Deriving new evolution equations for microstructures via relaxation of variational incremental problems[*]

Alexander Mielke[†]

10 August 2003; revised 23 December 2003

Abstract

We study material models for 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 nontrivial 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. The approach is based on a derivative free, energetic formulation using one functional for elastic energy storage and one for the dissipation distance. The latter is derived from a dissipation potential defining the dissipation distance between internal states and hence the Wasserstein distance between Young measures. This approach is strongly linked to an associated time-incremental problem which is a minimization problem of the type used for several years now in the engineering literature. The update algorithm for the incremental problem is discussed in detail, and two simple examples are 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 elasto-plasticity, [SO85, Sim88, MS92, Mie96, MAL02]. Here we especially restrict our attention to models which can be described by purely energetic consideration, i.e., there is an elastic

---

[†]Institut für Analysis, Dynamik und Modellierung; Universität Stuttgart; Germany.

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 by minimizing an update functional. This approach was studied extensively within the last five years, having applications in elasto-plasticity [OR99, ORS00, CHM02, LMD03, Mie02a, Mie03a, HH03, Mie03b], in shape-memory alloys [KMS99, Gol01, MTL02, Kru02, GMH02, Rou02, MR03], in ferro-magnetism [RK02], and in damage models [FM98, HSB02, KMR03].

Of course, such variational, incremental updates where used in the setting of infinitesimal strains much earlier, for instance in Hencky plasticity and in the mathematical existence theory of Moreau [Mor76, Joh76]. For multiplicative plasticity the symmetry of the consistent tangent operators in the incremental problems was also used a lot. Thus, it was more or less implicitly clear that these incremental problems are in fact Euler-Lagrange equations to an update functional. In general one should only expect that the solutions are local minimizers of such a functional. Only for the convex case local minimizers are also global minimizers. Our theory is based completely on the global minimization of the update functionals. This is a major change in the model which leads to certain principal advantages for the modeling and, especially, for the mathematical analysis.

Most prominently is the difficulty of finding a suitable definition for the term "local", a fact which is already discussed in the classical theory of calculus of variations where "weak local minimizers" and "strong local minimizers" are quite different, see [Mor66, Dac89]. Since we want to describe the spontaneous formation of microstructure we certainly need a notion of locality which allows for jumps in the deformation gradient which relates to strong local minimizers. Unfortunately, until now no theory of strong local minimizers is developed. Moreover, the notion of locality should be motivated on physical grounds, for instance through viscous regularizations. For first mathematical attempts in this direction we refer to [EM04], however, there is still a long way to make this theory accessible to nonlinear continuum mechanics. Meanwhile, we use the more phenomenological approach using global minimization and a suitably chosen dissipation potential, as in [MTL02, Sect.3.1] it was shown that for simple examples local minimization can be replaced by global minimization if the dissipation is increased in a corresponding way. Nevertheless, the equations derived here form a new set of equations which sometimes provide different solutions, see the discussion in Section 2.2.

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 elasto-plasticity, see [OR99, ORS00, CHM02, LMD03, Mie02a, Mie03a, HH03]. 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 $\mathbf{F} = \mathrm{D}\boldsymbol{\varphi}$ 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 elasto-plasticity where the plastic deformation $\mathbf{F}_\mathrm{p}$ with $\det \mathbf{F}_\mathrm{p} = 1$ 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 [Mie03b] 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 equality (E), see [MT99, MT03, MTL02] for its first occurrence. This framework has the major advantage that it does not 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 $\varphi \in \mathcal{F}$ 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, Mie02a, CHM02].

The formation of microstructure is now seen in the non-attainment 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 $\mathbf{F} = \mathrm{D}\varphi$ 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 micro-pattern changes, see Section 6.1. 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 a formulation in terms of the macroscopic deformation φ 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, BCH04]. 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, BCH04]. Closely related rate-independent models for ferro-magnetism 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 the control of the correlations between the Young measures on subsequent time steps. Moreover, 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 two examples for our general abstract evolution law. The first is concerned with a one-dimensional example of nonlinear elasto-plasticity involving the multiplicative split of the deformation gradient. We will show that formation of microstructure occurs in this model. The second example is taken from [CT03] and concerns two-dimensional rigid plasticity with a single slip system. For applications in N-phase shape-memory alloys we refer to [MTL02, CP01, GMH02]. Numerical algorithms using similar ideas as proposed here are implemented in two- and three-dimensional elasto-plasticity, see [OR99, HH03, Mie02a, Mie03a, ML03].

2 Rate-independent material models

First we describe the general setup of our material models. They are given as the rate-independent case of generalized standard materials described in terms of the deformation and additional inelastic variables. Second we give a derivative-free energetic formulation, which is based on two functionals and can be generalized to systems with microstructure. Third we provide a time-incremental problem which consists of a sequence of minimization problem with respect to the deformation and the inelastic variables.

2.1 Energy and dissipation functionals

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 $z = (z^1, \dots, z^n) \in Z \subset \mathbb{R}^n$ 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 $\widehat{\psi}(x, \cdot, z) \colon \mathrm{GL}_+(\mathbb{R}^d) \mapsto [0, \infty)$ 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.) Note that $\widehat{\Delta}(x, z, v) = \infty$ is allowed to forbid softening or to include cases where $\dot{z}$ has to lie in a strict subspace of $\mathrm{T}_z Z$, like for $\dot{\mathbf{P}}\mathbf{P}^{-1} = \mathrm{dev}\,\mathrm{sym}\,\mathbf{Q}$.

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

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

Often $\widehat{\Delta}(x, \cdot, \cdot) : \mathrm{T}Z \to [0, \infty]$ is called an infinitesimal (Finsler) metric on 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 $\mathcal{Z} \stackrel{\mathrm{def}}{=} \{ \, z : \Omega \to Z \mid z \text{ measurable } \}$. We define a (possibly unsymmetric) dissipation metric $\mathcal{D}$ on $\mathcal{Z}$ via

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

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

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

which does not contain any derivative.

2.2 The derivative-free energetic formulation

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 2.1 *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*

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

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

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

It should be noted here that this energetic formulation does no longer depend on the derivation of $\mathcal{D}$ from the dissipation potential $\widehat{\Delta}$. It suffices to take any dissipation distance $\mathcal{D} : \mathcal{Z} \times \mathcal{Z} \to [, \infty]$ which satisfies the triangle inequality (2.1) and to define the dissipation of a process via (2.2).

Our energetic formulation using (S) & (E) is a weak form of the classical local flow rules for generalized standard materials ([HN75]). If a solution of (S) & (E) is sufficiently smooth (i.e., $\boldsymbol{\varphi} : [0, T] \to \mathcal{F}$ is continuous and the rate $\dot{z}$ is defined), then the flow laws are satisfied as well. The opposite is not true, see [Mie03b, Sect.5.3] for a counterexample. Using the thermodynamically conjugate variables

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

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

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

$$(\mathrm{E})_{\mathrm{loc}} \qquad \int_\Omega \widehat{\Delta}(x, z(t, x), \dot{z}(t, x)) \, \mathrm{d}x \le \int_\Omega q(t, x) \cdot \dot{z}(t, x) \, \mathrm{d}x.$$

Here $\mathrm{T}_z Z$ denotes the tangent space to Z in z and $\mathrm{T}_z^* Z$ the dual space.

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

$$\partial_v^{\mathrm{sub}} \widehat{\Delta}(x, z, v) = \{ q \in \mathrm{T}_z^* Z \mid \widehat{\Delta}(x, z, v + \widetilde{v}) \ge \widehat{\Delta}(x, z, v) + q \cdot \widetilde{v} \text{ for all } \widetilde{v} \in \mathrm{T}_z Z \}, \quad (2.3)$$

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

$$q(t, x) \in \partial_v^{\mathrm{sub}} \widehat{\Delta}(x, z, \dot{z}) \quad \Longleftrightarrow \quad 0 \in \partial_v^{\mathrm{sub}} \widehat{\Delta}(x, z, \dot{z}) + \frac{\partial}{\partial z} \widehat{\psi}(x, \mathbf{F}, z). \quad (2.4)$$

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

$$\dot{z}(t, x) \in \partial^{\mathrm{sub}} \mathcal{X}_{\mathbb{Q}(x, z(t,x))}(q(t, x)) = \mathrm{N}_{q(t,x)} \mathbb{Q}(x, z(t, x)), \quad (2.5)$$

6

where $\mathcal{X}_{\mathbb{Q}}$ denotes the characteristic function of the set $\mathbb{Q}$ and $\mathrm{N}_q\mathbb{Q}(x,z) \subset \mathrm{T}_z Z$ 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 defining the elastic domain $\mathbb{Q}(x,z) = \{ q \mid \Phi(x,z,q) \leq 0 \}$.

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) relates to the "principle of maximal dissipation". The energy equality (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.

2.3   The variational incremental problem

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

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

Here "$\arg\min J$" denotes the set of all global minimizers of the functional J, i.e., with $\alpha = \min_{u \in \mathcal{U}} J(u)$ we have $\arg\min J \stackrel{\text{def}}{=} \{ u \in \mathcal{U} \mid J(u) = \alpha \}$.

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 equality; for the simple proof see [MT03, MTL02].

Theorem 2.2 *If $(\boldsymbol{\varphi}_0, z_0) \in \mathcal{F} \times \mathcal{Z}$ is stable, i.e., $(\boldsymbol{\varphi}_0, z_0) \in \arg \min \mathcal{E}(0, \widetilde{\boldsymbol{\varphi}}, \widetilde{z}) + \mathcal{D}(z_0, \widetilde{z})$ and 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., $\mathcal{E}(t_k, \boldsymbol{\varphi}_k, z_k) \leq \mathcal{E}(t_k, \widehat{\boldsymbol{\varphi}}, \widehat{z}) + \mathcal{D}(z_k, \widehat{z})$ for all $(\widehat{\boldsymbol{\varphi}}, \widehat{z})$, and

(ii) the two-sided discretized energy estimate

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

For this theorem we just need that the (possibly unsymmetric) dissipation distance $\mathcal{D}$ satisfies the triangle inequality (2.1).

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 condensed potential Ψ^{cond} and the update mapping $\widehat{Z}^{\mathrm{update}}$ via

$$
\begin{aligned}
\Psi^{\mathrm{cond}}(z_{\mathrm{old}}; x, \mathbf{F}) &\overset{\mathrm{def}}{=} \min_{z \in Z} \left[\widehat{\psi}(x, \mathbf{F}, z) + \widehat{D}(x, z_{\mathrm{old}}, z) \right], \\
\widehat{Z}^{\mathrm{update}}(x, \mathbf{F}, z_{\mathrm{old}}) &\overset{\mathrm{def}}{=} \arg \min_{z \in Z} \left[\widehat{\psi}(x, \mathbf{F}, z) + \widehat{D}(x, z_{\mathrm{old}}, z) \right].
\end{aligned}
\tag{2.6}
$$

and choose $z = \widehat{Z}^{\mathrm{update}}(x, \mathbf{F}, z_{\mathrm{old}})$ such that this is a minimizer in the definition of Ψ^{cond}. The new constitutive function Ψ^{cond} 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. See Section 8 for an explicit one-dimensional example and [Mie03b, Mie03c] for an explicit two-dimensional example. Now we have

$$
\begin{aligned}
\mathcal{E}^{\mathrm{cond}}_{z_{k-1}}(t_k, \boldsymbol{\varphi}) &\overset{\mathrm{def}}{=} \min_{z \in \mathcal{Z}} \left(\mathcal{E}(t_k, \boldsymbol{\varphi}, z) + \mathcal{D}(z_{k-1}, z) \right) \\
&= \int_\Omega \Psi^{\mathrm{cond}}(z_{k-1}(x); x, \mathrm{D}\boldsymbol{\varphi}(x)) \, \mathrm{d}x - \langle \ell(t_k), \boldsymbol{\varphi} \rangle.
\end{aligned}
$$

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

The minimization problem for $\mathcal{E}^{\mathrm{cond}}_{z_{k-1}}(t_k, \cdot)$ has the standard form of a problem of nonlinear elasticity, where $\Psi^{\mathrm{cond}}(z_{k-1}(x); x, \cdot)$ plays the role of the elastic potential. Important features of $\widehat{\psi}$ and $\widehat{\Delta}$ are inherited by Ψ^{cond}. For instance, if $\widehat{\psi}$ is frame indifferent, then so is Ψ^{cond}, i.e. $\Psi^{\mathrm{cond}}(z_{\mathrm{old}}; x, \mathbf{F}) = \Psi^{\mathrm{cond}}(z_{\mathrm{old}}; x, \mathbf{R}\mathbf{F})$ 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. [Mie03b]. Moreover, quasi- and rank-one convexity may no longer hold for $\Psi^{\mathrm{cond}}(z; x, \cdot)$ and we have to expect the formation of microstructure in infimizing sequences, see [OR99, CHM02].

> *If $\Psi^{\mathrm{cond}}(z; x, \cdot)$ 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.

3   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 matrices with determinant 1. Other possible applications involve damage models as treated in [FM98, HSB02, KMR03].

3.1   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^{cond} is not quasiconvex. Obviously we have

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

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 $\Psi^{\mathrm{cond}}(i; \cdot)$ 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 $\theta_j \in [0, 1]$. Thus, Z is replaced by the polyhedron

$$P_n = \{ \theta = (\theta_1, \dots, \theta_n) \in \mathbb{R}^n \mid \theta_i \geq 0, \ \textstyle\sum_{j=1}^n \theta_j = 1 \}. \tag{3.1}$$

We will show in Section 6.3 that this model is nothing as a special case of our general relaxation by Young measures.

