Diplomarbeit

Reverse approximation of rate-independent evolution processes

Institut für Mathematik
Technische Universität Berlin

Betreuer: Professor Dr. Petra Wittbold

Johan Filip Rindler

Berlin, den 24. September 2008

Abstract

Rate-independent evolution systems are a class of nonlinear evolutionary problems which lie in between certain pseudoparabolic partial differential equations and (stationary) variational problems. They incorporate both elastic (reversible) and dissipative (non-reversible) effects and have widespread applications in the material sciences, most notably in elastoplasticity, shape-memory alloys and ferroelectricity. Such systems display dynamic effects only through the change of an external loading (force, electric field, ...) and are characterized by the property of rate-independence. This means that transforming the dynamics of the external loading (e.g. stretching or compressing the time development) transforms all solutions by the same law. A modern approach to such problems has been developed in recent years focusing on the notion of "energetic solutions".

This thesis aims to investigates the relationship between this energetic formulation and its time-discrete counterparts. Such time-discretizations have been used to construct energetic solutions, but it is also possible to reverse this process and investigate how one can get solutions to the time-discrete problems from a given energetic solutions. These investigations are of interest in relaxation theory, numerical approximations and also give more physical insight into rate-independent problems, since the time-discrete problem is often more physically relevant than the time-continuous problem.

In this work, such "reverse approximation" theorems are proved both for a single problem and Γ-converging sequences thereof. To this aim, some generalized results in the theory of Γ-convergence are presented. As a complement to the positive reverse approximation results, couterexamples are given to show the optimality of the obtained assertions.

Besides this principal aim, short overviews of the theory of Γ-convergence and of the theory of energetic solutions to rate-independent problems are presented. Most notably, the complete existence proof for the rate-independent theory is carried through in the case of Γ-converging problems. Also, uniquness is briefly touched upon.

In a separate course of thought, contraction properties of sublevel sets are investigated and some results on their behavior are devised. In conjunction with the reverse approximation results, they also imply a new stability estimate on energetic solutions in the case of nonunique solvability.

Finally, the obtained reverse approximation results are applied to a concrete energy functional rooted in the theory of phase transitions and with a double-well structure.

Zusammenfassung

Ratenunabhängige Evolutionssysteme bilden eine Klasse von nichtlinearen Evolutionsproblemen, welche zwischen gewissen pseudoparabolischen partiellen Differentialgleichungen und (stationären) Variationsproblemen angesiedelt sind. Sie beinhalten sowohl elastische (reversible), als auch dissipative (nicht reversible) Effekte und haben viele Anwendungen in den Materialwissenschaften, insbesondere bei Elastoplastizität, Formgedächtnislegierungen und Ferroelektrizität. Solche Systeme zeigen dynamische Effekte nur bei Änderung einer äußeren Last (Kraft, elektrisches Feld, ...) und sind durch die Eigenschaft der Ratenunabhängigkeit charakterisiert. Diese bedeutet, dass eine Transformation des zeitlichen Verlaufs der äußeren Last (z.B. durch Dehnen oder Stauchen der Zeit) auch die Lösungen in gleicher Weise transformiert. Für solche Probleme wurde in jüngster Zeit ein neuer Zugang entwickelt, der auf dem Begriff der "Energielösung" fußt.

Diese Arbeit möchte die Beziehung zwischen dieser energetischen Formulierung und ihrem zeitdiskreten Gegenstück untersuchen. Solche Zeitdiskretisierungen werden verwendet, um Energielösungen zu konstruieren, aber auch der umgekehrte Weg ist möglich und man kann untersuchen, wie sich Lösungen der Zeitdiskretisierung aus einer gegebenen Energielösung gewinnen lassen. Diese Untersuchungen sind von Interesse in der Relaxationstheorie, in der Theorie der numerischen Approximation und geben auch zusätzliches physikalisches Verständnis ratenunabhängiger Probleme, da das zeitdiskrete Problem oft das physikalisch relevantere ist.

In dieser Arbeit werden Sätze über die "Rückwärtsapproximation" sowohl für ein einzelnes, als auch für eine Folge von Γ-konvergenten Problemen bewiesen. Zu diesem Ziel werden einige verallgemeinerte Resultate aus der Theorie der Γ-Konvergenz dargelegt. Als Ergänzung zu den positiven Resultaten zur Rüchwärtsapproximation werden Gegenbeispiele vorgeführt, die die Optimalität der erzielten Resultate belegen.

Neben diesem erstrangigen Ziel wird auch ein kurzer Überblick über die Theorie der Γ-Konvergenz und der Theorie der Energielösungen zu ratenunabhängigen Systemen gegeben. Insbesondere wird der vollständige Existenzbeweis der ratenunabhängigen Theorie im Falle von Γ-konvergenten Problemen präsentiert.

Als getrennter Gedankengang werden zunächst Kontraktionseigenschaften von Subniveaumengen untersucht und einige Resultate über deren Verhalten entwickelt. In Verbindung mit den Resultaten über die Rückwärtsapproximation implizieren diese dann eine neue Stabilitätsabschätzung im Falle von nicht eindeutigen Lösungen.

Abschließend werden die gewonnenen Resultate zur Rückwärtsapproximation auf ein konkretes Energiefunktional mit einer Doppelmuldenstruktur, welches aus der Theorie der Phasenübergänge entlehnt ist, angewandt.

gewidmet

Jörg Kraftmeier (4.8.1942–12.6.2005)

Preface

> Das Denken gehört zu den größten Vergnügungen der menschlichen Rasse.
>
> *(Bertolt Brecht)*

This thesis originated from a project undertaken under the supervision of Professor Dr. Alexander Mielke of Humboldt-Universität zu Berlin, and Weierstraß Institut für angewandte Analysis und Stochastik (WIAS), Berlin. He gave me the opportunity to work on a "Project in the Berlin Mathematical School", which then was converted into this thesis.

Professor Dr. Mielke vividly supported my interest to investigate the particular direction presented in this thesis and spent a lot of time discussing all kinds of mathematical topics with me. For all this, I wish to express my sincere thanks. I am also grateful for the possibility to present my work on the Autumn School of the Research Training Group 1128 "Analysis, Numerics, and Optimization of Multiphase Problems" at the ÚTIA in Prague.

I also wish to thank my supervisor, Professor Dr. Petra Wittbold of Technische Universität Berlin for supporting this thesis project and many encouraging and interesting discussions.

Personally, I wish to express my deepest gratitude towards my girlfriend Claudia Sas who was not only very tolerant and understanding when I worked first on the project and then on this thesis, but always has been a fountain of encouragement and support. Also, I wish to thank my mother Karin Rindler and her partner Michael Reinelt, who always supported me with intellectual stimulation and much encouragement. Finally, much of my curiosity into science was stimulated by my friend Dr. Günter von Häfen with whom I had the pleasure to discuss so many different mathematical, physical and other scientific matters throughout my schooltime.

I hope that some of the joy and excitement that I experienced while working on this project is reflected in this thesis.

Contents

Introduction

Rate-independence is a fascinating property of nonlinear evolutionary systems—both from a theoretical and from a practical point of view. Such systems have many applications in the material sciences, most notably in elastoplasticity [HR99, OR99, CHM02], phase transitions in shape-memory alloys [AP02, MTL02, AS05, KMR05], crack formation in brittle materials [FM98, DFT05] and ferroelectricity and -magnetism [MB89, KW03, MT06]. Rate-independent systems have two decisive (related) features: First, they are "quasi-static" meaning that they do not possess own dynamics, but only react to changes of an external loading (e.g. applied forces). In particular, if the external conditions do not change, the system rests in static equilibrium. The second decisive property is their behavior under time-reparametrizations: If the exterior loading is reparametrized, the solutions corresponding to this new loading also are reparametrizations of the old solutions, reparametrized by the same law. This means that the system is characterized by its *path* through the state space alone, and not by its *speed*.

Usually, such systems are formulated as an evolutionary doubly-nonlinear differential inclusion on a (state) Banach space $\mathcal{Q}$ as

$$0 \in \partial\mathcal{R}(\dot{q}(t)) + \mathrm{D}\mathcal{E}(t, q(t)) \qquad (\text{in } \mathcal{Q}^*) \qquad \text{for almost all } t \in [0,T], \tag{DI}$$

or the equivalent evolutionary variational inequality

$$\langle \mathrm{D}\mathcal{E}(t, q(t)), v - \dot{q}(t)\rangle + \mathcal{R}(v) - \mathcal{R}(\dot{q}(t)) \geq 0 \tag{EVI}$$

for a (sought) process $q : [0,T] \to \mathcal{Q}$. Here, the functional $\mathcal{E} : [0,T] \times \mathcal{Q} \to \mathbb{R}_\infty$ models elastic (reversible, potential) energy storage and $\mathcal{R} : \mathcal{Q} \to [0,\infty]$ is a convex dissipation potential. In this formulation, rate-independence manifests itself as the algebraic property of positive 1-homogeneity of $\mathcal{R}$. Similarly structured (but not necessarily rate-independent) problems have for example been studied in [CV90] and as "pseudoparabolic equations" in [GGZ74, V.§2]. The book [Rou05] considers many different types of these doubly-nonlinear problems in Chapter 11, where it is also remarked that most traditional results do not cover the rate-independent case with $\mathcal{R}$ 1-homogeneous.

A whole new theory based on so-called "energetic solutions" was introduced by A. Mielke and F. Theil in the works [MT99, MT04] and further developed for example in [MM05, FM06]; a recent survey is [Mie05]. Their method replaces (DI) and (EVI) with an integrated energy balance

$$\mathcal{E}(t, q(t)) + \int_0^t \mathcal{R}(\dot{q}(\tau))\,\mathrm{d}\tau = \mathcal{E}(0, q(0)) + \int_0^t \partial_t\mathcal{E}(\tau, q(\tau))\,\mathrm{d}\tau, \tag{E}$$

which is to be satisfied for all $t \in [0,T]$, and the stability condition

$$\mathcal{E}(t,q(t)) \le \mathcal{E}(t,\hat{q}) + \mathcal{D}(q(t),\hat{q}) \qquad \text{for all } \hat{q} \in \mathcal{Q},\ t \in [0,T]. \tag{S}$$

Under certain convexity and smoothness conditions on the involved functionals, the different formulations are equivalent, but the framework of energetic solutions allows for much more general systems, in particular ones with non-potential dissipation and a nondifferentiable stored energy functional (also, the solution processes need not be differentiable in the sense of Sobolev spaces).

One recurring object in the study of energetic solutions to rate-independent systems is the associated time-incremental problem on a partition $(0 = t_0, \ldots, t_N = T)$.

$$\begin{cases} \text{For } j = 1, \ldots, N \text{ inductively find } q_j \in \mathcal{Q} \text{ such that} \\ y_j \in \operatorname{Argmin} \left\{ \mathcal{E}(t_j, \hat{q}) + \mathcal{R}(\hat{q} - q_{j-1}) \ : \ \hat{q} \in \mathcal{Q} \right\}, \end{cases} \tag{IP}$$

or weaker variants thereof involving only approximate minimizers. In the proofs of the energetic existence results, one uses (IP) as a temporal semi-discretization to construct solutions of (S) & (E) by considering interpolants of discrete solutions to (IP) for finer and finer partitions and passing to the limit (in this context, a temporal semidiscretization is also called a Rothe method).

Since rate-independent problems are highly nonlinear, we cannot expect unique solvability. The aim of this work is to better understand the solution set of such rate-independent systems by showing that all energetic solutions in a certain sense "originate" from (IP) (or one of its variants). This can be seen as going the opposite direction as in the existence proof and is therefore called "reverse approximation". The relevance of this question really is a consequence of the nonuniqueness of solutions, since if there was only one solution, this one would be the one constructed in the existence proof and reverse approximation would be trivial.

When numerically calculating solutions to rate-independent problems as in [MR06], it is an interesting question whether *all* solutions of the problem can be approximated. Trivially, if the problem is uniquely solvable, this is clear by the standard existence theory as then all approximate solutions must converge to the solution of (S) & (E). Reverse approximation shows that even in the case of nonuniqueness, every solution can be approximated. Of course, for this to be relevant in practice, the numerical method has to be able to cope with the nonunique solvability of the problem.

A further application of the reverse approximation theory to optimal control problems for rate-independent systems, see [Rin08b], is under investigation in [Rin08a].

Recently, in [MRS08] sequences of rate-independent systems have been studied in the framework of Γ-convergence. In the study of phase transitions, for example, one is faced with the difficulty that the original minimization problem might admit any distribution of phases as mathematical solutions, but of course, not all such partitions are physically feasible. Rather, only partitions that satisfy some additional constraints on the minimality of surface tension will occur in nature. This can be modelled by setting up a sequence of approximate energy functionals which reflect an increasingly weaker influence of surface tension. Assuming that these approximate energy functionals Γ-converge to the original energy functional, one can then study convergence of the microstructure solutions to a solution of the original problem. Conversely, approximability by microstructure solutions can be chosen as a selection criterion on solutions of the original (limit) problem. In the stationary case the situation is well-studied and a rather complete theory can be set up in the framework of Γ-convergence, cf. Chapter 7 of [Dal93] and Chapter 6 of [Bra02]. For evolutionary problems, however, only very little is known. Partly, this is due to the fact that Γ-convergence is a purely static concept and no fully equivalent theory of a variational convergence for time-dependent problems has been developed so far. See, however, [SS04] for gradient flows and [Mie08] for semilinear Hamiltonian systems.

The work [MRS08] discussed under which conditions solutions of the approximate problems admit a limit point solving the limit problem. As for single problems, one can be interested in the reverse question: To a given solution to (S_∞) & (E_∞) can one find solutions to the corresponding approximate incremental problems (IP)? This question is answered positively (up to the fact that we have to allow approximate minimizers) in this work. Such studies draw their principal motivation from the increased understanding between the coupling of time-discrete and time-continuous problems they provide. This is, for example, of interest in relaxation theory and numerical approximations. Also, the time-discrete problem is from a certain physical point of view the more natural point to start with and hence physical insight might be gained from such investigations.

The text is parted into four chapters. The first chapter starts by introducing Γ-convergence and proving some of its main properties. This theory is standard and the presentation mainly follows [Bra02], but some material has been rearragend and some different proofs are given. Then, the chapter moves on to more special results which constitute the basis for the developments to come in later chapters. First, results on sets of approximate minimizers are given which are presumably new, but related to the corresponding results for true minimizers in [Dal93]. Most of this section is taken from [MR07a]. Second, contraction properties of minimizer sets of convex and strictly convex functionals are investigated. In particular, some results on the behavior of the diameter of sets of approximate minimizers are shown. These results should already be known, but the author is not aware of any references containing them.

Chapter 2 moves to the heart of the matter and introduces rate-independent systems, starting in an abstract algebraic fashion and using a little bit of terminology from category theory. Then, more concrete formulations using differential inclusions and variational inequalities are studied before the energetic approach is presented in (more or less) full generality and its equivalence to the traditional formulations is shown in the case of convex problems. Also, the time-incremental problem and its variants are introduced. After many preparatory results, most notably a generalized Helly selection principle, the main existence theorem of the theory (for a sequence of systems) is proved. The presentation of proof follows [MRS08], but again some rearrangement has taken place. Finally, we briefly touch on uniqueness issues.

The third chapter contains the main contributions of this work: The reverse approximation theorems for single systems and for sequences thereof. These sections are taken from [MR07a]. As a side track, some contraction properties of the incremental problems are shown in the third section (based on the results from the first chapter). Then, also using the reverse approximation results, a new stability estimate on the maximal distance between two solutions to the same initial value is given in terms of a "modulus of contraction" of the approximate minimizer sets of the energy storage functional.

The last chapter applies the presented theory to a particular functional and presents reverse approximation in relation to relaxation and regular approximation.

1 Γ-convergence

This chapter gives a short introduction to Γ-convergence and presents some of its main properties among which the convergence of minima and of the corresponding minimizers is the most important for this thesis (albeit not used directly). After the basic definitions of Γ-convergence in Section 1.1 and the aforementioned study of minima in Section 1.2, we move on to examine in more detail the behavior of approximate and true minimizers in Section 1.3. These not only serve as illustrations for the concept of reverse approximation, but also comprise important building blocks for the much more involved situation of reverse approximation for (sequences of) rate-independent systems, which is developed in Chapter 3. Finally, as a side track, Section 1.4 discusses contraction properties of approximate minimizer sets.

