Title
The free-energy of mixing for n-variant martensitic phase transformations using quasi-convex analysis

Authors
Govindjee, Sanjay
Mielke, Alexandar
Hall, Garrett

Publication Date
2001-05-01

Report No.
UCB/SEMM-2001/04

Structural Engineering
Mechanics and Materials

The Free-Energy of Mixing for n-Variant Martensitic Phase Transformations using Quasi-Convex Analysis

By

Sanjay Govindjee

Alexandar Mielke

and

Garrett J. Hall

May 2001

Department of Civil and Environmental Engineering
University of California, Berkeley

The Free-Energy of Mixing for n-Variant Martensitic Phase Transformations using Quasi-Convex Analysis

Sanjay Govindjee *
Alexandar Mielke †
Garrett J. Hall ‡

May 4, 2001

Abstract

The construction of effective models for materials that undergo martensitic phase transformations requires usable and accurate functional representations for the free energy density. The general representation of this energy is known to be highly non-convex; it even lacks the property of quasi-convexity. A quasi-convex relaxation, however, does permit one to make make certain estimates and powerful conclusions regarding phase transformation. The general expression for the relaxed free energy is however not known in the n-variant case. Analytic solutions are known only for up to 3 variants; whereas cases of practical interests involve 7 to 13 variants. In this study we examine the n-variant case utilizing relaxation theory and produce a seemingly obvious but very powerful observation regarding a lower-bound to the quasi-convex relaxation that make practical evolutionary computations possible. We also examine in detail the 4-variant case where

*Associate Professor of Civil Engineering, University of California Berkeley, govindjee@ce.berkeley.edu

†Professor of Mathematics, Institut A für Mathematik, Universität Stuttgart, mielke@mathematik.uni-stuttgart.de

‡Graduate Student Civil Engineering, University of California Berkeley, gjhall@ce.berkeley.edu

we explicitly show the relation between three different forms of the free-energy of mixing: Kohn's upper-bound by lamination, the Reuß lower-bound, and a lower estimate of the $\mathbb{H}$-measure bound. A discussion of the bounds and their utility is discussed; sample computations are presented for illustrative purposes.

Contents

1 Introduction

In the theory of shape memory alloys and other martensitic problems it is the basic assumption, that each mesoscopic part of the crystal can choose to be in one of the n allowed phases. We distinguish these phases by their stored–energy densities

$$W(\boldsymbol{\varepsilon}, \mathrm{e}_i) = \frac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_i):\mathbb{C}_i:(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_i) + \alpha_i, \qquad i = 1, \ldots, n,$$

where $\mathbf{e}_i \in \mathbb{R}^n$ is the ith unit vector, $\boldsymbol{\varepsilon} = \nabla_{\text{sym}}\mathbf{u} = \frac{1}{2}(\nabla\mathbf{u}+\nabla\mathbf{u}^T) \in \mathbb{R}^{d\times d}_{\text{sym}}$ is the linearized strain tensor, $\mathbb{C}_i$ is the rank-4 elasticity tensor for phase i, α_i is a temperature dependent term that defines each "well-height", and $\boldsymbol{\varepsilon}_i$ denotes the transformation strain of phase i. It has been argued by Ball and James

[2, 3] that the overall macroscopic energy density of such a material can be defined in terms of the individual phase energy densities in the following fashion:

$$W(\boldsymbol{\varepsilon}) = \min_{i=1,\dots,n} [W(\boldsymbol{\varepsilon}, \mathbf{e}_i)] \ . \tag{1}$$

The essential concept that is being advanced is that at each point in the body the material will convert to the phase that will generate the lowest local energy. This, in turn, generates what can appropriately be termed a process of equilibrium (reversible) phase transformation. If one considers the minimization of the potential energy of a body under Dirichlet boundary conditions, then one is faced with the problem of minimizing the integral of the energy density (1) over the body of interest. This integral is a functional of the deformation and is well-known to not be weakly lower semi-continuous (wlsc). Lack of the wlsc property, implies non-existence of a minimizing deformation for given boundary data and hence indicates the formation of microstructure; see e.g. [2, 3]. This primary difficulty can be elegantly circumvented by utilizing the quasi-convex relaxation of the energy density (1) in the integrand of the total potential energy of the body [17, 28, 11]. The result of this procedure is a potential energy functional that is wlsc. With wlsc and the addition of a mild coercivity condition on the functional one has a well established existence theorem for minimizing deformations [11]. We may call such deformations macroscopic as micro- and mesoscopic structure is averaged out. Nevertheless analyzing the macroscopic deformation gradients and the way the quasi-convex hull is formed a good prediction of various aspects of fine structure in martensitic alloys can be given – including for example observed twin structures, habit planes, etc. [2, 3, 5, 4, 8, 9, 7, 6, 27].

The aforementioned procedure is now well established in the mathematical literature concerning non-convex analysis and it has been applied successfully to the notions of phase transformation. One "detraction", however, of this procedure is that it generates an equilibrium model of phase transformation. Thus the model generated is incapable of predicting progressive evolution in homogeneous states of deformation and hysteresis effects – both of which are often of great interest. Macroscopic models that attempt to describe phase transformation in an evolutionary manner with hysteresis often employ as an internal variable the phase fractions $\mathbf{c} \in \mathcal{P}^n = \operatorname{conv}(\mathcal{P}^n_{\text{pure}})$ where $\mathcal{P}^n_{\text{pure}} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$; see e.g. [18, 19, 21, 23, 13, 10, 15, 12]. Geometrically, $\mathcal{P}^n$ is a convex polytope whose extremal points are given by $\mathcal{P}^n_{\text{pure}}$.

In this paper, we do not directly address the evolutionary problem, rather we consider in some detail the free energy density of the phase transforming material as parameterized by the phase fractions. In particular we adopt the highly successful formalism of quasi-convex analysis to describe a mathematically well-motivated free energy density $QW(\varepsilon, \mathbf{c})$ which is defined in (5). This function is derived mathematically solely from the functions $W(\cdot, \mathbf{e}_i)$, $i = 1, \dots, n$, using the notion of quasi-convex relaxation at fixed volume fraction. Hence, $QW(\varepsilon, \mathbf{c})$ is given as the minimal energy in a representative volume element where the minimization is done over all possible mesoscopic arrangements of the phases which are compatible with a given volume fraction $\mathbf{c} \in \mathcal{P}^n$ and all mesoscopic deformations with macroscopic strain ε; see e.g. [1] or [28]. The relevance of $QW(\varepsilon, \mathbf{c})$ for the modeling of the rate-independent evolution of microstructure was first emphasized in [23, 21] and further investigated in [22] where also a mathematical existence theorem for the two-phase situation is given. The mathematical justification of the up-scaling from the micro- to the macroscale is addressed in [30].

If the elasticity tensors are assumed identical (i.e., $\mathbb{C}_i = \mathbb{C}$), then one can make considerable progress and the macroscopic free energy density takes the form

$$QW(\varepsilon, \mathbf{c}) = \sum_{i=1}^{n} c_i W(\varepsilon, \mathbf{e}_i) + w_{\text{mix}}(\mathbf{c}), \quad \mathbf{c} = (c_1, \dots, c_n)^T,$$

where the free energy of mixing $w_{\text{mix}} : \mathcal{P}^n \to (-\infty, 0]$ is convex and satisfies $w_{\text{mix}}(\mathbf{e}_i) = 0$. While $w_{\text{mix}}(\theta \mathbf{e}_i + (1-\theta)\mathbf{e}_j)$ can be given explicitly for $\theta \in [0, 1]$ and all $i \neq j$ (see [17]), it is not known how to compute $w_{\text{mix}}(\mathbf{c})$ in the general case; for a discussion see e.g. [28] for $n = 3$ and [20] for general n. Here we provide upper and lower estimates w^{upp} and w^{low} for $w_{\text{mix}}(\mathbf{c})$ in terms of the function $\psi : \mathbb{R}^n_* = \{\, \boldsymbol{\eta} \in \mathbb{R}^n \mid \boldsymbol{\eta} \cdot \mathbf{e}_* = \sum \eta_j = 0 \,\} \to \mathbb{R}$ given by

$$\psi(\boldsymbol{\eta}) = \min\{ -\frac{1}{2}[\boldsymbol{\omega} \cdot \mathbb{C}{:}\boldsymbol{\varepsilon}^{\boldsymbol{\eta}}] \cdot \mathbf{T}^{-1}[\boldsymbol{\omega} \cdot \mathbb{C}{:}\boldsymbol{\varepsilon}^{\boldsymbol{\eta}}] \mid \boldsymbol{\omega} \in \mathbb{S}^{d-1} \} \leq 0 \qquad (2)$$

with

$$\boldsymbol{\varepsilon}^{\boldsymbol{\eta}} = \sum_{j=1}^{n} \eta_j \boldsymbol{\varepsilon}_j \,,$$

where $\mathbf{e}_* = (1, \dots, 1)^T \in \mathbb{R}^n$, $\mathbf{T}(\boldsymbol{\omega})$ is the acoustic tensor, and $\mathbb{S}^{d-1}$ is the unit sphere embedded in $\mathbb{R}^d$. With $\mathbf{M}(\mathbf{c}) = \operatorname{diag}(\mathbf{c}) - \mathbf{c} \otimes \mathbf{c}$ and $m = (n^2 - n)/2$

one obtains the lower bound

$$w^{\text{low}}(\mathbf{c}) = \min\{\sum_{j=1}^{m}\psi(\boldsymbol{\eta}_j) \mid \boldsymbol{\eta}_j \in \mathbb{R}^n_*, \sum_{j=1}^{m}\boldsymbol{\eta}_j\otimes\boldsymbol{\eta}_j = \mathbf{M}(\mathbf{c})\}. \tag{3}$$

An upper bound is every function $w^{\text{upp}} : \mathcal{P}^n \to \mathbb{R}$ satisfying $w^{\text{upp}}(\mathbf{e}_i) \geq 0$ and

$$\begin{aligned} w^{\text{upp}}(\theta\mathbf{c}^{(1)} + (1-\theta)\mathbf{c}^{(2)}) \;\geq\; & \theta w^{\text{upp}}(\mathbf{c}^{(1)}) + (1-\theta)w^{\text{upp}}(\mathbf{c}^{(2)}) \\ & +\theta(1-\theta)\psi(\mathbf{c}^{(1)}-\mathbf{c}^{(2)}) \end{aligned}$$

for any two $\mathbf{c}^{(1)}, \mathbf{c}^{(2)} \in \mathcal{P}^n$.

The theory for computing the bounds mentioned above uses $\mathbb{H}$–measures for minimizing sequences of optimal arrangements of the phase mixture as was proposed in [17, 28]. The lower estimate is obtained by relaxing the set of $\mathbb{H}$–measures whereas the upper estimate relies on the mixing formula for laminated $\mathbb{H}$–measures. Since these are only bounds, there is interest in estimating their preciseness. To this end we pursue two avenues in this paper: (1) we examine some special cases analytically to find subsets of $\mathcal{P}^n$ where some of the bounds are exact and (2) we consider the numerical computation of the bounds for the four variant case. With respect to analytical results, we are able show that the function $w_{\text{mix}} : P^n \to \mathbb{R}$ takes the so-called Reuß form,

$$w_{\text{mix}}(\mathbf{c}) = w_{\text{Reuß}}(\mathbf{c}) = -\frac{1}{2}\sum_{j=1}^{n} c_j\boldsymbol{\varepsilon}_j{:}\mathbb{C}{:}\boldsymbol{\varepsilon}_j + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} c_j c_k \boldsymbol{\varepsilon}_j{:}\mathbb{C}{:}\boldsymbol{\varepsilon}_k$$

whenever $\mathbf{c} \in \mathcal{P}^n$ can be expressed as $\mathbf{c} = \theta\mathbf{c}^{(1)}+(1-\theta)\mathbf{c}^{(2)}$ with $\theta \in (0,1)$ and the points $\mathbf{c}^{(l)}$ satisfy $w_{\text{mix}}(\mathbf{c}^{(l)}) = w_{\text{Reuß}}(\mathbf{c}^{(l)})$ for $l \in \{1,2\}$ and their averaged transformation strains are symmetrically rank-one connected (compatible), that is $\boldsymbol{\varepsilon}^{\boldsymbol{\eta}} = \mathbf{a}\otimes_{\text{sym}}\mathbf{b}$ where $\boldsymbol{\eta} = \mathbf{c}^{(2)}-\mathbf{c}^{(1)}$, see Corollary 5.4. The importance of this result emanates from the fact that the set of such phase fractions is rather large. Thus, we have a closed form approximation for the free energy of mixing in the n-variant case. An assessment of the accuracy of the Reuß bound is made by computing in the 4-variant case an upper bound and an improved lower bound utilizing a relaxed set of $\mathbb{H}$–measures. We observe from this exercise that the Reuß bound provides a computationally useful approximation to the actual mixing energy. The predictive power of using this bound in evolutionary models has been demonstrated in [12, 14] through comparison to the single crystal experiments of Shield [27]. Here we present an illustrative example employing the evolutionary model in [14].

2 Quasi-Convex Relaxation at Fixed Volume Fraction

As a point of departure consider n stored–energy densities $W(\mathbf{F}, \mathbf{e}_j)$, $j = 1, \dots, n$, which are assumed to be quasi-convex in their first argument, where $\mathbf{F}$ is the deformation gradient. An appropriate stored–energy density of the material is then given as:

$$W(\mathbf{F}) = \min_{i=1,\dots,n} [W(\mathbf{F}, \mathbf{e}_i)] \,. \tag{4}$$

As noted in the introduction, this energy density is not quasi-convex and thus does not lead to a wlsc potential energy. By relaxing this energy density we can recover the essential property of wlsc. The appropriate relaxation in this case is a quasi-convex relaxation which for given volume fraction $\mathbf{c} \in \mathcal{P}^n$ is defined via

$$QW(\mathbf{F}, \mathbf{c}) = \inf \Big\{ \int_{Q^d} W(\mathbf{F} + \nabla\boldsymbol{\varphi}(\mathbf{y}), \boldsymbol{\chi}(\mathbf{y}))\, \mathrm{d}\mathbf{y} \mid \boldsymbol{\varphi} \in W^{1,p}_{\mathrm{per}}(Q^d), \\ \boldsymbol{\chi}(\mathbf{y}) \in \mathcal{P}^n_{\mathrm{pure}}, \int_{Q^d} \boldsymbol{\chi}(\mathbf{y})\, \mathrm{d}\mathbf{y} = \mathbf{c} \Big\}, \tag{5}$$

where $Q^d = (0,1)^d \subset \mathbb{R}^d$. It can be shown (cf. [23]) that for each $\mathbf{c} \in \mathcal{P}^n$ the function $\mathbf{F} \mapsto QW(\mathbf{F}, \mathbf{c})$ is quasi-convex and for each $\mathbf{F}$ the function $\mathbf{c} \mapsto QW(\mathbf{F}, \mathbf{c})$ is convex. Moreover, the quasi-convex hull QW of the function W in (4) is given by $QW(\mathbf{F}) = \min_{\mathbf{c} \in \mathcal{P}^n} QW(\mathbf{F}, \mathbf{c})$, see [17].

