Evolution of Rate–Independent Inelasticity with Microstructure using Relaxation and Young Measures

Alexander Mielke

Universität Stuttgart, Mathematisches Institut A
Pfaffenwaldring 57, 70569 Stuttgart, Germany
mielke@mathematik.uni-stuttgart.de

Abstract: We use the energetic formulation for rate–independent inelasticity which is based in the dissipation distance on the internal state. From this the incremental problems inherit a variational form which can be used to derive suitable relaxations. As a result we obtain a similar energetic formulation on the set of Young measures which then describes the evolution of microstructure.

Keywords: Rate–independent inelasticity, finite–strain elastoplasticity, dissipation distance, quasi–convexity, relaxation, Young measures.

1 Introduction

The theory of elastoplasticity at finite strain has undergone a rapid development in the last few decades, see for instance [SO85, Sim88, MS92, Mie96, OS99]. It can be understood as a special case of general inelastic material behavior (cf. [ZW87, Mie00]). For this general setting of inelasticity a new energetic formulation was introduced in [MTL98, MTL02, CHM01, Mie03]. This framework has the major advantage that it doesn't make any assumption on the smoothness of the processes to be described; hence, it is particularly suited to describe systems where formation of microstructure takes place. We note that this phenomenon can be seen as a phase transformation where the deformation gradients as well as the internal variables have jumps in space and time. To minimize the macroscopic energy in such situations the systems develop spatial oscillations on microscopic scales which need to be described by macroscopic quantities. In general, simple averages are not enough and therefore we use Young measures to describe these oscillations. The measures describe for each material point (or representative volume element) the joint distribution of the deformation gradient and the internal variables. The major task is then to find an evolution equation for these Young measures. Here we derive such equations via relaxation of the original model. Thus we provide a rather abstract evolution law which has all

Published in:
IUTAM Symposium on Computational Mechanics of Solid Materials at Large Strains,
C. Miehe (edr), Kluwer, 2003, pp. 33-44.

mathematically desirable properties but which is very complicated to use in practice. For more details, in particular to the specific application to elastoplasticity at finite strain, we refer to [Mie02, Mie03, Mie01].

2 Rate–independent inelasticity

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 properties 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, z) we assume that the function $\widehat{\psi}(x, \cdot, z)$: $\mathrm{GL}_+(\mathbb{R}^d) \mapsto [0, \infty)$ is coercive (i.e., $\widehat{\psi}(x, \mathbf{F}, z) \to \infty$ if $(\det \mathbf{F})^{-1} + |\mathbf{F}| \to \infty$) and quasiconvex. 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

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

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

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

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

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

It is assumed 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 $c > 0$. Considering a process $z : [0, T] \times \Omega \mapsto Z$ the dissipation on an interval $[t_0, t_1]$ is then

$$\mathrm{Diss}(z; [t_0, t_1]) = \int_{t_0}^{t_1} \int_\Omega \Delta(x, z(t, x), \dot{z}(t, x)) \,\mathrm{d}x \,\mathrm{d}t.$$

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

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

This defines also a metric on the set of internal states $\mathcal{Z} = \{\ z : \Omega \mapsto \mathbb{R}^n \mid z$ measurable $\}$ by setting $\mathcal{D}(z_0, z_1) = \int_\Omega \widehat{D}(x; z_0(x), z_1(x))\,\mathrm{d}x$ for $z_0, z_1 \in \mathcal{Z}$.

3 Energetic formulation of the model

Our model is completely described by the two constitutive relations $\psi = \widehat{\psi}(x, \mathbf{F}, z)$ and $\Delta = \widehat{\Delta}(x, z, \dot{z})$ and by the loadings $\ell(t)$.