1.1 Definition of Γ-convergence

The theory of Γ-convergence was founded by Ennio De Giorgi in the 1970s (based on earlier works of Umberto Mosco and others) and is widely used nowadays in the Calculus of Variations. It is a convergence of functionals which implies convergence of minima and ensures that the limit is lower semicontinuous. The book [Bra02] is an illustrative introduction to Γ-convergence with many applications to homogenization theory, phase transition and free discontinuity problems. The encyclopedic work [Dal93] is a thorough and detailed treatise of the subject (despite denying this in its preface). In this section, most of the results are modified versions of the ones in [Bra02].

In the following, let $(\mathcal{X}, d)$ be a metric space. Later, when we deal with convexity, we will also require $\mathcal{X}$ to be a linear space, but this will then be stated explicitly. In fact, in subsequent chapters we will work in arbitrary Hausdorff topological spaces that might not be first or second countable and topological and sequential concepts may differ. Still, in the Calculus of Variations, sequences are also employed in these contexts. Following this philosophy, the concepts introduced in this chapter are taken as "sequential" versions and this suffices for our purposes[1].

[1] This is so partly because of the following considerations (also cf. [Con90]): In finite-dimensional vector spaces, there is only one (Hausdorff) topology turning the space into a topological vector space and in separable, reflexive Banach spaces the weak topology is metrizable on norm-bounded sets. If one uses the weak*-topology on a dual space with a separable primal space, then norm-bounded sets in the dual space are also metrizable. So in all these cases, one can work with sequences (under suitable coerciveness assumptions). The only case where one has to be careful is when one uses the

In this chapter—contrary to [Bra02]— all results have "sequential" proofs as to make them applicable in nonmetrizable spaces. The book [Dal93] contains a full treatment of how one should define Γ-convergence in arbitrary topological spaces if one does not want to resort to sequences.

The main definition is:

Definition 1.1. *The functional*[2] $F_\infty : \mathcal{X} \to \mathbb{R}_\infty$ *is called the* **(sequential) Γ-limit** *of the functionals* $F_k : \mathcal{X} \to \mathbb{R}_\infty$ *(*$k \in \mathbb{N}$*) if it satisfies for all* $x \in \mathcal{X}$ *the following two conditions:*

(i) *For all sequences* $(x_k)_k \subseteq \mathcal{X}$ *with* $x_k \to x$*, the* **lim inf-inequality** *holds:* $F_\infty(x) \le \liminf_{k\to\infty} F_k(x_k)$.

(ii) *There exists a* **recovery sequence** $(x_k)_k \subseteq \mathcal{X}$*, i.e.* $x_k \to x$ *and* $F_\infty(x) = \lim_{k\to\infty} F_k(x_k)$.

The Γ*-limit of the sequence* $(F_k)_k$ *is denoted by* $\Gamma\text{-}\lim_k F_k$.

Using the lim inf inequality (i), condition (ii) in the definition of Γ-convergence can be replaced by the weaker condition that there exists a sequence $x_k \to x$ such that

(ii') $F_\infty(x) \ge \limsup_{k\to\infty} F_k(x_k)$.

This is then called the **lim sup-inequality**.

Often useful is the following generalization of Γ-limit:

Definition 1.2. *The* **(sequential) Γ-limes inferior** $\Gamma\text{-}\liminf_k F_k$ *and the* **(sequential) Γ-limes superior** $\Gamma\text{-}\limsup_k F_k$*, respectively, of the sequence of functionals* $F_k : \mathcal{X} \to \mathbb{R}_\infty$ *(*$k \in \mathbb{N}$*) are the functionals defined for all* $x \in \mathcal{X}$ *through*

$$\Gamma\text{-}\liminf_k F_k(x) := \inf\{ \liminf_{k\to\infty} F_k(x_k) \ : \ x_k \to x \},$$
$$\Gamma\text{-}\limsup_k F_k(x) := \inf\{ \limsup_{k\to\infty} F_k(x_k) \ : \ x_k \to x \}.$$

The first lemma shows that if the Γ-limes inferior and Γ-limes superior are equal, then we have Γ-convergence.

Lemma 1.3. $\Gamma\text{-}\liminf_k F_k = \Gamma\text{-}\limsup_k F_k = F_\infty$ *if and only if* $F_\infty = \Gamma\text{-}\lim_k F_k$.

weak topology on a Banach space with nonseparable dual space (then the primal space cannot be both separable and reflexive), because then the weak topology might not be metrizable. This was demonstrated by Schur who showed that in the sequence space ℓ^1 (with nonseparable dual space ℓ^∞), weak convergence of *sequences* is equivalent to strong convergence, but the weak and strong *topologies* still differ, cf. [Con90, V.§5] for more details on such considerations.

[2] We here use "function" and "functional" synonymously since the distinction depends on whether one considers elements of $\mathcal{X}$ to be "points" or "functions".

Proof. Let first $\Gamma\text{-}\liminf_k F_k = \Gamma\text{-}\limsup_k F_k = F_\infty$. Then, for all $x_k \to x$ in $\mathcal{X}$ it holds that $F_\infty(x) = \Gamma\text{-}\liminf_k F_k(x) \le \liminf_{k\to\infty} F_k(x_k)$, i.e. the lim inf-inequality (i) is satisfied. On the other hand, for each $x \in \mathcal{X}$ and all $n \in \mathbb{N}$, let $x_k^n \to x$ in $\mathcal{X}$ $(k \to \infty)$ with

$$F_\infty(x) \ge \limsup_{k\to\infty} F_k(x_k^n) - \frac{1}{n};$$

such sequences always exist, because $F_\infty(x) = \Gamma\text{-}\limsup_k F_k(x)$. Now, inductively combine the sequences (x_k^n) into one sequence $\tilde{x}_k$ in the following way: Let $N(n) \in \mathbb{N}$, $n \in \mathbb{N}$, denote a growing sequence of indices such that

$$d(x_j^n, x) \le \frac{1}{n} \qquad \text{and} \qquad F_\infty(x) \ge F_j(x_j^n) - \frac{2}{n} \qquad \text{for all } j \ge N(n),$$

which is possible by the definition of the limes superior. Then set

$$(\tilde{x}_k)_k := (x^1_{N(1)}, x^1_{N(1)+1}, \dots, x^2_{N(2)}, x^2_{N(2)+1}, \dots, x^l_{N(l)}, x^l_{N(l)+1}, \dots)$$

and observe that $\tilde{x}_k \to x$ and $F_\infty(x) \ge \limsup_{k\to\infty} F_k(\tilde{x}_k)$; this is (ii').

For the converse implication, it suffices to note that by (i) and (ii) for all $x \in \mathcal{X}$

$$F_\infty(x) \le \Gamma\text{-}\liminf_k F_k(x) \le \Gamma\text{-}\limsup_k F_k(x) \le F_\infty(x),$$

since a recovery sequence is admissible in the definition of the Γ-limes superior. □

As an immediate consequence we have that the Γ-limit is unique.

If a sequence of functionals $(F_k)_k$ converges locally uniformly to some F_∞, i.e. for all $x \in \mathcal{X}$, there exists some open neighborhood $U \Subset \mathcal{X}$ of x such that $\|(F_k - F_\infty)|_U\|_\infty \to 0$ as $k \to \infty$, and if F_∞ is lower semicontinuous, then the F_k also Γ-converge to F_∞: The lim inf-inequality holds, because for any $x_k \to x$, $x_k \in U$ for k sufficiently large and hence

$$\liminf_{k\to\infty} F_k(x_k) \ge \liminf_{k\to\infty} F_\infty(x_k) - \lim_{k\to\infty} \|(F_k - F_\infty)|_U\|_\infty \ge F_\infty(x)$$

by the lower semicontinuity of F_∞. The lim sup-inequality holds for the constant recovery sequence. Pointwise convergence, however does not in general imply Γ-convergence:

Example 1.4. Set $\mathcal{X} = \mathbb{R}$ and for $k \in \mathbb{N}$ define

$$F_k(x) := \begin{cases} \pm 1 & \text{if } x = \pm 1/k, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad F_\infty(x) := \begin{cases} -1 & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then, $\Gamma\text{-}\lim_k F_k = F_\infty$. To verify this, on each open set $U \Subset \mathbb{R} \setminus \{0\}$, the convergence to the zero function is uniform and this together with the lower semicontinuity of F_∞

implies Γ-convergence on $\mathbb{R} \setminus \{0\}$ as seen above. For $x = 0$, we have $\Gamma\text{-}\lim F_k(0) = -1$ by taking the recovery sequence $x_k := -1/k$ (the lim inf-inequality is trivial).

This example shows that the Γ-limit does not necessarily coincide with the pointwise limit (which here is the zero function). In fact, even for the constant sequence consisting only of F_1, the Γ-limit is $-\delta_{-1}$ (the Kronecker delta) by a similar reasoning as above. Further, it is analogous to show that $\Gamma\text{-}\lim_k -F_k = F_\infty$, hence it might occur that $\Gamma\text{-}\lim_k -F_k \neq -\Gamma\text{-}\lim_k F_k$ and consequently also $\Gamma\text{-}\lim_k(F_k+G_k) \neq \Gamma\text{-}\lim_k F_k + \Gamma\text{-}\lim_k G_k$ is possible.

One property that makes Γ-convergence directly fitted for applications in the Calculus of Variations is that Γ-limits are always lower semicontinuous:

Proposition 1.5. *$F_\infty = \Gamma\text{-}\lim_k F_k$ is lower semicontinuous.*

Proof. We need to show $F_\infty(x) \leq \liminf_{k\to\infty} F_\infty(x_k)$ for all $x_k \to x$ in $\mathcal{X}$. Let $(x_k^j)_j$ for all $k \in \mathbb{N}$ be a recovery sequence for x_k, i.e. $x_k^j \to x_k$ as $j \to \infty$ and $F_\infty(x_k) = \lim_{j\to\infty} F_j(x_k^j)$. Also let $(j(k))_k$ be a strictly increasing sequence of indices with

$$d(x_k^{j(k)}, x_k) \leq \frac{1}{k} \qquad \text{and} \qquad \left| F_{j(k)}(x_k^{j(k)}) - F_\infty(x_k) \right| \leq \frac{1}{k}.$$

Define

$$\tilde{x}_l := \begin{cases} x_k^{j(k)}, & \text{if } l = j(k) \text{ for some } k, \\ x, & \text{otherwise.} \end{cases}$$

Then, $\tilde{x}_k \to x$ and the lim inf-inequality implies

$$F_\infty(x) \leq \liminf_{k\to\infty} F_k(\tilde{x}_k) \leq \liminf_{k\to\infty} F_{j(k)}(x_k^{j(k)}) = \liminf_{k\to\infty} F_\infty(x_k),$$

i.e. the lower semicontinuity of F_∞. □

It can be shown that also the Γ-limes inferior and the Γ-limes superior are lower semicontinuous, cf. [Bra02, Proposition 1.28].

One striking consequence of the preceding proposition is that in general the notion of Γ-convergence is not generated by a topology, because then necessarily the constant sequence would have to converge to the only element of the sequence, but for a not lower semicontinuous function, this is not the case. In spaces of lower semicontinuous functions, however, Γ-convergence is generated by a topology, cf. Chapter 10 of [Dal93] for details.

The following proposition is very useful, because it says that the Γ-limit of a constant sequence coincides with the **lower semicontinuous envelope**, which for a functional $F : \mathcal{X} \to \mathbb{R}_\infty$ is defined to be

$$\operatorname{lsc} F := \sup \{ G \,:\, G \leq F \text{ and } G \text{ lower semicontinuous} \}.$$

This result is a cornerstone for the applications of Γ-convergence in relaxation theory, cf. [Dal93, Chapter 3].

Proposition 1.6. $\Gamma\text{-}\lim F = \operatorname{lsc} F$.

Proof. For any functional $H : \mathcal{X} \to \mathbb{R}_\infty$, it always holds that $\Gamma\text{-}\limsup H \leq H$ since the constant sequence is admissible in the definition of the Γ-limes superior. If additionally H is lower semicontinuous, then even $H = \Gamma\text{-}\lim H$ since for all $x_k \to x$ in $\mathcal{X}$

$$H(x) \leq \inf \Big\{ \liminf_{k\to\infty} H(x_k) \ : \ x_k \to x \Big\} = \Gamma\text{-}\liminf H(x) \leq \Gamma\text{-}\limsup H(x) \leq H(x)$$

and thus $H = \Gamma\text{-}\lim H$ by Lemma 1.3. Even if H is not lower semicontinuous, we still have that $\Gamma\text{-}\lim H$ exists, since from the definition of the Γ-limes inferior we can for all $x \in \mathcal{X}$ and all $\varepsilon > 0$ find a sequence $x_k \to x$ such that $\lim_{k\to\infty} F(x_k) \leq \Gamma\text{-}\liminf_k F(x) + \varepsilon$. Inserting this sequence into the definition of the Γ-limes superior and letting $\varepsilon \to 0$, we get $\Gamma\text{-}\limsup_k F(x) \leq \Gamma\text{-}\liminf_k(x)$ and we can again invoke Lemma 1.3.

Consequently, for any lower semicontinuous $G \leq F$, we have

$$G = \Gamma\text{-}\lim G \leq \Gamma\text{-}\lim F$$

and by taking the supremum over all such G, we see $\operatorname{lsc} F \leq \Gamma\text{-}\lim F$.

To prove the converse inequality, it suffices to note that $\Gamma\text{-}\lim F$ is lower semicontinuous by Proposition 1.5 and $\Gamma\text{-}\lim F \leq F$ (this has already been shown at the beginning of the proof), because then it follows immediately from the definition of $\operatorname{lsc} F$ that $\Gamma\text{-}\lim F \leq \operatorname{lsc} F$. □

1.2 Γ-convergence and minimizers

The most important property of Γ-convergence is the fact that it implies convergence of minima (infima). This is the main justification for using Γ-convergence, because this property directly corresponds to the main topic of the Calculus of Variations, the search for minimizers. In order to state and prove this result, we need the following notion of uniform coercivity: A family $(F_k)_k$ of functionals $F_k : \mathcal{X} \to \mathbb{R}_\infty$ is called **equi-mildly coercive** if there exists a compact set $K \subseteq \mathcal{X}$ with

$$\inf_{\mathcal{X}} F_k = \inf_K F_k \qquad \text{for all } k \in \mathbb{N}. \tag{1.1}$$

Of course, without coercivity no convergence of infima can be expected as one easily sees from the sequence of functionals $F_k := -\delta_k$.

Theorem 1.7. *Let $F_k : \mathcal{X} \to \mathbb{R}_\infty$, $k \in \mathbb{N}$, be equi-mildly coercive functionals and assume $F_\infty = \Gamma\text{-}\lim_k F_k$ exists. Then F_∞ admits a minimizer and it holds that*

$$\min_{\mathcal{X}} F_\infty = \lim_{k\to\infty} \inf_{\mathcal{X}} F_k.$$

In addition, all accumulation points of a precompact sequence $(x_k)_k \subseteq \mathcal{X}$ with the property that $\liminf_{k\to\infty} F_k(x_k) = \liminf_{k\to\infty} \inf_{\mathcal{X}} F_k$ are minimizers of F_∞.

Proof. Let K be the compact set from the equi-mild coercivity of the F_k. Then, for all $k \in \mathbb{N}$ choose $x_k \in K$ such that $|F_k(x_k) - \inf_{\mathcal{X}} F_k| \leq 1/k$. In particular,

$$\liminf_{k\to\infty} F_k(x_k) = \liminf_{k\to\infty} \inf_{\mathcal{X}} F_k.$$

Because all x_k lie in the compact set K, we can select a subsequence $(x_{k(j)})_j$ with $x_{k(j)} \to x_* \in \mathcal{X}$ as $j \to \infty$ and satisfying

$$\lim_{j\to\infty} F_{k(j)}(x_{k(j)}) = \liminf_{k\to\infty} \inf_{\mathcal{X}} F_k.$$

Define the sequence $(\tilde{x}_k)_k$ as follows:

$$\tilde{x}_k := \begin{cases} x_{k(j)} & \text{if } k = k(j) \text{ for some } j \in \mathbb{N}, \\ x_* & \text{otherwise.} \end{cases}$$

Then, $\tilde{x}_k \to x_*$ and by the lim inf-inequality

$$\inf_{\mathcal{X}} F_\infty \leq F_\infty(x_*) \leq \liminf_{k\to\infty} F_k(\tilde{x}_k) \leq \liminf_{j\to\infty} F_{k(j)}(x_{k(j)}) = \liminf_{k\to\infty} \inf_{\mathcal{X}} F_k. \tag{1.2}$$

Now, let $x \in \mathcal{X}$ be such that $F_\infty(x) \leq \inf_{\mathcal{X}} F_\infty + \varepsilon$ and let $x_k \to x$ be a recovery sequence for x. Then,

$$\limsup_{k\to\infty} \inf_{\mathcal{X}} F_k \leq \limsup_{k\to\infty} F_k(x_k) = F_\infty(x) \leq \inf_{\mathcal{X}} F_\infty + \varepsilon$$

and by letting $\varepsilon \to 0$, we get

$$\limsup_{k\to\infty} \inf_{\mathcal{X}} F_k \leq \inf_{\mathcal{X}} F_\infty. \tag{1.3}$$

Consequently, combining (1.2) and (1.3), the claim of the theorem follows. The addition is clear if we take the given sequence instead of the constructed one in the first step of the proof. □

If F is lower semicontinuous, in particular $F = \Gamma\text{-}\lim_k F$ by Proposition 1.6, then the method of constructing a minimizer through a minimizing sequence as employed in the preceding proof ($\lim_{k\to\infty} F(x_k) = \inf_{\mathcal{X}} F$) is called the **direct method of the Calculus of Variations**.