In this work we completely restrict our attention to the case of linearized elasticity where additionally each phase has the same elastic tensor $\mathbb{C}$:

$$W(\mathbf{F}, \mathbf{e}_i) \approx W(\boldsymbol{\varepsilon}, \mathbf{e}_i) = \frac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_i){:}\mathbb{C}{:}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_i) + \alpha_i \,. \tag{6}$$

As mentioned above $\boldsymbol{\varepsilon} = \nabla_{\mathrm{sym}} \mathbf{u}$ is the linearized strain, $\boldsymbol{\varepsilon}_i$ are the transformation strains, and α_i are the heights of the wells which usually depends on the temperature which is assumed to be constant here.

Proposition 2.1 *If all $W(\mathbf{F}, \mathbf{e}_j)$, $j = 1, \dots, n$, have the form (6) then the relaxed energy density takes the form $QW(\boldsymbol{\varepsilon}, \mathbf{c}) = \sum_{j=1}^n c_j W(\boldsymbol{\varepsilon}, \mathbf{e}_j) + w_{\mathrm{mix}}(\mathbf{c})$, where the mixture term $w_{\mathrm{mix}} : \mathcal{P}^n \to \mathbb{R}$ is convex and given by*

$$w_{\mathrm{mix}}(\mathbf{c}) = \inf \Big\{ \int_{Q^d} \nabla_{\mathrm{sym}}\boldsymbol{\varphi}{:}\mathbb{C}{:}\big(\tfrac{1}{2}\nabla_{\mathrm{sym}}\boldsymbol{\varphi} - \textstyle\sum_{j=1}^n \chi_j(y)\boldsymbol{\varepsilon}_j\big)\, \mathrm{d}\mathbf{y} \mid \\ \int_{Q^d} \boldsymbol{\chi}(\mathbf{y})\, \mathrm{d}\mathbf{y} = \mathbf{c}, \boldsymbol{\chi}(\mathbf{y}) \in \mathcal{P}^n_{\mathrm{pure}}, \boldsymbol{\varphi} \in W^{1,2}_{\mathrm{per}}(Q^d) \Big\}. \tag{7}$$

This form follows immediately using the quadratic structure of the energy density, $\int \nabla_{\text{sym}}\boldsymbol{\varphi}\,\mathrm{d}\mathbf{y} = \mathbf{0}$, $\int \boldsymbol{\chi}(\mathbf{y})\,\mathrm{d}\mathbf{y} = \mathbf{c}$ and $\sum_{j=1}^{n} \chi_j(\mathbf{y}) \equiv 1$.

Since $\boldsymbol{\varphi}$ appears quadratically in $w_{\min}(\mathbf{c})$ we may eliminate it by first minimizing with respect to $\boldsymbol{\varphi}$ and keeping the phase indicator field $\boldsymbol{\chi}$ fixed. This is done most easily using Fourier series

$$\boldsymbol{\varphi}(\mathbf{y}) = \sum_{\boldsymbol{\xi}\in\Gamma_*} \boldsymbol{\varphi}_{\boldsymbol{\xi}} \mathrm{e}^{\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{y}}, \quad \boldsymbol{\chi}(\mathbf{y}) = \sum_{\boldsymbol{\xi}\in\Gamma} \boldsymbol{\chi}_{\boldsymbol{\xi}} \mathrm{e}^{\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{y}}, \quad \boldsymbol{\chi}_0 = \mathbf{c},$$

where $\Gamma = (2\pi\mathbb{Z})^d$ and $\Gamma_* = \Gamma \setminus \{\mathbf{0}\}$. Inserting this form into $w_{\text{mix}}(\mathbf{c})$ we find a decoupling between different $\boldsymbol{\xi}$'s and thus $\boldsymbol{\varphi}_{\boldsymbol{\xi}} = \overline{\boldsymbol{\varphi}}_{-\xi} \in \mathbb{C}^d$ can be found by minimizing the following term with respect to the Fourier coefficients, $\boldsymbol{\varphi}_{\boldsymbol{\xi}}$,

$$\frac{1}{2}[\mathrm{i}\boldsymbol{\xi}\otimes_{\text{sym}}\boldsymbol{\varphi}_{\boldsymbol{\xi}}]{:}\mathbb{C}{:}[\mathrm{i}\boldsymbol{\xi}\otimes_{\text{sym}}\boldsymbol{\varphi}_{\boldsymbol{\xi}}] - \operatorname{Re}\left\{[\mathrm{i}\boldsymbol{\xi}\otimes_{\text{sym}}\boldsymbol{\varphi}_{\boldsymbol{\xi}}]{:}\mathbb{C}{:}[\sum_{j=1}^{n}(\chi_{\boldsymbol{\xi}})_j\boldsymbol{\varepsilon}_j]\right\}. \tag{8}$$

After some elementary calculations we arrive at

$$w_{\text{mix}}(\mathbf{c}) = \inf\{\, I(\boldsymbol{\chi}) \mid \int_{Q^d} \boldsymbol{\chi}(\mathbf{y})\,\mathrm{d}\mathbf{y} = \mathbf{c}, \boldsymbol{\chi}(\mathbf{y}) \in \mathcal{P}^n_{\text{pure}}\}$$

with

$$I(\boldsymbol{\chi}) = -\frac{1}{2}\sum_{\boldsymbol{\xi}\in\Gamma_*}\left(\boldsymbol{\xi}\cdot\mathbb{C}{:}[\sum_{j=1}^{n}(\chi_{\boldsymbol{\xi}})_j\boldsymbol{\varepsilon}_j]\right)\cdot\mathbf{T}(\boldsymbol{\xi})^{-1}\left(\boldsymbol{\xi}\cdot\mathbb{C}{:}[\sum_{j=1}^{n}(\chi_{\boldsymbol{\xi}})_j\boldsymbol{\varepsilon}_j]\right).$$

Here the acoustic tensor $\mathbf{T}(\boldsymbol{\xi})$ is defined via $\mathbf{u}\cdot\mathbf{T}(\boldsymbol{\xi})\mathbf{u} = [\boldsymbol{\xi}\otimes_{\text{sym}}\mathbf{u}]{:}\mathbb{C}{:}[\boldsymbol{\xi}\otimes_{\text{sym}}\mathbf{u}]$ and thus is homogeneous of degree 2 in $\boldsymbol{\xi} \in \mathbb{R}^d$. Introducing $\widehat{\mathbf{G}}, \mathbf{G}(\boldsymbol{\xi}) \in \mathbb{R}^{n\times n}_{\text{sym}}$ via

$$G(\boldsymbol{\xi})_{jk} = \frac{1}{2}(\boldsymbol{\xi}\cdot\mathbb{C}{:}\boldsymbol{\varepsilon}_j)\cdot\mathbf{T}(\xi)^{-1}(\boldsymbol{\xi}\cdot\mathbb{C}{:}\boldsymbol{\varepsilon}_k), \quad \widehat{G}_{jk} = \frac{1}{2}\boldsymbol{\varepsilon}_j{:}\mathbb{C}{:}\boldsymbol{\varepsilon}_k$$

we obtain the compact formula

$$I(\boldsymbol{\chi}) = -\sum_{\boldsymbol{\xi}\in\Gamma_*}\mathbf{G}(\boldsymbol{\xi}){:}(\overline{\boldsymbol{\chi}}_{\boldsymbol{\xi}}\otimes\boldsymbol{\chi}_{\boldsymbol{\xi}}).$$

The computation of w_{mix} from this form of $I(\boldsymbol{\chi})$ is non-trivial due to the complicated character of the optimization constraints associated with the Fourier transform of the microstructure indicator field $\boldsymbol{\chi}$. In the next section, we review the notion of $\mathbb{H}$–measures to re-characterize this problem and discuss two useful means of approaching w_{mix} in an approximate sense.

3 The $\mathbb{H}$–measure associated with the micro-structure χ

The function $\chi : Q^d \to \mathcal{P}^n_{\text{pure}}$ defines a microstructure of pure phases having the volume average $\mathbf{c} = \int_{Q^d} \chi(\mathbf{y})\,\mathrm{dy}$. We associate to each function χ a matrix–valued measure $\boldsymbol{\mu} = \widetilde{\boldsymbol{\mu}}(\chi)$ on the unit sphere $\mathbb{S}^{d-1}$. The values of $\boldsymbol{\mu}$ will lie in the set $\mathbb{H}^n_{\geq 0}$ of positive semidefinite Hermitian matrices. For any measurable $\Sigma \subset \mathbb{S}^{d-1}$ we let

$$\boldsymbol{\mu}(\Sigma) = \sum_{\xi \in \Gamma_*,\, \xi/|\xi| \in \Sigma} \chi_\xi \otimes \overline{\chi}_\xi \in \mathbb{H}^n_{\geq 0}.$$

The measure is a homogeneous $\mathbb{H}$–measure in the sense of [29] when the generating sequence $\mathbf{u}^{(n)}(\mathbf{y}) = \chi(n\mathbf{y}) - \mathbf{c}$ is considered.

Introducing the linear functional

$$\begin{aligned} &J : \text{meas}(\mathbb{S}^{d-1}, \mathbb{H}^n_{\geq 0}) \to \mathbb{R} \\ &\boldsymbol{\mu} \mapsto J(\boldsymbol{\mu}) = -\int_{\omega \in \mathbb{S}^{d-1}} \mathbf{G}(\omega) : \boldsymbol{\mu}(\mathrm{d}\omega) \end{aligned} \tag{9}$$

we obtain the relation $I(\chi) = J(\widehat{\boldsymbol{\mu}}(\chi))$, where $\widehat{\boldsymbol{\mu}}(\chi) = \mathrm{Re}\widetilde{\boldsymbol{\mu}}(\chi)$. This follows from the symmetry $\mathbf{G}(-\omega) = \mathbf{G}(\omega)$ which tells us that J depends only on the real part of $\boldsymbol{\mu}$. We formally define $\widehat{\boldsymbol{\mu}}(\cdot)$ as:

$$\begin{aligned} &\widehat{\boldsymbol{\mu}} : \mathrm{L}^\infty(Q^d, \mathcal{P}^n_{\text{pure}}) \to \text{meas}(\mathbb{S}^{d-1}, \mathbb{P}^n) \\ &\chi \mapsto \widehat{\boldsymbol{\mu}}(\chi) = \mathrm{Re}\widetilde{\boldsymbol{\mu}}(\chi), \end{aligned}$$

where $\mathbb{P}^n = \{\, \mathbf{B} \in \mathbb{R}^{n \times n} \mid \mathbf{B} = \mathbf{B}^T,\ \mathbf{B} \text{ positive semidefinite} \,\}$. It then follows that each $\boldsymbol{\mu} = \widehat{\boldsymbol{\mu}}(\chi)$ satisfies

$$\begin{array}{ll} \text{(i)} & \boldsymbol{\mu}(\Sigma)\mathbf{e}_* = \mathbf{0} \text{ for all } \Sigma \subset \mathbb{S}^{d-1} \\ \text{(ii)} & \boldsymbol{\mu}(\mathbb{S}^{d-1}) = \mathbf{M}(\mathbf{c}) \text{ where } \mathbf{c} = \int \chi(\mathbf{y})\,\mathrm{dy} \\ \text{(iii)} & \boldsymbol{\mu}(-\Sigma) = \boldsymbol{\mu}(\Sigma) \in \mathbb{P}^n \text{ for all } \Sigma \subset \mathbb{S}^{d-1}. \end{array} \tag{10}$$

For each $\mathbf{c} \in \mathcal{P}^n$ we define the subsets of matrix–valued measures

$$\begin{aligned} \mathcal{H}(\mathbf{c}) &= \text{closure}\{\, \widehat{\boldsymbol{\mu}}(\chi) \mid \textstyle\int_{Q^d} \chi(\mathbf{y})\,\mathrm{dy} = \mathbf{c} \,\}, \\ \overline{\mathcal{H}}(\mathbf{c}) &= \{\, \boldsymbol{\mu} \in \text{meas}(\mathbb{S}^{d-1}, \mathbb{P}^n) \mid (10) \text{ holds} \,\}. \end{aligned}$$

Clearly, $\mathcal{H}(\mathbf{c}) \subset \overline{\mathcal{H}}(\mathbf{c})$ and $\overline{\mathcal{H}}(\mathbf{c})$ is convex. The function w_{mix} can now be expressed as

$$w_{\text{mix}}(\mathbf{c}) = \inf\{ J(\boldsymbol{\mu}) \mid \boldsymbol{\mu} \in \mathcal{H}(\mathbf{c}) \}. \tag{11}$$

The difficulty in calculating $w_{\text{mix}}(\mathbf{c})$ is the same as in calculating $\mathcal{H}(\mathbf{c})$. There is at present no good characterizations of $\mathcal{H}(\mathbf{c})$. To make progress one is faced with a situation where only bounds can be computed. Constructing $\mathcal{H}^{\text{inn}}$ and $\mathcal{H}^{\text{out}}$ with $\mathcal{H}^{\text{inn}}(\mathbf{c}) \subset \mathcal{H}(\mathbf{c}) \subset \mathcal{H}^{\text{out}}(\mathbf{c})$, we obtain upper and lower bounds for w_{mix} if $\mathcal{H}(\mathbf{c})$ in the infimum (11) is replaced by $\mathcal{H}^{\text{inn}}(\mathbf{c})$ and $\mathcal{H}^{\text{out}}(\mathbf{c})$, respectively. In section 4 the lamination mixture formula for $\mathbb{H}$–measures (see Theorem 4.1) is employed to find upper bounds. In section 5.1 we will use $\mathcal{H}^{\text{out}} = \overline{\mathcal{H}}(\mathbf{c})$ to obtain a lower bound.