3.2   Finite strain elasto-plasticity

In the case of elasto-plasticity 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 [Mie03b] for more details.

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

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

where $p \in \mathbb{R}^m$ denotes suitable hardening parameters. As shown in [Mie03b], $\mathbf{P}$ can be taken from any matrix subgroup of $\mathrm{GL}(\mathbb{R}^d)$. In, particular, we consider $\mathrm{GL}(\mathbb{R}^1)$ in Section 8.1 and $\{\, \mathbf{I} + \gamma \mathbf{e}_2 \otimes \mathbf{e}_1 \,|\, \gamma \in \mathbb{R} \,\}$ in Section 8.2.

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

(Sy2) Plastic indifference:
$\widehat{\psi}(\mathbf{F}\mathbf{G}, \mathbf{P}\mathbf{G}, p) = \widehat{\psi}(\mathbf{F}, \mathbf{P}, p)$, $\widehat{\Delta}(\mathbf{P}\mathbf{G}, p, \dot{\mathbf{P}}\mathbf{G}, \dot{p}) = \widehat{\Delta}(\mathbf{P}, p, \dot{\mathbf{P}}, \dot{p})$ 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:
$\widehat{\psi}(\mathbf{F}, \mathbf{S}\mathbf{P}, \tau_{\mathbf{S}}p) = \widehat{\psi}(\mathbf{F}, \mathbf{P}, p)$ and $\widehat{\Delta}(\mathbf{S}\mathbf{P}, \tau_{\mathbf{S}}p, \mathbf{S}\dot{\mathbf{P}}, \tau_{\mathbf{S}}\dot{p}) = \widehat{\Delta}(\mathbf{P}, p, \dot{\mathbf{P}}, \dot{p})$ for all $\mathbf{S} \in S$.

The special assumption for elasto-plasticity 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
$$\widehat{\psi}(x, \mathbf{F}, \mathbf{P}, p) = \widetilde{\psi}(x, \mathbf{F}\mathbf{P}^{-1}, p), \qquad \widehat{\Delta}(x, \mathbf{P}, p, \dot{\mathbf{P}}, p) = \widetilde{\Delta}(x, p, \dot{\mathbf{P}}\mathbf{P}^{-1}, \dot{p}), \qquad (3.2)$$
and the thermodynamically conjugated forces take the form

$$\mathbf{Q} = -\tfrac{\partial}{\partial \mathbf{P}}\widehat{\psi}(x, \mathbf{F}, \mathbf{P}, p) = \mathbf{P}^{-\top}\mathbf{F}^{\top}\tfrac{\partial}{\partial \mathbf{F}_{\mathrm{e}}}\widetilde{\psi}(x, \mathbf{F}\mathbf{P}^{-1}, p)\mathbf{P}^{-\top} \in \mathbb{R}^{d \times d},$$
$$q = -\tfrac{\partial}{\partial p}\widehat{\psi}(x, \mathbf{F}, \mathbf{P}, p) = -\tfrac{\partial}{\partial p}\widetilde{\psi}(x, \mathbf{F}\mathbf{P}^{-1}, p) \in \mathbb{R}^m.$$

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 (2.3). Using (Sy2) and the Lie-group structure of $\mathrm{SL}(\mathbb{R}^d)$, which gives

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

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

$$(\mathbf{Q}, q) \in \mathbb{Q}(x, \mathbf{P}, p) \iff (\mathbf{Q}\mathbf{P}^{\top}, q) \in \widetilde{\mathbb{Q}}(x, p) \overset{\text{def}}{=} \mathbb{Q}(x, \mathbf{I}, p). \qquad (3.3)$$

The flow rule (2.5) takes the form

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

10

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

In [CHM02, Sect.7] the following example for a condensed energy potential in the case of classical isotropic von Mises plasticity with hardening is given. Let

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

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, $\Psi^{\mathrm{cond}}((\mathbf{I}, 0); \cdot) : \mathbb{R}^{d\times d} \to \mathbb{R}$ is not rank-one convex. In particular, the function $\alpha \mapsto \Psi^{\mathrm{cond}}((\mathbf{I}, 0); \mathbf{I} + \alpha\mathbf{n}\otimes\mathbf{n})$ is nonconvex.

Thus, for this example the formation of microstructure is to be expected analytically. In [HH03] the occurrence and evolution of microstructure is observed numerically.

4 Relaxation

The general philosophy of relaxation (see [Rou97]) is that of replacing a difficult (non-convex) 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 do not 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.

4.1 Basic ideas of relaxation

We illustrate the concepts of relaxation with the help of a simple example, which however relates to a single, static problem. Afterwards we return to incremental problems and hence time-dependent problems. More details can be found in [Dac89, Rou97].

Consider the minimization for I on the set $\mathcal{F} = \mathrm{H}^1((0, 2))$, where

$$I(\varphi) = \int_0^2 f(x, \varphi(x), \varphi'(x))\,\mathrm{d}x - \mathrm{e}\varphi(2),$$
$$\text{with } f(x, \varphi, F) = \min\{(F-1)^2, (F+1)^2\}\} + (\varphi - \tfrac{1}{4}x^2)^2.$$

We first note that I does not have a minimizer, i.e., $\arg\min I$ is empty. For this purpose we use the convexification $I^{\mathrm{conv}}(\varphi) = \int_0^2 f^{\mathrm{conv}}(x, \varphi(x), \varphi'(x))\,\mathrm{d}x - \mathrm{e}\varphi(2)$, where

$$f^{\mathrm{conv}}(x, \varphi, F) = (\varphi - \frac{1}{4}x^2)^2 + \begin{cases} (F-1)^2 & \text{for } F \geq 1, \\ 0 & \text{for } |F| \leq 1, \\ (F+1)^2 & \text{for } F \leq -1. \end{cases}$$

Since I^{conv} is strictly convex, it has a unique minimizer φ^* on $\mathcal{F}$ which is given by

$$\varphi^*(x) = \begin{cases} \frac{1}{4}x^2 & \text{for } x \in [0,1], \\ \frac{1}{4}x^2 + \frac{1}{2}(1-\mathrm{e}^{1-x}) & \text{for } x \in [1,2]. \end{cases}$$

In particular, we have the identity $I^{\text{conv}}(\varphi^*) = \min_{\varphi \in \mathcal{F}} I^{\text{conv}}(\varphi)$.

However, we have also $I^{\text{conv}}(\varphi^*) = \inf_{\varphi \in \mathcal{F}} I(\varphi) \lneqq I(\varphi^*)$. To see the first identity it suffices to construct a suitable infimizing sequence $(\varphi_k)_{k \in \mathbb{N}}$ in $\mathcal{F}$ with $I(\varphi_k) \to I^{\text{conv}}(\varphi^*)$. One such sequence is obtained by letting $\varphi_k(x) = \varphi(x)^*$ for $x \in [1,2]$ and taking $\varphi_k(x) \in [\varphi^*(x), \varphi^*(x)+\frac{1}{k}]$ with the additional restriction $|\varphi_k'(x)| = 1$ almost everywhere in $(0,1)$.

The point is that all function $\varphi \in \mathcal{F}$ satisfy $I(\varphi) > \inf I = I^{\text{conv}}(\varphi^*)$. Each test-function stays above the desired value, but one can get as close as one likes. To approach this infimum the functions have to develop more and more structure. It can be shown that every infimizing sequence converges uniformly to φ, i.e., $\sup_{[0,2]} |\varphi_k(x)-\varphi^*(x)| \to 0$ for $k \to \infty$. However, the gradients $\varphi_k'(x)$ will oscillate wildly for $x \in (0,1)$, since we want to have simultaneously $|\varphi_k(x)| \approx 1$ and $\int_{x_1}^{x_2} \varphi_k'(x)\,\mathrm{d}x \approx \varphi^*(x_2)-\varphi^*(x_1) = \frac{x_2+x_1}{2}(x_2-x_1)$.

> A triple $(\mathbb{F}, \mathcal{J}, \mathbb{I})$ is called a **relaxation** of $(\mathcal{F}, I)$, if the following conditions hold.
> (r1) $\mathcal{J}$ is an embedding of $\mathcal{F}$ into $\mathbb{F}$, such that $\mathcal{J}(\mathcal{F})$ is dense in $\mathbb{F}$.
> (r2) $\mathbb{I}(\psi) = \inf\{\, \liminf_{k \to \infty} J(\varphi_k) \mid \mathcal{J}(\varphi_k) \xrightarrow{\mathbb{F}} \psi \,\}$.

For a given problem there are many relaxations. In our above example I^{conv} is the relaxation associated with the choice $\mathbb{F} = \mathcal{F}$. In this case $\mathbb{I} = I^{\text{conv}}$ is simply called the lower semicontinuous envelope. This standard relaxation is especially useful if we are only interested in the macroscopic quantities of almost minimizers, since they are given via φ^*. However, if we need information on the microstructure other relaxations are better.

One such relaxation is the Young measure relaxation which will be discussed in Section 5 in more detail. For the present example we consider $\mathbb{F}$ to be the set of pairs (φ, ν) such that $\varphi \in \mathcal{F}$ and $\nu : [0,2] \to \mathrm{Prob}(\mathbb{R})$ (set of probability measures on $\mathbb{R}$) with the compatibility condition $\varphi'(x) = \int_{\mathbb{R}} F\nu(x, \mathrm{d}F)$ (mean value). Denoting by $\delta_A \in \mathrm{Prob}(\mathbb{R})$ the Dirac distribution located in $A \in \mathbb{R}$ (i.e., $\int_{\mathbb{R}} f(F)\delta_A(\mathrm{d}F) = f(A)$), we define the embedding $\mathcal{J}$ in the canonical way

$$\mathcal{J} : \begin{cases} \mathcal{F} \to & \mathbb{F}, \\ \varphi \mapsto (\varphi, \delta_{\varphi'}). \end{cases}$$

Moreover, we define the relaxed functional $\mathbb{I}$ via

$$\mathbb{I}((\varphi, \nu)) = \int_{x=0}^{2} \int_{F \in \mathbb{R}} f(x, \varphi(x), F)\nu(x, \mathrm{d}F)\,\mathrm{d}x.$$

The important point is now that convergence in $\mathbb{F}$ can lead to ν-components which are no longer Dirac distributions, which may be thought of a functions with uncertain values. In particular, for the above problem every infimizing sequence (φ_k) has the property that it $\mathcal{J}(\varphi_k)$ converges, namely

$$\mathcal{J}(\varphi_k) = (\varphi_k, \delta_{\varphi_k'}) \xrightarrow{\mathbb{F}} (\varphi^*, \nu^*) \quad \text{with } \nu^*(x) = \begin{cases} \frac{2-x}{4}\delta_{-1} + \frac{2+x}{4}\delta_{+1} & \text{for } x \in (0,1), \\ \delta_{\varphi^{*\prime}(x)} & \text{for } x \in (1,2). \end{cases}$$

Of course, we have $(\varphi^*, \nu^*) \in \arg\min_{(\varphi, \nu) \in \mathbb{F}} \mathbb{I}((\varphi, \nu))$.

The solution (φ^*, ν^*) to this relaxed problem contains information on the microstructure which is needed to realize solutions φ_k which are almost minimizing. For $x \in (0,1)$ we see that in a microscopically small neighborhood of x, i.e., $|y - x| \leq \delta \ll 1$ the gradients $\varphi_k'(y)$ have to be either close to -1 or to $+1$, where the proportion for being close to -1 is $\frac{2-x}{4}$ and that for $+1$ is $\frac{2+x}{4}$. It is exactly this kind of information whose evolution we want to understand in the time-dependent setting.

4.2 Relaxation of incremental problems

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, CHM02]. This effect is due to material instabilities. We say that the incremental problem (IP) has no solution, if the infimum

$$\alpha_k = \inf_{(\varphi, z) \in \mathcal{F} \times \mathcal{Z}} \left(\mathcal{E}(t_k, \varphi, z) + \mathcal{D}(z_{k-1}, z) \right)$$

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

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

$$\mathcal{E}(t_k, \varphi_k^\varepsilon, z_k^\varepsilon) + \mathcal{D}(z_{k-1}^\varepsilon, z_k^\varepsilon) \leq \varepsilon + \mathcal{E}(t_k, \varphi, z) + \mathcal{D}(z_{k-1}^\varepsilon, z) \text{ for all } (\varphi, z) \in \mathcal{F} \times \mathcal{Z}.$$

Obviously, this problem has solutions for all $\varepsilon > 0$. The difficult, remaining question is how the solutions $(\varphi_k^\varepsilon, z_k^\varepsilon)$ behave for $\varepsilon \to 0$. As we have seen in the above example, 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 $\mathbb{W}$. As above, 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]
Let $w_0 \in \mathbb{W}$ be given. For $k = 1, \ldots, N$ find iteratively $w_k \in \mathbb{W}$ such that

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

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 $\mathbb{E}(t, \mathcal{J}(\varphi, z)) = \mathcal{E}(t, \varphi, z)$. Hence, in general the relaxation will not be an extension.