The preceding results allows for various refinements and extensions and indeed we will be occupied with such extensions during most of this chapter. One simple but later fact is concerned with the situation in which we only have the Γ-limes inferior:

Proposition 1.8. *Let $F_k : \mathcal{X} \to \mathbb{R}_\infty$, $k \in \mathbb{N}$, be functionals and denote by $F_* = \Gamma\text{-}\liminf_k F_k$ the Γ-limes inferior. Then, for any open set $U \subseteq \mathcal{X}$:*

$$\liminf_{k\to\infty} \inf_U F_k \leq \inf_U F_*.$$

Proof. Let $\varepsilon > 0$ and choose $x \in U$ with $F_*(x) \leq \inf_U F_* + \varepsilon$. Since

$$F_*(x) = \inf \Big\{ \liminf_{k\to\infty} F_k(x_k) \ \colon \ x_k \to x \Big\},$$

there exists a sequence $x_k \to x$ in $\mathcal{X}$ with $\liminf_{k\to\infty} F_k(x_k) \leq F_*(x) + \varepsilon$. Because U is open, $x_k \in U$ for k large enough, whence we get

$$\inf_U F_* \geq F_*(x) - \varepsilon \geq \liminf_{k\to\infty} F_k(x_k) - 2\varepsilon \geq \liminf_{k\to\infty} \inf_U F_k - 2\varepsilon$$

and the claim follows by letting $\varepsilon \to 0$. □

If additionally the F_k are equi-mildly coercive, we can repeat the whole argument in Theorem 1.7 until (1.2) for the Γ-limes inferior in place of the full Γ-limit (note that there we only used the lim inf-inequality, which also holds for the Γ-limes inferior) to get

$$\inf_{\mathcal{X}} F_* \leq F_*(x_*) \leq \liminf_{k\to\infty} \inf_{\mathcal{X}} F_k$$

for some $x_* \in \mathcal{X}$. Hence, applying the preceding proposition (with $U = \mathcal{X}$), we have shown:

Corollary 1.9. *Let $F_k : \mathcal{X} \to \mathbb{R}_\infty$, $k \in \mathbb{N}$, be equi-mildly coercive functionals and let $F_* := \Gamma\text{-}\liminf_k F_k$. Then F_* admits a minimizer and it holds that*

$$\min_{\mathcal{X}} F_* = \liminf_{k\to\infty} \inf_{\mathcal{X}} F_k.$$

By an analogous argument as in Proposition 1.8 one can also show that

$$\limsup_{k\to\infty} \inf_U F_k \leq \inf_U F^*,$$

where $F^* := \Gamma\text{-}\limsup_k F_k$ is now the Γ-limes superior. But contrary to the situation for the Γ-limes inferior, the Γ-limes superior does not always fulfill the reverse inequality (cf. [Dal93, Example 7.3] for a counterexample).

1.3 Approximate minimizers

This section presents results concerning approximate minimizers of sequences of functionals. First, we investigate the behavior of the set of ε-minimizers of functionals in the case of Γ-convergence (Propositions 1.10 and 1.14). Then, a characterization of the limit functional's set of ε-minimizers in terms of the sequence functionals' sets of $(\varepsilon + \delta)$-minimizers is given (Proposition 1.15). All this can be seen as generalizations of Theorem 1.7. Results in this spirit for the set of true minimizers instead of the set of ε-minimizers and with different proofs (for the topological definition of Γ-convergence) can be found in Chapter 7 of the book [Dal93].

For $\varepsilon \geq 0$, define the **set of ε-minimizers** of a functional $F : \mathcal{X} \to \mathbb{R}_\infty$ to be

$$\operatorname{Argmin}_\varepsilon(F) = \operatorname{Argmin}_\varepsilon \{ F(x) \ : \ x \in \mathcal{X} \} := \{ x \in \mathcal{X} \ : \ F(x) \leq \inf_{\mathcal{X}} F + \varepsilon \}.$$

For $\varepsilon = 0$ we recover the set of true minimizers and omit the index "$_0$".

The first result shows that ε-minimizers of F_∞ can be approximated arbitrarily well by $(\varepsilon + \delta)$-minimizers of the F_k, where $\delta > 0$.

Proposition 1.10. *Let $F_k : \mathcal{X} \to \mathbb{R}_\infty$, $k \in \mathbb{N}$, be equi-mildly coercive functionals and assume $F_\infty = \Gamma\text{-}\lim_k F_k$ exists. Then, for all $\varepsilon \geq 0$, $\delta > 0$, $x \in \operatorname{Argmin}_\varepsilon(F_\infty)$ and for all recovery sequences $x_k \to x$, it holds that $x_k \in \operatorname{Argmin}_{\varepsilon+\delta}(F_k)$ for all $k \geq k_0$ with $k_0 = k_0(\delta)$ sufficiently large.*

Proof. Owing to the equi-mild coercivity, from Theorem 1.7 we know that $\min_{\mathcal{X}} F_\infty = \lim_{k\to\infty} \inf_{\mathcal{X}} F_k$. Hence, for k sufficiently large, it holds that

$$|\min_{\mathcal{X}} F_\infty - \inf_{\mathcal{X}} F_k| \leq \frac{\delta}{2}.$$

As $(x_k)_k$ is a recovery sequence for x, we have, again for k big enough,

$$|F_k(x_k) - F_\infty(x)| \leq \frac{\delta}{2}.$$

Combining these two estimates with $|F_\infty(x) - \inf_{\mathcal{X}} F_\infty| \leq \varepsilon$ yields

$$\begin{aligned} |F_k(x_k) - \inf_{\mathcal{X}} F_k| &\leq |F_k(x_k) - F_\infty(x)| + |F_\infty(x) - \min_{\mathcal{X}} F_\infty| + |\min_{\mathcal{X}} F_\infty - \inf_{\mathcal{X}} F_k| \\ &\leq \frac{\delta}{2} + \varepsilon + \frac{\delta}{2} = \varepsilon + \delta, \end{aligned}$$

i.e. $x_k \in \operatorname{Argmin}_{\varepsilon+\delta}(F_k)$ for all k sufficiently large. □

Remark 1.11. An inspection of the proof reveals that if $\inf_{\mathcal{X}} F_k = \min_{\mathcal{X}} F_\infty$ for all $k \in \mathbb{N}$, then we do not need the assumptions of equi-mild coerciveness as it is only needed for convergence of infima (through Theorem 1.7). This equality of infima is indeed easy to fulfill in the Calculus of Variations (even if no coercivity is given): We can always set $F_k' := F_k + c_k$, where c_k is chosen precisely to ensure equality of infima. This translation does not change the minimization problem associated with F_k.

The preceding proposition is a first example of what here is called "reverse approximation". In general this is the search for approximating sequences to given (approximate) solutions of some problem. The last result established the existence of a sequence of approximate minimizers of F_k to a given (approximate) minimizer of F_∞. This gives additional information on the minimizer we started with and allows, for example, to work with the sequence of approximating minimizers, which might be more favorable at times. One can also view this, in a certain sense, as a generalization of the concept of recovery sequence.

One could hope to avoid the usage of the sequence $(x_k)_k$ and conjecture that an ε-minimizer of F_∞ is also an $n\varepsilon$-minimizer of F_k for some $n \in \mathbb{N}$ and for k sufficiently large. Even if all the F_k are lower semicontinuous, however, this is not the case as shown by the following counterexample.

Counterexample 1.12. Let $\mathcal{X} = \mathbb{R}$ and for $k \in \mathbb{N}$ define

$$F_k(x) := \begin{cases} -1 & \text{if } x = 1/k, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad F_\infty(x) := \begin{cases} -1 & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

It is clear that $F_\infty = \Gamma\text{-}\lim_k F_k$ (local uniform convergence in $\mathcal{X} \setminus \{0\}$ and use the recovery sequence $x_k = 1/k$ for $x = 0$) and that all F_k, F_∞ are lower semicontinuous. The only ε-minimizer, $\varepsilon \in [0, 1)$, of F_∞ is $x = 0$, but $x = 0$ is no ε-minimizer of any F_k (in fact, $x = 0$ is not an $n\varepsilon$-minimizer for any $n \in \mathbb{N}$ and sufficiently small ε, either).

The next counterexample shows that in Proposition 1.10 one cannot replace the Γ-limit by the Γ-limes inferior.

Counterexample 1.13. Let again $\mathcal{X} = \mathbb{R}$ and for all $k \in \mathbb{N} \setminus \{1\}$ define

$$G_k(x) := \begin{cases} (-1)^k & \text{if } x = 1/k, \\ -1/2 & \text{if } x = 1, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad G_*(x) := \begin{cases} -1 & \text{if } x = 0, \\ -1/2 & \text{if } x = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then $G_* = \Gamma\text{-}\liminf_k G_k$ by a similar reasoning as in the previous counterexample. The G_k, however, do not Γ-converge. In fact, for $x = 0$, one would need to construct a sequence $(x_k)_k$ with $-1 = G_*(0) \geq \limsup_{k\to\infty} G_k(x_k)$. But as $G_{2l} \geq -1/2$, also $\limsup_{k\to\infty} G_k(x_k) \geq -1/2$, a contradiction, and hence no such sequence $(x_k)_k$ exists.

The only, say, 1/6-minimizer of G_* is $x_* = 0$ ($\varepsilon = 1/6$), but any $(1/6 + 1/6)$-minimizer ($\delta = 1/6$) for G_{2l} must be $x_{2l} = 1$ and hence the sequence $(x_k)_k$ cannot converge to $x_* = 0$. Indeed, the conclusion of Proposition 1.10 fails, because we cannot find a "whole" sequence $x_k \to x_*$ with $x_k \in \operatorname{Argmin}_{\varepsilon+\delta}(F_k)$. For the *subsequence* $(x_{2l+1})_l$ with $x_{2l+1} = 1/(2l + 1) \to 0 = x_*$, however, we even have $x_{2l+1} \in \operatorname{Argmin}_\varepsilon(F_{2l+1})$ (without δ).

As seen in the previous counterexample, if we only have $F_* = \Gamma\text{-}\liminf_k F_k$, we need to rely on a construction of subsequences, which additionally depends on the choice of δ. This is made precise in the following proposition.

Proposition 1.14. *Let $F_k : \mathcal{X} \to \mathbb{R}_\infty$, $k \in \mathbb{N}$, be equi-mildly coercive functionals and let $F_* = \Gamma\text{-}\liminf_k F_k$. Then, for all $x \in \operatorname{Argmin}_\varepsilon(F_*)$, $\varepsilon \geq 0$, and for all $\delta > 0$, there exists a subsequence $(x_{k_l})_l$ with $x_{k_l} \to x$ such that $x_{k_l} \in \operatorname{Argmin}_{\varepsilon+\delta}(F_{k_l})$.*

Proof. Corollary 1.9 implies that $\min_{\mathcal{X}} F_* = \liminf_{k\to\infty} \inf_{\mathcal{X}} F_k$. Hence, for k sufficiently large,

$$\min_{\mathcal{X}} F_* - \inf_{\mathcal{X}} F_k \leq \frac{\delta}{2}. \tag{1.4}$$

Let now $x \in \operatorname{Argmin}_\varepsilon(F_*)$. Because $F_*(x) = \inf\{\liminf_{k\to\infty} F_k(x_k) \,:\, x_k \to x\}$, there exists a sequence $x_k \to x$ with $\liminf_{k\to\infty} F_k(x_k) \leq F_*(x) + \delta/4$. Hence, we can find a subsequence $(x_{k_l})_l$ such that

$$F_{k_l}(x_{k_l}) - F_*(x) \leq \frac{\delta}{2}$$

for all $l \in \mathbb{N}$. Without loss of generality we can assume that k_1 is already so big that (1.4) holds. Combining the last two estimates with $F_*(x) - \min_{\mathcal{X}} F_* \leq \varepsilon$ yields

$$\begin{aligned} F_{k_l}(x_{k_l}) - \inf_{\mathcal{X}} F_{k_l} &\leq (F_{k_l}(x_{k_l}) - F_*(x)) + (F_*(x) - \min_{\mathcal{X}} F_*) + (\min_{\mathcal{X}} F_* - \inf_{\mathcal{X}} F_{k_l}) \\ &\leq \frac{\delta}{2} + \varepsilon + \frac{\delta}{2} = \varepsilon + \delta, \end{aligned}$$

i.e. $x_{k_l} \in \operatorname{Argmin}_{\varepsilon+\delta}(F_{k_l})$. □

The language of topological set convergence allows us to nicely describe the set of ε-minimizers of F_* in terms of the sets of $(\varepsilon+\delta)$-minimizers of the F_k. For this, define the **Kuratowski lower limit** and the **Kuratowski upper limit**, respectively, of a sequence of sets $A_k \subseteq \mathcal{X}$ through

$$\begin{aligned} \text{K-}\liminf_k A_k &:= \left\{ x \in \mathcal{X} \,:\, \exists (x_k)_k \subseteq \mathcal{X} : x_k \to x,\ \forall k \in \mathbb{N} : x_k \in A_k \right\}, \\ \text{K-}\limsup_k A_k &:= \left\{ x \in \mathcal{X} \,:\, \exists (x_{k_l})_l \subseteq \mathcal{X} : k_l \uparrow \infty,\ x_{k_l} \to x,\ \forall l \in \mathbb{N} : x_{k_l} \in A_{k_l} \right\}. \end{aligned}$$

The Kuratowski upper and lower limits are always (sequentially) closed.

With these definitions, we can characterize the set of ε-minimizers as follows:

Theorem 1.15. *Let $F_k : \mathcal{X} \to \mathbb{R}_\infty$, $k \in \mathbb{N}$, be equi-mildly coercive functionals and let $F_* = \Gamma\text{-}\liminf_k F_k$. Then, for all $\varepsilon \geq 0$ it holds that*

$$\operatorname{Argmin}_\varepsilon(F_*) = \bigcap_{\delta>0} \text{K-}\limsup_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k).$$

If in addition $F_\infty = \Gamma\text{-}\lim_k F_k$, then

$$\operatorname{Argmin}_\varepsilon(F_\infty) = \bigcap_{\delta>0} \text{K-}\liminf_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k) = \bigcap_{\delta>0} \text{K-}\limsup_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k).$$

Proof. We commence with the result for the Γ-limes inferior. To prove this, let $x \in \bigcap_{\delta>0} \text{K-}\limsup_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k)$. This means that for all $\delta > 0$ we can find a sequence $x_k^\delta \to x$ (as $k \to \infty$) such that $x_k^\delta \in \operatorname{Argmin}_{\varepsilon+\delta}(F_k)$, i.e.

$$F_k(x_k^\delta) \leq \inf_{\mathcal{X}} F_k + \varepsilon + \delta.$$

Using the sequences $(x_k^\delta)_k$ in the definition of the Γ-limes inferior, we get

$$F_*(x) = \inf \Big\{ \liminf_{k\to\infty} F_k(x_k) \ : \ x_k \to x \Big\} \leq \liminf_{k\to\infty} \inf_{\mathcal{X}} F_k + \varepsilon + \delta.$$

As δ was arbitrary, it follows that

$$F_*(x) \leq \liminf_{k\to\infty} \inf_{\mathcal{X}} F_k + \varepsilon.$$

Proposition 1.8 (with $U = \mathcal{X}$) implies $\liminf_{k\to\infty} \inf_{\mathcal{X}} F_k \leq \inf_{\mathcal{X}} F_*$ and hence $F_*(x) \leq \inf_{\mathcal{X}} F_* + \varepsilon$, i.e. $x \in \operatorname{Argmin}_\varepsilon(F_*)$. Thus we have shown

$$\bigcap_{\delta>0} \text{K-}\limsup_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k) \subseteq \operatorname{Argmin}_\varepsilon(F_*). \tag{1.5}$$

In the language of topological set convergence, the result of Proposition 1.14 reads

$$\operatorname{Argmin}_\varepsilon(F_*) \subseteq \bigcap_{\delta>0} \text{K-}\limsup_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k) \tag{1.6}$$

and the claim for the Γ-limes inferior follows.