4 Upper Bounds

Since w_{mix} is non-positive we have the trivial bound $w_{\text{mix}}(c) \leq 0$; to improve this bound we need to find a suitable inner approximation of $\mathcal{H}(\mathbf{c})$. In fact, in light of the Krein–Milman theorem it is sufficient to characterize (or approximate) only extremal points of such a subset.

There is no general theory to describe $\mathcal{H}(\mathbf{c})$ suitably, however the *lamination mixture formula*, which is due to [1, 29, 17], allows us to construct new $\mathbb{H}$–measures as mixture of two given ones.

Theorem 4.1 *Assume $\boldsymbol{\mu}^{(1)} \in \mathcal{H}(\mathbf{c}^{(1)}), \boldsymbol{\mu}^{(2)} \in \mathcal{H}(\mathbf{c}^{(2)}), \theta \in [0,1]$ and $\boldsymbol{\omega} \in \mathbb{S}^{d-1}$, then there exists a $\boldsymbol{\mu} \in \mathcal{H}(\theta\mathbf{c}^{(1)}+(1-\theta)\mathbf{c}^{(2)})$ of the form*

$$\boldsymbol{\mu} = \theta\boldsymbol{\mu}^{(1)} + (1-\theta)\boldsymbol{\mu}^{(2)} + \theta(1-\theta)(\mathbf{c}^{(2)}-\mathbf{c}^{(1)})\otimes(\mathbf{c}^{(2)}-\mathbf{c}^{(1)})\widehat{\boldsymbol{\delta}}_{\omega}. \tag{12}$$

Note that the statement in (12) is consistent with

$$\begin{aligned} \mathbf{M}(\theta\mathbf{c}^{(1)}+(1-\theta)\mathbf{c}^{(2)}) &= \theta\mathbf{M}(\mathbf{c}^{(1)}) + (1-\theta)\mathbf{M}(\mathbf{c}^{(2)}) \\ &\quad +\theta(1-\theta)(\mathbf{c}^{(2)}-\mathbf{c}^{(1)})\otimes(\mathbf{c}^{(2)}-\mathbf{c}^{(1)}). \end{aligned} \tag{13}$$

This theorem implies convexity of $\mathcal{H}(\mathbf{c})$, since mixing with $\mathbf{c}^{(1)} = \mathbf{c}^{(2)} = \mathbf{c}$ does not generate an additional quadratic term. Another immediate consequence is the estimate

$$\begin{aligned} w_{\text{mix}}(\theta\mathbf{c}^{(1)}+(1-\theta)\mathbf{c}^{(2)}) &\leq \theta w_{\text{mix}}(\mathbf{c}^{(1)}) + (1-\theta)w_{\text{mix}}(\mathbf{c}^{(2)}) \\ &\quad +\theta(1-\theta)\psi(\mathbf{c}^{(1)}-\mathbf{c}^{(2)}), \end{aligned} \tag{14}$$

where $\psi : \mathbb{R}^n \to \mathbb{R}$ is given as

$$\psi(\boldsymbol{\eta}) = \inf\{ -\mathbf{G}(\boldsymbol{\omega}):(\boldsymbol{\eta}\otimes\boldsymbol{\eta}) \mid \boldsymbol{\omega} \in \mathbb{S}^{d-1} \}.$$

Note this expression is fully compatible with the relation given in Section 1 as Eq. (2).

To see the validity of (14) choose $\boldsymbol{\mu}^{(j)} \in \mathcal{H}(\mathbf{c}^{(j)})$ with $w_{\text{mix}}(\mathbf{c}^{(j)}) = J(\boldsymbol{\mu}^{(j)})$ for $j = 1$ and 2. Further choose $\boldsymbol{\omega} \in \mathbb{S}^{d-1}$ such that $\psi(\mathbf{c}^{(1)}-\mathbf{c}^{(2)}) = -\mathbf{G}(\boldsymbol{\omega}):[(\mathbf{c}^{(1)}-\mathbf{c}^{(2)})\otimes(\mathbf{c}^{(1)}-\mathbf{c}^{(2)})]$. Then, $\boldsymbol{\mu}$ defined in (12) lies in $\mathcal{H}(\mathbf{c})$ and (11) gives (14). As $\psi(\boldsymbol{\eta}) \leq 0$, this implies better bounds than simple convexity in most cases. Starting with $\mathbf{0} \in \mathcal{H}(\mathbf{e}_j)$ we may inductively apply this formula to obtain nontrivial bounds.

It was first observed by Kohn [17] that this approach gives an exact expression for w_{mix} in the case of two phases only, viz.

$$w_{\text{mix}}(\theta\mathbf{e}_j + (1-\theta)\mathbf{e}_k) = \theta(1-\theta)\psi(\mathbf{e}_j-\mathbf{e}_k). \tag{15}$$

This result follows immediately since the lower and upper bounds for $w_{\text{mix}}(\mathbf{c})$ coincide. Our theorem 6.2 generalizes this fact to the case of n phases, if all transformation strains lie on a single straight line.

By applying (14) with $\mathbf{c}^{(1)} = \theta\mathbf{e}_j + (1-\theta)\mathbf{e}_k$ and $\mathbf{c}^{(2)} = \mathbf{e}_l$ we obtain a bound for all points in $\text{conv}\{\mathbf{e}_j, \mathbf{e}_k, \mathbf{e}_l\}$. Proceeding like this we obtain a first upper bound.

Proposition 4.2 *For* $\mathbf{c} \in \mathcal{P}^n$ *define* $\mathbf{c}^{(k)} = (c_1+\ldots+c_k)^{-1}(c_1\mathbf{e}_1+\ldots+c_k\mathbf{e}_k)$, *then*

$$w_{\text{mix}}(\mathbf{c}) \leq \sum_{j=2}^{n} \tfrac{(c_1+\ldots+c_{j-1})c_j}{c_1+\ldots+c_j}\, \psi(\mathbf{c}^{(j-1)}-\mathbf{c}^{(j)}). \tag{16}$$

By permutation of the points $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *we can generate other upper bounds.*

Of course, the upper bound in the right–hand side of (16) may be further reduced by applying (14) again. In fact, in typical cases it needs infinitely many applications of the lamination formula to reach the lowest level. This corresponds to infinite sequential lamination, see [28].

In principle, the definition $w^{\text{upp}}(\mathbf{c}) = \max\{ \widetilde{w}(\mathbf{c}) \mid \widetilde{w}$ satisfies (14), $\widetilde{w}(\mathbf{e}_j) \leq 0,\ j = 1,\ldots,n \}$ defines the best upper bound for $w_{\text{mix}}(\mathbf{c})$ which can be obtained from lamination. For practical purposes it is better to write w^{upp} as a minimum, since then each candidate $\overline{w}$ provides a true upper bound

$$w^{\text{upp}}(\mathbf{c}) = \min\{ \overline{w}(\mathbf{c}) \mid \overline{w} \text{ satisfies (17)},\ \overline{w}(\mathbf{e}_j) \geq 0 \text{ for } j = 1,\ldots,n \},$$

where now the condition reads

$$\overline{w}(\theta\mathbf{c}^{(1)}+(1-\theta)\mathbf{c}^{(2)}) \geq \theta\overline{w}(\mathbf{c}^{(1)}) + (1-\theta)\overline{w}(\mathbf{c}^{(2)}) + \theta(1-\theta)\psi(\mathbf{c}^{(1)}-\mathbf{c}^{(2)}). \quad (17)$$

In Section 6 we will utilize this lamination bound to compute the mixing energy for a four variant material and compare it to the lower bounds to be discussed next.

5 Lower bounds

In this section we consider two lower bounds to the mixing energy. First we construct a lower bound by looking at a particular outer approximation to $\mathcal{H}(\mathbf{c})$. Second we consider a further relaxation of the problem by constructing a lower bound directly from the quasi-convex relaxation process by ignoring the issue of compatibility. It will be seen that this bound directly corresponds to the notion of a stress-ensemble and thus this second lower bound is termed a Reuß bound. The corresponding Taylor bound which is obtained by assuming constant strain leads to an uninteresting upper bound $w_{\text{mix}}(\mathbf{c}) \leq 0$, which follows trivially from $w_{\text{mix}}(\mathbf{e}_j) = 0$ and convexity.

5.1 $\mathbb{H}$-measure lower bound

To find a lower bound consider $\mathcal{H}(\mathbf{c}) \subset \overline{\mathcal{H}}(\mathbf{c})$ so that one may write

$$w_{\text{mix}}(\mathbf{c}) \geq w^{\text{low}}(\mathbf{c}) = \inf\{ -\int_{\mathbb{S}^{d-1}} \mathbf{G}(\boldsymbol{\omega}):\boldsymbol{\mu}(\mathrm{d}\boldsymbol{\omega}) \mid \boldsymbol{\mu} \in \overline{\mathcal{H}}(\mathbf{c}) \}.$$

It is possible to characterize $\overline{\mathcal{H}}(c)$ solely by algebraic conditions; thus it is easier to calculate $w^{\text{low}}(\mathbf{c})$ than $w_{\text{mix}}(\mathbf{c})$. In particular, following [28] we will show that the calculation of $w^{\text{low}}(\mathbf{c})$ can be reduced to a finite dimensional minimization problem.

Since the mapping $J : \text{meas}(\mathbb{S}^{d-1}, \mathbb{H}^m_{\geq 0}) \to \mathbb{R}$ given in (9) is linear and since $\overline{\mathcal{H}}(\mathbf{c})$ is convex, closed and bounded, we can apply the Krein–Milman theorem which states that J attains its minimum on the extremal points of $\overline{\mathcal{H}}(c)$:

$$\text{ex}(\overline{\mathcal{H}}(\mathbf{c})) = \{ \mu \in \overline{\mathcal{H}}(\mathbf{c}) \mid \overline{\mathcal{H}} \setminus \{\mu\} \text{ is convex} \}.$$

Recalling $m = (n^2-n)/2$ and using the Dirac measure $\boldsymbol{\delta}_\omega \in \text{meas}(\mathbb{S}^{d-1}, \mathbb{R})$ this set can be characterized as follows.

Proposition 5.1 *We have* $\boldsymbol{\mu} \in \mathrm{ex}(\overline{\mathcal{H}}(\mathbf{c}))$ *if and only if there exist directions* $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_m \in \mathbb{S}^{d-1}$ *and* $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_m \in \mathbb{R}^n_*$ *such that* $\boldsymbol{\mu} = \sum_{j=1}^m \widehat{\boldsymbol{\delta}}_{\boldsymbol{\omega}_j} \boldsymbol{\eta}_j \otimes \boldsymbol{\eta}_j$, *where* $\widehat{\boldsymbol{\delta}}_\omega = \frac{1}{2}(\boldsymbol{\delta}_\omega + \boldsymbol{\delta}_{-\omega})$.

Proof. Assume that $\mathbb{S}^{d-1}$ contains $m+1$ pairwise disjoint symmetric Σ_j (i.e. $\Sigma_j = -\Sigma_j$), such that $\boldsymbol{\mu}(\Sigma_j) \neq 0$. Since all $\boldsymbol{\mu}(\Sigma_j)$ lie in the m–dimensional linear space $\{\mathrm{M} \in \mathbb{R}^{n\times n}_{\mathrm{sym}} \mid \mathrm{M}\mathbf{e}_* = \mathbf{0}\}$ the equation $\sum_{j=1}^{m+1} \alpha_j \boldsymbol{\mu}(\Sigma_j) = \mathbf{0}$ has a nontrivial solution $(\alpha_1, \ldots, \alpha_{m+1}) \in \mathbb{R}^{m+1} \setminus \{\mathbf{0}\}$. We may assume $|\alpha_j| \leq 1$ and define the two measures $\boldsymbol{\mu}^\pm \in \overline{\mathcal{H}}(\mathbf{c})$ via $\boldsymbol{\mu}^\pm(\Sigma) = (1 \pm \alpha_j)\boldsymbol{\mu}(\Sigma)$ if $\Sigma \subset \Sigma_j$ and $\boldsymbol{\mu}^\pm(\Sigma) = \boldsymbol{\mu}(\Sigma)$ if $\Sigma \cap \bigcup_{j=1}^m \Sigma_j = \emptyset$. Clearly, $\boldsymbol{\mu}^+ \neq \boldsymbol{\mu}^-$ and $\boldsymbol{\mu} = \frac{1}{2}(\boldsymbol{\mu}^+ + \boldsymbol{\mu}^-)$. Thus, $\boldsymbol{\mu}$ cannot be extremal. We conclude that extremal measures take the form $\boldsymbol{\mu} = \sum_{j=1}^m \widehat{\boldsymbol{\delta}}_{\omega_j} \mathrm{M}_j$ with $\mathrm{M}_j \in \mathbb{P}^n_* = \{\mathrm{M} \in \mathbb{P}^n \mid \mathrm{M}\mathbf{e}_* = \mathbf{0}\}$.