Definition 4.1 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)$_{\mathbf{low}}$ *[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)$_{\mathbf{upp}}$ *[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 in the sense that every $w \in \mathbb{W}$ can be approximated by a sequence $(\boldsymbol{\varphi}^\varepsilon, z^\varepsilon)_{\varepsilon > 0}$ of classical elements in $\mathcal{F} \times \mathcal{Z}$, i.e., $\mathcal{J}(\boldsymbol{\varphi}^\varepsilon, z^\varepsilon) \xrightarrow{\mathbb{W}} w$ for $\varepsilon \to 0$. 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)$_{\mathbf{low}}$** and **(R4)$_{\mathbf{upp}}$** 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)$_\varepsilon$**.

4.3 The relaxed time-continuous problem

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 equality** For all $0 \leq t_1 < t_2 \leq T$ we have

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

Here $\boldsymbol{\varphi} = \boldsymbol{\Phi}(w)$ 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 N and all discretizations $t_0 \leq \tau_0 < \ldots < \tau_N \leq t_1$.

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 $\mathcal{J}(\boldsymbol{\varphi}, z) : [0, T] \mapsto \mathbb{W}$ 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 [The02]. 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 $w_{\text{old}}, w_{\text{new}} \in \mathbb{W}$ we have $\mathcal{D}(\mathcal{S}_n(w_{\text{old}}), \mathcal{S}_n(w_{\text{new}})) \to \mathbb{D}(w_{\text{old}}, w_{\text{new}})$ for $n \to \infty$.

An application of this theory to phase transformations is given in [The02].

5 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 (5.2). After introducing the main notations for Young measure, we introduce the so-called Wasserstein distance on measures which measures the minimal dissipation distance between two Young measures. Finally we will study gradient Young measures which arise as Young measures for the deformation gradient.

5.1 Basic notations for Young measures

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 $\text{SL}(\mathbb{R}^d) \subset \mathbb{R}^{d \times d}$) and let $\mathcal{Z}^p = \{\, z \in \text{L}^p(\Omega) \mid z(x) \in Z \text{ a.e.} \,\}$ 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 $z^\infty \in \text{L}^p(\Omega, \mathbb{R}^n)$ such that for all $h \in \text{C}_0(\Omega, \mathbb{R}^n)$ we have

$$\int_\Omega z^j(x) \cdot h(x) \, \mathrm{d}x \to \int_\Omega z^\infty(x) \cdot h(x) \, \mathrm{d}x \text{ for } j \to \infty. \tag{5.1}$$

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, (5.1) 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 (contained in the set of Radon measures $\mathrm{C}_0(Z)^* = \mathrm{Lin}(\mathrm{C}(Z)_0, \mathbb{R})$):

$$\mathrm{Prob}(Z) \stackrel{\mathrm{def}}{=} \{\, \nu \text{ measure on } Z \mid \nu \geq 0, \ \int_Z \nu(\mathrm{d}z) = 1 \,\}.$$

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 $z^j \xrightarrow{\mathrm{YM}} \mu$) if for each $x_0 \in \Omega$ and sufficiently small ε and every function $g \in C_0(Z, \mathbb{R})$ we have

$$\int_{x_0 + \varepsilon Q} g(z^j(x)) \, \mathrm{d}x \quad \rightarrow \quad \int_{x_0 + \varepsilon Q} \int_Z g(z)\mu(x, \mathrm{d}z) \, \mathrm{d}x \quad \text{for } j \to \infty. \tag{5.2}$$

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 $z^j \xrightarrow{\mathrm{YM}} \mu$ we have, for all measurable $W \subset Z$ and almost every $x \in \Omega$, the identity

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

In the sequel we will often consider probability measures on a product space $Z_1 \times Z_2$ which then is denoted by $\mathrm{Prob}(Z_1 \times Z_2)$ (or even $\mathrm{Prob}(Z_1 \times Z_2 \times Z_3)$ for a triple product space; the associated set of YMs is denoted by $\mathrm{YM}(\Omega, Z_1 \times Z_2)$. A measure $\nu \in \mathrm{Prob}(Z_1 \times Z_2)$ denotes a joint distribution of the two variables $z_1 \in Z_1$ and $z_2 \in Z_2$. Often it is necessary to extract the so-called **marginal distributions** $\nu_j = \mathcal{M}_j \nu \in \mathrm{Prob}(Z_j)$ which give the simple distribution of $z_j \in Z_j$ neglecting the correlations to z_{3-j}. Explicitly ν_1 is defined via its action on test functions $g : Z_1 \to \mathbb{R}$, i.e.,

$$\int_{Z_1} g(z_1)\nu_1(\mathrm{d}z_1) = \int_{Z_1 \times Z_2} g(z_1)\nu(\mathrm{d}z_1, \mathrm{d}z_2) \text{ for } g \in \mathrm{C}_0(Z_1, \mathbb{R}).$$

Similarly $\nu_2 = \mathcal{M}_2 \nu \in \mathrm{Prob}(Z_2)$ is defined via integrating out the variable z_1. We may also use the more intuitive form

$$\mathcal{M}_1 \nu = \nu_1 = \int_{z_2 \in Z_2} \nu(\cdot, \mathrm{d}z_2) \quad \text{and} \quad \mathcal{M}_2 \nu = \nu_2 = \int_{z_1 \in Z_1} \nu(\mathrm{d}z_1, \cdot).$$

Clearly $\mathcal{M}_j$ can be seen as a projection on the space of all measures. For a linear combination of Dirac masses $\nu = \sum_{k=1}^n \alpha_k \delta_{(z_1^k, z_2^k)}$ we simply obtain $\nu_j = \sum_{k=1}^n \alpha_k \delta_{z_j^k}$.

16

5.2 The Wasserstein distance between measures

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 $\mu_{k-1}(x) \in \text{Prob}(Z)$ 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 $z_{k-1}^\varepsilon, z_k^\varepsilon : x + \varepsilon Q \to Z$, 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^ε inside of Z but not on the microstructural arrangement of z_j^ε 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 the Wasserstein metric on $\text{Prob}(Z)$ which associates to $\widehat{D}$:

$$D^{\text{Wass}}(x, \nu_0, \nu_1) = \inf \left\{ \int_{Z \times Z} \widehat{D}(x, z_0, z_1) \eta(\mathrm{d}z_0, \mathrm{d}z_1) \mid \eta \in \text{Prob}(Z \times Z), \right.$$
$$\left. \nu_0 = \mathcal{M}_1 \eta, \ \nu_1 = \mathcal{M}_2 \eta \right\}. \tag{5.3}$$

For given $\nu_0, \nu_1 \in \text{Prob}(Z)$ we have to transport the mass from the start distribution ν_0 to the final distribution ν_1. Here $\eta \in \text{Prob}(Z \times Z)$ is a probability measure such that $\eta(\mathrm{d}z_0, \mathrm{d}z_1)$ 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 η under the given initial and final distributions.

We illustrate this by calculating the distance between two measures which have mass only in finitely many points, i.e., $\nu_j = \sum_{k=1}^{m_j} \alpha_k^j \delta_{z_k^j}$ with $z_k^j \in Z$, $\alpha_k^j \geq 0$ and $\sum_{k=1}^{m_j} \alpha_k^j = 1$. For the distance the associated transport measures take the form $\eta = \sum_{k=1}^{m_0} \sum_{l=1}^{m_1} \beta_{k,l} \delta_{z_k^0, z_l^1}$, where $\beta_{k,l} \in [0, 1]$ is the mass which is transported from z_k^0 to z_l^1. Thus, we find

$$D^{\text{Wass}}(x, \nu_0, \nu_1) = \min \{ \sum_{k=1}^{m_0} \sum_{l=1}^{m_1} \beta_{k,l} \widehat{D}(x, z_k^0, z_l^1) \mid \beta_{k,l} \geq 0, \ \sum_{k=1}^{m_0} \beta_{k,l} = \alpha_l^1, \ \sum_{l=1}^{m_1} \beta_{k,l} = \alpha_m^2 \}.$$

Thus, calculating the Wasserstein distance is a classical linear programming problem for the case of measures with finitely many mass points.

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

$$D^{\text{Wass}}(x, \nu_0, \nu_1) = \sup \{ \int_Z g(z) \nu_1(\mathrm{d}z) - \int_Z g(z) \nu_0(\mathrm{d}z) \mid g \in C_0(Z, \mathbb{R}), \text{Lip}_{\widehat{D}}(g) \leq 1 \}$$

where $\text{Lip}_{\widehat{D}}(g) = \sup_{z_1, z_2 \in Z} |g(z_1) - g(z_2)| / \widehat{D}(x, z_1, z_2)$. Hence, D^{Wass} can be understood as the restriction of a Banach-space norm to the convex set $\text{Prob}(Z)$. This definition of the norm is very useful in theoretical considerations but is less useful for the actual computation of the distance.

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

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

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

5.3 Gradient Young measures

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]):

$$\gamma \in \mathrm{GYM}(\Omega, \mathbb{R}^{d \times d}) \iff \begin{cases} \text{There exists } \boldsymbol{\varphi} : \Omega \mapsto \mathbb{R}^d \text{ such that for almost} \\ \text{every (a.e.) } x \in \Omega \text{ we have } \gamma(x) \in \mathrm{Prob}^{\mathrm{grad}}_{\mathrm{D}\boldsymbol{\varphi}(x)}(\mathbb{R}^{d \times d}), \end{cases}$$

where

$$\mathrm{Prob}^{\mathrm{grad}}(\mathbb{R}^{d \times d}) \stackrel{\mathrm{def}}{=} \Big\{ \nu \in \mathrm{Prob}(\mathbb{R}^{d \times d}) \,\Big|\, \text{for all quasiconvex } h : \mathbb{R}^{d \times d} \mapsto \mathbb{R}$$

$$\text{we have } h(\textstyle\int_{\mathbb{R}^{d \times d}} \mathbf{G}\, \nu(\mathrm{d}\mathbf{G})) \le \int_{\mathbb{R}^{d \times d}} h(\mathbf{G})\, \nu(\mathrm{d}\mathbf{G}) \Big\}$$

(5.4)

and $\mathrm{Prob}^{\mathrm{grad}}_{\mathbf{F}}(\mathbb{R}^{d \times d}) \stackrel{\mathrm{def}}{=} \{\, \nu \in \mathrm{Prob}^{\mathrm{grad}}(\mathbb{R}^{d \times d}) \mid \int_{\mathbb{R}^{d \times d}} \mathbf{G}\, \nu(\mathrm{d}\mathbf{G}) = \mathbf{F} \,\}.$

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. $x \in \Omega$.[1]

The above condition $\gamma(x) \in \mathrm{Prob}^{\mathrm{grad}}_{\mathrm{D}\boldsymbol{\varphi}(x)}(\mathbb{R}^{d \times d})$ 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 $\mathrm{rank}(A - B) \le 1$. 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 welcome feature of GYM is that the the condition $\gamma(x) \in \mathrm{Prob}^{\mathrm{grad}}_{\mathrm{D}\boldsymbol{\varphi}(x)}(\mathbb{R}^{d \times d})$ is local in the macroscopic space variable x. The only coupling between the different measures $\gamma(x)$ occurs through the mean value $\mathrm{D}\boldsymbol{\varphi}(x)$ which has to be a gradient. Moreover, all sets $\mathrm{Prob}^{\mathrm{grad}}_{\mathbf{F}}(\mathbb{R}^{d \times d})$ can be obtained by translating $\mathrm{Prob}^{\mathrm{grad}}_{\mathbf{0}}(\mathbb{R}^{d \times d})$, i.e., $\mathrm{Prob}^{\mathrm{grad}}_{\mathbf{F}}(\mathbb{R}^{d \times d}) = T_{\mathbf{F}}\mathrm{Prob}^{\mathrm{grad}}_{\mathbf{0}}(\mathbb{R}^{d \times d})$ where $(T_{\mathbf{F}}\nu)(A) = \nu(A - \mathbf{F})$ for all $A \subset \mathbb{R}^{d \times d}$. This can be used in finite-element calculations, since most computations can be done parallel in each element.

A major problem is that the set of quasiconvex functions appearing in the definition of $\mathrm{Prob}^{\mathrm{grad}}$ is not 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.

5.4 Sequential laminates

There is one nontrivial construction for GYM which are called *sequential laminates*. It is known that the closure of the set of sequential laminates (infinite sequential laminates) is a

---

[1]For mathematical reasons the above definitions have to be made more precise, such that the integrals $\int_{\mathbb{R}^{d \times d}} h(\mathbf{G})\, \nu(\mathrm{d}\mathbf{G})$ are well-defined. For instance, one can restrict to measures ν such that $\int_{\mathbb{R}^{d \times d}} 1 + |\mathbf{G}|^p \nu(\mathrm{d}\mathbf{G}) < \infty$ and then considers only quasiconvex functions h with $|h(\mathbf{G})| \le C(1 + |\mathbf{G}|^p)$.

strict subset of all GYM. However, so far it is the only set which is computationally accessible. Moreover, experimental observations show that almost all observed microstructures have the form of sequential laminates with one up to three levels of laminates.

