \cdot \boldsymbol{u}_1$ for variant i in orientation j.

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
orientation 1	0.0416	0.0415	0.0391	0.0459	−0.0813	−0.0813
orientation 2	−0.0305	−0.0308	0.0234	0.0616	−0.0092	−0.0089

twin laminates that are compatible with the austenite. Those twin laminates are obtained by solving the equation

$$\det(\theta\boldsymbol{\tau}_i + (1-\theta)\boldsymbol{\tau}_j) = 0 \tag{92}$$

with $\theta \in [0,1]$ and $i \neq j$. For any (i,j,θ) solving (92), the effective transformation strain of the twin laminate is $\theta\boldsymbol{\tau}_i+(1-\theta)\boldsymbol{\tau}_j$. Calculating the effective transformation strain of all possible twin laminates shows that the most favorably oriented twin laminates in orientation 1 are $(4,1,\tilde{\theta})$ and $(4,2,\tilde{\theta})$ with $\tilde{\theta} \simeq 0.6926$. The values of $\boldsymbol{u}_1 \cdot (\tilde{\theta}\boldsymbol{\tau}_i + (1-\tilde{\theta})\boldsymbol{\tau}_j) \cdot \boldsymbol{u}_1$ are the same for both twins. For a mix of those two twins (in any proportion), the volume fractions of variants 1,2 and 4 therefore satisfy

$$\frac{\theta_4}{\theta_1+\theta_2} = \frac{\tilde{\theta}}{1-\tilde{\theta}} \simeq 2.25.$$

The numerical results in Fig. 10 give $\frac{\theta_4}{\theta_1+\theta_2} \simeq 2.58$ which is relatively close to the value 2.25. In orientation 2, the most favorably twin laminates are found to be $(4,5,\tilde{\theta})$ and $(4,6,\tilde{\theta})$, so that $\frac{\theta_4}{\theta_5+\theta_6} \simeq 2.25$ in any mix of those twin laminates. The numerical results obtained from the energy estimate Ψ_{eff} give indeed that only variants 4,5,6 appear in orientation 2, but the ratio $\frac{\theta_4}{\theta_5+\theta_6}$ is found to be approximatively equal to 3. The active variants predicted by the energy estimate thus coincide with those expected from the consideration of twin-compatible laminates, but the proportions of the variants are not the same. We note that the energy estimate does not assume any type of microstructure. Twin-compatible second-rank laminates do not always achieve the minimum energy and more complex microstructures can arise.

7. Recoverable strains and low-temperature states in NiTiCu

Our final example is related to recoverable strains and austenite-free mixtures of martensite variants. The material we consider is NiTiCu, which obeys a cubic to monoclinic-I transformation. There are 12 martensitic variants and the corresponding transformations strains $\boldsymbol{\tau}_1, \cdots, \boldsymbol{\tau}_{12}$ are listed in Table A.3. Zhao et al. (1998)

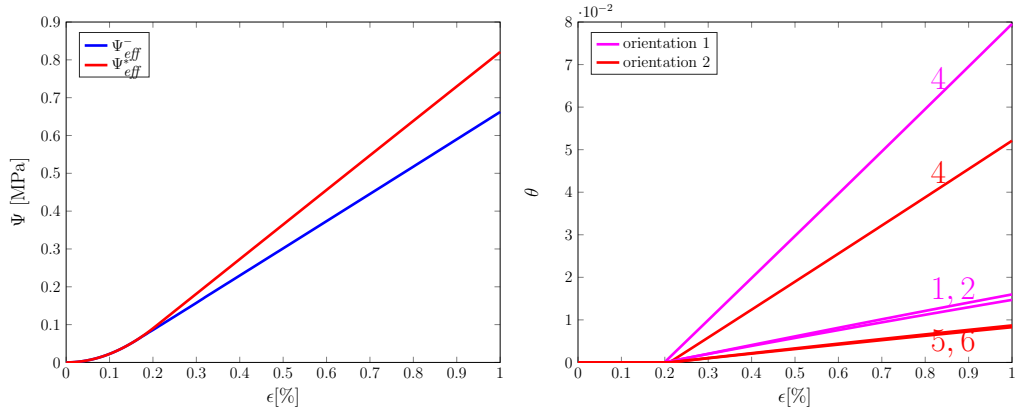

Figure 10: Uniaxial tension in a γ_1'CuAlNi polycrystal: effective energy (left) and martensitic volume fractions (right) as functions of the applied unaxial strain ϵ.

measured the texture and the (uniaxial) maxium recoverable strains in rolled NiTiCu sheets. The experiments of Zhao et al. (1998) show that the texture has mainly 2 orientations in proportions $(c_1, c_2) = (0.6, 0.4)$ and defined by the rotations

$$\boldsymbol{R}^1 = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & -1 & 0 \end{pmatrix}, \quad \boldsymbol{R}^2 = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & 1/\sqrt{6} & \sqrt{2/3} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}. \tag{93}$$

The matrix representations in (93) are relative to a orthonormal basis $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$ where $\boldsymbol{u}_1$ is the rolling direction, $\boldsymbol{u}_2$ is the transverse direction and $\boldsymbol{u}_3$ is the normal to the sheet. Let $\boldsymbol{u}(z) = \cos z\boldsymbol{u}_1 + \sin z\boldsymbol{u}_2$ be the unit vector in the sheet that makes an angle z with the rolling direction. In Fig. 11 (dotted line) is shown the maximum recoverable strain $s(z)$ measured by Zhao et al. (1998) for several values of the angle z. The set $\mathcal{S}_{\mathit{eff}}$ of recoverable strains for a polycrystal is the set of energy-minimizing strains at low temperature ($w_1 = \cdots = w_N < 0$), defined in a way similar to (18) by

$$\mathcal{S}_{\mathit{eff}} = \{\boldsymbol{\varepsilon} : \Psi_{\mathit{eff}}(\boldsymbol{\varepsilon}) = w_1\}. \tag{94}$$

As detailed by Shu and Bhattacharya (1998), the scalar $s(z)$ is related to the set $\mathcal{S}_{\mathit{eff}}$ in (94) by

$$s(z) = \max_{\boldsymbol{\varepsilon} \in \mathcal{S}_{\mathit{eff}}} \boldsymbol{u}(z) \cdot \boldsymbol{\varepsilon} \cdot \boldsymbol{u}(z). \tag{95}$$

