Analysis, Modeling and Simulation of Multiscale Problems

Relaxation via Young measures of material models for rate–independent inelasticity

Alexander Mielke

Preprint 35

Schwerpunktprogramm der Deutschen Forschungsgemeinschaft

Relaxation via Young measures of material models for rate–independent inelasticity*

Alexander Mielke†

20 February 2002

Abstract

We study material models rate–independent inelasticity in situations where no internal length scale is given and formation of microstructure for the deformation gradients and the internal variables may occur. We develop a rational procedure for deriving consistent macroscopic models which allow for the computation of non-trivial effective quantities without resolving the fine scales. The method involves the relaxation of variational incremental problems which are derived from an elastic and a plastic potential. We use Young measures to describe the microscopic distribution of the internal variables as well as the quasiconvexification of the elastic stored–energy density with respect to the deformation gradient. The resulting model provides a new rate–independent model in terms of the deformation and the Young measure. It uses generalized elastic and plastic potentials which defined a dissipation distance on Young measures which is related to the Wasserstein distance. The update algorithm for the incremental problem is discussed in detail, and a simple one–dimensional example is given.

1 Introduction

The aim of this paper is to give a theoretical foundation to material models for elastic bodies which are described by inelastic internal variables whose evolution is governed by a rate–independent evolution law of the type of flow rules in plasticity. We refer to [Ric71, HN75, ZW87] for the general setup of such models for generalized standard materials. A special and important class of such models are those of elastoplasticity, [SO85, Sim88, MS92, Mie96]. Here we especially restrict our attention to models which can be described by purely energetic consideration, i.e., there is an elastic potential which describes energy storage and there is a dissipation potential which describes how energy is dissipated. Such energetic descriptions lead in a natural way to variational incremental update and are studied extensively within the last five years, having applications in elastoplasticity [OR99,

*Research partially supported by SFB 404 (Multifield problems) and SPP 1095 (Analysis, Modeling and Simulation of Multiscale Problems) under Mi 459/3-1.

†Mathematisches Institut A, Universität Stuttgart, Germany

ORS00, CHM01, HH01, Mie02a], in shape–memory alloys [KMS99, MTL02, GMH02, Rou02], in ferromagnetism [RK02], and in damage models [FM98, HSB02, KMR02].

In recent years the occurrence and evolution of microstructure has attracted the attention of quite a body of research, see e.g., [BJ87, MTL02, GMH02] for work in the field of shape memory alloys. The present work was mainly motivated by experimental and numerical observations in finite–strain elastoplasticity, see [OR99, ORS00, CHM01, HH01]. Our aim is to develop a rational procedure for the derivation of consistent macroscopic models which capture essential features of the microstructure and thus allow for the computation of effective macroscopic quantities without resolving the small length scales. In particular, we hope that our rather theoretical considerations help to understand the general structure behind the modeling of microstructure evolution in rate–independent models. Thus, we expect that this work can help to make numerical computations more efficient.

We restrict our treatment to models were no material length scale is given, i.e., the elastic stored energy ψ depends solely locally on the deformation gradient $\mathbf{F} = \mathrm{D}\boldsymbol{\phi}$ and internal variables $z \in Z$ but not on their derivatives: $\psi = \widehat{\psi}(x, \mathbf{F}, z)$. We allow for models which are fully nonlinear to include cases like finite–strain elastoplasticity where the plastic deformation $\mathbf{F}_{\mathrm{pl}}$ with $\det \mathbf{F}_{\mathrm{pl}} = 1$ is part of the internal variables; hence $Z \subset \mathbb{R}^m$ may be a nonconvex set such that simple averages of microstructures are meaningless.

In addition to the elastic stored–energy density $\widehat{\psi}$ we postulate the existence of a dissipation potential $\Delta = \widehat{\Delta}(z, \dot{z})$ which is homogeneous of degree 1 in the rate $\dot{z}$. This infinitesimal metric on Z defines a dissipation distance D on Z such that $D(z_0, z_1)$ is the minimal amount of energy dissipated when changing z from z_0 to z_1 along a continuous path, see [Mie02a] for the first general description of this concept.

The rate–independent model is described by the flow rule associated with the dissipation potential $\widehat{\Delta}$ and the elastic equilibrium condition. We will show that this is just the class of generalized standard materials as defined in [HN75, Hac97].

Using the global dissipation distance D it is possible to formulate the model in terms of a stability condition (S) and the energy inequality (E), see [MT99, MT01, MTL02] for its first occurrence. This framework has the major advantage that it doesn't make any assumption on the smoothness of the constitutive functions or of the processes to be described. Hence, it is particularly suited to describe systems where formation of microstructure takes place. Moreover, there is a natural variational incremental problem, which is defined via simultaneous minimization with respect to the deformation $\boldsymbol{\phi} \in \mathcal{F}$ and the internal state $z \in \mathcal{Z}$ of the sum of the Gibbs' energy and the dissipation distance. The relevance of variational incremental problems was realized in several works independently, cf. [MT99, OR99, Mie01, CHM01].

The formation of microstructure is now seen in the nonattainment of minimizers of the variational problem. The relaxation we propose relies on the well–developed theory of relaxation of one variational problem. However several new features appear. First, the minimization is done here with respect to the deformation whose gradient $\mathbf{F} = \mathrm{D}\boldsymbol{\phi}$ occurs in the elastic potential $\widehat{\psi}$ and the internal variable which appears in $\widehat{\psi}$ and $D(z_{\mathrm{old}}, \cdot)$ locally and without any derivatives. Second, the coupling of the internal states z_{old} and z_{new} in one incremental step leads to a coupling of the microstructures between different time levels. This is the only point where we make an additional modeling assumption. Our approach will account only for dissipational losses if the distribution functions associated with the microstructure changes but not if the distributions stay fixed while the texture of the micropattern changes, see Remark 7.2. All the other steps of our approach follows directly from mathematical considerations and do not need further assumptions.

The underlying mathematical theory is that of Young measures where $\mathrm{YM}(\Omega, Z)$ replaces the classical set $\mathcal{Z} \stackrel{\text{def}}{=} \{ z : \Omega \to Z \mid z \text{ measurable} \}$ of internal states. A Young measure μ is a map from Ω into the set $\mathrm{Prob}(Z)$ of probability distributions on Z and hence can be seen as a function with takes distributed values occuring in a mesoscopic representative volume elements. Hence, we consider the Young measure as a macroscopic object. Our final model is then formulation in terms of the macroscopic deformation $\boldsymbol{\phi}$ and the Young measure μ by suitably extending the elastic functional and the dissipation distance to $\mathcal{F}\times\mathrm{YM}(\Omega, Z)$.

Relaxations for one single minimization problem are well established and there is a huge literature on this subject, including possible numerical implementations, see e.g. [Dac89, Rou97, Kru98, DH98, AP00, Kru02, Gol01]. However, there are almost no results for the relaxation of time dependent problems. Our approach follows that developed for shape–memory alloys in [MT99, MTL02]. A numerical implementation is given in [CP01]. Closely related rate–independent models for ferromagnetism are studied and used numerically in [RK02, Rou02]. A similar approach was applied in [FM98] to fracture in brittle materials.

A major point in our construction is that the relaxed problem on $\mathcal{F}\times\mathrm{YM}(\Omega, Z)$ has again a natural incremental variational problem. We give a detailed discussion of the update algorithm associated with this incremental problem. The crucial part is to control the correlations between the Young measures on subsequent time steps. Moreover, to the Young measure for z corresponds to a gradient Young measure for the microscopic deformation gradient which has to be taken into account. We show how the calculation of all the correlations between different Young measures can be reduced as much as possible. Nevertheless, even after discretization the remaining minimization problems remain extremely complex.

Finally we provide a simple example for our general abstract evolution law. It is concerned with a one–dimensional example of nonlinear elastoplasticity involving the multiplicative split of the deformation gradient. We will show that formation of microstructure occurs in this model. For applications in N–phase shape–memory alloys we refer to [MTL02, CP01, GMH02]. Numerical algorithms using similar ideas than proposed in our method are implemented in two– and three–dimensional elastoplasticity, see [OR99, HH01, Mie01].

2 The rate–independent material model

We start by describing our general model for rate–independent inelastic material behavior and provide its energetic formulation in the following section. The modeling concerns the two distinct physical effects of energy storage and energy dissipation. The first is attributed to an elastic potential which may also include hardening or latent energies and the second is attributed to frictional effects due to changes of the internal variables.

We consider a body $\Omega \subset \mathbb{R}^d$ which undergoes a deformation $\boldsymbol{\varphi} : \Omega \mapsto \mathbb{R}^d$ such that the deformation gradient $\mathbf{F}(x) = \mathrm{D}\boldsymbol{\varphi}(x) \in \mathbb{R}^{d\times d}$ lies in $\mathrm{GL}_+(\mathbb{R}^d) = \{ \mathbf{F} \in \mathbb{R}^{d\times d} \mid \det \mathbf{F} >$

$0\,\}$. Additionally there is a vector–valued internal variable $z = (z^1, \ldots, z^n) \in Z \subset \mathbb{R}^n$ which describes the internal state of the material. The elastic properties are given by the dependence of the elastic potential ψ on $(x, \mathbf{F}, z)$ via $\psi = \widehat{\psi}(x, \mathbf{F}, z)$. For fixed x and z we assume that the function $\widehat{\psi}(x, \cdot, z)$: $\mathrm{GL}_+(\mathbb{R}^d) \mapsto [0, \infty)$ is coercive (i.e., $\widehat{\psi}(x, \mathbf{F}, z) \to \infty$ if $(\det \mathbf{F})^{-1} + |\mathbf{F}| \to \infty$) and poly– or quasiconvex, see [Dac89, Rou97]. In addition, we describe boundary conditions and external loadings such that the total energy of a given state $(\boldsymbol{\varphi}, z) : \Omega \mapsto \mathbb{R}^d \times \mathbb{R}^n$ at time t is given by the functional

$$\mathcal{E}(t, \boldsymbol{\varphi}, z) = \int_\Omega \widehat{\psi}(x, \mathrm{D}\boldsymbol{\varphi}(x), z(x))\,\mathrm{d}x - \langle \ell(t), \boldsymbol{\varphi} \rangle,$$

where $\ell(t)$ denotes the external loading and $\boldsymbol{\varphi}$ lies in the set $\mathcal{F}$ of admissible deformations

$$\mathcal{F} = \{\, \boldsymbol{\varphi} : \Omega \mapsto \mathbb{R}^d \mid \boldsymbol{\varphi}|_{\Gamma_{\mathrm{Dir}}} = \boldsymbol{\varphi}_{\mathrm{Dir}},\ \mathrm{D}\boldsymbol{\varphi}(x) \in \mathrm{GL}_+(\mathbb{R}^d) \text{ on } \Omega \,\}.$$

Throughout our notation is such that upper case calligraphic letters like $\mathcal{E}, \mathcal{F}, \mathcal{Z}, \ldots$ associate with quantities which are defined over the spatial domain Ω.

A basic assumption of our rate–independent model will be that a solution process has to be in a stable elastic equilibrium for all $t \in [0, T]$, i.e. $\boldsymbol{\varphi}(t, \cdot) : \Omega \mapsto \mathbb{R}^d$ is a (local) minimizer of $\mathcal{E}(t, \cdot, z(t, \cdot))$ on $\mathcal{F}$.

Changes of the internal variables during a slow loading or unloading process will give rise to internal friction which dissipates energy via the dissipation rate $\Delta \geq 0$. We make the constitutive assumption

$$\Delta = \widehat{\Delta}(x, z, \dot{z}) \geq 0 \text{ where } \dot{z}(t, x) = \tfrac{\partial}{\partial t} z(t, x),$$

which also means that there are no other dissipation mechanisms in the model.

We call $\Delta : \Omega \times \mathrm{T}Z \mapsto [0, \infty]$ the dissipation potential [ZW87]. Rate–independency is obtained by assuming homogeneity in $\dot{z}$ of degree 1, namely $\widehat{\Delta}(x, z, \alpha\dot{z}) = \alpha\widehat{\Delta}(x, z, \dot{z})$ for $\alpha \geq 0$. Furthermore, we assume that $\widehat{\Delta}(x, z, \cdot) : \mathrm{T}_z Z \mapsto [0, \infty]$ is convex and that $\widehat{\Delta}$ satisfies $\widehat{\Delta}(x, z, v) \geq c|v|$ for some $c > 0$. (Here $\mathrm{T}_z Z$ denotes the tangent space of Z in the point z.)

For each $x \in \Omega$ the dissipation potential $\widehat{\Delta}(x, \cdot, \cdot)$ defines a distance metric on Z via

$$\widehat{D}(x; z_0, z_1) \stackrel{\text{def}}{=} \inf\{\, \int_0^1 \widehat{\Delta}(x, z(s), \dot{z}(s))\,\mathrm{d}s \mid z \in \mathrm{C}^1([0, 1], Z), z(0) = z_0, z(1) = z_1 \,\}.$$

Often $\widehat{\Delta}(x, \cdot, \cdot) : \mathrm{T}Z \to [0, \infty)$ is called an infinitesimal (Finsler) metric and Z whereas $\widehat{D} : Z \times Z \to [0, \infty)$ is called the associated global metric or distance. To indicate the globality, we will subsequently use the name dissipation distance for $\widehat{D}$.

The set of internal states is given by $\mathcal{Z} \stackrel{\text{def}}{=} \{\, z : \Omega \to Z \mid z \text{ measurable } \,\}$. We define a (possibly unsymmetric) dissipation metric $\mathcal{D}$ on $\mathcal{Z}$ via

$$\mathcal{D}(z_0, z_1) \stackrel{\text{def}}{=} \int_\Omega \widehat{D}(x; z_0(x), z_1(x))\,\mathrm{d}x \qquad \text{for } z_0, z_1 \in \mathcal{Z}.$$

From the above definition of $\widehat{D}$ it follows that $\mathcal{D}$ satisfies the triangle inequality

$$\mathcal{D}(z_{\mathrm{old}}, z_{\mathrm{new}}) \leq \mathcal{D}(z_{\mathrm{old}}, z) + \mathcal{D}(z, z_{\mathrm{new}}) \quad \text{for all } z, z_{\mathrm{old}}, z_{\mathrm{new}} \in \mathcal{Z}. \tag{2.1}$$

Considering a process $z : [0,T] \to \mathcal{Z}$ the dissipation on an interval $[t_0,t_1]$ is usually defined by $\mathrm{Diss}(z;[t_0,t_1]) = \int_{t_0}^{t_1}\int_\Omega \Delta(x,z(t,x),\dot z(t,x))\,\mathrm{d}x\,\mathrm{d}t$. This notion relies on the differentiability with respect to t. However, using the dissipation distance $\mathcal{D}$ from above we have the alternative definition

$$\mathrm{Diss}(z;[t_0,t_1]) \stackrel{\text{def}}{=} \sup\{ \sum_{j=1}^{N} \mathcal{D}(z(\tau_{j-1}),z(\tau_j)) \mid N\in\mathbb{N},\ t_0\le\tau_0<\tau_1<\ldots<\tau_N\le t_1 \}, \qquad (2.2)$$

which does not contain any derivative.

3 The energetic formulation of the model

As described above, our model is completely described by the elastic stored–energy density $\psi = \widehat\psi(x,\mathbf{F},z)$, and the dissipation potential $\Delta = \widehat\Delta(x,z,\dot z)$, and the external loadings $\ell(t)$.