The basic lamination construction works as follows. For $j = 1, 2$ consider $\mu_j \in$ $\mathrm{GYM}_{\varphi_j}(\Omega, \mathbb{R}^{d \times d})$ such that $\mathrm{rank}(\mathrm{D}\varphi_1(x) - \mathrm{D}\varphi_2(x)) \leq 1$. Then, for all functions $\theta \in$ $\mathrm{L}^\infty(\Omega, [0, 1])$ the YM $\mu = \theta\mu_1 + (1-\theta)\mu_2$ lies in $\mathrm{GYM}_\varphi(\Omega, \mathbb{R}^{d \times d})$, where $\varphi = \theta\varphi_1 + (1-\theta)\varphi_2$. To show this one has to combine the two different microstructures in a laminated fashion to form a new microstructure. Using $\mathbf{F}_1(x) - \mathbf{F}_2(x) = \mathbf{a}(x) \otimes \mathbf{n}(x)$, one forms on mesoscopic patches fine laminates with normal vector $\mathbf{n}(x)$ and relative volumes $\theta(x)$ and $1 - \theta(x)$. The condition $\mathbf{F}_1(x) - \mathbf{F}_2(x) = \mathbf{a}(x) \otimes \mathbf{n}(x)$ is needed to guarantee compatibility.

Starting with a trivial GYM of the form $\mu_j = \delta_{\mathrm{D}\varphi_j}$ and applying this process several times leads to more and more complicated GYMs. This process is called *sequential lamination*. It is known that this procedure does not yield all possible GYM, however, the reachable set seems to be sufficiently rich for applications. We refer to [BJ87, Kru98, OR99, ORS00, AP00, Gol01, GMH02] for more details.

Finally, we indicate how this procedure can be used for measures which have mass only at finitely many points. Obviously we have $\delta_\mathbf{F} \in \mathrm{Prob}_\mathbf{F}^{\mathrm{grad}}(\mathbb{R}^{d \times d})$ and for each $\alpha \in (0, 1)$, $\mathbf{a} \in \mathbb{R}^d$ and $\mathbf{n} \in \mathbb{R}^d$ with $|\mathbf{n}| = 1$ the two matrices

$$\mathbf{F}_1 = \mathbf{F} + (1-\alpha)\mathbf{a} \otimes \mathbf{n} \quad \text{and} \quad \mathbf{F}_2 = \mathbf{F} - \alpha\mathbf{a} \otimes \mathbf{n}$$

are compatible since $\mathbf{F}_1 - \mathbf{F}_2 = \mathbf{a} \otimes \mathbf{n}$. This splitting procedure of $\delta_\mathbf{F}$ into $\alpha\delta_{\mathbf{F}_1} + (1-\alpha)\delta_{\mathbf{F}_2} \in$ $\mathrm{Prob}_\mathbf{F}^{\mathrm{grad}}(\mathbb{R}^{d \times d})$ is called *branching* in [ORS00]. Replacing $\mathbf{F}_1$ and $\mathbf{F}_2$ by similar laminates we reach a GYM which has mass in four points, also called leaves. After finitely many steps we reach a general measure $\nu = \sum_1^N \alpha_k \delta_{\mathbf{F}_k}$ which lies in $\mathrm{Prob}_\mathbf{F}^{\mathrm{grad}}(\mathbb{R}^{d \times d})$. Moreover, such sequential lamination can be associated with a tree, where every branch point corresponds to a replacement of one gradient by a compatible pair. Since $\mathrm{Prob}_\mathbf{F}^{\mathrm{grad}}(\mathbb{R}^{d \times d})$ is a convex
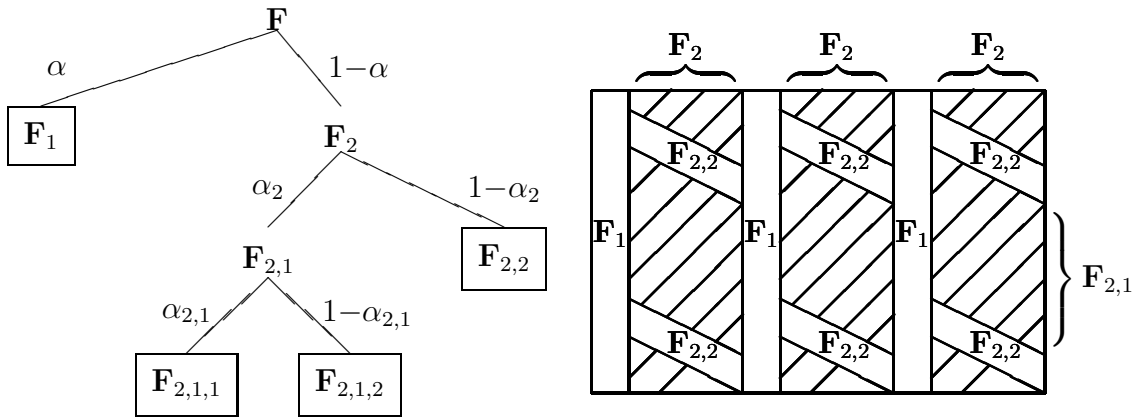

Figure 5.1: A lamination tree with four leaves and the associated sequential laminate

set, we may also take convex combinations of several lamination trees associated with $\mathbf{F}$.

For a given $\mathbf{F}$ a tree with k branchings lies in a $2kd$-dimensional manifold, since at each branch-point the weight $\alpha \in (0, 1)$ and the vectors $\mathbf{a}, \mathbf{n} \in \mathbb{R}^d$ with $|\mathbf{n}| = 1$ can be chosen freely. The construction of such measures is embedded in minimization problems arising from the incremental problems; hence it will be of importance to understand how the trees can be modified by changing the weights α, the jump vector $\mathbf{a}$ and the normal

19

vector $\mathbf{n}$ as well as by adding further branches or, to keep complexity as small as possible, by cutting of branches (called *pruning*) if their weight is too small, see Section 7.4.

6 Separate relaxation

6.1 Abstract separate relaxation

Whereas relaxation of single minimization problems is a standard tools nowadays, the relaxation of time-dependent or incremental problems is not developed as much. There are rather ad hoc methods in certain fields, but no general strategy seems to be available. We refer to [Ott98, Bre99, Bre00, MTL02] for such results.

Here we suggest a formal mathematical approach, which has the advantage that it is relatively simple and stays in the framework of variational incremental problems associated with energetic formulations. In particular, when this approach is made concrete within the field of Young measures, it leads to the models treated by engineers [OR99, ORS00, Mie03a, ML03].

The separate relaxation is obtain by embedding $\mathcal{F} \times \mathcal{Z}$ via $\mathcal{J}$ densely into a larger space $\mathbb{W}$. The relaxed functionals $\mathbb{E}$ and $\mathbb{D}$ are abstractly defined as relaxation of $\mathcal{E}$ and $\mathcal{D}$, respectively, in the sense of Section 4.1:

$$\mathbb{E}(t, w) = \inf\{ \liminf_{m \to \infty} \mathcal{E}(t, \boldsymbol{\varphi}_m, z_m) \mid \mathcal{J}(\boldsymbol{\varphi}_m, z_m) \xrightarrow{\mathbb{W}} w \},$$

$$\mathbb{D}(w_0, w_1) = \inf\{ \liminf_{m \to \infty} \mathcal{D}(z_{0,m}, z_{1,m}) \mid \mathcal{J}(\boldsymbol{\varphi}, z_{0,m}) \xrightarrow{\mathbb{W}} w_0, \ \mathcal{J}(\boldsymbol{\varphi}, z_{1,m}) \xrightarrow{\mathbb{W}} w_1 \}.$$

We call this separate relaxation, since in the relaxation process the energy storage and the energy dissipation are treated completely separately from each other. 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 6.1 *If for all $z_0, z_1 \in \mathcal{Z}$ we have $\mathbb{D}(\mathcal{J}(0, z_0), \mathcal{J}(0, z_1)) = \mathcal{D}(z_0, z_1)$ and if further technical assumptions hold, then the relaxation defined via $\mathbb{W}$, $\mathcal{J}$, $\mathbb{E}$ and $\mathbb{D}$ satisfies the properties (R1)–(R3) in Definition 4.1 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 consistencies (R3) and (R5) follow from the fact that the relaxation satisfies

$$\inf\{ \mathcal{E}(t, \boldsymbol{\varphi}, z) + \mathcal{D}(z_{\mathrm{old}}, z) \mid (\boldsymbol{\varphi}, z) \in \mathcal{F} \times \mathcal{Z} \} = \inf\{ \mathbb{E}(t, w) + \mathbb{D}(\mathcal{J}(0, z_{\mathrm{old}}), w) \mid w \in \mathbb{W} \}$$

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 $\mathbb{W}$. ∎

A major problem in the theory is that we do not understand under what conditions the separate relaxation also satisfies the properties $(R4)_{\mathrm{low}}$ or $(R4)_{\mathrm{upp}}$. The problems arise

from the fact that, depending on the choice of $\mathbb{W}$, we lose some or more information on the microstructure in step $k-1$ and just remember $w_{k-1} \in \mathbb{W}$. For instance, if $\mathbb{W}$ is the Young measure relaxation discussed in the next subsection, we only remember mesoscopic volume fractions of the infimizing sequence $(z_{k-1}^\varepsilon)_{\varepsilon>0}$ and lose information 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 (elasto-plasticity or shape-memory alloys) there is strong evidence that changes of the micro pattern 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 [The02] it is proved mathematically for a two-phase shape-memory model that the separate relaxation is an exact relaxation satisfying $(R4)_{\mathrm{low}}$.

6.2   Separate Young-measure relaxation

The abstract separate relaxation can be make explicit by choosing $\mathbb{W}$ suitably. We propose a relaxation which uses the weak convergence in $\mathcal{F} \subset W^{1,p}(\Omega, \mathbb{R}^d)$ 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

$$\begin{cases} \boldsymbol{\varphi}^j \rightharpoonup \boldsymbol{\varphi}^\infty \text{ in } W^{1,p}(\Omega, \mathbb{R}^d) \quad \text{and} \\ \forall\, g \in C_0(Z): \int_Z g(z)\, \mu^j(\mathrm{d}z) \rightharpoonup \int_Z g(z)\, \mu^\infty(\mathrm{d}z) \text{ in } L^1(\Omega). \end{cases}$$

The convergence $\mu^j \xrightarrow{\mathrm{YM}} \mu^\infty$ as defined above is also called Young measure convergence and it makes the set $\{\delta_z \,|\, z \in \mathcal{Z}\}$ dense in $\mathrm{YM}(\Omega, Z)$, cf. Section 4.1 and [Dac82, Bal89, Rou97].

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

$$\mathbb{E}(t, \boldsymbol{\varphi}, \mu) = \inf\{\, \liminf_{m\to\infty} \mathcal{E}(t, \boldsymbol{\varphi}_m, z_m) \,|\, \mathcal{J}(\boldsymbol{\varphi}_m, z_m) \xrightarrow{\mathbb{W}} (\boldsymbol{\varphi}, \mu) \,\},$$

$$\mathbb{D}(\mu_0, \mu_1) = \inf\{\, \liminf_{m\to\infty} \mathcal{D}(z_{0,m}, z_{1,m}) \,|\, \delta_{z_{0,m}} \xrightarrow{\mathrm{YM}} \mu_0, \ \delta_{z_{1,m}} \xrightarrow{\mathrm{YM}} \mu_1 \,\}.$$

This relaxation is useful since it allows for integral representations of the relaxed functionals. These follow from abstract results as given in [Dac82, Rou97, MTL02].

Theorem 6.2 *Under suitable technical assumptions the relaxed functionals have the form*

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

$$\mathbb{D}(\mu_{\mathrm{old}}, \mu_{\mathrm{new}}) = \int_\Omega D^{\mathrm{Wass}}(x, \mu_{\mathrm{old}}(x), \mu_{\mathrm{new}}(x))\, \mathrm{d}x,$$

with D^{Wass} from (5.3) and the relaxed potential Ψ_{rel} is given by

$$\Psi_{\mathrm{rel}}(x, \mathbf{F}, \nu) = \inf \left\{ \int_{(0,1)^d} \widehat{\psi}(x, \mathbf{F}{+}\mathrm{D}\boldsymbol{\varphi}(y), z(y))\, \mathrm{d}y \,\middle|\, \boldsymbol{\varphi} \in W_0^{1,\infty}((0,1)^d), \right.$$
$$\left. \nu(A) = \mathrm{vol}(\{\, y \,|\, z(y) \in A \,\}) \text{ for all } A \subset Z \right\}. \tag{6.1}$$

Moreover, for each $\mu_0 \in \text{YM}(\Omega, Z)$ the relaxed incremental problem

$$(\mathbb{RIP}) \qquad (\boldsymbol{\varphi}_k, \nu_k) \in \underset{(\boldsymbol{\varphi}, \mu) \in \mathcal{F} \times \text{YM}(\Omega, Z)}{\arg\min} \mathbb{E}(t_k, \boldsymbol{\varphi}, \mu) + \mathbb{D}(\mu_{k-1}, \mu)$$

is solvable, i.e., the solutions $(\boldsymbol{\varphi}_k, \nu_k) \in \mathcal{F} \times \text{YM}(\Omega, Z)$ exist for $k = 1, \dots, N$.

The computational results presented in [OR99, ORS00, MAL02, LMD03, Mie03a, ML03] are based on incremental problems of a very similar type as ($\mathbb{RIP}$), the major difference being that the dissipation distance $\mathbb{D}$ is approximated by a semi-implicit exponential ansatz. In [CP01] a careful numerical treatment of ($\mathbb{RIP}$) is given in the case small-strain model for a two-phase shape-memory alloys. For a mathematical analysis of some special cases we refer to [MTL02, MR03].