It remains to show that $\operatorname{rank} \mathrm{M}_j \leq 1$. Each M_j has the form $\mathrm{M}_j = \sum_{i=1}^{r_j} \boldsymbol{\eta}_{j,i} \otimes \boldsymbol{\eta}_{j,i}$ where $r_j = \operatorname{rank} \mathrm{M}_j$ and $\boldsymbol{\eta}_{j,i} \neq \mathbf{0}$. To find a contradiction we assume $\sum_{j=1}^m r_j > m$, then as above $\sum_{j=1}^m \sum_{i=1}^{r_j} \alpha_{j,i} \boldsymbol{\eta}_{j,i} \otimes \boldsymbol{\eta}_{j,i} = \mathbf{0}$ has a nontrivial solution with $|\alpha_{j,i}| \leq 1$. The measures $\boldsymbol{\mu}^\pm = \sum_{j=1}^m \sum_{i=1}^{r_j} (1 \pm \alpha_{j,i}) \widehat{\boldsymbol{\delta}}_{\omega_j} \boldsymbol{\eta}_{j,i} \otimes \boldsymbol{\eta}_{j,i}$ are different, lie in $\overline{\mathcal{H}}(\mathbf{c})$ and satisfy $\boldsymbol{\mu} = \frac{1}{2}(\boldsymbol{\mu}^+ + \boldsymbol{\mu}^-)$. Thus, $\boldsymbol{\mu}$ cannot be an extremal measure which is the desired contradiction. ∎

To formulate the main result of this section we define $\psi : \mathbb{R}^n \to \mathbb{R}$ via

$$\psi(\boldsymbol{\eta}) = \inf\{\, -\mathbf{G}(\boldsymbol{\omega}) : (\boldsymbol{\eta} \otimes \boldsymbol{\eta}) \mid \boldsymbol{\omega} \in \mathbb{S}^{d-1} \}.$$

Moreover, introduce the sets $\mathbb{E}^n_* = \{\boldsymbol{\eta} \otimes \boldsymbol{\eta} \in \mathbb{P}^n_* \mid \boldsymbol{\eta} \in \mathbb{S}^{n-1}, \boldsymbol{\eta} \cdot \mathbf{e}_* = 0\}$ and $\mathbb{K}^n_* = \{\mathrm{M} \in \mathbb{P}^n_* \mid \operatorname{tr}[\mathrm{M}] = 1\}$. Then $\mathbb{K}^n_* = \operatorname{conv}(\mathbb{E}^n_*)$ and $\mathbb{E}^n_* = \operatorname{ex}(\mathbb{K}^n_*)$. The set $\mathbb{E}^n_*$ is a smooth manifold of dimension $(n-2)$. The set $\mathbb{K}^n_*$ lies in an affine subspace of $\mathbb{R}^{n\times n}$ of dimension $(n^2 - n)/2 - 1$ and has nonempty interior with respect to this subspace.

Finally we define $\Psi : \mathbb{K}^n_* \to \mathbb{R}$ to be the largest convex function which satisfies $\Psi(\boldsymbol{\eta} \otimes \boldsymbol{\eta}) = \psi(\boldsymbol{\eta})$ for $\boldsymbol{\eta} \otimes \boldsymbol{\eta} \in \mathbb{E}^n_*$, viz.

$$\begin{aligned}\Psi(\mathrm{M}) &= \min\Big\{ \sum_{j=1}^m \alpha_j \psi(\boldsymbol{\eta}_j) \mid \alpha_j \geq 0, \sum_{j=1}^m \alpha_j = 1, \boldsymbol{\eta}_j \otimes \boldsymbol{\eta}_j \in \mathbb{E}^n_*, \quad (18)\\ &\qquad \sum_{j=1}^m \alpha_j \boldsymbol{\eta}_j \otimes \boldsymbol{\eta}_j = \mathrm{M} \Big\}\\ &= \min\Big\{ \sum_{j=1}^m \psi(\boldsymbol{\eta}_j) \mid \boldsymbol{\eta}_j \in \mathbb{R}^n_*, \sum_{j=1}^m \boldsymbol{\eta}_j \otimes \boldsymbol{\eta}_j = \mathrm{M} \Big\} \quad (19)\end{aligned}$$

Here again $m = (n^2-n)/2$ by the standard theory of finite-dimensional convexity and the fact that $\mathbb{K}_*^n$ lies in a $(m-1)$-dimensional affine subspace of $\mathbb{R}^{n\times n}$. In fact, we may extend ψ to $\psi : \mathbb{R}_*^n \to \mathbb{R}$ naturally such that $\psi(\alpha\boldsymbol{\eta}) = \alpha^2\psi(\boldsymbol{\eta})$, and similarly we may extend Ψ into $\Psi : \mathbb{P}_*^n \to \mathbb{R}$ such that $\Psi(\alpha\mathrm{M}) = \alpha\Psi(\mathrm{M})$ for $\alpha \geq 0$.

Theorem 5.2 *We have the formula* $w^{\mathrm{low}}(\mathbf{c}) = \Psi(\mathbf{M}(\mathbf{c}))$.

The proof follows directly from the Krein–Milman theorem and the definition of Ψ.

It should be noted that we do not need the function Ψ on the whole set $\mathbb{K}_*^n$. Since

$$\mathbf{M}(\mathbf{c}) = \mathrm{diag}(\mathbf{c})-\mathbf{c}\otimes\mathbf{c} = \frac{1}{2}\sum_{i,j=1}^{n} c_i c_j(\mathrm{e}_i-\mathrm{e}_j)\otimes(\mathrm{e}_i-\mathrm{e}_j), \tag{20}$$

we have $[\,\mathrm{tr}\,\mathbf{M}(\mathbf{c})]^{-1}\mathbf{M}(\mathbf{c}) \in \mathrm{conv}\{\,\frac{1}{2}(\mathrm{e}_i-\mathrm{e}_j)\otimes(\mathrm{e}_i-\mathrm{e}_j) \in \mathbb{E}_*^n : i \neq j\,\} \subset \mathbb{K}_*^n$. In the case $n = 3$ the dimension $m-1$ equals 2 and $\mathbb{K}_*^n$ can be identified with a circle, see [28]. Formula (19) [cf. also (3)] is obtained by extending (18) homogeneously.

5.2 The Reuß bound

Above we discussed the lower bound w^{low} to w_{mix} by considering an outer approximation to $\mathcal{H}(\mathbf{c})$. In the examples section we will examine the quality of this bound. Before doing so however, we will briefly consider another lower bound. The bounds that have been discussed so far have all respected the notion of compatibility of the microstructural displacement field through the explicit presence of $\boldsymbol{\varphi} \in W^{1,2}_{\mathrm{per}}(Q^d)$ in Eq. (7). If however we explicitly relax the compatibility restriction and perform the minimization over the (larger) set of all symmetric gradient fields, then we will obtain a lower bound to the free energy of mixing – a different one than that obtained in the previous section. We call this lower bound the Reuß bound; note $w_{\mathrm{mix}}(\mathbf{c}) \geq w^{\mathrm{low}}(\mathbf{c}) \geq w_{\mathrm{Reuß}}(\mathbf{c})$.

In practical terms we can carry out this minimization by considering the term in (8) and minimizing over $\boldsymbol{\gamma} = \overline{\boldsymbol{\gamma}}^T \in \mathbb{C}^{d\times d}$ instead of over $\mathrm{i}\boldsymbol{\xi}\otimes_{\mathrm{sym}}\boldsymbol{\varphi}_{\boldsymbol{\xi}}$. If we use the fact that $\boldsymbol{\chi}(\mathrm{y}) \in \mathcal{P}^n_{\mathrm{pure}}$ gives $\int_{Q^d} \boldsymbol{\chi}(\mathrm{y})\otimes\boldsymbol{\chi}(\mathrm{y})\,\mathrm{dy} = \sum_{\boldsymbol{\xi}\in\Gamma} \boldsymbol{\chi}_{\boldsymbol{\xi}}\otimes\boldsymbol{\chi}_{-\boldsymbol{\xi}} =$

diag($\mathbf{c}$), then we arrive at the result

$$\begin{aligned} w_{\text{Reuß}}(\mathbf{c}) &= -\widehat{\mathbf{G}}{:}\mathbf{M}(\mathbf{c}) \\ &= -\tfrac{1}{2}\sum_{j=1}^{n} c_j \boldsymbol{\varepsilon}_j{:}\mathbb{C}{:}\boldsymbol{\varepsilon}_j + \tfrac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} c_j c_k \boldsymbol{\varepsilon}_j{:}\mathbb{C}{:}\boldsymbol{\varepsilon}_k \end{aligned}$$

The importance of this bound is that: (1) it is given in explicit form, and (2), as we shall see in this section, it is exact in certain circumstances.

The exactness of the bound is closely related with the notion of compatible phases, this means that their transformation strains are symmetrically rank–one connected (denoted for short: sr1c). This is expressed as $\boldsymbol{\varepsilon}_j - \boldsymbol{\varepsilon}_i = \mathbf{a}\otimes_{\text{sym}}\mathbf{b}$ for some vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ and we recall (cf. e.g. [4])

$$\mathbb{R}^{3\times 3}_{\text{sym}} \ni \mathbf{A} = \mathbf{a}\otimes_{\text{sym}}\mathbf{b} \quad \Longleftrightarrow \quad \text{spec}(\mathbf{A}) = \{\lambda_1, 0, \lambda_3\} \text{ with } \lambda_1 \le 0 \le \lambda_3. \tag{21}$$

The following proposition shows that sr1c transformation strains play a central role as they achieve the lowest possible value for the function ψ.

Proposition 5.3 *If* $\boldsymbol{\varepsilon}^{\boldsymbol{\eta}} = \sum_{j=1}^{n} \eta_j \boldsymbol{\varepsilon}_j$ *satisfies* $\boldsymbol{\varepsilon}^{\boldsymbol{\eta}} = \mathbf{a}\otimes_{\text{sym}}\mathbf{b}$, *then*

$$\psi(\boldsymbol{\eta}) = -\boldsymbol{\eta}\cdot\widehat{\mathbf{G}}\boldsymbol{\eta} = -\widehat{\mathbf{G}}{:}(\boldsymbol{\eta}\otimes\boldsymbol{\eta}) = -\frac{1}{2}\boldsymbol{\varepsilon}^{\boldsymbol{\eta}}{:}\mathbb{C}{:}\boldsymbol{\varepsilon}^{\boldsymbol{\eta}}.$$

Proof. Using the Reuß bound we know $\psi(\boldsymbol{\eta}) \ge -\frac{1}{2}\boldsymbol{\varepsilon}^{\boldsymbol{\eta}}{:}\mathbb{C}{:}\boldsymbol{\varepsilon}^{\boldsymbol{\eta}}$. We define $g(\boldsymbol{\omega}) = -\mathbf{G}{:}(\boldsymbol{\eta}\otimes\boldsymbol{\eta}) = -\boldsymbol{\eta}\cdot\mathbf{G}(\boldsymbol{\omega})\boldsymbol{\eta} = -\frac{1}{2}\left(\boldsymbol{\omega}\cdot\mathbb{C}{:}\boldsymbol{\varepsilon}^{\boldsymbol{\eta}}\right)\cdot\mathbf{T}(\boldsymbol{\omega})^{-1}\left(\boldsymbol{\omega}\cdot\mathbb{C}{:}\boldsymbol{\varepsilon}^{\boldsymbol{\eta}}\right)$, then clearly $\psi(\boldsymbol{\eta}) = \inf\{\, g(\boldsymbol{\omega}) \mid \boldsymbol{\omega}\in\mathbb{S}^{d-1}\,\}$. We now will show that when $\boldsymbol{\omega} = \mathbf{a}/|\mathbf{a}|$ the equality holds and thus the result will be proved.

For the symmetric acoustic tensor $\mathbf{T}(\boldsymbol{\xi})$ we have the identity

$$\mathbf{w}\cdot\mathbf{T}(\boldsymbol{\xi})\mathbf{v} = [\boldsymbol{\xi}\otimes_{\text{sym}}\mathbf{w}]{:}\mathbb{C}{:}[\boldsymbol{\xi}\otimes_{\text{sym}}\mathbf{v}] \quad \text{for all } \boldsymbol{\xi}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^d.$$

For arbitrary $\mathbf{h}\in\mathbb{R}^d$ we choose $\boldsymbol{\xi} = \mathbf{a}$, $\mathbf{v} = \mathbf{b}$ and $\mathbf{w} = \mathbf{T}(\mathbf{a})^{-1}\mathbf{h}$ and find

$$\begin{aligned} \mathbf{b}\cdot\mathbf{h} &= \mathbf{T}(\mathbf{a})\mathbf{b}\cdot\mathbf{T}(\mathbf{a})^{-1}\mathbf{h} = [\mathbf{a}\otimes_{\text{sym}}(\mathbf{T}(\mathbf{a})^{-1}\mathbf{h})]{:}\mathbb{C}{:}[\mathbf{a}\otimes_{\text{sym}}\mathbf{b}] \\ &= [(\mathbf{T}(\mathbf{a})^{-1}\mathbf{h})]\cdot(\mathbf{a}\cdot\mathbb{C}{:}[\mathbf{a}\otimes_{\text{sym}}\mathbf{b}]) \\ &= \mathbf{h}\cdot\left(\mathbf{T}(\mathbf{a})^{-1}(\mathbf{a}\cdot\mathbb{C}{:}[\mathbf{a}\otimes_{\text{sym}}\mathbf{b}])\right). \end{aligned}$$

As $\mathbf{h}$ was arbitrary we conclude $\mathbf{T}(\mathbf{a})^{-1}(\mathbf{a}\cdot\mathbb{C}{:}[\mathbf{a}\otimes_{\text{sym}}\mathbf{b}]) = \mathbf{b}$, and thus, we find

$$\begin{aligned} g(\mathbf{a}/|\mathbf{a}|) &= -\tfrac{1}{2}\left(\mathbf{a}\cdot\mathbb{C}{:}[\mathbf{a}\otimes_{\text{sym}}\mathbf{b}]\right)\cdot\mathbf{T}(\mathbf{a})^{-1}\left(\mathbf{a}\cdot\mathbb{C}{:}[\mathbf{a}\otimes_{\text{sym}}\mathbf{b}]\right) \\ &= -\tfrac{1}{2}\mathbf{b}\cdot(\mathbf{a}\cdot\mathbb{C}{:}[\mathbf{a}\otimes_{\text{sym}}\mathbf{b}]) = -\tfrac{1}{2}[\mathbf{a}\otimes_{\text{sym}}\mathbf{b}]{:}\mathbb{C}{:}[\mathbf{a}\otimes_{\text{sym}}\mathbf{b}] \\ &= -\tfrac{1}{2}\boldsymbol{\varepsilon}^{\boldsymbol{\eta}}{:}\mathbb{C}{:}\boldsymbol{\varepsilon}^{\boldsymbol{\eta}} \end{aligned}$$

and the minimum is attained as desired. ∎

From this we derive a nice corollary concerning the achievement for the Reuß bound. If we assume that the Reuß bound is achieved at two points $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$ and if $\boldsymbol{\varepsilon}^{c^{(1)}}$ and $\boldsymbol{\varepsilon}^{c^{(2)}}$ are sr1c then, the Reuß bound holds in fact on the whole segment between $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$.