For the second assertion, note that the Γ-limit F_∞ is also the Γ-limes inferior of the F_k (by Lemma 1.3) and hence the first step of the proof includes this case as well. This time, however, we finish differently and use Proposition 1.10, which gives the stronger assertion

$$\operatorname{Argmin}_\varepsilon(F_\infty) \subseteq \bigcap_{\delta>0} \text{K-}\liminf_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k)$$

instead of (1.6). Combining this with (1.5), we have established

$$\bigcap_{\delta>0} \text{K-}\limsup_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k) \subseteq \operatorname{Argmin}_\varepsilon(F_\infty) \subseteq \bigcap_{\delta>0} \text{K-}\liminf_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k).$$

Because trivially $\text{K-}\liminf_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k) \subseteq \text{K-}\limsup_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k)$, these inclusions are in fact equalities. □

Note how this result complements Theorem 1.7 since it gives us information about the *minimizers* and not only on the *minimum*.

1.4 Contraction of minimizer sets

This section examines the behavior of the diameter of the sets of ε-minimizers with respect to changes in ε. In particular, continuity properties of the diameter as a function of ε depend both on (lower semi-)continuity and convexity properties of the functional (Propositon 1.16). We will also see that the contraction of the approximate minimizer sets of the Γ-limit of a sequence of functionals also implies contraction for the functionals in the sequence (Proposition 1.20).

For the concept of convexity to make sense, let in all of the following $\mathcal{X}$ be an vector space such that the metric d turns $\mathcal{X}$ into a topological vector space, i.e. such that vector addition and scalar multiplication are continuous. As usual, a functional $F : \mathcal{X} \to \mathbb{R}_\infty$ is called **convex** if

$$F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y) \qquad \text{for all } \lambda \in (0,1),\ x, y \in \mathcal{X}.$$

It is called **strictly convex** if "$<$" holds instead of "$\leq$".

For a subset $A \subseteq \mathcal{X}$, define the **diameter** $\operatorname{diam}(A) \in [0,\infty]$ to be

$$\operatorname{diam}(A) := \sup \left\{ d(x,y) \ : \ x, y \in A \right\}.$$

To each functional $F : \mathcal{X} \to \mathbb{R}_\infty$ (not necessarily convex), we associate a **modulus of contraction** $\omega_F : [0,\infty) \to [0,\infty]$ through

$$\omega_F(\varepsilon) := \operatorname{diam}(\operatorname{Argmin}_\varepsilon(F)).$$

Obviously, F is increasing[3]. For strictly convex F, $\omega_F(0) = 0$ since the minimizer set is (at most) a singleton.

A functional $F : \mathcal{X} \to \mathbb{R}_\infty$ is called **coercive** $\operatorname{Argmin}_\varepsilon(F)$ is relatively compact for all $\varepsilon > 0$.

Proposition 1.16. *Let $F : \mathcal{X} \to \mathbb{R}_\infty$ be a functional. If F is convex, then ω_F is left-continuous (in $(0,\infty)$). If F is lower semicontinuous and coercive, then ω_F is right-continuous (in $[0,\infty)$).*

Proof. Left-continuity.
Fix $\varepsilon_0 > 0$. To show

$$\lim_{\varepsilon \uparrow \varepsilon_0} \operatorname{diam}(\operatorname{Argmin}_\varepsilon(F)) = \lim_{\varepsilon \uparrow \varepsilon_0} \omega_F(\varepsilon) = \omega_F(\varepsilon_0) = \operatorname{diam}(\operatorname{Argmin}_{\varepsilon_0}(F))$$

(the limit exists owing to the monotonicity of ω_F), let $\varepsilon_j \uparrow \varepsilon_0$ ($\varepsilon_j < \varepsilon_0$) and assume to the contrary that $\lim_{j\to\infty} \omega_F(\varepsilon_j) < \operatorname{diam}(\operatorname{Argmin}_{\varepsilon_0}(F))$ for all $j \in \mathbb{N}$. For

$$A := \bigcup_{j=1}^{\infty} \operatorname{Argmin}_{\varepsilon_j}(F)$$

[3] "Increasing" here is to be understood as synonymous to "nondecreasing".

it holds that $\operatorname{diam}(A) = \lim_{j\to\infty} \omega_F(\varepsilon_j)$ since clearly $\operatorname{diam}(A) \geq \operatorname{diam}(\operatorname{Argmin}_{\varepsilon_j}(F))$ for all $j \in \mathbb{N}$, and for any $x, y \in A$ it follows that $x, y \in \operatorname{Argmin}_{\varepsilon_j}(F)$ for all $j \in \mathbb{N}$ large enough ($\operatorname{Argmin}_{\varepsilon_j}(F) \subseteq \operatorname{Argmin}_{\varepsilon_k}(F)$ for $j \leq k$), whence it holds that $\operatorname{diam}(A) \leq \lim_{j\to\infty} \operatorname{diam}(\operatorname{Argmin}_{\varepsilon_j}(F))$.

By assumption, $\operatorname{diam}(A) = \lim_{j\to\infty} \omega_F(\varepsilon_j) < \operatorname{diam}(\operatorname{Argmin}_{\varepsilon_0}(F))$ and we can infer that there exists $b \in \operatorname{Argmin}_{\varepsilon_0}(F) \setminus A$ with positive distance from A. Also, let $a \in A$. With $m := \inf_{\mathcal{X}} F$ we have $F(a) < m + \varepsilon_0$ ($a \in \operatorname{Argmin}_{\varepsilon_j}(F)$ for some $j \in \mathbb{N}$) and $F(b) = m + \varepsilon_0$ ("$\leq$" by $b \in \operatorname{Argmin}_{\varepsilon_0}(F)$ and "$\geq$" by $b \notin A$). On one hand, since b has positive distance from A, there exists some $\lambda \in (0,1)$ such that $c := \lambda a + (1-\lambda)b \in \operatorname{Argmin}_{\varepsilon_0}(F) \setminus A$. But on the other hand, by convexity, $F(c) \leq \lambda F(a) + (1-\lambda)F(b) < m + \varepsilon_0$, whence $c \in A$, a contradiction.

Right-continuity.
Fix $\varepsilon_0 \geq 0$. We have to show

$$\lim_{\varepsilon\downarrow\varepsilon_0} \operatorname{diam}(\operatorname{Argmin}_{\varepsilon}(F)) = \lim_{\varepsilon\downarrow\varepsilon_0} \omega_F(\varepsilon) = \omega_F(\varepsilon_0) = \operatorname{diam}(\operatorname{Argmin}_{\varepsilon_0}(F))$$

(again, the limit exists by monotonicity). To the contrary assume that there exists a sequence $\varepsilon_j \downarrow \varepsilon_0$ ($\varepsilon_j > \varepsilon_0$) and points $x_j, y_j \in \operatorname{Argmin}_{\varepsilon_j}(F)$ such that $d(x_j, y_j) \geq \operatorname{diam}(\operatorname{Argmin}_{\varepsilon_0}(F)) + \mu$ for some fixed $\mu > 0$. By the coercivity of F, we can find subsequences (not relabelled) with $x_j \to x$ and $y_j \to y$. For these it holds that $d(x,y) \geq \operatorname{diam}(\operatorname{Argmin}_{\varepsilon_0}(F)) + \mu$ since by the triangle inequality

$$\operatorname{diam}(\operatorname{Argmin}_{\varepsilon_0}(F)) + \mu \leq d(x_j, y_j) \leq d(x_j, x) + d(x, y) + d(y, y_j) \tag{1.7}$$

and the first and third term converge to zero. By the lower semicontinuity, we infer $x, y \in \operatorname{Argmin}_{\varepsilon_0}(F)$. But then $\operatorname{diam}(\operatorname{Argmin}_{\varepsilon_0}(F)) \geq d(x,y) \geq \operatorname{diam}(\operatorname{Argmin}_{\varepsilon_0}(F)) + \mu$, a contradiction. □

Some counterexamples show the necessity of the assumptions.

Counterexample 1.17. For nonconvex functionals the modulus of contraction may not be left-continuous: Take the nonconvex, continuous function $F : \mathbb{R} \to \mathbb{R}$ defined as

$$F(x) := \begin{cases} -x-2 & \text{if } x \in (-\infty, -1], \\ x & \text{if } x \in [-1, 0], \\ -x & \text{if } x \in [0, 2], \\ x-4 & \text{if } x \in [2, +\infty) \end{cases}$$

(see Figure 1). Then,

$$\omega_F(\varepsilon) = \begin{cases} 2\varepsilon & \text{if } \varepsilon \in [0, 1), \\ 2\varepsilon + 2 & \text{if } \varepsilon \in [1, +\infty), \end{cases}$$

which is not left-continuous at $\varepsilon = 1$.

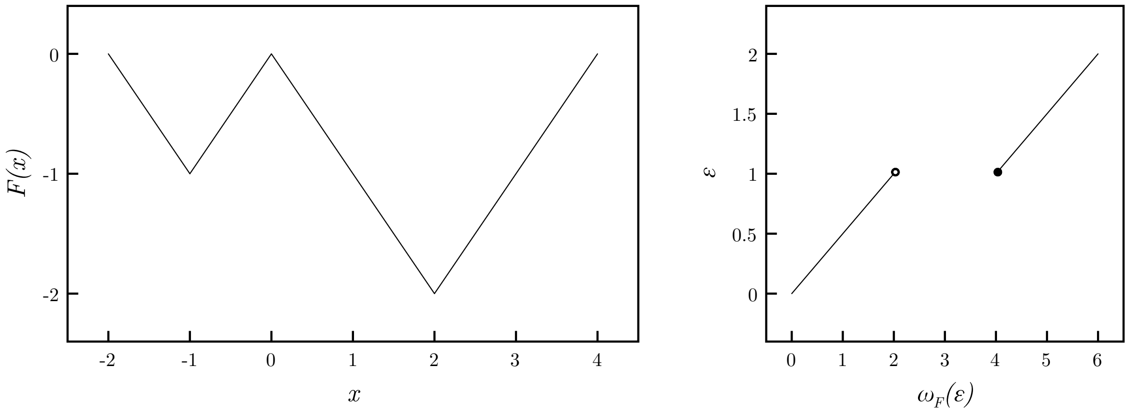

Figure 1: A nonconvex F such that ω_F is not left-continuous.

Counterexample 1.18. If a functional is not lower semicontinuous[4], its modulus of contraction may not be right-continuous: The non-lower semicontinuous function $F : \mathbb{R} \to \mathbb{R}$ defined as

$$F(x) := \begin{cases} -x-2 & \text{if } x \in (-\infty, -1), \\ 0 & \text{if } x = -1, \\ x & \text{if } x \in (-1, 0], \\ -x & \text{if } x \in [0, 2], \\ x-4 & \text{if } x \in [2, +\infty) \end{cases}$$

(see Figure 2) has a modulus of contraction

$$\omega_F(\varepsilon) = \begin{cases} 2\varepsilon & \text{if } \varepsilon \in [0, 1], \\ 2\varepsilon + 2 & \text{if } \varepsilon \in (1, +\infty) \end{cases}$$

which is not right-continuous at $\varepsilon = 1$.

Counterexample 1.19. If no coercivity assumption is made, right-continuity of ω_F in 0 may fail, even if the functional is continuous and convex: For $F : \mathbb{R} \to \mathbb{R}$ given by $F(x) := \mathrm{e}^x$, we have $\omega_F(\varepsilon) = \operatorname{diam}((-\infty, \ln \varepsilon]) = \infty$ for any $\varepsilon > 0$, but $\omega_F(0) = \operatorname{diam}(\emptyset) = 0$.

[4]This counterexample presents a function which is neither lower semicontinuous nor convex. This is so, because in many situations convexity implies lower semicontinuity: In finite-dimensional spaces, convex functions are automatically continuous at all points where the function is bounded from above, see [Dac89, Theorem 2.3]; in infinite-dimensional Banach spaces this also holds for the strong topology while for the weak topology, they are at least lower semicontinuous since convex strongly closed sets are also weakly closed and if the space is additionally reflexive and separable then it is also sequentially weakly closed (see [Con90, Theorem V.1.5], also see the counterexample for the non-reflexive case there).

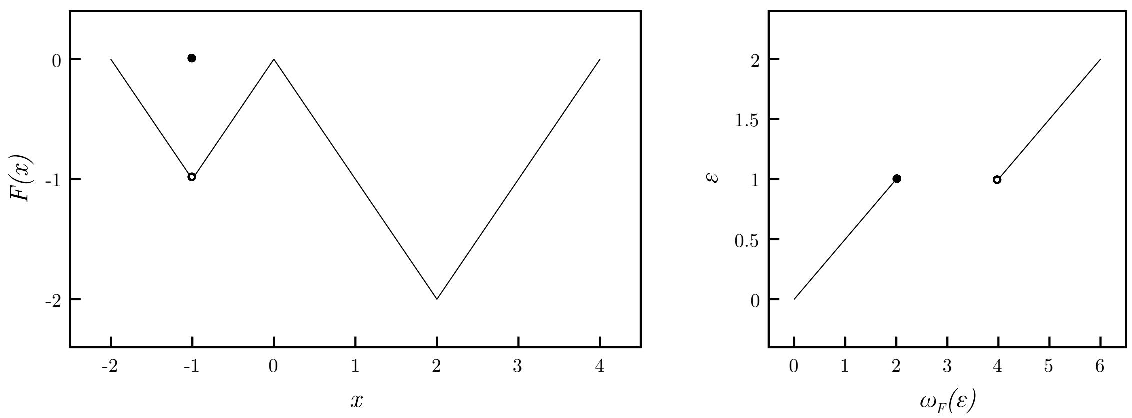

Figure 2: A non-lower semicontinuous F such that ω_F is not right-continuous.

Finally, Proposition 1.15 allows us to deduce that if the sets of ε-minimizers for $F_* = \Gamma\text{-}\liminf_k F_k$ contract, the sets of ε-minimizers of F_k contract as well.

Proposition 1.20. *Let $F_k : \mathcal{X} \to \mathbb{R}_\infty$, $k \in \mathbb{N}$, be equi-coercive functionals such that there exist $K \subseteq \mathcal{X}$ compact and $\varepsilon_0 > 0$ with $\bigcup_{k\in\mathbb{N}} \operatorname{Argmin}_\varepsilon(F_k) \subseteq K$ for all $\varepsilon \in [0, \varepsilon_0]$. Denote by $F_* = \Gamma\text{-}\liminf_k F_k$ the Γ-limes inferior of the F_k and fix $\eta > 0$. Then, for all $\varepsilon \in [0, \varepsilon_0]$ with*

$$\omega_{F_*}(\varepsilon) := \operatorname{diam}(\operatorname{Argmin}_\varepsilon(F_*)) \leq \frac{\eta}{2},$$

there exists a $k_0 = k_0(\eta, \varepsilon) \in \mathbb{N}$ such that

$$\omega_{F_k}(\varepsilon) := \operatorname{diam}(\operatorname{Argmin}_\varepsilon(F_k)) \leq \eta \qquad \text{for all } k \geq k_0.$$

Proof. We prove this by contradiction and assume that there exist $\eta > 0$ and $\varepsilon \in [0, \varepsilon_0]$ such that $\operatorname{diam}(\operatorname{Argmin}_\varepsilon(F_*)) \leq \eta/2$, but we can find sequences $(x_{k_l})_l$, $(y_{k_l})_l$ ($k_l \uparrow \infty$) in K with $x_{k_l}, y_{k_l} \in \operatorname{Argmin}_\varepsilon(F_{k_l})$ and $d(x_{k_l}, y_{k_l}) > \eta$. Because K is compact, we can choose subsequences (not relabelled) with $x_{k_l} \to x \in K$ and $y_{k_l} \to y \in K$. For these x, y we have $d(x, y) \geq \eta$ as seen in the same fashion as in (1.7) (by the triangle inequality).

By the definition of the Kuratowski upper limit and Proposition 1.15, we have

$$\begin{aligned} x, y &\in \text{K-}\limsup\nolimits_k \operatorname{Argmin}_\varepsilon(F_k) \\ &\subseteq \bigcap_{\delta>0} \text{K-}\limsup\nolimits_k \operatorname{Argmin}_{\varepsilon+\delta}(F_k) = \operatorname{Argmin}_\varepsilon(F_*). \end{aligned}$$

Hence $\operatorname{diam}(\operatorname{Argmin}_\varepsilon(F_*)) \geq \eta$. But this is a contradiction. □

Remark 1.21. It is obvious that, by using exactly the same arguments, one can change the premise $\omega_{F_*}(\varepsilon) \leq \eta/2$ of the preceding proposition to read as $\omega_{F_*}(\varepsilon) \leq \alpha\eta$ with some fixed $\alpha \in (0, 1)$.