As given in the Section 4.3 there is now a time-continuous, energetic rate-independent evolutionary problem $(\mathbb{S}) \& (\mathbb{E})$ associated to the function space $\mathbb{W} = \mathcal{F} \times \text{YM}(\Omega, Z)$ 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 $\text{YM}(\Omega, Z)$ and generalized the potentials accordingly. (This corresponds to a locally compact convexification in the sense of [Rou97, Rou00].) However, the solvability of $(\mathbb{S}) \& (\mathbb{E})$ is much more complicated than that of ($\mathbb{RIP}$). Having solvability of the latter for a sequence of time steps tending to 0, one tries to show convergence (of a subsequence) of the time-discrete, incremental solutions towards a solution of $(\mathbb{S}) \& (\mathbb{E})$. This convergence question will not be addressed here, and we refer to [MTL02, MR03, MM03] for first results in this direction.

It should be noted that it would have been possible to define a larger relaxation by including the fluctuations of the deformation gradient as well. We will see later that our smaller relaxation is in fact equivalent, since we need to calculate the combined measure of the gradient $D\boldsymbol{\varphi}$ and z anyway in order to find the relaxed potential Ψ_{rel}. However, since the only coupling between the time levels in the incremental problem occurs via the internal variable z, it suffices to keep track only of the Young measure ν as given above.

It remains the question how the relaxed incremental problem ($\mathbb{RIP}$)

$$(\boldsymbol{\varphi}_k, \mu_k) \in \underset{(\boldsymbol{\varphi}, \mu) \in \mathbb{W}}{\arg\min} \int_\Omega \Big(\Psi_{\text{rel}}(D\boldsymbol{\varphi}(x), \mu(x)) + D^{\text{Wass}}(\mu_{k-1}(x), \mu(x)) \Big) \, \mathrm{d}x - \langle \ell(t_k), \boldsymbol{\varphi} \rangle \quad (6.2)$$

can be solved algorithmically. Here again the locality of the integrand will play a crucial rôle. For ($\mathbb{RIP}$) the important theorem 2.2 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 $\nu(\mathrm{d}z)$. We always have $\Psi_{\text{rel}}(\mathbf{F}, \nu) \leq \int_Z \widehat{\psi}(\mathbf{F}, z) \nu(\mathrm{d}z)$, and for nontrivial measure ν we have strict inequality in most cases. This becomes clear if we realize that $\mathbf{F}$ 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 $D\boldsymbol{\varphi}$ is necessary. We illustrate this in the following example.

Example 6.3 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 (6.1) we may choose any $z : (0, 1) \mapsto \mathbb{R}$ generating ν and then define φ via $\varphi(0) = 0$ and $\varphi'(y) = z(y) - \int_0^1 z(s) \, \mathrm{d}s$. Using $\int_0^1 z(s) \, \mathrm{d}s = \int_{\mathbb{R}} z \, \nu(\mathrm{d}z)$ the result follows. Moreover, we have

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

which shows that Ψ_{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 as

$$\Psi_{\mathrm{rel}}(x, \mathbf{F}, \nu) = \inf \left\{ \int_{\mathbb{R}^{d \times d} \times Z} \widehat{\psi}(x, \mathbf{G}, z) \rho(\mathrm{d}\mathbf{G}, \mathrm{d}z) \, \Big| \, \rho \in \mathcal{P}_{\mathbf{F}, \nu} \right\}, \text{ where}$$
$$\mathcal{P}_{\mathbf{F}, \nu} \stackrel{\mathrm{def}}{=} \left\{ \rho \in \mathrm{Prob}(\mathbb{R}^{d \times d} \times Z) \, \Big| \, \mathcal{M}_1 \rho \in \mathrm{Prob}_{\mathbf{F}}^{\mathrm{grad}}(\mathbb{R}^{d \times d}), \, \mathcal{M}_2 \rho = \nu \right\}. \tag{6.3}$$

See [Rou97] for the equivalence of (6.1) and (6.3). This construction is still very involved, and it is necessary to find suitable algorithms to approximate Ψ_{rel}. Since the sets $\mathcal{P}_{\mathbf{F}, \nu}$ are convex this is a classical linear convex programming problem, yet in infinite dimensions. We will discuss this further in Section 7.4.

Remark 6.4 *The minimizing measure ρ in (6.3) 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 $(\varphi_k, \mu_k) \in \mathcal{F} \times \mathrm{YM}(\Omega, Z)$ of $(\mathbb{RIP})$ we may consider the associated $\gamma_k \in \mathrm{GYM}_{\varphi_k}(\Omega, \mathbb{R}^{d \times d})$ which is obtained via $\gamma_k(x) = \mathcal{M}_1 \rho_k(x)$ from the minimizer $\rho_k(x)$ in (6.3) for calculating $\Psi_{\mathrm{rel}}(x, \mathrm{D}\varphi_k(x), \mu_k(x))$. In the case that $\gamma_k(x)$ is a sequential laminate the micropattern can be reconstructed as explained in Section 5.4.*

Often in computations one needs to replace the exact relaxed potential Ψ_{rel} by an approximate function; then it is important to keep its most important features, particularly its convexity properties.

Proposition 6.5 *The relaxed potential Ψ_{rel} is **cross-quasiconvex**, i.e. for all $(\mathbf{F}, \nu) \in \mathbb{R}^{d \times d} \times \mathrm{Prob}(Z)$ we have*

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

for all $\varphi \in \mathrm{W}_0^{1,\infty}((0,1)^d)$ 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 $\Psi_{\mathrm{rel}}(\cdot, \nu) : \mathbb{R}^{d \times d} \to [0, \infty]$ is quasiconvex for each ν and that $\Psi_{\mathrm{rel}}(\mathbf{F}, \cdot) : \mathrm{Prob}(Z) \to [0, \infty]$ is convex for each $\mathbf{F}$.

Cross-quasiconvexity is a special case of the more general A-quasiconvexity, see [FM99]. 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 Ψ_{rel} and the calculation of the Wasserstein distance $D^{\mathrm{Wass}}(\nu_{\mathrm{old}}, \nu_{\mathrm{new}})$ have to be done simultaneously. And therefore it is especially useful to employ the condensed energy density Ψ^{cond}, cf. Remark 7.2(1).

6.3   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 3.1. The finite set $Z = \{1, \ldots, n\}$ for the values 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 (3.1) 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 (probability) 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 stored-energy densities $\widehat{\psi}(x, \cdot, j) : \mathbb{R}^{d \times d} \to [0, \infty]$ 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

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

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 $h : \mathbb{R}_\circ^n \to [0, \infty)$ which is convex and homogeneous of degree one (see Prop. 4.7 in [MTL02]). Here $\mathbb{R}_\circ^n = \{\, \theta \in \mathbb{R}^n \mid \sum_1^n \theta_j = 0 \,\}$.

Thus, the relaxed function space $\mathbb{W} = \mathcal{F} \times \mathrm{YM}(\Omega, Z)$ takes here the form of the classical function space $\mathbb{W} = \mathrm{H}_\Gamma^1(\Omega) \times \mathrm{L}^1(\Omega, P_n)$. Moreover, the relaxed time-continuous problem (S) & (E) takes the form of a classical standard generalized material. However, we emphasize that the relaxed energy Ψ_{rel} (also called *mixture function* in [MTL02] and *free-energy of mixing* in [GMH02]) as well as the dissipation potential h are derived on purely mathematical grounds from abstract coarse-graining principles. For this process no further physical or mechanical knowledge about phase mixtures are needed.

Denoting by $\mathbf{u} = \boldsymbol{\varphi} - \mathrm{id}$ the displacement, the relaxed energy $\mathbb{E}$ and the relaxed dissipation distance $\mathbb{D}$ take the form

$$\mathbb{E}(t, \mathbf{u}, \theta) = \int_\Omega \Psi_{\mathrm{rel}}(x, \mathbf{I} + \mathrm{D}\mathbf{u}, \theta) - \mathbf{f}_{\mathrm{ext}}(t, x) \cdot \mathbf{u}(x)\, \mathrm{d}x \text{ and } \mathbb{D}(\theta_0, \theta_1) = \int_\Omega h(\theta_1(x) - \theta_0(x))\, \mathrm{d}x.$$

The flow formulation (cf. (2.4)) of the associated relaxed problem was first derived in [MTL02] and is implemented in [GM01]. It takes the form

$$\left.\begin{aligned} -\,\mathrm{div}\, \partial_{\mathbf{F}} \Psi_{\mathrm{rel}}(x, \mathbf{I} + \mathrm{D}\mathbf{u}, \theta) &= \mathbf{f}_{\mathrm{ext}}(t, x), \\ 0 \in \partial^{\mathrm{sub}} h(\dot{\theta}) + \partial_\theta \Psi_{\mathrm{rel}}(x, \mathbf{I} + \mathrm{D}\mathbf{u}, \theta) &+ \mathrm{N}_\theta P_n. \end{aligned}\right\} \text{ in } \Omega.$$

In the case $n = 2$ and in certain cases also for $n \geq 3$ we have $w_{\mathrm{mix}}(\theta) = -\mathbf{B} : \mathbf{M}(\theta)$ with $\mathbf{B} = \mathbf{B}^\top \geq 0$ and $\mathbf{M}(\theta) = \mathrm{diag}\,\theta - \theta \otimes \theta$, see [Mie00, GMH02].

24

7 Algorithms for updating the Young measures

In this section we want to address the question how the incremental problem associated to (RIP) can be solved. For this purpose we cannot use the abstract form using $\mathbb{E}$ and $\mathbb{D}$ nor the more concrete version (6.2) which still uses $\Psi_{\mathrm{rel}}(x, \mathbf{F}, \nu_{\mathrm{new}})$ and $D^{\mathrm{Wass}}(\nu_{\mathrm{old}}, \nu_{\mathrm{new}})$. The definition of the relaxed elastic potential Ψ_{rel} involves the joint fluctuation measure $\rho \in \mathrm{Prob}(\mathbb{R}^{d \times d} \times Z_{\mathrm{new}})$ and the definition of the Wasserstein distance D^{Wass} involves a transport measure $\eta \in \mathrm{Prob}(Z_{\mathrm{old}} \times Z_{\mathrm{new}})$. Here $Z_{\mathrm{old}} = Z_{\mathrm{new}} = Z$ and the subscripts just indicate what time level is associated with the corresponding internal variable z.

The calculation of all these measures must be part of the algorithm for solving (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, Mie03b] for exact results and [ML03] for semi-implicit approximations). Later on, we will assume that also the two derived constitutive functions are available, namely the condensed potential $\Psi^{\mathrm{cond}}(z_{\mathrm{old}}; x, \mathbf{F})$ and the update function $z_{\mathrm{new}} = \widehat{Z}^{\mathrm{update}}(x, \mathbf{F}, z_{\mathrm{old}})$.

This means that we assume that everything which is needed for a numerical calculation of classical solutions (i.e., without microstructure) is easy to calculate and already available. This does not mean that these algorithm are simpler or easier. Here we only study what we have to do additionally to be able to calculate the evolution of microstructures described by YMs.

7.1 Calculation of the full YM of $\mathbf{F}$, z_{new} and z_{old}

We start with the formulation of (RIP) which is formulated in terms of the joint measure $\tau \in \mathrm{Prob}(\mathbb{R}^{d \times d} \times Z_{\mathrm{old}} \times Z_{\mathrm{new}})$ which contains the joint distributions of $\mathbf{F}$, z_{old} and z_{new}. In particular, τ includes the transport measure $\eta = \mathcal{M}_{2,3}\tau = \int_{\mathbf{G} \in \mathbb{R}^{d \times d}} \tau(\mathrm{d}\mathbf{G}, \cdot, \cdot)$, which is needed to calculate the dissipation distance between the old and the new distribution of the internal variable, and the joint distribution of present gradients and internal variables $\rho = \mathcal{M}_{1,3}\tau = \int_{z_{\mathrm{old}} \in Z_{\mathrm{old}}} \tau(\cdot, \mathrm{d}z_{\mathrm{old}}, \cdot)$ which is needed to calculate $\Psi_{\mathrm{rel}}(x, \mathbf{F}, \nu)$ with $\nu = \mathcal{M}_3\tau$. It is essential for deriving a reasonable evolution for the microstructure to maintain the full generality that τ contains all the correlations between these distributions of $\mathbf{F} = \mathrm{D}\varphi$, z_{new} and z_{old}.

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

(i) Find a minimizer $(\varphi_k, \tau_k) \in \mathcal{F} \times \mathrm{YM}(\Omega, \mathbb{R}^{d \times d} \times Z_{\mathrm{old}} \times Z_{\mathrm{new}})$ of the functional

$$
\mathbb{I}(t_k, \varphi, \tau) \overset{\mathrm{def}}{=} -\langle \ell(t_k), \varphi \rangle +
$$
$$
+ \int_{\Omega} \int_{\mathbb{R}^{d \times d} \times Z_{\mathrm{old}} \times Z_{\mathrm{new}}} [\widehat{\psi}(x, \mathbf{G}, z_{\mathrm{new}}) + \widehat{D}(x, z_{\mathrm{old}}, z_{\mathrm{new}})] \, \tau(x; \mathrm{d}\mathbf{G}, \mathrm{d}z_{\mathrm{old}}, \mathrm{d}z_{\mathrm{new}}) \mathrm{d}x
$$

amongst all $(\boldsymbol{\varphi}, \tau) \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^d) \times \mathrm{YM}(\Omega, \mathbb{R}^{d \times d} \times Z_{\text{old}} \times Z_{\text{new}})$ satisfying

$$
\begin{aligned}
\mathcal{M}_1 \tau &= \int \int_{Z_{\text{old}} \times Z_{\text{new}}} \tau(\cdot; \cdot, \mathrm{d}z_{\text{old}}, \mathrm{d}z_{\text{new}}) \in \mathrm{GYM}_{\boldsymbol{\varphi}}(\Omega, \mathbb{R}^{d \times d}) \quad (\text{cf. } (5.4)), \\
\mathcal{M}_3 \tau &= \int_{\mathbb{R}^{d \times d} \times Z_{\text{new}}} \tau(\cdot; \mathrm{d}\mathbf{G}, \cdot, \mathrm{d}z_{\text{new}},) = \mu_{k-1} \in \mathrm{YM}(\Omega, Z_{\text{old}}).
\end{aligned}
\tag{7.1}
$$

(ii) Let $\mu_k(x) = \int_{\mathbb{R}^{d \times d} \times Z_{\text{old}}} \tau_k(x; \mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}}, \cdot)$, i.e., $\mu_k = \mathcal{M}_3 \tau_k \in \mathrm{YM}(\Omega, Z)$

(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 (7.1) and the new distribution μ_k is transported to the next step via (ii) and (iii). Note that we have written $\mathrm{YM}(\Omega, Z)$ in (ii) while our notation would give $\mathcal{M}_2 \tau_k \in \mathrm{YM}(\Omega, Z_{\text{new}})$. However, in the next time step, the present Z_{new} will be identified with the future Z_{old}.