The energy estimate Ψ^*_{eff} generates the estimate $\mathcal{S}^*_{\mathit{eff}} = \{\boldsymbol{\varepsilon} : \Psi^*_{\mathit{eff}}(\boldsymbol{\varepsilon}) = w_1\}$ of the set of recoverable strains and therefore an estimate of $s(z)$ given by $\max_{\boldsymbol{\varepsilon} \in \mathcal{S}^*_{\mathit{eff}}} \boldsymbol{u}(z) \cdot$

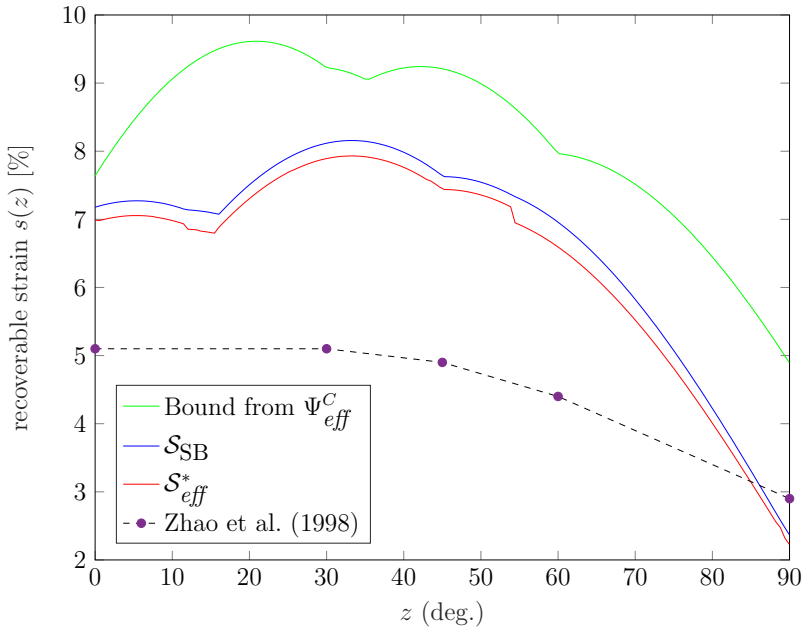

Figure 11: Recoverable strains in NiTiCu with the Zhao-Beyer texture.

$\boldsymbol{\varepsilon} \cdot \boldsymbol{u}(z)$. We are interested here in comparing $\mathcal{S}^*_{eff}$ with an other estimate of the recoverable strains, proposed by Shu and Bhattacharya (1998) and defined as

$$\mathcal{S}_{\text{SB}} = \bigcap_r \boldsymbol{R}^r \mathcal{K}^c \boldsymbol{R}^{r,T} \tag{96}$$

where $\mathcal{K}^c$ is the convex hull of the set $\mathcal{K} = \{\boldsymbol{\tau}_1, \cdots, \boldsymbol{\tau}_{12}\}$. In Fig. 11 (blue line) is shown the estimate of $s(z)$ obtained by replacing $\mathcal{S}_{eff}$ with $\mathcal{S}_{\text{SB}}$ in (95). The estimate of $s(z)$ obtained from the proposed approach is shown in red. The proposed estimate is close to that of Shu and Bhattacharya (1998) but gives slightly smaller values which are overall closer to the experimental results of Zhao et al. (1998). As a comparison, the upper bound on $s(z)$ obtained from the convex energy bound Ψ^C_{eff} in (63) is shown in green in Fig. 11.

We note that the proposed estimate is not restricted to recoverable strains and also gives some information on the energy of any martensite state. To illustrate that point, we consider volume fractions $\boldsymbol{\Theta}$ varying between two special states $\boldsymbol{A}$ and $\boldsymbol{B}$. The state $\boldsymbol{A} = \{A^r_i\}^{1\le r\le 2}_{1\le i\le 12}$ is defined by $A^r_i = 1/12$ and can be interpreted as a self-accommodated state, i.e. the state reached by the material when cooled down from

a high temperature in stress-free conditions. The state $\boldsymbol{B} = \{B_i^r\}_{1\leq i\leq 12}^{1\leq r\leq 2}$ correspond to the volume fractions realized the convex bound on $s(z)$ (green curve in Fig. 11) at $z = 0$. It is defined by $B_2^1 = B_2^2 = 1$, the other components B_i^r being equal to zero. In Fig. 12(red curve) is shown the values taken by the function h for martensitic volume fractions $\boldsymbol{\Theta} = (1-r)\boldsymbol{A} + r\boldsymbol{B}$ for $0 \leq r \leq 1$. The function g corresponding to the lower energy bound Ψ^-_{eff} is shown in blue. In green is shown the function l corresponding to the upper energy bound Ψ^+_{eff} of Hackl and Heinen (2008), see Eq. (58). We can observe in Fig. 12 that $g \leq h \leq l$ as expected from the general results of Sect. 3. More specifically, the plots in Fig. 12 show that h "interpolates" between g and l, behaving as g for low values of r and behaving as l for high values or r. For $r = 0$, it can proved the exact value of the energy is zero. This is captured by the functions g and h but not by l because second-rank lamination is not sufficient to obtain a zero-energy. For $r = 1$, the three functions g, h and l coincide. This actually occurs whenever every crystalline orientation is in a pure phase, i.e. for volume fractions $\boldsymbol{\Theta} = \{\theta_i^r\}$ of the form $\theta_i^r = \sum_s \delta_{rs}\delta_{ij(r)}$. In that case, it is proved in Peigney (2009) that

$$g(\boldsymbol{\Theta}) = l(\boldsymbol{\Theta}) = \frac{1}{4}\sum_{r,s} c_r c_s (\boldsymbol{\tau}^r_{j(r)} - \boldsymbol{\tau}^s_{j(s)}) : (\mathbb{L} - \mathbb{L} : \mathbb{P} : \mathbb{L}) : (\boldsymbol{\tau}^r_{j(r)} - \boldsymbol{\tau}^s_{j(s)})$$

That value is recovered from h as can be verified from (60).