The results of this section will be applied in Chapter 3.

2 Rate-independent systems

Rate-independent systems could roughly be described as “minimization problems with time-dependence” for they possess properties of both evolutionary and of stationary systems. While they describe processes with time-evolution, they do not possess own dynamics meaning that if the exterior conditions (for example, an external mechanical loading) do not change, the system rests in static equilibrium. This is caused by the (possibly idealized) fact that the time evolution is on a much larger time-scale than the system response. Therefore, they are sometimes called **quasi-static** (also cf. [Mie05, Chapter 1] for a clarification of this terminology).

First, we investigate several different formulations of rate-independent systems in Section 2.1; in particular, energetic solutions are introduced. After a look at the corresponding time-discrete case in Section 2.2, the standard assumptions of the theory and the main existence theorems are presented in Sections 2.3 and 2.4 for single systems and for sequences of systems, respectively. Finally, we briefly touch on uniqueness issues in Section 2.5.

2.1 Definition of rate-independence

There are different paths from which one can approach rate-independent systems. We will first look at a rather abstract axiomatic definition of rate-independence and then investigate more concrete situations. After looking at the most commonly known statement with variational inequalities or (sub-)differential inclusions, we present the more recent concept of energetic solutions, which is the most important formulation for this work. In the convex case, this formulation turns out to be equivalent to the usual formulations (see Proposition 2.6).

2.1.1 Axiomatics of rate-independence

This section wants to define rate-independent systems in a rather abstract (axiomatic) and formal fashion. For this, let $\mathcal{Q}$ be the state space of the system[1] and denote by $\mathcal{Q}[s,t]$ the space of all functions defined on the interval $[s,t]$, $s < t$, with values in $\mathcal{Q}$. Further, let $\mathcal{L}[s,t]$ contain all admissible external loadings in the interval $[s,t]$ (e.g. external forces with some regularity). The nature of the space $\mathcal{L}[s,t]$ is deliberately left

[1] Here, we do not assume any additional structure on $\mathcal{Q}$, but later $\mathcal{Q}$ will at least be a Hausdorff topological space.

vague here, because we stay in an abstract setting (but in many cases $\mathcal{L}[s,t]$ consists of functions defined on $[s,t]$ with values in the dual space $\mathcal{Q}^*$ of $\mathcal{Q}$). Now, an evolutionary system H on the interval $[s,t]$ is represented by its **solution operator** $\mathsf{S}_\mathsf{H}[s,t]$ (this is sometimes also called an **input-output operator**) which takes a loading l from $\mathcal{L}[s,t]$ and an initial value $q_0 \in \mathcal{Q}$ to the corresponding solution set in $\mathcal{Q}[s,t]$ (we do not assume unique solvability), so we have the signature

$$\mathsf{S}_\mathsf{H}[s,t] : \mathcal{L}[s,t] \times \mathcal{Q} \rightrightarrows \mathcal{Q}[s,t],$$

and for concise notation we denote by $\mathcal{O}[s,t]$ the space of all such functions with signature $\mathcal{L}[s,t] \times \mathcal{Q} \rightrightarrows \mathcal{Q}[s,t]$; in particular $\mathsf{S}_\mathsf{H} \in \mathcal{O}[s,t]$.

Now, let INT denote the set of all intervals $[s,t] \subseteq \mathbb{R}$, $s < t$, and define

$$\mathsf{SOP} := \bigcup_{[s,t] \in \mathsf{INT}} \mathcal{O}[s,t],$$

the **set of all solution operators**. Then, we can view S_H as a mapping from INT to SOP which is given through $\mathsf{S}_\mathsf{H} : [s,t] \mapsto \mathsf{S}_\mathsf{H}[s,t] \in \mathcal{O}[s,t]$. We call S_H the **solution functor** (see below for an explanation of this terminology).

First, we want the rate-independent system to be a (nonlinear) evolutionary system, meaning that concatenations and restrictions of solutions remain solutions. To state this precisely, let $\bowtie$ denote the **concatenation** of two adjacent intervals: $[r,s] \bowtie [s,t] := [r,t]$ ($r < s < t$) and also denote the concatenation of two solution operators $\mathsf{G} \in \mathcal{O}[r,s]$ and $\mathsf{R} \in \mathcal{O}[s,t]$ by $\mathsf{G} \bowtie \mathsf{R} \in \mathcal{O}[r,t]$,

$$(\mathsf{G} \bowtie \mathsf{R})(l, q_0)(\tau) := \begin{cases} \mathsf{G}\big(l|_{[r,s]}, q_0\big)(\tau) & \text{if } \tau \in [r,s], \\ \mathsf{R}\big(l|_{[s,t]}, \mathsf{G}(l|_{[r,s]}, q_0)(s)\big)(\tau) & \text{if } \tau \in (s,t], \end{cases}$$

where $l \in \mathcal{L}[r,t]$, $q_0 \in \mathcal{Q}$ and $\tau \in [r,t]$. Note that for set-valued maps $A : \mathcal{X} \rightrightarrows \mathcal{Y}$ and $B : \mathcal{Y} \rightrightarrows \mathcal{Z}$ the composition $B \circ A : \mathcal{X} \rightrightarrows \mathcal{Z}$ is defined as

$$(B \circ A)(x) = B(A(x)) := \big\{\, B(y) \in \mathcal{Z} \;\; \textbf{:} \;\; y \in A(x) \,\big\} \qquad \text{for all } x \in \mathcal{X}.$$

Then the fact that we want concatenations of solutions to be solutions as well can be simply expressed as naturality of the solution functor S_H with respect to concatenations:

Definition 2.1. *A system* H *given through its solution functor* $\mathsf{S}_\mathsf{H} : \mathsf{INT} \to \mathsf{SOP}$ *is called a* **multi-valued evolutionary system** *if*

$$\mathsf{S}_\mathsf{H}[r,t] = \mathsf{S}_\mathsf{H}([r,s] \bowtie [s,t]) = \mathsf{S}_\mathsf{H}[r,s] \bowtie \mathsf{S}_\mathsf{H}[s,t] \tag{2.1}$$

for all $r < s < t$.

This definition also includes the property that restrictions of solutions are again solutions since for $r < s < t < u$, we can write

$$\mathsf{S}_{\mathsf{H}}[r,u] = \mathsf{S}_{\mathsf{H}}[r,s] \bowtie \mathsf{S}_{\mathsf{H}}[s,t] \bowtie \mathsf{S}_{\mathsf{H}}[t,u]$$

and hence $\mathsf{S}_{\mathsf{H}}[r,u](l,q_0)|_{[s,t]} = \mathsf{S}_{\mathsf{H}}[s,t](l,\tilde{q}_0)$, where $\tilde{q}_0 = \mathsf{S}_{\mathsf{H}}[r,s](l|_{[r,s]},q_0)(s)$.

The decisive property of rate-independent evolution systems is that reparametrizations of the time interval do not change the dynamics, i.e. the new solution is just a reparametrization of the old solution. Again, to make this precise, let REP denote the set of all orientation-preserving C^1-**reparametrizations of intervals**, i.e. all $\alpha \in \mathrm{C}^1([s,t])$ with $\dot{\alpha}(\tau) > 0$ for all $\tau \in [s,t]$ and $[s,t] \in \mathsf{INT}$ (such α are automatically injective, i.e. bijective onto their image, by Rolle's theorem)[2]. Note that $\alpha_*[s,t] := \alpha([s,t]) = [\alpha(s),\alpha(t)]$ is connected, i.e. an interval, because α is continuous. Now, each $\alpha : [s,t] \to [\alpha(s),\alpha(t)]$ induces a **transformation** $\alpha_* : \mathcal{O}[s,t] \to \mathcal{O}[\alpha(s),\alpha(t)]$ of solution operators through

$$(\alpha_*\mathsf{G})(l,q_0)(\tau) := \mathsf{G}(l \circ \alpha, q_0)(\alpha^{-1}(\tau)),$$

where $l \in \mathcal{L}[\alpha(s),\alpha(t)]$ (hence $l \circ \alpha \in \mathcal{L}[s,t]$), $q_0 \in \mathcal{Q}$ and $\tau \in [\alpha(s),\alpha(t)]$. With this notion at hand, rate-independence can be formally defined:

Definition 2.2. *A multi-valued evolutionary system* H *given through its solution functor* $\mathsf{S}_{\mathsf{H}} : \mathsf{INT} \to \mathsf{SOP}$ *is called a* **rate-independent system** *if for all* $\alpha \in \mathsf{REP}$, S_{H} *is natural (or homogeneous) with respect to* α_*, *i.e. if for all* $s < t$

$$\mathsf{S}_{\mathsf{H}}[\alpha(s),\alpha(t)] = \mathsf{S}_{\mathsf{H}}(\alpha_*[s,t]) = \alpha_*\mathsf{S}_{\mathsf{H}}[s,t] \tag{2.2}$$

or, equivalently, if

$$\mathsf{S}_{\mathsf{H}}[\alpha(s),\alpha(t)](l,q_0) = \mathsf{S}_{\mathsf{H}}[s,t](l \circ \alpha, q_0) \circ \alpha^{-1} \tag{2.3}$$

for all $l \in \mathcal{L}[\alpha(s),\alpha(t)]$ *and* $q_0 \in \mathcal{Q}$.

In other words, if the external conditions change twice as fast, the system will follow exactly the same path as before, but now with twice the speed. Of course, in many applications, this is an idealization. For example, inertia might have to be taken into account and for very steep α, (2.2) might seize to hold.

From an algebraic point of view, INT can be seen as a (concrete, small) category[3] with the intervals as objects, with the maps $\alpha \in \mathsf{REP}$ as morphisms and with the composition of functions as the categorial composition. Similarly, SOP can be seen as a (concrete, small) category of solution operators with the (induced) transformations α_* of solution

[2] In fact, in this abstract setting it suffices to require α to be a topological embedding, i.e. a homeomorphism onto its image and $\alpha(s) < \alpha(t)$. Later, however, we will need differentiability, so we require it here once and for all.

[3] See, for example, [Mac98] or [Sch70] for categorial terminology.

operators ($\alpha \in \mathsf{REP}$) as morphisms and the composition of transformations as categorial composition. For $\alpha : [s,t] \to [\alpha(s), \alpha(t)]$ and $\beta : [\alpha(s), \alpha(t)] \to [\beta(\alpha(s)), \beta(\alpha(t))]$, we observe that $(\beta \circ \alpha)_* = \beta_* \alpha_*$ by the chain of equalities

$$\begin{aligned}((\beta \circ \alpha)_* \mathsf{G})(l, q_0) &= \mathsf{G}(l \circ \beta \circ \alpha, q_0) \circ \alpha^{-1} \circ \beta^{-1} = (\alpha_* \mathsf{G})(l \circ \beta, q_0) \circ \beta^{-1} \\ &= (\beta_* \alpha_* \mathsf{G})(l, q_0),\end{aligned}$$

where $\mathsf{G} \in \mathcal{O}[s,t]$, $l \in \mathcal{L}[\beta(\alpha(s)), \beta(\alpha(t))]$ and $q_0 \in \mathcal{Q}$. Hence, the mapping S_H extended by $\alpha \mapsto \alpha_*$ is a covariant functor between INT and SOP and rate-independence is the property that S_H commutes with α_* (one time thought of as a transformation of intervals, the other times as a transformation of solution operators).

2.1.2 Subdifferential inclusions and variational inequalities

After these abstract contemplations, it is time to consider a more concrete situation: Let, for now, $\mathcal{Q}$ be a Banach space and consider a **stored energy functional** $\mathcal{E} : [0,T] \times \mathcal{Q} \to \mathbb{R}_\infty$, which for all $t \in [0,T]$ is **Gateâux–differentiable** with respect to the second (state) variable, i.e. the directional derivative (with respect to the second argument)

$$\mathrm{D}\mathcal{E}(t,q)[v] := \lim_{h \downarrow 0} \frac{\mathcal{E}(t, q + hv) - \mathcal{E}(t,q)}{h}$$

exists for all $q, v \in \mathcal{Q}$ and $\mathrm{D}\mathcal{E}(t,q)[\bullet] \in \mathcal{Q}^*$; for all $q \in \mathcal{Q}$, the functional $\mathrm{D}\mathcal{E}(t,q)[\bullet]$ is called the **Gâteaux-differential** of $\mathcal{E}(t, \bullet)$ at q. Note that we allow $\mathcal{E}(t, \bullet)$ to be nonconvex. Further, let $\mathcal{R} : \mathcal{Q} \to [0, \infty]$ be a convex **dissipation potential**. In concrete applications, $\mathcal{E}$ measures reversible or potential energy and $\mathcal{R}$ measures energy dissipation.

We will now formulate the governing equation of a rate-independent process $q \in \mathcal{Q}[0,T]$ (for the moment, we will assume as much smoothness as necessary for all expressions to make sense). One form of this equation is that of an **evolutionary doubly-nonlinear (sub-)differential inclusion**

$$0 \in \partial\mathcal{R}(\dot{q}(t)) + \mathrm{D}\mathcal{E}(t, q(t)) \qquad (\text{in } \mathcal{Q}^*) \qquad \text{for almost all } t \in [0,T]. \tag{DI}$$

Here, the **(convex) subdifferential** ∂F of a functional $F : \mathcal{X} \to \mathbb{R}_\infty$ on a Banach space $\mathcal{X}$ is—as usual—defined to be the set of all supporting hyperplanes, i.e.

$$\partial F(u) := \{\, x^* \in \mathcal{X}^* \;:\; F(u) + \langle x^*, v - u \rangle \leq F(v) \text{ for all } \; v \in \mathcal{X} \,\},$$

cf. for example [Roc70, §23]. Consequently, we can rewrite (DI) as an **evolutionary variational inequality**

$$\langle \mathrm{D}\mathcal{E}(t, q(t)), v - \dot{q}(t) \rangle + \mathcal{R}(v) - \mathcal{R}(\dot{q}(t)) \geq 0 \tag{EVI}$$

for all $v \in \mathcal{Q}$ and almost all $t \in [0,T]$.

The governing equations (DI) and (EVI), repectively, are complemented by an initial condition

$$y(0) = y_0 \in \mathcal{Q}.$$

However, not all initial values are possible, we will come back to this later.

The special and decisive property of rate-independent systems now is the **positive 1-homogeneity** of $\mathcal{R}$, i.e.

$$\mathcal{R}(\lambda q) = \lambda \mathcal{R}(q) \qquad \text{for all } \lambda \geq 0 \text{ and } q \in \mathcal{Q}.$$

Hence, for all $q, v \in \mathcal{Q}$, $q^* \in \mathcal{Q}^*$ and $\lambda > 0$

$$\mathcal{R}(\lambda q) + \langle q^*, v - \lambda q \rangle - \mathcal{R}(v) = \lambda \left[\mathcal{R}(q) + \left\langle q^*, \frac{v}{\lambda} - q \right\rangle - \mathcal{F}\Big(\frac{v}{\lambda}\Big) \right],$$

and substituting $v' := v/\lambda$, one can cancel λ to see that $q^* \in \partial\mathcal{R}(\lambda q)$ if and only if $q^* \in \partial\mathcal{R}(q)$, i.e.

$$\partial\mathcal{R}(\lambda q) = \partial\mathcal{R}(q).$$

This expresses that $\mathcal{R}$ is 0-homogeneous and is the mathematical reflection of rate-independence. Since $\mathcal{R}$ takes as an argument the time-derivative of the process q, this means that the energetic balance is independent of the length of $\dot{q}$, i.e. the process speed and depends only on the direction of $\dot{q}$. This property has profound implications for the behavior of the system. As mentioned earlier, the system should be quasi-static, but this cannot be seen directly from the above equations. From a physical point of view it is therefore natural to switch to a purely energetic formulation, which also easily extends to more general situations. We will pursue this direction in the next section.

The case above is not the most general, one can also think of $\mathcal{R}$ additionally depending on the point $q(t)$, cf. for example [Mie05, Section 4.3]. More generally, one can consider a (Banach- or finite-dimensional) manifold $\mathcal{M}$ instead of $\mathcal{Q}$. Then, one should be more careful and switch to the point of view that $\dot{q}(t)$ takes values in the tangent space $T_{q(t)}\mathcal{M}$ to $\mathcal{M}$ at $q(t)$, cf. [Mie07] for such developments.