7.2 Calculating the reduced YM for F and z_{old}

The complexity of this minimization problem can be reduced by minimizing internally with respect to the variable z_{new}. For this, it is essential to have a good algorithm to calculate the condensed potential $\Psi^{\text{cond}} : Z_{\text{old}} \times \Omega \times \mathbb{R}^{d \times d} \to [0, \infty)$ and the update function $\widehat{Z}^{\text{update}} : \Omega \times \mathbb{R}^{d \times d} \times Z_{\text{old}} \to Z_{\text{new}}$. 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 recall the definition $\mathcal{P}_{\mathbf{F}, \nu} \subset \mathrm{Prob}(\mathbb{R}^{d \times d} \times Z)$ from (6.3).

Proposition 7.1 *Using Ψ^{cond} and $\widehat{Z}^{\text{update}}$ (cf. (2.6)) the steps (i) and (ii) in the algorithm for $(\mathbb{RIP})$ (see (7.1)) can be replaced by the following equivalent algorithm:*

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

$$
\widetilde{\mathbb{I}}(t_k, \boldsymbol{\varphi}, \sigma) \stackrel{\text{def}}{=} \int_{\Omega} \int_{\mathbb{R}^{d \times d} \times Z_{\text{old}}} \Psi^{\text{cond}}(z_{\text{old}}; x, \mathbf{G}) \, \sigma(x; \mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}}) \mathrm{d}x - \langle \ell(t_k), \boldsymbol{\varphi} \rangle
$$

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

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

Here the push forward $\widetilde{\nu} = \Phi|^* \nu \in \mathrm{Prob}(V)$ of a measure $\nu \in \mathrm{Prob}(U)$ by a measurable mapping $\Phi : U \to V$ is defined via $\widetilde{\nu}(A) \stackrel{\text{def}}{=} \nu(\Phi^{-1}(A))$ 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 $\tau \in \mathrm{Prob}(\mathbb{R}^{d \times d} \times Z_{\text{old}} \times Z_{\text{new}})$ satisfying the constraints $\int_{\mathbb{R}^{d \times d} \times Z_{\text{new}}} \tau(\mathrm{d}\mathbf{G}, \cdot, \mathrm{d}z_{\text{new}}) = \mu_{k-1}(x)$ and $\int_{Z_{\text{old}} \times Z_{\text{new}}} \tau(\cdot, \mathrm{d}z_{\text{old}}, \mathrm{d}z_{\text{new}}) \in \mathrm{Prob}^{\text{grad}}_{\mathrm{D}\boldsymbol{\varphi}_k(x)}(\mathbb{R}^{d \times d})$. First assume that τ is a finite convex combination of Dirac masses:

26

$\tau = \sum_{i=1}^{n} \lambda_i \delta_{(\mathbf{G}_i, z_{\text{old}}^{(i)}, z_{\text{new}}^{(i)})}$. After adding suitable points with mass 0 we may write τ in the form

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

Minimizing $a \stackrel{\text{def}}{=} \int_{\mathbb{R}^{d \times d} \times Z_{\text{old}} \times Z_{\text{new}}} [\widehat{\psi}(x, \mathbf{G}, z_{\text{new}}) + \widehat{D}(x, z_{\text{old}}, z_{\text{new}})] \tau(\mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}}, \mathrm{d}z_{\text{new}})$ with respect to $z_{\text{new}}^{(j,l)}$, while keeping $\theta_{j,l}$, $\mathbf{G}_j$, and $z_{\text{old}}^{(l)}$ fixed, leads to

$$a = \sum_{j,l=1}^{m} \theta_{j,l} \Psi^{\text{cond}}(z_{\text{old}}^{(l)}; \mathbf{G}_j) \quad \text{and} \quad z_{\text{new}}^{(j,l)} = \widehat{Z}^{\text{update}}(\mathbf{G}_j, z_{\text{old}}^{(j)}).$$

With $\sigma \stackrel{\text{def}}{=} \mathcal{M}_{1,2}\tau = \sum_{j,l=1}^{m} \theta_{j,l} \delta_{(\mathbf{G}_j, z_{\text{old}}^{(l)})} \in \mathcal{P}_{\mathrm{D}\boldsymbol{\varphi}_k(x), \mu_{k-1}(x)}$ we find

$$a = \int_{\mathbb{R}^{d \times d} \times Z_{\text{old}}} \Psi^{\text{cond}}(z_{\text{old}}; \mathbf{G})\, \sigma(\mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}}) \text{ and } \mu_k(x) = \sum_{j,l=1}^{m} \theta_{j,l} \delta_{\widehat{Z}^{\text{update}}(\mathbf{G}_j, z_{\text{old}}^{(l)})} = \widehat{Z}^{\text{update}}|^* \sigma.$$

Thus, the desired forms of a and $\mu_k(x)$ are established if τ consists of finitely many 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 $\boldsymbol{\varphi}_k$ and σ_k due to the restriction $\int_{Z_{\text{old}}} \sigma_k(\cdot, \mathrm{d}z_{\text{old}}) \in \mathrm{GYM}_{\boldsymbol{\varphi}_k}(\Omega, \mathbb{R}^{d \times d})$. However, conceptually we may define the following condensed relaxed potential $\Psi_{\text{rel}}^{\text{cond}}$ and obtain a reduced, purely macroscopic minimization problem for $\boldsymbol{\varphi}_k$ which is a classical minimization problem for nonlinear elasticity.

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

$$\Psi_{\text{rel}}^{\text{cond}}(\nu_{\text{old}}; x, \mathbf{F}) \stackrel{\text{def}}{=} \min_{\sigma \in \mathcal{P}_{\mathbf{F}, \nu_{\text{old}}}} \int_{\mathbb{R}^{d \times d} \times Z} \Psi^{\text{cond}}(z_{\text{old}}; x, \mathbf{G})\, \sigma(\mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}}). \tag{7.2}$$

If $\Psi_{\text{rel}}^{\text{cond}}(\nu_{\text{old}}; x, \cdot)$ is known, then $\boldsymbol{\varphi}_k$ can be obtained by solving the classical (pseudo-) elastic minimization problem

$$\boldsymbol{\varphi}_k \in \arg\min_{\boldsymbol{\varphi} \in \mathcal{F}} \mathbb{E}^{\text{cond}}(t_k, \boldsymbol{\varphi}) \quad \text{with}$$
$$\mathbb{E}^{\text{cond}}(t, \boldsymbol{\varphi}) = \int_{\Omega} \Psi_{\text{rel}}^{\text{cond}}(\mu_{k-1}(x); x, \mathrm{D}\boldsymbol{\varphi}(x))\, \mathrm{d}x - \langle \ell(t), \boldsymbol{\varphi} \rangle. \tag{7.3}$$

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 $\sigma = \widehat{\Sigma}(x, \mathbf{F}, \nu_{\text{old}}) \in \mathcal{P}_{\mathbf{F}, \nu_{\text{old}}}$, i.e.,

$$\widehat{\Sigma}(x, \mathbf{F}, \nu_{\text{old}}) \in \arg\min_{\sigma \in \mathcal{P}_{\mathbf{F}, \nu_{\text{old}}}} \int_{\mathbb{R}^{d \times d} \times Z} \Psi^{\text{cond}}(z_{\text{old}}; x, \mathbf{G})\, \sigma(\mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}}). \tag{7.4}$$

Having a minimizer $\widehat{\Sigma}(x, \mathbf{F}, \mu_{k-1}(x))$ a the macroscopic deformation $\boldsymbol{\varphi}_k$ as solution of (7.3) we define $\sigma_k(x) = \widehat{\Sigma}(x, \mathrm{D}\boldsymbol{\varphi}_k(x), \mu_{k-1}(x))$ and obtain μ_k via step $\widetilde{\text{(ii)}}$ in Proposition 7.1, i.e., $\mu_k(x) = \widehat{Z}^{\text{update}}(x, \cdot)|^* \sigma_k(x)$. This finishes the construction of the solution for the kth step.

Remark 7.2
(1) Finding a minimizer $\widehat{\Sigma}$ is in fact a simultaneous relaxation and calculation of the

Wasserstein distance from ν_{old}. However, we have contracted the effort to the minimal information needed for realizing the relaxed incremental problem ($\mathbb{RIP}$).

(2) The minimization in (7.4) can't be decomposed into minimizations with respect to Dirac masses for z_{old}. More precisely, assume $\nu_{\text{old}} = \sum_{j=1}^{m} \theta_j \delta_{z_j}$ 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 $\mathcal{P}_{\mathbf{F}, \nu_{\text{old}}}$; however, in general it is not a minimizer, cf. Example 6.3. The point is that the measures σ_j may be associated with different macroscopic gradients $\mathbf{F}_j$ as long as they are compatible and have the correct average.

7.3 First and second derivative of the incremental potential

For solving the macroscopic problem (7.3) we need $\Psi_k(x, \mathbf{F}) \stackrel{\text{def}}{=} \Psi_{\text{rel}}^{\text{cond}}(\mu_{k-1}(x); x, \mathbf{F})$ 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 $\sigma_k^{\mathbf{F}}(x) \in \mathcal{P}_{\mathbf{0}, \mu_{k-1}(x)}$ via $\sigma_k^{\mathbf{F}}(x; S) \stackrel{\text{def}}{=} \widehat{\Sigma}_k^{\mathbf{F}}(x, \mathbf{F}, \mu_{k-1})(T_{\mathbf{F}} S)$ for all measurable subsets S of $\mathbb{R}^{d \times d} \times Z$. Note that the Young measures $\sigma_k^{\mathbf{F}}(x)$ have mean value $\mathbf{0}$ in the $\mathbf{G}$-component

Proposition 7.3 *With the above notations we have the following formulae:*

$$
\begin{aligned}
\Psi_k(x, \mathbf{F}) &= \int_{\mathbb{R}^{d \times d} \times Z} \Psi^{\text{cond}}(z_{\text{old}}, x, \mathbf{F}+\mathbf{G})\, \sigma_k^{\mathbf{F}}(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}}), \\
\mathrm{D}_{\mathbf{F}} \Psi_k(x, \mathbf{F}) &= \int_{\mathbb{R}^{d \times d} \times Z} \mathrm{D}_{\mathbf{F}} \Psi^{\text{cond}}(z_{\text{old}}, x, \mathbf{F}+\mathbf{G})\, \sigma_k^{\mathbf{F}}(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}}).
\end{aligned}
\tag{7.5}
$$

If $\sigma_k^{\mathbf{F}}(x)$ 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

$$
\begin{aligned}
\mathrm{D}_{\mathbf{F}}^2 \Psi_k(x, \mathbf{F})[\mathbf{H}, \mathbf{H}] &= \int_{\mathbb{R}^{d \times d} \times Z} \mathrm{D}_{\mathbf{F}}^2 \Psi^{\text{cond}}(z_{\text{old}}, x, \mathbf{F}+\mathbf{G})[\mathbf{H}, \mathbf{H}]\, \sigma_k^{\mathbf{F}}(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}}) \\
&\quad + \int_{\mathbb{R}^{d \times d} \times Z} \mathrm{D}_{\mathbf{F}} \Psi^{\text{cond}}(z_{\text{old}}, x, \mathbf{F}+\mathbf{G})[\mathbf{H}]\, \mathrm{D}_{\mathbf{F}} \sigma_k^{\mathbf{F}}(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}})[\mathbf{H}] \\
&= \int_{\mathbb{R}^{d \times d} \times Z} \mathrm{D}_{\mathbf{F}}^2 \Psi^{\text{cond}}(z_{\text{old}}, x, \mathbf{F}+\mathbf{G})[\mathbf{H}, \mathbf{H}]\, \sigma_k^{\mathbf{F}}(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}}) \\
&\quad - \int_{\mathbb{R}^{d \times d} \times Z} \Psi^{\text{cond}}(z_{\text{old}}, x, \mathbf{F}+\mathbf{G}) \mathrm{D}_{\mathbf{F}}^2 \sigma_k^{\mathbf{F}}(x, \mathrm{d}\mathbf{G}, \mathrm{d}z_{\text{old}})[\mathbf{H}, \mathbf{H}].
\end{aligned}
\tag{7.6}
$$

We refer to [BKK00] for a rigorous proof of the second formula in (7.5).