Corollary 5.4 *Assume we have $\mathbf{c}^{(j)} \in \mathcal{P}^n$ with $w_{\text{mix}}(\mathbf{c}^{(j)}) = -\widehat{\mathbf{G}}:\mathbf{M}(\mathbf{c}^{(j)})$ for $j = 1, 2$. Moreover, assume $\boldsymbol{\varepsilon}^{c^{(2)}-c^{(1)}} = \boldsymbol{\varepsilon}^{c^{(2)}} - \boldsymbol{\varepsilon}^{c^{(1)}} = \mathbf{a}\otimes_{\text{sym}}\mathbf{b}$. For $\theta \in [0,1]$ let $\mathbf{c}_\theta = \theta\mathbf{c}^{(1)} + (1-\theta)\mathbf{c}^{(2)}$, then*

$$w_{\text{mix}}(\mathbf{c}_\theta) = w_{Reu\beta}(\mathbf{c}_\theta) = -\widehat{\mathbf{G}}:\mathbf{M}(\mathbf{c}_\theta) \quad \text{for} \quad \theta \in [0,1].$$

Proof. We apply (14) and Proposition 5.3 to obtain

$$\begin{aligned} w_{\text{mix}}(\mathbf{c}_\theta) &\leq \theta w_{\text{mix}}(\mathbf{c}^{(1)}) + (1-\theta) w_{\text{mix}}(\mathbf{c}^{(2)}) + \theta(1-\theta)\psi(\mathbf{c}^{(2)}-\mathbf{c}^{(1)}) \\ &= -\widehat{\mathbf{G}}:[\theta\mathbf{M}(\mathbf{c}^{(1)}) + (1-\theta)\mathbf{M}(\mathbf{c}^{(2)}) + \theta(1-\theta)(\mathbf{c}^{(2)}-\mathbf{c}^{(1)})\otimes(\mathbf{c}^{(2)}-\mathbf{c}^{(1)})]. \end{aligned}$$

With (13) we conclude $w_{\text{mix}}(\mathbf{c}_\theta) \leq -\widehat{\mathbf{G}}:\mathbf{M}(\mathbf{c}_\theta)$. However, the opposite estimate always holds since this is also the Reuß bound; the result is thus established. ∎

As a second corollary we obtain a result which seems to be well–known in the community, but it is not stated like this.

Corollary 5.5 *Assume that all transformation strains $\boldsymbol{\varepsilon}_i$ are pairwise sr1c. Then, the mixture function satisfies $w_{\text{mix}}(\mathbf{c}) = w_{Reu\beta}(\mathbf{c}) = -\widehat{\mathbf{G}}:\mathbf{M}(\mathbf{c})$ for all $\mathbf{c} \in \mathcal{P}^n$.*

6 Examples

In this section we present three examples demonstrating the use of the analysis developed to this point. In the first example we consider an analytical application of our developments to a particular case that has appeared in the literature for 3-variants and extend it to the case of n-variants. In the second example, we examine the goodness of the bounds for a particular material that undergoes a cubic to tetragonal phase transformation; thus this is a problem with 4-variants. Lastly, we demonstrate for illustrative purposes the application of the Reuß bound in an equilibrium evolutionary model.

6.1 Co-linear transformation strains

Proposition 6.1 *If $\psi(\boldsymbol{\eta}) = \mathrm{B}{:}(\boldsymbol{\eta}\otimes\boldsymbol{\eta})$ for all $\boldsymbol{\eta} \in \mathbb{E}^n_*$, then, $w_{\mathrm{mix}}(\mathbf{c}) = \mathrm{B}{:}\mathrm{M}(\mathbf{c})$.*

To prove this we first note that $w^{\mathrm{low}}(\mathbf{c}) = \mathbf{B}{:}\mathbf{M}(\mathbf{c})$ since in (3) every linear combination gives the same result by linearity. Similarly the upper bound (16) gives the same result. In fact, the assumption can be weakened severely by assuming $\psi(\boldsymbol{\eta}) \geq \mathbf{B}{:}(\boldsymbol{\eta}\otimes\boldsymbol{\eta})$ with equality only for those $\boldsymbol{\eta}$ which are needed in (16).

The n–well problem can now be solved explicitly, when all transformation strains lie on one straight line in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$, i.e., $\boldsymbol{\varepsilon}_j = \boldsymbol{\varepsilon}_0 + a_j\widehat{\boldsymbol{\varepsilon}}$ with $a_j \in \mathbb{R}$, and $\boldsymbol{\varepsilon}_0, \widehat{\boldsymbol{\varepsilon}} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ given constants. The case $\boldsymbol{\varepsilon}_1 = -\boldsymbol{\varepsilon}_3$ and $\boldsymbol{\varepsilon}_2 = \mathbf{0}$ is considered in [26]. For $\boldsymbol{\varepsilon}^{\boldsymbol{\eta}} = \sum_j \eta_j \boldsymbol{\varepsilon}_j$ with $\boldsymbol{\eta} \in \mathbb{E}^n_*$ we find $\boldsymbol{\varepsilon}^{\boldsymbol{\eta}} = (\boldsymbol{\eta}\cdot \mathrm{a})\widehat{\boldsymbol{\varepsilon}}$ and

$$\begin{aligned} \psi(\boldsymbol{\eta}) &= -\tfrac{\gamma}{2}(\mathrm{a}\cdot\boldsymbol{\eta})^2 = -\tfrac{\gamma}{2}(\mathrm{a}\otimes\mathrm{a}){:}(\boldsymbol{\eta}\otimes\boldsymbol{\eta}), \text{ where} \\ \gamma &= \max\{\, (\boldsymbol{\omega}\cdot\mathbb{C}{:}\widehat{\boldsymbol{\varepsilon}})\cdot T(\omega)^{-1}(\boldsymbol{\omega}\cdot\mathbb{C}{:}\widehat{\boldsymbol{\varepsilon}}) \mid \boldsymbol{\omega}\in\mathbb{S}^{d-1}\,\}. \end{aligned}$$

With Proposition 6.1 we conclude the following exact formula.

Theorem 6.2 *If all $W(\boldsymbol{\varepsilon}, \mathbf{e}_j)$ have the form (6) with $\boldsymbol{\varepsilon}_j = \boldsymbol{\varepsilon}_0 + a_j\widehat{\boldsymbol{\varepsilon}}$, then*

$$w_{\mathrm{mix}}(\mathbf{c}) = -\frac{\gamma}{2}(\mathrm{a}\otimes\mathrm{a}){:}\mathrm{M}(\mathbf{c}) = -\frac{\gamma}{2}\Big[\sum_{j=1}^{n} a_j^2 c_j - (\mathrm{a}\cdot\mathbf{c})^2\Big].$$

6.2 Cubic to tetragonal (isotropic case)

We finally consider a special case which relates to the cubic to tetragonal phase transformation in shape memory alloys, see [9]. There are four phases, three variants of martensite ($i = 1, 2, 3$) and one austenite ($i = 4$) with transformation strains

$$\boldsymbol{\varepsilon}_1 = \mathrm{diag}(\beta,\alpha,\alpha),\ \boldsymbol{\varepsilon}_2 = \mathrm{diag}(\alpha,\beta,\alpha),\ \boldsymbol{\varepsilon}_3 = \mathrm{diag}(\alpha,\alpha,\beta),\ \boldsymbol{\varepsilon}_4 = \mathbf{0}.$$

Here the constants $\alpha, \beta \in \mathbb{R}$ usually are such that $2\alpha+\beta$ is much smaller than $|\alpha|+|\beta|$. For instance, for some shape memory alloys we have $(\alpha,\beta) =$

$(-0.0608, 0.1302)$ for Ni-36.8Al, $(0.0868, -0.1497)$ for Fe–25Pt, and $(0.1241, -0.1941)$ for Fe–30Ni–0.3C (see e.g. [9]).

By (21) we easily see that the martensite phases as defined above are srlc with each other for these materials. Moreover, certain combinations of martensites may be srlc to the austenite phase. This is the case if and only if

$$\alpha\beta < 0 \quad \text{and} \quad 0 < |\alpha| < |\beta|.$$

These conditions seem to be satisfied for all cubic to tetragonal phase transformations. Under them, we have that $\theta\boldsymbol{\varepsilon}_i + (1-\theta)\boldsymbol{\varepsilon}_j$ is srlc to $\boldsymbol{\varepsilon}_4$ for $\theta \in \{\alpha/(\alpha-\beta), \beta/(\beta-\alpha)\}$ and $i \neq j \in \{1,2,3\}$. Note that the polytope $\mathcal{P} \subset \mathbb{R}^4$ in this situation can be viewed as an element of $\mathbb{R}^3$ embedded in $\mathbb{R}^4$. If we consider it as an object in $\mathbb{R}^3$, then the six points (θ) define a hexagon H on the face of the polytope $\mathcal{P}$ which is spanned by the three corners $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$. Taking into account the additional point $\mathbf{e}_4$ we can then define a hexagonal pyramid inside the polytope $\mathcal{P}$. Applying Corollary 5.4 we find

Proposition 6.3 *Inside the hexagonal pyramid we have* $w_{\text{mix}} = w_{Reuß}$.

Although the above results holds for general (but identical) elasticity tensors $\mathbb{C}$, we assume that the elasticity tensors is isotropic: $\mathbb{C}{:}\boldsymbol{\varepsilon} = \lambda\,\mathrm{tr}\,(\boldsymbol{\varepsilon})\mathbf{1} + 2\mu\boldsymbol{\varepsilon}$. This assumption makes our example slightly academic but nevertheless we are able to obtain interesting features and are able to give explicit formulae. The acoustic tensor reads

$$\mathbf{T}(\boldsymbol{\omega}) = \mu\mathbf{1} + (\lambda+\mu)\boldsymbol{\omega}\otimes\boldsymbol{\omega}, \qquad \mathbf{T}(\boldsymbol{\omega})^{-1} = \frac{1}{\mu}\mathbf{1} - \frac{\lambda+\mu}{\mu(\lambda+2\mu)}\boldsymbol{\omega}\otimes\boldsymbol{\omega}.$$

For convenience, we define $a = -\frac{1}{2\mu}$ and $b = \frac{\lambda+\mu}{2\mu(\lambda+2\mu)}$ and note that $a+b = -1/(2\lambda+4\mu) < 0$.

The function ψ can now be evaluated as follows:

Proposition 6.4 *Let $\sigma_j \in \mathbb{R}$, $j = 1,2,3$, be the eigenvalues of the symmetric tensor $\mathbb{C}{:}\boldsymbol{\varepsilon}^{\boldsymbol{\eta}}$ and $\sigma_{\min}$ and $\sigma_{\max}$ the smallest and largest eigenvalue. Then, $\psi(\boldsymbol{\eta}) = \gamma(\sigma_{\min}, \sigma_{\max})$ with $\gamma(s,t) = (a+b)\max\{s^2, t^2\}$ if $b \leq 0$ and*

$$\begin{aligned}\gamma(s,t) &= \tfrac{a+b}{2}(s^2+t^2) - \tfrac{(a+b)^2}{4b}(s+t)^2 - \tfrac{b}{4}(s-t)^2 \\ &+ \tfrac{1}{4b}\Big(\max\{0, |a+b||s+t| - b|s-t|\}\Big)^2,\end{aligned}$$

for $b > 0$.

Proof. We have $\psi(\boldsymbol{\eta}) = \min\{\, t(\boldsymbol{\omega},\boldsymbol{\eta}) \;:\; \boldsymbol{\omega} \in \mathbb{S}^2\}$ with $t(\boldsymbol{\omega},\boldsymbol{\eta}) = -\frac{1}{2}(\boldsymbol{\omega}\cdot \mathbb{C}{:}\boldsymbol{\varepsilon}^{\eta})\cdot \mathrm{T}(\boldsymbol{\omega})^{-1}(\boldsymbol{\omega}\cdot\mathbb{C}{:}\boldsymbol{\varepsilon}^{\eta})$. By isotropy it is sufficient to only consider the case $\mathbb{C}{:}\boldsymbol{\varepsilon}^{\eta} = \mathrm{diag}(\boldsymbol{\sigma}) = \mathrm{diag}(\sigma_1,\sigma_2,\sigma_3)$. Then,

$$t(\boldsymbol{\omega},\boldsymbol{\eta}) = a|\mathrm{diag}(\boldsymbol{\sigma})\boldsymbol{\omega}|^2 + b(\boldsymbol{\omega}\cdot\mathrm{diag}(\boldsymbol{\sigma})\boldsymbol{\omega})^2 = a\sum_{i=1}^{3}\sigma_i^2\omega_i^2 + b\Big[\sum_{i=1}^{3}\sigma_i\omega_i^2\Big]^2.$$

With $\Delta = \{\mathbf{y}\in\mathbb{R}^3 \mid y_i \geq 0,\ y_1{+}y_2{+}y_3 = 1\}$ we now define

$$g(\boldsymbol{\sigma},\mathbf{y}) = a\sum_{i=1}^{3}\sigma_i^2 y_i + b\Big[\sum_{i=1}^{3}\sigma_i y_i\Big]^2 \quad\text{and}\quad \Gamma(\boldsymbol{\sigma}) = \min\{g(\boldsymbol{\sigma},\mathbf{y}) \mid \mathbf{y}\in\Delta\}$$

such that it remains to show $\gamma(\sigma_{\min},\sigma_{\max}) = \Gamma(\boldsymbol{\sigma})$. The case $b \leq 0$ is treated easily as $g(\boldsymbol{\sigma},\cdot) : \Delta \to \mathbb{R}$ is concave and, hence, assumes its minimum in one of the extremal points by the Krein–Milman theorem. Clearly $\Gamma(\boldsymbol{\sigma}) = \min\{g(\boldsymbol{\sigma},\mathbf{e}_j) \mid j = 1,2,3\}$ is the desired result (recall $a{+}b < 0$).