2.1.3 Energetic formulation

Energetic solutions to rate-independent problems were introduced in [MT99, MT04] and further developed for example in [MM05, FM06]; a recent survey is [Mie05]. This framework allows for a mathematical treatment of a variety of evolution problems in the material sciences, for example phase transitions in shape-memory alloys [MTL02, KMR05] and crack formation in brittle materials [FM98, DFT05].

We now describe the general framework in order to introduce the main ideas and postpone precise technical assumptions until Sections 2.3 and 2.4.

Let the state-space $\mathcal{Q} = \mathcal{F} \times \mathcal{Z}$ of the system be the product of two Hausdorff topological spaces $\mathcal{F}$ and $\mathcal{Z}$ corresponding to the elastic (or, more generally, non-dissipative) and the internal (or dissipative) variables. This splitting is typical in continuum mechanics with dissipation [HN75, ZW87, Fré02]. Despite the fact that $\mathcal{F}$ and $\mathcal{Z}$ might not be first or second countable, we will deal with sequences rather than with general topology tools and all topological notions are to be understood in a sequential sense; for example, *compactness* always means *sequential compactness*. In particular, we will use the sequential flavor of Γ-convergence introduced in Chapter 1.

Parallel to the separation of the state-space into two parts, the system itself is modelled by two functionals: an **energy-storage functional** $\mathcal{E} : [0,T] \times \mathcal{Q} \to \mathbb{R}_\infty$ and a **dissipation distance** $\mathcal{D} : \mathcal{Z} \times \mathcal{Z} \to [0,\infty]$. The energy $\mathcal{E}$ depends on the process time via a time-dependent loading. The value $\mathcal{D}(z_0, z_1)$ denotes the minimal dissipated energy when the state is changed from $z_0 \in \mathcal{Z}$ to $z_1 \in \mathcal{Z}$. Because of this physical interpretation, we require $\mathcal{D}$ to be a **quasimetric**, i.e. we require the triangle inequality and the positivity $\mathcal{D}(z_1, z_2) = 0$ if and only if $z_1 = z_2$ to hold[4]. We do not, however, require $\mathcal{D}$ to be symmetric as the physical dissipation might not have this property (consider, for example, crack formation in brittle materials [FM98, DFT05] or damage and delamination [KMR06]). Although $\mathcal{D}$ acts on the dissipative part $\mathcal{Z}$ of the underlying state-space $\mathcal{Q}$ only, for $q_1 = (\varphi_1, z_1), q_2 = (\varphi_2, z_2) \in \mathcal{Q}$ we also write $\mathcal{D}(q_1, q_2)$ when in fact we mean $\mathcal{D}(z_1, z_2)$.

For a process $z \in \mathcal{Z}[0,T]$ (only in the dissipation part of the state-space) and $s, t \in [0,T]$, define the **total dissipation** $\mathrm{Diss}_\mathcal{D}(z; [s,t])$ of z in the subinterval $[s,t]$ to be the total variation of z with respect to the quasimetric $\mathcal{D}$, i.e.

$$\mathrm{Diss}_\mathcal{D}(z; [s,t]) := \sup \left\{ \sum_{j=1}^{N} \mathcal{D}(z(\tau_{j-1}), z(\tau_j)) \;:\; s = \tau_0 < \cdots < \tau_N = t,\ N \in \mathbb{N} \right\}.$$

Again, for a process $q : [0,T] \to \mathcal{F} \times \mathcal{Z} = \mathcal{Q}$ with $t \mapsto (\varphi(t), z(t))$, we also write $\mathrm{Diss}_\mathcal{D}(q; [s,t])$ when we really mean $\mathrm{Diss}_\mathcal{D}(z; [s,t])$.

Lemma 2.3. *The dissipation is additive, i.e. for all $r < s < t$ and $z \in \mathcal{Z}[0,T]$, it holds that*

$$\mathrm{Diss}_\mathcal{D}(z; [r,t]) = \mathrm{Diss}_\mathcal{D}(z; [r,s]) + \mathrm{Diss}_\mathcal{D}(z; [s,t]).$$

Proof. By combining any two partitions of $[r,s]$ and $[s,t]$, respectively, into one, we get the inequality with "$\geq$". The other direction follows with the triangle inequality since for any partition $r = \tau_0 < \cdots < \tau_N = t$ of $[r,t]$ such that there exists $m \in \{0, \ldots, N-1\}$

[4] In fact, we only need semi-positivity, i.e. $\mathcal{D}(z,z) = 0$.

with $\tau_m < s < \tau_{m+1}$ (this is the only case we need to explicitly consider), we have

$$\begin{aligned}\sum_{j=1}^{N} \mathcal{D}(z(\tau_{j-1}), z(\tau_j)) &\le \sum_{j=1}^{m} \mathcal{D}(z(\tau_{j-1}), z(\tau_j)) + \mathcal{D}(z(\tau_m), z(s)) \\ &\qquad + \mathcal{D}(z(s), z(\tau_{m+1}) + \sum_{j=m+2}^{N} \mathcal{D}(z(\tau_{j-1}), z(\tau_j)) \\ &\le \mathrm{Diss}_{\mathcal{D}}(z; [r,s]) + \mathrm{Diss}_{\mathcal{D}}(z; [s,t])\end{aligned}$$

and hence, by taking the supremum over all such partitions, we see the converse inequality with "$\le$". □

If the dissipation is given by a potential, the total dissipation takes a simple form for sufficiently regular functions:

Proposition 2.4. *Let $\mathcal{Q}$ be a Banach space and assume that the dissipation distance is given through $\mathcal{D}(q_1, q_2) = \mathcal{R}(q_2 - q_1)$ for a convex, (strongly) continuous and positively 1-homogeneous dissipation potential $\mathcal{R} : \mathcal{Q} \to [0, \infty]$. Then, for $q \in \mathrm{W}^{1,1}(s, t; \mathcal{Q})$ it holds that*

$$\mathrm{Diss}_{\mathcal{D}}(q; [s,t]) = \int_s^t \mathcal{R}(\dot q(\tau))\, \mathrm{d}\tau.$$

Proof. By density of $\mathrm{C}^1([s,t]; \mathcal{Q})$ in $\mathrm{W}^{1,1}(s,t;\mathcal{Q})$ [Rou05, Lemma 7.2] and by the strong continuity of the integral $\int_s^t \mathcal{R}(\dot q(\tau))\,\mathrm{d}\tau$ (one can assume that a subsequence converges strongly almost everywhere) it suffices to consider the case $q \in \mathrm{C}^1([s,t];\mathcal{Q})$. For such q, $\tau \mapsto \mathcal{R}(\dot q(\tau))$ is Riemann-integrable and by the 1-homogeneity of $\mathcal{R}$ and the mean value theorem we get for any partition $\Pi = (s = \tau_0, \ldots, \tau_N = t)$

$$D(\Pi) := \sum_{j=1}^{N} \mathcal{D}(q(\tau_{j-1}), q(\tau_j)) = \sum_{j=1}^{N} \mathcal{R}\left(\frac{q(\tau_j) - q(\tau_{j-1})}{\Delta\tau_j}\right)\Delta\tau_j = \sum_{j=1}^{N} \mathcal{R}(\dot q(\xi_j))\Delta\tau_j,$$

where $\Delta\tau_j := \tau_j - \tau_{j-1}$ and $\xi_j \in (\tau_{j-1}, \tau_j)$. Now, let Π_k be a sequence of partitions with $D(\Pi_k) \uparrow \mathrm{Diss}_{\mathcal{D}}(q; [s,t])$. Without loss of generality we can assume that the fineness of Π_k converges to zero, since by the triangle inequality insertion of points may only increase the value of $D(\Pi_k)$. Hence, the $D(\Pi_k)$ are Riemann sums, which implies $D(\Pi_k) \to \int_s^t \mathcal{R}(\dot q(\tau))\,\mathrm{d}\tau$ and the claim follows. □

An **energetic solution** to the evolution system associated with $\mathcal{E}$ and $\mathcal{D}$ is a process $q \in \mathcal{Q}[0,T]$ that satisfies the **stability condition** (S) and the **energy balance** (E) for all $t \in [0,T]$:

$$\text{(i)}\quad \mathcal{E}(t, q(t)) \le \mathcal{E}(t, \hat q) + \mathcal{D}(q(t), \hat q) \qquad \text{for all } \hat q \in \mathcal{Q}, \tag{S}$$

$$\text{(ii)}\quad \mathcal{E}(t, q(t)) + \mathrm{Diss}_{\mathcal{D}}(q; [0,t]) = \mathcal{E}(0, q(0)) + \int_0^t \partial_t \mathcal{E}(\tau, q(\tau))\,\mathrm{d}\tau. \tag{E}$$

It is obvious how one can generalize the time inteval to $[S,T]$ where $S < T$ (to conform to the formal definition of a multi-valued evolutionary system, cf. Proposition 2.5 below.

The so-called **stability sets**

$$\mathcal{S}(t) := \left\{ q \in \mathcal{Q} \;:\; \mathcal{E}(t,q) < \infty \text{ and } \mathcal{E}(t,q) \le \mathcal{E}(t,\hat{q}) + \mathcal{D}(q,\hat{q}) \text{ for all } \hat{q} \in \mathcal{Q} \right\}$$

play a vital role in the theory and allow condition (S) to be rephrased into

$$q(t) \in \mathcal{S}(t) \qquad \text{for all } t \in [0,T]. \tag{S'}$$

Additionally to (S) & (E), we prescribe a stable initial value $q(0) = q_0 \in \mathcal{S}(0)$.

We can now prove that the energetic solution functor S_H for a system $\mathsf{H} = (\mathcal{Q}, \mathcal{E}, \mathcal{D})$ indeed defines a rate-independent multi-valued evolutionary system as specified in Definitions 2.1 and 2.2 if we think of the loadings to be the only time-dependent part in $\mathcal{E}$ (or we need to consider the full energy functional as "loading", but this is only a formal problem). For this proposition, we consider the problem on the time interval $[S,T]$ and not only on $[0,T]$.

Proposition 2.5. *The energetic formulation of the system* $\mathsf{H} = (\mathcal{Q}, \mathcal{E}, \mathcal{D})$ *defines a rate-independent multi-valued evolutionary system.*

Proof. Condition (2.1) follows from the fact that (S) is a pointwise definition and (E) is additive in the solution interval by additivity of the integral and the dissipation (Lemma 2.3).

To see the rate-independence, let $\alpha \in \mathrm{C}^1([S,T])$ with $\dot{\alpha} > 0$ and first observe that $\mathrm{Diss}(q \circ \alpha; [S,t]) = \mathrm{Diss}(q; [\alpha(S), \alpha(t)])$ for all $t \in [S,T]$ since α induces a bijection between all partitions $\Pi = (\tau_j)_j$ on $[S,t]$ and those on $[\alpha(S), \alpha(t)]$ through $\alpha_* \Pi := (\alpha(\tau_j))_j$ and hence

$$\begin{aligned} \mathrm{Diss}_\mathcal{D}(q \circ \alpha; [S,t]) &= \sup \left\{ \sum_{j=1}^N \mathcal{D}(q(\alpha(\tau_{j-1})), q(\alpha(\tau_j))) \;:\; \Pi = (\tau_j)_j \right\} \\ &= \sup \left\{ \sum_{j=1}^N \mathcal{D}(q(\zeta_{j-1}), q(\zeta_j)) \;:\; \alpha_* \Pi = (\zeta_j)_j \right\} \\ &= \mathrm{Diss}_\mathcal{D}(q; [\alpha(S), \alpha(t)]). \end{aligned} \tag{2.4}$$

For a process $q \in \mathcal{Q}[\alpha(S), \alpha(T)]$, the energy balance (E) reads as

$$\mathcal{E}(\hat{t}, q(\hat{t})) + \mathrm{Diss}_\mathcal{D}(q; [\alpha(S), \hat{t}]) = \mathcal{E}(\alpha(S), q(\alpha(S))) + \int_{\alpha(S)}^{\hat{t}} \partial_t \mathcal{E}(\hat{\tau}, q(\hat{\tau})) \,\mathrm{d}\hat{\tau}$$

for all $\hat{t} \in [\alpha(S), \alpha(T)]$. Writing $\hat{t} = \alpha(t)$ and using (2.4) as well as the substitution $\hat{\tau} = \alpha(\tau)$, we can transform the last equality into the following statement, where the composition $\mathcal{E} \circ \alpha$ is understood to be acting on the time argument:

$$\begin{aligned}
&(\mathcal{E} \circ \alpha)(t, (q \circ \alpha)(t)) + \mathrm{Diss}_{\mathcal{D}}(q \circ \alpha; [S, t]) \\
&\quad = (\mathcal{E} \circ \alpha)(S, (q \circ \alpha)(S)) + \int_S^t \partial_t \mathcal{E}(\alpha(\tau), q(\alpha(\tau)))\dot{\alpha}(\tau)\,\mathrm{d}\tau \\
&\quad = (\mathcal{E} \circ \alpha)(S, (q \circ \alpha)(S)) + \int_S^t \partial_t (\mathcal{E} \circ \alpha)(\tau, q(\alpha(\tau)))\,\mathrm{d}\tau
\end{aligned}$$

This is nothing else than (E) in the interval $[S, T]$ for the process $q \circ \alpha$ and the transformed functional $\mathcal{E} \circ \alpha$. Furthermore, stability for $q \in \mathcal{Q}[\alpha(S), \alpha(T)]$ implies those of $q \circ \alpha$ for the modified energy functional $\mathcal{E} \circ \alpha$ by a simple transformation of variables. Reversing these transformations, we see that $q \in \mathsf{S}_{\mathsf{H}}[\alpha(S), \alpha(T)](l, q_0)$ if and only if $q \circ \alpha \in \mathsf{S}_{\mathsf{H}}[S, T](l \circ \alpha, q_0)$, i.e. condition (2.3) and H is rate-independent (note that as mentioned above, we consider the loading to be the only time-dependent part of $\mathcal{E}$ or the loading to be identified with the whole energy functional). □

In the following, we will show that if $\mathcal{E}$ is *convex* (and the dissipation has a potential), the energetic formulation indeed is an extension of the traditional formulations via the differential inclusion (DI) or, equivalently, via the evolutionary variational inequality (EVI). In contrast to (DI) and (EVI), however, the energetic formulation (S) & (E) is derivative-free (neither $\mathrm{D}\mathcal{E}$ nor $\dot{q}$ occur) and no linear structure of $\mathcal{Q}$ needs to be assumed. This allows for a treatment of more general problems in continuum mechanics, cf. Section 7 of [Mie05] for a survey.

The following proposition is adapted from [Mie07, Section 2.1], but more advanced results are presented in [MT04], where also functions of bounded variation are considered as possible solutions of (DI) or (EVI) and a fine network of relations between many different problem formulations is established.

Proposition 2.6. *Let $\mathcal{Q}$ be a Banach space and assume that $\mathcal{E}(\bullet, \bullet)$ is (strongly) continuous (jointly in both arguments) and $\mathcal{E}(t, \bullet)$ is Gâteaux-differentiable for all $t \in [0, T]$. Further, let the dissipation distance be given through $\mathcal{D}(q_1, q_2) = \mathcal{R}(q_2 - q_1)$ for a convex, (strongly) continuous and positively 1-homogeneous dissipation potential $\mathcal{R} : \mathcal{Q} \to [0, \infty]$. Then, an energetic solution $q \in \mathrm{W}^{1,1}(0, T; \mathcal{Q})$ of* (S) *&* (E) *also solves* (EVI) *(or, equivalently,* (DI)*). If additionally $\mathcal{E}(t, \bullet)$ is convex for all $t \in [0, T]$, then the converse implication holds as well.*

Proof. Let $q \in \mathrm{W}^{1,1}(0, T; \mathcal{Q})$ be an energetic solution. The stability condition (S) for $\hat{q} := q(t) + hv$ with $v \in \mathcal{Q}$ and $h > 0$ gives for all $t \in [0, T]$

$$\frac{\mathcal{E}(t, q(t) + hv) - \mathcal{E}(t, q(t))}{h} + \frac{\mathcal{R}(hv)}{h} \geq 0.$$

Using the 1-homogeneity of $\mathcal{R}$ and letting $h \downarrow 0$, we get

$$\langle \mathrm{D}\mathcal{E}(t, q(t)), v\rangle + \mathcal{R}(v) \geq 0. \tag{2.5}$$

By Proposition 2.4, $\mathrm{Diss}_{\mathcal{D}}(q;[0,t]) = \int_0^t \mathcal{R}(\dot q(\tau))\,\mathrm{d}\tau$, whence the energy balance (E) reads

$$\mathcal{E}(t,q(t)) + \int_0^t \mathcal{R}(\dot q(\tau))\,\mathrm{d}\tau = \mathcal{E}(0,q(0)) + \int_0^t \partial_t\mathcal{E}(\tau,q(\tau))\,\mathrm{d}\tau.$$

Differentiating with respect to t and cancelling terms, we have (by the chain rule)

$$\langle \mathrm{D}\mathcal{E}(t,q(t)),\dot q(t)\rangle + \mathcal{R}(\dot q(t)) = 0.$$

Subtracting this from (2.5) gives

$$\langle \mathrm{D}\mathcal{E}(t,q(t)),v-\dot q(t)\rangle + \mathcal{R}(v) - \mathcal{R}(\dot q(t)) \geq 0,$$

which is (EVI).