8. Concluding remarks

The presented estimate takes both intra-grain compatibility conditions and inter-grain constraints into account, without making any assumption on the type of austenite-martensite microstructures in each grain. That estimate is guaranteed to satisfy known lower and upper bounds on the effective energy. In that regard, a pattern seems to emerge from the examples presented: the estimate behaves as the upper bound of Hackl and Heinen (2008) for high strains (i.e. in situations where known lower bounds are expected to underestimate the effective energy significantly) and tends to behave as the lower bound of Peigney (2009) for low strains (i.e. in situations where the expected microstructures are not necessarily second-rank twin-compatible laminates). We close with a final remark related to hysteresis effects and energy dissipation. Both thermal and mechanical hysteresis are frequently observed in SMAs. In the framework of energy minimization, using a convex function to estimate the effective energy (such as the lower bound Ψ^C_{eff}) rules out any hysteresis. Then it is necessary to introduce an *ad hoc* dissipation ansatz to capture

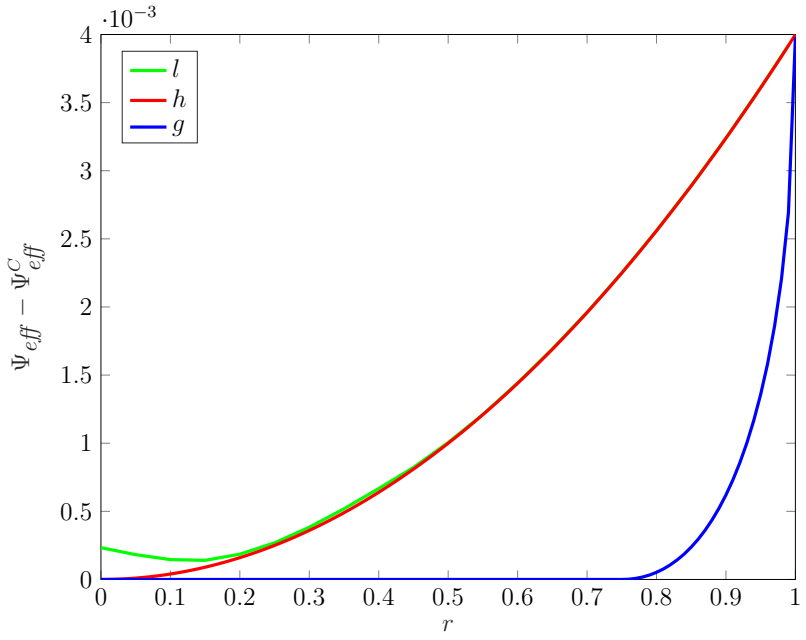

Figure 12: Bounds and estimate on the mixing energy in NiTiCu for martensitic states with volume fractions $(1-r)\boldsymbol{A} + r\boldsymbol{B}$.

Table A.2: Transformation strains for the cubic to orthorhombic transformation.

$\boldsymbol{\tau}_1$	$\boldsymbol{\tau}_2$	$\boldsymbol{\tau}_3$
$\begin{pmatrix} \alpha & 0 & \delta \\ 0 & \beta & 0 \\ \delta & 0 & \alpha \end{pmatrix}$	$\begin{pmatrix} \alpha & 0 & -\delta \\ 0 & \beta & 0 \\ -\delta & 0 & \alpha \end{pmatrix}$	$\begin{pmatrix} \alpha & \delta & 0 \\ \delta & \alpha & 0 \\ 0 & 0 & \beta \end{pmatrix}$
$\boldsymbol{\tau}_4$	$\boldsymbol{\tau}_5$	$\boldsymbol{\tau}_6$
$\begin{pmatrix} \alpha & -\delta & 0 \\ -\delta & \alpha & 0 \\ 0 & 0 & \beta \end{pmatrix}$	$\begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & \delta \\ 0 & \delta & \alpha \end{pmatrix}$	$\begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & -\delta \\ 0 & -\delta & \alpha \end{pmatrix}$

hysteresis in the material model, see e.g. Govindjee and Miehe (2001); Hackl et al. (2008); Peigney et al. (2011); Peigney and Seguin (2013) for examples of such an approach. The situation is different when using the proposed estimate (or any non convex estimate for that matter). Being non convex, that estimate may indeed give rise to metastability and hysteresis effects, without introducing a dissipation ansatz. In particular, mechanical hysteresis in the superelastic regime is investigated in the companion paper in relation with Luders-type instabilities.

Appendix A. Transformation strains

Appendix A.1. Cubic to orthorhombic transformation

For the cubic to orthorhombic transformation, there are 6 martensitic variants with transformation strains listed in Table A.2. Values of the lattice parameters for γ_1'CuAlNi are $\alpha = 0.0425$, $\beta = -0.0822$, $\delta = 0.0194$ (Otsuka and Shimizu, 1974).

Appendix A.2. Cubic to monoclinic-I transformation

The transformation strains of the 12 martensitic variants in the cubic to monoclinic-I transformation are listed in Table A.3. Values of the lattice parameters for TiNiCu are $\alpha = 0.0232$, $\beta = -0.0410$, $\delta = 0.0532$, $\epsilon = 0.0395$ (Nam et al., 1990).

Appendix B. Properties of the $\mathbb{P}$ tensor

In all that follows we use the notation $\mathbb{A} \succeq \mathbb{B}$ to indicate that two given fourth-order tensors $\mathbb{A}$ and $\mathbb{B}$ satisfy $\boldsymbol{\varepsilon} : \mathbb{A} : \boldsymbol{\varepsilon} \geq \boldsymbol{\varepsilon} : \mathbb{B} : \boldsymbol{\varepsilon}$ for all non-zero symmetric second-order tensor $\boldsymbol{\varepsilon}$. Let $\tilde{\mathbb{L}}$ and $\boldsymbol{a}$ satisfying the relations (66). Setting $\tilde{\mathbb{K}} = \tilde{\mathbb{L}} + \mathbb{K}(\boldsymbol{a})$, we

Table A.3: Transformation strains for the cubic to monoclinic-I transformation.