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

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

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

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

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

The stability condition (S) has a clear mechanical interpretation. Letting $\widehat{z} = z(t)$ we have $\mathcal{D}(z(t), \widehat{z}) = 0$ and the condition implies that $\boldsymbol{\varphi}(t)$ is the global minimizer of $\mathcal{E}(t, \cdot, z(t))$ on $\mathcal{F}$, which gives the elastic equilibrium. Moreover, changing the internal variable from $z(t)$ to $\widehat{z}$ (and adjusting $\widehat{\boldsymbol{\varphi}}$ optimally) the dissipation must be at least as large as the elastic energy release. The internal variable z will change (and dissipate energy) as soon as the elastic energy release is large enough to compensate for the dissipation. Thus, (S) is equivalent to a principle of maximal dissipation. The energy inequality (E) has an obvious interpretation, since the work of the external forces is given by the last term. This abstract setting has applications in many rate–independent continuum models. We refer to [MTL98, KMS99, GMH01, MTL02] for applications to phase transformations in shape memory alloys. Finally we note that the energetic formulation using (S) & (E) is equivalent to the classical local flow rules, when the solution processes are sufficiently smooth, see [Mie02, Mie03].

One of the 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 $k = 1, \dots, N$ find iteratively $(\boldsymbol{\varphi}_k, z_k) \in \mathcal{F} \times \mathcal{Z}$ such that

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

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

Theorem 3.1 *If $(\boldsymbol{\varphi}_k, z_k)_{k=1,\dots,N}$ is a solution of (IP), then for $k = 1, \dots, N$ we have*
(i) stability of $(\boldsymbol{\varphi}_k, z_k)$ at time t_k, i.e.,
$\mathcal{E}(t_k, \boldsymbol{\varphi}_k, z_k) \le \mathcal{E}(t_k, \widehat{\boldsymbol{\varphi}}, \widehat{z}) + \mathcal{D}(z_k, \widehat{z})$ for all $(\widehat{\boldsymbol{\varphi}}, \widehat{z})$,
and (ii) the discretized energy estimate

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

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

$$\Psi^{\mathrm{red}}(z_{\mathrm{old}}; x, \mathbf{F}) = \min_{z \in Z} \Big[\widehat{\psi}(x, \mathbf{F}, z) + \widehat{D}(x, z_{\mathrm{old}}, z) \Big]$$

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

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

eeq Thus, the kth step of (IP) is solved if we find a minimizer of $\mathcal{E}^{\mathrm{red}}_{z_{k-1}}(t_k, \cdot)$ on $\mathcal{F}$. Then, $(\boldsymbol{\varphi}_k, z_k)$ with $z_k(x) = \widehat{Z}_{\mathrm{new}}(x, \mathrm{D}\boldsymbol{\varphi}_k(x), z_{k-1}(x))$ is the desired minimizer in (IP).
The minimization problem for $\mathcal{E}^{\mathrm{red}}_{z_{k-1}}(t_k, \cdot)$ has the standard form of a problem of nonlinear elasticity, where $\Psi^{\mathrm{red}}(z_{k-1}(x); x, \cdot)$ plays the role of the elastic potential. By construction we have frame indifference, i.e. $\Psi^{\mathrm{red}}(z; x, \mathbf{F}) = \Psi^{\mathrm{red}}(z; x, \mathbf{RF})$ for all $\mathbf{R} \in \mathrm{SO}(\mathbb{R}^d)$. However, other properties like the coercivity with respect to $\mathbf{F}$ are not 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. [Mie03].

Similarly, quasi– and rank–one convexity may no longer hold for $\Psi^{\rm red}(z;x,\cdot)$ and we have to expect the formation of microstructure in infimizing sequences, see [OR99, CHM01]. As a result, (IP) and consequently (S) & (E) may not have a solution. It is this case for which our theory below is developed.

4 Relaxation

Formation of microstructure occurs, if the minimization problem (IP) has no solution, cf. [BJ87, CHM01]. If the incremental problem (IP) has no solution, we may consider the following approximate incremental problem.

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

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

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

Obviously, this problem has, for all $\varepsilon > 0$, a solution and the question is how the solutions $(\boldsymbol{\varphi}_k^\varepsilon, z_k^\varepsilon)$ behave for $\varepsilon \to 0$. We cannot expect pointwise convergence but certain macroscopic quantities should have limits for $\varepsilon \to 0$. To define an abstract notion of relaxation we introduce a generalized convergence "$\xrightarrow{\mathbb{W}}$" on an enlarged space $\mathbb{W}$. This space is connected to $\mathcal{F}\times\mathcal{Z}$ via a continuous embedding $\mathcal{J} : \mathcal{F}\times\mathcal{Z} \mapsto \mathbb{W}$. 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 inital datum $w_0 \in \mathbb{W}$ and the time discretization $0 = t_0 < t_1 < \ldots < t_N = T$ as follows.