In the case $b > 0$ the function $g(\boldsymbol{\sigma},\cdot)$ is convex, however it is easy to see (as $\mathrm{D}_y^2 g$ has rank one) that the minimum is also attained on the boundary of Δ. This leads to a minimization on three intervals, each leading to the result $\gamma(\sigma_i,\sigma_j)$ with γ as define above. We use the explicit formula

$$\min\{s\theta{+}t\theta^2 \mid \theta\in[0,1]\} = \frac{1}{2}\big[s{+}t{+}tw(\frac{t+s}{t})\big]$$

with

$$w(\tau) = \frac{1}{2}[\tau^2{+}1 - \max\{0,|\tau|{-}1\}^2],$$

and apply it to the restriction of $g(\sigma,\cdot)$ to each of the boundary segments.

It remains to show that the minimum is achieved on the side connecting the largest and the smallest eigenvalue. To this end note that $\gamma(\cdot,t)$ is increasing on $(-\infty,t]$ and decreasing on $[t,\infty)$. This implies, for $\sigma_i \leq \sigma_j$, the inequality

$$\gamma(\sigma_i,\sigma_j) \geq \max\{\gamma(\sigma_{\min},\sigma_j),\gamma(\sigma_i,\sigma_{\max})\}.$$

This establishes the result. ∎

In our special situation the eigenvalues σ_j of $\mathbb{C}{:}\boldsymbol{\varepsilon}^{\eta}$ are easily obtained, since $\mathbb{C}{:}\boldsymbol{\varepsilon}^{\eta}$ is always diagonal: diag($\boldsymbol{\sigma}$). Employing the matrix

$$\mathbf{B} = \lambda(\beta{+}2\alpha)\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} + 2\mu\begin{pmatrix} \beta & \alpha & \alpha & 0 \\ \alpha & \beta & \alpha & 0 \\ \alpha & \alpha & \beta & 0 \end{pmatrix}$$

we find $\boldsymbol{\sigma} = \mathbf{B}\boldsymbol{\eta}$. Together with $\overline{\sigma} = [\lambda(\beta{+}2\alpha){+}2\mu\alpha](\eta_1{+}\eta_2{+}\eta_3)$ we obtain the formula

$$\begin{aligned} \psi(\boldsymbol{\eta}) &= \gamma(\sigma_{\min}, \sigma_{\max}) \quad \text{with} \\ \sigma_{\min} &= \overline{\sigma} + 2\mu \min\{(\beta{-}\alpha)\eta_j \mid j = 1,2,3\}, \\ \sigma_{\min} &= \overline{\sigma} + 2\mu \max\{(\beta{-}\alpha)\eta_j \mid j = 1,2,3\}. \end{aligned}$$

This provides an explicit formula for the evaluation of the function ψ. Interestingly enough we see that ψ is defined piecewise by quadratic functions.

To assess the quality of the expressions for the bounds we have utilized this expression for $\Psi(\boldsymbol{\eta})$ to compute the $\mathbb{H}$–measure lower bound derived in Section 5.1 and then have compared it to the lamination upper bound (using 2-lamination steps) and to the Reuß bound. These results are shown in Figures 1–3 for a Ni-36.8Al alloy assuming $\lambda = 97.8$ kN/mm^2 and $\mu = 53.6$ kN/mm^2. Each figure represents a fixed value of the austenite phase fraction. In the upper left-hand corner of each figure, one finds a representation of the polytope $\mathcal{P}$ with the top point representing austenite $\mathbf{e}_4$ and a cutting plane at a fixed value of phase fraction as indicated. In the upper right-hand corner of each figure, one finds a plot of the intersection of the cutting plane with the polytope. On the intersection, we have also plotted the intersection of the cutting plane with the hexagonal pyramid as a dotted line. The 12 labelled points are used to define rays on the cutting plane over which we plot the mixing energy. Six such plots are shown in the lower half of each figure. The top curve in each plot corresponds to the lamination upper bound, the middle curve corresponds to the $\mathbb{H}$–measure bound, and the lower curve corresponds to the Reuß bound. As can be seen from the figures, the Reuß bound is indeed exact inside the hexagonal pyramid; in fact all three bounds are exact there. Outside of the hexagonal pyramid, we can see that there is some deterioration of the Reuß bound but it is still quite reasonable. The $\mathbb{H}$–measure lower bound is very close to the lamination upper bound inside the pyramid (where it is exact) and also outside the pyramid – indicating that these two bounds are quite sharp for the case examined.

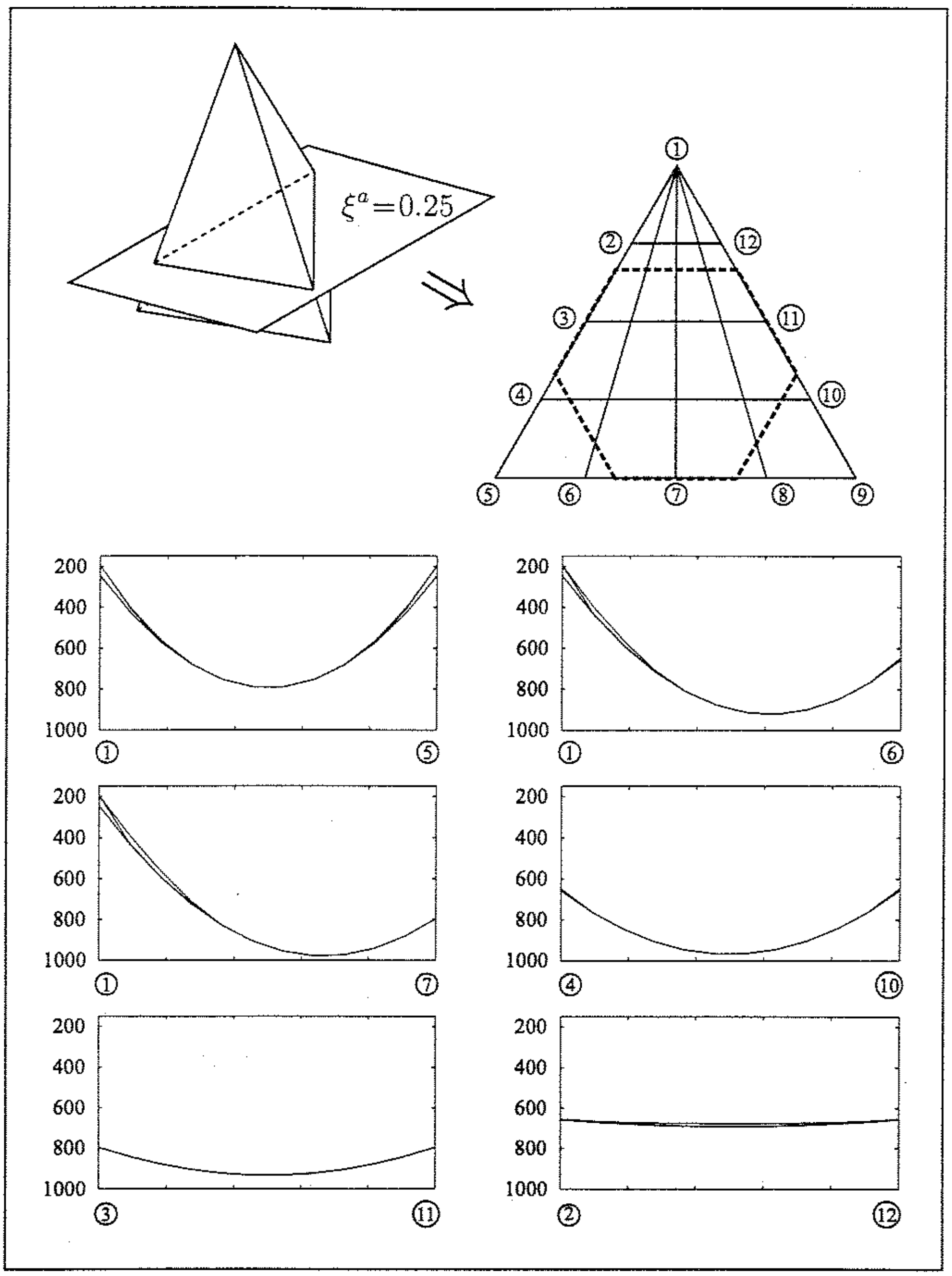

Figure 1: Comparison of mixing energy bounds for a fixed austenite phase fraction $c_4 = 0.25$. The top curve is the lamination bound, the middle curve is the $\mathbb{H}$–measure lower bound, and the lower curve is the Reuß lower bound.

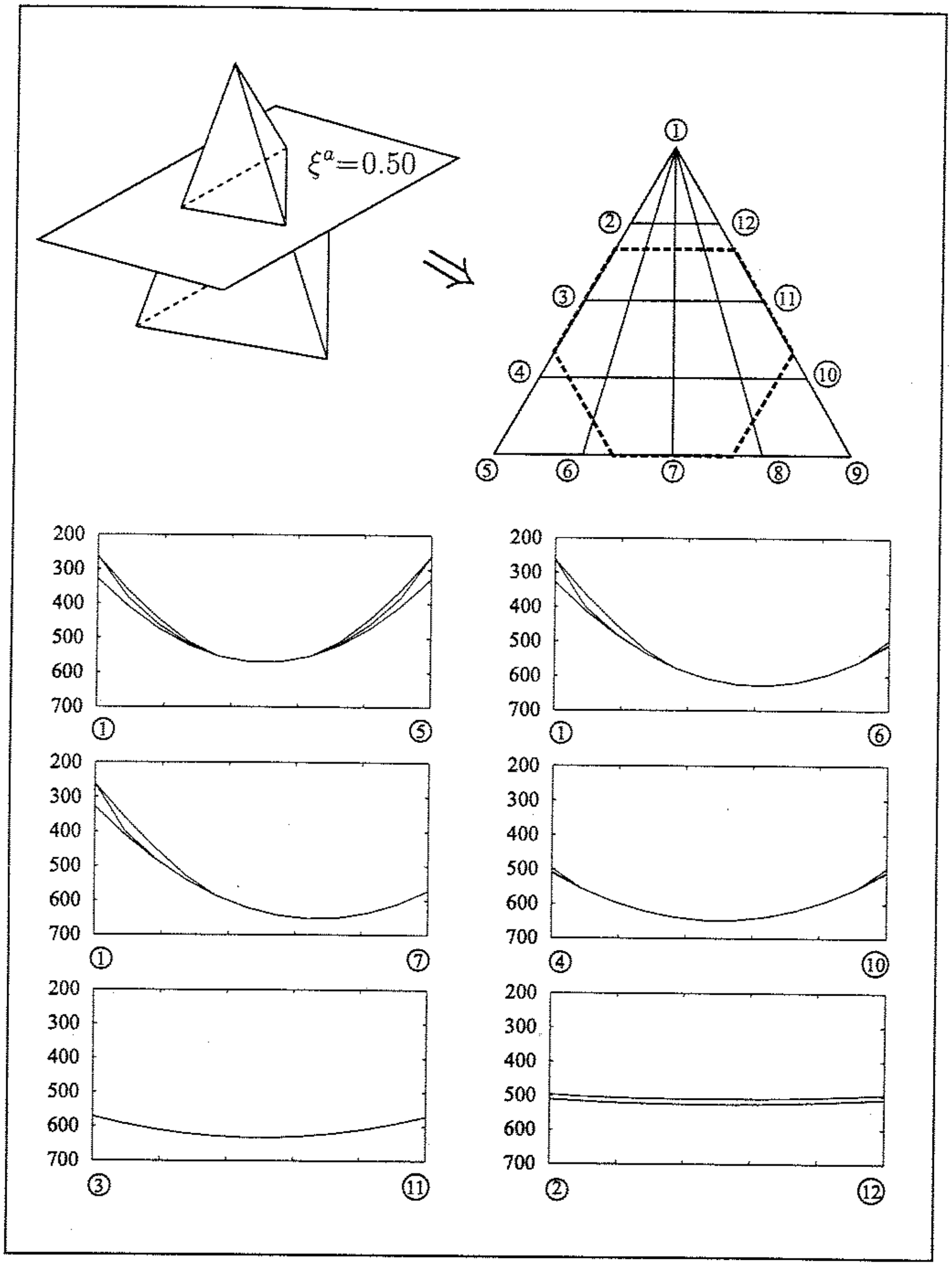

Figure 2: Comparison of mixing energy bounds for a fixed austenite phase fraction $c_4 = 0.50$. The top curve is the lamination bound, the middle curve is the $\mathbb{H}$–measure lower bound, and the lower curve is the Reuß lower bound.