$\boldsymbol{\tau}_1$	$\boldsymbol{\tau}_2$	$\boldsymbol{\tau}_3$
$\begin{pmatrix} \beta & \epsilon & \epsilon \\ \epsilon & \alpha & \delta \\ \epsilon & \delta & \alpha \end{pmatrix}$	$\begin{pmatrix} \beta & -\epsilon & -\epsilon \\ -\epsilon & \alpha & \delta \\ -\epsilon & \delta & \alpha \end{pmatrix}$	$\begin{pmatrix} \beta & -\epsilon & \epsilon \\ -\epsilon & \alpha & -\delta \\ \epsilon & -\delta & \alpha \end{pmatrix}$
$\boldsymbol{\tau}_4$	$\boldsymbol{\tau}_5$	$\boldsymbol{\tau}_6$
$\begin{pmatrix} \beta & \epsilon & -\epsilon \\ \epsilon & \alpha & -\delta \\ -\epsilon & -\delta & \alpha \end{pmatrix}$	$\begin{pmatrix} \alpha & \epsilon & \delta \\ \epsilon & \beta & \epsilon \\ \delta & \epsilon & \alpha \end{pmatrix}$	$\begin{pmatrix} \alpha & -\epsilon & \delta \\ -\epsilon & \beta & -\epsilon \\ \delta & -\epsilon & \alpha \end{pmatrix}$
$\boldsymbol{\tau}_7$	$\boldsymbol{\tau}_8$	$\boldsymbol{\tau}_9$
$\begin{pmatrix} \alpha & -\epsilon & -\delta \\ -\epsilon & \beta & \epsilon \\ -\delta & \epsilon & \alpha \end{pmatrix}$	$\begin{pmatrix} \alpha & \epsilon & -\delta \\ \epsilon & \beta & -\epsilon \\ -\delta & -\epsilon & \alpha \end{pmatrix}$	$\begin{pmatrix} \alpha & \delta & \epsilon \\ \delta & \alpha & \epsilon \\ \epsilon & \epsilon & \beta \end{pmatrix}$
$\boldsymbol{\tau}_{10}$	$\boldsymbol{\tau}_{11}$	$\boldsymbol{\tau}_{12}$
$\begin{pmatrix} \alpha & \delta & -\epsilon \\ \delta & \alpha & -\epsilon \\ -\epsilon & -\epsilon & \beta \end{pmatrix}$	$\begin{pmatrix} \alpha & -\delta & \epsilon \\ -\delta & \alpha & -\epsilon \\ \epsilon & -\epsilon & \beta \end{pmatrix}$	$\begin{pmatrix} \alpha & -\delta & -\epsilon \\ -\delta & \alpha & \epsilon \\ -\epsilon & \epsilon & \beta \end{pmatrix}$

prove in this Appendix that

$$(\mathbb{L}-\tilde{\mathbb{K}})^{-1}-\mathbb{P}\succeq(\mathbb{L}-\tilde{\mathbb{K}}+(\mathbb{L}-\tilde{\mathbb{K}}):\tilde{\mathbb{P}}:(\mathbb{L}-\tilde{\mathbb{K}}))^{-1} \tag{B.1}$$

where $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are the polarization tensors associated to $\mathbb{L}$ and $\tilde{\mathbb{L}}$, as defined by (56). Property (B.1) plays an important role in proving that the proposed estimate Ψ_{eff} complies with the lower bound Ψ^{-}_{eff} in (64). For later reference, we first note that the quadratic function $\boldsymbol{\varepsilon}\mapsto\boldsymbol{\varepsilon}:\mathbb{K}(\boldsymbol{a}):\boldsymbol{\varepsilon}$ is quasiconvex for any $\boldsymbol{a}\succeq 0$, i.e. satisfies

$$\frac{1}{|\Omega|}\int_\Omega \boldsymbol{\varepsilon}:\mathbb{K}(\boldsymbol{a}):\boldsymbol{\varepsilon}d\boldsymbol{x}\geq\bar{\boldsymbol{\varepsilon}}:\mathbb{K}(\boldsymbol{a}):\bar{\boldsymbol{\varepsilon}} \tag{B.2}$$

for any $\boldsymbol{\varepsilon}\in\mathcal{K}(\bar{\boldsymbol{\varepsilon}})$ (Peigney, 2008). Taking $\bar{\boldsymbol{\varepsilon}}=0$, $N=2$ and $\boldsymbol{\eta}^2=0$ in (50) and (51) shows that

$$-\frac{1}{2}c_1(1-c_1)\boldsymbol{\eta}^1:\mathbb{P}:\boldsymbol{\eta}^1=\inf_{\boldsymbol{\varepsilon}\in\mathcal{K}_{(0)}}\frac{1}{|\Omega|}\int_\Omega\sum_{r=1}^{2}\chi^r(\boldsymbol{x})w_r(\boldsymbol{\varepsilon})d\boldsymbol{x} \tag{B.3}$$

where $w_1(\boldsymbol{\varepsilon})=\frac{1}{2}\boldsymbol{\varepsilon}:\mathbb{L}:\boldsymbol{\varepsilon}+\boldsymbol{\eta}^1:\boldsymbol{\varepsilon}$ and $w_2(\boldsymbol{\varepsilon})=\frac{1}{2}\boldsymbol{\varepsilon}:\mathbb{L}:\boldsymbol{\varepsilon}$. For any given $\tilde{\boldsymbol{\eta}}^1$ we have similarly

$$-\frac{1}{2}c_1(1-c_1)\tilde{\boldsymbol{\eta}}^1:\tilde{\mathbb{P}}:\tilde{\boldsymbol{\eta}}^1=\inf_{\boldsymbol{\varepsilon}\in\mathcal{K}_{(0)}}\frac{1}{|\Omega|}\int_\Omega\sum_{r=1}^{2}\chi^r(\boldsymbol{x})\tilde{w}_r(\boldsymbol{\varepsilon})d\boldsymbol{x} \tag{B.4}$$

where $\tilde{w}_1(\boldsymbol{\varepsilon})=\frac{1}{2}\boldsymbol{\varepsilon}:\tilde{\mathbb{L}}:\boldsymbol{\varepsilon}+\tilde{\boldsymbol{\eta}}^1:\boldsymbol{\varepsilon}$ and $\tilde{w}_2(\boldsymbol{\varepsilon})=\frac{1}{2}\boldsymbol{\varepsilon}:\tilde{\mathbb{L}}:\boldsymbol{\varepsilon}$. We introduce the functions $\hat{w}_r=\tilde{w}_r+\frac{1}{2}\boldsymbol{\varepsilon}:\mathbb{K}(\boldsymbol{a}):\boldsymbol{\varepsilon}$ and consider the Legendre transform

$$(w_r-\hat{w}_r)^*(\boldsymbol{\sigma})=\sup_{\boldsymbol{\varepsilon}}\boldsymbol{\sigma}:\boldsymbol{\varepsilon}-w_r(\boldsymbol{\varepsilon})+\hat{w}_r(\boldsymbol{\varepsilon}). \tag{B.5}$$