Definition 3.1 *A process $(\boldsymbol{\varphi},z) : [0,T] \mapsto \mathcal{F}\times\mathcal{Z}$ is called a solution of the above rate–independent model, if (S) and (E) hold:*

(S) [Stability] *For all $t\in[0,T]$ we have*

$$\mathcal{E}(t,\boldsymbol{\varphi}(t),z(t)) \le \mathcal{E}(t,\widehat{\boldsymbol{\varphi}},\widehat z) + \mathcal{D}(z(t),\widehat z) \text{ for all } (\widehat\varphi,\widehat z)\in\mathcal{F}\times\mathcal{Z}.$$

(E) [Energy inequality] *For all t_0,t_1 with $0\le t_0<t_1\le T$ we have*

$$\mathcal{E}(t_1,\boldsymbol{\varphi}(t_1),z(t_1)) + \mathrm{Diss}(z,[t_0,t_1]) \le \mathcal{E}(t_0,\boldsymbol{\varphi}(t_0),z(t_0)) - \int_{t_0}^{t_1}\langle\dot\ell(t),\boldsymbol{\varphi}(t)\rangle\,\mathrm{d}t.$$

Remark 3.2 *Our energetic formulation using (S) & (E) is equivalent to the classical local flow rules for generalized standard materials ([HN75]), if the solution processes are sufficiently smooth, i.e., the rates need to be well defined. Using the thermodynamically conjugate variables*

$$\mathbf{T} \stackrel{\text{def}}{=} \tfrac{\partial}{\partial\mathbf{F}}\widehat\psi(x,\mathbf{F},z)\in\mathbb{R}^{d\times d} \text{ and } q \stackrel{\text{def}}{=} -\tfrac{\partial}{\partial z}\widehat\psi(x,\mathbf{F},z)\in\mathrm{T}_z^*Z$$

we find the local versions of (S) and (E):

$$\begin{array}{ll} (S)_{\mathrm{loc}} & \left\{\begin{array}{l} \int_\Omega \mathbf{T}{:}\mathrm{D}\mathbf{u}\,\mathrm{d}x = \langle\ell(t),\mathbf{u}\rangle \text{ for all } \mathbf{u} \text{ with } \mathbf{u}|_{\Gamma_{\mathrm{Dir}}}=0; \\ q(t,x)\cdot v \le \widehat\Delta(x,z(t,x),v) \text{ for all } v\in\mathrm{T}_{z(t,x)}Z. \end{array}\right. \\ (E)_{\mathrm{loc}} & \int_\Omega \widehat\Delta(x,z(t,x),\dot z(t,x))\,\mathrm{d}x \le \int_\Omega q(t,x)\cdot\dot z(t,x)\,\mathrm{d}x. \end{array}$$

Here T_zZ denotes the tangent space to Z in z and T_z^*Z the dual space.

Defining the subdifferential of the convex function $\widehat\Delta(x,z,\cdot):\mathrm{T}_zZ\mapsto[0,\infty)$ via

$$\partial_v\widehat\Delta(x,z,v) = \{ q\in\mathrm{T}_z^*Z \mid \widehat\Delta(x,z,v{+}\widetilde v) \ge \widehat\Delta(x,z,v) + q\cdot\widetilde v \text{ for all } \widetilde v\in\mathrm{T}_zZ \}, \qquad (3.1)$$

we obtain the dual form of the flow rule from $(E)_{\mathrm{loc}}$ and the second equation on $(S)_{\mathrm{loc}}$:

$$q(t,x)\in\partial_v\widehat\Delta(x,z,\cdot z) \quad\Longleftrightarrow\quad 0\in\partial_v\widehat\Delta(x,z,\cdot z) + \tfrac{\partial}{\partial z}\widehat\psi(x,\mathbf{F},z). \qquad (3.2)$$

The elastic region is given via $\mathbb{Q}(x,z) = \partial_v \widehat{\Delta}(x,z,0) \subset \mathrm{T}_z^* Z \subset \mathbb{R}^n$ *and the primal form of the flow rule takes the form*

$$\dot{z}(t,x) \in \partial \mathcal{X}_{\mathbb{Q}(x,z(t,x))}(q(t,x)) = \mathrm{N}_{q(t,x)}\mathbb{Q}(x,z(t,x)), \tag{3.3}$$

where $\mathrm{N}_q\mathbb{Q}(x,z) \subset \mathrm{T}_z Z$ *denotes the (outer) normal cone. This is an equivalent, mathematical form of the other two well–known formulations of flow rules used more often in continuum mechanics. First there is the variational inequality*

$$\langle \dot{z}(t,x), q(t,x) - \widehat{q} \rangle \geq 0 \quad \text{for all } \widehat{q} \in \mathbb{Q}(x,z(t,x))$$

and, second, there are the Karush–Kuhn–Tucker conditions

$$\dot{z} = \dot{\lambda} \tfrac{\partial}{\partial q} \Phi(z,q), \quad \Phi(z,) \leq 0, \ \dot{\lambda} \geq 0, \ \dot{\lambda}\, \Phi(z,q) = 0,$$

where $\Phi : (x,z,q) \mapsto \mathbb{R}$ *is the yield function such that* $\mathbb{Q}(x,z) = \{\, q \mid \Phi(x,z,q) \leq 0 \,\}$*. Our formulation (3.3), which is based on the subdifferential or the normal cone, has the conceptual advantage that it is equally valid for single as well as for multiple yield surfaces (*$\mathbb{Q}(x,z)$ *convex but not necessarily smooth). We refer to [Mie02b, MT01] and to the applications below for more details.*

The first obvious advantage of our energetic formulation for rate–independent material models is that neither the functionals $\mathcal{E}$ and $\mathcal{D}$ (and hence $\widehat{\psi}$ and $\widehat{\Delta}$) nor the desired solutions have to be differentiable. The only appearing derivative is that of the loading ℓ.

The stability condition (S) has a clear mechanical interpretation. Letting $\widehat{z} = z(t)$ we have $\mathcal{D}(z(t),\widehat{z}) = 0$ and the condition implies that $\boldsymbol{\varphi}(t)$ is the global minimizer of $\mathcal{E}(t,\cdot,z(t))$ on $\mathcal{F}$, which gives the elastic equilibrium. Moreover, changing the internal variable from $z(t)$ to $\widehat{z}$ (and adjusting $\widehat{\boldsymbol{\varphi}}$ optimally) the dissipation must be at least as large as the elastic energy release. The internal variable z will change (and dissipate energy) as soon as the elastic energy release is large enough to compensate for the dissipation. Thus, (S) is equivalent to a principle of maximal dissipation. The energy inequality (E) has an obvious interpretation, since the work of the external forces is given by the last term. This abstract setting has applications in many rate–independent continuum models. We refer to [KMS99, GMH02, MTL02] for applications to phase transformations in shape–memory alloys.

The second major advantages of the energetic formulation (S) & (E) is that it immediately gives rise to a natural incremental algorithm where each step is realized as a variational minimization problem. We discretize the time interval $[0,T]$ via $0 = t_0 < t_1 < \ldots < t_N = T$ and give a stable initial condition $(\boldsymbol{\varphi}_0, z_0) \in \mathcal{F}\times\mathcal{Z}$:

(IP) [Incremental Problem]
For $k = 1,\ldots,N$ find iteratively $(\boldsymbol{\varphi}_k, z_k) \in \mathcal{F}\times\mathcal{Z}$ such that

$$(\boldsymbol{\varphi}_k, z_k) \in \operatorname*{arg\,min}_{(\boldsymbol{\varphi},z)\in\mathcal{F}\times\mathcal{Z}} \Big(\mathcal{E}(t_k,\boldsymbol{\varphi},z) + \mathcal{D}(z_{k-1},z) \Big)\,.$$

Here "argmin J" denotes the set of minimizers of the functional J.

The fact that (IP) is very useful is manifested through the following result which states that incremental solutions are always stable and satisfy a discretized version of the energy inequality; for the simple proof see [MT01, MTL02].

Theorem 3.3 *If $(\boldsymbol{\varphi}_k, z_k)_{k=1,\dots,N}$ is a solution of (IP), then for $k = 1, \dots, N$ we have (i) stability of $(\boldsymbol{\varphi}_k, z_k)$ at time t_k, i.e., $\mathcal{E}(t_k, \boldsymbol{\varphi}_k, z_k) \le \mathcal{E}(t_k, \widehat{\boldsymbol{\varphi}}, \widehat{z}) + \mathcal{D}(z_k, \widehat{z})$ for all $(\widehat{\boldsymbol{\varphi}}, \widehat{z})$, and (ii) the two–sided discretized energy estimate*

$$\begin{aligned}
&\mathcal{E}(t_k, \boldsymbol{\varphi}_k, z_k) - \mathcal{E}(t_{k-1}, \boldsymbol{\varphi}_k, z_k) = \textstyle\int_{t_{k-1}}^{t_k} \frac{\partial}{\partial t}\mathcal{E}(s, \boldsymbol{\varphi}_k, z_k)\,\mathrm{d}s \\
&\le \mathcal{E}(t_k, \boldsymbol{\varphi}_k, z_k) - \mathcal{E}(t_{k-1}, \boldsymbol{\varphi}_{k-1}, z_{k-1}) + \mathcal{D}(z_{k-1}, z_k) \\
&\le \textstyle\int_{t_{k-1}}^{t_k} \frac{\partial}{\partial t}\mathcal{E}(s, \boldsymbol{\varphi}_{k-1}, z_{k-1})\,\mathrm{d}s = \mathcal{E}(t_k, \boldsymbol{\varphi}_{k-1}, z_{k-1}) - \mathcal{E}(t_{k-1}, \boldsymbol{\varphi}_{k-1}, z_{k-1}).
\end{aligned}$$

For this theorem we essentially just need that the (possibly unsymmetric) dissipation distance $\mathcal{D}$ satisfies the triangle inequality (2.1). However, for connecting the time–discrete incremental problem to the time–continuous problem (S) & (E) we need that $\mathcal{D}$ is derived from the local infinitesimal metric $\Delta = \widehat{\Delta}(x, z, \dot{z})$, such that (2.2) equals the usual definition of dissipation.

An important feature of (IP) is the local occurrence of z (i.e. no gradients appear in the integrand defining $\mathcal{E}+\mathcal{D}$). This can be used to work out the minimization in z pointwise. We define the reduced potential Ψ^{red} and the update mapping $\widehat{Z}_{\mathrm{new}}$ via

$$\begin{aligned}
\Psi^{\mathrm{red}}(z_{\mathrm{old}}; x, \mathbf{F}) &\stackrel{\mathrm{def}}{=} \min_{z\in Z} \Big[\widehat{\psi}(x, \mathbf{F}, z) + \widehat{D}(x, z_{\mathrm{old}}, z)\Big], \\
\widehat{Z}_{\mathrm{new}}(x, \mathbf{F}, z_{\mathrm{old}}) &\stackrel{\mathrm{def}}{=} \arg\min_{z\in Z} \Big[\widehat{\psi}(x, \mathbf{F}, z) + \widehat{D}(x, z_{\mathrm{old}}, z)\Big].
\end{aligned} \tag{3.4}$$

and choose $z = \widehat{Z}_{\mathrm{new}}(x, \mathbf{F}, z_{\mathrm{old}})$ such that this is a minimizer in the definition of Ψ^{red}. The new constitutive function Ψ^{red} is uniquely defined by $\widehat{\psi}$ and $\widehat{\Delta}$ and it contains the most important information on the combined effect of the elastic storage and the dissipational behavior of the material. Now we have

$$\begin{aligned}
\mathcal{E}^{\mathrm{red}}_{z_{k-1}}(t_k, \boldsymbol{\varphi}) &\stackrel{\mathrm{def}}{=} \min\nolimits_{z\in\mathcal{Z}} \Big(\mathcal{E}(t_k, \boldsymbol{\varphi}, z) + \mathcal{D}(z_{k-1}, z)\Big) \\
&= \textstyle\int_\Omega \Psi^{\mathrm{red}}(z_{k-1}(x); x, \mathrm{D}\boldsymbol{\varphi}(x))\,\mathrm{d}x - \langle \ell(t_k), \boldsymbol{\varphi}\rangle.
\end{aligned}$$

Thus, the kth step of (IP) is solved if we find a minimizer of $\mathcal{E}^{\mathrm{red}}_{z_{k-1}}(t_k, \cdot)$ on $\mathcal{F}$. If $\boldsymbol{\varphi}_k$ is such a minimizer, then $(\boldsymbol{\varphi}_k, z_k)$ with $z_k(x) = \widehat{Z}_{\mathrm{new}}(x, \mathrm{D}\boldsymbol{\varphi}_k(x), z_{k-1}(x))$ is the desired minimizer in (IP).

The minimization problem for $\mathcal{E}^{\mathrm{red}}_{z_{k-1}}(t_k, \cdot)$ has the standard form of a problem of nonlinear elasticity, where $\Psi^{\mathrm{red}}(z_{k-1}(x); x, \cdot)$ plays the role of the elastic potential. Important features of $\widehat{\psi}$ and $\widehat{\Delta}$ are inherited by Ψ^{red}. For instance, if $\widehat{\psi}$ is frame indifferent, then so is Ψ^{red}, i.e. $\Psi^{\mathrm{red}}(z_{\mathrm{old}}; x, \mathbf{F}) = \Psi^{\mathrm{red}}(z_{\mathrm{old}}; x, \mathbf{RF})$ for all $\mathbf{R} \in \mathrm{SO}(\mathbb{R}^d)$. Similarly, material symmetries are inherited.

However, other properties like the coercivity with respect to $\mathbf{F}$ are nontrivial and depend on hardening properties. If coercivity fails, the solvability of (IP) is not guaranteed, and we have to expect fracture or localization (shear bands), cf. [Mie02a]. Moreover, quasi– and rank–one convexity may no longer hold for $\Psi^{\mathrm{red}}(z; x, \cdot)$ and we have to expect the formation of microstructure in infimizing sequences, see [OR99, CHM01].

> *If $\Psi^{\mathrm{red}}(z; x, \cdot)$ is not quasiconvex, then there is a loading $t \mapsto \ell(t)$ such that (IP) and consequently (S) & (E) does not have a solution, because of formation of microstructure.*

It is this problem which is the reason for developing our theory below.

4 Two applications

We give two possible applications of the above theory which are well known to give rise to the formation of microstructure. The first is the case of shape memory alloys, where the internal variable z is a phase indicator taking values in the finite set of possible variants of phases. In the second application we treat a model for finite strain plasticity, where Z contains the matrix group of matrixes with determinant 1. Other possible applications involve damage models as treated in [FM98, HSB02, KMR02].

4.1 Shape–memory alloys

The phase transformations in shape–memory alloys can be described most easily by assuming that the material can choose be be in one of n different phases which are characterized by the phase indicator $z \in Z = \{1, \dots, n\}$ and the associated stored energy densities $\psi = \widehat{\psi}(x, \mathbf{F}, i)$. For instance, in the case of a cubic to tetragonal phase transformation we have $n = 4$ where one phase is the austenite and there are three martensitic phases.

In addition to the elastic properties we also need to prescribe the dissipational behavior. Since the internal variables are discrete, we cannot prescribe an infinitesimal metric $\widehat{\Delta}$ but have to give the dissipation distance $\widehat{D} : Z \times Z \to [0, \infty)$ directly. Hence, we have to give constants $\widehat{D}(i, j) \geq 0$ such that $\widehat{D}(i, i) = 0$ and the triangle inequality holds:

$$\widehat{D}(i, k) \leq \widehat{D}(i, j) + \widehat{D}(j, k) \quad \text{for all } i, j, k \in Z.$$

The value $\widehat{D}(i, j)$ has the interpretation of an energetic threshold value which has to be overcome in order to propagate an interface between phase i and j by a unit volume, see [MTL02, GMH02] for more details.

In particular, it is easy to obtain simple examples where Ψ^{red} is not quasiconvex. Obviously we have

$$\Psi^{\mathrm{red}}(i; x, \mathbf{F}) = \min\{\, \widehat{\psi}(x, \mathbf{F}, j) + \widehat{D}(i, j) \mid j \in Z \,\}.$$

For instance, in the case $\widehat{\psi}(\mathbf{F}, j) = [\mathbf{E} - \boldsymbol{\varepsilon}_j]{:}\mathbb{C}[\mathbf{E} - \boldsymbol{\varepsilon}_j] + \gamma_j$, where $\mathbf{E} = \frac{1}{2}(\mathbf{F} + \mathbf{F}^\top) - \mathbf{I}$, rank–one convexity does not hold whenever $\Psi^{\mathrm{red}}(i; \cdot)$ is not identical $\widehat{\psi}(x, \mathbf{F}, i)$.

In these two references a relaxed, macroscopic problem is studied which is based on phase fractions $\theta \in [0, 1]$. Thus, Z is replaced by the polytope

$$P_n = \{\, \theta = (\theta_1, \dots, \theta_n) \in \mathbb{R}^n \mid \theta_i \geq 0,\ \textstyle\sum_{j=1}^n \theta_j = 1 \,\}. \tag{4.1}$$

We will show below that this model is nothing as a special case of our general relaxation by Young measures.

