Universität Stuttgart
Mathematisches Institut A
1. Lehrstuhl

Angewandte Analysis
Pfaffenwaldring 57
D–70569 Stuttgart

http://www.mathematik.uni-stuttgart.de/mathA/lst1/

A variational formulation of rate–independent phase transformations using an extremum principle

by

Alexander Mielke
Mathematisches Institut A, Universität Stuttgart,
Pfaffenwaldring 57, 70569 Stuttgart, Germany

Florian Theil
Mathematical Institute, University of Oxford,
24-29 St Giles', OX1 3LB Oxford, U.K.

and

Valery I. Levitas
Department of Mechanical Engineering, Texas Tech University,
Lubbock, Texas 79409-1021, U.S.A.

28 June 2000

A variational formulation of rate–independent phase transformations using an extremum principle*

Alexander Mielke†, Florian Theil‡, Valery I. Levitas§

June 28, 2000

Dedicated to Erwin Stein on the occasion of his seventeeth birthday

Contents

---

*Research supported by VW–Stiftung under the grant *Strain induced phase transitions in engineering materials*, no. I 70/282 and I 70/284.

†Mathematisches Institut A, Universität Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany
‡University of Oxford, Mathematical Institute, 24-29 St Giles', OX1 3LB Oxford, U.K.
§Texas Tech University, Department of Mechanical Engineering, Lubbock, Texas 79409-1021, U.S.A.

1 Introduction

In this paper we present a mathematical approach to the modeling of phase transformations (PT) in certain elastic materials, like martensitic PT in shape memory alloys. However, the aims of this paper are wider and there are three different lines of thought which lead somehow parallel through this work. First, we want to present a simple multi–dimensional model for the slow evolution of PT processes in solids which are able to describe hysteresis phenomena. Second, we discuss a general relaxation method for rate–independent time evolution processes which are ill–posed in their original form due to the formation of microstructure. Third, we show that the derived evolution problem can be analyzed in some special nontrivial cases.

Elastic materials which allow for PT (like single or poly–crystalline shape memory alloys) display hysteresis effects under quasistatic loading cycles. While there are many engineering models for this phenomenon for both uniaxial and multiaxial loading (see e.g. [Wil93, HuM93, Lev94, LeS97, BoL97, PAM98, HaG99, LIS99, KMS99, SP*99] and the references therein), mathematically rigorous models are only developed for simplified situations. On the one hand there is an extensive body of mathematical work on spatially one–dimensional situations (uniaxial case), see e.g. [KLMS97, KuM00, MüS96]. On the other hand there is a rich mathematical literature on static problems for the multi–dimensional situation. Most of that work is restricted to zero–stress or zero–energy solutions where the crystallographic properties and the transformation strain tensors determine the set of reachable macroscopic strains, see e.g. [BaJ87, BaJ92, Bha93, BCJ95, BhK96, Mül98].

Sofar there is no mathematical treatment of models which combine the hysterestic behavior with three–dimensional models of elasticity. It is here where we make a first proposal which combines these two features. Our model is truly three–dimensional, it has hysteretic behavior and can deal with nonzero stresses. The model is derived from very simple mechanical postulates, namely the thermodynamic PT criterion and definition of stable thermodynamic states. In the present version we neglect all thermal effects (isothermal case) and also all plastic effects.

Of course, our model is crude and should be considered as the simplest nontrivial model rather than a realistic model for specific shape memory alloys. It certainly will be a cornerstone for more elaborate models.

The second theme of this paper is the relaxation of time–evolution processes which are ill posed. This illposedness is due to the formation of microstructure which usually arises from neglecting microscopical effects which would bound the smallest scales. On a mesoscopic scale ones sees formation of very small scales which have to be described in an averaged sense by relaxation. Here we treat rate–independent processes and approximate them by an fully implicit incremental problem (as is often used in engineering for plasticity problems) which for each step takes the form of a minimization problem. These problems can be relaxed introducing a phase–mixture theory and the results can be understood of the incremental formulation of a time–continuous evolution problem for the phase mixtures.

This part is written in a very general and informal way to see that it is a general programme which opens up new ways to relax evolutionary problems derived from energy principles. In this part we claim no rigorosity in the analysis but try to convey the methodology.

The third theme in the paper is devoted to one specific special case of solid–solid PT, namely that of linearized elasticity with two phases only, both having the same elasticity tensor but different transformation strains. For this case we show that most steps in the relaxation procedure can be made rigorous. Moreover, we finally obtain an existence theorem for the relaxed time–continuous model. Here we use the general theory of rate–independent hysteresis as developed in [MiT00] which provides a incremental solutions and thus we obtain a limit function by Helly's principle. To show that this limit is a solution we need a further compactness argument in the spatial direction which is derived using pseudo–differential operators and H–measures.

Having established a mathematical model which enables us to study simultaneously hysteretic and multi–dimensional effects, it is possible to study the question of their interaction. In particular one should investigate the question whether multi–dimensional models display a richer class of hysteresis effects and whether it is really necessary to introduce more internal variables to explain the more complicated hysteresis loops observed in uniaxial experiments, cf. [Wil93, Lev94, MüS96, KuM00].

To be more specific, we sketch the underlying mechanical model. In many cases martensitic PT even in elastic materials (shape memory alloys) are accompanied by dissipation due to resistance against interface motion. We are mainly interested in athermal, rate–independent resistence [OlC86, GhO94]. Athermal resistance is similar to dry friction, i.e. PT can start when the driving force exceeds some dissipative threshold and this friction is independent of the interface velocity. It is caused by the interaction of interfaces with obstacles having long-range stress field (point defects, dislocations, grains, twins and subgrain boundaries) which cannot be overcome by thermal fluctuations. Short-range resistance which can be overcome by thermal activation gives rise to rate-dependent or viscous friction.

There are quasistatic models coupling rate–independent PT rules to an equation for the energy balance which contains rate–dependent heat flow. However their mathematical analysis needs still to be developed, cf. [BoL97, Lev98, KLMS97].

Our theory is based on the assumption that a PT is possible whenever the associated gain in the energy of the whole system is larger than the energy which is dissipated through the PT. This energy loss might also be understood as the height of the potential barrier between two potential wells: as soon as we reach the saddle point between the wells the energy difference between the saddle point and the next well is lost, see Section 4.1 for some of these modeling questions.

Moreover, we use an *extremum principle* to determine which PT (what phases are created in what regions) out of the set of all possible PT does occur. This principle is formulated in [Lev95, Lev97]) and says that PT occur as soon as they are possible (i.e. when the thermodynamic PT criterion is met), which results in the conclusion that always those PT actually occur which minimize the sum of the dissipation and the energy increment. In the context of irreversible thermodynamics or plasticity theory similar extremum principles are well known, see e.g. [Zie63] and the related literature.

The global version of the extremum principle appears first in [Lev98]. While in these works the theory is presented quite generally including plastic effects as well as surface energies, we here restrict ourself to the purely elastic case and neglect surface energies between the phase boundaries. This limits the applicability but makes the problem mathematically accessible since the extremum principle reduces in this case to the stability condition for the phase distributions.

Every material point x in the body $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, can be in one of n possible phases e_j, where $P^{\mathrm{p}} = \{e_1, \dots, e_n\}$ is the set of all possible pure phases. The material properties of each phase are given by the energy density $W_{e_j}(F)$ where $F = \nabla u$ is the deformation gradient. In martensitic PT one often assumes $W_{e_j}(F) = W(FT_j^{-1}) + \beta_j$ with a fixed energy density W, $T_1 = I$ and $\beta_1 = 0$ for the austenite phase, and T_j are the transformation strains for the martensitic phases (e.g. $n = 4$ for the cubic to tetragonal PT and $n = 13$ for the cubic to monoclinic PT, see [BhK96]). The matrices T_j are typically related to each other by the crystallographic symmetry group, and $\beta_j = \beta$ for $j = 2, \dots, n$ where β is positive (negative) if we are below (above) the critical temperature. The whole analysis works for general inhomogeneous materials with $W_{e_j}(x, F)$ but for notational convinience we restrict ourself to the homogeneous case.

A function $c^{\mathrm{p}} : \Omega \to P^{\mathrm{p}}$ is called a state of the system and for each (process) time dependent loading $G(t)$ the energy functional is given by

$$\mathcal{I}^{\mathrm{p}}(t, c) = \inf\{ \textstyle\int_\Omega W_{c^{\mathrm{p}}(x)}(\nabla u(x))\,\mathrm{d}x - \langle G(t), u\rangle \;\Big|\; u \in \mathcal{V} \},$$

where $\mathcal{V}$ is the set of kinematically admissible deformations. The minimization with respect to $u \in \mathcal{V}$ reflects the fact that the system is in elastic equilibrium. The superscript ${}^{\mathrm{p}}$ denotes "pure" and indicates the fact that at the present stage phase mixtures are not taken into account, i.e. at every point $x \in \Omega$ the material is in a **pure** phase. Under the assumptions that at most one PT occurs at each material point the dissipation of a PT from state $a^{\mathrm{p}} : \Omega \to P^{\mathrm{p}}$ into $c^{\mathrm{p}} : \Omega \to P^{\mathrm{p}}$ is

$$\mathcal{D}^{\mathrm{p}}(a^{\mathrm{p}}, c^{\mathrm{p}}) = \int_\Omega D^{\mathrm{p}}(a^{\mathrm{p}}(x), c^{\mathrm{p}}(x))\,\mathrm{d}x, \quad \text{where } D^{\mathrm{p}}(e_i, e_j) = \kappa_{i\to j}.$$

Here $\kappa_{j\to j} = 0$, and $\kappa_{i\to j} > 0$ for $j \neq i$ are the dissipation coefficients for a PT from e_i into e_j. They serve as a threshold which makes PT more difficult. A family of states $(c^{\mathrm{p}}(t))_{t\in[0,T]}$ is called a process, and the total amount of dissipation due to PT in the time interval $[t_1, t_2]$ is denoted by $\mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2)$.

We say that a state c^{p} is **stable at time** t if

$$\mathcal{I}^{\mathrm{p}}(t, b^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}, b^{\mathrm{p}}) \geq \mathcal{I}^{\mathrm{p}}(t, c^{\mathrm{p}}) \quad \text{for all } b^{\mathrm{p}} : \Omega \to P^{\mathrm{p}}.$$

This stability means exactly that there is no admissible PT from the given state $c^{\mathrm{p}}(t)$ which has a gain in the energy which is larger than the associated dissipation. With this notion we give the following time–continuous formulation of the extremum principle.

> ***(CP)*** *For given c_0^{p}, find a process $c^{\mathrm{p}} : [0, T] \times \Omega \to P^{\mathrm{p}}$ with $c^{\mathrm{p}}(0, x) = c_0^{\mathrm{p}}(x)$ such that for all t_1, t_2 with $0 \leq t_1 \leq t_2 \leq T$ the state $c^{\mathrm{p}}(t_1)$ is stable at t_1 and the energy inequality $\mathcal{I}^{\mathrm{p}}(t_2, c^{\mathrm{p}}(t_2)) + \mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2) \leq \mathcal{I}^{\mathrm{p}}(t_1, c^{\mathrm{p}}(t_1)) - \int_{t_1}^{t_2} \langle \dot{G}(s), u(s)\rangle\,\mathrm{d}s$ holds.*

We might as well have imposed energy equality, but due to the relaxation process done below we would have to give up such an equality anyway. The importance of the dissipation by friction is that it is rate–independent and nonnegative. Hence, this will be the major source of hysteretic behavior.

In general we cannot expect to find solutions of (CP) since it is not even clear whether for a given loading $G(t)$ there is any stable state $c^{\mathrm{p}} : \Omega \to P^{\mathrm{p}}$. A natural way to find stable states at time t is to minimize $\mathcal{I}^{\mathrm{p}}(t, c^{\mathrm{p}})$ over all c^{p}. Clearly such a minimizer must be stable, however, in general we can only expect to find an infimizing sequence c_n, that is $\mathcal{I}^{\mathrm{p}}(t, c_n^{\mathrm{p}}) \to \alpha = \inf\{ \mathcal{I}^{\mathrm{p}}(t, b^{\mathrm{p}}) \;\Big|\; b^{\mathrm{p}} : \Omega \to P \}$. The weak limit c of the sequence c_n^{p} will no longer be a simple phase function but gives rise to infinitesimally fine phase mixtures. The goal of this work is to discuss rigorous extensions of the problem from classical states c^{p} with $c^{\mathrm{p}}(t,x) \in P^{\mathrm{p}}$ to phase mixtures c with $c(t,x) \in P = \mathrm{conv}(P^{\mathrm{p}})$. In mathematical terms such a procedure is called relaxation. We insist however, that we do not include new modeling assumptions into a mixture theory, but we want to see how far the given functions $W_{e_j}(F)$ can be used to derive a consistent mixture theory.

At this point we want to connect our theory to the purely static theory of PT without dissipation which an active mathematical subject, see [BaJ87, BaJ92] and the survey [Mül98]. For our mechanical situation it would mean to consider only one stored energy function

$$\overline{W}(F) = \min\{\, W_{e_j}(F) \,\big|\, j = 1, \ldots, n \,\},$$

which now is no longer quasiconvex but has potential wells in the sets $\mathrm{SO}(3)T_j$. In infimizing the functional $\overline{\mathcal{E}}(t,u) = \int_\Omega \overline{W}(\nabla u(x))\,\mathrm{d}x - \langle G(t), u\rangle$ the infimum of $\overline{\mathcal{E}}(t,\cdot)$ is not achieved but infimizing sequences show the development of microstructure. For a function u we say that the point $x \in \Omega$ is the phase e_j if $\overline{W}(\nabla u(x)) = W_{e_j}(\nabla u(x))$. Of course, when always looking for the global minimizer we cannot obtain hysteretic behavior in any quasistatic loading process as no time direction is prefered. Nevertheless hysteretic behavior can occur (see [BCJ95]) due to local strong minimizers which have a true potential barrier in a suitable function space. To describe this phenomenon with our approach we have to choose the dissipation functional $\mathcal{D}^{\mathrm{p}}(a^{\mathrm{p}}, c^{\mathrm{p}})$ such that for each such local strong minimizer u_j the associated phase function $c^{\mathrm{p}}_j : \Omega \to \mathcal{P}^{\mathrm{p}}$ is a global minimum of $a^{\mathrm{p}} \mapsto \mathcal{I}^{\mathrm{p}}(t, a^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}_j, a^{\mathrm{p}})$. In Section 4.1 we discuss the relation between the model with internal variables c^{p} and that without but having potential barriers.

The main idea to find the proper relaxation is to replace the time–continuous problem (CP) by an incremental problem (IP) which has a variational structure arising from the extremum principle. The variational problem can be relaxed using mathematical tools leading to the relaxed incremental problem (RIP). However, here we propose a separately relaxed incremental problem (SRIP) which might not coincide with the mathematical relaxation in general, but see [Th00b] for a rigorous justifications in special cases. In general the separate relaxation will neglect incompatibilities between the microstructures of two phase mixtures and therefore underestimate the true dissipation.

Its major advantage is that it is explicit enough to show existence and give a solution algorithm for (SRIP). From this relaxation we can even state a time–continuous relaxed problem (SRCP). Another successful relaxation of a time-continuous problem for pattern formation in magnetic fluids was obtained in [Ott98]; there again an incremental problem in variational form was formulated whose relaxation appears to be the incremental problem of the time–continuous relaxed problem.

To be specific, for $0 < t_1 < \ldots < t_{N-1} < t_N = T$ we define the *incremental problem*:

(IP) *For given c^{p}_{k-1}, find c^{p}_k such that $\mathcal{I}^{\mathrm{p}}(t_k, c^{\mathrm{p}}_k) + \mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}_{k-1}, c^{\mathrm{p}}_k)$ is minimal.*

The main problem in treating (IP) as a mathematical or numerical problem is that in general this minimization process does not have a solution. For given c^{p}_{k-1} the infimum α_k of $\mathcal{J}^{\mathrm{p}}(t_k, c^{\mathrm{p}}_{k-1}, c^{\mathrm{p}}_k) = \mathcal{I}^{\mathrm{p}}(t_k, c^{\mathrm{p}}_k) + \mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}_{k-1}, c^{\mathrm{p}}_k)$ with respect to c^{p}_k is not attained. This means that we can find $c^{\mathrm{p}}_{k,\varepsilon}$ with $\mathcal{J}^{\mathrm{p}}(t_k, c^{\mathrm{p}}_{k-1}, c^{\mathrm{p}}_{k,\varepsilon}) \le \varepsilon + \alpha_k$ for each $\varepsilon > 0$ but not for $\varepsilon = 0$. This problem arises from the physical fact that the sequence $c^{\mathrm{p}}_{k,\varepsilon}$ has to develop finer and finer structure such that the sequence $c^{\mathrm{p}}_{k,\varepsilon}$ does not converge pointwise in Ω.

However, subsequences will converge in a weak sense to a function c_k which now does no longer take values in P^{p} but its convex hull

$$P = \{\, c \in \mathbb{R}^n \,\big|\, c = \textstyle\sum_{j=1}^n \theta_j e_j,\ \theta_j \geq 0,\ \sum_{j=1}^n \theta_j = 1 \,\}.$$

Thus, we are automatically led to consider microscopically fine phase mixtures $c_{k,\varepsilon}$, $\varepsilon \ll 1$ which can be described by macroscopical functions $c_k(x)$ which contain only information on the volume fractions $\theta_j(x)$ of the phases.

Since many physically interesting quantities only depend on the limit c_k it is desirable to find c_k directly through a new incremental problem. The mathematical tool to be used here is the method of relaxation (see [Dac89, Rou97]) which is well–known in the theory of nonconvex variational problems like in optimal design theory. There the aim is to find an extension $\mathcal{J}(t_k, a, c)$ of $\mathcal{J}^{\mathrm{p}}(t_k, a^{\mathrm{p}}, c^{\mathrm{p}})$ to functions $a, c : \Omega \to P$ such the solutions $c_1, \ldots, c_N$ of the following *relaxed incremental problem* (RIP) are in fact the weak limits of infimizing sequences $c^{\mathrm{p}}_{k,\varepsilon}$ for (IP). The question of the proper relaxation for such incremental problems is not yet understood since it is not enough to consider the relaxation of $\mathcal{J}(t_k, \cdot, \cdot)$, see [Th00b] for a more detailed investigation and positive results for this problem.

Here we propose a separate relaxation for $\mathcal{I}^{\mathrm{p}}$ and $\mathcal{D}^{\mathrm{p}}$ which is based and the **mixture function** $\mathbb{W}(c, F)$ and the dissipation function $\Delta : \mathbb{R}^n \to \mathbb{R}_{\geq}$. The mixture function is defined via

$$\mathbb{W}(c, F) \stackrel{\text{def}}{=} \inf\{\, \textstyle\fint_B W_{c^{\mathrm{p}}(y)}(F+\nabla\phi(y))\,\mathrm{d}y \,\big|\, \phi \in \mathrm{W}_0^{1,p}(B),\ c^{\mathrm{p}}(x) \in P^{\mathrm{p}},\ \fint_B c^{\mathrm{p}}(y)\,\mathrm{d}y = c \,\}$$

and describes the deformation energy of mixtures. Note that this mixture function is not postulated but follows from an exact mathematical theory which is similar to homogenization theory. In fact there is a big literature on methods to how to calculate or estimate $\mathbb{W}(c, F)$ which relates to optimal bounds for composite materials with given volume fractions, see [Koh91, GrK95, SmW98, Mie00].

Defining the relaxed dissipation functional $\boldsymbol{\Delta}(z) = \int_\Omega \Delta(z(x))\,\mathrm{d}x$ and the relaxed energy functional

$$\mathcal{I}(t, c) = \inf\{\, \textstyle\int_\Omega \mathbb{W}(c(x), \nabla u(x))\,\mathrm{d}x - \langle G(t), u\rangle \,\big|\, u \in \mathcal{V} \,\},$$

the **separately relaxed incremental problem** reads

> ***(SRIP)*** *For given $c_{k-1} : \Omega \to P$, find $c_k : \Omega \to P$ such that $\mathcal{I}(t_k, c_k) + \boldsymbol{\Delta}(c_k - c_{k-1})$ is minimal.*

The minimization in (SRIP) is a simultaneous minimization on c and u which has the advantage that c appear only locally under the integral. Hence the minimization with respect to c can be done pointwise first and then u can be determined, see Section 5.5. The structure of (SRIP) shows that it is nothing else than the incremental version of the following time–continuous problem.

> ***(SRCP)*** *For given c_0, find $c : [0, T] \to \mathcal{P}$ with $c(0, x) = c_0(x)$ such that for all t_1, t_2 the state $c(t_1)$ is stable at time t_1, i.e., $\mathcal{I}(t_1, c(t_1)) \leq \mathcal{I}(t_1, a) + \boldsymbol{\Delta}(a - c(t_1))$ for all $a \in \mathcal{P}$, and that the energy inequality $\mathcal{I}(t_2, c(t_2)) + \int_{t_1}^{t_2} \boldsymbol{\Delta}(\mathrm{d}c) \leq \mathcal{I}(t_1, c(t_1)) - \int_{t_1}^{t_2} \langle \dot{G}(s), u(s)\rangle\,\mathrm{d}s$ holds.*

In Section 6 we present an existence theorem for a special case of (SRCP) which relates to a PT problem with two phases which both consist of linearized elastic materials with the same elastic tensor. This theory is based on the abstract methods for rate–independent problems which are worked out in Section 3 and given in more detail in [MiT00]. The basic idea is to solve the incremental problem (SRIP) for finer and finer time discretizations and then use the a priori bound on

$$\int_0^T \mathbf{\Delta}(\mathrm{d}c) = \sup\{\, \textstyle\sum_{k=1}^N \mathbf{\Delta}(c(t_k)-c(t_{k-1})) \,\big|\, N\in\mathbb{N}, 0\le t_0 < t_1 < \ldots < t_N \le T\,\}$$

to extract a subsequence which converges weakly for each $t\in[0,T]$. A special analysis is then needed to show that the stability property is weakly closed. We use pseudo–differential operators and the H–measure there.

Similar incremental approaches to rate–independent processes were developed in [HaR95] for problems in elastoplasticity. See also [ACZ99, CHM00] for a more general incremental formulation in elastoplasticity, including the case of finite deformations, which again has the form of a minimization problem.

Acknowledgments

This work emanated within the research program *Strain induced phase transitions in engineering materials* supported by the Volkswagen–Stiftung under grant no. I 70/282 and I 70/284. The authors are thankful to Gilles Francfort, Sanjay Govindjee, Mietek Kuczma, Markus Kunze, Stefan Müller and Augusto Visintin for helpful discussions.