Definition (B.5) implies that $w_r(\boldsymbol{\varepsilon})\geq\boldsymbol{\sigma}:\boldsymbol{\varepsilon}+\hat{w}_r(\boldsymbol{\varepsilon})-(w_r-\hat{w}_r)^*(\boldsymbol{\sigma})$ for any $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$. It follows that

$$\begin{aligned}\int_\Omega\sum_{r=1}^{2}\chi^r(\boldsymbol{x})w_r(\boldsymbol{\varepsilon})d\boldsymbol{x}\geq\ &\int_\Omega\boldsymbol{\sigma}:\boldsymbol{\varepsilon}d\boldsymbol{x}+\int_\Omega\frac{1}{2}\boldsymbol{\varepsilon}:\mathbb{K}(\boldsymbol{a}):\boldsymbol{\varepsilon}d\boldsymbol{x}+\int_\Omega\sum_{r=1}^{2}\chi^r(\boldsymbol{x})\tilde{w}_r(\boldsymbol{\varepsilon})d\boldsymbol{x}\\&-\int_\Omega(w_r-\hat{w}_r)^*(\boldsymbol{\sigma})d\boldsymbol{x}\end{aligned} \tag{B.6}$$

for any $\boldsymbol{\varepsilon}\in\mathcal{K}(0)$. Choosing $\boldsymbol{\sigma}$ as independent of the location $\boldsymbol{x}$, the first integral in the RHS of (B.6) is equal to 0 because the mean value of $\boldsymbol{\varepsilon}$ vanishes for any

$\boldsymbol{\varepsilon} \in \mathcal{K}(0)$. Moreover, Eq. (B.2) shows that the second integral on the RHS of (B.6) is positive. Eq. (B.4) implies that the third integral in the RHS is bounded from below by $-\frac{1}{2}c_1(1-c_1)\tilde{\boldsymbol{\eta}}^1 : \tilde{\mathbb{P}} : \tilde{\boldsymbol{\eta}}^1 |\Omega|$. We thus arrive at

$$\frac{1}{\Omega}\int_\Omega \sum_{r=1}^{2} \chi^r(\boldsymbol{x}) w_r(\boldsymbol{\varepsilon}) \geq -\frac{1}{2}c_1(1-c_1)\tilde{\boldsymbol{\eta}}^1 : \tilde{\mathbb{P}} : \tilde{\boldsymbol{\eta}}^1 - \sum_{r=1}^{2} c_r (w_r - \hat{w}_r)^*(\boldsymbol{\sigma}). \tag{B.7}$$

Taking the infimum over $\boldsymbol{\varepsilon} \in \mathcal{K}(0)$ yields, in view of (B.3),

$$-\frac{1}{2}c_1(1-c_1)\boldsymbol{\eta}^1 : \mathbb{P} : \boldsymbol{\eta}^1 \geq -\frac{1}{2}c_1(1-c_1)\tilde{\boldsymbol{\eta}}^1 : \tilde{\mathbb{P}} : \tilde{\boldsymbol{\eta}}^1 - \sum_{r=1}^{2} c_r (w_r - \hat{w}_r)^*(\boldsymbol{\sigma}). \tag{B.8}$$

The terms $(w_r - \hat{w}_r)^*(\boldsymbol{\sigma})$ can be calculated directly from (B.5). Substituting the result in (B.8) gives

$$\begin{aligned} -\frac{1}{2}c_1(1-c_1)\boldsymbol{\eta}^1 : \mathbb{P} : \boldsymbol{\eta}^1 \geq \; & -\frac{1}{2}c_1(1-c_1)\tilde{\boldsymbol{\eta}}^1 : \tilde{\mathbb{P}} : \tilde{\boldsymbol{\eta}}^1 - \frac{1}{2}\boldsymbol{\sigma} : (\mathbb{L} - \tilde{\mathbb{K}})^{-1} : \boldsymbol{\sigma} \\ & + c_1 \boldsymbol{\sigma} : (\mathbb{L} - \tilde{\mathbb{K}})^{-1} : (\boldsymbol{\eta}_1 - \tilde{\boldsymbol{\eta}}_1) - \frac{c_1}{2}(\boldsymbol{\eta}_1 - \tilde{\boldsymbol{\eta}}_1) : (\mathbb{L} - \tilde{\mathbb{K}})^{-1} : (\boldsymbol{\eta}_1 - \tilde{\boldsymbol{\eta}}_1). \end{aligned}$$

The right-hand side of the inequality above is quadratic in $\boldsymbol{\sigma}$. Maximizing with respect to $\boldsymbol{\sigma}$ gives

$$-\boldsymbol{\eta}^1 : \mathbb{P} : \boldsymbol{\eta}^1 \geq -\tilde{\boldsymbol{\eta}}^1 : \tilde{\mathbb{P}} : \tilde{\boldsymbol{\eta}}^1 - (\boldsymbol{\eta}_1 - \tilde{\boldsymbol{\eta}}_1) : (\mathbb{L} - \tilde{\mathbb{K}})^{-1} : (\boldsymbol{\eta}_1 - \tilde{\boldsymbol{\eta}}_1). \tag{B.9}$$

Choosing in particular $\tilde{\boldsymbol{\eta}}^1 = 0$ in (B.9) shows that

$$(\mathbb{L} - \tilde{\mathbb{K}})^{-1} \succeq \mathbb{P} \tag{B.10}$$

which notably implies (by taking $\boldsymbol{a} = 0$ and $\tilde{\mathbb{L}} \to 0$) the property

$$\mathbb{L}^{-1} \succeq \mathbb{P} \tag{B.11}$$

as mentioned in Sect. 3.

Note that (B.9) is quadratic in $\boldsymbol{\eta}^1$ and can be rewritten as

$$\boldsymbol{\eta}^1 : ((\mathbb{L} - \tilde{\mathbb{K}})^{-1} - \mathbb{P}) : \boldsymbol{\eta}^1 - 2\boldsymbol{\eta}^1 : (\mathbb{L} - \tilde{\mathbb{K}})^{-1} : \tilde{\boldsymbol{\eta}}^1 \geq -\tilde{\boldsymbol{\eta}}^1 : ((\mathbb{L} - \tilde{\mathbb{K}})^{-1} + \tilde{\mathbb{P}}) : \tilde{\boldsymbol{\eta}}^1. \tag{B.12}$$

Eq. (B.10) ensures that the LHS of (B.12) remains bounded from below when $\boldsymbol{\eta}^1$ varies. Minimizing (B.12) with respect to $\boldsymbol{\eta}^1$ yields

$$-\tilde{\boldsymbol{\eta}}^1 : (\mathbb{L} - \tilde{\mathbb{K}})^{-1} : ((\mathbb{L} - \tilde{\mathbb{K}})^{-1} - \mathbb{P})^{-1} : (\mathbb{L} - \tilde{\mathbb{K}})^{-1} : \tilde{\boldsymbol{\eta}}^1 \geq -\tilde{\boldsymbol{\eta}}^1 : ((\mathbb{L} - \tilde{\mathbb{K}})^{-1} + \tilde{\mathbb{P}}) : \tilde{\boldsymbol{\eta}}^1.$$