4.2 Finite strain elastoplasticity with hardening

In the case of elastoplasticity we need to give specific constitutive assumptions, in particular for the internal variables $z \in Z$. We follow here the recent papers [OR99, Mie02b] and refer to [Mie02a] for more details.

Multiplicative elastoplasticity uses the splitting $\mathbf{F} = \mathrm{D}\boldsymbol{\varphi} = \mathbf{F}_{\mathrm{el}}\mathbf{F}_{\mathrm{pl}}$, where $\mathbf{F}_{\mathrm{pl}}$ is an internal variable and $\mathbf{F}_{\mathrm{el}} = \mathbf{F}\mathbf{F}_{\mathrm{pl}}^{-1}$ is the part which accounts for elastic energy and stresses. As internal variables we use

$$z = (\mathbf{P}, p) \in \mathrm{SL}(\mathbb{R}^d)\times\mathbb{R}^m, \quad \text{with } \mathbf{P} = \mathbf{F}_{\mathrm{pl}}^{-1} \text{ and } \mathrm{SL}(\mathbb{R}^d) = \{\, \mathbf{P} \in \mathbb{R}^{d\times d} \mid \det \mathbf{P} = 1 \,\},$$

where $p \in \mathbb{R}^m$ denotes suitable hardening parameters.

For the constitutive functions $\widehat{\psi}$ and $\widehat{\Delta}$ we now specify the associated symmetry conditions. They involve the material symmetry group $S \subset \mathrm{SO}(\mathbb{R}^d)$ and they are supposed to hold for all $(\mathbf{F}, \mathbf{P}, p) \in \mathrm{GL}_+(\mathbb{R}^d)\times\mathrm{SL}(\mathbb{R}^d)\times\mathbb{R}^m$ (for notational simplicity we suppress the possible dependence on $x \in \Omega$):

(Sy1) Frame indifference:
$\widehat{\psi}(\mathbf{RF}, \mathbf{P}, p) = \widehat{\psi}(\mathbf{F}, \mathbf{P}, p)$ for all $\mathbf{R} \in \mathrm{SO}(\mathbb{R}^d)$;

(Sy2) Plastic indifference:
$\widehat{\psi}(\mathbf{F}\mathbf{G}^{-1}, \mathbf{GP}, p) = \widehat{\psi}(\mathbf{F}, \mathbf{P}, p)$, $\widehat{\Delta}(\mathbf{GP}, p, \mathbf{G}\dot{\mathbf{P}}, \dot{p}) = \widehat{\Delta}(\mathbf{P}, p, \dot{\mathbf{P}}, \dot{p})$ for all $\mathbf{G} \in \mathrm{SL}(\mathbb{R}^d)$;

(Sy3) Rate independency:
$\widehat{\Delta}(\mathbf{P}, p, \alpha\dot{\mathbf{P}}, \alpha\dot{p}) = \alpha\widehat{\Delta}(\mathbf{P}, p, \dot{\mathbf{P}}, \dot{p})$ for all $\alpha \geq 0$;

(Sy4) Material symmetry:
$\widehat{\psi}(\mathbf{F}, \mathbf{PS}, \tau_{\mathbf{S}}p) = \widehat{\psi}(\mathbf{F}, \mathbf{P}, p)$ and $\widehat{\Delta}(\mathbf{PS}, \tau_{\mathbf{S}}p, \dot{\mathbf{P}}\mathbf{S}, \tau_{\mathbf{S}}\dot{p}) = \widehat{\Delta}(\mathbf{P}, p, \dot{\mathbf{P}}, \dot{p})$ for all $\mathbf{S} \in S$.

The special assumption for elastoplasticity is the "plastic indifference" (Sy2) which leads to the multiplicative split in $\widehat{\psi}$ as well as to the correct time rates in the flow rules. We find

$$\widehat{\psi}(x, \mathbf{F}, \mathbf{P}, p) = \widetilde{\psi}(x, \mathbf{FP}, p), \qquad \widehat{\Delta}(x, \mathbf{P}, p, \dot{\mathbf{P}}, p) = \widetilde{\Delta}(x, p, \mathbf{P}^{-1}\dot{\mathbf{P}}, \dot{p}), \tag{4.2}$$

and the thermodynamically conjugated forces take the form

$$\begin{aligned} \mathbf{Q} &= -\tfrac{\partial}{\partial \mathbf{P}}\widehat{\psi}(x, \mathbf{F}, \mathbf{P}, p) = -\mathbf{F}^\top \tfrac{\partial}{\partial \mathbf{F}_{\mathrm{el}}}\widetilde{\psi}(x, \mathbf{FP}, p) \in \mathbb{R}^{d\times d}, \\ q &= -\tfrac{\partial}{\partial p}\widehat{\psi}(x, \mathbf{F}, \mathbf{P}, p) = -\tfrac{\partial}{\partial p}\widetilde{\psi}(x, \mathbf{FP}, p) \in \mathbb{R}^m. \end{aligned}$$

The elastic domain $\mathbb{Q}(x, \mathbf{P}, p)$ associated to $\widehat{\Delta}(x, \mathbf{P}, p, \cdot, \cdot)$ is the set of all thermodynamic forces $(\mathbf{Q}, q)$ which are not strong enough to overcome the dissipational friction: $\mathbb{Q}(x, \mathbf{P}, p) = [\partial_{(\dot{\mathbf{P}}, \dot{p})}\widehat{\Delta}(x, \mathbf{P}, p, \cdot, \cdot)](\mathbf{0}, 0)$, see (3.1). Using (Sy2) and the Lie–group structure of $\mathrm{SL}(\mathbb{R}^d)$, which gives

$$\mathbf{V} \in \mathrm{T}_{\mathbf{P}}\mathrm{SL}(\mathbb{R}^d) \iff \boldsymbol{\xi} = \mathbf{P}^{-1}\mathbf{V} \in \mathrm{sl}(\mathbb{R}^d) \stackrel{\text{def}}{=} \mathrm{T}_{\mathbf{I}}\mathrm{SL}(\mathbb{R}^d) = \{\, \boldsymbol{\xi} \in \mathbb{R}^{d\times d} \mid \operatorname{trace}\boldsymbol{\xi} = 0 \,\}.$$

This leads us to the plastically indifferent form of the elastic domain:

$$(\mathbf{Q}, q) \in \mathbb{Q}(x, \mathbf{P}, p) \iff (\mathbf{P}^\top\mathbf{Q}, q) \in \widetilde{\mathbb{Q}}(x, p) \stackrel{\text{def}}{=} \mathbb{Q}(x, \mathbf{I}, p). \tag{4.3}$$

The flow rule (3.3) takes the form

$$(\mathbf{P}^{-1}\dot{\mathbf{P}}, \dot{p}) \in \partial\mathcal{X}_{\widetilde{\mathbb{Q}}(x,p)}(\mathbf{P}^\top\mathbf{Q}, q) = \mathrm{N}_{(\mathbf{P}^\top\mathbf{Q}, q)}\widetilde{\mathbb{Q}}(x, p)$$

which is the well–known associative flow rule of multiplicative elastoplasticity. It contains the "plastically indifferent" plastic rate $\mathbf{P}^{-1}\dot{\mathbf{P}}$ as well as the "plastically indifferent" conjugate force $\mathbf{P}^\top\mathbf{Q} = -\mathbf{F}_{\mathrm{el}}^\top \frac{\partial}{\partial \mathbf{F}_{\mathrm{el}}}\widetilde{\psi}(x, \mathbf{F}_{\mathrm{el}}, p)$.

In Section 7 of [CHM01] the following example for a reduced energy potential in the case of classical isotropic von Mises plasticity with hardening is given. Let

$$\psi = \widetilde{\psi}(\mathbf{F}_{\mathrm{el}}, p) = h(\det \mathbf{F}_{\mathrm{el}}) + \tfrac{\mu}{2} \operatorname{tr} \mathbf{F}_{\mathrm{el}}^\top \mathbf{F}_{\mathrm{el}} + \tfrac{a}{2} p^2 \quad \text{with } h(\delta) = \lambda\delta^2/4 - (\lambda/2+\mu)\log\delta.$$

For the dissipation metric we choose

$$\Delta = \widetilde{\Delta}(p, \boldsymbol{\xi}, \dot{p}) = r\|\boldsymbol{\xi}\| \text{ for } \boldsymbol{\xi} = \boldsymbol{\xi}^\top,\ \operatorname{tr}\boldsymbol{\xi} = 0 \text{ and } \dot{p} + \|\boldsymbol{\xi}\| \leq 0$$

and $\widetilde{\Delta}(p, \boldsymbol{\xi}, \dot{p}) = \infty$ else. Then, $\Psi^{\mathrm{red}}((\mathbf{I}, 0); \cdot) : \mathbb{R}^{d\times d} \to \mathbb{R}$ is not rank–one convex. In particular, the function $\alpha \mapsto \Psi^{\mathrm{red}}((\mathbf{I}, 0); \mathbf{I} + \alpha \mathbf{n}\otimes\mathbf{n})$ is nonconvex.

Thus, for this example the formation of microstructure is to be expected analytically. In [HH01] the occurrence and evolution of microstructure is observed numerically.

5 Relaxation

The general philosophy of relaxation (see [Rou97]) is that of replacing a difficult (nonconvex) problem, which may be unsolvable, by a more general problem, which can be solved easier than the original problem but still maintains the essential features of the original problem. In particular, in problems developing microstructure one is interested in deriving macroscopic models for generalized objects such that their spatial behavior is better and small scales don't need to be resolved. The mathematical process is abstract enough to allow for the identification of new macroscopic quantities and, hopefully, for the derivation of evolution laws for these new quantities. Thus, an efficient calculation of the macroscopic behavior of the essential features is possible.

We return to the abstract setting of inelasticity with the internal variable $z \in Z$ and the energetic formulation of the rate–independent model. Based on the associated incremental problem we formulate basic properties of suitable relaxations. Formation of microstructure occurs, if minimization problems have no solution, cf. [BJ87, CHM01]. This effect is due to material instabilities. We say that the incremental problem (IP) has no solution, if the infimum

$$\alpha_k = \inf\nolimits_{(\boldsymbol{\varphi}, z)\in\mathcal{F}\times\mathcal{Z}} \Big(\mathcal{E}(t_k, \boldsymbol{\varphi}, z) + \mathcal{D}(z_{k-1}, z)\Big)$$

is not attained, i.e., there is no $(\boldsymbol{\varphi}, z) \in \mathcal{F}\times\mathcal{Z}$ with $\alpha_k = \mathcal{E}(t_k, \boldsymbol{\varphi}, z) + \mathcal{D}(z_{k-1}, z)$. In this situation we may consider the following approximate incremental problem.

(AIP)$_\varepsilon$ Given $\varepsilon > 0$ and $z_0 \in \mathcal{Z}$, find $(\boldsymbol{\varphi}_k^\varepsilon, z_k^\varepsilon) \in \mathcal{F}\times\mathcal{Z}$ with

$$\mathcal{E}(t_k, \boldsymbol{\varphi}_k^\varepsilon, z_k^\varepsilon) + \mathcal{D}(z_{k-1}^\varepsilon, z_k^\varepsilon) \leq \varepsilon + \mathcal{E}(t_k, \boldsymbol{\varphi}, z) + \mathcal{D}(z_{k-1}^\varepsilon, z)$$

for all $(\boldsymbol{\varphi}, z) \in \mathcal{F}\times\mathcal{Z}$.

Obviously, this problem has solutions for all $\varepsilon > 0$. The difficult, remaining question is how the solutions $(\boldsymbol{\varphi}_k^\varepsilon, z_k^\varepsilon)$ behave for $\varepsilon \to 0$. We cannot expect pointwise convergence but certain macroscopic quantities should have limits for $\varepsilon \to 0$. To define an abstract notion of relaxation we introduce a generalized convergence "$\xrightarrow{\mathbb{W}}$" on an enlarged space $\mathbb{W}$. This space is connected to $\mathcal{F}\times\mathcal{Z}$ via a continuous embedding $\mathcal{J} : \mathcal{F}\times\mathcal{Z} \mapsto \mathbb{W}$. Prototypes of such embeddings are convex locally compact embeddings as given in [Rou97, Rou00] Moreover, generalized functionals $\mathbb{E} : [0,T]\times\mathbb{W} \to \mathbb{R}$ and $\mathbb{D} : \mathbb{W}\times\mathbb{W} \to [0,\infty]$ replace the elastic functional $\mathcal{E}$ and the dissipation distance $\mathcal{D}$. We define the associated incremental problem for the initial datum $w_0 \in \mathbb{W}$ and the time discretization $0 = t_0 < t_1 < \ldots < t_N = T$ as follows.

($\mathbb{RIP}$) [Relaxed incremental problem]
For $k = 1, \ldots, N$ find iteratively $w_k \in \mathbb{W}$ such that

$$w_k \in \arg\min_{w\in\mathbb{W}} \Big(\mathbb{E}(t_k, w)+\mathbb{D}(w_{k-1}, w)\Big).$$

We do not ask for the conditions $\mathbb{D}(\mathcal{J}(0, z_0), \mathcal{J}(0, z_1)) = \mathcal{D}(z_0, z_1)$ and $\mathbb{E}(t, \mathcal{J}(\boldsymbol{\varphi}, z)) = \mathcal{E}(t, \boldsymbol{\varphi}, z)$. Hence, in general the relaxation will not be an extension.

Definition 5.1 *A 4–tuple* $(\mathbb{W}, \mathcal{J}, \mathbb{E}, \mathbb{D})$ *as defined above is called a* ***lower (or upper) incremental relaxation*** *of* $(\mathcal{F}\times\mathcal{Z}, \mathcal{E}, \mathcal{D})$ *if the following four conditions hold:*

(R1) *[Solvability] For each* $w_0 \in \mathbb{W}$ *the relaxed incremental problem* ($\mathbb{RIP}$) *has a solution.*

(R2) *[Approximation]* $\mathcal{J}(\mathcal{F}\times\mathcal{Z})$ *is dense in* $\mathbb{W}$.

(R3) *[Incremental consistency] If* $(\boldsymbol{\varphi}_k, z_k)_{k=1,\ldots,N}$ *solves (IP), then* $\mathcal{J}(\boldsymbol{\varphi}_k, z_k)_{k=1,\ldots,N}$ *solves* ($\mathbb{RIP}$)*; and vice versa, if* $(w_k)_{k=1,\ldots,N}$ *satisfies* $w_k = \mathcal{J}(\boldsymbol{\varphi}_k, z_k)$ *and solves* ($\mathbb{RIP}$)*, then* $(\boldsymbol{\varphi}_k, z_k)_{k=1,\ldots,N}$ *solves (IP).*

(R4)$_{\text{low}}$ *[for a lower incremental relaxation] For each solution* $(w_k)_{k=1,\ldots,N}$ *of* ($\mathbb{RIP}$)*, there exist solutions* $(\boldsymbol{\varphi}_k^\varepsilon, z_k^\varepsilon)_{k=1,\ldots,N}$ *of (AIP)*$_\varepsilon$ *with* $\mathcal{J}(\boldsymbol{\varphi}_k^\varepsilon, z_k^\varepsilon) \xrightarrow{\mathbb{W}} w_k$ *for* $\varepsilon \to 0$.

(R4)$_{\text{upp}}$ *[for an upper incremental relaxation] If* $\mathcal{J}(\boldsymbol{\varphi}_k^\varepsilon, z_k^\varepsilon) \xrightarrow{\mathbb{W}} w_k$ *and* $(\boldsymbol{\varphi}_k^\varepsilon, z_k^\varepsilon)_{k=1,\ldots,N}$ *solves (AIP)*$_\varepsilon$*, then* $(w_k)_{k=1,\ldots,N}$ *solves* ($\mathbb{RIP}$).

Our definition implies that the relaxed problem has to be of the same energetic kind as the original one; we just give up the clear distinction between $\boldsymbol{\varphi} \in \mathcal{F}$ and $z \in \mathcal{Z}$. Condition (R1) forces us to consider only useful relaxations, namely those which have solutions. If the original problem is already solvable, then we can choose $\mathbb{W} = \mathcal{F}\times\mathcal{Z}$, $\mathbb{E} = \mathcal{E}$ and $\mathbb{D} = \mathcal{D}$, since no relaxation is necessary. Condition (R2) says that the new state space $\mathbb{W}$ should not be unnecessarily big. Condition (R3) is very important as it says that the relaxation must maintain classical solutions, if they exist for (IP) or if they are found by solving ($\mathbb{RIP}$). Conditions (R4)$_{\text{low}}$ and (R4)$_{\text{upp}}$ link the rate–independent evolution of $(\mathcal{F}\times\mathcal{Z}, \mathcal{E}, \mathcal{D})$ to that of $(\mathbb{W}, \mathbb{E}, \mathbb{D})$ via the approximate incremental problem (AIP)$_\varepsilon$.