2 Thermomechanic Theory

In this mechanically oriented part we will use notations which are common in the thermomechanical literature. This should enable readers with a mechanical background to follow the derivation of the model. Tensors and vectors are written with bold letters, the Helmholtz energy is denoted as ψ. In the remainder of the paper we will use the established notation from the mathematical literature. For easy comparison we give a list of the most important symbols in their two different notations

Notations in	thermomechanics	mathematics
space dimension	3	d
deformation	$\boldsymbol{u}$	u
deformation gradient	$\boldsymbol{F}=\nabla\boldsymbol{u}$	$F=\nabla u$
Helmholtz energy	$\psi=\widehat{\psi}(e_j,\boldsymbol{F})$	$W_{e_j}(F)$
loading	$\int_S \boldsymbol{p}\cdot\boldsymbol{u}\,\mathrm{d}a+\int_\Omega \boldsymbol{f}\cdot\boldsymbol{u}\,\mathrm{d}x$	$\langle G,u\rangle$
Gibbs free energy	$\mathcal{G}$	$\mathcal{I}$

We consider an elastic body $\Omega\subset\mathbb{R}^d$, $d=1,2,3$ in its reference configuration. Its surface is denoted by $S=\partial\Omega$; on the part $S_{\boldsymbol{p}}$ of the surface the stress vector $\boldsymbol{p}$ (surface traction) is prescribed and on the other part $S_{\boldsymbol{u}}=S\setminus S_{\boldsymbol{p}}$ the deformation $\boldsymbol{u}$ is given. Of course, more complicated mixed boundary conditions are also possible. In the reference configuration, every material point $x \in \Omega$ can be in one of n possible phases e_j, where $P^{\mathrm{p}} = \{e_1, \dots, e_n\}$ is the set of all possible phases. A mapping $c^{\mathrm{p}} : \Omega \to P^{\mathrm{p}}$ is called a phase distribution. We use the superscript $^{\mathrm{p}}$ to indicate that at the present stage the modeling is only concerned with **pure** phase distributions, i.e. phase mixtures are not taken into account.

The material properties of each phase are given by the constitutive relation $\psi = \widehat{\psi}(c^{\mathrm{p}}, \boldsymbol{F})$ where $\boldsymbol{F} = \nabla \boldsymbol{u}$ is the deformation gradient. In martensitic PT one often assumes $\widehat{\psi}(e_j, \boldsymbol{F}) = \widetilde{\psi}(\boldsymbol{F}\boldsymbol{T}_j^{-1}) + \beta_j$ with a fixed energy density $\widetilde{\psi}$, $\boldsymbol{T}_1 = \boldsymbol{I}$ and $\beta_1 = 0$ for the austenite phase, and $\boldsymbol{T}_j$ are the transformation strains for the martensitic phases (e.g. $n = 4$ for the cubic to tetragonal PT and $n = 13$ for the cubic to monoclinic PT, see [BhK96]). The tensors $\boldsymbol{T}_j$ are typically related to each other by the crystallographic symmetry group, and $\beta_j = \beta$ for $j = 2, \dots, n$ where β is positive (negative) if we are below (above) the critical temperature. The whole analysis works for general inhomogeneous materials with $\psi = \widehat{\psi}(c^{\mathrm{p}}, x, \boldsymbol{F})$ but for notational convenience we restrict ourself to the homogeneous case.

A function $c^{\mathrm{p}} : \Omega \to P^{\mathrm{p}}$ is called a state of the system; it gives for each point in the body the associated phase. A family of states $(c^{\mathrm{p}}(t))_{t \in [0,T]}$ is called a **transformation process**. Hence, a state contains only the information on the phase distribution and not on the deformation $\boldsymbol{u}$. This simplification is possible since we will later make the constitutive assumption on the dissipation that it does not depend on $\boldsymbol{F} = \nabla \boldsymbol{u}$. This implies that the deformation $\boldsymbol{u}$ can be computed from the phase distribution c^{p}. Therefore we can oppress the dependence of the Gibbs energy $\mathcal{G}$ on $\boldsymbol{u}$. This simplification is done only to reduce the notational complexity; the influence of the elastic deformation is by no means negligible.

The second law of thermodynamics, in the global form of the Planck inequality for isothermal processes, reads as follows

$$\mathcal{X} := \int_S \boldsymbol{p} \cdot \boldsymbol{v} \, \mathrm{d}S + \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} \, \mathrm{d}x - \int_\Omega \dot{\psi} \, \mathrm{d}x \geq 0. \tag{2.1}$$

Here ψ is the Helmholtz free energy per unit volume, $\boldsymbol{v} = \frac{\partial}{\partial t} \boldsymbol{u}$ is the velocity vector and $\boldsymbol{f}$ is the body force. The important quantity is $\mathcal{X}$ which is the total dissipation rate of the process.

Reversible PT without dissipation can be described with the help of the principle of minimizing the Gibbs free energy. This case is mathematically well–investigated [BaJ87, BaJ92, Bha93, BCJ95, BhK96, Mül98]. In reality martensitic PTs, even in elastic materials like shape memory alloys, are accompanied by dissipation and thus are non–reversible. The dissipation comes from a threshold resistance for the appearance of a new nucleus and from the resistance against the movement of phase boundaries. There are two types of resistance, namely athermal or rate–independent and rate–dependent [OlC86, GhO94]. Athermal resistance is similar to dry friction, i.e. PT can start when the thermomechanical driving force exceeds some dissipative threshold, and this friction is independent of interface velocity. It is caused by the interaction of interfaces with obstacles having a long–range stress fields (point defects, dislocations, grains, twins and subgrain boundaries) which cannot be overcome by thermal fluctuations. The Peierls barrier arising from the discrete crystal lattice and the emission of acoustic waves are other sources of athermal friction. Short–range resistance which can be overcome by thermal activation gives rise to rate–dependent or viscous friction. Limiting ourself to quasi–equilibrium consideration (rate–independence) we have to neglect viscous friction. Accounting for dry friction results in very interesting phenomena, namely in appearance of hysteresis and metastable states (e.g. some finite change of external parameters does not lead to PT in multiphase mixture) and history dependence in the material behavior. In the framework of pure mechanical models the energy loss might not be due to dissipation only, but in addition energy might be transferred into latent heat, which is radiated through the boundary of the body. The dissipative threshold in our time–independent model includes also a thermally–activated part of the driving force above which PT occurs with visible velocity, see e.g. [Lev00].

Thus, it is reasonable to model the loss of energy due to changes in the phase distribution by constitutive assumptions. Consider a given transformation process $(c^{\mathrm{p}}(t))_{t\in[0,T]}$ and a time interval $[t_1, t_2]$, then the total energy loss (over the whole volume) is given by

$$\begin{aligned}\widetilde{X}(c^{\mathrm{p}}; t_1, t_2) := \textstyle\int_{t_1}^{t_2} \mathcal{X}(t)\,\mathrm{d}t = \;& -\textstyle\int_\Omega (\psi(t_2,x)-\psi(t_1,x))\,\mathrm{d}x \\ & + \textstyle\int_{t_1}^{t_2} \Big[\int_S \boldsymbol{p}\cdot\dot{\boldsymbol{u}}\,\mathrm{d}a + \int_\Omega \boldsymbol{f}\cdot\dot{\boldsymbol{u}}\,\mathrm{d}x \Big]\,\mathrm{d}t \geq 0.\end{aligned}$$

Introducing the Gibbs energy

$$\mathcal{G}(t, c^{\mathrm{p}}) := \int_\Omega \psi\,\mathrm{d}x - \int_{S_p} \boldsymbol{p}(t,x)\cdot\boldsymbol{u}(x)\,\mathrm{d}a - \int_\Omega \boldsymbol{f}(t,x)\cdot\boldsymbol{u}(x)\,\mathrm{d}x,$$

the energy loss can be presented in the equivalent form

$$\begin{aligned}\widetilde{X}(c^{\mathrm{p}}; t_1, t_2) \;=\;& -(\mathcal{G}(t_2, c^{\mathrm{p}}(t_2)) - \mathcal{G}(t_1, c^{\mathrm{p}}(t_1))) \\ & - \textstyle\int_{t_1}^{t_2} \Big[\int_{S_p} \dot{\boldsymbol{p}}\cdot\boldsymbol{u}\,\mathrm{d}a - \int_{S_u} \boldsymbol{p}\cdot\dot{\boldsymbol{u}}\,\mathrm{d}a + \int_\Omega \dot{\boldsymbol{f}}\cdot\boldsymbol{u}\,\mathrm{d}x \Big]\,\mathrm{d}t.\end{aligned}$$

The value $\widetilde{X}(c^{\mathrm{p}}, t_1, t_2)$ is called the **calculated** value of dissipated energy along the transformation process $(c^{\mathrm{p}}(t))_{t\in[0,T]}$ during the time interval $[t_1, t_2]$.

Let us consider an idealized case of elastic materials without dissipation due to PT. Clearly, a transformation process from $c^{\mathrm{p}}(t_1)$ to $c^{\mathrm{p}}(t_2) \neq c^{\mathrm{p}}(t_1)$ with $\widetilde{X}(c^{\mathrm{p}}, t_1, t_2) < 0$ is impossible, because it contradicts the second law of thermodynamics (2.1). The condition $\widetilde{X}(c^{\mathrm{p}}, t_1, t_2) = 0$ is the criterion for PT without dissipation; in principle PT is possible (no contradiction to the second law of thermodynamics) and the dissipation is 0.

In order to take thermodynamic hysteresis and athermal components of resistance to PT into account, we introduce the **actual** or **experimentally determined** dissipation due to PT. This will be a constitutive material law whose coefficients can be determined experimentally. Assume that a subvolume $\omega \subset \Omega$ transforms from the phase e_i into phase e_j. Then, this change is associated to a dissipation of energy of the amount $\mathrm{vol}(\omega)\kappa_{i\to j}$. The coefficients $\kappa_{i\to j}$ are considered as experimentally determined material parameters and satisfy $\kappa_{j\to j} = 0$, and $\kappa_{i\to j} > 0$ for $j \neq i$. Of course, the parameters $\kappa_{i\to j}$ might also vary in space. The consideration of inhomogeneous friction is conceptually quite useful in situations, where discontinuities in time might occur. An important example is the appearance of an infinitesimal nucleus which triggers the spontaneous transformation of a macroscopic volume.

The proportionality of the dissipation to the volume of the transformed region can be understood, for instance, by moving interfaces which have a resistance due to a constant distribution of dislocations. Then the volume which is coated by the moving interface is relevant for the dissipation.

Except for the elastic properties of the phases (elastic tensors and transformation strains) the only additional parameters in our model are the threshold values $\kappa_{i\to j}$. However, these values are directly related to the dissipated energy which can be measured in experiments via the area of the hysteresis loop in the stress–strain diagram. Thus, it will be possible to compare the prediction of our model with those from experiments, see [KMS99] and [HaG99] for related results in that direction.

To express this dissipation quantitatively throughout the body we define a function $D^{\mathrm{p}} : P^{\mathrm{p}} \times P^{\mathrm{p}} \to [0, \infty)$ such that $D^{\mathrm{p}}(e_i, e_j) = \kappa_{i\to j}$ for all $i, j \in \{1, \ldots, n\}$. Under the assumption that the transformation process c^{p} is such, that in the time interval $[t_1, t_2]$ the function $c^{\mathrm{p}}(\cdot, x)$ changes at most once (that is, at most one PT takes place in each $x \in \Omega$), the total dissipation in the body Ω is given by

$$\mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}(t_1), c^{\mathrm{p}}(t_2)) = \int_{\Omega} D^{\mathrm{p}}(c^{\mathrm{p}}(t_1, x), c^{\mathrm{p}}(t_2, x))\,\mathrm{d}x \quad \text{where } D^{\mathrm{p}}(e_i, e_j) = \kappa_{i\to j}.$$

If at some material points several phase transformations occur in this time interval then the total amount of dissipated energy is equal to sum of dissipation due to each of these PT, i.e.

$$\mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2) = \lim_{K\to\infty} \sum_{k=1}^{K} \mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}(t_1{+}(k{-}1)\tau_K), c^{\mathrm{p}}(t_1{+}k\tau_K)) \qquad \text{where } \tau_K = (t_2{-}t_1)/K.$$

We derive an evolution law from thermomechanical postulates which is similar in nature to the principle of maximal dissipation in elastoplasticity. For every transformation process $c^{\mathrm{p}}(t)$ which actually occurs the calculated value of dissipated energy $\widetilde{X}(c^{\mathrm{p}}; t_1, t_2)$ has to be equal to experimentally determined value $\mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2)$, otherwise such a transformation process is impossible. We will use the weaker assumption, namely inequality

$$\widetilde{X}(c^{\mathrm{p}}; t_1, t_2) \geq \mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2), \tag{2.2}$$

and derive the equality later. In the case $c^{\mathrm{p}}(t_1) \neq c^{\mathrm{p}}(t_2)$, and consequently $\mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2) > 0$, we refer to (2.2) as the **phase transformation criterion**. Note, however, that the validity of (2.2) does not depend on the positivity of $\mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2)$. We will use it to find a unique evolution law even if the behavior of the material is purely elastic.

The explicit form of (2.2) can be written as

$$\begin{aligned} &-(\mathcal{G}(t_2, c^{\mathrm{p}}(t_2)) - \mathcal{G}(t_1, c^{\mathrm{p}}(t_1))) - \textstyle\int_{t_1}^{t_2} \Big[\int_{S_{\boldsymbol{p}}} \dot{\boldsymbol{p}} \cdot \boldsymbol{u}\,\mathrm{d}a - \int_{S_{\boldsymbol{u}}} \boldsymbol{p} \cdot \dot{\boldsymbol{u}}\,\mathrm{d}a + \int_{\Omega} \dot{\boldsymbol{f}} \cdot \boldsymbol{u}\,\mathrm{d}x \Big]\,\mathrm{d}t \\ &\geq \mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2). \end{aligned} \tag{2.3}$$

The derivation of this inequality relies purely on energetic considerations and does not require any modeling assumption. For this reason inequality (2.3) is denoted as **energy inequality**.

The energy inequality (2.3) is only one scalar equation which is not sufficient for the determination of the transformation process. In the rate–independent limit the principle of minimal Gibbs free energy cannot be applied for dissipative materials, e.g. for elastoplastic or elastic with dissipation due to PT. For this purpose a new thermodynamic assumption named the **postulate of realizability** was formulated and applied in [Lev95, Lev97, Lev98, Lev00]. The essence of the postulate of realizability is very simple: as soon as some dissipative process (plastic flow, PT) can occur from the viewpoint of thermodynamics, it will occur (or be realized), i.e. the first fulfillment of the necessary energetic condition is sufficient for the beginning of a dissipative process. This assumption can be considered as the concept of stability. In fact, if PT or some other dissipative process can occur from an energetic point of view, but does not occur, such a situation is not stable. Various fluctuations provoke the beginning of a dissipative process and when all the equations are satisfied, the dissipative process will occur.

To formalize this idea we assume that every pair of states c^{p}, b^{p} can be connected via a virtual process $(\widetilde{c^{\mathrm{p}}}(s))_{s\in[t,t+\tau]}$ so that the dissipation $\mathrm{Diss}^{\mathrm{p}}(\widetilde{c^{\mathrm{p}}};t,t{+}\tau)$ is minimal, i.e. $\widetilde{c^{\mathrm{p}}}(t)=c^{\mathrm{p}}$, $\widetilde{c^{\mathrm{p}}}(t{+}\tau)=b^{\mathrm{p}}$ and $\mathrm{Diss}^{\mathrm{p}}(\widetilde{c^{\mathrm{p}}};t,t{+}\tau)=\mathcal{D}^{\mathrm{p}}(\widetilde{c^{\mathrm{p}}}(t),\widetilde{c^{\mathrm{p}}}(t{+}\tau))$. Virtual processes are not required to satisfy the energy inequality (2.3). We emphasize that until now no assumption, that the external forcing, which is given by the functions $\boldsymbol{f}$ and $\boldsymbol{p}$, is independent of time, was made.

The postulate of realizability now states that for every observable state c^{p} and every comparison state $b^{\mathrm{p}}\neq c^{\mathrm{p}}$ there exists a time increment Δt and the corresponding load increment so that for each $\tau\leq\Delta t$ and every virtual process $(\widetilde{c^{\mathrm{p}}}(s))_{s\in[t,t+\tau]}$ connecting c^{p} to b^{p} violates the energy inequality, i.e.

$$\widetilde{X}(\widetilde{c^{\mathrm{p}}};t,t{+}\tau)<\mathrm{Diss}^{\mathrm{p}}(\widetilde{c^{\mathrm{p}}};t,t{+}\tau). \tag{2.4}$$

The mechanical interpretation of this statement is as follows. Starting with an observable state c^{p} we choose some finite virtual increment of all interfaces which includes appearance of new nuclei, i.e. a comparison state b^{p}. Then we argue that we can find such an infinitesimal time increment Δt and the corresponding load increment, that we cannot find any transformation process $(\widetilde{c^{\mathrm{p}}}(s))_{s\in[t,t+\Delta t]}$ connecting c^{p} to b^{p} which satisfies the energy inequality. The increment Δt and corresponding increment of external conditions depend on the state b^{p} which was previously selected.

This statement expresses mathematically our concept of stability. In fact, if for some state b^{p} which differs from c^{p} we cannot find the infinitesimal time increment Δt and corresponding infinitesimal increment of external conditions, for which inequality (2.4) is valid, then for every infinitesimal $\tau>0$ there exists a virtual process connecting state c^{p} to b^{p} which satisfies the PT criterion. Therefore the transformation from the state c^{p} to b^{p} can occur, and according to the postulate of realizability, will occur. Since τ is arbitrarily small, the transformation of the state c^{p} into state b^{p} happens instantaneously, i.e. at fixed external conditions. Consequently, the state $c^{\mathrm{p}}(t)$ is unstable.

Since the time increment τ can be arbitrarily small, inequality (2.4) is local in time (for fixed t let τ tend to 0). The external loading changes only very slowly in time and in particular does not jump. Hence, the work integrals in (2.4) can be made as small as we like; and we end up with the local version

$$-(\mathcal{G}(t,b^{\mathrm{p}})-\mathcal{G}(t,c^{\mathrm{p}}(t)))<\mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}(t),b^{\mathrm{p}})\quad\text{ for } b^{\mathrm{p}}\neq c^{\mathrm{p}}, \tag{2.5}$$

where $c^{\mathrm{p}}(t)$ is the value of the thermodynamically admissible process while b^{p} is an arbitrary comparison function. We call the state $c^{\mathrm{p}}(t)$ **strictly stable** if (2.5) holds, we we ask for strict inequality whenever $b^{\mathrm{p}}\neq c^{\mathrm{p}}(t)$.

The postulate of realizability implies that the energy inequality (2.3) is in fact an identity,

$$
\begin{aligned}
&-(\mathcal{G}(t_2, c^{\mathrm{p}}(t_2)) - \mathcal{G}(t_1, c^{\mathrm{p}}(t_1))) - \int_{t_1}^{t_2} \left[\int_{S_{\boldsymbol{p}}} \dot{\boldsymbol{p}} \cdot \boldsymbol{u} \, \mathrm{d}a - \int_{S_{\boldsymbol{u}}} \boldsymbol{p} \cdot \dot{\boldsymbol{u}} \, \mathrm{d}a + \int_{\Omega} \dot{\boldsymbol{f}} \cdot \boldsymbol{u} \, \mathrm{d}x \right] \mathrm{d}t \\
&= \mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2).
\end{aligned} \tag{2.6}
$$

A proof of (2.6) is provided by Theorem 3.3 for a special case, which can be generalized easily. Note that the stability criterion (2.5) is given by a scalar equation. However, it involves all possible comparison functions $b^{\mathrm{p}} : \Omega \to P^{\mathrm{p}}$, thus it is strong enough to generate an evolution law for the transformation process $(c^{\mathrm{p}}(t))_{t\in[0,T]}$. We now summarize the mechanical conditions which determine the admissible transformation processes:

Given the initial state $c_0^{\mathrm{p}} = c^{\mathrm{p}}(0)$ we have to find a transformation process $c^{\mathrm{p}}(t), t \in [0, T]$, which satisfies the energy inequality (2.2) (or the identity (2.6)) and which is strictly stable in the sense of (2.5) for all $t \in (0, T]$.

A solution of this problem displays PT in a time interval $[t_1, t_2]$ only if the PT criterion $\mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2) > 0$ is satisfied. For general loadings we have to allow for solutions which do not lead to any PT (i.e. $\mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; 0, T) = 0$), for instance in the case of small loading where the hysteresis loop is not even entered. For elastic materials without dissipation we can use the same problem formulation with $\mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2) = 0$. However, as the final state is history independent and is determined only by the final load, we can find the solution for any given external condition with the help of inequality (2.5) if $\mathcal{D}^{\mathrm{p}}(a^{\mathrm{p}}, b^{\mathrm{p}}) = 0$. In this case, inequality (2.5) is nothing but the principle of minimal Gibbs free energy.

To formulate a mathematical version of the problem of finding admissible processes we have to drop the notation of strict stability, since our approximation techniques are such that in the limit we are not able to conclude strict stability. Throughout this paper this problem is called the time–continuous problem (CP) and it is concerned with pure phases only.

> ***(CP)*** *For given c_0^{p}, find a transformation process $c^{\mathrm{p}} : [0, T] \times \Omega \to P^{\mathrm{p}}$ with $c^{\mathrm{p}}(0, x) = c_0^{\mathrm{p}}(x)$ such that for all t_1, t_2 with $0 \le t_1 < t_2 \le T$ we have the two conditions*
>
> *(Energy inequality)* $\mathcal{G}(t_2, c^{\mathrm{p}}(t_2)) + \mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2) \le \mathcal{G}(t_1, c^{\mathrm{p}}(t_1))$*+work of external forces;*
>
> *(Stability)* $\mathcal{G}(t_2, c^{\mathrm{p}}(t_2)) \le \mathcal{G}(t_2, b^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}(t_2), b^{\mathrm{p}})$ *for all* $b^{\mathrm{p}} : \Omega \to P^{\mathrm{p}}$.

Note that the nonstrict stability criterion can produce nonuniqueness in certain situations. For instance, on a horizontal part of a hysteresis loop every point is stable but only one of the end points is strictly stable. Nevertheless, strict stability is not enough to enforce uniqueness in all situations.