Setting $\boldsymbol{\eta} = (\mathbb{L} - \tilde{\mathbb{K}})^{-1} : \tilde{\boldsymbol{\eta}}^1$, we have

$$-\boldsymbol{\eta} : ((\mathbb{L} - \tilde{\mathbb{K}})^{-1} - \mathbb{P})^{-1} : \boldsymbol{\eta} \geq -\boldsymbol{\eta} : ((\mathbb{L} - \tilde{\mathbb{K}}) + (\mathbb{L} - \tilde{\mathbb{K}}) : \tilde{\mathbb{P}} : (\mathbb{L} - \tilde{\mathbb{K}})) : \boldsymbol{\eta}$$

for all $\boldsymbol{\eta}$, i.e.

$$(\mathbb{L} - \tilde{\mathbb{K}}) + (\mathbb{L} - \tilde{\mathbb{K}}) : \tilde{\mathbb{P}} : (\mathbb{L} - \tilde{\mathbb{K}}) \succeq ((\mathbb{L} - \tilde{\mathbb{K}})^{-1} - \mathbb{P})^{-1} \tag{B.13}$$

Eq. (B.1) follows from taking the inverse in (B.13) (we recall that $\mathbb{B} \succeq \mathbb{A} \succeq 0 \Longrightarrow \mathbb{A}^{-1} \succeq \mathbb{B}^{-1}$, see e.g. Zhan (2002)).

Remark: Some additional manipulations show that (B.13) can also be rewritten in the more compact form $\mathbb{P}^{-1} - \tilde{\mathbb{P}}^{-1} \succeq \mathbb{L} - \tilde{\mathbb{K}}$. However, formulation (B.1) will be more convenient for our purpose.

Appendix C. Proof of the inequality $h \geq g$

We show in this Appendix that

$$h(\boldsymbol{\Theta}) \geq g(\boldsymbol{\Theta}) \tag{C.1}$$

for any $\boldsymbol{\Theta} \in \mathcal{T}_n^N$. We first note that definition (65) can be rewritten as

$$g(\boldsymbol{\Theta}) = \sup_{\boldsymbol{a}, \tilde{\mathbb{L}}} H(\boldsymbol{\Theta}, \boldsymbol{a}, \tilde{\mathbb{L}}) \tag{C.2}$$

where the supremum is taken over tensors $(\boldsymbol{a}, \tilde{\mathbb{L}})$ satisfying (66) and

$$H(\boldsymbol{\Theta}, \boldsymbol{a}, \tilde{\mathbb{L}}) = \frac{1}{4} \sum_{r,s,i,j} c_r c_s \theta_i^r \theta_j^s (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^s) : \tilde{\mathbb{M}} : (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^s) + \frac{1}{2} \sum_r c_r \boldsymbol{h}^r : \tilde{\mathbb{Q}} : \boldsymbol{h}^r \tag{C.3}$$

with

$$\tilde{\mathbb{M}} = \mathbb{L} - \mathbb{L} : (\mathbb{L} - \tilde{\mathbb{K}})^{-1} : \mathbb{L}, \quad \tilde{\mathbb{Q}} = \Big(\mathbb{L} - \tilde{\mathbb{K}} + (\mathbb{L} - \tilde{\mathbb{K}}) : \tilde{\mathbb{P}} : (\mathbb{L} - \tilde{\mathbb{K}})\Big)^{-1}.$$

For any given tensors $(\boldsymbol{a}, \tilde{\mathbb{L}})$ satisfying (16), we show that $H(\boldsymbol{\Theta}, \boldsymbol{a}, \tilde{\mathbb{L}}) \leq h(\boldsymbol{\Theta})$, from which (C.1) will follow directly. Recalling that $\sum_{i=0}^n \theta_i^r = 1$ and $\sum_{r=1}^N c_r = 1$, some straightforward manipulations show that

$$\frac{1}{4} \sum_{r,s,i,j} c_r c_s \theta_i^r \theta_j^s (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^s) : \tilde{\mathbb{M}} : (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^s) = \frac{1}{2} \sum_{r,i} c_r \theta_i^r \boldsymbol{\tau}_i^r : \tilde{\mathbb{M}} : \boldsymbol{\tau}_i^r - \frac{1}{2} \bar{\boldsymbol{\tau}} : \tilde{\mathbb{M}} : \bar{\boldsymbol{\tau}}$$

with $\bar{\boldsymbol{\tau}} = \sum_{r,i} c_r \theta_i^r \boldsymbol{\tau}_i^r$. In a similar fashion, we have, for any given r,

$$\frac{1}{4}\sum_{i,j} \theta_i^r \theta_j^r (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^r) : \tilde{\mathbb{M}} : (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^r) = \frac{1}{2}\sum_i \theta_i^r \boldsymbol{\tau}_i^r : \tilde{\mathbb{M}} : \boldsymbol{\tau}_i^r - \frac{1}{2}\boldsymbol{\tau}^r : \tilde{\mathbb{M}} : \boldsymbol{\tau}^r$$

where $\boldsymbol{\tau}^r = \sum_i \theta_i^r \boldsymbol{\tau}_i^r$. It follows that

$$\begin{aligned} H(\boldsymbol{\Theta}, \boldsymbol{a}, \tilde{\mathbb{L}}) = \; & \frac{1}{4}\sum_{r,i,j} c_r \theta_i^r \theta_j^r (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^r) : \tilde{\mathbb{M}} : (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^r) \\ & + \frac{1}{2}\sum_r c_r \boldsymbol{\tau}^r : \tilde{\mathbb{M}} : \boldsymbol{\tau}^r - \frac{1}{2}\bar{\boldsymbol{\tau}} : \tilde{\mathbb{M}} : \bar{\boldsymbol{\tau}} + \frac{1}{2}\sum_r c_r \boldsymbol{h}^r : \tilde{\mathbb{Q}} : \boldsymbol{h}^r. \end{aligned} \tag{C.4}$$