Here the distributional derivatives $\mathrm{D}_{\mathbf{F}} \widehat{\sigma}[\mathbf{H}]$ and $\mathrm{D}_{\mathbf{F}}^2 \widehat{\sigma}[\mathbf{H}, \mathbf{H}]$ of a measures $\widehat{\sigma}$ with respect to $\mathbf{F}$ in the direction $\mathbf{H}$ are again measures and are defined via the derivatives of the linear functional $f \mapsto L_f(\mathbf{F}) \stackrel{\text{def}}{=} \int_{\mathbb{R}^{d \times d} \times Z} f(\mathbf{G}, z) \widehat{\sigma}(\mathrm{d}\mathbf{G}, \mathrm{d}z)$ as follows:

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

for all sufficiently smooth test functions $f : \mathbb{R}^{d \times d} \times Z \to \mathbb{R}$.

These relations are best understood by assuming that the measure is given as a 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 $\widetilde{\mathbf{z}}_j$.

7.4   Updates using sequential laminations

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 (7.4). At present there is no easy method for solving this problem. However, the procedure called **sequential lamination** (see [Kru98, AP00, Gol01, GMH02] and Section 5.4) 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 $\sum_{j=1}^{J} \theta_j \delta_{z_j}$. A general $\sigma \in \mathcal{P}_{\mathbf{F}, \nu}$ which is a convex combination of finitely many Dirac masses has the form $\sigma = \sum_{k=1}^{K} s_k \delta_{(\mathbf{G}_k, \widetilde{z}_k)}$. However, $\nu = \mathcal{M}_2 \sigma$ implies $\{ \widetilde{z}_k \mid k = 1, ..., K \} \subset \{ z_j \mid j = 1, ..., J \}$. Hence, we can use the ansatz

$$\sigma = \sum_{j=1}^{J} \sum_{l=1}^{L_j} s_{l,j} \delta_{(\mathbf{G}_{l,j}, z_j)} \quad \text{with } s_{l,j} \geq 0 \text{ and } \sum_{l=1}^{L_j} s_{l,j} = \theta_j \text{ for } j = 1, \dots, J. \quad (7.7)$$

We have to guarantee that $\gamma = \sum_{j=1}^{J} \sum_{l=1}^{L_j} s_{l,j} \delta_{\mathbf{G}_{l,j}}$ lies in $\mathrm{Prob}_{\mathbf{F}}^{\mathrm{grad}}(\mathbb{R}^{d \times d})$, which is the most difficult restriction and is now replaced by the somewhat stronger restriction that γ is a convex combination of sequential laminates, as defined in Section 5.4.

The major difficulty here is that the partial measures $\gamma_j = \sum_{l=1}^{L_j} \frac{s_{l,j}}{\theta_l} \delta_{\mathbf{G}_{l,j}}$ need not lie in $\mathrm{Prob}_{\mathbf{F}_j}^{\mathrm{grad}}(\mathbb{R}^{d \times d})$ for any $\mathbf{F}_j$. Thus, the compatibility of the microstructural deformation gradients can be achieved by using different internal variables z_k, see Example 6.3.

Since the condition on γ being a gradient YM is the most severe one, a typical algorithmical approach is the following. One assumes that the minimum defining $\Psi_{\mathrm{rel}}^{\mathrm{cond}}(\nu; x, \mathbf{F})$ is attained at a measure $\sigma \in \mathcal{P}_{\mathbf{F}, \nu}$ such that $\gamma = \mathcal{M}_1 \sigma$ is a single lamination tree and to each leaf there is associated a single value of z_j, see Figure 7.1. In practical application it is of course sufficient, that the minimum over the restricted class is very close to the true minimum. (However, it is to be expected that in general situations the minimal measure is a more complex measure which consists of several independent lamination trees or of a tree which has more than two branches coming out of one gradient, like two different sub-laminates.)

To start the minimization in this class, it is common to construct a candidate in this class from the data of the previous iteration step. Since this candidate is not a minimizer, we have to adjust the parameters in the lamination tree to make it minimal. However, the freedom in the lamination tree is restricted by the condition $\mathcal{M}_2 \sigma = \nu = \sum_{1}^{J} \theta_j z_j$.
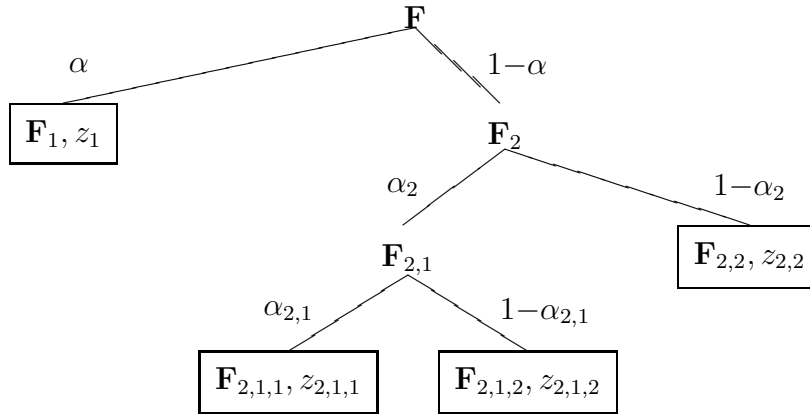
Figure 7.1: A measure in $\mathcal{P}_{\mathbf{F},\nu}$ with four leaves:

$$\sigma = \alpha\,\delta_{(\mathbf{F}_1,z_1)} + (1-\alpha)\Big[\alpha_2\big(\alpha_{2,1}\delta_{(\mathbf{F}_{2,1,1},z_{2,1,1})}(1-\alpha_{2,1})\delta_{(\mathbf{F}_{2,1,2},z_{2,1,2})}\big) + (1-\alpha_2)\delta_{(\mathbf{F}_{2,2},z_{2,2})}\Big].$$

Since, only one of the values z_j is associated to each leaf, we obtain J conditions on the weights in the branches of the tree. Thus, we need more branchings than the number J to be able to modify the weights. Recall that a lamination tree with m branchings contains $m{+}1$ leaves and m weights. Thus, the weights are restricted, whereas the m lamination normals $\mathbf{n}_\rho \in \mathbb{S}^{d-1}$ and jump vectors $\mathbf{a}_\rho \in \mathbb{R}^d$ are free to be modified.

During the minimization process one has to test for each leaf whether it is advantageous to replace this leaf by a simple laminate. This process is then called *branching*. After this process the two new leaves will have the same value of z. The opposite of branching is called *pruning*, i.e., we replace two leaves by one. This can be done if the weight of one of the leaves is very small or if the jump vector $\mathbf{a}$ is small and the two values of z are the same. In exceptional cases, even a longer branch may be pruned because its weight is very small.

Now, assume that the minimization is finished, i.e., we have found $\sigma_k \in \mathcal{P}_{\mathrm{D}\boldsymbol{\varphi}_k(x),\mu_{k-1}(x)}$ in the form (7.7) as a lamination tree with $m = \sum_1^J L_j$ leaves. Then, $\mu_k(x)$ has to be defined as in $\widetilde{(\mathrm{ii})}$ of Proposition 7.1, namely $\mu_k(x) = \sum_1^J \sum_1^{L_j} s_{l,j}\delta_{\widehat{\mathcal{Z}}^{\mathrm{update}}(\mathbf{G}_{l,j},z_j)}$. One problem is now that in general the number of values attained in μ_k is actually m, the number of leaves, and is, hence, much bigger than J, the number of values in $\mu_{k-1}(x)$. Thus, in a post-processing step one has to reduce the number again to keep the complexity of the problem in a reasonable range.

We refer to [ORS00, ML03] for implementations of quite similar algorithms.

8 Examples

Here we present two examples which can be treated analytically. For further applications we refer to [MT99, GM01, MTL02, CP01, GMH02, MR03], where models for N-phase shape-memory alloys are studied. Numerical algorithms using similar ideas as proposed here are implemented in two- and three-dimensional elastoplasticity, see [OR99, HH03, Mie02a, Mie03a, ML03].

8.1 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 3

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

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

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

We apply external forcing by imposing the boundary conditions $\boldsymbol{\varphi}(t, 0) = 0$ and $\boldsymbol{\varphi}(t, 1) = \Phi_{\mathrm{Dir}}(t) \overset{\text{def}}{=} t + 1$ 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 condensed potential can be derived explicitly. We have

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

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

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 [Mie03b].)

The most important feature of this model for our present purposes is that the condensed potential $\widetilde{\Psi}^{\mathrm{cond}}(p; \cdot)$ 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 loss of 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 $\widetilde{\Psi}^{\mathrm{cond}}(p; \cdot)$. It coincides with $\widetilde{\Psi}^{\mathrm{cond}}(p; \cdot)$ 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 $(\mathbf{F}_0, \widetilde{\Psi}^{\mathrm{cond}}(-1; \mathbf{F}_0))$ and $(\mathbf{F}_1, \widetilde{\Psi}^{\mathrm{cond}}(-1; \mathbf{F}_1))$, which is the tangent to $\widetilde{\Psi}^{\mathrm{cond}}$ at both points, see Figure 8.1. We define $(\mathbf{P}_1, p_1) = \widehat{Z}^{\mathrm{update}}(\mathbf{F}_1, (\mathbf{P}_0, p_0))$ which is the unique pair for which $\widetilde{\Psi}^{\mathrm{cond}}(p_0; \mathbf{F}_1) = W(\mathbf{F}_1\mathbf{P}_1) + h(p_1) + \widehat{D}((\mathbf{P}_0, p_0), (\mathbf{P}_1, p_1))$.
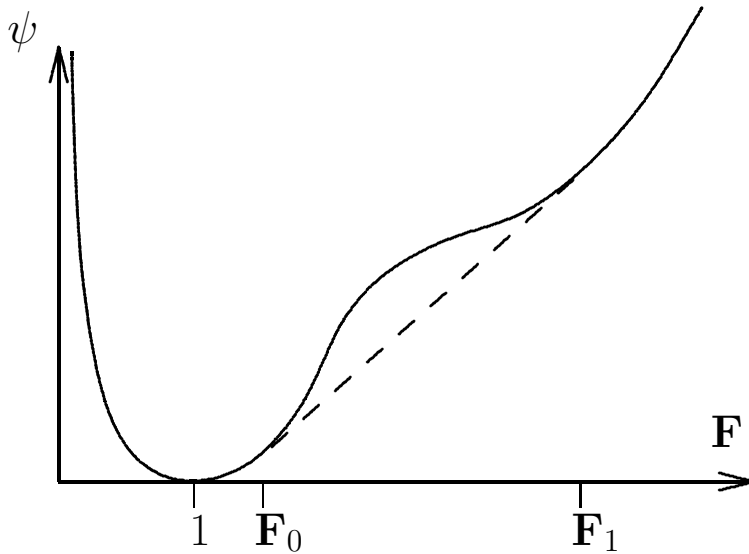
Figure 8.1: The function $\widetilde{\Psi}^{\mathrm{cond}}(-1; \cdot)$ (full line) and its convexification (broken line).

We look for a spatially homogeneous Young-measure solution $w(t) = (\boldsymbol{\varphi}(t), \nu(t))$ with a trivial macroscopic deformation $\boldsymbol{\varphi}(t, x) = \Phi_{\mathrm{Dir}}(t)x$ and $\nu(t) \in \mathrm{Prob}(\mathbb{R}^2)$. Using the times t_0 and t_1 given by $\Phi_{\mathrm{Dir}}(t_j) = \mathbf{F}_j$ 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 $(\mathbf{P}(t), p(t)) = \widehat{Z}^{\mathrm{update}}((\mathbf{1}, -1), \Phi_{\mathrm{Dir}}(t))$.

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

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

defines a solution as well.

8.2 The rigid plasticity model of CONTI & THEIL

In [CT03] a model for plasticity is presented which is fully nonlinear and can be analyzed analytically in quite some detail. It is motivated by [OR99] where microstructure formation in crystal plasticity was studied for the first time. We describe the model in our notation of Section 3.2 and give the results which concern our theory. In particular, the example treated in [CT03, Sect.4] provides a special case of our relaxation in Section 4 where (R1) to (R5) can be proved rigorously. For more details see the original paper.

We consider a body $\Omega \in \mathbb{R}^2$ (related results are also obtained for $d = 3$). The set Z of internal variables a one-dimensional subgroup of $\mathrm{GL}(\mathbb{R}^2)$ which corresponds to slip in a single slip system. For simplicity we choose the slip system as $\mathbf{S} = \mathbf{e}_1 \otimes \mathbf{e}_2$, i.e.,

$$\mathbf{F}_{\mathrm{p}} = \mathbf{P}_\beta = \mathbf{I} + \beta \mathbf{S} \in Z \stackrel{\text{def}}{=} \left\{ \left. \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \right| \alpha \in \mathbb{R} \right\} \subset \mathrm{SL}(\mathbb{R}^2).$$

There is no hardening assumed, the dissipation potential is $\Delta(\mathbf{P}_\beta, \dot{\mathbf{P}}_\beta) = |\dot{\beta}|$ which gives the simple dissipation distance $\widehat{D}(\mathbf{P}_\alpha, \mathbf{P}_\beta) = |\beta - \alpha|$. The stored-energy density is such that it has infinitely hard elastic response, i.e., $\mathbf{F}_{\mathrm{e}} = \mathrm{D}\boldsymbol{\varphi} \mathbf{P}^{-1}$ must lie in $\mathrm{SO}(\mathbb{R}^2)$.

$$\widehat{\psi}(\mathbf{F}, \mathbf{P}) = \begin{cases} 0 \ \text{ for } \mathbf{FP}^{-1} \in \mathrm{SO}(\mathbb{R}^d), \\ \infty \ \text{else.} \end{cases}$$

This defines the full model if we additionally specify suitable loadings and boundary conditions. For simplicity we stay with the example given in [CT03] which imposes time-dependent homogeneous boundary conditions $\boldsymbol{\varphi}(t, x) = \Phi_{\mathrm{Dir}}(t)x$ for all $x \in \partial\Omega$ for a given curve $\Phi_{\mathrm{Dir}} : [0, T] \to \mathrm{GL}(\mathbb{R}^2)$. Moreover, we impose the initial condition $\mathbf{P}(0, \cdot) = \mathbf{I}$ on all of Ω.