We now explain some of the results in mechanical notations. First we emphasize that the given problem (CP) has in general no solutions. This is due to the fact that the system prefers to develop microstructure rather than staying in pure phases, (i.e. $c^{\mathrm{p}}(t, x) \in P^{\mathrm{p}}$). That is the reason why we have to relax the problem by allowing for a larger set of states including phase mixtures in each material point, namely $c(t, x) \in P = \mathrm{conv}(P^{\mathrm{p}})$. The major task is to extend the constitutive functions $\psi_{e_j}(\boldsymbol{F})$ and D^{p} consistently. By the relaxation procedure we find that a unique extension of $\widehat{\psi} : P^{\mathrm{p}} \times \mathbb{R}^{3\times 3} \to \mathbb{R}$ is given by the mixture function $\mathbb{W} : P \times \mathbb{R}^{3\times 3} \to \mathbb{R}$ and the function D^{p} extends to $D(a,c) = \Delta(c-a)$ for a suitable convex function Δ which is homogeneous of degree one. The important point here is that we obtain a mixture theory without further mechanical assumptions.

Second, we show that a natural time discretization of rate–independent problems of the type (CP) is obtained by the following time–discrete incremental problem.

(IP) *For given initial datum c_0^{p} and time instants $0 = t_0 < t_1 < \ldots < t_N = T$ find $(c_k^{\mathrm{p}})_{k=1,\ldots,N}$ from the minimization problem*

$$c_k^{\mathrm{p}} \in \arg\min_{b^{\mathrm{p}}} \Big[\mathcal{G}(t_k, b^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c_{k-1}^{\mathrm{p}}, b^{\mathrm{p}}) \Big].$$

Note that the Gibbs energy is obtained from the elastic energy and the external loadings via minimization with respect to the deformation. Hence, this minimization problem really involves $\boldsymbol{u} : \Omega \to \mathbb{R}^3$ as well as $c^{\mathrm{p}} : \Omega \to P^{\mathrm{p}}$.

Third, we mention that the structure of the minimization problem in (IP) is very special. We can write

$$\mathcal{G}(t_k, b) + \mathcal{D}(c_{k-1}, b) = \int_\Omega \Big[f_k(x, b(x), \nabla \boldsymbol{u}(x)) - \boldsymbol{f} \cdot \boldsymbol{u} \Big] \,\mathrm{d}x - \int_{S_{\boldsymbol{p}}} \boldsymbol{p} \cdot \boldsymbol{u} \,\mathrm{d}a,$$

where $f_k(x, c, \boldsymbol{F}) = \mathbb{W}(c, \boldsymbol{F}) + \Delta(c - c_{k-1}(x))$ depends only locally on $b \in P$. Thus, it is possible to minimize with respect to $b \in P$ pointwise under the integral leading to the reduced density $f_k^{\mathrm{red}}(x, \boldsymbol{F}) = \min_{c \in P} f_k(x, c, \boldsymbol{F})$ and the reduced functional

$$\mathcal{F}_k^{\mathrm{red}}(\boldsymbol{u}) = \int_\Omega \Big[f_k^{\mathrm{red}}(x, \nabla \boldsymbol{u}(x)) - \boldsymbol{f} \cdot \boldsymbol{u} \Big] \,\mathrm{d}x - \int_{S_{\boldsymbol{p}}} \boldsymbol{p} \cdot \boldsymbol{u} \,\mathrm{d}a.$$

This is a minimization problem of the same type as classical nonlinear elasticity problems.

Fourth, we mention that the above procedure leads in a natural way to flow rules for PT problems, see the local flow formulation in (5.20). In [GoM00] it is shown that these flow rules can predict the evolution of the different phase fractions qualitatively very well.

3 Rate–independent processes

Our existence result for the final relaxed problem relies on an abstract theory for rate–independent hysteretic processes which are defined via an energy functional $\mathcal{I}(t, \cdot)$ and a dissipation functional $\boldsymbol{\Delta}$. Moreover this theory, which is further developed in [MiT00], plays a central role in the motivation of our relaxation procedure which is based on the incremental time–stepping algorithm.

Our approach to hysteresis is quite different from previous ones which involve differential inclusions and monotone operator theory, see e.g. [Vis94, BrS96]. It seems that these methods are not easily adjustable to the special features on our model. First, we do not assume that $\mathcal{I}(t, \cdot)$ is convex, or equivalently that $\mathrm{D}\mathcal{I}(t, \cdot)$ is a (set–valued) monotone operator. Second, the restriction of $c \in X$ to the set $\mathcal{P}$, which has a nontrivial boundary, together with the nonsmooth function $\boldsymbol{\Delta}$ leads to a doubly nonlinear system in the sense of [CoV90], where related problems are studied under the assumption of convexity of $\mathcal{I}(t, \cdot)$ and X being a Hilbert space. The new structure in our approach is the energetic viewpoint which gives us more flexibility than classical approaches using differential inclusions.

3.1 The functional analytic setup

Let X be a separable Banach space with dual space X^* and let $\mathcal{P}$ be a weakly closed subset of X. For every time $t \in [0,T]$ where T is a fixed positive number and every $c \in \mathcal{P}$ the energy of the state c is given by the **energy functional** $\mathcal{I}(t,c) \in \mathbb{R}_{\geq}$. For $\mathcal{I}$ we assume the following properties:

$$\mathcal{I} \in \mathrm{C}^0([0,T] \times \mathcal{P}), \qquad \sup_{c\in\mathcal{P}, t\in[0,T]} |\partial_t \mathcal{I}(t,c)| < \infty, \tag{3.1}$$

for every $t \in [0,T]$ the mapping $c \mapsto \mathcal{I}(t,c)$ is weakly lower semicontinuous. (3.2)

In the PT problem the restriction of the phase space to $\mathcal{P} \subset X$ accounts for the fact that the variable c measures the concentration of phases, i.e. $c(x) \in P \subset \mathbb{R}^n$. This restriction is by far more than a technical modification of the usual setting e.g. in the theory of plasticity where $\mathcal{P} = X$. If $\mathcal{P}$ is bounded, the set of tangential direction may become abruptly smaller as the state c reaches the boundary of $\mathcal{P}$. As a positive effect, the set $\mathcal{P}$ may be weakly compact although X is not reflexive, this simplifies central steps of the analysis.

Most of our work will take place in spaces of bounded variation in time,

$$\mathrm{BV}([0,T],\mathcal{P}) = \left\{ f : [0,T] \to \mathcal{P} \;\middle|\; \sup_{0=t_0<\ldots<t_N=T} \sum_{k=0}^{N-1} \|f(t_{k+1}) - f(t_k)\| < \infty \right\}.$$

Observe that our definition of BV–functions does not neglect sets of measure 0, therefore $\mathrm{BV}([0,T],X)$ is larger than $\mathrm{C}^0([0,T],X^*)^*$, even if X is reflexive or finite dimensional.

The dissipation generated by c changing from one state to a different state is given by a **dissipation functional** $\boldsymbol{\Delta} : X \to \mathbb{R}_{\geq}$ which is convex and homogeneous of degree 1. We assume that there exists $C > 0$ so that

$$\frac{1}{C}\|v\| \leq \boldsymbol{\Delta}(v) \leq C\|v\|. \tag{3.3}$$

It will become clear later that in our model the solutions are rate independent if and only if the $\boldsymbol{\Delta}$ is homogeneous of degree one. An immediate consequence of the convexity and the homogeneity of $\boldsymbol{\Delta}$ is the triangle inequality

$$\boldsymbol{\Delta}(v_1 + v_2) \leq \boldsymbol{\Delta}(v_1) + \boldsymbol{\Delta}(v_2) \ \text{ for every } v_1, v_2 \in X. \tag{3.4}$$

A state c is said to be **stable** at time t if

$$\mathcal{I}(t,c) \leq \mathcal{I}(t,a) + \boldsymbol{\Delta}(a - c) \ \text{ for every } a \in \mathcal{P} \tag{2.S}$$

holds. By $\mathrm{Var}_{\boldsymbol{\Delta}}(f)$ we denote the variation of a BV–function f with respect to the functional $\boldsymbol{\Delta}$:

$$\mathrm{Var}_{\boldsymbol{\Delta}}(f) = \sup_{0=t_0<\ldots<t_N=T} \sum_{k=0}^{N-1} \boldsymbol{\Delta}(f(t_{k+1}) - f(t_k)) = \int_0^T \boldsymbol{\Delta}(\mathrm{d}f)\,.$$

A process $c \in \mathrm{BV}([0,T],\mathcal{P})$ satisfies the **weak energy inequality** if

$$\mathcal{I}(T, c(T)) + \mathrm{Var}_{\boldsymbol{\Delta}}(c) \leq \mathcal{I}(0, c(0)) + \int_0^T \partial_t \mathcal{I}(t, c(t))\,\mathrm{d}t \tag{2.E}$$

holds. Under sufficient smoothness the weak thermodynamic PT criterion is nothing else but a integrated version of the local thermodynamic PT criterion

$$\mathrm{D}_c\mathcal{I}(t,c)[\dot{c}] + \boldsymbol{\Delta}(\dot{c}) \leq 0.$$

The weak formulation requires no differentiability from $\mathcal{I}(t,\cdot)$ and $c(\cdot)$. The central problem in this work is the construction of processes which satisfy both (2.S) and (2.E).

Definition 3.1 *For every $t \in [0,T]$ we denote by $\mathcal{S}(t) \subset \mathcal{P}$ the **set of stable points**,*

$$\mathcal{S}(t) = \{\, c \in \mathcal{P} \,\big|\, \mathcal{I}(t,c) \leq \mathcal{I}(t,a) + \boldsymbol{\Delta}(a-c) \;\forall a \in \mathcal{P} \,\}.$$

*A process $c \in \mathrm{BV}([0,T],\mathcal{P})$ is called **stable** if $c(t) \in \mathcal{S}(t)$ for every $t \in (0,T]$. Stable processes which satisfy (2.E) are called **admissible**.*

We note that the stable sets $\mathcal{S}(t)$ are usually not convex even when $\mathcal{I}(t,\cdot)$ is strictly convex (unless $\mathcal{I}(t,\cdot)$ is quadratic). Using the (strong) continuity of $\mathcal{I}$ and $\boldsymbol{\Delta}$ we obtain the (strong) closedness of the set of stable states. However, for our application we will need also weak closedness which is more difficult to obtain.

When taking limits with respect to time we have to be careful about temporal discontinuities. By general theorems on BV–functions we know that a BV–function c has at most a countable number of discontinuities and that for each $t \in [0,T]$ the limit from the right and from the left exists. For c we define its projection c^- into the space of left–continuous BV–functions via $c^-(0) = c(0)$ and $c^-(t) = \lim_{s\nearrow t} c(s)$ for $t \in (0,T]$ and find $\|c^-\|_{\mathrm{BV}([0,T],X)} \leq \|c\|_{\mathrm{BV}([0,T],X)}$.

Lemma 3.2 *Under the assumptions (3.1) and (3.3) the stable set $\mathcal{S}(t)$ is a closed subset of $\mathcal{P}$ for all $t \in [0,T]$. Moreover, $\cup_{t\in[0,T]}(t,\mathcal{S}(t))$ is a closed subset of $[0,T]\times\mathcal{P}$.*

If a process $c \in \mathrm{BV}([0,T],\mathcal{P})$ is stable for all s in a dense subset of $[0,T]$, then c^- is a stable process, where $c^-(t) = \lim_{s\nearrow t} c(s)$ for $t \in (0,T]$.

Proof. We prove the second statement since the first is a consequence of the second. Assume $c^k \in \mathcal{S}(t_k)$ with $c^k \to c^*$ and $t_k \to t_*$. By the continuity of $\mathcal{I}$ and $\boldsymbol{\Delta}$ we have, for arbitrary $a \in \mathcal{P}$,

$$\mathcal{I}(t_*,c^*) = \lim_{k\to\infty} \mathcal{I}(t_k,c^k) \leq \liminf_{k\to\infty}(\mathcal{I}(t_k,a) + \boldsymbol{\Delta}(a-c^k)) = \mathcal{I}(t,a) + \boldsymbol{\Delta}(a-c^k).$$

To prove the third statement let $\mathcal{T}$ be the dense subset with $c(t) \in \mathcal{S}(t)$ for $t \in \mathcal{T}$. For $t \in (0,T]$ choose a sequence $(t_k)_{k\in\mathbb{N}}$ with $t_k \nearrow t$. Then, $c^-(t) = \lim_{k\to\infty} c(t_k)$ and the result follows from the second statement. ∎

In the next three subsections we give an overview of the properties of admissible processes, for more details we refer to [MiT00].

3.2 Qualitative properties of admissible processes

At the jumps a process can take rather arbitrary values without violating (2.S) or (2.E). To obtain definiteness we restrict ourselves to functions which are continuous from the left, i.e., $c = c^-$. Assumption (3.3) gives a bound on the variation of c. Therefore the space $\mathrm{BV}([0,T],\mathcal{P})$ is the natural space to work within. From (2.S) and (2.E) can be inferred that an energy identity holds; not only for the whole time interval, but also on every subinterval. This is the direct analogue of equation 2.6.

Theorem 3.3 *Assume that the dissipation functional $\mathbf{\Delta}$ satisfies the coercivity assumption (3.3) for some positive constant C and that (3.1) holds. If c is an admissible process with $c = c^-$, then the local energy equality*

$$\mathcal{I}(t, c(t)) + \int_s^t \mathbf{\Delta}(\mathrm{d}c) = \mathcal{I}(s, c(s)) + \int_s^t \tfrac{\partial \mathcal{I}}{\partial t}(\tau, c(\tau))\,\mathrm{d}\tau \tag{3.5}$$

holds for every $0 \le s \le t \le T$.

Proof. We set for $t_1, t_2 \in [0, T]$ the function $\mathcal{J}(t_1, t_2) = \mathcal{I}(t_1, c(t_2))$. Now for every $0 \le s \le t \le T$ one has that

$$\mathcal{I}(s, c(s)) - \mathcal{I}(t, c(t)) + \int_s^t \{\partial_t \mathcal{I}(\tau, c(\tau)) - \mathbf{\Delta}(\dot{c}(\tau))\}\,\mathrm{d}\tau = -\int_s^t \{\partial_{t_2} \mathcal{J}(\tau, \tau) + \mathbf{\Delta}(\dot{c}(\tau))\}\,\mathrm{d}\tau.$$

By (2.E) the left–hand side is nonnegative if we let $s = 0$ and $t = T$, therefore the claim follows if we can show that the integrand on the right–hand side is nonnegative in $C([0,T])^*$. The stability inequality (2.S) implies, for every $t \in [0, T)$ and $h \in (0, T-t)$,

$$\tfrac{1}{h}[\mathcal{I}(t, c(t+h)) - \mathcal{I}(t, c(t))] + \mathbf{\Delta}(\tfrac{1}{h}[c(t+h) - c(t)]) \ge 0,$$

and this expression converges weak* to $t \mapsto \partial_{t_2} \mathcal{J}(t,t) + \mathbf{\Delta}(\dot{c}(t))$ as $h \to 0$. Since the set of positive functionals is weak* closed, the proof is finished. ■

3.3 Existence of admissible processes

So far we have only discussed the qualitative properties of admissible processes without knowing whether they exist. The typical way to achieve existence is to discretize (2.S) and (2.E) fully implicitly in time. This gives a variational problem and in certain cases it is possible to show that that solutions to the discretized problem converge to an admissible process as the fineness of the time discretization tends to 0. It is surprising that the discretized problems seem to have more structure than the original system (2.E), (2.S) although no assumption on the differentiability of the potential $\mathcal{I}$ and the dissipation functional $\mathbf{\Delta}$ is made. Major parts of the work in the following chapters are inspired by the time discretized problem. Replacing $\dot{c}$, $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}(t, c(t))$ and $\frac{\partial}{\partial t}\mathcal{I}(t, c)$ by the appropriate finite differences, one arrives immediately at the incremental problem

> ***(IP)*** *For given $c_0 \in \mathcal{P}$ and a time discretization $0 = t_0 < \ldots < t_N = T$, find for $k = 1 \ldots N$ states $c_k^N \in \mathcal{P}$ such that $\mathcal{I}(t_k, c_k^N) + \mathbf{\Delta}(c_k^N - c_{k-1}^N)$ is minimal.*

To guarantee the existence of solutions to (IP) we need some compactness:

$$\textit{for all } r > 0 \textit{ the intersection } \mathcal{P} \cap \{\, c \in X \;\big|\; \|c\| \le r \,\} \textit{ is weakly compact in } X. \tag{3.6}$$

Defining $c^N : [0, T] \to \mathcal{P}$ as the left–continuous piecewise constant function satisfying $c^N(t) = c_k^N$ for $t \in (t_{k-1}, t_k]$ and $c^N(0) = c_0$ we obtain the following existence result.

Theorem 3.4 *If the assumptions (3.1) to (3.6) as well as (3.3) hold, then (IP) has a solution and for every solution to (IP) the following statements hold true:*

1. *For every $t \in \{t_1, \ldots, t_N = T\}$ the state $c^N(t)$ is stable,*

2. c^N satisfies the discrete energy inequality

$$\mathcal{I}(T, c^N(T)) + \mathrm{Var}_{\boldsymbol{\Delta}}(c^N) \leq \mathcal{I}(0, c_0) + \int_0^T \partial_t \mathcal{I}(t, c^N(t))\,\mathrm{d}t. \tag{3.7}$$

Proof. First we observe that the weak lower semicontinuity of $\mathcal{I}(t_k, \cdot)$ and the convexity of $\boldsymbol{\Delta}(\cdot - c_{k-1}^N)$ implies that the sum is weakly lower semicontinuous. The positivity of $\mathcal{I}$ implies coercivity such that for every $c \in \mathcal{P}$ with $\|c - c_k^N\| > C \cdot \mathcal{I}(t_{k+1}, c_k^N)$ we have that

$$\mathcal{I}(t_{k+1}, c_k^N) \leq \mathcal{I}(t_k, c_k^N) + \mathcal{I}(t_{k+1}, c) < \mathcal{I}(t_{k+1}, c) + \boldsymbol{\Delta}(c - c_k^N),$$

therefore $\|c_{k+1}^N - c_k^N\| \leq C \cdot \mathcal{I}(t_{k+1}, c_k^N)$ holds. The weak compactness (3.6) implies now the existence of solutions to (IP).

The stability of $c^N(t_k) = c_k^N$ arises from its minimizing property. We have that

$$\begin{aligned} \mathcal{I}(t_k, a) + \boldsymbol{\Delta}(a - c_k^N) &= \mathcal{I}(t_k, a) + \boldsymbol{\Delta}(a - c_{k-1}^N) + \boldsymbol{\Delta}(a - c_k^N) - \boldsymbol{\Delta}(a - c_{k-1}^N) \\ &\geq \mathcal{I}(t_k, c_k^N) + \boldsymbol{\Delta}(c_k^N - c_{k-1}^N) + \boldsymbol{\Delta}(a - c_k^N) - \boldsymbol{\Delta}(a - c_{k-1}^N) \geq \mathcal{I}(t_k, c_k^N), \end{aligned}$$

where the triangle inequality (3.4) has been used for the last inequality. To obtain the discrete energy inequality we first observe that (3.7) is equivalent to

$$\sum_{k=1}^N \Big(\mathcal{I}(t_k, c^N(t_k)) - \mathcal{I}(t_k, c^N(t_{k-1})) + \boldsymbol{\Delta}(c^N(t_k) - c^N(t_{k-1})) \Big) \geq 0. \tag{3.8}$$

From the fact that c_k^N minimizes $\mathcal{I}(t_k, \cdot) + \boldsymbol{\Delta}(\cdot - c_{k-1}^N)$ it follows that each term in (3.8) is positive and hence (3.7) holds. ■

It remains to discuss the question under what condition the functions $c^N \in \mathrm{BV}([0,T], \mathcal{P})$ converge to a limit and whether this limit is an admissible process. The essential tool is a version of Helly's theorem for functions which take values in Banach spaces.

Theorem 3.5 ([BaP86], Thm. 3.5, Ch. 1) *Let X be a reflexive separable Banach space with separable dual X^*. Let $\{w_n\} \subset \mathrm{BV}([a,b]; X)$ be such that $\|w_n(t)\| \leq C$ for $t \in [a,b]$ and $\mathrm{Var}(w_n; [a,b]) \leq C$ for all n. Then there exists a subsequence $\{w_{n_k}\} \subset \{w_n\}$ and a function $w \in \mathrm{BV}([a,b]; X)$ such that, as $k \to \infty$, $w_{n_k} \rightharpoonup w(t)$ in X for all $t \in [a,b]$.*

Since in our application the Banach space equals $\mathrm{L}^1(\Omega, \mathbb{R}^n)$ we need a version without reflexivity of X and separability of X^*. Here we take advantage of the subset $\mathcal{P}$ and property (3.6). We have the following simple generalization.

Corollary 3.6 *Let X be a separable Banach space and $\mathcal{P}$ a weakly closed subset satisfying (3.6). Then for each sequence $(c^N)_{N\in\mathbb{N}}$ in $\mathrm{BV}([0,T], X)$ with $\sup_{N\in\mathbb{N}}(\|c^N(0)\| + \|c^N\|_{\mathrm{BV}([0,T],X)}) < \infty$ there exists a subsequence N_k and a limit function $c^\infty \in \mathrm{BV}([0,T], \mathcal{P})$ such that for all $t \in [0,T]$ we have $c^{N_k}(t) \rightharpoonup c^\infty(t)$ as $k \to \infty$.*

Clearly the existence theorem 3.4 provides a sequence (c^N) which is bounded in $\mathrm{BV}([0,T], X)$ and thus we can extract a limit $c^\infty \in \mathrm{BV}([0,T], \mathcal{P})$. From (3.7) it is not so difficult to show that c^∞ satisfies the weak energy inequality, hence it remains to check whether the process c^∞ is stable. There are two ways to derive stability. If the convergence is not only weak, but strong, then the continuity of $\boldsymbol{\Delta}$ suffices to prove stability since $\mathcal{S}(t)$ is strongly closed. If on the other hand the set of stable states is weakly closed, then the stability of the weak limit follows. Note however, that important cases exist where neither approach is successful. We believe however, that both approaches have the potential to be extended to more complicated cases, see [MiT00].

Theorem 3.7 *Let the assumptions of Theorem 3.4 be satisfied. Assume further that*

$$\text{for all } t \in [0,T] \text{ the mapping } c \mapsto \partial_t \mathcal{I}(t,c) \text{ is weakly continuous} \tag{3.9}$$

and that for all $t \in [0,T]$ the stable sets $\mathcal{S}(t)$ are weakly closed. Consider a sequence of hierarchical discretizations (specified in the proof) and let c^∞ be the limit as constructed above. Then, $c = (c^\infty)^- \in \mathrm{BV}([0,T],\mathcal{P})$ is an admissible process with $c(0) = c_0$.