Moreover the relaxed incremental problem ($\mathbb{RIP}$) can be interpreted as the incremental problem associated to the following relaxed energetic formulation of a rate–independent time–continuous problem: The function $w : [0,T] \mapsto \mathbb{W}$ is called a **solution of the relaxed problem** $(\mathbb{W}, \mathbb{E}, \mathbb{D})$, if ($\mathbb{S}$) and ($\mathbb{E}$) are satisfied:

($\mathbb{S}$) Stability For all $t \in [0,T]$ and all $\widetilde{w} \in \mathbb{W}$ we have

$$\mathbb{E}(t, w(t)) \leq \mathbb{E}(t, \widetilde{w}) + \mathbb{D}(w(t), \widetilde{w}).$$

($\mathbb{E}$) Energy inequality For all $0 \leq t_1 < t_2 \leq T$ we have

$$\mathbb{E}(t_1, w(t_2)) + \mathrm{Diss}^{\mathrm{rel}}(w; [t_1, t_2]) \leq \mathbb{E}(t_1, w(t_1)) - \int_{t_1}^{t_2} \langle \dot{\ell}(s), \mathbf{\Phi}(w(s)) \rangle \, \mathrm{d}s.$$

Here $\boldsymbol{\varphi} = \mathbf{\Phi}(w)$ is the macroscopic deformation $\boldsymbol{\varphi}$ associated to the generalized limit $w \in \mathbb{W}$. The relaxed dissipation is given by $\mathrm{Diss}^{\mathrm{rel}}(w; [t_0, t_1]) = \sup \sum_{j=1}^{N} \mathbb{D}(w(\tau_{j-1}), w(\tau_j))$ where the supremum is taken over all $M \in \mathbb{N}$ and all discretizations $t_0 \leq \tau_0 < \tau_1 < \ldots < \tau_{N-1} < \tau_N \leq t_1$.

A further desirable property for relaxations is the consistency for the time continuous problem:

(R5) [Time–continuous consistency] If $(\boldsymbol{\varphi}, z) : [0,T] \mapsto \mathcal{F} \times \mathcal{Z}$ solves (S) & (E), then $\mathcal{J}(\boldsymbol{\phi}, z) : [0,T] \mapsto \mathbb{W}$ solves ($\mathbb{S}$) & ($\mathbb{E}$); and vice versa, if $w : [0,T] \mapsto \mathbb{W}$ satisfies $w(t) = \mathcal{J}(\boldsymbol{\varphi}(t), z(t))$ and solves ($\mathbb{S}$) & ($\mathbb{E}$), then $(\boldsymbol{\varphi}, z) : [0,T] \mapsto \mathcal{F} \times \mathcal{Z}$ solves (S) & (E).

Another way to define relaxations for rate independent problems of the type (S) & (E) is proposed in [The01]. This definition avoids totally the usage of incremental problems but needs instead a sequence of approximation operators $\mathcal{S}_n : \mathbb{W} \mapsto \mathcal{F} \times \mathcal{Z}$ such that:
(R.i) For all $(\boldsymbol{\varphi}, z) \in \mathcal{F} \times \mathcal{Z}$ we have $\mathcal{S}_n(\mathcal{J}(\boldsymbol{\varphi}, z)) \xrightarrow{\mathbb{W}} (\boldsymbol{\varphi}, z)$ for $n \to \infty$.
(R.ii) For all $t \in [0,T]$ and all $w \in \mathbb{W}$ we have $\mathcal{E}(t, \mathcal{S}_n(w)) \to \mathbb{E}(t, w)$ for $n \to \infty$.
(R.iii) For all $w_{\mathrm{old}}, w_{\mathrm{new}} \in \mathbb{W}$ we have $\mathcal{D}(\mathcal{S}_n(w_{\mathrm{old}}), \mathcal{S}_n(w_{\mathrm{new}})) \to \mathbb{D}(w_{\mathrm{old}}, w_{\mathrm{new}})$ for $n \to \infty$.

An application of this theory to phase transformations is given in [The01].

6 Young measures

A special relaxation can be given by using Young measures which were introduced into the field of continuum mechanics in [BJ87] and were further developed in the last 15 years, see e.g., [Rou97]. One particular instance of the relaxation developed below was studied in [MT99, GMH02, MTL02], where phase transformations in shape–memory alloys were studied. In our terminology of micro–, meso– and macroscopic scales we consider the Young measure as macroscopic object which is obtained by averaging microscopic fluctuations over mesoscopic representative volume elements, see the discussion in the paragraph after (6.2).

We begin by recalling the definition and properties of Young measures, for more details see [Rou97, Mie99]. Recall that Z is assumed to be a closed subset of $\mathbb{R}^n$ (which is not necessarily convex, like $\mathrm{SL}(\mathbb{R}^d) \subset \mathbb{R}^{d \times d}$) and let $\mathcal{Z}^p = \{\, z \in \mathrm{L}^p(\Omega) \mid z(x) \in Z \text{ a.e.} \,\}$ for $p > 1$. We consider sequences of functions $(z^j)_{j \in \mathbb{N}}$ with $z^j \in \mathcal{Z}^p$, which may be obtained from an infimizing sequence in some incremental problem. In general, we may only expect weak convergence, i.e., there exist $z^\infty \in \mathrm{L}^p(\Omega, \mathbb{R}^n)$ such that for all $h \in \mathrm{C}_0(\Omega, \mathbb{R}^n)$ we have

$$\int_\Omega z^j(x) \cdot h(x) \, \mathrm{d}x \to \int_\Omega z^\infty(x) \cdot h(x) \, \mathrm{d}x \text{ for } j \to \infty. \tag{6.1}$$

A major problem with weak convergence is that it is not compatible with nonconvexity of Z. In general the weak limit z^∞ does not lie in $\mathcal{Z}^p$. For instance, $\det \mathbf{P}^j \equiv 1$ does not imply $\det \mathbf{P}^\infty \equiv 1$.

Weak limits are obtained by simple averages over mesoscopic test regions $x_0+\varepsilon Q$, where $Q = (0,1)^d$, and the subsequent limit $j \to \infty$. In particular, (6.1) is equivalent to the following convergence: for all $x_0 \in \Omega$ and $\varepsilon > 0$ small enough we have $\int_{x_0+\varepsilon Q} z^j(x)\,\mathrm{d}x \to \int_{x_0+\varepsilon Q} z^\infty(x)\,\mathrm{d}x$ for $j \to \infty$.

The Young measure does not average over these test regions but rather measures how the values $z(x)$ are distributed throughout the test regions $x_0+\varepsilon Q$. Denote by $\mathrm{Prob}(Z)$ the set of all probability measures on Z. The **Young measures** on Ω with values in Z are given by

$$\mathrm{YM}(\Omega, Z) \stackrel{\mathrm{def}}{=} \{\, \mu : \Omega \mapsto \mathrm{Prob}(Z) \mid \mu \text{ is weakly measurable} \,\}.$$

We say that the sequence $(z^j)_{j\in\mathbb{N}}$ generates the Young measure μ (written $z^j \stackrel{\mathrm{YM}}{\longrightarrow} \mu$) if for each $x_0 \in \Omega$ and sufficiently small ε and every function $g \in C_0(Z,\mathbb{R})$ we have

$$\int_{x_0+\varepsilon Q} g(z^j(x))\,\mathrm{d}x \quad \to \quad \int_{x_0+\varepsilon Q} \int_Z g(z)\mu(x,\mathrm{d}z)\,\mathrm{d}x \quad \text{ for } j \to \infty. \tag{6.2}$$

A function $x \mapsto z(x)$ can be considered as a Young measure by letting $\mu(x) = \delta_{z(x)}$ where δ_a denotes the point mass in a, i.e. $\delta_a(A) = 1$ if $a \in A$ and 0 else.

In the above considerations we consider the small cubes $x+\varepsilon Q$ as mesoscopic representative volume elements which have side length $\varepsilon > 0$. On the one hand, the value $\varepsilon > 0$ is bigger than the microscopical fluctuations which appear in the sequence $z^j : \Omega \to Z$ for j sufficiently large. On the other hand ε is smaller than the typical length scale on which the Young measure $\mu(x) \in \mathrm{Prob}(Z)$ varies with the macroscopic variable $x \in \Omega$. In mathematical terms this is expressed in the fact that the limit $j \to \infty$ has to be done before the limit $\varepsilon \to 0$. If $z^j \stackrel{\mathrm{YM}}{\to} \mu$ we have, for all measurable $W \subset Z$ and almost every $x \in \Omega$, the identity

$$\mu(x, W) = \int_{z\in W} \mu(x,\mathrm{d}z) = \lim_{\varepsilon\to 0} \varepsilon^{-d} \lim_{j\to\infty} \mathrm{vol}(\{\, y \in x+\varepsilon Q \mid z^j(y) \in W \,\}).$$

Given a (dissipation) metric $\widehat{D} : Z\times Z \mapsto [0,\infty]$, we want to generalize it to a metric $\mathbb{D}$ on the space of Young measures which satisfies $\mathbb{D}(\delta_{z_0}, \delta_{z_1}) = \mathcal{D}(z_0, z_1)$ for all $z_0, z_1 \in \mathcal{Z}$. Moreover, it should fit to the microscopic observation that microscopic changes from z_0 to z_1 can be arranged in the mesoscopic region $x_0+\varepsilon Q$ in an optimal way. Here optimality is meant in the sense of minimal dissipation. Assume that at time level t_{k-1} at a macroscopic point $x \in \Omega$ the microscopic distribution of z is given via $\mu_{k-1}(x) \in \mathrm{Prob}(Z)$ and that this distribution has to be changed into $\mu_k(x)$. Assuming that both distributions are realized on the mesoscopic region $x+\varepsilon Q$ via functions $z_{k-1}, z_k : x+\varepsilon Q \to Z$ we have a lot of freedom to choose these functions to make the dissipation minimal. Note that the dissipation only depends only the distribution of the values z_j inside of Z but not on the texture of z_j as a function from $x+\varepsilon Q$ into Z.

This leads directly to the so–called Wasserstein distance between two probability measures. By D^{Wass} we denote Wasserstein metric on $\mathrm{Prob}(Z)$ which associates to $\widehat{D}$:

$$\begin{aligned} D^{\mathrm{Wass}}(x, \nu_0, \nu_1) = \inf \Big\{ \int_{Z\times Z} \widehat{D}(x, z_0, z_1)\sigma(\mathrm{d}z_0, \mathrm{d}z_1) \;\Big|\; & \sigma \in \mathrm{Prob}(Z\times Z), \\ & \int_Z \sigma(\cdot, \mathrm{d}z_1) = \nu_0,\ \int_Z \sigma(\mathrm{d}z_0, \cdot) = \nu_1 \Big\}. \end{aligned} \tag{6.3}$$

For given $\nu_0, \nu_1 \in \mathrm{Prob}(Z)$ we have to transport the mass from the start distribution ν_0 to the final distribution ν_1. Here $\sigma \in \mathrm{Prob}(Z\times Z)$ is a probability measure such that $\sigma(\mathrm{d}z_0, \mathrm{d}z_1)$ gives the amount of mass which from a neighborhood of z_0 is transported into a neighborhood of z_1. We have to infimize the associated dissipation $\widehat{D}$ with respect to σ under the given initial and final distributions.

The Wasserstein metric has a dual representation (which is well–known in probability theory) using Lipschitz continuous functions, cf. [Rac91, Mie99]:

$$D^{\mathrm{Wass}}(x, \nu_0, \nu_1) = \sup\{ \int_Z g(z)\, \nu_1(\mathrm{d}z) - \int_Z g(z)\, \nu_0(\mathrm{d}z) \mid g \in \mathrm{C}_0(Z, \mathbb{R}), \mathrm{Lip}_{\widehat{D}}(g) \leq 1 \}$$

where $\mathrm{Lip}_{\widehat{D}}(g) = \sup_{z_1, z_2 \in Z} |g(z_1) - g(z_2)|/\widehat{D}(x, z_1, z_2)$. Hence, D^{Wass} can be understood as the restriction of a Banach–space norm to the convex set $\mathrm{Prob}(Z)$.

For two Young measures $\mu_0, \mu_1 \in \mathrm{YM}(\Omega, Z)$ we define the dissipation distance via

$$\mathbb{D}(\mu_0, \mu_1) = \int_{x\in\Omega} D^{\mathrm{Wass}}(x, \mu_0(x), \mu_1(x))\, \mathrm{d}x,$$

which is an integral over a local function in the macroscopic variable $x \in \Omega$.

There is a special case of Young measures, namely those which are generated by sequences which are gradients of functions, like the deformation gradients in continuum mechanics $\mathbf{F}^j = \mathrm{D}\boldsymbol{\varphi}^j \in \mathbb{R}^{d\times d}$. The Young measures obtained as limits from gradients form a subset of all Young measures and are called **gradient Young measures** (GYM). They are characterized as follows (cf. [KP94, Rou97]):

$$\begin{aligned}
&\gamma \in \mathrm{GYM}(\Omega, \mathbb{R}^{d\times d}) \iff \begin{cases} \text{There exists } \boldsymbol{\varphi} : \Omega \mapsto \mathbb{R}^d \text{ such that} \\ \text{for almost every (a.e.) } x \in \Omega \text{ we have} \\ \gamma(x) \in \mathrm{Prob}^{\mathrm{grad}}_{\mathrm{D}\boldsymbol{\varphi}(x)}(\mathbb{R}^{d\times d}), \end{cases} \\
&\text{where} \\
&\mathrm{Prob}^{\mathrm{grad}}(\mathbb{R}^{d\times d}) \stackrel{\mathrm{def}}{=} \Big\{ \nu \in \mathrm{Prob}(\mathbb{R}^{d\times d}) \,\Big|\, \text{for all quasiconvex } h : \mathbb{R}^{d\times d} \mapsto \mathbb{R} \\
&\qquad\qquad \text{we have } h\big(\textstyle\int_{\mathbb{R}^{d\times d}} \mathbf{G}\, \nu(\mathrm{d}\mathbf{G})\big) \leq \int_{\mathbb{R}^{d\times d}} h(\mathbf{G})\, \nu(\mathrm{d}\mathbf{G}) \Big\} \\
&\text{and } \mathrm{Prob}^{\mathrm{grad}}_{\mathbf{F}}(\mathbb{R}^{d\times d}) \stackrel{\mathrm{def}}{=} \{ \nu \in \mathrm{Prob}^{\mathrm{grad}}(\mathbb{R}^{d\times d}) \mid \textstyle\int_{\mathbb{R}^{d\times d}} \mathbf{G}\, \nu(\mathrm{d}\mathbf{G}) = \mathbf{F} \}.
\end{aligned} \tag{6.4}$$

We shortly write $\gamma \in \mathrm{GYM}_{\boldsymbol{\varphi}}(\Omega, \mathbb{R}^{d\times d})$ if $\gamma(x) \in \mathrm{Prob}^{\mathrm{grad}}_{\mathrm{D}\boldsymbol{\varphi}(x)}(\mathbb{R}^{d\times d})$ for a.e. $x \in \Omega$.

The above condition (b) is essential as it reflects the information of compatibility of different gradients. For instance, $\nu = \theta\delta_A + (1-\theta)\delta_B$ with $\theta \in (0,1)$ and $A, B \in \mathbb{R}^{d\times d}$ lies in $\mathrm{Prob}^{\mathrm{grad}}(\mathbb{R}^{d\times d})$ if and only if $\mathrm{rank}(A-B) \leq 1$. In contrast to $\mathrm{Prob}(\mathbb{R}^{d\times d})$ and $\mathrm{YM}(\Omega, \mathbb{R}^{d\times d})$, the sets $\mathrm{Prob}^{\mathrm{grad}}(\mathbb{R}^{d\times d})$ and $\mathrm{GYM}(\Omega, \mathbb{R}^{d\times d})$ are no longer convex which makes calculations with GYM much more difficult than those with simple YM. However, for fixed $\mathbf{F}$ the set $\mathrm{Prob}^{\mathrm{grad}}_{\mathbf{F}}(\mathbb{R}^{d\times d})$ is convex, and consequently $\mathrm{GYM}_{\boldsymbol{\varphi}}(\Omega, \mathbb{R}^{d\times d})$ is convex for fixed $\boldsymbol{\varphi}$.