Observing that $\sum_r c_r(\boldsymbol{\tau}^r - \bar{\boldsymbol{\tau}}) : \tilde{\mathbb{M}} : (\boldsymbol{\tau}^r - \bar{\boldsymbol{\tau}}) = \sum_r c_r \boldsymbol{\tau}^r : \tilde{\mathbb{M}} : \boldsymbol{\tau}^r - \bar{\boldsymbol{\tau}} : \tilde{\mathbb{M}} : \bar{\boldsymbol{\tau}}$ and recalling from (62) that $\boldsymbol{h}^r = -\mathbb{L} : (\boldsymbol{\tau}^r - \bar{\boldsymbol{\tau}})$, we obtain

$$H(\boldsymbol{\Theta}, \boldsymbol{a}, \tilde{\mathbb{L}}) = \frac{1}{4}\sum_{r,i,j} c_r \theta_i^r \theta_j^r (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^r) : \tilde{\mathbb{M}} : (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^r) + \frac{1}{2}\sum_r c_r \boldsymbol{h}^r : (\mathbb{L}^{-1} - (\mathbb{L} - \tilde{\mathbb{K}})^{-1} + \tilde{\mathbb{Q}}) : \boldsymbol{h}^r. \tag{C.5}$$

We now proceed to bound the two sums in the RHS of (C.5). For any given r, we have by definition (14)

$$\frac{1}{4}\sum_{i,j=0}^{n} \theta_i^r \theta_j^r (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) : \mathbb{M}(\boldsymbol{a}) : (\boldsymbol{\tau}_i - \boldsymbol{\tau}_j) \leq f(\boldsymbol{\theta}^r) \tag{C.6}$$

for any $\boldsymbol{a} \succeq 0$ such that $\mathbb{L} - \mathbb{K}(\boldsymbol{a}) \succ 0$. The elasticity tensor $\mathbb{L}$ being isotropic, an important observation is that the function f in (14) is rotationally invariant with respect to the transformation strains. This implies that we may as well replace $\boldsymbol{\tau}_i$ with $\boldsymbol{\tau}_i^r = \boldsymbol{R}^r \boldsymbol{\tau}_i \boldsymbol{R}^{r,T}$ in (C.6), so that

$$\frac{1}{4}\sum_{i,j=0}^{n} \theta_i^r \theta_j^r (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^r) : \mathbb{M}(\boldsymbol{a}) : (\boldsymbol{\tau}_i^r - \boldsymbol{\tau}_j^r) \leq f(\boldsymbol{\theta}^r) \tag{C.7}$$

for any $\boldsymbol{a} \succeq 0$ such that $\mathbb{L} - \mathbb{K}(\boldsymbol{a}) \succ 0$. Consider given tensors $(\boldsymbol{a}, \tilde{\mathbb{L}})$ verifying (16). We have $\mathbb{L} \succ \tilde{\mathbb{K}} = \tilde{\mathbb{L}} + \mathbb{K}(\boldsymbol{a}) \succ \mathbb{K}(\boldsymbol{a})$, hence $\mathbb{L} - \mathbb{K}(\boldsymbol{a}) \succ 0$. Eq. (C.7) is thus satisfied. The relation $\tilde{\mathbb{K}} \succ \mathbb{K}(\boldsymbol{a})$ also implies that $\mathbb{L} - \mathbb{K}(\boldsymbol{a}) \succ \mathbb{L} - \tilde{\mathbb{K}}$. Since $\mathbb{L} - \tilde{\mathbb{K}} \succ 0$, it follows that $(\mathbb{L} - \tilde{\mathbb{K}})^{-1} \succ (\mathbb{L} - \mathbb{K}(\boldsymbol{a}))^{-1}$ and consequently that

$$\mathbb{M}(\boldsymbol{a}) \succ \mathbb{L} - \mathbb{L} : (\mathbb{L} - \tilde{\mathbb{K}})^{-1} : \mathbb{L} = \tilde{\mathbb{M}} \tag{C.8}$$

where $\mathbb{M}(\boldsymbol{a})$ is defined as in (15). It follows that

$$\frac{1}{4}\sum_{i,j}\theta_i^r\theta_j^r(\boldsymbol{\tau}_i^r-\boldsymbol{\tau}_j^r):\tilde{\mathbb{M}}:(\boldsymbol{\tau}_i^r-\boldsymbol{\tau}_j^r)\leq\frac{1}{4}\sum_{i,j}\theta_i^r\theta_j^r(\boldsymbol{\tau}_i^r-\boldsymbol{\tau}_j^r):\mathbb{M}(\boldsymbol{a}):(\boldsymbol{\tau}_i^r-\boldsymbol{\tau}_j^r)\leq f(\boldsymbol{\theta}^r)$$

and consequently that

$$H(\boldsymbol{\Theta},\boldsymbol{a},\tilde{\mathbb{L}})\leq\sum_r c_r f(\boldsymbol{\theta}^r)+\frac{1}{2}\sum_r c_r\boldsymbol{h}^r:(\mathbb{L}^{-1}-(\mathbb{L}-\tilde{\mathbb{K}})^{-1}+\tilde{\mathbb{Q}}):\boldsymbol{h}^r.$$

Eq. (B.1) proved in Appendix B entails that $\mathbb{L}^{-1}-\mathbb{P}\succeq\mathbb{L}^{-1}-(\mathbb{L}-\tilde{\mathbb{K}})^{-1}+\tilde{\mathbb{Q}}$, so that $\boldsymbol{h}^r:(\mathbb{L}^{-1}-(\mathbb{L}-\tilde{\mathbb{K}})^{-1}+\tilde{\mathbb{Q}}):\boldsymbol{h}^r\leq\boldsymbol{h}^r:(\mathbb{L}^{-1}-\mathbb{P}):\boldsymbol{h}^r$ and

$$H(\boldsymbol{\Theta},\boldsymbol{a},\tilde{\mathbb{L}})\leq\sum_r c_r f(\boldsymbol{\theta}^r)+\frac{1}{2}\sum_r c_r\boldsymbol{h}^r:(\mathbb{L}^{-1}-\mathbb{P}):\boldsymbol{h}^r.$$

The right-hand of that inequality is equal to $h(\boldsymbol{\Theta})$, see. Eq (61). We thus have $H(\boldsymbol{\Theta},\boldsymbol{a},\tilde{\mathbb{L}})\leq h(\boldsymbol{\Theta})$ for all $(\boldsymbol{a},\tilde{\mathbb{L}})$ verifying (16). In view of (C.2), taking the supremum over $(\boldsymbol{a},\tilde{\mathbb{L}})$ gives $g(\boldsymbol{\Theta})\leq h(\boldsymbol{\Theta})$.

References

Andersen, N., Lebech, B., Poulsen, H., 1990. The structural phase diagram and oxygen equilibrium partial pressure of $YBa_2Cu_3O_{6+x}$ studied by neutron powder diffraction and gas volumetry. Physica C: Superconductivity 172, 31–42.