Proof. We have to show that the process c^∞ which is constructed in Theorem 3.4 satisfies the energy inequality and is stable. By the weak lower semicontinuity of $\mathcal{I}(t,\cdot)$ and $\boldsymbol{\Delta}$ and the weak continuity $\partial_t\mathcal{I}(t,\cdot)$ (cf. (3.9)) we find that the discrete energy inequality (3.7) is stable under pointwise weak convergence, hence c^∞ satisfies the weak energy inequality (2.E). To show that $c = (c^\infty)^-$ also satisfies (2.E), we first observe that the boundedness of the variation of c^∞ implies that the set $\{\, t \in [0,T] \,\big|\, c(t) \neq c^\infty(t) \,\}$ is at most countable, thus $\int_0^T \partial_t \mathcal{I}(t,c(t))\,\mathrm{d}t = \int_0^T \partial_t\mathcal{I}(t,c^\infty(t))\,\mathrm{d}t$. Together with $c(0) = c^\infty(0) = c_0$ and $\mathrm{Var}_{\boldsymbol{\Delta}}(c) \leq \mathrm{Var}_{\boldsymbol{\Delta}}(c^\infty)$, this gives the result.

Now consider a sequence $(\Pi_r)_{r\in\mathbb{N}}$ of partitions with $\Pi_r = \{t_0^{N_r}, t_1^{N_r}, \ldots, t_{N_r-1}^{N_r}, t_{N_r}^{N_r}\}$ and $t_k^{N_r} < t_{k+1}^{N_r}$. This sequence is called hierarchical if $\{0,T\} \subset \Pi_r \subset \Pi_{r+1}$ and if the fineness $\Phi(\Pi_r) = \max\{\, t_k^{N_r} - t_{k-1}^{N_r} \,\big|\, k = 1,\ldots,N_r \,\}$ of Π_r tends to 0. Let $\mathcal{T} = \cup_{r\in\mathbb{N}}\Pi_r$, then $\mathcal{T}$ is dense in $[0,T]$ and for each $t \in \mathcal{T}$ we have $c^{N_r}(t) \in \mathcal{S}(t)$ for all sufficiently large r. By the weak convergence to $c^\infty(t)$ and the weak closedness of $\mathcal{S}(t)$ we obtain $c^\infty(t) \in \mathcal{S}(t)$ for all $t \in \mathcal{T}$. Using Lemma 3.2 we conclude that $(c^\infty)^-$ is a stable process. ∎

Remark 3.8 *For $\mathcal{I}(t,c) = \widetilde{\mathcal{I}}(c) - \langle G(t), c\rangle$ with $G \in \mathrm{C}^1([0,T],X^*)$ assumption (3.9) is satisfied .*

3.4 Implications of regularity of $\mathcal{I}$

If the energy functional $\mathcal{I}$ is differentiable, the global extremum principle implies a local extremum principle. If $\mathcal{I}(t,\cdot)$ is convex then the local and the global principle will be equivalent. In this case we are able to write the problem as a doubly nonlinear differential inclusion in the sense of [CoV90]. The local formulation of the problem is very useful for the derivation of qualitative properties of the solutions.

We denote by $\sigma(t,c) = \mathrm{D}_c\mathcal{I}(t,c)$ the gradient of $\mathcal{I}(t,\cdot)$ at c. Moreover, we assume that $\mathcal{P}$ is a closed convex subset of X. The set $\mathrm{TC}_c\mathcal{P}$ is the interior tangent cone of $\mathcal{P}$ at the state c, i.e.,

$$\mathrm{TC}_c\mathcal{P} = \overline{\{\, v \in X \,\big|\, \exists \lambda > 0 :\; c + \lambda v \in \mathcal{P} \,\}}.$$

The following result can be interpreted in the line with the thermomechanical extremum principle formulated in [Lev97]: Within all possible directions v a process can evolve from the given state $c(t)$ the direction $\dot{c}(t)$ is chosen such that it minimizes the dissipation plus the energy increment. See also [Zie63] and the references there.

Proposition 3.9 *Let $\mathcal{I}(t,\cdot)$ be differentiable and c be an admissible process. Then we have the local extremum principle*

$$\langle \sigma(t,c(t)), v\rangle + \boldsymbol{\Delta}(v) \geq 0 \quad \text{for all } v \in \mathrm{TC}_{c(t)}\mathcal{P}, \tag{3.10}$$

$$\langle \sigma(t,c(t)), \dot{c}\rangle + \boldsymbol{\Delta}(\dot{c}) = 0 \quad \text{whenever } \dot{c}(t) = \lim_{h\nearrow 0}\frac{1}{h}(c(t+h) - c(t))\,\text{exists.} \tag{3.11}$$

If $\mathcal{I}(t,\cdot)$ is convex, the local system (3.10, 3.11) is satisfied if and only if c is admissible.

Proof. It is clear that (3.11) is equivalent to the local energy equality (3.5). By inserting $a = c+\theta\lambda v$ with $\theta \in (0,1]$ in the stability definition (2.S), dividing by θ and taking the limit $\theta \to 0$ we obtain (3.10).

On the other hand if $\mathcal{I}(t,\cdot)$ is convex, we have for every $a \in \mathcal{P}$ that

$$\mathcal{I}(t,a) - \mathcal{I}(t,c) \geq \langle \sigma(t,c), a-c \rangle \geq -\mathbf{\Delta}(a-c)$$

by (3.10), hence the stability inequality (2.S) holds. ∎

Definition 3.10 *By $\mathbf{\Delta}^* : X^* \to \mathbb{R} \cup \{\infty\}$ we denote the conjugate function or the Legendre–Fenchel transform of $\mathbf{\Delta}$, i.e.,*

$$\mathbf{\Delta}^*(\sigma) = (\mathcal{L}\mathbf{\Delta})(\sigma) = \sup\{\, \langle \sigma, v \rangle - \mathbf{\Delta}(v) \,\big|\, v \in X \,\}.$$

For a convex function $\mathcal{J} : X \to \mathbb{R} \cup \{\infty\}$ the subdifferential at c is the set

$$\partial \mathcal{J}(c) = \{\, \sigma \in X^* \,\big|\, \mathcal{J}(a) \geq \mathcal{J}(c) + \langle \sigma, a-c \rangle \text{ for all } a \in X \,\}.$$

By definition $\mathbf{\Delta}^*$ is convex, lower semicontinuous and nonnegative, therefore the set $\mathbf{\Sigma} = \partial\mathbf{\Delta}(0) = (\mathbf{\Delta}^*)^{-1}(\{0\}) \subset X^*$ is convex, closed and, by hypothesis (3.3), satisfies $B(0,1/C) \subset \mathbf{\Sigma} \subset B(0,C)$. In fact, the homogeneity of $\mathbf{\Delta}$ implies $\mathbf{\Delta}^* = \mathcal{X}_{\mathbf{\Sigma}}$, where $\mathcal{X}_A$ denotes the characteristic function of the set A which means $\mathcal{X}_A(w) = 0$ for $w \in A$ and $\mathcal{X}_A(w) = \infty$ for $w \notin A$.

Remark 3.11 *In the case of convex $\mathcal{I}(t,\cdot)$, the subdifferential allows for a formulation of the problem as a differential inclusion:*

$$0 \in \partial\mathbf{\Delta}(\dot{c}(t)) + \partial(\mathcal{I}(t,\cdot) + \mathcal{X}_{\mathcal{P}})(c(t)) \text{ for a.e. } t \in [0,T]. \tag{3.12}$$

By definition and with the relation $\partial(\mathcal{I}(t,\cdot)+\mathcal{X}_{\mathcal{P}}) = \partial\mathcal{I}(t,\cdot) + \partial\mathcal{X}_{\mathcal{P}}$ this means

$$\exists\, \tilde{\sigma}(t) \in \partial\mathcal{I}(t,c(t))\ \exists\, \tilde{n}(t) \in \partial\mathcal{X}_{\mathcal{P}} :\ -\tilde{\sigma}(t) - \tilde{n}(t) \in \partial\mathbf{\Delta}(\dot{c}(t)).$$

Here $-\tilde{\sigma}(t)$ can be understood as the conjugate force to the variable c, and $-\tilde{n}(t)$ is the reaction of the boundary of $\mathcal{P}$. Clearly the sum of both has to be equal to the dissipative force. We have the following explicit formulae for the subdifferentials

$$\begin{aligned}
\partial\mathbf{\Delta}(z) &= \begin{cases} \mathbf{\Sigma} & \text{for } z = 0, \\ \{\, \sigma \in \mathbf{\Sigma} \,\big|\, \langle \sigma, z \rangle = \mathbf{\Delta}(z) \,\} & \text{for } z \neq 0; \end{cases} \\
\partial\mathcal{X}_{\mathcal{P}}(c) &= \mathrm{NC}_c\mathcal{P} = \{\, \sigma \in X^* \,\big|\, \langle \sigma, v \rangle \leq 0 \text{ for all } v \in \mathrm{TC}_c\mathcal{P} \,\},
\end{aligned}$$

where $\mathrm{NC}_c\mathcal{P}$ is the (outer) normal cone in the point $c \in \mathcal{P}$. We will use this formulation in Section 5.6 to derive a flow formulation for the phase transformation problem which has a lot of similarities with flow rules in plasticity theory. However, since these flow formulations are not needed in the sequel we omit the details and refer to [MiT00] for more information.

Assuming convexity Proposition 3.9 allows us to characterize the set of stable states $\mathcal{S}(t)$.

Proposition 3.12 *Let $\mathcal{P} = X$ and assume that the mapping $c \mapsto \mathcal{I}(t, c)$ is convex and differentiable. Then the relation $\mathcal{S}(t) = \sigma(t, \cdot)^{-1}(-\mathbf{\Sigma})$ holds.*

With these results at hand we are able to show a variety of existence, uniqueness and regularity results for admissible processes. First we give examples where the second assumption of Theorem 3.7, the weak closedness of the set of stable states holds. Since the time dependency does not matter, we neglect the time variable t for the moment.

Theorem 3.13 *If $\mathcal{P} = X$, $\mathcal{I}(\cdot)$ is convex and differentiable and the mapping $c \mapsto \sigma(c)$ is weakly continuous, then the set of stable states $\mathcal{S}$ is weakly closed.*

Proof. This follows directly from the fact that $\mathbf{\Sigma}$ is weakly closed (as it is strongly closed and convex) and therefore $\sigma^{-1}(-\mathbf{\Sigma})$ is weakly closed. ∎

Corollary 3.14 *If $\mathcal{P} = X$ and if the energy functional is quadratic, i.e., $\mathcal{I}(c) = \frac{1}{2}\langle Lc, c\rangle + \langle G, c\rangle + h$ for some bounded, positive semi–definite linear operator $L : X \to X^*$, $G \in X^*$ and $h \in \mathbb{R}$, then $\mathcal{S}$ is weakly closed.*

The following theorem gives some general properties of admissible processes under stronger convexity assumptions. Strict convexity implies continuity in time of admissible processes; and uniform convexity implies even Lipschitz continuity. Finally we give conditions which guarantee uniqueness and strong convergence of the solutions to the incremental problem. However, our applications in the PT problem always lead to the case where uniform convexity is violated, only strict convexity will be possible, see Remark 6.6.

Theorem 3.15 *Let c_0 be a stable initial condition.*

1. *If $\mathcal{I}$ is strictly convex in c (i.e. $\mathcal{I}(t, \theta c_1 + (1-\theta)c_2) < \theta\mathcal{I}(t, c_1) + (1-\theta)\mathcal{I}(t, c_2)$ if $0 < \theta < 1$) then every admissible process c with $c(0) = c_0$ is continuous on $[0, T]$.*
2. *If $\mathcal{I}(t, \cdot)$ is uniformly convex (i.e. $\mathcal{I}(t, \theta c_1 + (1-\theta)c_2) \leq \theta\mathcal{I}(t, c_1) + (1-\theta)\mathcal{I}(t, c_2) - \frac{\alpha}{2}\theta(1-\theta)\|c_1 - c_2\|^2$ for some $\alpha > 0$ independent of c_1, c_2, and t), then every admissible process c with $c(0) = c_0$ is Lipschitz continuous on $[0, T]$.*
3. *If $\mathcal{I}(t, \cdot)$ is uniformly convex, $\mathcal{P} = X$ and $\mathcal{I}(t, \cdot) \in \mathrm{C}^3(X, \mathbb{R})$ holds, then there exists a unique admissible process c with $c(0) = c_0$. Moreover, the solutions to the incremental problem converge strongly to c:*

$$\sup_{t\in[0,T]} \|c(t) - c^N(t)\| \leq K \cdot \Big(\max_{k=1\ldots N} (t_k^N - t_{k-1}^N) \Big)^{1/2}$$

for some positive constant K.

The proof of this theorem, which will not be needed in the sequel, can be found in [MiT00]. A version of the uniqueness result in a slightly different context can be found in in [HaN75].

4 Application to martensitic phase transformations

4.1 Modeling hysteresis via metastability or via internal variables

Before starting with the PT problem we relate the above model to the general theory of systems with hysteretic behavior. Hysteresis is a well–known effect which is either attributed to rate–independent friction (like dry friction) or to metastability in nonlinear systems. We want to show that there is a close connection between both situations when we consider slow or quasistatic processes, their extreme being rate–independent processes as considered here. In particular, we motivate how the usage of an internal variable enables us to use global minimization instead of studying local stability.

We show this connection at a simple model for a mass point moving slowly (without inertia) in the nonconvex potential

$$E(t,w) = \rho(w) - g(t)w \quad \text{with } \rho(w) = \tfrac{1}{4}(w^2-3)^2$$

where $g : [0,T] \to \mathbb{R}$ is the slowly varying loading. Of course we could consider more general functions ρ as well. Since the position $w(t) \in \mathbb{R}$ of the mass point always minimizes $E(t,\cdot)$ locally we have $\rho'(w(t)) = g(t)$ and $\rho''(w(t)) \geq 0$ which may have no or one solution lying in $(-\infty,-1]$ or $[1,\infty)$, respectively. The first situation is called phase e_1 and the second is called phase e_2. We assume that the mass point remains in a locally stable state (metastable) as long as possible. If stability is lost then it drops into the nearest stable point. The energy loss will be dissipated.

Now take that $g(t) = \min\{t-2, 8-t\}$, $t \geq 0$, and $w(0) = -2$. Then $\rho'(w(t)) = g(t)$ with $w(t) \leq -1$ for $t \in [0,4] \cup (10,\infty)$ and $w(t) \in [1,\infty)$ for $t \in (4,10]$. Hence there is a sudden phase change at $t = 4$ (with $g(t) = 2$) from e_1 to e_2 where w jumps from -1 to 2 and the energy drops down from 3 down to $-15/4$. At $t = 7$ (with $g(t) = -2$) the phase changes back from e_2 to e_1 where w jumps from 1 to -2 and the energy drops again from 3 to $-15/4$. Note that the global minimizer changes at $g(t) = 0$ without any energy loss; this relates to the Maxwell line having no hysteresis.

The same hysteresis effect can be obtained by using the internal variable $c^{\mathrm{p}} \in \{e_1, e_2\}$ and the dissipation coefficients $\kappa_{1\to2} = \kappa_{2\to1} = 3-(-15/4) = 27/4$. We choose a strictly convex function $\widehat{\rho} : \mathbb{R} \to \mathbb{R}$ with $\widehat{\rho}(w) = \rho(w)$ for $w \leq -1$ and define

$$\widetilde{E}(t,w,c^{\mathrm{p}}) = \begin{cases} \widehat{\rho}(w) - g(t)w & \text{for } c^{\mathrm{p}} = e_1, \\ \widehat{\rho}(-w) - g(t)w & \text{for } c^{\mathrm{p}} = e_2; \end{cases}$$

and $D^{\mathrm{p}} : P^{\mathrm{p}} \times P^{\mathrm{p}} \to \mathbb{R}_{\geq}$ using $\kappa_{i\to j}$ as above. Then, $(w(t), c^{\mathrm{p}}(t))$ is stable if

$$\widetilde{E}(t, w(t), c^{\mathrm{p}}(t)) \leq \widetilde{E}(t, \widehat{w}, a^{\mathrm{p}}) + D^{\mathrm{p}}(c^{\mathrm{p}}(t), a^{\mathrm{p}}) \quad \text{ for all } \widehat{w} \in \mathbb{R} \text{ and } a^{\mathrm{p}} \in \{e_1, e_2\}.$$

Using $a^{\mathrm{p}} = c^{\mathrm{p}}(t) = e_j$ we obtain the necessary local criterion $\widehat{\rho}'(w(t)) = g(t)$ which gives a unique solution $w_{g(t)}$ by strict convexity. Let $\eta(g) = \widehat{\rho}(w_g) - gw_g = \min\{\, \widehat{\rho}(w) - gw \,\big|\, w \in \mathbb{R}\}$; then $-\eta(\cdot)$ is the Legendre transform of $\widehat{\rho}$ with $\eta(2) = 3$ and $\eta(-2) = -15/4$. Additionally, we may assume, by choosing $\widehat{\rho}$ suitably for $w > 1$, that $\eta(g) < 27/4 + \eta(-g)$ for $g > 2$. Now, asking for stability of $(w(t), c^{\mathrm{p}}(t)) = (w_{g(t)}, e_1)$ with respect to $(\widehat{w}, e_2)$ we find the additional restriction $\eta(g(t)) \leq 15/4 + \eta(-g(t))$ which by construction is equivalent to $g(t) \leq 2$.

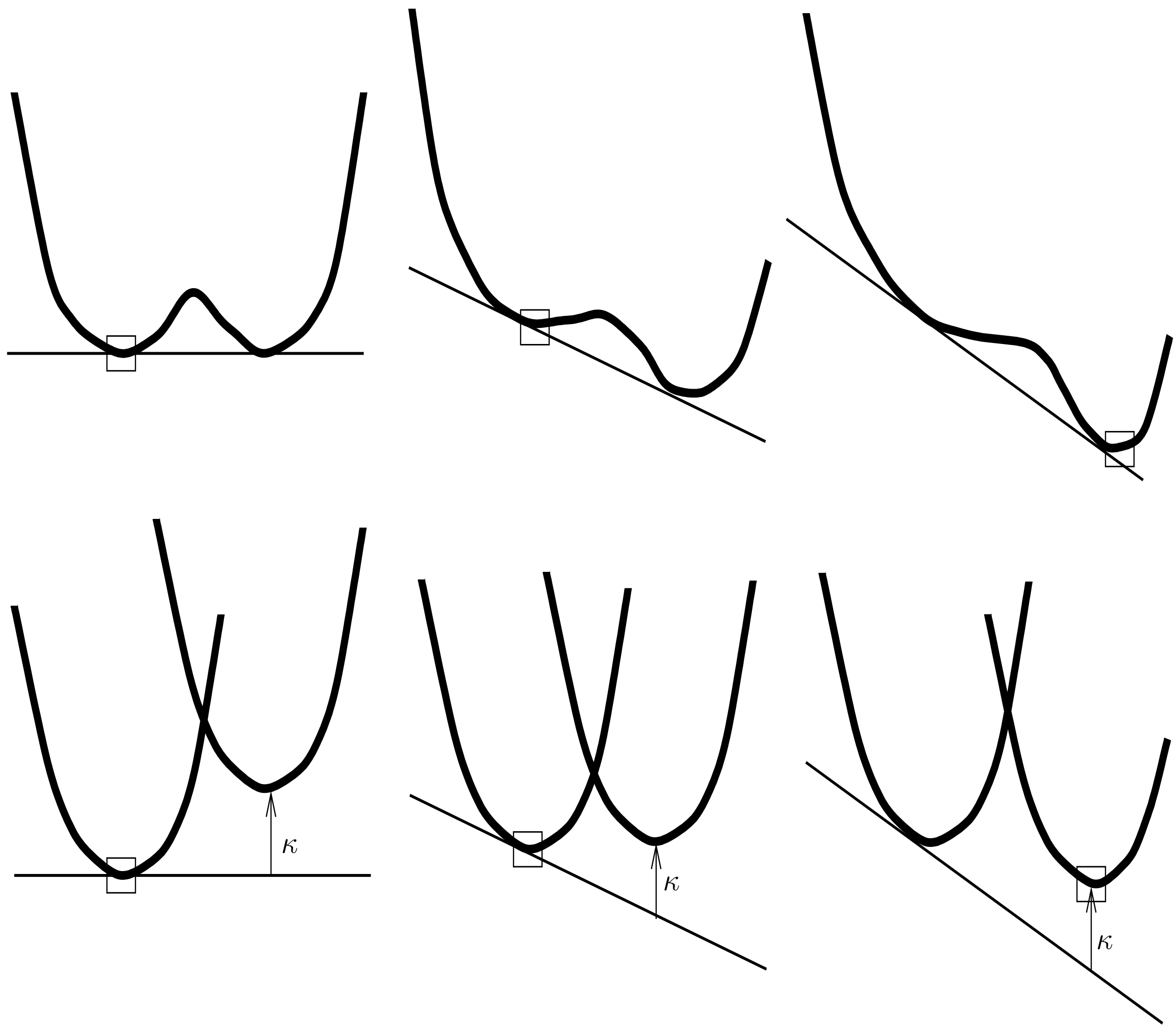

Figure 1: The upper three graphs show $E(t,\cdot)$ for the three time levels $t=2$, $t=3.8$ and $t=4.2$ (from left to right). The lower graphs display $\widetilde{E}(t,\cdot,c^{\mathrm{p}})$ with $c^{\mathrm{p}}=e_1$ to the right and $c^{\mathrm{p}}=e_2$ to the left for the same time levels. The squares $\square$ indicate the (meta) stable positions.

Thus, using the same loading $g(t)=\min\{t-2,8-t\}$, $t\geq 0$, as above and the initial condition $w(0)=-2$, we obtain exactly the same solution if we set additionally $c^{\mathrm{p}}(t)=e_1$ for $t\in[0,4]\cup(10,\infty)$ and $c^{\mathrm{p}}(t)=e_2$ for $t\in(4,10]$. The major difference to the approach without internal variable is that in this situation the solution pair $(w(t),c^{\mathrm{p}}(t))$ is a global minimizer rather than $w(t)$ alone being a local minimizer only.

4.2 Mathematical formulation of the model

We repeat the description of the model in Section 2 so that readers without a strong background in mechanics can follow the mathematical analysis.

We denote by $\Omega\subset\mathbb{R}^d$ $d\in\{1,2,3\}$ the reference configuration of an elastically deformable body. For $n>1$ the discrete set $P^{\mathrm{p}}=\{e_1,\ldots,e_n\}$ is the set of possible **pure phases**, e_j is the j–th unit vector in $\mathbb{R}^n$. The purpose of this notation is to identify a very natural relaxation with the convexification of the set $\mathcal{P}$ in the following section. The set $\mathcal{P}^{\mathrm{p}}=\{c^{\mathrm{p}}:\Omega\to P^{\mathrm{p}}\}$ is denoted as set of pure phase configurations, the superscript

$^{\mathrm{p}}$ is needed to distinguish between unrelaxed (or pure–phase) models and relaxed models where c takes its values in a continuous set.