We now turn to the converse implication and additionally assume $\mathcal{E}(t,\bullet)$ convex. Then, (EVI) with $v := h^{-1}w$, where $w \in \mathcal{Q}$ and $h > 0$, reads after multiplication with h

$$\langle \mathrm{D}\mathcal{E}(t,q(t)),w-h\dot q(t)\rangle + \mathcal{R}(w) - h\mathcal{R}(\dot q(t)) \geq 0$$

for almost all $t \in [0,T]$. Letting $h \downarrow 0$,

$$\langle \mathrm{D}\mathcal{E}(t,q(t)),w\rangle + \mathcal{R}(w) \geq 0 \tag{2.6}$$

and the convexity of $\mathcal{E}(t,\bullet)$ implies

$$\mathcal{E}(t,v) - \mathcal{E}(t,q(t)) + \mathcal{R}(v-q(t)) \geq \langle \mathrm{D}\mathcal{E}(t,q(t)),v-q(t)\rangle + \mathcal{R}(v-q(t)) \geq 0,$$

were the last inequality follows from (2.6) with $w := v - q(t)$ for $v \in \mathcal{Q}$. This is (S) for $t \in A \subseteq [0,T]$, where A has full measure in $[0,T]$. To show (S) for the remaining $t \in [0,T] \setminus A$, choose a sequence $(t_j)_j \subseteq A$ with $t_j \to t$. Since $q \in \mathrm{W}^{1,1}(0,T;\mathcal{Q})$, (a member of the equivalence class of) this function is actually continuous, see [Rou05, lemma 7.1]. For $\hat q \in \mathcal{Q}$ arbitrary, it therefore holds that $\hat q_j := q(t_j) + \hat q - q(t) \to \hat q$ by continuity. Because $q(t_j) \in \mathcal{S}(t_j)$

$$\mathcal{E}(t_j,q(t_j)) \leq \mathcal{E}(t_j,\hat q_j) + \mathcal{R}(\hat q_j - q(t_j))$$

and passing with $j \to \infty$ in this inequality yields

$$\mathcal{E}(t,q(t)) \leq \mathcal{E}(t,\hat q) + \mathcal{R}(\hat q - q(t)),$$

because $\hat q_j - q(t_j) = \hat q - q(t)$ for all $j \in \mathbb{N}$. Thus, (S) holds for all $t \in [0,T]$.

Now, (EVI) with $v := 0$ and (2.6) with $w := \dot q(t)$ are the "$\leq$"- and the "$\geq$"-part, respectively, of

$$\langle \mathrm{D}\mathcal{E}(t,q(t)),\dot q(t)\rangle + \mathcal{R}(\dot q(t)) = 0.$$

Adding $\partial_t\mathcal{E}(t,q(t))$ on both sides and using the chain rule backwards, the equality

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t,q(t)) + \mathcal{R}(\dot q(t)) = \partial_t\mathcal{E}(t,q(t)).$$

is established, which integrated from 0 to $t \in [0,T]$ is (E). $\square$

In this work, we also need to consider sequences of rate-independent systems. These are given through sequences of energy functionals $(\mathcal{E}_k)_k$ and corresponding dissipation distances $(\mathcal{D}_k)_k$. For the kth problem, $k \in \mathbb{N}_\infty$, we denote by (S_k) and (E_k) the solution conditions corresponding to (S) and (E), respectively. We also abbreviate Diss_k for $\mathrm{Diss}_{\mathcal{D}_k}$.

2.2 Time-incremental problems

Construction of solutions, but other theoretical investigations as well rely on a time-incremental problem. This can be seen as a modified Rothe method for the variational inequality (EVI), cf. [Rou05, Section 8.2] and [Kač85], and consists of a semidiscretization in time. As discretization scheme, we use the implicit Euler formula, cf. [Kre98, Chapter 10] or [QV94, Section 11.3] for other applications to the numerical approximation of parabolic ordinary and partial differential equations. Of course, here the methods are adapted to the setting of a variational *inequality* (cf. [Rou05, Section 11.1] for this scenario).

For a partition $\Pi = (0 = t_0, \dots, t_N = T)$ of the interval $[0, T]$, define the **fineness** $\|\Pi\|$ of the partition Π as

$$\|\Pi\| := \max_{j=1,\dots,N} \Delta t_j, \qquad \text{where } \Delta t_j := t_j - t_{j-1}.$$

To heuristically derive the time-incremental problem, we can set up a discretization scheme for the variational inequality (EVI),

$$\langle \mathrm{D}\mathcal{E}(t, q(t)), v - \dot{q}(t)\rangle + \mathcal{R}(v) - \mathcal{R}(\dot{q}(t)) \geq 0 \qquad \text{for all } v \in \mathcal{Q},$$

where $\mathcal{R}(q_2 - q_1) = \mathcal{D}(q_1, q_2)$, on an equidistant partition $\Pi = (t_j = jh)_{j=0,\dots,N}$ with timestep $h = T/N$ for some $N \in \mathbb{N}$. Writing the variational inequality at time $t_j = jh$, $j \in \{0, 1, \dots, N\}$, and replacing $q(t_j)$ by the approximation q_j and the derivative $\dot{q}(t_j)$ by the backward difference $h^{-1}(q_j - q_{j-1})$, we get

$$\Big\langle \mathrm{D}\mathcal{E}(t_j, q_j), v - \frac{q_j - q_{j-1}}{h} \Big\rangle + \mathcal{R}(v) - \mathcal{R}\Big(\frac{q_j - q_{j-1}}{h}\Big) \geq 0 \qquad \text{for all } v \in \mathcal{Q}.$$

Now, inserting $v := (\hat{q} - q_{j-1})/h$, $\hat{q} \in \mathcal{Q}$, and using the 1-homogeneity of $\mathcal{R}$, this is equivalent to

$$\langle \mathrm{D}\mathcal{E}(t_j, q_j), \hat{q} - q_j\rangle + \mathcal{R}(\hat{q} - q_{j-1}) - \mathcal{R}(q_j - q_{j-1}) \geq 0 \qquad \text{for all } \hat{q} \in \mathcal{Q}.$$

Rewriting this with $\mathcal{D}$ instead of $\mathcal{R}$, we get

$$\langle \mathrm{D}\mathcal{E}(t_j, q_j), \hat{q} - q_j\rangle + \mathcal{D}(q_{j-1}, \hat{q}) \geq \langle \mathrm{D}\mathcal{E}(t_j, q_j), \hat{q} - q_j\rangle + \mathcal{D}(q_{j-1}, \hat{q}) \geq \mathcal{D}(q_{j-1}, q_j).$$

Adding $\mathcal{E}(t_j, q_j)$ on both sides of this inequality we arrive at

$$\mathcal{E}(t_j, \hat{q}) + \mathcal{D}(q_{j-1}, \hat{q}) \gtrapprox \mathcal{E}(t_j, q_j) + \mathcal{D}(q_{j-1}, q_j) \qquad \text{for all } \hat{q} \in \mathcal{Q},$$

because $\mathcal{E}(t_j, \hat{q}) \approx \mathcal{E}(t_j, q_j) + \langle \mathrm{D}\mathcal{E}(t_j, q_j), \hat{q} - q_j \rangle$ up to second order by Taylor's formular (in fact, if $\mathcal{E}$ is convex, then even "$\geq$" holds, which fits nicely in the above derivation and we can replace "$\gtrapprox$" by "$\geq$" in the last inequality). Hence, the last inequality suggests the following **time-incremental problem** for a partition $\Pi = (0 = t_0, \ldots, t_N = T)$ of the interval $[0, T]$ (now not necessarily equidistant anymore[5]):

$$\begin{cases} \text{For } j = 1, \ldots, N \text{ inductively find } q_j \in \mathcal{Q} \text{ such that} \\ q_j \in \operatorname{Argmin} \big\{ \mathcal{E}(t_j, \hat{q}) + \mathcal{D}(q_{j-1}, \hat{q}) \ : \ \hat{q} \in \mathcal{Q} \big\}. \end{cases} \tag{IP}$$

Under standard assumptions (see Section 2.4.1), these problems admit at least one solution.

Contrary to the above heuristics that started with the abstract variational inequality, the time-discrete problem is from a physical point of view the more natural place to begin with. It expresses the system's instanteneous movement to the (at the current time) energetically best point in the state space, i.e. the system tries to minimize its total energy, but doing this, it takes into account the necessary dissipation $\mathcal{D}(q_{j-1}, \hat{q})$ when changing the state from q_{j-1} to $\hat{q}$.

Later, we will rather work with the **approximate time-incremental problem**

$$\begin{cases} \text{For } j = 1, \ldots, N \text{ inductively find } q_j \in \mathcal{Q} \text{ such that} \\ q_j \in \operatorname{Argmin}_{\varepsilon \Delta t_j} \big\{ \mathcal{E}(t_j, \hat{q}) + \mathcal{D}(q_{j-1}, \hat{q}) \ : \ \hat{q} \in \mathcal{Q} \big\}. \end{cases} \tag{AIP}$$

Note that the grade of minimization depends on the local fineness of the partition. Obviously, this problem always has a solution. These approximte problems are more realistic from a physical point of view, because they can express *instablity* of the system (see Counterexample 3.1 on page 53).

To treat sequences of problems, **approximate time-incremental problems for sequences of systems** (also depending on $k \in \mathbb{N}$) are employed: Let $\varepsilon > 0$ and consider:

$$\begin{cases} \text{For } j = 1, \ldots, N \text{ inductively find } q_j^k \in \mathcal{Q} \text{ such that} \\ q_j^k \in \operatorname{Argmin}_{\varepsilon \Delta t_j} \big\{ \mathcal{E}_k(t_j, \hat{q}) + \mathcal{D}_k(q_{j-1}^k, \hat{q}) \ : \ \hat{q} \in \mathcal{Q} \big\}. \end{cases} \tag{AIP$_k$}$$

Most of the work in this thesis is about clarifying the relationship between the approximate time-incremental problem and the energetic formulation. The Existence Theorems 2.9 and 2.19 will show that from (discrete) solutions to (AIP) or (AIP$_k$) on finer and finer partitions (and simultaneously letting $\varepsilon \to 0$), one can construct solutions to (S) & (E) or (S$_\infty$) & (E$_\infty$), respectively. The reverse approximation results from Sections 3.1 and 3.2 will indicate that the other direction is possible as well.

[5] In fact, the above derivations work just the same if h is replaced by the time step $\Delta t_j = t_j - t_{j-1}$.

2.3 Existence for single systems

In this section, the main existence theorem for rate-independent systems is formulated. It will not be proved, though, because it is a special case of the existence theorem for sequences of systems in the next section. Other versions (with slightly different assumptions) can be found in [MT04, Mie05, FM06, MR07b].

Recall that all topological notions are to be understood in a sequential sense. Also, let the **approximate stability sets** be defined through

$$\mathcal{S}_k^\alpha(t) := \big\{\, q \in \mathcal{Q} \;\mathbf{:}\; \mathcal{E}_k(t,q) < \infty \text{ and } \mathcal{E}_k(t,q) \le \mathcal{E}_k(t,\hat{q}) + \mathcal{D}(q,\hat{q}) + \alpha \text{ for all } \hat{q} \in \mathcal{Q} \,\big\}.$$

We assume:

Uniform control of the power[6]*:*
There exist $c_0^E \in \mathbb{R}, c_1^E > 0$ such that :
If $q \in \mathcal{Q}$ satisfies $\mathcal{E}(s,q) < \infty$ for some $s \in [0,T]$, then (A1')
(i) $\mathcal{E}(\bullet, q) \in \mathrm{C}^1([0,T])$ and
(ii) $|\partial_t \mathcal{E}(t,q)| \le c_1^E(\mathcal{E}(t,q) + c_0^E)$ for all $t \in [0,T]$.

Uniform time-continuity of the power:
For all $\varepsilon > 0$ and all $E \in \mathbb{R}$, there exists $\delta > 0$ such that:
If for $q \in \mathcal{Q}$ with $\mathcal{E}(0,q) < E$ and $t_1, t_2 \in [0,T]$ it holds that $|t_1 - t_2| < \delta$, (A2')
then $|\partial_t \mathcal{E}(t_1,q) - \partial_t \mathcal{E}(t_2,q)| < \varepsilon$.

Compactness of energy sublevels:
For all $t \in \mathcal{Q}$ and $E \in \mathbb{R}$: The set $\big\{\, q \in \mathcal{Q} \;\mathbf{:}\; \mathcal{E}(t,q) \le E \,\big\}$ is compact. (A3')

Conditioned joint continuity of the power:
For all sequences $t_k \to t$ and $q_k \to q$ in $\mathcal{Q}$ with $q \in \mathcal{S}^{\alpha_k}(0), \alpha_k \downarrow 0$ (A4')
and $\sup_k \mathcal{E}(t_k, q_k) < \infty : \partial_t \mathcal{E}(t_k, q_k) \to \partial_t \mathcal{E}_\infty(t,q)$.

Quasimetric:
For all $z_1, z_2, z_3 \in \mathcal{Z}$: (A5')
(i) $\mathcal{D}(z_1, z_2) = 0$ if and only if $z_1 = z_2$ **(positivity)** and
(ii) $\mathcal{D}(z_1, z_3) \le \mathcal{D}(z_1, z_2) + \mathcal{D}(z_2, z_3)$ **(triangle inequality)**.

Continuity of the dissipation:
For all sequences $(q_k)_k, (\tilde{q}_k)_k \subseteq \mathcal{Q}$ with $\sup_{k\in\mathbb{N}}(\mathcal{E}(s,q) + \mathcal{E}(s,\tilde{q})) < \infty$ (A6')
for one (hence all) $s \in [0,T]$:
If $q_k \to q$ and $\tilde{q}_k \to \tilde{q}$, then $\mathcal{D}(q_k, \tilde{q}_k) \to \mathcal{D}_\infty(q, \tilde{q})$.

[6] In some applications, $\partial_t \mathcal{E}$ can be interpreted as a (physical) power.

Example 2.7. Let $\mathcal{Q}$ be a Banach space endowed with its weak topology and let $\mathcal{Q}$ be compactly embedded into another Banach space $\mathcal{Q}_1$. Since $\mathcal{Q} \subseteq \mathcal{Q}_1$, we can choose the $\mathcal{Q}_1$–Norm $\|\bullet\|_1$ as our dissipation distance, i.e. $\mathcal{D}(u, v) := \|v - u\|_1$. Assumption (A5') holds by the properties of a norm while the compact embedding ensures the continuity (A6') of $\mathcal{D}$.

Example 2.8. Let Ω be a bounded Lipschitz domain. By the Rellich–Kondrachov Compactness Theorem (see e.g. [GT98, Theorem 7.26]), the last example shows that we can use the $\mathrm{L}^1(\Omega)$–Norm as the continuous dissipation distance in $\mathrm{H}^1(\Omega)$ (or more generally, in $\mathrm{W}^{1,p}(\Omega)$ for any $p \geq 1$), which is a common situation in rate-independent continuum mechanics [Mie05].