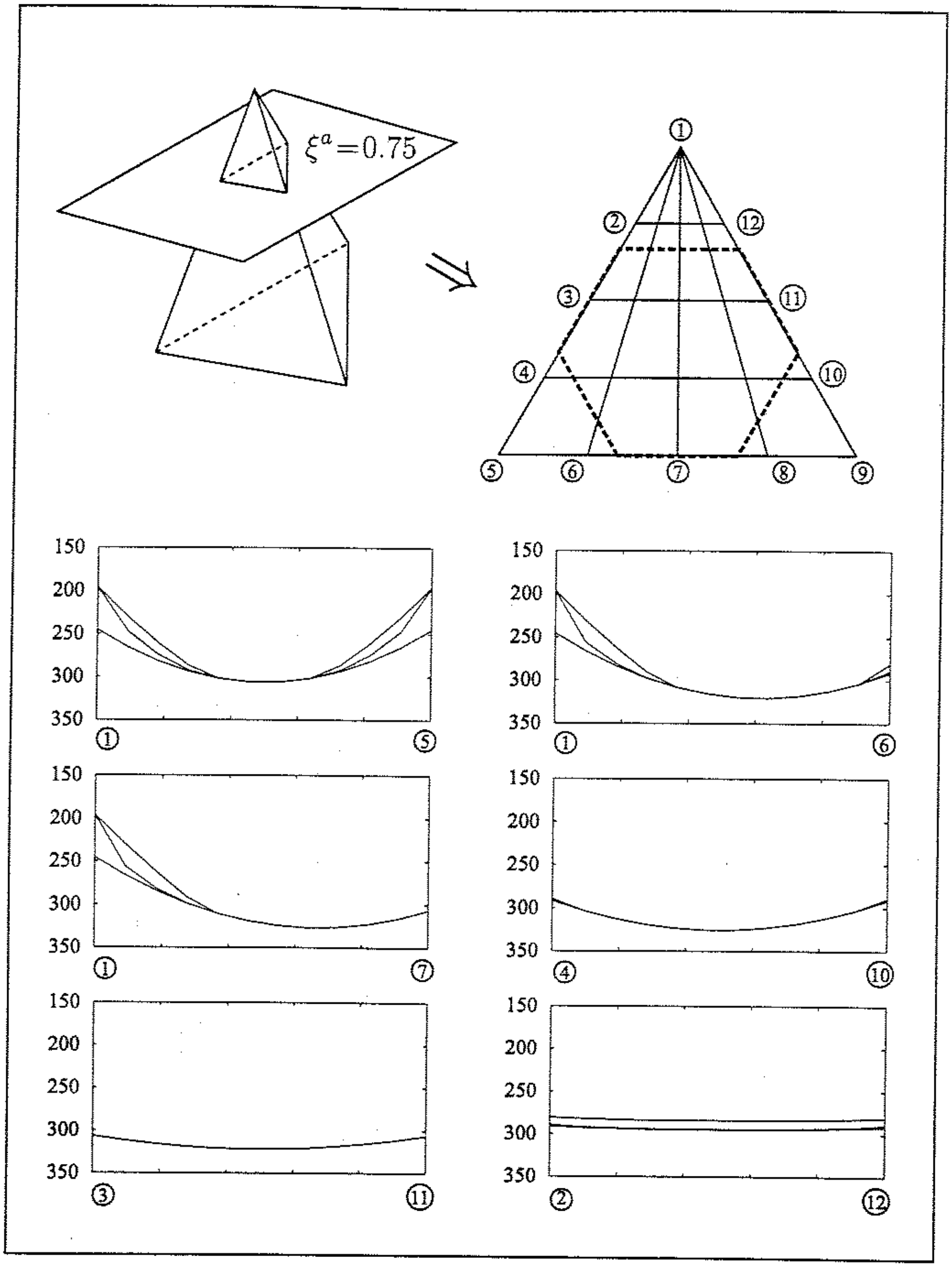

Figure 3: Comparison of mixing energy bounds for a fixed austenite phase fraction $c_4 = 0.75$. The top curve is the lamination bound, the middle curve is the $\mathbb{H}$–measure lower bound, and the lower curve is the Reuß lower bound.

6.3 Evolutionary example utilizing Reuß bound

To demonstrate a practical application of mixture energy bounds, the Reuß estimate is employed in the numerical realization of a fully relaxed shape memory alloy model. The predictions based on this model correspond well with results obtained from experimental tests of single crystal specimens. To begin, the model of [14] is briefly described.

6.3.1 Relaxed Single Crystal Model

With given bounds on the mixture energy, it is possible to construct a model strictly from the viewpoint of elastic stability. In this section the equations for one such model are briefly presented. The model is summarized in Table 1, where the partially relaxed free energy is given in Eq. (22) and employs the Reuß term for the mixing energy. In [14], the thermal terms are given explicitly to allow for the simulation of thermomechanical problems. Here, we simplify the presentation and note that all examples will be conducted at a fixed temperature above the austenite finish temperature where $\alpha_n < \alpha_i$ for $i \in \{1, \cdots, n-1\}$; n is understood to correspond to the austenite variant and $\alpha_j = \alpha_i$ for $i, j \in \{1, \cdots, n-1\}$. This partially relaxed free energy allows for the determination of the volume fractions as a function of the strain path through the optimization problem of Eq. (23). Finally, the stress is specified in Eq.'s(24-25) as a function of the strain and phase fractions. Details of the conditions under which a solution to the relaxation step exists, treatment of the numerical implementation issues, and the overall behavior of the model is more fully discussed in [14].

6.3.2 Single Crystal Simulation

One of the most striking aspects of experimental data on shape memory single crystals is the strong dependence of the phase transformation properties on the crystallographic orientation of the test specimen. It is shown below that this physically observed behavior is captured by the mixture energy bounds as incorporated into the model of Table 1. This is demonstrated by simulating the response of uniaxial tensile specimens over a representative set of directions with respect to the parent lattice. The material under consideration is a Cu-Al-Ni alloy undergoing a $\beta_1 \to \gamma_1'$ stress induced phase

Table 1: Constitutive Equations.

1. Free energy functions:

$$QW = \sum_{i=1}^{n} c_i W(\boldsymbol{\varepsilon}, \mathbf{e}_i) + w_{\text{Reuß}}(\mathbf{c}) \tag{22}$$

2. Volume Fractions:

$$\mathbf{c}^*(\boldsymbol{\varepsilon}, \theta) = \arg\left(\inf_{\mathbf{c}}\left\{QW(\boldsymbol{\varepsilon}, \mathbf{c}) \;\middle|\; \sum_{i=1}^{n} c_i = 1 \;,\; c_i \geq 0\right\}\right) \tag{23}$$

3. Stress:

$$\boldsymbol{\sigma}_i = \mathbb{C}\colon(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_i) \tag{24}$$

$$\boldsymbol{\sigma} = \sum_{i=1}^{n} c_i^* \boldsymbol{\sigma}_i \tag{25}$$

change[1]. In this case, the symmetry change is from a cubic (DO_3) parent to six variants of an orthorhombic martensite where each of the martensite variants are twin compatible with one another in the usual rank-one sense. The estimated material parameters for this material are those specified in [14].

The simulation procedure is as follows. First, specimen dimensions were chosen to be of comparable dimensions to those typically used in the experimental literature: Figure 4 shows the dimensions and the mesh used for all of the examples. To reproduce the most common experimental boundary conditions, one end of the mesh was fixed while the other underwent an imposed displacement along the major axis of the specimen. The examination of orientation dependence followed by creating a discretization of approximately eighty points in the $(0,0,1)\beta_1 - (0,0,1)\beta_1 - (\bar{1},1,1)\beta_1$ region of a standard stereographic projection. At each point a pseudoelastic tensile test at constant temperature above the "austenite finish" point was simulated using the projected direction of the point in the stereogram.

To give an idea of the variation of the overall mechanical response with respect to orientation, Figure 5 shows the predicted behavior of specimens whose tensile axes coincide with the indicated points in the stereographic plane of the parent lattice. In this plot, the infinitesimal strain and stress data were computed from the end response of the bars in each simulation; this is of importance because of the strong effect of the end conditions on the response of the specimen. Note that both the "transformation stress" and the "apparent recoverable strain" vary with respect to orientation. The term "apparent transformation strain" is used here to distinguish the experimental and simulated response from the recoverable strain predictions made by the crystallographic theory.

An interesting aspect of the predicted response is the inverse relation between the length of the apparent yield plateau and the transformation stress level. This can be seen by comparing Figure 6 and Figure 7 which show the variation of apparent transformation strain with orientation and the variation of transformation stress with orientation respectively. For example, the $(\bar{1},1,1)\beta_1$ direction was predicted to have the highest transformation stress in the region and a low transformation strain. Similarly, the greatest apparent recoverable strain simulated was located in the $(0,1,1)\beta_1$ direction

[1]The twinning characteristic of this particular material has been studied in detail in the work of [16, 25]

which had a low transformation stress.

In Table 2 the simulations are compared with well established data regarding the recoverable strain. The first data row is reproduced from Table IV of [24], wherein the maximum recoverable strain of single crystal (orthorhombic) Cu-Al-Ni is listed for the indicated directions based on the established lattice parameters of each phase. Although based on experimentally determined single crystal data, their recoverable strain values rely on idealized homogeneous tensile conditions, and hence should serve as an upper bound for our simulations. While this is in fact the case, it is more noteworthy that the simulations follow the trends established by the upper bound with regard to direction and the relative magnitudes of the recoverable strain.

Table 2: Recoverable strain.

Direction	$(0,0,1)\beta_1$	$(0,1,1)\beta_1$	$(\bar{1},1,1)\beta_1$
Bound	4.3%	6.2%	1.6%
Simulation	3.6%	5.1%	1.0%

A similar comparison is made for transformation stress levels in Table 3 wherein the transformation stress levels are compared for three directions for which experimental data is available. The transformation stress was estimated from the response curves for both the experimental and simulated data by using a 0.1% strain offset to estimate the start of the phase transformation. The close match for initiation of the phase change is an indicator that the Reuß mixture energy bound has captured the essential features of the transformation energetics[2].

Table 3: Transformation stress (MPa), 0.1% offset.

Direction	$(0.925, 0.380, 0.000)\beta_1$	$(-0.447, -0.447, 0.775)\beta_1$	$(-0.577, -0.577, 0.577)\beta_1$
Experiment	105	170	400
Simulation	100	180	395

[2]More detailed comparisons with the experimental work of [27] are made in [14].

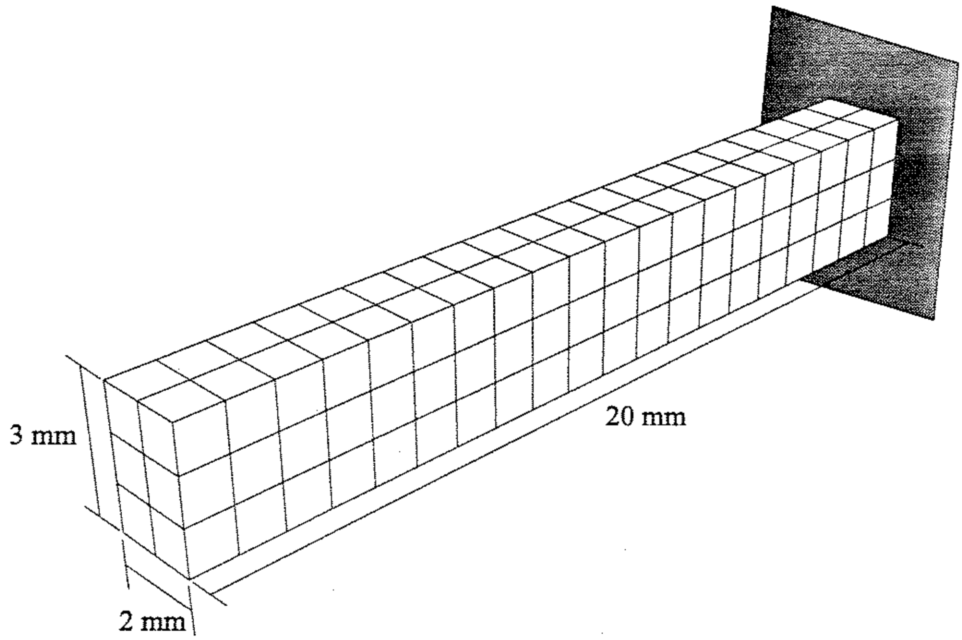

Figure 4: Dimensions and mesh of the simulated specimen.

Acknowledgements

SG dankt der Alexander von Humboldt–Stiftung für die Verleihung eines Forschungsstipendium an der Universität Stuttgart. SG also acknowledges the support of the Lawrence Livermore National Laboratory through subcontract B502681 of DOE prime contract W-7405-ENG-48. The research of AM was partially supported by DFG through project C7 of SFB 404 (Multifield Problems).

References

[1] M. Avellaneda and G.W. Milton, *Bounds on the effective elasticity tensor of composites based on two-point correlations*, Proceedings of the ASME Energy–Technology Conference and Exposition (D. Hui and T. Koszic, eds.), ASME, 1989.

[2] J.M. Ball and R.D. James, *Fine phase mixtures as minimizers of energy*, Archive for Rational Mechanics and Analysis **100** (1987), 13–52.

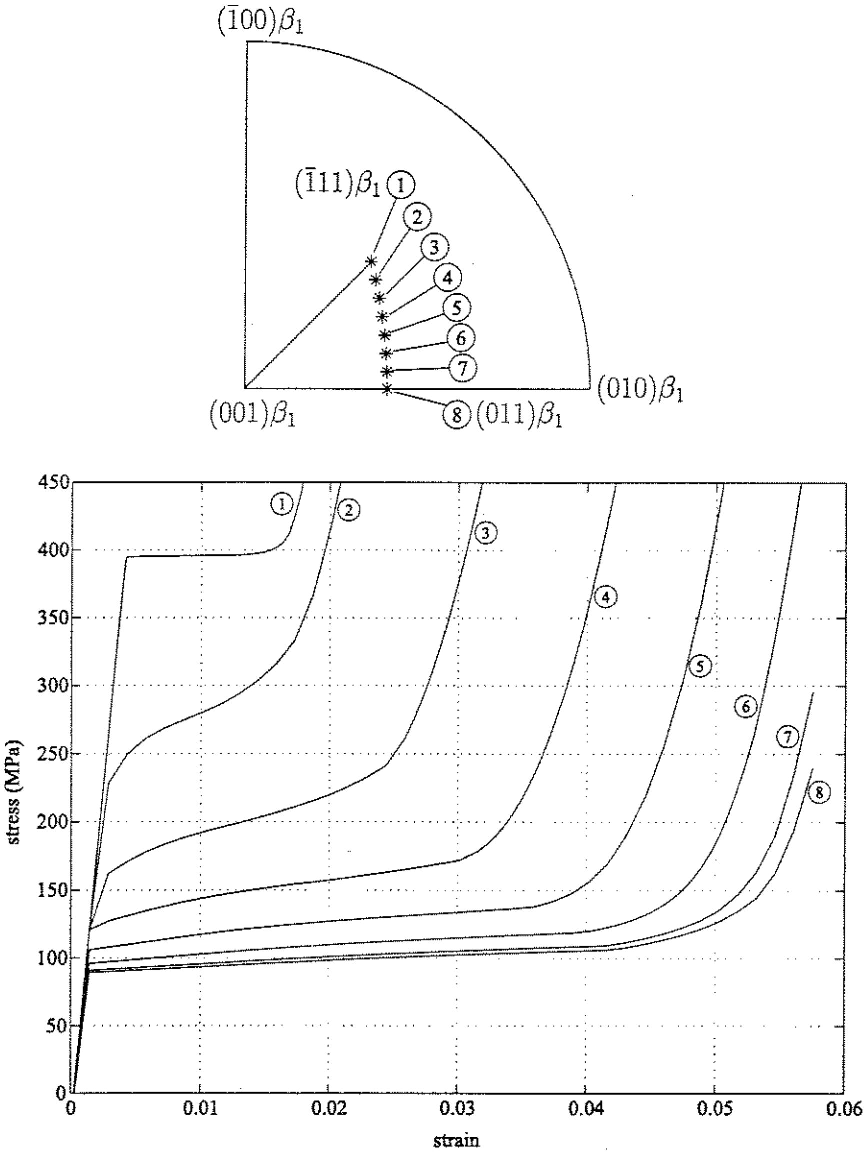

Figure 5: Variation of mechanical response along the indicated ray in the stereographic plane.