The elastic material properties of phase e_j are given determined by the stored energy density $W_{e_j}(F)$ where F replaces the deformation gradient ∇u. For given $c^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}} = \mathrm{L}^1(\Omega, P^{\mathrm{p}})$ and $u \in \mathrm{W}^{1,p}(\Omega)$ the elastic bulk energy is given by

$$\mathcal{E}^{\mathrm{p}}_{\mathrm{elast}}(c^{\mathrm{p}}, u) = \int_\Omega W_{c^{\mathrm{p}}(x)}(x, \nabla u(x))\,\mathrm{d}x.$$

Our aim is to describe time–dependent processes which are driven by external forces and boundary conditions. The boundary conditions are imposed by a closed (affine) subspace $\mathcal{V}$ of $\mathrm{W}^{1,p}(\Omega)$. Without going into the details we assume that the change of the Neumann–boundary data and the bulk force can be represented by a time dependent functional $G \in \mathcal{V}^*$, the dual space of $\mathcal{V}$. Thus, at time t the system is described by $(c^{\mathrm{p}}, u) \in \mathcal{P}^{\mathrm{p}} \times \mathcal{V}$ and has the total energy (Gibbs' energy)

$$\mathcal{E}^{\mathrm{p}}(t, c^{\mathrm{p}}, u) = \textstyle\int_\Omega W_{c^{\mathrm{p}}(x)}(\nabla u(x))\,\mathrm{d}x - \langle G(t), u\rangle.$$

Here $\langle\cdot,\cdot\rangle$ is the duality pairing in $\mathcal{V}^* \times \mathcal{V}$.

After the introduction of the potential energy we proceed to the second ingredient of the theory, the dissipation potential. In the literature (see e.g. [GNS83]) it is also denoted as *pseudo potential*, since it takes a rate as an argument. We find it more convenient, not to impose a linear structure on the state space. In Section 2 we assumed that each PT dissipates an amount of energy which is proportional to the volume of the region occupied by the transformed material times the constant $\kappa_{i\to j}$ depending only on the state e_i before and the state e_j after the PT. The friction coefficients $\kappa_{i\to j}$, $i = 1\ldots n$, $j = 1\ldots n$ generate a dissipation potential similar to the functional $\mathbf{\Delta}$ in Section 3. Since the dissipated energy should be positive and finite we have that $\kappa_{i\to j} \geq 0$ and $\kappa_{i\to i} = 0$ for every i, j. The general theory requires that $\mathbf{\Delta}$ is coercive and convex. These two requirements translate directly into

$$\textbf{coercivity:} \qquad \kappa_{i\to j} > 0 \text{ if } i \neq j \tag{4.1}$$

$$\textbf{triangle inequality:} \qquad \kappa_{i\to j} \leq \kappa_{i\to k} + \kappa_{k\to j} \tag{4.2}$$

for every i, j, k. Note that our state space is not endowed with a linear structure since phase distribution functions cannot be multiplied with a scalar. With these assumptions we can postulate our dissipation function which instead of taking rates as an argument is defined on pairs of states.

$$\mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}(t_1), c^{\mathrm{p}}(t_2)) = \int_\Omega D^{\mathrm{p}}(c^{\mathrm{p}}(t_1, x), c^{\mathrm{p}}(t_2, x))\,\mathrm{d}x \quad \text{where } D^{\mathrm{p}}(e_i, e_j) = \kappa_{i\to j}.$$

Consequently the total dissipation associated to a process is given the sum of the individual dissipations.

$$\mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2) = \sup\Big\{ \sum_{k=1}^{m} \mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}(\tau_{k-1}), c^{\mathrm{p}}(\tau_k)) \,\Big|\, m{\in}\mathbb{N},\ t_1{\leq}\tau_0{<}\tau_1{<}\ldots{<}\tau_m{\leq}t_2 \Big\}. \tag{4.3}$$

The central concept which determines the evolution of the phase configuration and the deformation is the stability criterion (2.5). Using the notation introduced so far we obtain the representation

$$\mathcal{E}^{\mathrm{p}}(t, c^{\mathrm{p}}, u) \leq \mathcal{E}^{\mathrm{p}}(t, b^{\mathrm{p}}, v) + \mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}(t), b^{\mathrm{p}}) \tag{4.4}$$

for every $b^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$, $v \in \mathcal{V}$. Since the dissipated energy $\mathcal{D}^p$ does not depend on the deformation u, the concept of stability implies that for given t and c^{p} the deformation u is a minimizer of $\mathcal{E}$. For notational simplicity we introduce the energy of a phase configuration

$$\mathcal{I}^{\mathrm{p}}(t, c^{\mathrm{p}}) = \inf\{\, \mathcal{E}^{\mathrm{p}}(t, c^{\mathrm{p}}, v) \,\Big|\, v \in \mathcal{V} \,\} \tag{4.5}$$

so that the stability inequality now reads

$$\mathcal{I}^{\mathrm{p}}(t, c^{\mathrm{p}}(t)) \leq \mathcal{I}^{\mathrm{p}}(t, b^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}(t), b^{\mathrm{p}})$$

for every $b^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$.

Similarly the energy inequality (2.3) takes the form

$$\mathcal{I}^{\mathrm{p}}(t_2, c^{\mathrm{p}}(t_2)) + \mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2) \leq \mathcal{I}^{\mathrm{p}}(t_1, c^{\mathrm{p}}(t_1)) - \int_{t_1}^{t_2} \langle \dot{G}(t), u(t) \rangle \,\mathrm{d}t, \tag{4.6}$$

where the last term contains the energy contribution of the work done by the external loading. Note that the original formulation was local in time. Similarly to the stability inequality the energy inequality generates a restriction in the space of stresses and strain–rates. However, we demonstrated in Section 3 (Theorem 3.3) that the local version is not really needed. Hence, to simplify the mathematical structure we ask only for the integrated form to be satisfied. The term $\mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; t_1, t_2)$ includes the dissipation which occurs through phase transformations in the time interval $[t_1, t_2]$. This implies that at time t a PT from c_- to c_+ is energetically possible whenever the gain in the energy $\mathcal{I}^{\mathrm{p}}$ is at least as large as the associated dissipation:

$$\mathcal{I}^{\mathrm{p}}(t, c^{\mathrm{p}}_-) - \mathcal{I}^{\mathrm{p}}(t, c^{\mathrm{p}}_+) \geq \mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}_-, c^{\mathrm{p}}_+). \tag{4.7}$$

After these considerations we see, that the mechanical derivation in Section 2 and the abstract theory in Section 3 can merged into a simple mathematical model.

Definition 4.1 *A process $c^{\mathrm{p}} : [0,T] \to \mathcal{P}^{\mathrm{p}}$ is called **stable** if for every $t \in (0,T]$*

$$\mathcal{I}^{\mathrm{p}}(t, a^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c^{\mathrm{p}}(t), a^{\mathrm{p}}) \geq \mathcal{I}^{\mathrm{p}}(t, c^{\mathrm{p}}(t)) \quad \textit{for all } a^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}.$$

*The process c^{p} satisfies the **weak energy inequality** on $[0,T]$ if*

$$\mathcal{I}^{\mathrm{p}}(T, c^{\mathrm{p}}(T)) + \mathrm{Diss}^{\mathrm{p}}(c^{\mathrm{p}}; 0, T) \leq \mathcal{I}^{\mathrm{p}}(0, c^{\mathrm{p}}(0)) - \textstyle\int_0^T \langle \dot{G}(t), u(t) \rangle \,\mathrm{d}t.$$

*A process is called **admissible** if it is stable and satisfies the weak energy inequality.*

The definition of stability gives a time–dependent restriction for $c^{\mathrm{p}}(t) \in \mathcal{P}^{\mathrm{p}}$. It does not see the (rate–independent) dynamics of the process. The stability implies that for no time t_1 the process can be changed on $(t_1, t_2]$ such that the difference between dissipation and energy gain is negative. We will use the terms stability principle and extremum principle synonymously.

In agreement with the abstract theory in Section 3 in which admissible processes are defined via (2.S) and (2.E) we postulate the following **time–continuous problem**.

> ***(CP)*** *For given $c^{\mathrm{p}}_0 \in \mathcal{P}^{\mathrm{p}}$, find an admissible process $c^{\mathrm{p}} : [0,T] \to \mathcal{P}^{\mathrm{p}}$ with $c^{\mathrm{p}}(0) = c^{\mathrm{p}}_0$.*

The class of stored energy density functions W_{e_i} admitted in our general approach is still very large. Rigorous results we only obtain for a much smaller, yet important, subclass, namely quadratic energy density functions for linearized elasticity. In this case $\mathcal{U}(G, c^{\mathrm{p}})$ can be found by solving a linear elliptic system, see Theorem 5.2. This restriction leads of course to the known weakness that large deformations are not described correctly. For nonquadratic energy densities only existence of energy minimizing deformations is known, this suffices for the abstract formulation of the time discretized approximation scheme. Most questions which are beyond existence of time discrete solutions are open. The reason we formulate the model in such a great generality is that the abstraction clarifies the structure of the algorithm and the arising problems.

The stability condition is posed globally over the whole body Ω, however it involves all possible variations of the state a^{p}. Hence, it is possible to derive local conditions for material points $x \in \Omega$. For the following two results we refer to [MTL98]. The first one can be interpreted as the stability condition of a pure phase e_i with respect to formation of a nucleus including different phases c.

Theorem 4.2 *(a) Let $c^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$ be stable and let $x_0 \in \Omega$ be such that c^{p} and $F = \nabla u$ are continuous in $x = x_0$. Then, with $\mathbb{W}(a, F)$ and $\mathbf{\Delta}$ from Sections (5.3,5.4) we have*

$$W_{c^{\mathrm{p}}(x_0)}(F(x_0)) \leq \mathbf{\Delta}(a - c(x_0)) + \mathbb{W}(a, F(x_0)) \quad \textit{for all } a \in P = \mathrm{conv}(P^{\mathrm{p}}).$$

(b) Let $c^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$ be stable and let $x_0 \in \Omega$ be a point where c^{p} has a phase boundary with normal vector ν. Let $(e^{\pm}, F^{\pm})$ be the left and right limits of $(c^{\mathrm{p}}(x), F(x))$ which satisfy by strain and stress–compatibility $F^+ - F^- = a \otimes \nu$ and $T_0 \stackrel{\mathrm{def}}{=} \partial_F W_{e^+}(F^+)\nu = \partial_F W_{e^-}(F^-)\nu$. Then, we have

$$-D^{\mathrm{p}}(e^+, e^-) \leq W_{e^-}(F^-) - W_{e^+}(F^+) + T_0 \cdot ([F^+ - F^-]\nu) \leq D^{\mathrm{p}}(e^-, e^+).$$

The scalar quantity $\xi = W_{e^-}(F^-) - W_{e^+}(F^+) + T_0 \cdot ([F^+ - F^-]\nu)$ is the normal component $\nu \cdot (E_0 \nu)$ of the Eshelby tensor E_0. It is well known in the literature [Grinfeld91, KaR88, Lev95, Lev97] and is frequently denoted as the driving force for the interface. As long as ξ lies strictly between $-D^{\mathrm{p}}(e^+, e^-)$ and $D^{\mathrm{p}}(e^-, e^+)$ the interface cannot move in a time–continuous process since the dissipation plays the role of a threshold. If ξ reaches the value $D^{\mathrm{p}}(e^-, e^+)$ the interface can move such that phase e^+ grows, and analogously phase e^- will start to grow if $\xi = -D^{\mathrm{p}}(e^+, e^-)$.

From the abstract theory in Section 3, especially the incremental approach in Subsection 3.3, it is natural to try to find solutions by time discretization. Choosing a discrete set of times $0 = t_0 < t_1 < \ldots < t_k < \ldots < t_N = T$ we are lead to consider the associated **incremental problem (IP)**:

(IP) *For given $c_0^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$, find $c_1^{\mathrm{p}}, \ldots, c_N^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$ such that, for $k = 1, \ldots, N$,*

$$\mathcal{I}^{\mathrm{p}}(t_k, c_k^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c_{k-1}^{\mathrm{p}}, c_k^{\mathrm{p}}) = \inf\{\, \mathcal{I}^{\mathrm{p}}(t_k, a^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c_{k-1}^{\mathrm{p}}, a^{\mathrm{p}}) \,\big|\, a^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}} \,\}.$$

In contrast to Section 3 we don't have the special form $\mathcal{D}^{\mathrm{p}}(a^{\mathrm{p}}, c^{\mathrm{p}}) = \mathbf{\Delta}(c^{\mathrm{p}} - a^{\mathrm{p}})$. However, the triangle inequality (4.2) takes over the role of convexity in Section 3 and we are still able to conclude that the solutions of of (IP) are stable for $t = t_0, \ldots, t_N$. In complete analogy to the abstract Theorem 3.4 we obtain the desired properties for the solutions of (IP).

Lemma 4.3 *Assume that $D^{\mathrm{p}} : P^{\mathrm{p}} \times P^{\mathrm{p}} \to \mathbb{R}_{\geq}$ satisfies (4.2), then the solutions of (IP) are stable, i.e., $\mathcal{I}^{\mathrm{p}}(t_k, c_k^{\mathrm{p}}) \leq \mathcal{I}^{\mathrm{p}}(t_k, a^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c_k^{\mathrm{p}}, a^{\mathrm{p}})$ for all $a^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$, and satisfy the discrete energy inequality $\mathcal{I}^{\mathrm{p}}(t_k, c_k^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c_{k-1}^{\mathrm{p}}, c_k^{\mathrm{p}}) \leq \mathcal{I}^{\mathrm{p}}(t_{k-1}, c_{k-1}^{\mathrm{p}}) + \int_{t_{k-1}}^{t_k} \partial_t \mathcal{I}^{\mathrm{p}}(s, c_{k-1}^{\mathrm{p}})\,\mathrm{d}s$.*

We refer to Remark 5.4 for the necessity of the triangle inequality in this context.

5 Coarse graining via relaxation

The formulation we derived so far is not satisfactory because of two reasons:

(1) The existence of admissible processes cannot be assured, not even for the incremental problem. Since the state space $\mathcal{P}^{\mathrm{p}} = \mathrm{L}^1(\Omega, P^{\mathrm{p}})$ is **not** weakly closed, it is unclear whether, for fixed $t \in [0, T]$ and $a^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$, the functional

$$c^{\mathrm{p}} \mapsto \mathcal{I}^{\mathrm{p}}(t, c^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(a^{\mathrm{p}}, c^{\mathrm{p}})$$

has minimizers. Therefore we cannot apply the existence Theorem 3.7.

(2) The numerical solution of the incremental problem would exhibit strong oscillations, cf. [CaP97]. Energetic considerations show that if the data is chosen in order to achieve deformations in the "soft" directions of a shape memory alloy, strong oscillations are needed to avoid large stresses.

For these two reasons it is necessary to derive a coarse grained model which relies on effective quantities like phase portions instead of being restricted to pure phase distributions. It is not particularly difficult to find extensions from pure phase models to models with phase mixtures so that the existence of admissible processes can be guaranteed. The problem is to establish a clear link between the extended and the original model. Such links can be obtained in several ways. One possibility is that of regularizing the problem, e.g. by adding a small surface energy to $\mathcal{I}$. Another abstract approach which is closely related to the spirit of Γ–convergence can be found in [Th00b].

Here we indicate a third possibility which remains solely on the level of the incremental problem (IP). Since (IP) is a variational problem we may study its relaxation. Neither in a real physical process nor in a mathematical approximation of (IP) we can expect that exact minimization of $\mathcal{I}^{\mathrm{p}}(t_k, \cdot) + \mathcal{D}^{\mathrm{p}}(c_{k-1}^{\mathrm{p}}, \cdot)$ occurs. Thus, we propose the following **approximate incremental problem** $(\mathrm{AIP})_\varepsilon$ where quasiminimal solutions are allowed.

(AIP)$_\varepsilon$ *For given $\varepsilon > 0$ and $c_0^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$, find $c_1^{\mathrm{p}}, \ldots, c_N^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$ such that*

$$\mathcal{I}^{\mathrm{p}}(t_k, c_k^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c_{k-1}^{\mathrm{p}}, c_k^{\mathrm{p}}) \leq \varepsilon + \inf\{\, \mathcal{I}^{\mathrm{p}}(t_k, a^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(c_{k-1}^{\mathrm{p}}, a^{\mathrm{p}}) \;\big|\; a^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}} \,\}.$$

Clearly, this problem always has a solution $(c_{1,\varepsilon}^{\mathrm{p}}, \ldots, c_{N,\varepsilon}^{\mathrm{p}})$ and we may extract suitable limits $(c_1, \ldots, c_N)$. Then, the aim is to find a relaxed incremental problem (RIP) which has these limits as a solution.

5.1 Weak limits and relaxation

Consider solutions $(c_{1,\varepsilon}^{\mathrm{p}}, \ldots, c_{N,\varepsilon}^{\mathrm{p}})$ of the approximate incremental problem $(\mathrm{AIP})_\varepsilon$ with $\varepsilon \ll 1$. In practice, one is often not interested in the details of $c_{k,\varepsilon}$, for example the exact geometry of the microstructure of the phases. More relevant are global or averaged quantities like the energy $\mathcal{I}^{\mathrm{p}}(t_k, c_{k,\varepsilon}^{\mathrm{p}})$, the deformation $u_{k,\varepsilon} = \mathcal{U}(G(t_k), c_{k,\varepsilon})$, or the macroscopic distribution of the volume fractions of the phases. Even when $c^{\mathrm{p}}_{k,\varepsilon}(x)$ does not converge for $\varepsilon \to 0$, these quantities may converge when the weak limit exists, $c^{\mathrm{p}}_{k,\varepsilon} \rightharpoonup c_k$. Weak convergence is defined via

$$\lim_{\varepsilon\to 0} \fint_{B(x_0,r)} c^{\mathrm{p}}_{k,\varepsilon}(x)\,\mathrm{d}x = \fint_{B_r(x_0)} c_k(x)\,\mathrm{d}x \quad \text{for all } B(x_0,r) \subset \Omega.$$

(Throughout we use the notation $\fint_A f(x)\,\mathrm{d}x = \frac{1}{\mathrm{vol}(A)} \int_A f(x)\,\mathrm{d}x$.) The weak limit c_k takes its values no longer in $P^{\mathrm{p}} = \{e_1, \dots, e_n\}$ but in its convex hull $P = \mathrm{conv}(P^{\mathrm{p}})$, that is $c_k \in \mathcal{P} = \mathrm{L}^1(\Omega, P)$. Here $\mathcal{P}$ is the weak closure of $\mathcal{P}^{\mathrm{p}}$ or, equivalently, the closure in the norm topology of the convex hull of $\mathcal{P}^{\mathrm{p}}$. Thus, the functional

$$\mathcal{J}^{\mathrm{p}}_k(a^{\mathrm{p}}, c^{\mathrm{p}}) = \mathcal{I}^{\mathrm{p}}(t_k, c^{\mathrm{p}}) + \mathcal{D}^{\mathrm{p}}(a^{\mathrm{p}}, c^{\mathrm{p}}), \quad a^{\mathrm{p}}, c^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}},$$

has to be extended to a functional $\mathcal{J}_k : \mathcal{P} \times \mathcal{P} \to \mathbb{R}_{\geq}$ such that

$$\mathcal{J}_k(a^{\mathrm{p}}, c) = \inf\{ \liminf_{n\to\infty} \mathcal{J}^{\mathrm{p}}_k(a^{\mathrm{p}}, c^{\mathrm{p}}_n) \,\Big|\, c^{\mathrm{p}}_n \rightharpoonup c \text{ for } n \to \infty \}. \tag{5.1}$$

Such a functional is called the relaxation of $\mathcal{J}^{\mathrm{p}}_k(a^{\mathrm{p}}, \cdot)$, see [Dac89, Rou97]. While the relaxation of $\mathcal{J}^{\mathrm{p}}_k(a^{\mathrm{p}}, c^{\mathrm{p}})$ in the second argument c^{p} is clear, it is an open problem how the extension from $a^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$ to $a \in \mathcal{P}$ has to be done in order to have a proper connection between the following **relaxed incremental problem** (RIP) and the original incremental problem.

> ***(RIP)*** *For given $c^{\mathrm{p}}_0 \in \mathcal{P}^{\mathrm{p}}$, find $c_1, \dots, c_N \in \mathcal{P}$ such that*
>
> $$\mathcal{J}_k(c_{k-1}, c_k) = \inf\{ \mathcal{J}_k(c_{k-1}, a) \,\Big|\, a \in \mathcal{P} \}.$$

It remains to manifest the relations between (RIP) and (IP) (or $(\mathrm{AIP})_\varepsilon$ resp.). A minimal requirement for (RIP) is that at least one of the following two questions (or appropriate variants of them) can be answered affirmatively.

Question 1: Let $(c^{\mathrm{p}}_{1,\varepsilon}, \dots, c^{\mathrm{p}}_{N,\varepsilon})$ solve $(\mathrm{AIP})_\varepsilon$ such that $c^{\mathrm{p}}_{k,\varepsilon} \rightharpoonup c_k$ for $\varepsilon \to 0$. Does $(c_1, \dots, c_N)$ solve (RIP)?

Question 2: Let $(c_1, \dots, c_N)$ solve (RIP). Do there exist solutions $(c_{1,\varepsilon}, \dots, c_{N,\varepsilon})$ of $(\mathrm{AIP})_\varepsilon$ such that $c_{k,\varepsilon} \rightharpoonup c_k$ for $\varepsilon \to 0$?