A major problem is that the set of quasiconvex functions appearing in the definition of $\mathrm{Prob}^{\mathrm{grad}}$ is not very well understood. The construction of approximating sequences with classical functions or other approximations are rather easy for usual YM, but very difficult and not really understood for GYM, see [KP94, Rou97]. This is due to the fact that the set of quasiconvex functions $h : \mathbb{R}^{d\times d} \to \mathbb{R}$ is not understood at the first place. There is one nontrivial construction for GYM which is called subsequent lamination. We will discuss this procedure shortly in Section 8.

7 Separate relaxation

We propose a relaxation which uses the weak convergence in $\mathcal{F} \subset \mathrm{W}^{1,p}(\Omega, \mathbb{R}^d)$, and the Young measures $\mu \in \mathrm{YM}(\Omega, Z)$ to replace the states $z \in \mathcal{Z}$. Hence, in the abstract setting we let

$$w = (\boldsymbol{\varphi}, \mu) \in \mathcal{F} \times \mathrm{YM}(\Omega, Z) =: \mathbb{W} \quad \text{and} \quad \mathcal{J}(\boldsymbol{\varphi}, z) = (\boldsymbol{\varphi}, \delta_z).$$

For the convergence in $\mathbb{W}$ we choose the definition $(\boldsymbol{\varphi}^j, \mu^j) \xrightarrow{\mathbb{W}} (\boldsymbol{\varphi}^\infty, \mu^\infty)$ if and only if

$$\begin{cases} \boldsymbol{\varphi}^j \rightharpoonup \boldsymbol{\varphi}^\infty \text{ in } \mathrm{W}^{1,p}(\Omega, \mathbb{R}^d) \quad \text{and} \\ \forall\, g \in \mathrm{C}_0(Z) : \int_Z g(z)\, \mu^j(\mathrm{d}z) \overset{*}{\rightharpoonup} \int_Z g(z)\, \mu^\infty(\mathrm{d}z) \text{ in } \mathrm{L}^1(\Omega). \end{cases}$$

The convergence $\mu^j \overset{\mathrm{YM}}{\to} \mu^\infty$ as defined above is also called Young measure convergence and it makes the set $\{\, \delta_z \mid z \in \mathcal{Z} \,\}$ dense in $\mathrm{YM}(\Omega, Z)$, cf. (R2) and (6.2).

The relaxed functionals are abstractly defined as Young–measure relaxation separately for $\mathcal{E}$ and $\mathcal{D}$:

$$\begin{aligned} \mathbb{E}(t, \boldsymbol{\varphi}, \mu) &= \inf\{\, \liminf_{m\to\infty} \mathcal{E}(t, \boldsymbol{\varphi}_m, z_m) \mid \mathcal{J}(\boldsymbol{\varphi}_m, z_m) \xrightarrow{\mathbb{W}} (\boldsymbol{\varphi}, \mu) \,\}, \\ \mathbb{D}(\mu_0, \mu_1) &= \inf\{\, \liminf_{m\to\infty} \mathcal{D}(z_{0,m}, z_{1,m}) \mid \delta_{z_{0,m}} \overset{\mathrm{YM}}{\to} \mu_0,\ \delta_{z_{1,m}} \overset{\mathrm{YM}}{\to} \mu_1 \,\}. \end{aligned}$$

These functionals are given as the smallest possible limit of all sequences of classical functions producing the correct limit in $\mathbb{W}$. This abstract setting is not very useful for practical purposes unless the form of $\mathbb{E}$ and $\mathbb{D}$ is made more specific. However, we obtain a general result in terms of relaxation.

Theorem 7.1 *Under suitable technical assumptions the relaxation defined via $\mathbb{W}$, $\mathcal{J}$, $\mathbb{E}$ and $\mathbb{D}$ satisfies the properties (R1)—(R3) in Definition 5.1 as well as (R5).*

Proof: The existence of minimizers for the incremental problem $\mathbb{E}(t, z) + \mathbb{D}(z_{\mathrm{old}}, z) \to \min$ is solvable since by construction $\mathbb{E}(t, \cdot)$ and $\mathbb{D}(z_{\mathrm{old}}, \cdot)$ are weakly lower semicontinuous. Suitable assumptions give coercivity as well, hence infimizing sequences have a weak limit which is a minimizer. The density of $\mathcal{J}(\mathcal{F} \times \mathcal{Z})$ in $\mathbb{W}$ is a standard result in Young measure theory, cf. [Dac89, Rou97].

The consistencies (R3) and (R5) follow from the fact that the relaxation satisfies

$$\inf\{\, \mathcal{E}(t, \boldsymbol{\varphi}, z) + \mathcal{D}(z_{\mathrm{old}}, z) \mid (\boldsymbol{\varphi}, z) \in \mathcal{F} \times \mathcal{Z} \,\} = \inf\{\, \mathbb{E}(t, w) + \mathbb{D}(\delta_{z_{\mathrm{old}}}, w) \mid w \in \mathbb{W} \,\}$$

for all $t \in [0, T]$ and $z_{\mathrm{old}} \in \mathcal{Z}$. Hence, stability in $\mathcal{F} \times \mathcal{Z}$ already implies stability in $\mathbb{W}$. ∎

Remark 7.2 *A major problem in the theory is that we cannot show in general that the separate relaxation also satisfies the properties $(R4)_{\mathrm{low}}$ or $(R4)_{\mathrm{upp}}$. This problem arises from the fact that we neglect the full information on the microstructure in step $k-1$ and just remember the Young measure μ_{k-1}. Thus, we keep track only of the mesoscopic distribution (phase portions) of the fluctuations but not of their microscopic arrangements (micropattern). Hence, our generalized dissipation distance $\mathbb{D}$ may underestimate the dissipation if the distribution of the fluctuation changes only little but the micropattern* *changes a lot. However, in situations we have in mind (elastoplasticity or shape–memory alloys) there is strong evidence that changes of the micropattern are always dominated by changes in the mesoscopic distribution. For instance laminates will not rotate with given phase portions they rather grow by changes in the relative phase portion.*

Hence, we believe that the separate relaxation is a reasonably good model. In [The01] it is proved mathematically for a two–phase shape–memory model that the separate relaxation is an exact relaxation satisfying (R4)$_{\mathrm{low}}$.

To make the above relaxation useful we need integral representations of the relaxed functionals. These follow from abstract results as given in [Dac89, Rou97, MTL02].

Theorem 7.3 *Under suitable technical assumptions the relaxed functionals have the form*

$$\mathbb{E}(t,\boldsymbol{\varphi},\mu) = \int_\Omega \Psi_{\mathrm{rel}}(x,\mathrm{D}\boldsymbol{\varphi}(x),\mu(x))\,\mathrm{d}x - \langle \ell(t),\boldsymbol{\varphi}\rangle,$$
$$\mathbb{D}(\mu_0,\mu_1) = \int_\Omega D^{\mathrm{Wass}}(x,\mu_0(x),\mu_1(x))\,\mathrm{d}x,$$

with D^{Wass} *from (6.3) and the relaxed potential* Ψ_{rel} *is given by*

$$\Psi_{\mathrm{rel}}(x,\mathbf{F},\nu) = \inf\Big\{ \int_{(0,1)^d} \widehat{\psi}(x,\mathbf{F}{+}\mathrm{D}\boldsymbol{\phi}(y),z(y))\,\mathrm{d}y \,\Big|\, \boldsymbol{\phi}\in \mathrm{W}^{1,\infty}_0((0,1)^d), \ \nu(A) = \mathrm{vol}(\{\, y \,|\, z(y)\in A\,\}) \text{ for all } A\subset Z \Big\}. \tag{7.1}$$

As given in the previous section there is now a energetic rate independent model associated to the function space $\mathbb{W} = \mathcal{F}\times \mathrm{YM}\Omega, Z)$ and the generalized functionals $\mathbb{E}$ and $\mathbb{D}$. It is of the same type as the previous problem for $\mathcal{E}$ and $\mathcal{D}$ defined on the space $\mathcal{F}\times\mathcal{Z}$. We only have replaced the (nonconvex) space $\mathcal{Z}$ by the convexification $\mathrm{YM}\Omega, Z)$ and generalized the potentials accordingly. (This corresponds to a convexification in the sense of [Rou97, Rou00].)

It remains the question how the relaxed incremental problem ($\mathbb{R}\mathbb{I}\mathbb{P}$)

$$(\boldsymbol{\varphi}_k,\mu_k)\in \operatorname*{arg\,min}_{(\boldsymbol{\varphi},\mu)\in\mathbb{W}} \int_\Omega \Big(\Psi_{\mathrm{rel}}(\mathrm{D}\boldsymbol{\varphi}(x),\mu(x)){+}D^{\mathrm{Wass}}(\mu_{k-1}(x),\mu(x))\Big)\,\mathrm{d}x - \langle \ell(t_k),\boldsymbol{\varphi}\rangle \tag{7.2}$$

can be solved algorithmically. Here again the locality of the integrand will play a crucial rôle. For ($\mathbb{R}\mathbb{I}\mathbb{P}$) the important theorem 3.3 for the relevance of the incremental problem holds in exactly the same fashion.

Unfortunately the definition of Ψ_{rel} is rather complicated. In particular, it is not obtained by integrating $\psi(x,\mathbf{F},z)$ with respect to $\nu(\mathrm{d}z)$ (cf. Ex. 7.4), since $\mathbf{F}$ is the macroscopic strain while ν relates to microscopic fluctuations of $z\in Z$ which have a counterpart in microscopic fluctuations of the strain. The definition of Ψ_{rel} shows that minimization with respect to the combined microscopic arrangements in (the representative volume element) $(0,1)^d$ of the internal variable z and the fluctuation strain $\mathrm{D}\boldsymbol{\phi}$ is necessary.

Example 7.4 To illustrate the above concept we consider the case $d=1$, $Z=\mathbb{R}$, and $\widehat{\psi}(\mathbf{F},z) = (\mathbf{F}{-}z)^2$. We obtain $\Psi_{\mathrm{rel}}(\mathbf{F},\nu) = [\mathbf{F}{-}\int_{\mathbb{R}} z\,\nu(\mathrm{d}z)\,]^2$, since in (7.1) we may choose any $z:(0,1)\mapsto\mathbb{R}$ generating ν and then define $\boldsymbol{\phi}$ via $\boldsymbol{\phi}(0)=0$ and $\boldsymbol{\phi}'(y) = z(y){-}\int_0^1 z(s)\,\mathrm{d}s$. Using $\int_0^1 z(s)\,\mathrm{d}s = \int_{\mathbb{R}} z\,\nu(\mathrm{d}z)$ the result follows. Moreover, we have

$$\Psi_{\mathrm{rel}}(\mathbf{F},\nu) = \int_{\mathbb{R}} \widehat{\psi}(\mathbf{F},z)\,\nu(\mathrm{d}z) - \Big(\int_{\mathbb{R}} z^2\nu(\mathrm{d}z) - \big[\int_{\mathbb{R}} z\,\nu(\mathrm{d}z)\big]^2\Big)$$

which shows that $\Psi_{\rm rel}$ is strictly less than the "averaged" energy density, if ν is not a Dirac mass.

Using the notion of gradient Young measures as introduced in the previous section we may rewrite the relaxed elastic potential

$$\Psi_{\rm rel}(x,\mathbf{F},\nu) = \inf\Big\{ \int_{\mathbb{R}^{d\times d}\times Z} \widehat{\psi}(x,\mathbf{G},z)\rho(\mathrm{d}\mathbf{G},\mathrm{d}z) \;\Big|\; \rho\in \mathrm{Prob}(\mathbb{R}^{d\times d}\times Z), \\ \int_{\mathbb{R}^{d\times d}} \rho(\mathrm{d}\mathbf{G},\cdot) = \nu,\ \int_Z \rho(\cdot,\mathrm{d}z) \in \mathrm{Prob}^{\rm grad}_{\mathbf{F}}(\mathbb{R}^{d\times d}) \Big\}. \tag{7.3}$$

See [Rou97] for the equivalence of (7.1) and (7.3). This construction is still very involved, and it is necessary to find suitable algorithms to approximate $\Psi_{\rm rel}$. A major advantage of the latter form (7.3) will be worked out in the next section.

Remark 7.5 *Additionally, the minimizing measure ρ in (7.3) can be used for a post–processing to find the microscopic pattern associated with the Young measures (these are the patterns one expects to see in infimizing sequences as well as in experiments). Having the macroscopic solutions $(\boldsymbol{\varphi}_k,\mu_k)\in\mathcal{F}\times\mathrm{YM}(\Omega,Z)$ of (RIP) we may consider the associated $\gamma_k\in \mathrm{GYM}_{\boldsymbol{\varphi}_k}(\Omega,\mathbb{R}^{d\times d})$ which is obtained from the minimizer $\rho_k(x)$ in (7.3) for calculation $\Psi_{\rm rel}(x,\mathrm{D}\boldsymbol{\varphi}_k(x),\mu_k(x))$. Having the GYM $\gamma(x)$ it is possible in simple cases to reconstruct the way the microstructure has to be arranged. For instance if $\mu(x)=\lambda\delta_{\mathbf{A}}+(1{-}\lambda)\delta_{\mathbf{B}}$ with $\lambda\in(0,1)$ and $\mathbf{B}{-}\mathbf{A}=\mathbf{a}\otimes\boldsymbol{\nu}$ then we have laminates with the normal vector $\boldsymbol{\nu}$ with alternating layers of gradients $\mathbf{A}$ and $\mathbf{B}$ with volume fraction λ and $1{-}\lambda$, respectively.*

When replacing the relaxed potential by an approximate function it is important to keep its most important features, particularly its convexity properties.

Proposition 7.6 *The relaxed potential $\Psi_{\rm rel}$ is cross–quasiconvex, i.e. for all $(\mathbf{F},\nu)\in\mathbb{R}^{d\times d}\times\mathrm{Prob}(Z)$ we have*

$$\int_{(0,1)^d} \Psi_{\rm rel}(\mathbf{F}{+}\mathrm{D}\boldsymbol{\phi}(y),\mu(y))\,\mathrm{d}y \geq \Psi_{\rm rel}(\mathbf{F},\nu)$$

for all $\boldsymbol{\phi}\in \mathrm{W}^{1,\infty}_0((0,1)^d)$ and $\mu\in\mathrm{YM}((0,1)^d,Z)$ with $\int_{(0,1)^d}\mu(y)\,\mathrm{d}y=\nu$.

In particular, this implies that $\Psi_{\rm rel}(\cdot,\nu):\mathbb{R}^{d\times d}\to[0,\infty]$ is quasiconvex for each ν and that $\Psi_{\rm rel}(\mathbf{F},\cdot):\mathrm{Prob}(Z)\to[0,\infty]$ is convex for each $\mathbf{F}$.

The proof of these results follows closely those in [FKP94, LDR00].

An important feature of the update algorithm discussed below is that the calculation of the relaxed stored–energy density $\Psi_{\rm rel}$ and the calculation of the Wasserstein distance $D^{\rm Wass}(\nu_{\rm old},\nu_{\rm new})$ have to be done simultaneous. And therefore it is especially useful to employ the reduced energy density $\Psi^{\rm red}$, cf. Remark 8.2(1).

Relaxation of the shape–memory model

We close this section by returning to the rate–independent model for shape–memory alloys with n phases as introduced in Section 4.1. The set $Z = \{1, \dots, n\}$ of internal variables is just used to indicate in which phase the material is at a certain microscopic point. With P_n defined as in (4.1) we obtain

$$\mathrm{Prob}(Z) \cong P_n \quad \text{and} \quad \mathrm{YM}(\Omega, Z) \cong \mathrm{L}^1(\Omega, P_n).$$

Hence a Young measure $\mu \in \mathrm{YM}(\Omega, Z)$ can be identified with a function $\theta : \Omega \to P_n$ where $\theta_j(x) \in [0,1]$ now gives the phase fraction of the jth phase in the mesoscopic representative volume element $x{+}\varepsilon Q$.