With these definitions the incremental problem (IP) is well defined. The formulation of the time-continuous problem (S) & (E) is somewhat more complicated here since on the one hand the space of admissible deformations $\mathcal{F}(t)$ is time dependent via $\Phi_{\mathrm{Dir}}(t)$ and on the other hand $\widehat{\psi}$ is not continuous. Note that the energy functional satisfies

$$\mathcal{E}(t, \boldsymbol{\varphi}, \mathbf{P}_\beta) = \begin{cases} 0 \ \text{ if } \boldsymbol{\varphi} \in \mathcal{F}(t) \text{ and } \mathbf{FP}_\beta^{-1} \in \mathrm{SO}(\mathbb{R}^2) \text{ a.e. in } \Omega, \\ \infty \ \text{else.} \end{cases}$$

The condensed energy potential and the update function take the form

$$\left(\Psi^{\mathrm{cond}}(\mathbf{P}_\beta; \mathbf{F}), \widehat{Z}^{\mathrm{update}}(\mathbf{F}, \mathbf{P}_\beta) \right) = \begin{cases} \left(|\beta - \alpha|, \mathbf{P}_\alpha \right) \text{ for } \mathbf{FP}_\alpha \in \mathrm{SO}(\mathbb{R}^2), \\ \left(\infty, \emptyset \right) \quad \text{else.} \end{cases}$$

The separately relaxed functionals $\mathbb{E}$ and $\mathbb{D}$ on $\mathbb{W}(t) = \mathcal{F}(t) \times \mathrm{YM}(\Omega, Z)$ can also be calculated explicitly as stated in Theorem 6.2. For this we introduce

$$M^{(2)} \stackrel{\text{def}}{=} \{ \mathbf{RP}_\alpha \mid \mathbf{R} \in \mathrm{SO}(\mathbb{R}^2), \alpha \in \mathbb{R} \} = \{ \mathbf{G} \mid \det \mathbf{G} = 1, |\mathbf{Ge}_1| = 1 \} \text{ and}$$
$$N^{(2)} \stackrel{\text{def}}{=} \{ \mathbf{G} \mid \det \mathbf{G} = 1, |\det \mathbf{G}| \leq 1 \} \supset M^{(2)}.$$

The relaxed energetic potential Ψ_{rel} now takes the form

$$\Psi_{\mathrm{rel}}(\mathbf{F}, \nu) = \begin{cases} 0 \ \text{ if } \mathbf{F} \in N^{(2)} \text{ and } \nu \in \mathcal{N}(\mathbf{F}), \\ \infty \ \text{else,} \end{cases}$$

where $\mathcal{N}(\mathbf{F}) \stackrel{\text{def}}{=} \{ \nu \in \mathrm{Prob}(Z) \mid \exists \gamma \in \mathrm{Prob}_{\mathbf{F}}^{\mathrm{grad}}(M^{(2)}) : \nu = \mathcal{R}|^* \gamma \}$, where $\mathcal{R} : M^{(2)} \to Z$ is defined via $\mathcal{R}(\mathbf{G}) = \mathbf{P}_\beta$ with $\beta = (\mathbf{Ge}_1) \cdot (\mathbf{Ge}_2)$. It is shown in [CT03] that $N^{(2)}$ is the

quasiconvexification of $M^{(2)}$, hence $\mathcal{N}(\mathbf{F})$ is nonempty for each $\mathbf{F} \in N^{(2)}$. The Wasserstein distance $\mathbb{D}$ is easy, since D is simply the distance on the straight line $Z \subset \mathbb{R}^{2 \times 2}$.

Without going into detail with all the other constructions presented above, we now study a particular shear experiment, which shows that microstructure occurs through simple laminates. For this consider the Dirichlet boundary conditions defined via $\Phi_{\mathrm{Dir}}(t) = \mathbf{I} + t\mathbf{K}$ with $\mathbf{K} = \binom{1}{1} \otimes \binom{-1}{1}$ for $t \in [0, 1]$ (which is exactly the interval for which $\Phi_{\mathrm{Dir}}(t) \in N^{(2)}$ holds). Defining $\mathbf{F}_1 = \mathbf{I} = \Phi_{\mathrm{Dir}}(0)$ and $\mathbf{F}_2 = \mathbf{I} + \mathbf{K} = \Phi_{\mathrm{Dir}}(1)$, we immediately have

$$\mathbf{F}_j \in M^{(2)}, \quad \mathbf{F}_2 - \mathbf{F}_1 = 2\mathbf{K} = \mathbf{a} \otimes \mathbf{n} \quad \text{and} \quad \Phi_{\mathrm{Dir}}(t) = (1-t)\mathbf{F}_1 + t\,\mathbf{F}_2$$

Thus, the simple laminate $\gamma(t) = (1-t)\delta_{\mathbf{F}_1} + t\,\delta_{\mathbf{F}_2}$ lies in $\mathrm{Prob}^{\mathrm{grad}}_{\Phi_{\mathrm{Dir}}(t)}(M^{(2)})$ and $\nu(t) = \mathcal{R}|^*\gamma$ lies in $\mathcal{N}(\Phi_{\mathrm{Dir}}(t))$ which implies $\mathbb{E}(\Phi_{\mathrm{Dir}}(t), \nu(t)) = 0$. Since $\mathcal{R}(\mathbf{F}_2) = \mathbf{P}_{-2}$ we find $D(\mathbf{F}_1, \mathbf{F}_2) = 2$ and the Wasserstein distance gives, for $0 \le t_1 < t_2 \le 1$, the identity

$$\mathbb{D}(\nu(t_1), \nu(t_2)) = \mathrm{Diss}_{\mathrm{rel}}(\nu, [t_1, t_2]) = 2|t_1 - t_2|.$$

It is shown in [CT03] that this solution is the unique Young measure limit of our approximate incremental problem $(\mathrm{AIP})_\varepsilon$ of Section 4.2, whereas the classical incremental problem (IP) has no solution. Moreover, it is also the solution of the relaxed time-continuous problem $(\mathbb{S}) \& (\mathbb{E})$, and thus provides, in addition to [The02], another nontrivial example where the separate relaxation leads to a useful relaxed problem satisfying (R4) of Section 4.2.

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.

[Bal89]   J. Ball. A version of the fundamental theorem for Young measures. In *PDEs and continuum models of phase transitions (Nice, 1988)*, pages 207–215. Springer, Berlin, 1989.

[BCH04]   S. Bartels, K. Carstensen, C. Hackl, and U. Hoppe. Effective relaxation for microstructure simulations: algorithms and applications. *Comput. Methods Appl. Mech. Engrg.*, 2004. To appear.

[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.

[Bre99] Y. BRENIER. Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations. *Comm. Pure Appl. Math.*, 52, 411–452, 1999.

[Bre00] Y. BRENIER. Derivation of the Euler equations from a caricature of Coulomb interaction. *Comm. Math. Phys.*, 212, 93–104, 2000.

[CHM02] C. CARSTENSEN, K. HACKL, and A. MIELKE. Non–convex potentials and microstructures in finite–strain plasticity. *Proc. Royal Soc. London, Ser. A*, 458, 299–317, 2002.

[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.

[CT03] S. CONTI and F. THEIL. Single-slip elastoplastic microstructures. *MPI MIS Preprint 64*, 2003.

[Dac82] B. DACOROGNA. *Weak continuity and weak lower semicontinuity of nonlinear functionals.* Springer-Verlag, Berlin, 1982.

[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.

[EM04] M. EFENDIEV and A. MIELKE. On the rate–independent limit of systems with dry friction and small viscosity. *J. Convex Analysis*, 2004. To appear.

[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.

[FM99] I. FONSECA and S. MÜLLER. $\mathcal{A}$-quasiconvexity, lower semicontinuity, and Young measures. *SIAM J. Math. Anal.*, 30, 1355–1390 (electronic), 1999.

[GM01] S. GOVINDJEE and C. MIEHE. A multi–variant martensitic phase transformation model: Formulation and numerical implementation. *Computer Meth. Applied Mech. Eng.*, 191, 215–238, 2001.

[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*, 50, 1897–1922, 2002. Erratum and Correct Reprinting: 51(4) 2003, pp. 763 & I-XXVI.

[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.

[HH03] K. HACKL and U. HOPPE. On the calculation of microstructures for inelastic materials using relaxed energies. In C. Miehe, editor, *IUTAM Symposium on Computational Mechanics of Solids at Large Strains*, pages 77–86. Kluwer, 2003.

[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.

[Joh76] C. JOHNSON. Existence theorems for plasticity problems. *J. Math. Pures Appl. (9)*, 55(4), 431–444, 1976.

[KMR03] M. KOČVARA, A. MIELKE, and T. ROUBÍČEK. A rate–independent approach to the delamination problem. *Math. Mech. Solids*, 2003. Submitted.

[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.

[LMD03] M. LAMBRECHT, C. MIEHE, and J. DETTMAR. Energy relaxation of nonconvex incremental stress potentials in a strain-softening elastic-plastic bar. *Int. J. Solids Structures*, 40, 1369–1391, 2003.

[MAL02] C. MIEHE, N. APEL, and M. LAMBRECHT. Anisotropic additive plasticity in the logarithmic strain space: modular kinematic formulation and implementation based on incremental minimization principles for standard materials. *Comput. Methods Appl. Mech. Engrg.*, 191(47-48), 5383–5425, 2002.

[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, 85–123, 1999.

[Mie00]   A. MIELKE. Estimates on the mixture function for multiphase problems in elasticity. In A.-M. Sändig, W. Schiehlen, and W. Wendland, editors, *Multifield problems*, pages 96–103, Berlin, 2000. Springer–Verlag.

[Mie02a] C. MIEHE. Strain-driven homogenization of inelastic microstructures and composites based on an incremental variational formulation. *Internat. J. Numer. Methods Engrg.*, 55(11), 1285–1322, 2002.

[Mie02b] A. MIELKE. Finite elastoplasticity, Lie groups and geodesics on SL(d). In P. Newton, A. Weinstein, and P. Holmes, editors, *Geometry, Dynamics, and Mechanics*, pages 61–90. Springer–Verlag, 2002.

[Mie03a] C. MIEHE. Computational micro-to-macro transitions for discretized microstructures of heterogeneous materials at finite strains based on the minimization of averaged incremental energy. *Comput. Methods Appl. Mech. Engrg.*, 192(5-6), 559–591, 2003.

[Mie03b] A. MIELKE. Energetic formulation of multiplicative elasto–plasticity using dissipation distances. *Cont. Mech. Thermodynamics*, 15, 351–382, 2003.

[Mie03c] A. MIELKE. Existence of minimizers in incremental elasto–plasticity with finite strains. *SIAM J. Math. Analysis*, 2003. Submitted.

[ML03]   C. MIEHE and M. LAMBRECHT. Analysis of microstructure development in shearbands by energy relaxation of incremental stress potentials: large–strain theory for generalized standard solids. *Int. J. Numer. Meth. Engrg.*, 58, 1–41, 2003.

[MM03]   A. MAINIK and A. MIELKE. Existence results for energetic models for rate–independent systems. SFB 404, Preprint 2003/34, 2003.

[Mor66]   C. MORREY. *Multiple Integrals in the Calculus of Variations.* Springer, 1966.

[Mor76]   J.-J. MOREAU. Application of convex analysis to the treatment of elastoplastic systems. In P. Germain and B. Nayroles, editors, *Applications of methods of functional analysis to problems in mechanics*, pages 56–89. Springer-Verlag, 1976. Lecture Notes in Mathematics, 503.

[MR03]   A. MIELKE and T. ROUBÍČEK. A rate–independent model for inelastic behavior of shape–memory alloys. *Multiscale Model. Simul.*, 2003. In print.

[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.

[MT03]   A. MIELKE and F. THEIL. On rate–independent hysteresis models. *Nonl. Diff. Eqns. Appl. (NoDEA)*, 2003. In print.

[MTL02] A. MIELKE, F. THEIL, and V. LEVITAS. A variational formulation of rate–independent phase transformations using an extremum principle. *Arch. Rational Mech. Anal.*, 162, 137–177, 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.

[Ott98] F. OTTO. Dynamics of labyrinthine pattern formation in magnetic fluids: a mean-field theory. *Arch. Rational Mech. Anal.*, 141(1), 63–103, 1998.

[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 micro-magnetics. *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.*, 4, 111–136, 2002.

[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.

[The02] F. THEIL. Relaxation of rate-independent evolution problems. *Proc. Roy. Soc. Edinburgh Sect. A*, 132, 463–481, 2002.

[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.