5.2 Separate relaxation

Here we propose a relaxation of $\mathcal{J}^{\mathrm{p}}$ which uses the special structure of $\mathcal{J}^{\mathrm{p}}(a^{\mathrm{p}}, c^{\mathrm{p}})$, namely that it is a sum of two terms. The first term $\mathcal{I}^{\mathrm{p}}(t_k, c^{\mathrm{p}})$ does not depend on a^{p} and the second term $\mathcal{D}^{\mathrm{p}}(a^{\mathrm{p}}, c^{\mathrm{p}})$ depends on a^{p} and c^{p} only pointwise under the integral over Ω. For $a, c \in \mathcal{P}$ we now define the relaxations

$$\mathcal{I}(t, c) = \inf\{ \liminf_{m\to\infty} \mathcal{I}^{\mathrm{p}}(t, c^{\mathrm{p}}_m) \,\Big|\, c^{\mathrm{p}}_m \rightharpoonup c \text{ for } m \to \infty \}, \tag{5.2}$$

$$\mathcal{D}(a, c) = \inf\{ \liminf_{m\to\infty} \mathcal{D}^{\mathrm{p}}(a^{\mathrm{p}}_m, c^{\mathrm{p}}_m) \,\Big|\, a^{\mathrm{p}}_m \rightharpoonup a,\ c^{\mathrm{p}}_m \rightharpoonup c \text{ for } m \to \infty \}, \tag{5.3}$$

$$\mathcal{J}_k(a, c) = \mathcal{I}(t_k, c) + \mathcal{D}(a, c).$$

Note that $\mathcal{D}^{\mathrm{p}}(\cdot,\cdot)$ is relaxed in both variables simultaneously. It is not so difficult to see that $\mathcal{J}_k$ is in fact the simultaneous relaxation of $\mathcal{J}_k^{\mathrm{p}}$ in both its arguments, viz.,

$$\mathcal{J}_k(a,c) = \inf\{ \liminf_{m\to\infty} \mathcal{J}_k^{\mathrm{p}}(a_m^{\mathrm{p}}, c_m^{\mathrm{p}}) \,\big|\, a_m^{\mathrm{p}} \rightharpoonup a,\ c_m^{\mathrm{p}} \rightharpoonup c \text{ for } m\to\infty \}.$$

However, it is unclear whether using this simultaneous relaxation is of help in answering Question 1 or 2 from above. We refer to [Th00a, Th00b] and Theorem 5.6 for more details.

The separate relaxation is compatible with the requirement (5.1) for the second argument of $\mathcal{J}$, but concerning the first argument it contains a modeling assumptions. This assumption can be understood best by considering infinimizing sequences $(c_{k,\varepsilon}^{\mathrm{p}})$ obtained from $(\mathrm{AIP})_\varepsilon$ with weak limits c_k for $\varepsilon\to 0$. Our assumption reduces to

$$\mathcal{D}^{\mathrm{p}}(c_{k-1,\varepsilon}^{\mathrm{p}}, c_{k,\varepsilon}^{\mathrm{p}}) \to \mathcal{D}(c_{k-1}, c_k) \text{ for } \varepsilon\to 0. \tag{5.4}$$

Clearly this is not correct for general sequences. The problem is that the weak limits c_k contain only information on the macroscopic volume fractions of the pure phases but no information on their microscopic arrangements. However, for infinimizing sequences the modeling error may be small as these sequences try to minimize $\mathcal{D}^{\mathrm{p}}$ which means that changes in the microscopic arrangement are energetic less favorable. Yet the microscopic arrangements are enforced by the first term $\mathcal{I}^{\mathrm{p}}(t_k,\cdot)$ and might differ at t_{k-1} and t_k. Then our relaxation underestimates the dissipated energy.

5.3 Relaxation of the energy $\mathcal{I}^{\mathrm{p}}$

First we recall the definition of quasiconvexity which plays a central role in the existence theory for minimizers in the calculus of variations. A function W is called quasiconvex if for all $F\in\mathbb{R}^{d\times d}$ we have

$$⨍_B W(F+\nabla\phi(y))\,\mathrm{d}y \geq W(F) \quad \text{for all } \phi\in \mathrm{W}_0^{1,p}(B),$$

where $B = B(0,1) \subset \mathbb{R}^d$. Further on we always assume that the energy densities W_{e_j} of the pure phases are quasiconvex in this sense. Considering a small ball of material $B(x_0,r) = \{ x\in\Omega \,\big|\, |x-x_0| < r \}$ where the phase distribution is given by $c = ⨍_{B(x_0,r)} c^{\mathrm{p}}(x)\,\mathrm{d}x$, we now allow for suitable rearrangements $c^{\mathrm{p}}\in \mathrm{L}^1(B(x_0,r),P^{\mathrm{p}})$ of the phases without changing the average and we also allow for perturbations ϕ of the deformation which vanishes on the boundary of the ball. For small r the perturbation can be understood as a microscopic adjustment which does not affect the macroscopic deformation. Letting $x = x_0 + ry$ with $y\in B$ we are lead to the **mixture function**

$$\mathbb{W}(c,F) \stackrel{\mathrm{def}}{=} \inf\Big\{ ⨍_B W_{c^{\mathrm{p}}(y)}(F+\nabla\phi(y))\,\mathrm{d}y \,\Big|\, \phi\in \mathrm{W}_0^{1,p}(B),\ c^{\mathrm{p}}\in \mathrm{L}^1(B,P^{\mathrm{p}}),\ ⨍_B c^{\mathrm{p}}(y)\,\mathrm{d}y = c \Big\} \tag{5.5}$$

which was already defined in [Koh91] as a tool to study the quasiconvexification of $\min\{ W_{e_j}(F) \,\big|\, j=1,\ldots,n \}$. In this work, $\mathbb{W}(c,F)$ itself plays a central role and is called the *quasiconvexification with fixed phase fractions*, denoted by $Q_cW(F)$.

By the general theory the set B can in fact be replaced by any open subset of $\mathbb{R}^d$. We now have defined a function $\mathbb{W} : P\times\mathbb{R}^{d\times d} \to \mathbb{R}; (c,F) \mapsto \mathbb{W}(c,F)$ which satisfies

$\mathbb{W}(e_j, F) = W_{e_j}(F)$ since W_{e_j} was assumed to be quasiconvex. Thus, $\mathbb{W}$ contains information on the mixture theory without the addition of new physics. In general the minimum $\mathbb{W}(c, F)$ is not achieved, but one needs a infimizing sequence $(c^{\mathrm{p}}_\varepsilon)_{\varepsilon>0}$ which tells us that certain microstructures are needed to minimize the energy under given phase fractions.

In the general case it is possible to show that $\mathbb{W}(\cdot, F) : P \to \mathbb{R}$ is convex for fixed F and that $W(c, \cdot) : \mathbb{R}^{d\times d} \to \mathbb{R}$ is quasiconvex for fixed $c \in P$, cf. [Th00b]. Here we establish this fact only for some special cases where $\mathbb{W}$ can be given explicitly. The first such case is the one–dimensional case. There we have the explicit formula

$$\mathbb{W}(c, F) = \mathcal{L}\Big(\textstyle\sum_{j=1}^n c^{(j)} \mathcal{L} W_{e_j} \Big)(F) \quad \text{with } c = (c^{(1)}, \dots , c^{(n)})^T,$$

where $\mathcal{L}$ is the Legendre–Fenchel transform with $(\mathcal{L}W)(\sigma) = \inf\{\, \sigma F - W(F) \,\big|\, F \in \mathbb{R}\,\}$, cf. Definition 3.10. In the quadratic case with $W_{e_i}(F) = \frac{1}{\eta^{(i)}}(F - A^{(i)})^2 + \beta^{(i)}$ with $\eta^{(i)} > 0$ this gives $\mathbb{W}(c, F) = \frac{(F-A\cdot c)^2}{\eta\cdot c} + \beta \cdot c$.

The second case where the mixture function can be given explicitly is that of linearized elasticity with an identical elastic tensor $\mathbf{C} \in \mathrm{Lin}(\mathbb{R}^{d\times d}_{\mathrm{sym}}, \mathbb{R}^{d\times d}_{\mathrm{sym}})$ which is positive definite, i.e., there exists $\alpha > 0$ such that

$$\mathbf{C}E{:}E \overset{\mathrm{def}}{=} \textstyle\sum_{i,j=1}^d (\mathbf{C}E)_{ij} E_{ij} \geq \alpha E{:}E \quad \text{for all } E \in \mathbb{R}^{d\times d}_{\mathrm{sym}}. \tag{5.6}$$

We continue to use the abbreviation $E = \frac{1}{2}(F+F^T) - I$ for the linearized strain tensor. To $\mathbf{C}$ and a direction $\omega \in \mathbb{R}^d$ we associate the so–called acoustic tensor $\mathbb{A}(\omega) \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ via

$$[\mathbb{A}(\omega)a] \cdot a = \tfrac{1}{4}[\mathbf{C}(a\otimes\omega + \omega\otimes a)]{:}(a\otimes\omega + \omega\otimes a) \quad \text{for } a \in \mathbb{R}^d.$$

Proposition 5.1 *Assume that $\mathbf{C}$ satisfies (5.6) and that*

$$W_{e_i}(F) = \tfrac{1}{2}[\mathbf{C}(E - A_i)]{:}[E - A_i] + \beta^{(i)} \quad \textit{with } A_i \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \textit{ and } \beta^{(i)} \in \mathbb{R}. \tag{5.7}$$

Then we have

$$\mathbb{W}(c, F) = \textstyle\sum_{i=1}^n c^{(i)} W_{e_i}(F) + w_{\mathrm{mix}}(c), \tag{5.8}$$

where $w_{\mathrm{mix}} : P \to \mathbb{R}$ is convex and satisfies the bounds

$$0 \geq w_{\mathrm{mix}}(c) \geq -\tfrac{1}{2} \textstyle\sum_{j=1}^n c^{(j)} [\mathbf{C}A_j]{:}A_j + \tfrac{1}{2}\big[\mathbf{C}A^c\big]{:}A^c \quad \textit{with } A^c = \textstyle\sum_{j=1}^n c^{(j)} A_j.$$

Moreover, if $A_j = A_0 + a^{(j)} A$ for $j = 1, \dots , n$, then we have the explicit formula

$$\begin{aligned} w_{\mathrm{mix}}(c) &= -\tfrac{\gamma}{2}\Big[\textstyle\sum_1^n \big(a^{(j)}\big)^2 c^{(j)} - (a\cdot c)^2 \Big] \textit{ with } a = (a^{(1)}, \dots , a^{(n)}) \textit{ and} \\ \gamma &= \max\{\, |\mathbb{A}(\omega)^{-1/2}(\mathbf{C}A)\omega|^2 \,\big|\, |\omega| = 1 \,\}. \end{aligned} \tag{5.9}$$

In the case $n = 2$ the explicit formula (5.9) is always applicable. The first part of the assertion follows from the quadratic nature of the problem, see e.g., [Kha67, Roi67, Kha83, Koh91] where also the case $n = 2$ was solved. Formula (5.9) for $n \geq 3$ was established in [Mie00] and applied to a particular PT problem with one austenite and two martensitic phases in [PAM98]. For partial results in more generals situations we refer to [SmW98, Mie00].

The importance of the mixture function is that it is exactly the right tool to characterize the relaxation $\mathcal{I}$ of the functional $\mathcal{I}^{\mathrm{p}}$. Define $\mathcal{E} : [0,T] \times \mathcal{P} \times \mathcal{V} \to \mathbb{R}$ via

$$\mathcal{E}(t,c,u) = \textstyle\int_\Omega \mathbb{W}(c(x), \nabla u(x))\,\mathrm{d}x - \langle G(t), u\rangle. \tag{5.10}$$

Then, under suitable assumptions technical assumptions (concerning continuity and growth) we have $\mathcal{E}(t,c,u) = \inf\{\, \liminf_{k\to\infty} \mathcal{E}^{\mathrm{p}}(t, c_k^{\mathrm{p}}, u_k) \,\big|\, c_k^{\mathrm{p}} \rightharpoonup c,\ u_k \rightharpoonup u \,\}$ and

$$\mathcal{I}(t,c) = \min\{\, \mathcal{E}(t,c,u) \,\big|\, u \in \mathcal{V} \,\}. \tag{5.11}$$

We formulate the corresponding result only in the special case treated below.

Theorem 5.2 *Let W_{e_j} be given as in Proposition 5.1 and $\mathcal{E}$ as in (5.10) with $\mathbb{W}$ from (5.8). Then, the relaxation $\mathcal{I}$ defined in (5.2) is given by*

$$\begin{aligned}\mathcal{I}(t,c) &= \min\{\, \mathcal{E}(t,c,u) \,\big|\, u \in \mathcal{V} \subset \mathrm{W}^{1,2}(\Omega) \,\} \\ &= \eta(t) + \textstyle\int_\Omega [(L_0 c)(x) + \widehat{g}(t,x)] \cdot c(x) + w_{\mathrm{mix}}(c(x))\,\mathrm{d}x,\end{aligned}$$

where $L_0 \in \mathrm{Lin}(Y,Y)$ with $Y = \mathrm{L}^2(\Omega, \mathbb{R}^n)$ is a selfadjoint operator and $\widehat{g}(t,\cdot) \in Y$.

Proof. The validity of (5.11) in this special case is proved in [Th00b]. The second identity follows immediately from the first by using the quadratic nature and the fact that $\mathrm{L}^1(\Omega, P)$ can be embedded continuously into Y. Minimization in u does not involve w_{mix} and it follows that the minimizer $u = \mathcal{U}(G(t), c)$ depends linearly on $G(t)$ and c, that is $\mathcal{U}(G,c) = K_1 G + K_2 c$ with $K_1 \in \mathrm{Lin}(\mathcal{V}^*, \mathcal{V})$ and $K_2 \in \mathrm{Lin}(Y, \mathcal{V})$. Inserting this linear expression gives the desired quadratic expression. ■

5.4 Relaxation of the dissipation functional

The relaxation $\mathcal{D} : \mathcal{P} \times \mathcal{P} \to \mathbb{R}_{\geq}$ of the function $\mathcal{D}^{\mathrm{p}} : \mathcal{P}^{\mathrm{p}} \times \mathcal{P}^{\mathrm{p}} \to \mathbb{R}_{\geq}$ as defined in (5.3) can be calculated much easier, since $\mathcal{D}^{\mathrm{p}}(a^{\mathrm{p}}, c^{\mathrm{p}})$ depends on a^{p} and c^{p} only through a simple integration over the point values of $D^{\mathrm{p}}(a^{\mathrm{p}}(x), c^{\mathrm{p}}(x))$. For $(a,c) \in P \times P$ we set

$$D(a,c) = \inf\{\, \textstyle\sum_{j,i=1}^n m_{ji}\kappa_{j\to i} \,\big|\, m_{ji} \geq 0,\ \sum_{i=1}^n m_{ji} = a^{(j)},\ \sum_{j=1}^n m_{ji} = c^{(i)} \,\}. \tag{5.12}$$

Clearly D is obtained by minimizing $D^{\mathrm{p}}(a^{\mathrm{p}}, c^{\mathrm{p}})$ over $a^{\mathrm{p}}, c^{\mathrm{p}} \in \mathrm{L}^1(B, P^{\mathrm{p}})$ under the constraints $a = ⨍_B a^{\mathrm{p}}(y)\,\mathrm{d}y$ and $c = ⨍_B c^{\mathrm{p}}(y)\,\mathrm{d}y$. The coefficients m_{ji} in (5.12) are simply the relative measures of the sets $\{\, y \in B \,\big|\, a(y) = e^{(j)},\ c(y) = e^{(i)} \,\}$. Using $\kappa_{j\to j} = 0$, $\kappa_{\min} = \min\{\, \kappa_{j\to i} \,\big|\, j \neq i \,\}$, and $\kappa_{\max} = \max\{\, \kappa_{j\to i} \,\big|\, j \neq i \,\}$, we immediately find

$$\tfrac{\kappa_{\min}}{2}|c{-}a|_1 \leq D(a,c) \leq \tfrac{\kappa_{\max}}{2}|c{-}a|_1 \quad \text{where } |b|_1 \stackrel{\text{def}}{=} \textstyle\sum_{j=1}^n |b^{(j)}|. \tag{5.13}$$

We also have $D(e_j, c) = \sum_{i=1}^n c^{(i)}\kappa_{j\to i}$. Since the difference $c{-}a$ for $a, c \in P$ will play a major role we define the linear space

$$\mathbb{R}^n_* = \{\, z \in \mathbb{R}^n \,\big|\, z{\cdot}e_* = 0 \,\} \quad \text{where } e_* = (1, \ldots, 1)^T.$$

Proposition 5.3 *The function $D : P \times P \to [0, \infty)$ is convex, that is*

$$D(\theta a_1{+}(1{-}\theta)a_2, \theta c_1{+}(1{-}\theta)c_2) \leq \theta D(a_1, c_1) + (1{-}\theta)D(a_2, c_2). \text{ for all } \theta \in [0, 1].$$

If D^{p} satisfies the triangle inequality (4.2), then so does D and there exists a unique function $\Delta : \mathbb{R}^n_* \to [0, \infty)$ which is homogeneous of degree 1 (that is $\Delta(\alpha b) = \alpha\Delta(b)$ for all $\alpha \geq 0$ and $b \in \mathbb{R}^n_*$) such that D has the form

$$D(a, c) = \Delta(c{-}a) \quad \text{for all } a, c \in P. \tag{5.14}$$

Remark 5.4 *The triangle inequality plays a twofold role. First, it guarantees the difference representation with Δ. Second, it guarantees the stability of the solutions of the incremental problem. To illustrate the necessity in both cases consider a situation with three phases. Let $\kappa_{1\to3} = \kappa$ and $\kappa_{j\to i} = 1$ else for $i \neq j$. For the volume fractions $a = (1/4, \alpha, 3/4{-}\alpha)^T$ and $c = (0, \alpha, 1{-}\alpha)^T$ we have $c{-}a = (-1/4, 0, 1/4)^T$ which is independent of α. For $\kappa \in (0, 2]$ the triangle inequality holds and we find $D(a, c) = \kappa/4$. For $\kappa > 2$ and $\alpha \in [0, 1/4]$ we have $D(a, c) = 2\alpha{+}(1{-}4\alpha)\kappa/4$ which clearly contradict the difference formula.*

Moreover, consider the trivial energies $\mathcal{I}^{\mathrm{p}}(t, e_j) = 6{-}2j$ together with the initial state $c_0^{\mathrm{p}} = e_1$. Then, $\mathcal{I}^{\mathrm{p}}(t, c^{\mathrm{p}}) + D^{\mathrm{p}}(e_1, c^{\mathrm{p}})$ is minimized with $c^{\mathrm{p}} = e_2$ for $\kappa \geq 3$ and with $c^{\mathrm{p}} = e_3$ for $\kappa \leq 3$. However, e_2 is never stable whereas e_3 is always stable.

Proof. Consider $a_i, c_i \in P$ and let $m_{jk}^{(i)}$ be the minimizers in the definition of $D(a_i, c_i)$ for $i = 1, 2$, that is $m_{jk}^{(i)} \geq 0$, $\sum_{l=1}^n m_{jl}^{(i)} = a_i^{(j)}$, and $\sum_{j=1}^n m_{jl}^{(i)} = c_i^{(l)}$. Then, $m_{jl}^{(3)} = \theta m_{jl}^{(1)} + (1{-}\theta)m_{jl}^{(2)}$ forms an admissible set in the definition of $D(\theta a_1{+}(1{-}\theta)a_2, \theta c_1{+}(1{-}\theta)c_2)$, which gives the convexity result after separating the minimum into the weighted sum of two minima. The triangle inequality for D follows by a straight forward use of its definition and (4.2).

The proof of the representation (5.14) involves the standard duality theory for linear transport problems as developed in the textbooks [Gal60, Gas58]. This theory implies the dual representation

$$D(a, c) = \max\{\, \mu \cdot c - \nu \cdot a \,\big|\, \mu, \nu \in \mathbb{R}^n \text{ with } \mu^{(j)}{-}\nu^{(i)} \leq \kappa_{i\to j} \text{ for all } i, j \,\}.$$

We have to show that the triangle inequality allows us to consider the case $\nu = \mu$ only, which gives the function

$$\Delta(c{-}a) \stackrel{\text{def}}{=} \max\{\, \mu \cdot (c{-}a) \,\big|\, \mu \in \mathbb{R}^n \text{ with } \mu^{(j)}{-}\mu^{(i)} \leq \kappa_{i\to j} \text{ for all } i, j \,\}. \tag{5.15}$$

To this end we employ the following standard optimality condition. If (m_{ij}) is a minimizer in (5.12) and (μ, ν) a maximizer in the dual formulation then

$$m_{ij} > 0 \quad \text{implies} \quad \kappa_{i\to j} = \mu^{(j)}{-}\nu^{(i)}. \tag{5.16}$$

Thus, using $\kappa_{j\to j} = 0$ we are finished if we can show that $m_{jj} > 0$ for all $j = 1, \ldots, n$.

Using the triangle inequality we show that it is always possible to find a minimizer (m_{ij}) in (5.12) which satisfies

$$m_{ii} = \sigma_i \quad \text{with } \sigma_i \stackrel{\text{def}}{=} \min\{a^{(i)}, c^{(i)}\} \text{ for } i = 1, \ldots, n. \tag{5.17}$$

Indeed, if this would not be the case for some i_0 then there exist j_0 and k_0 such that $\delta = \min\{m_{i_0j_0}, m_{k_0i_0}\} \in (0, \sigma_{i_0}-m_{i_0i_0}]$. By the triangle inequality it is advantageous, for minimizing $\sum \kappa_{i\to j}m_{ij}$, to increase $m_{i_0i_0}$ and $m_{k_0j_0}$ by $\delta > 0$ and decrease $m_{i_0j_0}$ and $m_{k_0i_0}$ by the same amount. Clearly after a finite number of such operations we achieve $\sigma_{i_0} = m_{i_0i_0}$.

Using identity (5.17) and the extremality condition (5.16) we conclude $\mu = \nu$ in the dual problem whenever a and c have positive coordinates. Then, the desired relation $D(a,c) = \Delta(c-a)$ follows. By continuity of D and Δ it immediately extends to all of $P\times P$. ■

We now define the integrated function $\mathbf{\Delta} : \mathrm{L}^1(\Omega, \mathbb{R}^n_*) \to \mathbb{R}_\geq$ via

$$\mathbf{\Delta}(z) = \int_\Omega \Delta(z(x))\,\mathrm{d}x. \tag{5.18}$$

Under the assumption that D^p satisfies the triangle inequality we have $D^\mathrm{p}(a^\mathrm{p}, c^\mathrm{p}) = \Delta(c^\mathrm{p}-a^\mathrm{p})$ and hence $\mathcal{D}^\mathrm{p}(a^\mathrm{p}, c^\mathrm{p}) = \mathbf{\Delta}(c^\mathrm{p}-a^\mathrm{p})$ for all $a^\mathrm{p}, c^\mathrm{p} \in \mathcal{P}^\mathrm{p}$.

Theorem 5.5 *Assume that $D^\mathrm{p} : P^\mathrm{p} \times P^\mathrm{p} \to \mathbb{R}_\geq$ satisfies the triangle inequality (4.2). Then, the relaxation $\mathcal{D} : \mathcal{P} \times \mathcal{P} \to \mathbb{R}_\geq$ defined in (5.3) is given by $\mathcal{D}(a,c) = \mathbf{\Delta}(c-a)$. It satisfies the triangle inequality and $\kappa_{\min}\|c-a\|_{\mathrm{L}^1} \le 2\mathcal{D}(a,c) \le \kappa_{\max}\|c-a\|_{\mathrm{L}^1}$.*

The proof can be found in [Th00b].