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

$$w_k \in \operatorname*{argmin}_{w\in\mathbb{W}} \ \mathbb{E}(t_k, w) + \mathbb{D}(w_{k-1}, w).$$

We do not ask for the conditions $\mathbb{D}(\mathcal{J}(0,z_0), \mathcal{J}(0,z_1)) = \mathcal{D}(z_0,z_1)$ and $\mathbb{E}(t, \mathcal{J}(\boldsymbol{\varphi}, z)) = \mathcal{E}(t, \boldsymbol{\varphi}, z)$. Hence, 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) *[Minimality] $\mathcal{J}(\mathcal{F}\times\mathcal{Z})$ is dense in $\mathbb{W}$.*

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

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

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

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

Moreover the relaxed incremental problem ($\mathbb{RIP}$) can be interpreted as the incremental problem associated to the following relaxed energetic formulation: The function $w : [0,T] \mapsto \mathbb{W}$ is called a solution process 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*

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

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

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

Here $\boldsymbol{\varphi} = \boldsymbol{\Phi}(w)$ is the macroscopic deformation $\boldsymbol{\varphi}$ associated to the limit $w \in \mathbb{W}$. The relaxed dissipation is given by

$$\mathrm{Diss}^{\mathrm{rel}}(w; [t_1, t_2]) = \sup \sum_{j=1}^{M} \mathbb{D}(w(\tau_{j-1}), w(\tau_j))$$

where the supremum is taken over all $M \in \mathbb{N}$ and all discretizations $t_1 \le \tau_0 < \tau_1 < \ldots < \tau_{M-1} < \tau_M \le t_2$.

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

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

Another way to define relaxations for rate independent problems of the type (S) & (E) is proposed in [The01]. This definition avoids totally the usage of incremental problems and works directly with (S) & (E) and ($\mathbb{S}$) & ($\mathbb{E}$.) An application of this theory to phase transformation is given in [MTL02, The01].

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 [Rou97, Mül99]. One particular instance of this relaxation was studied in [MT99, GMH01, MTL02], where phase transformations in shape memory alloys were studied.

We refer to [Bal89, Ped97, Rou97, Mie99] for the basic notions of Young measures and repeat here only the definitions. Denote by $\mathrm{Prob}(Z)$ the set of all probability measures on Z. The **Young measures** on Ω with values in Z are given by

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

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 $g \in \mathrm{C}_0(Z, \mathbb{R})$ we have

$$\int_{|x_0-x|<\varepsilon} g(z^j(x))\,\mathrm{d}x \to \int_{|x_0-x|<\varepsilon}\int_Z g(z)\mu(x,\mathrm{d}z)\,\mathrm{d}x \text{ for } j\to\infty.$$

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.

Given a (dissipation) metric $\widehat{D} : Z\times Z \mapsto [0,\infty]$, we have 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 in an optimal way. We start with the Wasserstein metric D^{Wass} on $\mathrm{Prob}(Z)$ which

associates to $\widehat{D}$. For given $\nu_0, \nu_1 \in \mathrm{Prob}(Z)$ we want to transport the mass, which at the beginning is distributed according to ν_0, in such a way that the final distribution is ν_1 and that the dissipation cost is minimal. Denoting by $\sigma \in \mathrm{Prob}(Z \times Z)$ a probability measure such that $\sigma(\mathrm{d}z_0, \mathrm{d}z_1)$ gives the amount of mass which from a neighborhood of z_0 is transported into a neighborhood of z_1, we have to infimize with respect to σ under the given initial and final distributions:

$$D^{\mathrm{Wass}}(\nu_0, \nu_1) = \inf \Big\{ \int_{Z\times Z} \widehat{D}(z_0, z_1) \sigma(\mathrm{d}z_0, \mathrm{d}z_1) \,\Big|\, \sigma \in \mathrm{Prob}(Z \times Z), \\ \int_Z \sigma(\cdot, \mathrm{d}z_1) = \nu_0,\ \int_Z \sigma(\mathrm{d}z_0, \cdot) = \nu_1 \Big\}.$$