The relaxed energy potential Ψ_{rel} has the form $\psi = \Psi_{\mathrm{rel}}(x, \mathbf{F}, \theta)$, where $\Psi_{\mathrm{rel}}(x, \cdot, \cdot) : \mathbb{R}^{d\times d} {\times} P_n \to [0, \infty]$ is cross–quasiconvex in the usual sense of [LDR00]. In [MTL02, GMH02] the form of Ψ_{rel} is specified in more detail under the additional assumption that the functions $\widehat{\psi}(x, \cdot, j) : \mathbb{R}^{d\times d} \to [0, \infty]$ are given via linearized elasticity in the form

$$\widehat{\psi}(x, \mathbf{F}, j) = \tfrac{1}{2}\big(\mathbb{C}(x)[\mathbf{E}{-}\mathbf{T}_j(x)]\big) : [\mathbf{E}{-}\mathbf{T}_j(x)] + c_j(x) \quad \text{with } \mathbf{E} = \tfrac{1}{2}\big(\mathbf{F}{+}\mathbf{F}^{\top}\big){-}\mathbf{I},$$

where the elasticity tensor $\mathbb{C}(x)$ is independent of the phase $j \in Z$. Then, one finds

$$\Psi_{\mathrm{rel}}(x, \mathbf{F}, \theta) = \textstyle\sum_{j=1}^{n} \theta_j \widehat{\psi}(x, \mathbf{F}, j) + w_{\mathrm{mix}}(\theta),$$

where the mixture term $w_{\mathrm{mix}} : P_n \to (-\infty, 0]$ is a convex function with $w_{\mathrm{mix}}(e_j) = 0$, see [GMH02] for more details.

Finally we also indicate that the Wasserstein distance associated to the metric $\widehat{D}$ in Z, which is given via $\widehat{D}(j,k) \geq 0$, can be calculated explicitly. One obtains the surprising result that

$$D^{\mathrm{Wass}}(\theta_{\mathrm{old}}, \theta_{\mathrm{new}}) = h(\theta_{\mathrm{new}}{-}\theta_{\mathrm{old}})$$

for a suitable function $h : \mathbb{R}^n \to [0, \infty)$ which is convex and homogeneous of degree one (see Prop. 4.7 in [MTL02]).

8 Algorithms for updating the Young measures

In this section we want to address the question how the incremental problem associated to ($\mathbb{RIP}$) can be solved. For this purpose we cannot use the abstract form using $\mathbb{E}$ and $\mathbb{D}$ nor the more concrete version (7.2) which still used $\Psi_{\mathrm{rel}}(x, \mathbf{F}$,
$nu_{\mathrm{new}})$ and $D^{\mathrm{Wass}}(\nu_{\mathrm{old}}, \nu_{\mathrm{new}})$. The definition of the relaxed elastic potential Ψ_{rel} involves the joint fluctuation measure $\rho \in \mathrm{Prob}(\mathbb{R}^{d\times d}{\times}Z)$ and the definition of the Wasserstein distance D^{Wass} involves a transport measure $\sigma \in \mathrm{Prob}(Z{\times}Z)$.

The calculation of all these measures must be part of the algorithm for solving ($\mathbb{RIP}$). However, we always assume that the original constitutive functions $\widehat{\psi}$ and the (global) dissipation distance $\widehat{D}$ can be calculated directly (in the latter case this might be difficult if only the infinitesimal metric $\widehat{\Delta}$ is given, cf. [Mie02b]). Later on, we will assume that also the two derived constitutive functions are available, namely the reduced potential $\Psi^{\mathrm{red}}(z_{\mathrm{old}}; x, \mathbf{F})$ and the update function $z_{\mathrm{new}} = \widehat{Z}_{\mathrm{new}}(x, \mathbf{F}, z_{\mathrm{old}})$.

We start with the formulation of ($\mathbb{RIP}$) which is formulated in terms of the joint measure $\tau \in \operatorname{Prob}(\mathbb{R}^{d\times d}\times Z\times Z)$ which contains the joint distributions of $\mathbf{F}$, $z_{\rm new}$ and $z_{\rm old}$, in particular $\int_{\mathbf{G}\in\mathbb{R}^{d\times d}} \tau(\mathrm{d}\mathbf{G},\cdot,\cdot) = \sigma$ and $\int_{z_0\in Z} \tau(\cdot,\cdot,\mathrm{d}z_0) = \rho$. Note that τ contains all the correlations between these distributions; this is necessary in order to obtain a reasonable evolution.

Algorithm for the relaxed incremental problem ($\mathbb{RIP}$).
Start: Choose $\mu_0 \in \mathrm{YM}(\Omega, Z)$, e.g., let $\mu_0(x) = \delta_{z_0(x)}$.
Iterate for $k = 1, 2, \ldots, N$:

(i) Find a minimizer $(\boldsymbol{\varphi}_k, \tau_k) \in \mathcal{F}\times\mathrm{YM}(\Omega, \mathbb{R}^{d\times d}\times Z\times Z)$ of the functional

$$\mathbb{I}(t_k, \boldsymbol{\varphi}, \tau) \stackrel{\rm def}{=} \int_\Omega \int_{\mathbb{R}^{d\times d}\times Z\times Z} [\widehat{\psi}(x, \mathbf{G}, z_1) + \widehat{D}(x, z_0, z_1)]\tau(x; \mathrm{d}\mathbf{G}, \mathrm{d}z_1, \mathrm{d}z_0)\mathrm{d}x - \langle \ell(t_k), \boldsymbol{\varphi}\rangle$$

over all $(\boldsymbol{\varphi}, \tau) \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^d)\times\mathrm{YM}(\Omega, \mathbb{R}^{d\times d}\times Z\times Z)$ satisfying

$$\begin{aligned} &\int_{Z\times Z} \tau(\cdot;\cdot, \mathrm{d}z_1, \mathrm{d}z_0) \in \mathrm{GYM}_{\boldsymbol{\varphi}}(\Omega, \mathbb{R}^{d\times d}) \qquad \text{(cf. (6.4))},\\ &\int_{\mathbb{R}^{d\times d}\times Z} \tau(\cdot; \mathrm{d}\mathbf{G}, \mathrm{d}z_1, \cdot) = \mu_{k-1} \in \mathrm{YM}(\Omega, Z). \end{aligned} \tag{8.1}$$

(ii) Let $\mu_k(x) = \int_{\mathbb{R}^{d\times d}\times Z} \tau_k(x; \mathrm{d}\mathbf{G}, \cdot, \mathrm{d}z_0)$

(iii) If $k = N$ exit, else increase k and return to (i).

The memory of the microstructure distribution μ_{k-1} is implemented in the second condition on τ_k in (8.1) and the new distribution μ_k is transported to the next step via (ii) and (iii).

The complexity of this minimization problem can be reduced by minimizing internally with respect to the variable z_1. For this it is essential to have a good algorithm to calculate the reduced potential $\Psi^{\rm red} : Z\times\Omega\times\mathbb{R}^{d\times d} \to [0, \infty)$ and the update function $\widehat{Z}_{\rm new} : \Omega\times\mathbb{R}^{d\times d}\times Z \to Z$. Note that these functions are classical function whose arguments are finite–dimensional, not involving any measure. We find the following reduced form for one time step of ($\mathbb{RIP}$). For its formulation we introduce the abbreviation

$$\mathcal{P}(\mathbf{F}, \nu) \stackrel{\rm def}{=} \Big\{\sigma \in \operatorname{Prob}(\mathbb{R}^{d\times d}\times Z) \,\Big|\, \int_Z \sigma(\cdot, \mathrm{d}z) \in \operatorname{Prob}^{\rm grad}_{\mathbf{F}}(\mathbb{R}^{d\times d}),\ \int_{\mathbb{R}^{d\times d}} \sigma(\mathrm{d}\mathbf{G}, \cdot) = \nu\Big\},$$

where $\mathbf{F} \in \mathbb{R}^{d\times d}$ and $\nu \in \operatorname{Prob}(Z)$ are given.

Proposition 8.1 *Using $\Psi^{\rm red}$ and $\widehat{Z}_{\rm new}$ (cf. (3.4)) the steps (i) and (ii) in the algorithm for ($\mathbb{RIP}$) (see (8.1)) can be replaced by the following equivalent algorithm:*

$\widetilde{\text{(i)}}$ *Find a minimizer $(\boldsymbol{\varphi}_k, \sigma_k) \in \mathcal{F}\times\mathrm{YM}(\Omega, \mathbb{R}^{d\times d}\times Z)$ of the functional*

$$\widetilde{\mathbb{I}}(t_k, \boldsymbol{\varphi}, \sigma) \stackrel{\rm def}{=} \int_\Omega \int_{\mathbb{R}^{d\times d}\times Z\times Z} \Psi^{\rm red}(z_0; x, \mathbf{G})\, \sigma(x; \mathrm{d}\mathbf{G}, \mathrm{d}z_0)\mathrm{d}x - \langle \ell(t_k), \boldsymbol{\varphi}\rangle$$

over all $(\boldsymbol{\varphi}, \sigma) \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^d)\times\mathrm{YM}(\Omega, \mathbb{R}^{d\times d}\times Z)$ satisfying $\sigma(x) \in \mathcal{P}(\mathrm{D}\boldsymbol{\varphi}_k(x), \mu_{k-1}(x))$ for a.e. $x \in \Omega$.

$\widetilde{\text{(ii)}}$ *Let* $\mu_k(x) = \widehat{Z}_{\mathrm{new}}(x,\cdot,\cdot)|^*\sigma_k(x)$.

Here the push forward $\widetilde{\nu} = \Phi|^*\nu \in \mathrm{Prob}(V)$ of a measure $\nu \in \mathrm{Prob}(U)$ by a measurable mapping $\Phi : U \mapsto V$ is defined via $\widetilde{\nu}(A) \overset{\mathrm{def}}{=} \nu(\Phi^{-1}(A))$ for all measurable subsets $A \subset V$. In particular, for $\nu = \sum_{j=1}^n \theta_j \delta_{u_j}$ we obtain $\widetilde{\nu} = \sum_{j=1}^n \theta_j \delta_{v_j}$ with $v_j = \Phi(u_j)$.

Proof: The essential point is the minimization of the inner integral in the definition of $\mathbb{I}(t_k, \boldsymbol{\varphi}, \tau)$ with respect to τ. Thus, we keep $x \in \Omega$ fixed and suppress it in the remainder of the proof. The minimum is attained for some $\kappa \in \mathrm{Prob}(\mathbb{R}^{d\times d}\times Z\times Z)$ satisfying the constraints $\int_{\mathbb{R}^{d\times d}\times Z} \kappa(\mathrm{d}\mathbf{G}, \mathrm{d}z_1, \cdot) = \nu_{\mathrm{old}} = \mu_{k-1}(x)$ and $\int_{Z\times Z} \kappa(\cdot, \mathrm{d}z_1, \mathrm{d}z_0) \in \mathrm{Prob}^{\mathrm{grad}}_{\mathrm{D}\boldsymbol{\varphi}_k(x)}(\mathbb{R}^{d\times d})$. First assume that κ is a finite convex combination of Dirac masses: $\kappa = \sum_{i=1}^n \lambda_i \delta_{(\mathbf{G}_i, z_{1,i}, z_{0,i})}$. After adding "zeros" we may write κ in the form

$$\kappa = \textstyle\sum_{j,l=1}^m \theta_{j,l}\delta_{(\mathbf{G}_j, z_{1,j,l}, z_{0,l})} \quad \text{with } \theta_{j,l} \geq 0 \text{ and } \textstyle\sum_{j,l=1}^m \theta_{j,l} = 1.$$

Minimizing $a \overset{\mathrm{def}}{=} \int_{\mathbb{R}^{d\times d}\times Z\times Z}[\widehat{\psi}(x, \mathbf{G}, z_1)+\widehat{D}(x, z_0, z_1)]\kappa(\mathrm{d}\mathbf{G}, \mathrm{d}z_1, \mathrm{d}z_0)$ with respect to $z_{1,j,l}$, while keeping $\theta_{j,l}$, $\mathbf{G}_j$, and $z_{0,l}$ fixed, leads to

$$a = \textstyle\sum_{j,l=1}^m \theta_{j,l}\Psi^{\mathrm{red}}(z_{0,l}; \mathbf{G}_j) \quad \text{and} \quad z_{1,j,l} = \widehat{Z}_{\mathrm{new}}(\mathbf{G}_j, z_{0,j}).$$

With $\sigma \overset{\mathrm{def}}{=} \sum_{j,l=1}^m \theta_{j,l}\delta_{(\mathbf{G}_j, z_{0,l})} \in \mathcal{P}(\mathrm{D}\boldsymbol{\varphi}_k(x), \nu_o ld(x))$ we find $a = \int_{\mathbb{R}^{d\times d}\times Z} \Psi^{\mathrm{red}}(z; \mathbf{G})\,\sigma(\mathrm{d}\mathbf{G}, \mathrm{d}z)$ and $\nu_{\mathrm{new}} = \sum_{j,l=1}^m \theta_{j,l}\delta_{\widehat{Z}_{\mathrm{new}}(\mathbf{G}_j, z_{0,j})} = \widehat{Z}_{\mathrm{new}}|^*\sigma$. Thus, the desired forms of a and $\nu_{\mathrm{new}} = \mu_k(x)$ are established for the case that κ consists of finite Dirac masses. Using the density of such measures the general result follows. ■

To solve the minimization in $\widetilde{\text{(i)}}$ it is not possible to decompose the problem into alternating minimization with respect to $\boldsymbol{\varphi}$ and σ_k due to the restriction $\int_Z \sigma_k(\cdot,\cdot,\mathrm{d}z) \in \mathrm{GYM}_{\boldsymbol{\varphi}_k}(\Omega, \mathbb{R}^{d\times d})$. However, conceptually we may define the following reduced relaxed potential $\Psi^{\mathrm{red}}_{\mathrm{rel}}$ and obtain a reduced, purely macroscopic minimization problem for $\boldsymbol{\varphi}_k$ which is a classical minimization problem for nonlinear elasticity as is well–known in finite–strain elasticity.

For fixed $\mathbf{F} = \mathrm{D}\boldsymbol{\varphi}(x)$ and $\nu_{\mathrm{old}} = \mu_{k-1}(x)$ we define the reduced relaxed energy density

$$\Psi^{\mathrm{red}}_{\mathrm{rel}}(\nu_{\mathrm{old}}; x, \mathbf{F}) \overset{\mathrm{def}}{=} \min_{\sigma\in\mathcal{P}(\mathbf{F},\nu_{\mathrm{old}})} \int_{\mathbb{R}^{d\times d}\times Z} \Psi^{\mathrm{red}}(z_{\mathrm{old}}; x, \mathbf{G})\,\sigma(\mathrm{d}\mathbf{G}, \mathrm{d}z_{\mathrm{old}}). \tag{8.2}$$

If $\Psi^{\mathrm{red}}_{\mathrm{rel}}(\nu_{\mathrm{old}}; x, \cdot)$ is known, then $\boldsymbol{\varphi}_k$ can be obtained by solving the classical (pseudo–) elastic minimization problem

$$\begin{aligned} &\boldsymbol{\varphi}_k \in \arg\min\nolimits_{\boldsymbol{\varphi}\in\mathcal{F}} \mathbb{E}^{\mathrm{red}}(t_k, \boldsymbol{\varphi}) \quad \text{with} \\ &\mathbb{E}^{\mathrm{red}}(t, \boldsymbol{\varphi}) = \textstyle\int_\Omega \Psi^{\mathrm{red}}_{\mathrm{rel}}(\mu_{k-1}(x); x, \mathrm{D}\boldsymbol{\varphi}(x))\,\mathrm{d}x - \langle \ell(t), \boldsymbol{\varphi}\rangle. \end{aligned} \tag{8.3}$$

However, we recall from step $\widetilde{\text{(ii)}}$ that μ_{k-1} and $\boldsymbol{\varphi}_k$ are not enough to find μ_k. Instead, we need a minimizer $\sigma = \widehat{\Sigma}(x, \mathbf{F}, \nu_{\mathrm{old}}) \in \mathcal{P}(\mathbf{F}, \nu_{\mathrm{old}})$, i.e.,

$$\widehat{\Sigma}(x, \mathbf{F}, \nu_{\mathrm{old}}) \in \underset{\sigma\in\mathcal{P}(\mathbf{F},\nu_{\mathrm{old}})}{\arg\min} \int_{\mathbb{R}^{d\times d}\times Z} \Psi^{\mathrm{red}}(z_{\mathrm{old}}; x, \mathbf{G})\,\sigma(\mathrm{d}\mathbf{G}, \mathrm{d}z_{\mathrm{old}}). \tag{8.4}$$

Remark 8.2

(1) Finding a minimizer $\widehat{\Sigma}$ is in fact a simultaneous relaxation and calculation of the Wasserstein distance from $\nu_{\rm old}$. However, we have contracted the effort to the minimal information needed for realizing the relaxed incremental problem ($\mathbb{RIP}$).