5.5 The separately relaxed incremental problem

Using these two relaxed functionals we obtain a relaxed formulation of the incremental problem (IP). With $\mathcal{I}$ from (5.11) and $\mathbf{\Delta}$ from (5.18) we postulate the following **separately relaxed incremental problem**:

(SRIP) *For given $c_0 \in \mathcal{P}$, find $c_1, \ldots, c_N \in \mathcal{P}$ such that*

$$\mathcal{I}(t_k, c_k) + \mathbf{\Delta}(c_k-c_{k-1}) = \inf\{\, \mathcal{I}(t_k, c) + \mathbf{\Delta}(c-c_{k-1}) \,\big|\, c \in \mathcal{P}\,\}.$$

Clearly, solutions $(c_k)_{k=1,\ldots,N}$ of (SRIP) are stable and satisfy the incremental energy inequality, viz.,

(S) $\quad \mathcal{I}(t_k, c_k) \le \mathcal{I}(t_k, a) + \mathbf{\Delta}(a-c_k) \quad$ for all $a \in \mathcal{P}$,

$(\mathrm{E})_{\mathrm{incr}}$ $\quad \mathcal{I}(t_k, c_k) + \mathbf{\Delta}(c_k-c_{k-1}) \le \mathcal{I}(t_k, c_{k-1}) = \mathcal{I}(t_{k-1}, c_{k-1}) + \int_{t_{k-1}}^{t_k} \partial_t \mathcal{I}(s, c_{k-1})\,\mathrm{d}s.$

There are two ways to see that (SRIP) always has a solution. In the first case we use the general construction of $\mathcal{I}$ and $\mathbf{\Delta}$ which implies that $\mathcal{J}_k(a,\cdot) : \mathcal{P} \to \mathbb{R}; c \mapsto \mathcal{I}(t_k, c) + \mathbf{\Delta}(c-a)$ is weakly lower semicontinuous. Since $\mathcal{P}$ is bounded in $\mathrm{L}^1(\Omega, \mathbb{R}^d)$ the existence of a minimizer follows. We will make this concrete for our special case in Section 6.

The second approach to the existence of minimizers $c \in \mathcal{P}$ is much more useful for doing actual numerics. We reintroduce $u \in \mathcal{V}$ and minimize the functional

$$\begin{aligned} \mathcal{J}_k(a,c,u) &= \mathcal{E}(t_k, c, u) + \mathbf{\Delta}(c-a) \\ &= \textstyle\int_\Omega \mathbb{W}(c(x), \nabla u(x)) + \Delta(c(x)-a(x))\,\mathrm{d}x - \langle G(t), u\rangle \end{aligned}$$

with respect to $(c, u) \in \mathcal{P} \times \mathcal{V}$. As $c(x)$ appears only locally in the integral we may minimize with respect to c first and do the minimization pointwise in $x \in \Omega$. For this purpose define

$$\Psi(a, F) = \min\{\, \mathbb{W}(c, F) + \Delta(c-a) \,\big|\, c \in P \,\}, \tag{5.19}$$

where, by continuity of $\mathbb{W}(\cdot, F) + \Delta(\cdot - a)$, the minimum is achieved at $c = C(a, F)$, i.e., $\Psi(c, F) = \mathbb{W}(C(a, F), F) + \Delta(C(a, F) - a)$. It is shown in [Th00b] that for each $a \in P$ the function $\Psi(a, \cdot)$ is quasiconvex such that the functional

$$\mathcal{F}(t, a, u) = \int_\Omega \Psi(a(x), \nabla u(x))\,\mathrm{d}x - \langle G(t), u\rangle$$

is weakly lower semicontinuous, cf. the general theory in [AcF84, Dac89]. For our special case of Proposition 5.1 with $n = 2$ the functions Ψ and C can be given explicitly, see [MTL98]. Now, (SRIP) can be reformulated as follows.

(SRIP)' *Given $c_0^{\mathrm{p}} \in \mathcal{P}^{\mathrm{p}}$, find $(c_1, u_1), \ldots, (c_N, u_N) \in \mathcal{P} \times \mathcal{V}$ such that*

$$\mathcal{F}(t_k, c_{k-1}, u_k) = \inf\{\, \mathcal{F}(t_k, c_{k-1}, v) \,\big|\, v \in \mathcal{V} \,\} \text{ and } c_k(x) = C(c_{k-1}(x), \nabla u_k(x)).$$

The formulation of (SRIP)' has the major advantage that it reduces to a simple variational problem for the variable $u_k \in \mathcal{V}$. By the quasiconvexity of $\Psi(a, \cdot)$ the existence of a minimizer is clear. The important message is that the function $\Psi(a, F)$ is in principle completely determined from the pure energy densities $W_{e_j}(F)$ and the dissipation coefficients $\kappa_{i\to j} = D(e_i, e_j)$. It can be calculated either analytically or numerically before even starting to solve the incremental problem (SRIP).

Of course, (SRIP) might not be a correct relaxation of (IP) but it has the advantage that it is explicit and still close enough at the proper relaxation (RIP) such that there is hope that the solutions behave similarly. So far, a rigorous connection between (IP) and (SRIP) via $(\mathrm{AIP})_\varepsilon$ is only known in the (trivial) one–dimensional case and in the case of linearized elasticity with two phases with the same elastic moduli. The following result which is established in [Th00a, Th00b] answers Question 2 in Section 5.1 affirmatively.

Theorem 5.6 *Consider the case of Proposition 5.1 with $n = 2$ and take any time discretization $0 = t_0 < t_1 < \ldots < t_N = T$. Then, (SRIP) has a solution and for every such solution $(c_k)_{k=1,\ldots,N} \in \mathcal{P}^N$ there exists $\varepsilon_0 > 0$ and solutions $(c_{k,\varepsilon}^{\mathrm{p}})_{k=1,\ldots,N} \in (\mathcal{P}^{\mathrm{p}})^N$ of $(AIP)_\varepsilon$ for $\varepsilon \in (0, \varepsilon_0)$ which satisfy $c_{k,\varepsilon}^{\mathrm{p}} \rightharpoonup c_k$.*

The main reason why this special case can be handled is that here the formation of microstructure can be controlled very well. It is sufficient to study lamination in one specific direction ω which is characterized as maximizer in (5.9). Then, it is possible to satisfy the convergence property (5.4).

5.6 The relaxed time–continuous problem

One reason for choosing the separate relaxation was to obtain rather explicit results. The second reason is that the structure of the problem remains the same. In particular, (SRIP) can be understood the incremental version of a time–continuous relaxed problem. The functionals $\mathcal{I}$ and $\mathbf{\Delta}$ constructed above are now the ones used in the abstract existence theory of Section 3. The **separately relaxed time–continuous problem** is exactly that of the existence of an admissible process as stated in the abstract section.

(SRCP) *Given $c_0 \in \mathcal{P}$, find $c : [0,T] \to \mathcal{P}$ with $c(0) = c_0$ such that $c(t) \in \mathcal{P}$ is stable for all $t \in (0,T]$ and that the weak energy inequality holds, viz.,*

$$\begin{array}{ll} (S) & \mathcal{I}(t, c(t)) \le \mathcal{I}(t, a) + \mathbf{\Delta}(a - c(t)) \quad \text{for all } t \in (0,T] \text{ and } a \in \mathcal{P}, \\ (E) & \mathcal{I}(T, c(T)) + \int_0^T \Delta(\mathrm{d}c) \le \mathcal{I}(0, c_0) - \int_0^T \langle \dot G(t), u(t)\rangle \,\mathrm{d}t, \end{array}$$

where $u(t) = \mathcal{U}(G(t), c(t))$ is a minimizer of $\mathcal{E}(t, c(t), \cdot)$.

Note that in our situation we have $\partial_t \mathcal{I}(t, c) = -\langle \dot G(t), \mathcal{U}(G(t), c(t))\rangle$ such that we are in accordance with (2.E).

According to Section 3.3 we can associate to (SRCP) an incremental problem to obtain solutions. By construction this is exactly the separately relaxed incremental problem (SRIP) we started with. It is the purpose of the next section to show that the assumptions of the abstract theory are satisfied for a nontrivial special case. Thus, we are able to conclude existence of solutions for (SRCP).

Finally we show that (SRCP) can be interpreted as a differential inclusion which looks similar to flow rules in plasticity. This will be a specification of the abstract flow rules discussed in Remark 3.11. In the engineering literature such flow rules are used for phase transformation problems as well, see [LIS99, HaG99, GoM00].

Here we use specifically that $\mathcal{P} = \mathrm{L}^1(\Omega, P)$ and $\mathbf{\Delta}(z) = \int_\Omega \Delta(z(x))\,\mathrm{d}x$ are given pointwise in $x \in \Omega$. This allows us to make the differential inclusion (3.12) quite explicit. Define $\Sigma \subset \mathbb{R}^n_* = \{\, z \in \mathbb{R}^n \,\big|\, z \cdot e_* = 0 \,\}$ via the subdifferential of $\Delta : \mathbb{R}^n_* \to \mathbb{R}$ as $\Sigma = \partial\Delta(0)$. According to (5.15) we have

$$\Sigma = \mathrm{conv}\{\sigma_1, \ldots, \sigma_N\} \quad \text{and} \quad \Delta(z) = \max\{\, z \cdot \sigma_j \,\big|\, j = 1, \ldots, N \,\}.$$

In particular we find $\mathbf{\Sigma} = \{\, \sigma \in \mathrm{L}^\infty(\Omega, \mathbb{R}^n_*) \,\big|\, \sigma(x) \in \Sigma \text{ for a.e. } x \in \Omega \,\}$ and

$$\partial\mathbf{\Delta}(z) = \{\, \sigma \in \mathbf{\Sigma} \,\big|\, \sigma(x) \cdot z(x) = \Delta(z(x)) \text{ for a.e. } x \in \Omega \,\}.$$

Similarly the outer normal cone $\partial\mathcal{X}_\mathcal{P}(c) = \mathrm{NC}_c\mathcal{P}$ can be characterized pointwise

$$\mathrm{NC}_c\mathcal{P} = \{\, \sigma \in \mathrm{L}^\infty(\Omega, \mathbb{R}^n_*) \,\big|\, \sigma(x) \in \mathrm{NC}_{c(x)}P \text{ for a.e. } x \in \Omega \,\},$$

where $\mathrm{NC}_{c(x)}P \subset \mathbb{R}^n_*$ is the finite–dimensional normal cone of the polytope P.

The derivative of the functional $\mathcal{I}(t, \cdot)$ is best represented by reintroducing the deformation u and the associated equation for the elastic equilibrium. Recall for this end that $\mathcal{I}(t, c)$ was defined by minimizing $\mathcal{E}(t, c, u)$ with respect to $u \in \mathcal{V}$:

$$\mathrm{D}_u\mathcal{E}(t, c(t), u(t)) = 0 \quad \text{and} \quad 0 \in \partial\mathbf{\Delta}(\dot c(t)) + \mathrm{D}_c\mathcal{E}(t, c(t), u(t)) + \mathrm{NC}_{c(t)}\mathcal{P}.$$

Here the first equation is the classical elliptic system describing the elastic equilibrium and the second relation is a flow rule for $c(t, x)$ which is completely local in $x \in \Omega$.

Putting everything together the total problem can be written as follows.

Local flow formulation: *For almost all $(t, x) \in [0,T] \times \Omega$ we have*

$$\begin{aligned} -\mathrm{div}_x \tfrac{\partial}{\partial F}\mathbb{W}(x, \nabla_x u(t,x), c(t,x)) &= f(t,x) \\ 0 \in \partial\Delta(\dot c(t,x)) + \tfrac{\partial}{\partial c}\mathbb{W}(\nabla_x u(t,x), c(t,x)) &+ \mathrm{NC}_{c(t,x)}P. \end{aligned} \tag{5.20}$$

together with suitable boundary conditions.

The second relation can be reformulated without the subdifferential of Δ as follows:

$$\exists\, n(t,x) \in \mathrm{NC}_{c(t,x)}P \text{ such that } \Delta(\dot c(t,x)) + \dot c(t,x)\cdot\Big[\frac{\partial}{\partial c}\mathbb{W}(\nabla_x u(t,x), c(t,x)) + n(t,x)\Big] = 0$$

for a.e. $(t,x) \in [0,T]\times\Omega$. Recall that $\mathbb{W}$ is convex in $c \in P$, however it may not be differentiable. In such a case the may replace the partial derivative $\frac{\partial}{\partial c}\mathbb{W}(F,c)$ by any element in the subdifferential $\partial_c\mathbb{W}(F,c)$.

6 Existence of admissible processes for a special case

We now restrict our view completely to the case of linearized elasticity with two phases ($n = 2$) whose elastic tensors are identical. For this case the mixture function $\mathbb{W}$ is given explicitly in Proposition 5.1. We simplify the notation by identification of $P = \mathrm{conv}\{e_1, e_2\} \subset \mathbb{R}^2$ with the interval $[0,1]$ via the mapping $[0,1] \ni \theta \mapsto (1-\theta)e_1 + \theta e_2$. Moreover, we will use the (infinitesimal) displacement $\tilde u : \Omega \to \mathbb{R}^d$ rather than the displacement $u : x \mapsto x + \varepsilon\tilde u(x)$, where ε is a small parameter. We will consider phase distributions $\theta \in \mathcal{P} = \mathrm{L}^1(\Omega,[0,1]) \subset X \overset{\mathrm{def}}{=} \mathrm{L}^1(\Omega,\mathbb{R})$ and displacements $\tilde u \in \mathcal{V} = \{\, v \in \mathrm{W}^{1,2}(\Omega,\mathbb{R}^d) \,\big|\, v|_\Gamma = 0 \,\}$.

The mixture function takes the form

$$\mathbb{W}(\theta, E) = \tfrac{1-\theta}{2}\mathbf{C}[E-A_1]{:}(E-A_1) + \tfrac{\theta}{2}\mathbf{C}[E-A_2]{:}(E-A_2) - \tfrac{\gamma}{2}\theta(1-\theta), \qquad (6.1)$$

where $\mathbf{C}$ satisfies (5.6) and γ is defined in (5.9) with $A = A_2 - A_1$. The dissipation functional now reads $\mathbf{\Delta} : X \to \mathbb{R}_{\geq};\, \zeta \mapsto \int_\Omega \Delta(\zeta(x))\,\mathrm{d}x$ with

$$\Delta : \mathbb{R} \to \mathbb{R}_{\geq};\, \zeta \mapsto \max\{\zeta\kappa_{1\to2}, -\zeta\kappa_{2\to1}\} \qquad (6.2)$$

with $\kappa_{1\to2}, \kappa_{2\to1} > 0$.

The aim of this section is to proof an existence theorem for the separately relaxed problem (SRCP) for this special case. The following result gives a typically situation, however slight generalizations are possible.

Theorem 6.1 *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with C^1 boundary and $\mathcal{P}$, X, $\mathcal{V}$, $\mathbf{\Delta}$, and $\mathbb{W}$ as above. Moreover, assume $g \in \mathrm{C}^1([0,T],\mathcal{V}^*)$ and let*

$$\mathcal{I}(t,\theta) = \min\{\, \textstyle\int_\Omega \mathbb{W}(\theta, \tfrac12(\nabla\tilde u + \nabla\tilde u^T))\,\mathrm{d}x + \langle g(t), \tilde u\rangle \,\big|\, \tilde u \in \mathcal{V}\,\}.$$

If additionally either (a) or (b) with

(a) $\Gamma = \partial\Omega$ (full Dirichlet conditions),

(b) $A_2 - A_1 = b\otimes\xi + \xi\otimes b$ for some $b, \xi \in \mathbb{R}^d$ (infinitesimally rank–one connected wells).

hold, then for each $c_0 \in \mathcal{P}$ there exists an admissible process in the sense of Definition 3.1, i.e., a solution of (SRCP).

The proof of this result is the remainder of this section. In fact, we can strengthen the result slightly. If (a) as well as certain algebraic conditions on $\mathbf{C}$ and $A_2 - A_1$ hold, it is possible to prove strict convexity of $\mathcal{I}(t,\cdot)$, see Remark 6.6 at the end of this section. Then, Part 1 of Theorem 3.15 implies continuity of $c : (0,T] \to \mathcal{P}$.

Our special case is dictated from the fact that $\mathbb{W}$ is quadratic in c and ∇u. Hence, it is possible to work out the form of $\mathcal{I}(t,\theta)$ more explicitly than in Theorem 5.2. We use also the Hilbert space $Y = \mathrm{L}^2(\Omega,\mathbb{R})$, which is possible since the weak and strong topologies on $\mathcal{P}$ which are induced by Y and X are the same.

Proposition 6.2 *If the above assumptions and if $G \in \mathrm{C}^1([0,T],\mathcal{V}^*)$ hold, then there exist $\eta \in \mathrm{C}^1([0,T],\mathbb{R})$, $\widehat{g} \in \mathrm{C}^1([0,T],Y)$ and a symmetric operator $L \in \mathrm{Lin}(Y,Y)$ such that*

$$\mathcal{I}(t,\theta) = \eta(t) + \textstyle\int_\Omega [(\frac{1}{2}L\theta)(x) + \widehat{g}(t,x)]\theta(x)\,\mathrm{d}x.$$

Moreover, L is a pseudo–differential operator whose symbol $\ell : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$ is given by

$$\ell(\omega) = \gamma - \lambda(\omega) \quad \textit{with } \lambda(\omega) = [\mathbb{A}(\omega)^{-1}(\mathbf{C}A)\omega]\cdot[(\mathbf{C}A)\omega] \quad \textit{where } A = A_2 - A_1, \tag{6.3}$$

and hence is nonnegative, cf. (5.9).

Proof. We have to minimize $\mathcal{E}(t,\theta,\widetilde{u})$, which is quadratic in $(\theta,\widetilde{u}) \in Y \times \mathcal{V}$, with respect to $\widetilde{u} \in \mathcal{V}$. We obtain the unique minimizer $\widetilde{u} = K_1 G(t) + K_2\theta$ where $K_1 \in \mathrm{Lin}(\mathcal{V}^*,\mathcal{V})$ and $K_2 \in \mathrm{Lin}(Y,\mathcal{V})$. Note that K_2 is a pseudo–differential operator of order -1. Inserting this result defines the pseudo–differential operator L of order 0 (as $E = \frac{1}{2}(\nabla\widetilde{u}+\nabla\widetilde{u}^T)$) and the differentiable functions $t \mapsto \eta(t) \in \mathbb{R}$ and $t \mapsto \widehat{g}(t) \in Y$.

It remains to calculate the symbol of L. This is done most easily by inserting $E = E_{a,\omega} \stackrel{\text{def}}{=} \frac{1}{2}(a\otimes\omega + \omega\otimes a) \in \mathbb{R}^{d\times d}_{\text{sym}}$ into the quadratic part of $\mathbb{W}$:

$$\tfrac{1}{2}\mathbf{C}E_{a,\omega}{:}E_{a,\omega} - \theta\mathbf{C}E_{a,\omega}{:}A + \theta^2\tfrac{\gamma}{2} = \tfrac{1}{2}\mathbb{A}(\omega)a\cdot a - \theta\mathbf{C}A\omega\cdot a + \theta^2\tfrac{\gamma}{2}.$$

A subsequent minimization with respect to $a \in \mathbb{R}^d$ gives the desired result. ∎

From the theory of pseudo–differential operators it follows that the range of the symbol ℓ constitutes the continuous spectrum whereas compact perturbations which are not seen in the symbol may generate additional discrete eigenvalues. Such compact perturbations arise from the different boundary conditions for the displacement $\widetilde{u}$. However, we conclude that the (strictly) negative part of L must be compact and it is exactly this property which guarantees weak lower semicontinuity of $\mathcal{I}(t,\cdot) : Y \to \mathbb{R}$ and hence that of $\mathcal{I}(t,\cdot) : \mathcal{P} \subset \mathrm{L}^1(\Omega,\mathbb{R}) \to \mathbb{R}$. It is remarkable that the continuous spectrum always includes 0, as the definition of γ is just such that the minimum of $\ell(\omega)$, $0 \neq \omega \in \mathbb{R}^d$ is 0. This implies that $\mathcal{I}$ can never be uniformly convex, however strict convexity is still possible depending on the compact perturbations through the boundary conditions.

Our subsequent existence result will work for all cases where L is positive semi–definite. Thus, we give some cases where this can be guaranteed. We now have to take into account the boundary conditions for $\widetilde{u}$. We recall that $\Gamma \subset \partial\Omega$ is that part of the boundary where we have imposed Dirichlet data, i.e., $\widetilde{u}|_\Gamma = 0$.

Lemma 6.3 *The operator L is positive semi–definite if either (a) or (b) in Theorem 6.1 holds.*

Proof. The idea of the proof is to return to the quadratic functional $\mathcal{E}(t,\theta,\widetilde{u})$ and to eliminate θ first, cf. Section 5.5 and the derivation of (SRIP)' for the analogous strategy. Consider the quadratic form

$$Q(\theta,\widetilde{u}) = \textstyle\int_\Omega \frac{1}{2}\mathbf{C}E(\widetilde{u}){:}E(\widetilde{u}) - \theta\mathbf{C}E(\widetilde{u}){:}A + \theta^2\frac{\gamma}{2}\,\mathrm{d}x \quad \text{where } E(\widetilde{u}) = \frac{1}{2}(\nabla\widetilde{u} + \nabla\widetilde{u}^T),$$

which is the homogeneous part of degree 2 of $\mathcal{E}$.

On the one hand we have $\int_\Omega (L\theta)\theta\,\mathrm{d}x = 2\min\{\, Q(\theta,\widetilde{u}) \,\big|\, u \in \mathcal{V} \,\}$. On the other hand we may eliminate θ first to obtain

$$\widehat{Q}(\widetilde{u}) = \min\{\, Q(\theta,\widetilde{u}) \,\big|\, \theta \in Y \,\} = \textstyle\int_\Omega \frac{1}{2}\widehat{\mathbf{C}}E(\widetilde{u}){:}E(\widetilde{u})\,\mathrm{d}x,$$

where the reduced elasticity tensor $\widehat{\mathbf{C}}$ is given by $\widehat{\mathbf{C}}E{:}E = \mathbf{C}E{:}E - [\mathbf{C}E{:}A]^2/\gamma$. Clearly we have $L \geq 0$ if and only if $\widehat{Q}(u) \geq 0$ for all $u \in \mathcal{V}$. Since $\widehat{Q}$ is homogeneous the last condition is equivalent to convexity of $\widehat{Q}$.