Ball, J.M., James, R.D., 1992. Proposed experimental tests of a theory of fine microstructure and the two-well problem. Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences 338, 389–450.

Bhattacharya, K., 2003. Microstructure of martensite: why it forms and how it gives rise to the shape-memory effect. Oxford University Press.

Bhattacharya, K., Kohn, R.V., 1997. Elastic energy minimization and the recoverable strains of polycrystalline shape-memory materials. Archive for Rational Mechanics and Analysis 139, 99–180.

Boussaid, O., Kreisbek, C., Schlömerkemper, A., 2019. Characterizations of symmetric polyconvexity. Archive for Rational Mechanics and Analysis 234, 417–451.

Bunge, H.J., 2013. Texture analysis in materials science: mathematical methods. Elsevier.

Castaneda, P.P., Suquet, P., 1997. Nonlinear composites, in: Advances in applied mechanics. Elsevier. volume 34, pp. 171–302.

Churchill, C., Shaw, J., Iadicola, M., 2009. Tips and tricks for characterizing shape memory alloy wire: part 3-localization and propagation phenomena. Experimental Techniques 33, 70–78.

Cisse, C., Zaki, W., Zineb, T.B., 2016. A review of constitutive models and modeling techniques for shape memory alloys. International Journal of Plasticity 76, 244–284.

Dacorogna, B., 2007. Direct methods in the calculus of variations. volume 78. Springer Science & Business Media.

Daly, S., Ravichandran, G., Bhattacharya, K., 2007. Stress-induced martensitic phase transformation in thin sheets of nitinol. Acta Materialia 55, 3593–3600.

Eucken, S., Hirsch, J., 1990. The effect of textures on shape memory behaviour, in: Materials Science Forum, Trans Tech Publ. pp. 487–492.

Govindjee, S., Hackl, K., Heinen, R., 2007. An upper bound to the free energy of mixing by twin-compatible lamination for n-variant martensitic phase transformations. Continuum Mechanics and Thermodynamics 18, 443–453.

Govindjee, S., Miehe, C., 2001. A multi-variant martensitic phase transformation model: formulation and numerical implementation. Computer Methods in Applied Mechanics and Engineering 191, 215–238.

Govindjee, S., Mielke, A., Hall, G., 2003. The free energy of mixing for n-variant martensitic phase transformations using quasi-convex analysis. Journal of the Mechanics and Physics of Solids 4, 763.

Hackl, K., Heinen, R., 2008. An upper bound to the free energy of $n-$variant polycrystalline shape memory alloys. J.Mech.Phys.Solids 56, 2832–2843.

Hackl, K., Heinen, R., Schmahl, W.W., Hasan, M., 2008. Experimental verification of a micromechanical model for polycrystalline shape memory alloys in dependence of martensite orientation distributions. Materials Science and Engineering: A 481, 347–350.

Kachanov, L.M., 2004. Fundamentals of the Theory of Plasticity. Courier Corporation.

Kohn, R.V., 1991. The relaxation of a double-well energy. Continuum Mechanics and Thermodynamics 3, 193–236.

Lurie, K.A., Cherkaev, A.V., 1984. Exact estimates of conductivity of composites formed by two isotropically conducting media taken in prescribed proportion. Proceedings of the Royal Society of Edinburgh Section A: Mathematics 99, 71–87.

Milton, G., 2002. The theory of composites. Cambridge University Press.

Murat, F., 1985. Calcul des variations et homogénéisation, in: Les Méthodes de l'Homogenéisation: Théorie et Applications en Physique, Coll. de la Direction des Etudes et Recherches de EDF, Eyrolles. pp. 319–370.

Nam, T.H., Saburi, T., Shimizu, K., 1990. Cu-content dependence of shape memory characteristics in ti–ni–cu alloys. Materials Transactions, JIM 31, 959–967.

Otsuka, K., Shimizu, K., 1974. Morphology and crystallography of thermoelastic cu–al–ni martensite analyzed by the phenomenological theory. Transactions of the Japan Institute of Metals 15, 103–108.

Peigney, M., 2008. Recoverable strains in composite shape memory alloys. J. Mech. Phys. Solids 56, 360–375.

Peigney, M., 2009. A non-convex lower bound on the effective free energy of polycrystalline shape memory alloys. J.Mech.Phys.Solids 57, 970–986.

Peigney, M., 2013. On the energy-minimizing strains in martensitic microstructures-part 2: Geometrically linear theory. J.Mech.Phys.Solids 61, 1511–1530.

Peigney, M., Seguin, J., 2013. An incremental variational approach to coupled thermo-mechanical problems in anelastic solids. application to shape-memory alloys. Int. J. Sol. Struct. 50, 4043–4054.

Peigney, M., Seguin, J., Hervé-Luanco, E., 2011. Numerical simulation of shape memory alloys structures using interior-point methods. Int. J. Sol. Struct. 48, 2791–2799.

Pipkin, A.C., 1991. Elastic materials with two preferred states. The Quarterly Journal of Mechanics and Applied Mathematics 44, 1–15.

Shaw, J.A., Kyriakides, S., 1997. On the nucleation and propagation of phase transformation fronts in a niti alloy. Acta materialia 45, 683–700.

Shu, Y., Bhattacharya, K., 1998. The influence of texture on the shape-memory effect in polycrystals. Acta Materialia 46, 5457–5473.

Smyshlyaev, V., Willis, J., 1998. A 'non-local' variational approach to the elastic energy minimalization of martensitic polycrystals. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 454, 1573–1613.

Smyshlyaev, V., Willis, J., 1999. On the relation of a three-well energy. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 455, 779–814.

Willis, J.R., 1977. Bounds and self-consistent estimates for the overall properties of anisotropic composites. Journal of the Mechanics and Physics of Solids 25, 185–202.

Willis, J.R., 1981. Variational and related methods for the overall properties of composites, in: Advances in applied mechanics. Elsevier. volume 21, pp. 1–78.

Zhan, X., 2002. Matrix inequalities. volume 1790. Springer Science & Business Media.

Zhang, Z., James, R.D., Müller, S., 2009. Energy barriers and hysteresis in martensitic phase transformations. Acta Materialia 57, 4332–4352.

Zhao, L., Willemse, P., Mulder, J., Beyer, J., Wei, W., 1998. Texture development and transformation strain of a cold-rolled ti50-ni45-cu5 alloy. Scripta materialia 39, 1317–1323.