(2) The minimization in (8.4) can't be decomposed into minimizations with respect to Dirac masses for $z_{\rm old}$. More precisely, assume $\nu_{\rm old} = \sum_{j=1}^m \theta_j \delta_{z_j}$ and that $\sigma_j = \widehat{\Sigma}(x, \mathbf{F}, \delta_{z_j}) \in \mathcal{P}(\mathbf{F}, \delta_{z_j})$ are known. Then, $\widetilde{\sigma} = \sum_{j=1}^m \theta_j \sigma_j$ lies in $\mathcal{P}(\mathbf{F}, \nu_{\rm old})$; however, in general it is not a minimizer, cf. Example 7.4. The point is that the measures σ_j may associated to different macroscopic gradients $\mathbf{F}_j$ as long as they are compatible and have the correct average.

Having minimizers $\widehat{\Sigma}(x, \mathbf{F}, \mu_{k-1}(x))$ we are able to find the macroscopic deformation $\boldsymbol{\varphi}_k$ as solution of (8.3). This defines $\sigma_k(x) = \widehat{\Sigma}(x, \mathrm{D}\boldsymbol{\varphi}_k(x), \mu_{k-1}(x))$ and we obtain μ_k via step $\widetilde{\text{(ii)}}$ in Proposition 8.1, i.e., $\mu_k(x) = \widehat{Z}_{\rm new}(x, \cdot)|^* \sigma_k(x)$. This finishes the construction of the solution for the kth step.

For solving the macroscopic problem (8.3) we need $\Psi_k(x, \mathbf{F}) \stackrel{\rm def}{=} \Psi^{\rm red}_{\rm rel}(\mu_{k-1}(x); x, \mathbf{F})$ as well as its first and second derivatives. They can be obtained if the minimizing Young measure $\widehat{\Sigma}(x, \mathbf{F}, \mu_{k-1}(x))$ is known. (Within a finite–element method we need this only for a finite number of nodes $x_j \in \Omega$ and a finite number of strains $\mathbf{F}_l \in \mathbb{R}^{d\times d}$.) To simplify the subsequent formulae we shift this measure in the $\mathbf{G}$–component by its mean value $\mathbf{F}$. Let $T_{\mathbf{F}} : (\mathbf{G}, z) \mapsto (\mathbf{G}{+}\mathbf{F}, z)$ and define $\sigma^{\mathbf{F}}_k(x) \in \mathcal{P}(\mathbf{0}, \mu_{k-1}(x))$ via $\sigma^{\mathbf{F}}_k(x; S) \stackrel{\rm def}{=} \widehat{\Sigma}^{\mathbf{F}}_k(x, \mathbf{F}, \mu_{k-1})(T_{\mathbf{F}} S)$ for all measurable subsets S of $\mathbb{R}^{d\times d}{\times}Z$. Note that the Young measures $\sigma^{\mathbf{F}}_k(x)$ have mean value $\mathbf{0}$ in the $\mathbf{G}$–component

Proposition 8.3 *With the above notations we have the following formulae:*

$$
\begin{aligned}
\Psi_k(x, \mathbf{F}) &= \int_{\mathbb{R}^{d\times d}\times Z} \Psi^{\rm red}(z_{\rm old}, x, \mathbf{F}{+}\mathbf{G})\, \sigma^{\mathbf{F}}_k(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\rm old}),\\
\mathrm{D}_{\mathbf{F}}\Psi_k(x, \mathbf{F}) &= \int_{\mathbb{R}^{d\times d}\times Z} \mathrm{D}_{\mathbf{F}}\Psi^{\rm red}(z_{\rm old}, x, \mathbf{F}{+}\mathbf{G})\, \sigma^{\mathbf{F}}_k(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\rm old}).
\end{aligned}
\tag{8.5}
$$

If $\sigma^{\mathbf{F}}_k(x)$ depends sufficiently smoothly on $\mathbf{F}$ (as a distribution on $\mathbb{R}^{d\times d}{\times}Z$), then for all $\mathbf{F}, \mathbf{H} \in \mathbb{R}^{d\times d}$ we have

$$
\begin{aligned}
\mathrm{D}^2_{\mathbf{F}}\Psi_k(x, \mathbf{F})[\mathbf{H}, \mathbf{H}] &= \int_{\mathbb{R}^{d\times d}\times Z} \mathrm{D}^2_{\mathbf{F}}\Psi^{\rm red}(z_{\rm old}, x, \mathbf{F}{+}\mathbf{G})[\mathbf{H}, \mathbf{H}]\, \sigma^{\mathbf{F}}_k(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\rm old})\\
&\quad + \int_{\mathbb{R}^{d\times d}\times Z} \mathrm{D}_{\mathbf{F}}\Psi^{\rm red}(z_{\rm old}, x, \mathbf{F}{+}\mathbf{G})[\mathbf{H}]\, \mathrm{D}_{\mathbf{F}}\sigma^{\mathbf{F}}_k(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\rm old})[\mathbf{H}]\\
&= \int_{\mathbb{R}^{d\times d}\times Z} \mathrm{D}^2_{\mathbf{F}}\Psi^{\rm red}(z_{\rm old}, x, \mathbf{F}{+}\mathbf{G})[\mathbf{H}, \mathbf{H}]\, \sigma^{\mathbf{F}}_k(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\rm old})\\
&\quad - \int_{\mathbb{R}^{d\times d}\times Z} \Psi^{\rm red}(z_{\rm old}, x, \mathbf{F}{+}\mathbf{G}) \mathrm{D}^2_{\mathbf{F}}\sigma^{\mathbf{F}}_k(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\rm old})[\mathbf{H}, \mathbf{H}].
\end{aligned}
\tag{8.6}
$$

We refer to [BKK00] for a rigorous proof of the second formula in (8.5).

The derivatives of the measures $\sigma^{\mathbf{F}} = \sigma^{\mathbf{F}}_k(x)$ with respect to $\mathbf{F}$ in the direction $\mathbf{H}$ are defined via the linear functional $L_f(\mathbf{F}) \stackrel{\rm def}{=} \int_{\mathbb{R}^{d\times d}\times Z} f(\mathbf{G}, z) \sigma^{\mathbf{F}}(\mathrm{d}\mathbf{G}, \mathrm{d}z)$ and

$$
\begin{aligned}
&\int_{\mathbb{R}^{d\times d}\times Z} f(\mathbf{G}, z) \mathrm{D}_{\mathbf{F}}\sigma^{\mathbf{F}}(\mathrm{d}\mathbf{G}, \mathrm{d}z)[\mathbf{H}] \stackrel{\rm def}{=} \mathrm{D}_{\mathbf{F}} L_f(\mathbf{F})[\mathbf{H}],\\
&\int_{\mathbb{R}^{d\times d}\times Z} f(\mathbf{G}, z) \mathrm{D}^2_{\mathbf{F}}\sigma^{\mathbf{F}}(\mathrm{d}\mathbf{G}, \mathrm{d}z)[\mathbf{H}, \mathbf{H}] \stackrel{\rm def}{=} \mathrm{D}^2_{\mathbf{F}} L_f(\mathbf{F})[\mathbf{H}, \mathbf{H}],
\end{aligned}
$$

where $f : \mathbb{R}^{d\times d}\times Z \to \mathbb{R}$ is an arbitrary smooth test function.

These relations are best understood by assuming that the measure is given as a linear linear combination of Dirac masses, i.e., $\sigma^{\mathbf{F}} = \sum_{j=1}^{m} \Lambda_j(\mathbf{F})\delta_{\widetilde{\mathbf{G}}_j(\mathbf{F}),\widetilde{\mathbf{z}}_j(\mathbf{F})}$. Then, we have $L_f(\mathbf{F}) = \sum_{j=1}^{m} \Lambda_j(\mathbf{F}) f(\widetilde{\mathbf{G}}_j(\mathbf{F}), \widetilde{\mathbf{z}}_j(\mathbf{F}))$ and obtain

$$\mathrm{D}_{\mathbf{F}} L_f(\mathbf{F})[\mathbf{H}] = \sum_{j=1}^{m} \Big(\mathrm{D}_{\mathbf{F}}\Lambda_j[\mathbf{H}] f(\widetilde{\mathbf{G}}_j, \widetilde{\mathbf{z}}_j) + \Lambda_j \mathrm{D}_{(\mathbf{G},z)} f(\widetilde{\mathbf{G}}_j, \widetilde{\mathbf{z}}_j)[(\mathrm{D}\widetilde{\mathbf{G}}_j[\mathbf{H}], \mathrm{D}\widetilde{\mathbf{z}}_j[\mathbf{H}])] \Big),$$

where all functions $\Lambda_j, \widetilde{\mathbf{G}}_j$ and $\widetilde{\mathbf{z}}_j$ and their derivatives are evaluated at $\mathbf{F}$. An analogous formula holds for the second derivative; it involves first and second derivatives of the functions $f, \Lambda_j, \widetilde{\mathbf{G}}_j$ and $\widetilde{\mathbf{z}}_j$.

Altogether we have reduced the incremental problem ($\mathbb{RIP}$) to finding the minimizer $\sigma = \widehat{\Sigma}(x, \mathbf{F}, \nu) \in \mathcal{P}(\mathbf{F}, \nu)$ of (8.4). At present there is no easy method for solving this problem. However, the procedure called **subsequent lamination** (see [Kru98, AP00, Gol01, GMH02]) seems to provide sufficiently good approximations of minimizing gradient Young measures. Its usage in incremental problems was first highlighted in [OR99, ORS00]. We indicate how this procedure can be adapted to the present context where we have to find probability measure σ in $\mathcal{P}(\mathbf{F}, \nu)$, and thus have to take care of the additional internal variable.

Assume that ν has the form $\sum_{j=1}^{J} \theta_j \delta_{z_j}$. We seek $\sigma \in \mathcal{P}(\mathbf{F}, \nu)$ in the form

$$\sigma = \sum_{l=1}^{L} \sum_{j=1}^{J} \lambda_{l,j} \delta_{(\mathbf{G}_l, z_j)} \quad \text{with } \sum_{l=1}^{L} \lambda_{l,j} = \theta_j \text{ for } j = 1, \ldots, J.$$

Letting $\rho_l = \sum_{j=1}^{J} \lambda_{l,j}$ we have to guarantee that $\widehat{\sigma} = \sum_{l=1}^{L} \rho_l \delta_{\mathbf{G}_l}$ lies in $\mathrm{Prob}^{\mathrm{grad}}_{\mathbf{F}}(\mathbb{R}^{d\times d})$, which is the most difficult restriction. Here one uses the idea of sequential lamination. Assume we have a measure σ satisfying all the above restrictions (this can be obtained by adaption from a previous iteration step or simply by letting $L = 1$ and $\mathbf{G}_1 = \mathbf{F}$). Then, a new measure $\widetilde{\sigma} \in \mathcal{P}(\mathbf{F}, \nu)$ can be obtained by replacing one gradient $\mathbf{G}_j$ by two new ones. After renumbering we may assume that we replace the $\delta_{\mathbf{G}_L}$ by $\alpha\delta_{\widetilde{\mathbf{G}}_L} + (1-\alpha)\delta_{\widetilde{\mathbf{G}}_{L+1}}$ where

$$\widetilde{\mathbf{G}}_L = \mathbf{G}_L - (1-\alpha)\mathbf{a}\otimes\mathbf{n} \quad \text{and} \quad \widetilde{\mathbf{G}}_{L+1} = \mathbf{G}_L + \alpha\, \mathbf{a}\otimes\mathbf{n}. \tag{8.7}$$

Here the parameters $\alpha \in (0, 1)$, $\mathbf{a} \in \mathbb{R}^d$ and $\mathbf{n} \in \mathbb{S}^{d-1}$ can be chosen freely. In addition, we have lots of freedom to adjust the new weights $\widetilde{\lambda}_{l,j} \geq 0$, $l = L, L{+}1$, such that all side restrictions hold.

This splitting procedure of $\delta_{\mathbf{G}_L}$ into $\alpha\delta_{\widetilde{\mathbf{G}}_L} + (1-\alpha)\delta_{\widetilde{\mathbf{G}}_{L+1}}$ is called *branching* in [ORS00]. Every sequential lamination can be associated with a tree, where every branch point corresponds to a replacement of one gradient by a compatible pair. A major problem arises in finding good criteria for deciding when such a splitting is necessary (for instance, in the case of loss of rank–one convexity). Moreover, the minimization of $\int_{\mathbb{R}^{d\times d}\times Z} \Psi^{\mathrm{red}}(z; x, \mathbf{G})\, \sigma(\mathrm{d}\mathbf{G}, \mathrm{d}z)$ in the set of all $\sigma \in \mathcal{P}(\mathbf{F}, \nu)$ having a given tree structure is highly nontrivial because to the nonlinear relations given in (8.7). Another aspect of this procedure is that the calculational effort can only be kept reasonable, if one also has good criteria under what conditions a branching can be reversed, i.e., when a laminate can be replaced by a simple gradient. This procedure is called *pruning* (cutting off the end branches). For more details we refer to [ORS00].

9 A one–dimensional example

To illustrate the above concepts we consider a one–dimensional example with $\Omega = (0,1) \subset \mathbb{R}$, i.e., $d=1$ in the above. For the internal variables we use $z = (\mathbf{P}, p) \in Z \overset{\text{def}}{=} (0,\infty)\times\mathbb{R}$ and let in accordance with Section 4

$$\widehat{\psi}(\mathbf{F}, \mathbf{P}, p) = W(\mathbf{F}\mathbf{P}) + h(p) \quad \text{and } \widehat{\Delta}(\mathbf{P}, p, \dot{\mathbf{P}}, \dot{p}) = \widetilde{\Delta}(\dot{\mathbf{P}}/\mathbf{P}, \dot{p}).$$

The hardening energy is $h(p) = \mathrm{e}^{4p} - p$ for $p \geq 0$ and $h(p) = 1+3p$ for $p \leq 0$, moreover

$$W(\mathbf{F}_{\mathrm{el}}) = \begin{cases} \frac{1}{4}(\mathbf{F}_{\mathrm{el}}^4 + \mathbf{F}_{\mathrm{el}}^{-4} - 2) & \text{for } \mathbf{F}_{\mathrm{el}} > 0, \\ \infty & \text{for } \mathbf{F}_{\mathrm{el}} \leq 0; \end{cases} \quad \text{and} \quad \widetilde{\Delta}(\boldsymbol{\xi}, v) = \begin{cases} |\boldsymbol{\xi}| & \text{for } v \geq |\boldsymbol{\xi}|, \\ \infty & \text{else.} \end{cases}$$

We apply external forcing by imposing the boundary conditions $\boldsymbol{\varphi}(t,0) = 0$ and $\boldsymbol{\varphi}(t,1) = \boldsymbol{\varepsilon}(t) \overset{\text{def}}{=} t+1$ and let $\ell(t) = 0$ for the loadings. As initial conditions for the plastic parameters we choose $(\mathbf{P}_0, p_0) \equiv (\mathbf{1}, -1)$.