The case (a) is now obtained by studying the symbol associated to $\widehat{Q}$ which takes the form $S(\omega) = \mathbb{A}(\omega) - \frac{1}{\gamma}[(\mathbf{C}A)\omega]\otimes[(\mathbf{C}A)\omega]$. Using $a = \mathbb{A}(\omega)^{-1/2}b$ and the definition of γ in the estimate

$$S(\omega)a\cdot a = |b|^2 - \tfrac{1}{\gamma}[(\mathbf{C}A)\omega\cdot\mathbb{A}(\omega)^{-1/2}b]^2 \geq |b|^2\Big(1-\tfrac{1}{\gamma}|\mathbb{A}(\omega)^{-1/2}(\mathbf{C}A)\omega|^2\Big)$$

shows that the symbol is always positive semidefinite. This means that the function $F \mapsto \frac{1}{4}\widehat{\mathbf{C}}(F{+}F^T){:}(F{+}F^T)$ is rank–one convex. Together with the full Dirichlet condition this implies convexity of $\widehat{Q}$, cf. [Dac89].

For the case (b) we cannot use the boundary conditions. Instead we show convexity of $\widehat{Q}$ by convexity of the integrand, i.e., $\widehat{\mathbf{C}}E{:}E \geq 0$ for all $E \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$. We define a scalar product on $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ via $\langle\!\langle E_1, E_2\rangle\!\rangle = \mathbf{C}E_1{:}E_2$. Then, on the one hand we have $\widehat{\mathbf{C}}E{:}E = \langle\!\langle E,E\rangle\!\rangle - \langle\!\langle E,A\rangle\!\rangle^2/\gamma$; and convexity holds, by the Cauchy–Schwartz inequality, if and only if $\langle\!\langle A,A\rangle\!\rangle \leq \gamma$. On the other hand γ takes the form

$$\gamma = \max\{\, \langle\!\langle E_{a,\omega}, A\rangle\!\rangle^2/\langle\!\langle E_{a,\omega}, E_{a,\omega}\rangle\!\rangle \,\Big|\, 0 \neq a,\omega \in \mathbb{R}^d \,\} \leq \langle\!\langle A,A\rangle\!\rangle.$$

By assumption (b), $A = E_{b,\xi}$ is an admissible candidate in the maximum and we conclude $\gamma = \langle\!\langle A,A\rangle\!\rangle$ and, hence, the desired convexity. ■

We now want to apply the abstract theory of Section 3. Clearly, the quadratic form of $\mathcal{I}$ and the continuous differentiability of in t implies that all necessary continuity assumptions are satisfied for $\mathcal{I}$ and $\mathbf{\Delta}$. Thus, Theorem 3.4 is applicable and we have solutions of the incremental problem (SRIP) which are stable and satisfy the discrete energy inequality. The compactness condition (3.6) for $\mathcal{P} \subset Y \subset X$ is also fulfilled and thus Corollary 3.6 allows us to extract a limit function as a suitable candidate for the solution of (SRCP). The final existence theorem has two additional assumptions. The first is the weak continuity of the mapping $\theta \mapsto \partial_t\mathcal{I}(t,\cdot) = \int_\Omega g(t,x)\theta(x)\,\mathrm{d}x$ which clearly holds.

The second condition is the weak closedness of the set of stable states

$$\mathcal{S}(t) = \{\, \theta\in\mathcal{P} \,\Big|\, \mathcal{I}(t,\theta) \leq \mathcal{I}(t,\zeta) + \mathbf{\Delta}(\zeta{-}\theta) \text{ for all } \zeta\in\mathcal{P} \,\}.$$

It is here where we need to restrict our analysis to the case where L is positive semi–definite. We do not believe that this condition is really needed, but we were unable to find a proof for the general case. Recall that convexity of $\mathcal{I}(t,\cdot)$ does not imply convexity of $\mathcal{S}(t)$, we rather have to use the theory of pseudo–differential operators or the H–measure to obtain the desired result.

Proposition 6.4 *If in addition to the above assumptions we also have $L \geq 0$, then the set of stable states $\mathcal{S}(t)$ is weakly closed for all $t \in [0,T]$.*

Proof. For simplicity we omit the dependence on t throughout the proof.

As $L \geq 0$ the energy $\mathcal{I}(\cdot)$ is convex, and by the standard theory of variational inequalities (Proposition 3.9) we have $\theta \in \mathcal{S}$ if and only if $\mathrm{D}\mathcal{I}(\theta)[\zeta{-}\theta] + \mathbf{\Delta}(\zeta{-}\theta) \geq 0$ for all $\zeta \in \mathcal{P}$. (In the nonconvex case this is necessary but not sufficient.) Written as integrals this gives

$$\int_\Omega [L\theta(x){+}\widehat{g}(x)](\zeta(x){-}\theta(x)){+}\Delta(\zeta(x){-}\theta(x))\,\mathrm{d}x \geq 0 \text{ for } \zeta\in\mathcal{P} = \mathrm{L}^1(\Omega,[0,1]). \qquad (6.4)$$

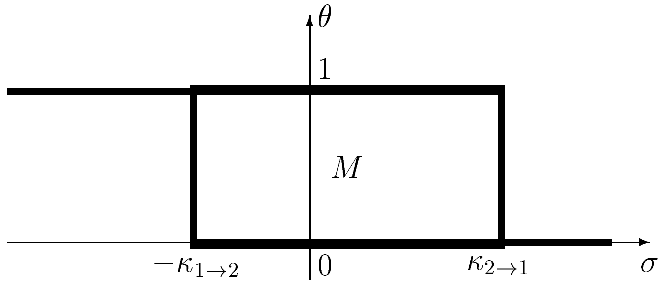

Figure 2: The set $M \subset \mathbb{R}^2$

This condition has the big advantage that the test function ζ appears only locally under the integral. Thus, we can vary ζ pointwise and obtain that $\theta \in \mathcal{S}$ if and only if

$$(\theta(x), (L\theta)(x) + \widehat{g}(x)) \in M \quad \text{for a.e. } x \in \Omega, \tag{6.5}$$

where $M \subset \mathbb{R}^2$ is defined via

$$M = \left\{ (\theta, \sigma) \in [0,1] \times \mathbb{R} \,\middle|\, (\sigma - \kappa_{2\to1})\theta \le 0,\ (\sigma + \kappa_{1\to2})(\theta - 1) \le 0 \right\}$$

and is depicted in Figure 2. Note that M is not convex; however roughly spoken, the proof works since intersections of M with lines of the form $(\theta, \ell\theta + \widehat{g})$ with $\ell \ge 0$ are convex.

Now consider a sequence $\theta_k \in \mathcal{S} \subset \mathcal{P}$ which satisfies $\theta_k \rightharpoonup \theta_*$ in $Y = \mathrm{L}^2(\Omega)$. Clearly we then have $L\theta_k + \widehat{g} \rightharpoonup L\theta_* + \widehat{g}$. The following lemma shows the implication

$$\theta_k(L\theta_k) \rightharpoonup w \text{ in } X = \mathrm{L}^1(\Omega) \quad \Longrightarrow \quad \theta_*(x)(L\theta_*)(x) \le w(x) \text{ for a.e. } x \in \Omega. \tag{6.6}$$

(Convexity only implies the much weaker statement $\int_\Omega w \,\mathrm{d}x \ge \int_\Omega \theta_*(L\theta_*)\,\mathrm{d}x$.) Thus, we are able to control the sign of the quadratic expressions which appear in the definition of M. By (6.5) we know that $\theta_k \in \mathcal{S}$ is equivalent to

$$\begin{aligned} &\theta_k(x) \in [0,1], \quad \theta_k(x)(L\theta_k)(x) \le (\widehat{g}(x) + \kappa_{2\to1})\theta_k(x), \\ &\theta_k(x)(L\theta_k)(x) \le (\widehat{g}(x) - \kappa_{1\to2})(\theta_k(x) - 1) + (L\theta_k)(x) \text{ for a.e. } x \in \Omega. \end{aligned}$$

Choosing a subsequence (k_l) such $\theta_{k_l}(L\theta_{k_l})$ converges to w weakly in X, we conclude for the limit θ_* the estimates

$$\begin{aligned} &\theta_*(x) \in [0,1], \quad w(x) \le (\widehat{g}(x) + \kappa_{2\to1})\theta_*(x), \\ &w(x) \le (\widehat{g}(x) - \kappa_{1\to2})(\theta_*(x) - 1) + (L\theta_*)(x) \text{ for a.e. } x \in \Omega. \end{aligned}$$

Together with $\theta_*(L\theta_*) \le w$ this implies $\theta_* \in \mathcal{S}$. ■

Theorem 6.1 is established if the implication (6.6) is proved, which is the content of the following lemma.

Lemma 6.5 *Let L be a scalar pseudo–differential operator of order 0 with nonnegative symbol ℓ. Assume that for $k \to \infty$ we have $\theta_k \rightharpoonup \theta_*$ in $\mathrm{L}^p(\Omega)$ and $\theta_k(L\theta_k) \rightharpoonup w$ in $\mathrm{L}^{p/2}(\Omega)$, where $p \ge 2$. Then $\theta_*(L\theta_*) \le w$ a.e. in Ω.*

Proof. We define $v_k = \theta_k - \theta_*$ such that $v_k \rightharpoonup 0$. Then it suffices to show that $v_k(Lv_k) \rightharpoonup z$ implies $0 \leq z$.

There are two closely related ways to prove this result. The first uses the H–measure as developed in [Tar90]: we apply Corollary 1.12 of [Tar90] to the sequences $U_{1,k} = U_{2,k} = v_k$ and the operators $A_1 = \mathrm{id}$ and $A_2 = L$ and use the positivity of the diagonal entries of the H–measure.

We describe the second proof in more detail. It is based solely on the theory of pseudo–differential operators. Let $\phi \geq 0$ be an arbitrary localization function and denote by M_ϕ the multiplication operator $v \mapsto \phi v$. Then $M_\phi L = LM_\phi + K_\phi$ where K_ϕ is a compact operator depending only on ϕ (see e.g. [Tar90, Lemma 1.7] and [Th00b]).

For a sequence $(v_k)_{k\in\mathbb{N}}$ with $v_k \rightharpoonup 0$ and $v_k\, Lv_k \rightharpoonup z$ one finds that

$$\begin{aligned}\textstyle\int_\Omega \phi^2\, z \,\mathrm{d}x &= \textstyle\lim_{k\to\infty} \int_\Omega \phi^2\, v_k\, (Lv_k)\,\mathrm{d}x = \lim_{k\to\infty} \int_\Omega (M_\phi v_k)\, M_\phi(Lv_k)\,\mathrm{d}x \\ &= \textstyle\lim_{k\to\infty} \int_\Omega (M_\phi v_k)\, L(M_\phi v_k)\,\mathrm{d}x + \lim_{k\to\infty} \int_{\mathbb{R}^d} (M_\phi v_k)\,(K_\phi v_k)\,\mathrm{d}x \geq 0.\end{aligned}$$

In the last estimate we have used $L \geq 0$ for the first term. The second term is 0 by the compactness of K_ϕ. Since ϕ was arbitrary, this gives the desired result. ∎

Remark 6.6 *With the above theory we can easily construct cases where the operator L is strictly positive. If we assume that $\Gamma = \partial\Omega$ (condition (a) from above), then Fourier transform shows that*

$$\langle \theta, L\theta\rangle = \int_\Omega \theta\,(L\theta)\,\mathrm{d}x = \int_{\mathbb{R}^d} \ell(\omega)|(\mathcal{F}\theta_{\mathrm{ext}})(\omega)|^2\,\mathrm{d}\omega \ \textit{with}\ \theta_{\mathrm{ext}}(x) = \begin{cases} \theta(x) & \textit{for } x \in \Omega, \\ 0 & \textit{for } x \in \mathbb{R}^d \setminus \Omega. \end{cases}$$

This implies $\langle \theta, L\theta\rangle > 0$ for all $\theta \neq 0$ if and only if $\ell(\omega) \neq 0$ for almost every $\omega \in \mathbb{R}^d$. Since $\ell|_{\mathbb{S}^{d-1}}$ is analytic the latter condition holds whenever ℓ is not identically 0.

The case $\ell \equiv 0$ may occur in degenerate or highly symmetric cases. Recall that ℓ always vanishes along at least one straight line. If now the elastic tensor is isotropic and $A_2 - A_1$ is a multiple of identity, then ℓ must be rotationally symmetric and hence vanishes everywhere.

References

[AcF84] E. Acerbi, N. Fusco. Semicontinuity problems in the calculus of variations. *Archive Rational Mechanics Analysis.* ***86*** (1984) 125–145.

[ACZ99] J. Alberty, C. Carstensen, D. Zarrabi. Adaptive numerical analysis in primal elastoplasticity with hardening. *Comp. Meth. Appl. Mech. Eng.* **171** (1999) 175–204.

[BaJ87] J.M. Ball, R.D. James. Fine phase mixtures as minimizers of energy. *Arch. Rational Mech. Analysis* **100** (1987) 13–52.

[BaJ92] J.M. Ball, R.D. James. Proposed experimental tests of a theory of fine microstructure and the two–well problem. *Phil. Trans. Royal Soc. London* **A338** (1992) 389–450.

[BaP86] V. Barbu, T. Precupanu. *Convexity and Optimization in Banach Spaces*, D. Reidel Pub. Co, Dordrecht, 2nd edition, 1986.

[BCJ95] J.M. Ball, C. Chu, R.D. James. Hysteresis in martensite phase transformations. In "Proceedings of ICOMAT-95, *J. de Physique* IV **5**, colloque C8", pp 245–251.

[BoL97] Z. Bo, C. D. Lagoudas. Thermomechanical modeling of polycrystalline SMAs under cyclic loading, Part IV: Modeling of minor hysteresis loops. *Int. J. Eng. Sci.* (1997). Submitted.

[Bha93] K. Bhattacharya. Comparison of the geometrically nonlinear and linear theories of martensitic transformations. *Continuum Mech. Thermodyn.* **5** (1993) 205–242.

[BhK96] K. Bhattacharya, R.V. Kohn. Symmetry, texture and recoverable strain of shape–memory polycrystals. *Acta Metallurgica* **44** (1996) 529–542.

[BrS96] M. Brokate, J. Sprekels. *Hysteresis and Phase Transitions.* Springer–Verlag, 1996.

[CaP97] C. Carstensen, P. Plecháč. Numerical solution of the scalar double–well problem allowing microstructure. *Math. Comp.* **66** (1997) 997–1026.

[Cha99] J. Chapman. A hierarchy of model for type–II superconductors. Part I: Bulk superconductors. University of Oxford, Preprint 1999.

[CHM00] C. Carstensen, K. Hackl, A. Mielke. Nonconvex potentials and microstructures in finite–strain plasticity. In preparation, 2000.

[CoV90] P. Colli, A. Visintin. On a class of doubly nonlinear evolution equations. *Comm. Partial Diff. Eqns.* **15** (1990) 737–756.

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

[Gas58] S.I. Gass. *Linear Programming. Methods and Applications.* McGraw–Hill 1958.

[Gal60] D. Gale. *The Theory of Linear Economic Models.* McGraw–Hill 1960.

[GNS83] P. Germain, Q.S. Nguyen, P. Suquet. Continuum Thermodynamics. *Trans. ASME–Applied Mechanics* **50** (1983) 1010–1020.

[GhO94] G. Ghosh, G.B. Olson. Kinetics of F.C.C. $\to$ B.C.C. heterogeneous martensitic nucleation. I. The critical driving force for athermal nucleation. *Acta Metall. Mater.* **42** (1994) 3361–3370.

[GoM00] S. Govindjee, A. Mielke. The free–energy of mixing for evolving microstructure in matrensitic materials by quasi–convex analysis. In preparation, 2000.

[GrK95] Y. Grabovsky, R.V. Kohn. Microstructures minimizing the energy of a two phase elastic composite in two space dimension. I: the confocal ellipse construction. And II: The Vidgergauz microstructure. *J. Mech. Physics Solids* **43** (1995) 933–947 and 949–972.

[Grinfeld91] Grinfeld, M. A. (1991). *Thermodynamic Methods in the Theory of Heterogeneous Systems*, Longman, Sussex.

[HaG99] G.J. Hall, S. Govindjee. A model and numerical framework for the simulation of solid–solid phase transformation. Preprint UCB/SEMM–1999/11, Univ. Calif. Berkeley, 1999.

[HaN75] B. Halphen, Q.S. Nguyen. Sur les matériaux standards généralisés. *J. Mécanique* **14**, No 1 (1975) 39–63.

[HaR95] W. Han, D. Reddy. Computational plasticity and numerical analysis. *Comp. Mech. Adv.* *2* (1995) 283–400.

[HuM93] Y. Huo, I. Müller. Nonequilibrium thermodynamics of pseudoelasticity. *Continuum Mech. Thermodyn.* **5** (1993) 163–204.

[KaR88] I.M. Kaganova, A.L. Roitburd. Equilibrium of elastically interacting phases. *Sov. Physics JETP* **67** (1988) 1174–1186.

[Kha67] A.G. Khachaturyan. Some questions concerning the theory of phase transformations in solids. *Soviet Physics: Solid State* **8** (1967) 2163–2168.

[Kha83] A.G. Khachaturyan. *Theory of Structural Transformations in Solids.* John Wiley and Sons 1983.

[KLMS97] M.S. Kuczma, V.I. Levitas, A. Mielke, E. Stein. Nonisothermal hysteresis loops in pseudoelasticity. In "Proceedings of the XIII Conference on *Computer Methods in Mechanics*, Poznan, 5–8 May 1997", pp. 711–718.

[KMS99] M.S. Kuczma, A. Mielke, E. Stein. Modeling of hysteresis in problems with phase transformations. *Archive of Mechanics* **51**(6)(1999) 693–715.

[Koh91] R.V. Kohn. The relaxation of a double–well energy. *Continuum Mech. Thermodyn.* **3** (1991) 193–236.

[KuM00] M.S. Kuczma, A. Mielke. Influence of hardening and inhomogeneity on internal loops in pseudoelasticity. *Zeits. Angew. Math. Mechanik* **80** (2000) 291–306.

[LeS97] V.I. Levitas, E. Stein Simple micromechanical model of thermoelastic martensitic transformations, *Mech. Res. Commun.* **24** (1997) 309–318.

[Lev94] Levitas V.I.: [1994] Thermomechanical description of pseudoelasticity – the threshold–type dissipative force with discrete memory. Mech. Res. Commun., Vol. 21, No. 3, pp. 273–280.

[Lev95] V.I. Levitas. The postulate of realizability: formulation and applications to post–bifurcation behavior and phase transitions in elastoplastic materials. *Int. J. Eng. Sci.* **33** (1995) 921–971.

[Lev97] V.I. Levitas. Phase transitions in elastoplastic materials: continuum thermomechanical theory and examples of control. Part I and II. *J. Mech. Physics Solids* **45** (1997) 923–947 and 1203–1222.

[Lev98] V.I. Levitas. Thermomechanical theory of martensitic phase transformations in inelastic materials. *Int. J. Solids Structures* **35** (1998) 889–940.

[Lev00] V.I. Levitas. Structural changes without stable intermediate state in inelastic material. Part I and II. *Int. J. Plasticity* **16**, No. 7-8, (2000) 805-849 and 851-892.

[LIS99] V.I. Levitas, A.V. Idesman, E. Stein. Shape memory alloys: micromechanical modeling and numerical analysis of structures. *J. Intelligent Mat. Systems Structures* **10** (1999).

[Mie00] A. Mielke. Estimates on the mixture function for multiphase problems in elasticity. In "Proceedings of the Conference on *Multifield Problems*, Stuttgart Oct. 6-8, 1999; Springer–Verlag 2000". To appear.

[MiT00] A. Mielke, F. Theil. Convex analysis and rate independent processes. In preparation.

[MTL98] A. Mielke, F. Theil, V.I. Levitas. Mathematical formulation of quasistatic phase transformations with friction using an extremum principle. Preprint A8, Institut für Angewandte Mathematik, Universität Hannover, September 1998.

[Mül98] S. Müller. Variational models for microstructure and phase transitions. MPI Leipzig. Lecture Notes no. 2 (1998) 127 pp.

[MüS96] I. Müller, S. Seelecke. Thermodynamic aspects of shape memory. Submitted to *Trans Tech Publications* (1996).

[OlC86] G.B. Olson, M. Cohen. Dislocation theory of martensitic transformations. In "*Dislocations in Solids*, F.R.N. Nabarro (edr), Elsevier Science Publishers B. V. (1986)," pp. 297–407.

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

[PAM98] S. Pagano, P. Alert, O. Maisonneuve. Solid–solid phase transition modelling. Local and global minimizations of non–convex and relaxed potentials. Isothermal case for shape memory alloys. *Int. J. Eng. Sci.* **36** (1998) 1143–1172.

[RaS99] K.R. Rajagopal, A.R. Srinivasa. On the thermomechanics of shape memory wires. *Zeits. angew. Math. Physik* **50** (1999) 459–496.

[Roi67] A.L. Roitburd. *Kristallografia* **12** (1967) 567ff (in Russian).

[Rou97] T. Roubíček. *Relaxation in Optimization Theory and Variational Calculus.* De Gruyter, 1997.

[SmW98] V.P. Smyshlyaev, J.R. Willis. On the relaxation of a three–well energy. *Proc. Royal Soc. London* **A 455** (1998) 779–814.

[SP*99] N. Siderey, E. Patoor, M. Berveiller, A. Eberhardt. Constitutive equations for polycrystalline thermoelastic shape memory alloys. Part I. Intragranular interactions and behavior of the grain. *Int. J. Solids Structures* **36** (1999) 4289–4315.

[Tar90] L. Tartar. H–measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations. *Proc. Royal Soc. Edinb.* **A 115** (1990) 193–230.

[Th00a] F. Theil. Approximation of quasistatic processes with friction by incremental problems in the theory of phase transitions. In "*Trends in Applications of Mathematics to Mechanics*, G. Iooss, O. Guès, A. Nouri (eds), Chapman & Hall/CRC, 2000", pp. 172–179.

[Th00b] F. Theil. Relaxation of rate independent evolution problems. *Proc. Roy. Soc. Edinb.* **A** (2000). Submitted.

[Vis94] A. Visintin. *Differential Models of Hysteresis.* Springer–Verlag 1994.

[Wil93] K. Wilmanski. Models of stress–strain hysteresis loops in shape memory alloys. *Int. J. Eng. Sci.* **31**(8) (1993) 1121–1138.

[Zie63] H. Ziegler. Some extremum principles in irreversible thermodynamics with application to continuum mechanics. *Progress in solid mechanics, Vol IV*, North Holland Publ. Comp., Amsterdam 1963.