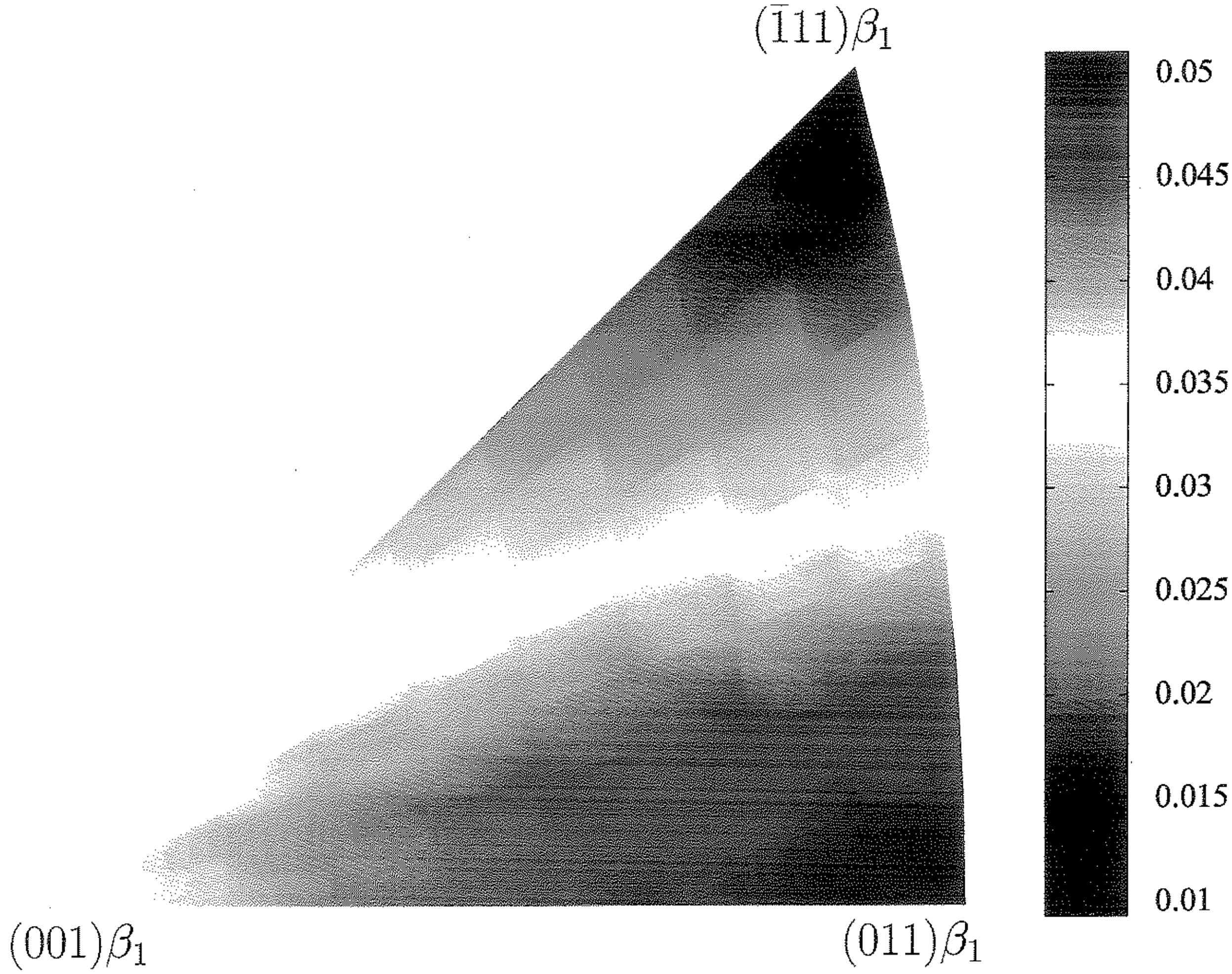

Figure 6: Apparent transformation strain as a function of orientation.

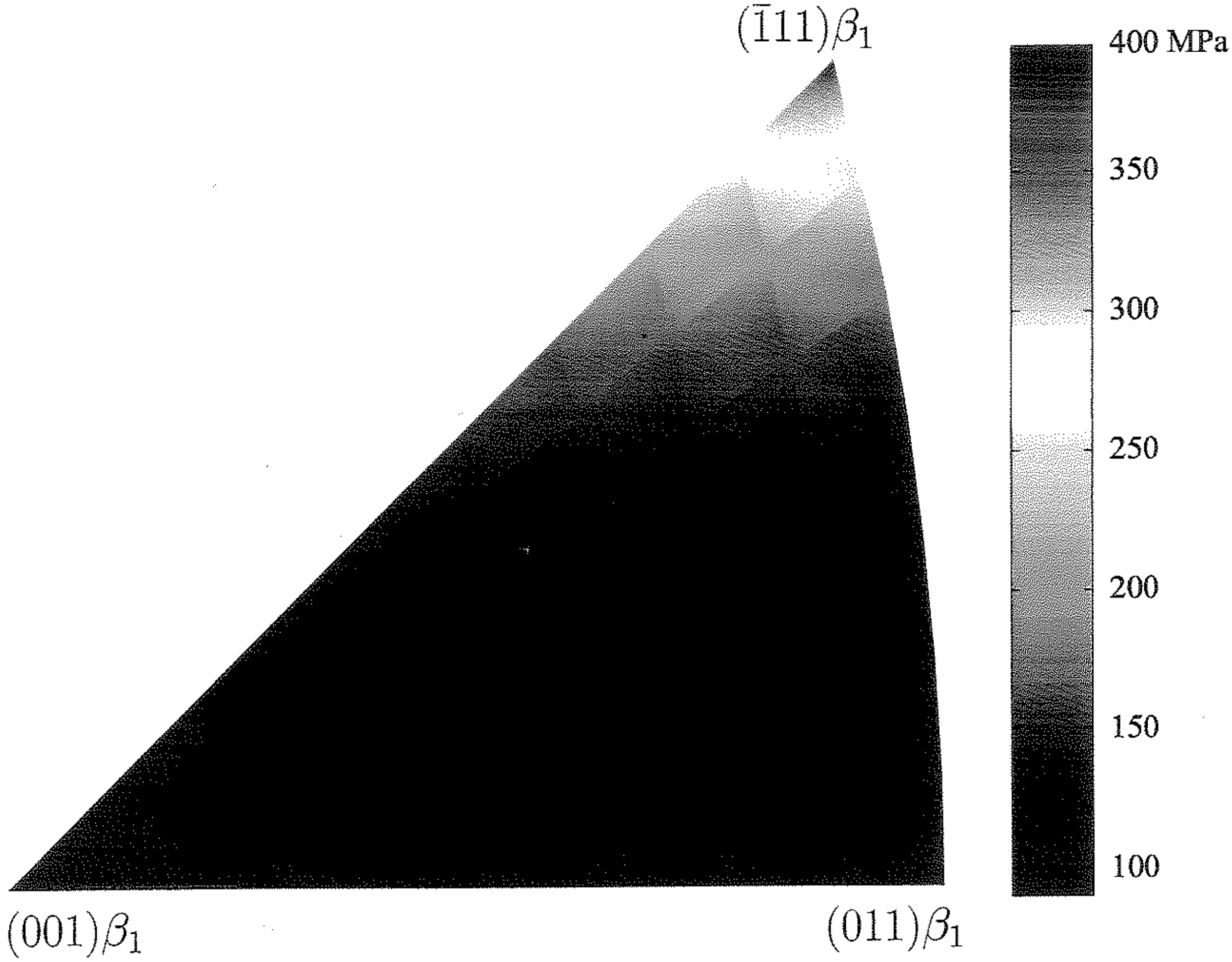

Figure 7: Transformation stress as a function of specimen orientation.

[3] ———, *Proposed experimental tests of a theory of fine microstructures and the 2-well problem*, Philosophical Transactions of the Royal Society of London **A338** (1992), 389–450.

[4] K. Bhattacharya, *Comparison of the geometrically nonlinear and linear theories of martensitic transformations*, Continuum Mechanics and Thermodynamics **5** (1993), 205–242.

[5] ———, *The microstructure of martensite and its implications for the shape-memory effect*, Microstructure and Phase Transition (D. Kinderlehrer, R.D. James, M. Luskin, and J.L. Eriksen, eds.), Springer-Verlag, New York, 1993, pp. 1–25.

[6] K. Bhattacharya and R.D. James, *A theory of thin films of martensitic materials with applications to microactuators*, (1998), (submitted for publication).

[7] K. Bhattacharya, R.D. James, and P.J. Swart, *Relaxation in shape-memory alloys: Part I: Basic model*, Acta Materialia **45** (1997), 4547–4560.

[8] K. Bhattacharya and R.V Kohn, *Recoverable strains in shape-memory alloys*, Journal de Physique IV **5** (1995), C8–261–C8–266.

[9] ———, *Symmetry, texture and the recoverable strain of shape-memory polycrystals*, Acta Materiallia **44** (1996), 529–542.

[10] J.G. Boyd and D.C. Lagoudas, *A thermodynamic constitutive model for the shape memory materials. Part I. The monolithic shape memory alloy*, International Journal of Plasticity **12** (1996), 805–842.

[11] B. Dacorogna, *Direct methods in the calculus of variations*, Springer-Verlag, 1989.

[12] S. Govindjee and C. Miehe, *A multi-variant martensitic phase transformation model: Formulation and numerical implementation*, Tech. Report UCB/SEMM-2000/06, University of California Berkeley, Department of Civil Engineering, 2000.

[13] G.J. Hall and S. Govindjee, *A model and numerical framework for the simulation of solid-solid phase transformations*, Tech. Report

UCB/SEMM-1999/11, University of California Berkeley, Department of Civil Engineering, 1999.

[14] ———, *Application of a paritally relaxed shape memory free energy function to estimate the phase diagram and predict global microstructure evolution*, Journal of the Mechanics and Physics of Solids (2001), in press.

[15] M. Huang and L.C. Brinson, *A multivariant model for single crystal shape memory alloy behavior*, Journal of the Mechanics and Physics of Solids **46** (1998), 1379–1409.

[16] S. Ichinose, Y. Funatsu, and K. Otsuka, *Type II deformation twinning in γ_1' martensite in a Cu-Al-Ni alloy*, Acta Metallurgica **33** (1985), 1613–1620.

[17] R.V. Kohn, *The relaxation of a double-well problem*, Continuum Mechanics and Thermodynamics **3** (1991), 193–236.

[18] M.S. Kuczma, A. Mielke, and E. Stein, *Modelling of hysteresis in two-phase systems*, Arch. Mech. **51** (1999), 693–715.

[19] V.I. Levitas, *Thermomechanical theory of martensitic phase transformations in inelastic materials*, International Journal of Solids and Structures **35** (1998), 889–940.

[20] A. Mielke, *Estimates on the mixture function for multiphase problems in elasticity*, Multifield problems, Springer, Berlin, 2000, pp. 96–103.

[21] A. Mielke and F. Theil, *A mathematical model for rate–independent phase transformations with hysteresis*, Models of Continuum Mechanics in Analysis and Engineering (H.-D. Alber, R.M. Balean, and R. Farwig, eds.), Shaker–Verlag, 1999, pp. 117–129.

[22] A. Mielke, F. Theil, and V.I. Levitas, *A variational formulation of rate-independent phase transformations with using an extremum principle*, Math. Models Methods Applied Sciences (2000), submitted.

[23] ———, *Mathematical formulation of quasistatic phase transformations with friction using an extremum principle*, Tech. Report Preprint IfAM, Universität Hannover, Sept. 1998.

[24] S. Nenno and T. Saburi, *The shape memory effect and related phenomena*, Proceedings of an International Conference on Solid Solid Phase Transformations, 1981, pp. 1455–1479.

[25] K. Okamoto, S. Ichinose, K. Morii, K. Otsuka, and K. Shimizu, *Stress-induced martensitic transformation in a Cu-Al-Ni alloy*, Acta Metallurgica **34** (1986), 2065–2073.

[26] S. Pagano, P. Alert, and O. Maisonneuve, *Solid-solid phase transition modelling. Local and global minimizations of non-convex and relaxed potentials. Isothermal case for shape memory alloys*, International Journal for Engineering Science **36** (1998), 1143–1172.

[27] T.W. Shield, *Orientation dependence of the pseudoelastic behavior of single crystals of Cu-Al-Ni in tension*, Journal of the Mechanics and Physics of Solids **43** (1995), 869–895.

[28] V.P. Smyshlyaev and J.R. Willis, *On the relaxation of a three-well energy*, Proceedings of the Royal Society of London A **455** (1998), 779–814.

[29] L. Tartar, *H-measures, a new approach for studying homogenization, oscillation and concentration effects in partial differential equations*, Proceedings of the Royal Society of Edinburgh A **115** (1990), 193–230.

[30] F. Theil, *Relaxation of rate independent evolution problems*, Proceedings of the Royal Society of Edinburg A (2000), submitted.