For the elastic domain $\widetilde{\mathbb{Q}}(p) = \mathbf{P}^\top \mathbb{Q}(\mathbf{P}, p)$ we find $\widetilde{\mathbb{Q}}(p) = \{\, (\boldsymbol{\eta}, q) \mid |\boldsymbol{\eta}| + q \leq 1,\ q \leq 0 \,\}$ and the dissipation distance $\widehat{D} : Z \times Z \mapsto [0, \infty]$ reads

$$\widehat{D}((\mathbf{P}_0, p_0), (\mathbf{P}_1, p_1)) = \begin{cases} |\log(\mathbf{P}_1/\mathbf{P}_0)| & \text{for } p_1 \geq p_0 + |\log(\mathbf{P}_1/\mathbf{P}_0)|, \\ \infty & \text{else.} \end{cases}$$

From this and the special choice of the elastic potential W and the hardening function h the reduced potential can be derived explicitly. We have

$$\Psi^{\mathrm{red}}((\mathbf{P}_0, p_0); \mathbf{F}) = \widetilde{\Psi}^{\mathrm{red}}(p; \mathbf{F}\mathbf{P}_0) \text{ with } \widetilde{\Psi}^{\mathrm{red}}(p, \mathbf{F}_{\mathrm{el}}) = \widetilde{\Psi}^{\mathrm{red}}(p, \mathbf{F}_{\mathrm{el}}^{-1}) \text{ and}$$

$$\widetilde{\Psi}^{\mathrm{red}}(p; \mathbf{F}) = \begin{cases} W(\mathbf{F}) + h(p) & \text{for } 1 \leq \mathbf{F}^4 \leq \mathbf{F}^{-4} + b_p s_p, \\ W(\mathbf{F}_*) + 4\log(\mathbf{F}/\mathbf{F}_*) + h(p) & \text{for } \mathbf{F}_*^4 \leq \mathbf{F}^4 \leq s_p \mathbf{F}_*^4, \\ \frac{1}{2}(\sqrt{1 + b_p \mathbf{F}^4} - 1) - p & \text{for } \mathbf{F}^4 \geq s_p^2(\mathbf{F}^{-4} + b_p); \end{cases}$$

where $b_p = 4\mathrm{e}^{4p}$, $s_p = \max\{1, \mathrm{e}^{-4p}\}$ and $\mathbf{F}_* = (\sqrt{5}+2)^{1/4} \approx 1.435$ is the unique solution of $\mathbf{F}_*^4 = \mathbf{F}_*^{-4} + 4$. (For more details of the calculation and related examples we refer to [Mie02a].)

The most important feature of this model for our present purposes is that the reduced potential $\widetilde{\Psi}^{\mathrm{red}}(p; \cdot)$ is not (quasi–) convex in $\mathbf{F}$ for $p < 0$ on the interval $[\mathbf{F}_*, \mathrm{e}^{-p}\mathbf{F}_*]$. Thus, the system may develop microstructure since the hardening in the corresponding regime is not strong enough. Note that the quasiconvexity occurs only for $\mathbf{F} > 1$, whereas convexity holds for $\mathbf{F} \in (0,1)$. This manifests the typical feature of nonlinear elasticity that compression and extension behave differently.

An important role in the analysis will be played by the convexification of $\widetilde{\Psi}^{\mathrm{red}}(p; \cdot)$. It coincides with $\widetilde{\Psi}^{\mathrm{red}}(p; \cdot)$ outside of an interval which is a strict superset of $[\mathbf{F}_*, \mathrm{e}^{-p}\mathbf{F}_*]$. For $p = -1$ this interval is given by $[\mathbf{F}_0, \mathbf{F}_1]$ with $\mathbf{F}_0 \approx 1.22 < \mathbf{F}_* < \mathrm{e}\mathbf{F} \approx 3.90 < \mathbf{F}_1 \approx 5.6$. On this interval the convexification is given by the segment connecting the two points $(\mathbf{F}_0, \widetilde{\Psi}^{\mathrm{red}}(-1; \mathbf{F}_0))$ and $(\mathbf{F}_1, \widetilde{\Psi}^{\mathrm{red}}(-1; \mathbf{F}_1))$, which is the tangent to $\widetilde{\Psi}^{\mathrm{red}}$ at both points, see

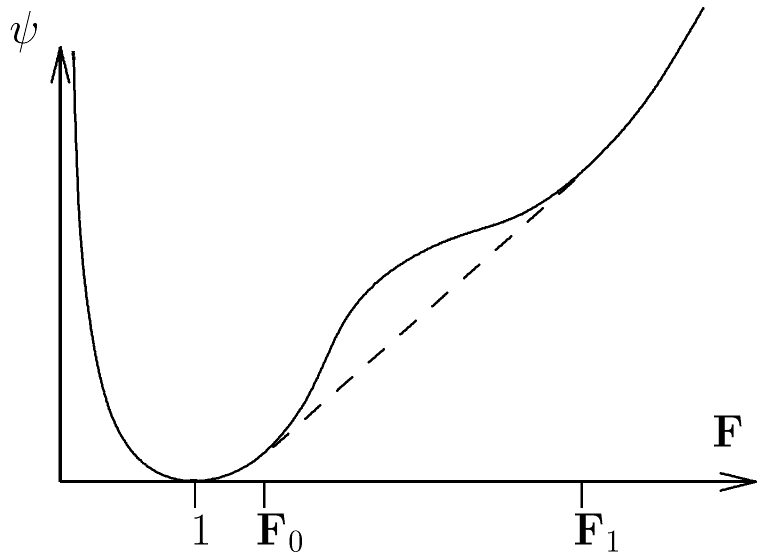

Figure 9.1: The function $\widetilde{\Psi}^{\mathrm{red}}(-1;\,\cdot\,)$ (full line) and its convexification (broken line).

Figure 9.1. We define $(\mathbf{P}_1, p_1) = \widehat{Z}_{\mathrm{new}}(\mathbf{F}_1, (\mathbf{P}_0, p_0))$ which is the unique pair for which $\widetilde{\Psi}^{\mathrm{red}}(p_0; \mathbf{F}_1) = W(\mathbf{F}_1\mathbf{P}_1) + h(p_1) + \widehat{D}((\mathbf{P}_0, p_0), (\mathbf{P}_1, p_1))$.

We look for a spatially homogeneous Young–measure solution $w(t) = (\boldsymbol{\varphi}(t), \nu(t))$ with a trivial macroscopic deformation $\boldsymbol{\varphi}(t, x) = \boldsymbol{\varepsilon}(t)x$ and $\nu(t) \in \mathrm{Prob}(\mathbb{R}^2)$. Using the times t_0 and t_1 given by $\boldsymbol{\varepsilon}(t_j) = \mathbf{F}_j$ the solution can be described a follows:

$t \in [0, t_0]$**:** No plastic flow occurs, the solution has no microstructure: $\nu(t) = \delta_{(\mathbf{1},-1)}$.

$t \in (t_0, t_1)$**:** Microstructure develops such that $\mathbf{F}(t)$ is a mixture of $\mathbf{F}_0$ and $\mathbf{F}_1$. We have the solution $\nu(t) = [1-\theta(t)]\delta_{(\mathbf{1},-1)} + \theta(t)\delta_{(\mathbf{P}_1, p_1)}$ with $\theta(t) = (t-t_0)/(t_1-t_0)$.

$t \in [t_1, \infty)$**:** The microstructure has disappeared and the solution is again classical: $\nu(t) = \delta_{(\mathbf{P}(t), p(t))}$ with $(\mathbf{P}(t), p(t)) = \widehat{Z}_{\mathrm{new}}((\mathbf{1}, -1), \boldsymbol{\varepsilon}(t))$.

We note that for this one–dimensional problem the solutions of ($\mathbb{S}$) & ($\mathbb{E}$) are highly nonunique. Choose any function $\Theta : [t_0, t_1]\times\Omega \to [0, 1]$ such that for all x the function $t \mapsto \Theta(t, x)$ is decreasing and such that $\int_\Omega \Theta(t, x)\,\mathrm{d}x = \theta(t) = (t-t_0)/(t_1-t_0)$. Then, for $t \in [t_0, t_1]$ the pair $(\boldsymbol{\varphi}(t, \cdot), \mu(t, \cdot))$ with

$$\begin{aligned} \boldsymbol{\varphi}(t, x) &= \textstyle\int_0^x \big([1-\Theta(t, y)]\mathbf{F}_0 + \Theta(t, y)\mathbf{F}_1\big)\,\mathrm{d}y \\ \text{and } \mu(t, x) &= [1-\Theta(t, x)]\delta_{(\mathbf{1},0)} + \Theta(t, x)\delta_{(\mathbf{P}_1, p_1)} \end{aligned}$$

defines a solution as well.

Acknowledgment: The research was partially supported by DFG through SFB 404 "Multifield Problems" and through the DFG Priority Programme Multiscale Problems under Mi 459/3–1. The author is grateful for stimulating discussions with Carsten Carstensen, Klaus Hackl, Florian Theil, Christian Miehe, Matthias Lambrecht, and Tomáš Roubíček.

References

[AP00] E. Aranda and P. Pedregal. On the computation of the rank-one convex hull of a function. *SIAM J. Sci. Comput.*, 22(5), 1772–1790 (electronic), 2000.

[BJ87] J. M. Ball and R. D. James. Fine phase mixtures as minimizers of energy. *Arch. Rational Mech. Anal.*, 100(1), 13–52, 1987.

[BKK00] J. M. Ball, B. Kirchheim, and J. Kristensen. Regularity of quasiconvex envelopes. *Calc. Var. Partial Differential Equations*, 11, 333–359, 2000.

[CHM01] C. Carstensen, K. Hackl, and A. Mielke. Nonconvex potentials and microstructure in finite–strain plasticity. *Proc. Royal Soc. London Ser. A*, 2001. In press.

[CP01] C. Carstensen and P. Plecháč. Numerical analysis of a relaxed variational model of hysteresis in two-phase solids. *M2AN Math. Model. Numer. Anal.*, 35(5), 865–878, 2001.

[Dac89] B. Dacorogna. *Direct Methods in the Calculus of Variations.* Springer-Verlag, Berlin, 1989.

[DH98] B. Dacorogna and J.-P. Haeberly. Some numerical methods for the study of the convexity notions arising in the calculus of variations. *RAIRO Modél. Math. Anal. Numér.*, 32(2), 153–175, 1998.

[FKP94] I. Fonseca, D. Kinderlehrer, and P. Pedregal. Energy functionals depending on elastic strain and chemical composition. *Calc. Var. Partial Differential Equations*, 2(3), 283–313, 1994.

[FM98] G. Francfort and J.-J. Marigo. Revisiting brittle fracture as an energy minimization problem. *J. Mech. Phys. Solids*, 46, 1319–1342, 1998.

[GMH02] S. Govindjee, A. Mielke, and G. Hall. The free-energy of mixing for n–variant martensitic phase transformations using quasi–convex analysis. *J. Mech. Physics Solids*, 2002. To appear.

[Gol01] G. H. Goldsztein. The effective energy and laminated microstructures in martensitic phase transformations. *J. Mech. Phys. Solids*, 49, 899–925, 2001.

[Hac97] K. Hackl. Generalized standard media and variational principles in classical and finite strain elastoplasticity. *J. Mech. Phys. Solids*, 45(5), 667–688, 1997.

[HH01] K. Hackl and U. Hoppe. On the calculation of material microstructures using relaxed energies. Talk on IUTAM Symposium, Univ. Stuttgart, 2001.

[HN75] B. Halphen and Q. S. Nguyen. Sur les matériaux standards généralisés. *J. Mécanique*, 14, 39–63, 1975.

[HSB02] K. Hackl and M. Schmidt-Baldassari. Incremental variational principles in damage mechanics. Talk on GAMM Seminar, TU Wien, 2002.

[KMR02] M. Kočvara, A. Mielke, and T. Roubíček. Modeling delamination as a rate-independent process. In preparation, 2002.

[KMS99] M. Kuczma, A. Mielke, and E. Stein. Modelling of hysteresis in two–phase systems. *Archive of Mechanics (Warzawa)*, 51, 693–715, 1999.

[KP94] D. Kinderlehrer and P. Pedregal. Gradient Young measures generated by sequences in Sobolev spaces. *J. Geom. Anal.*, 4, 59–90, 1994.

[Kru98] M. Kružík. Numerical approach to double well problems. *SIAM J. Numer. Anal.*, 35(5), 1833–1849 (electronic), 1998.

[Kru02] M. Kružík. Variational models for microstructure in shape memory alloys and in micromagnetics and their numerical treatment. In A. Ruffing and M. Robnik, editors, *Proceedings of the Bexbach Kolloquium on Science 2000 (Bexbach, 2000).* Shaker Verlag, 2002.

[LDR00] H. Le Dret and A. Raoult. Variational convergence for nonlinear shell models with directors and related semicontinuity and relaxation results. *Arch. Ration. Mech. Anal.*, 154(2), 101–134, 2000.

[Mie96] C. Miehe. Multisurface thermoplasticity for single crystals at large strains in terms of Eulerian vector updates. *Int. J. Solids Struct.*, 33, 1981–2004, 1996.

[Mie99] A. Mielke. Flow properties for Young-measure solutions of semilinear hyperbolic problems. *Proc. Roy. Soc. Edinburgh Sect. A*, 129(1), 85–123, 1999.

[Mie01] C. Miehe. Computational homogenization of materials with microstructures based on incremental variational formulations. Talk on IUTAM Symposium in Stuttgart, 2001.

[Mie02a] A. Mielke. Energetic formulation of multiplicative elastoplasticity using the dissipation metric. In preparation, 2002.

[Mie02b] A. Mielke. Finite elastoplasticity, Lie groups and geodesics on $\mathrm{SL}(d)$. In P. Newton, A. Weinstein, and P. Holmes, editors, *Geometry, Dynamics, and Mechanics.* Springer–Verlag, 2002. To appear.

[MS92] C. Miehe and E. Stein. A canonical model of multiplicative elasto–plasticity. Formulation and aspects of numerical implementation. *Europ. J. Mech. A/Solids*, 11, 25–43, 1992.

[MT99] A. Mielke and F. Theil. A mathematical model for rate–independent phase transformations with hysteresis. In H.-D. Alber, R. Balean, and R. Farwig, editors, *Proceedings of the Workshop on "Models of Continuum Mechanics in Analysis and Engineering"*, pages 117–129. Shaker–Verlag, 1999.

[MT01] A. Mielke and F. Theil. On rate–independent hysteresis models. *Nonl. Diff. Eqns. Appl. (NoDEA)*, 2001. To appear.

[MTL02] A. Mielke, F. Theil, and V. Levitas. A variational formulation of rate–independent phase transformations using an extremum principle. To appear, 2002.

[OR99] M. Ortiz and E. Repetto. Nonconvex energy minimization and dislocation structures in ductile single crystals. *J. Mech. Phys. Solids*, 47(2), 397–462, 1999.

[ORS00] M. Ortiz, E. Repetto, and L. Stainier. A theory of subgrain dislocation structures. *J. Mech. Physics Solids*, 48, 2077–2114, 2000.

[Rac91] S. Rachev. *Probability metrics and the stability of stochastic models.* John Wiley & Sons Ltd., Chichester, 1991.

[Ric71] J. Rice. Inelastic constitutive relations for solids: an internal-variable theory and its applications to metal plasticity. *J. Mech. Phys. Solids*, 19, 203–240, 1971.

[RK02] T. Roubíček and M. Kružík. Mircrostructure evolution model in micromagnetics. *Zeits. angew. math. Physik*, 2002. In print.

[Rou97] T. Roubíček. *Relaxation in Optimization Theory and Variational Calculus.* Walter de Gruyter & Co., Berlin, 1997.

[Rou00] T. Roubíček. Convex locally compact extensions of Lebesgue spaces and their applications. In *Calculus of variations and optimal control (Haifa, 1998)*, pages 237–250. Chapman & Hall/CRC, Boca Raton, FL, 2000.

[Rou02] T. Roubíček. Evolution model for martensitic phase transformation in shape–memory alloys. *Interfaces Free Bound.*, 2002. In print.

[Sim88] J. Simo. A framework for finite strain elastoplasticity based on maximum plastic dissipation and the multiplicative decomposition. I. Continuum formulation. *Comput. Methods Appl. Mech. Engrg.*, 66(2), 199–219, 1988.

[SO85] J. Simo and M. Ortiz. A unified approach to finite deformation elastoplastic analysis based on the use of hyperelastic constitutive relations. *Comput. Methods Appl. Mech. Engrg.*, 49, 221–245, 1985.

[The01] F. Theil. Relaxation of rate independent evolution problems. *Proc. Roy. Soc. Edinb. A*, 2001. To appear.

[ZW87] H. Ziegler and C. Wehrli. The derivation of constitutive relations from the free energy and the dissipation function. In *Advances in applied mechanics, Vol. 25*, pages 183–237. Academic Press, Orlando, FL, 1987.