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

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

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

For two Young measures $\mu_0, \mu_1 \in \mathrm{YM}(\Omega, Z)$ we set

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

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

There is a special case of Young measures, namely those which are generated by sequences which are gradients of functions, like the deformation gradients in continuum mechanics $\mathbf{F}^j = \mathrm{D}\boldsymbol{\varphi}^j \in \mathbb{R}^{d\times d}$. The Young measures obtained as limits from gradients form a subset of all Young measures and are called **gradient Young measures** (GYM). They are characterized as follows (cf. [KP94, Ped97, Rou97]):

$$\gamma \in \mathrm{GYM}(\Omega, \mathbb{R}^{d\times d}) \iff \\ \left\{ \begin{array}{l} \text{There exists } \boldsymbol{\varphi} : \Omega \mapsto \mathbb{R}^d \text{ such that for a.e. } x \in \Omega \text{ we have} \\ \text{(a) } \mathrm{D}\boldsymbol{\varphi}(x) = \int_{\mathbb{R}^{d\times d}} \mathbf{G}\, \gamma(x, \mathrm{d}G) \text{ and (b) } \gamma(x) \in \mathrm{Prob}^{\mathrm{grad}}(\mathbb{R}^{d\times d}), \end{array} \right.$$

where

$$\mathrm{Prob}^{\mathrm{grad}}(\mathbb{R}^{d\times d}) = \\ \Big\{ \nu \in \mathrm{Prob}(\mathbb{R}^{d\times d}) \,\Big|\, \text{for all quasiconvex } Q : \mathbb{R}^{d\times d} \mapsto \mathbb{R} \text{ we have} \\ Q(\textstyle\int_{\mathbb{R}^{d\times d}} \mathbf{G}\, \nu(\mathrm{d}\mathbf{G})) \le \int_{\mathbb{R}^{d\times d}} Q(\mathbf{G})\, \nu(\mathrm{d}\mathbf{G}) \Big\}$$

We shortly write $\gamma \in \mathrm{GYM}_{\boldsymbol{\varphi}}(\Omega, \mathbb{R}^{d\times d})$ if (a) and (b) hold. The above condition (b) is essential as it reflects the information of compatibility of different gradients. For instance, $\nu = \theta\delta_A + (1-\theta)\delta_B$ with $\theta \in (0,1)$ and $A, B \in \mathbb{R}^{d\times d}$ lies in $\mathrm{Prob}^{\mathrm{grad}}(\mathbb{R}^{d\times d})$ if and only if $\mathrm{rank}(A-B) \le 1$.

6 Separate relaxation

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

$$
\begin{aligned}
&w = (\boldsymbol{\varphi}, \mu) \in \mathcal{F}\times\mathrm{YM}(\Omega, Z) =: \mathbb{W}, \qquad \mathcal{J}(\boldsymbol{\varphi}, z) = (\boldsymbol{\varphi}, \delta_z),\\
&(\boldsymbol{\varphi}^j, \mu^j) \xrightarrow{\mathbb{W}} (\boldsymbol{\varphi}^\infty, \mu^\infty) \iff\\
&\begin{cases} \boldsymbol{\varphi}^j \rightharpoonup \boldsymbol{\varphi}^\infty \text{ in } \mathrm{W}^{1,p}(\Omega, \mathbb{R}^d),\\ \forall\, g \in \mathrm{C}_0(Z) : \int_Z g(z)\,\mu^j(\mathrm{d}z) \overset{*}{\rightharpoonup} \int_Z g(z)\,\mu^\infty(\mathrm{d}z) \text{ in } \mathrm{L}^1_{\mathrm{loc}}(\Omega). \end{cases}
\end{aligned}
$$

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

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

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

Theorem 6.1 *Under suitable technical assumptions the relaxation defined via $\mathbb{W}$, $\mathcal{J}$, $\mathbb{E}$ and $\mathbb{D}$ satisfies the properties (R1)—(R3) in Definition 4.1 as well as (R5).*

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

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

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

where the relaxed potential is given by $\Psi_{\rm rel}(x,\mathbf{F},\nu) =$

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

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

Example 6.1 To illustrate the above concept we consider the case $d=1$, $Z=\mathbb{R}$, and $\widehat{\psi}(\mathbf{F},z)=(\mathbf{F}{-}z)^2$. We obtain $\Psi_{\rm rel}(\mathbf{F},\nu)=[\mathbf{F}{-}\int_{\mathbb{R}} z\,\nu(\mathrm{d}z)\,]^2$, since in the definition of the relaxed potential we may choose any $z:(0,1)\mapsto\mathbb{R}$ generating ν and then define $\boldsymbol{\phi}$ via $\boldsymbol{\phi}(0)=0$ and $\boldsymbol{\phi}'(y)=z(y){-}\int_0^1 z(s)\,\mathrm{d}s$. Using $\int_0^1 z(s)\,\mathrm{d}s=\int_{\mathbb{R}} z\,\nu(\mathrm{d}z)$ the result follows. Moreover, we have

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

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

Remark 6.1 Up to now the relaxed elastic potential can be computed only in a few cases, e.g., for applications in phase transformations in shape memory alloys, where $Z=\{1,2,\dots,n\}$ is the discrete set of possible phases, cf. [MTL02, GMH01]. There $\mathrm{Prob}(Z)$ can be identified with the polytope $P^n=\{\,(\theta_1,\dots,\theta_n)\in\mathbb{R}^n\mid\theta_j\ge 0,\ \sum_1^n\theta_j=1\,\}$ and $\mathrm{YM}(\Omega,Z)=\mathrm{L}^\infty(\Omega,P^n)$. The dissipation distance $\mathbb{D}$ takes the form $\mathbb{D}(\mu_{\rm old},\mu_{\rm new})=\int_\Omega\|\theta_{\rm new}(x){-}\theta_{\rm old}(x)\|\,\mathrm{d}x$ for a suitable polyhedral vector norm $\|\cdot\|$ on $\mathbb{R}^n$.

Acknowledgments

The research was partially supported by DFG through SFB 404 "Multifield Problems". The author is grateful for stimulating discussions with Carsten Carstensen, Klaus Hackl, and Christian Miehe.

References

[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.

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

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

[GMH01] 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. 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.

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

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

[Mie00] C. Miehe. Elements of computational inelasticity at finite strains. In preparation, 2000.

[Mie01] A. Mielke. Evolution of microstructure in elastoplasticity using relaxation and Young measures. In preparation, 2001.

[Mie02] A. Mielke. Finite plasticity, Lie groups and geodesics on SL(d). In "P. Newton, A. Weinstein, and P. Holmes, editors, *Geometry, Dynamics, and Mechanics*. Springer–Verlag, 2002", pp. 61–90.

[Mie03] A. Mielke. Energetic formulation of multiplicative elastoplasticity using the dissipation metric. *Cont. Mech. Thermodynamics*, 2003. To appear.

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

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

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

[MTL98] A. Mielke, F. Theil, and V. Levitas. Mathematical formulation of quasistatic phase transformations with friction using an extremum principle. Preprint A8, Hannover, Sept. 1998.

[MTL02] A. Mielke, F. Theil, and V. Levitas. A variational formulation of rate–independent phase transformations using an extremum principle. *Archive Rational Mech. Analysis*, 162, 137–177, 2002.

[Mül99] S. Müller. Variational models for microstructure and phase transitions. In *Calculus of variations and geometric evolution problems (Cetraro, 1996)*, pages 85–210. Springer, Berlin, 1999.

[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.

[OS99] M. Ortiz and L. Stainier. The variational formulation of viscoplastic constitutive updates. *Comput. Methods Appl. Mech. Engrg.*, 171(3-4), 419–444, 1999.

[Ped97] P. Pedregal. *Parametrized measures and variational principles*. Birkhäuser Verlag, Basel, 1997.

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

[Rou97] T. Roubíček. *Relaxation in optimization theory and variational calculus*. Walter de Gruyter & Co., Berlin, 1997.

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

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

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

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