For a partition $\Pi = (0 = t_0, \ldots, t_N = T)$, recall the approximate incremental problem (also see page 32):

$$\begin{cases} \text{For } j = 1, \ldots, N \text{ inductively find } q_j \in \mathcal{Q} \text{ such that} \\ q_j \in \operatorname{Argmin}_{\varepsilon \Delta t_j} \big\{ \mathcal{E}(t_j, \hat{q}) + \mathcal{D}(q_{j-1}, \hat{q}) \ : \ \hat{q} \in \mathcal{Q} \big\}. \end{cases} \tag{$\mathrm{AIP}_{\varepsilon,\Pi}$}$$

Then, the main existence theorem reads as:

Theorem 2.9. *Let the above assumptions* (A1')–(A6') *be fulfilled. Assume further that we are given a sequence of partitions $\Pi_k = (0 = t_0^k, \ldots, t_{N(k)}^k = T)$ of the interval $[0, T]$ with $\|\Pi_k\| \to 0$, a falling sequence $(\varepsilon_k)_k$ of positive real numbers with $\varepsilon_k \to 0$, and a sequence of initial values $(q_0^k)_k \subseteq \mathcal{Q}$ satisfying the* **compatibility condition**

$$q_0^k \in \mathcal{S}^{\varepsilon_k \|\Pi_k\|}(0), \qquad q_0^k \to q_0 \ \text{in } \mathcal{Q}, \qquad \text{and} \qquad \mathcal{E}(0, q_0^k) \to \mathcal{E}(0, q_0). \tag{CC}$$

Then, the piecewise-constant, right-continuous interpolants $\bar{q}^k$ of discrete solutions $q^k = (q_j^k)_j$ to $(\mathrm{AIP}_{\varepsilon_k,\Pi_k})$ *admit a subsequence q^{k_n} and a solution $q = (\varphi, z) : [0, T] \to \mathcal{F} \times \mathcal{Z} = \mathcal{Q}$ of* (S) *&* (E) *to the initial value q_0 with*

(i) $\mathcal{E}(t, \bar{q}^{k_n}(t)) \to \mathcal{E}(t, q(t))$ *for all* $t \in [0, T]$,

(ii) $\mathrm{Diss}_{\mathcal{D}}(\bar{q}^{k_n}; [0, t]) \to \mathrm{Diss}_{\mathcal{D}}(q; [0, t]) < \infty$ *for all* $t \in [0, T]$,

(iii) $z^{k_n}(t) \to z(t)$ *in* $\mathcal{Z}$ *for all* $t \in [0, T]$,

(iv) $\partial_t \mathcal{E}(\bullet, \bar{q}^{k_n}(\bullet)) \to \partial_t \mathcal{E}(\bullet, q(\bullet))$ *in* $\mathrm{L}^1(0, T)$,

(v) For all $t \in [0, T]$, *there exists a* t*-dependent subsequence* $(K_n^t)_n$ *of* k_n *such that* $\bar{\varphi}^{K_n^t}(t) \to \varphi(t)$ *in* $\mathcal{F}$.

Further, any limit of a subsequence obtained in this way solves (S) *&* (E). *If additionally the topology on $\mathcal{Q}$ is separable and metrizable on compact sets, φ (hence q) can be chosen measurable.*

Example 2.10. Weak topologies on separable, reflexive Banach spaces are metrizable on compact sets, see [Con90, Theorem 5.1].

As mentioned before, this theorem is a special case of Theorem 2.19 below. In fact, it can be shown that assumption (A4') is not needed if $q_0 \in \mathcal{S}(0)$, but this requires a slightly different proof, which can be found in [Mie05, Theorem 5.2].

2.4 Existence for sequences of problems

This section is devoted to proving existence of solutions to the energetic formulation of a rate-independent system, which is the "limit" of a sequence of such systems. This result contains Theorem 2.9 from the previous section as a special case. The existence result was first established in [MT99, MT04] and subsequently improved in [MM05, FM06]. Existence for Γ-converging sequences of problems was first established in [MRS08]. We follow the argument from the latter source, but combine their Theorems 3.3 and 4.1 into one.

2.4.1 Assumptions

In all of the following, the stored energy functionals and the dissipation distances are assumed to satisfy the following requirements:

Uniform control of the power:
There exist $c_0^E \in \mathbb{R}, c_1^E > 0$ such that for all $k \in \mathbb{N}_\infty$:
If $q \in \mathcal{Q}$ satisfies $\mathcal{E}_k(s,q) < \infty$ for some $s \in [0,T]$, then (A1)
(i) $\mathcal{E}_k(\bullet, q) \in \mathrm{C}^1([0,T])$ and
(ii) $|\partial_t \mathcal{E}_k(t,q)| \le c_1^E(\mathcal{E}_k(t,q) + c_0^E)$ for all $t \in [0,T]$.

Uniform time-continuity of the power:
For all $\varepsilon > 0$ and all $E \in \mathbb{R}$, there exists $\delta > 0$ such that:
If for $q \in \mathcal{Q}$ with $\mathcal{E}_\infty(0,q) < E$ and $t_1, t_2 \in [0,T]$ it holds that $|t_1 - t_2| < \delta$, (A2)
then $|\partial_t \mathcal{E}_\infty(t_1,q) - \partial_t \mathcal{E}_\infty(t_2,q)| < \varepsilon$.

Compactness of energy sublevels:
For all $t \in \mathcal{Q}$ and $E \in \mathbb{R}$: The sublevel set (A3)

$$\bigcup_{k\in\mathbb{N}} \big\{ q \in \mathcal{Q} \ : \ \mathcal{E}_k(t,q) \le E \big\} \text{ is relatively compact.}$$

Conditioned continuous convergence of the power:
For all sequences $t_k \to t$ and $q_k \to q$ in $\mathcal{Q}$ with $q_k \in \mathcal{S}_k^{\alpha_k}(t_k), \alpha_k \downarrow 0$ (A4)
and $\sup_k \mathcal{E}_k(t_k,q_k) < \infty : \partial_t \mathcal{E}_k(t_k,q_k) \to \partial_t \mathcal{E}_\infty(t,q)$.

Quasimetric:

$$\begin{aligned}&\text{For all } k \in \mathbb{N}_\infty \text{ and } z_1, z_2, z_3 \in \mathcal{Z} :\\ &\text{(i)}\quad \mathcal{D}_k(z_1, z_2) = 0 \text{ if and only if } z_1 = z_2 \text{ (positivity) and}\\ &\text{(ii)}\quad \mathcal{D}_k(z_1, z_3) \le \mathcal{D}_k(z_1, z_2) + \mathcal{D}_k(z_2, z_3) \text{ (triangle inequality).}\end{aligned} \tag{A5}$$

By Gronwall's inequality applied to (A1),

$$\mathcal{E}_k(t, q) + c_0^E \le (\mathcal{E}_k(s, q) + c_0^E)\mathrm{e}^{c_1^E|t-s|}, \tag{2.7}$$

$$|\partial_t \mathcal{E}_k(t, q)| \le c_1^E(\mathcal{E}_k(s, q) + c_0^E)\mathrm{e}^{c_1^E|t-s|} \tag{2.8}$$

for all $k \in \mathbb{N}_\infty$, $q \in \mathcal{Q}$ and $s, t \in [0, T]$.

We also require the functionals $\mathcal{E}_k$ and $\mathcal{D}_k$ to converge to $\mathcal{E}_\infty$ and $\mathcal{D}_\infty$, respectively, in an appropriate sense:

Γ-*limit for the energy functionals:*

$$\text{For all } t \in [0, T] : \mathcal{E}_\infty(t, \bullet) = \Gamma\text{-}\lim_k \mathcal{E}_k(t, \bullet). \tag{Γ1}$$

Continuous convergence of the dissipation distances:

$$\begin{aligned}&\text{For all sequences } (q_k)_k, (\tilde{q}_k)_k \subseteq \mathcal{Q} \text{ with } \sup_{k\in\mathbb{N}}(\mathcal{E}_k(s, q_k) + \mathcal{E}_k(s, \tilde{q}_k)) < \infty\\ &\text{for one (hence all) } s \in [0, T] :\\ &\text{If } q_k \to q \text{ and } \tilde{q}_k \to \tilde{q}, \text{ then } \mathcal{D}_k(q_k, \tilde{q}_k) \to \mathcal{D}_\infty(q, \tilde{q}).\end{aligned} \tag{Γ2}$$

Note that conditions (Γ1) and (Γ2) together imply the joint Γ-convergence (on subsets of finite energy)

$$\mathcal{E}_\infty(t, \bullet) + \mathcal{D}_\infty(q, \bullet) = \Gamma\text{-}\lim_k(\mathcal{E}_k(t, \bullet) + \mathcal{D}_k(q, \bullet)) \tag{2.9}$$

for all $t \in [0, T]$ and $q \in \mathcal{Q}$ (this follows directly from the definition of Γ-convergence with the recovery sequence for the sequence $(\mathcal{E}_k(t, \bullet))_k$).

The Γ-convergence assumption (Γ1) is obviously satisfied if $\mathcal{E}_k(t, q^k) \to \mathcal{E}_\infty(t, q)$ for all $t \in [0, T]$ and all sequences $q^k \to q$ in $\mathcal{Q}$ (this is called **continuous convergence**).

Example 2.7 applies here *mutatis mutandis* if we take the dissipation distance to be the same for all $k \in \mathbb{N}_\infty$.

For all $E \in \mathbb{R}$, the the set $\{\, q \in \mathcal{Q} \,:\, \mathcal{E}_\infty(t, q) \le E \,\}$ is relatively compact by (A3). Hence, each sequence $(q_k)_k \subseteq K$ has a convergent subsequence (not relabelled) $q_k \to q \in \mathcal{Q}$ with corresponding $\mathcal{Z}$-components $z_k \to z \in \mathcal{Z}$. From ($\Gamma$2) and (A5), we get $\min\{\mathcal{D}_\infty(z_k, z), \mathcal{D}_\infty(z, z_k)\} \to 0$ and by the strict positivity of $\mathcal{D}_\infty(\bullet, \bullet)$, we can conclude that $\mathcal{D}$-convergence determines the limit, which can be rephrased into:

Positivity of the limit dissipation $\mathcal{D}_\infty$*:*

$$\begin{aligned}&\text{For all } E \in \mathbb{R} \text{ and } (z_k)_k \subseteq \{\, q \in \mathcal{Q} \,:\, \mathcal{E}_\infty(t, q) \le E \,\} :\\ &\text{If } \min\{\mathcal{D}_\infty(z_k, z), \mathcal{D}_\infty(z, z_k)\} \to 0, \text{ then } z_k \to z \text{ in } \mathcal{Z}.\end{aligned} \tag{2.10}$$

Note that it is possible to considerably weaken the convergence requirements ($\Gamma1$) and ($\Gamma2$), cf. [MRS08] (roughly, the Γ-limes inferior together with suitable energy bounds suffices), but here it is favorable to work with the above special case, because later when we want to investigate reverse approximation, we need the stronger conditions anyway. It is interesting to note that some "coupling" between the convergences for $\mathcal{E}_k$ and $\mathcal{D}_k$ must be present in order to conclude existence and convergence of solutions, as is shown in Example 3.2 of [MRS08]; in particular $\mathcal{E}_\infty(t,\bullet) = \Gamma\text{-lim}\,\mathcal{E}_k(t,\bullet)$ and $\mathcal{D}_\infty = \Gamma\text{-lim}\,\mathcal{D}_k$ is not enough.

2.4.2 A generalized Helly selection principle

As the crucial compactness result for functions with bounded dissipation ($\mathcal{D}$-variation), a refined Helly selection principle was employed in [MT04, MM05, MRS08] (the original Banach-space-valued version can be found in [BP86, Theorem 1.3.5]). Interestingly, the proof is related to the real-variable theory of functions of bounded variation, cf. [Nat75, VIII.§§3,4], and is not based on the "modern" functional analytic approach to functions of bounded variation, which is presented, for example, in [AFP00]. This is also imperative here, since in the modern treatment based on L^p-spaces, at most convergence almost everywhere can be expected. While this is no problem when additional regularity of solutions is to be expected, in the situation of rate-independent systems we need pointwise *everywhere* convergence results, because in general solutions will not have any additional regularity and changes on set of measure zero may change whether a function is a solution or not (in particular, the stability condition (S) is defined for *all* points of the time interval).

Following [Nat75, VIII.§§3,4], we first look at two results on real-valued functions of bounded variation as a preparation for the generalized Helly selection principle, Theorem 2.13.

Define the **pointwise variation** of a function $f : [a,b] \to \mathbb{R}$ to be the quantity

$$\mathrm{Var}(f;[a,b]) := \sup\left\{ \sum_{j=1}^{N} |f(\tau_j) - f(\tau_{j-1})| \;:\; a = \tau_0 < \cdots < \tau_N = b,\ N \in \mathbb{N} \right\}.$$

As one particular example, a bounded monotone function $f : [a,b] \to [-M,M]$ satisfies $\mathrm{Var}(f;[a,b]) \le 2M$. An interesting property of functions of bounded variations is the following:

Proposition 2.11. *A function $f : [a,b] \to \mathbb{R}$ with $\mathrm{Var}(f;[a,b]) < \infty$ has at most countably many discontinuity points.*

Proof. Let $a = \tau_0 < \cdots < \tau_N = b$, $N \in \mathbb{N}$ be any partition of $[a,b]$. Using the modified partition $a = \tau_0 < \tau_0 + \varepsilon < \tau_1 - \varepsilon < \tau_1 + \varepsilon < \cdots < \tau_N - \varepsilon < \tau_N = b$ with $\varepsilon > 0$ sufficiently small, we get

$$|f(\tau_0+\varepsilon)-f(\tau_0)|+\sum_{j=1}^{N}|f(\tau_j+\varepsilon)-f(\tau_j-\varepsilon)|+|f(\tau_N)-f(\tau_N-\varepsilon)|\le V,$$

where $V := \mathrm{Var}(f;[a,b]) < \infty$. Letting $\varepsilon \downarrow 0$ and writing $f(x \pm 0)$ for $\lim_{\varepsilon\downarrow 0} f(x\pm\varepsilon)$, this implies

$$|f(\tau_0+0)-f(\tau_0)|+\sum_{j=1}^{N}|f(\tau_j+0)-f(\tau_j-0)|+|f(\tau_N)-f(\tau_N-0)|\le V.$$

Thus, the set J_k of jump points $x \in [a,b]$ such that $|f(x+0)-f(x-0)| \ge 1/k$ (we let $a-0 := a$, $b+0 := b$) must be finite and their union $J = \bigcup_k J_k$ is countable. But J is nothing else than the set of discontinuity points of f. □

As noted above, monotone functions are of bounded variation and hence have only countably many discontinuity points, whence we can show the following compactness lemma for increasing functions[7]. This result is also one step in the proof of the classical Helly selection principle, cf. [Nat75, VIII.§4].

Lemma 2.12. *A uniformly bounded sequence $(f_k)_k$ of functions $f_k : [a,b] \to \mathbb{R}$ with each f_k increasing admits a subsequence converging pointwise to an increasing function $f : [a,b] \to \mathbb{R}$.*

Proof. Let $C := [a,b] \cap \mathbb{Q}$. Because C is countable, we can select a diagonal subsequence (not relabelled) such that $f_k(t) \to g(t)$ for all $t \in C$ with some function $g : C \to \mathbb{R}$. Since for any $t_1, t_2 \in C$ with $t_1 \le t_2$ we have $0 \le f_k(t_2) - f_k(t_1) \to g(t_2) - g(t_1)$, the function f is increasing and we can extend g on all of $[a,b]$ by defining

$$g(t) := \lim_{C\ni q\uparrow t} g(q) = \sup_{C\ni q<t} g(q) \qquad \text{for all } t \in [a,b]\setminus C.$$

The function $g : [a,b] \to \mathbb{R}$ is still increasing, which follows directly from the definition of the extension. Hence, by Proposition 2.11, g has an at most countable set of discontinuity points $D \subseteq [a,b]$. We claim that for all $t \in [a,b]\setminus D$, $f_k(t) \to g(t)$. To see this, take for fixed $t \in [a,b]\setminus D$ points $q_-, q_+ \in C$ with $q_- < t < q_+$ and $g(q_+) - g(q_-) \le \varepsilon/2$ (such q_-, q_+ exist by continuity). Because $f_k(q_\pm) \to g(q_\pm)$,

$$g(q_-) - \frac{\varepsilon}{2} \le f_k(q_-) \le f_k(t) \le f_k(q_+) \le g(q_+) + \frac{\varepsilon}{2}$$

for k sufficiently large. Because $g(q_+) - g(q_-) \le \varepsilon/2$, this implies

$$g(t) - \varepsilon \le f_k(t) \le g(t) + \varepsilon,$$

[7]Here, “increasing” means what often is called “nondecreasing”

whence $f_k \to g$ pointwise on $[a,b] \setminus D$. Since D is only countable, we can select a further (diagonal) subsequence (not relabelled) such that $f_k \to f$ on all of $[a,b]$, where $f|_{[a,b]\setminus D} = g|_{[a,b]\setminus D}$. By the same argument as used above for g, f is increasing. $\square$

Of course, the previous lemma also works *mutatis mutandis* in the case where all the f_k are decreasing.

The setup of the generalized Helly selection principle consists of a Hausdorff space $\mathcal{Z}$ and a sequence of dissipation distances $\mathcal{D}_k : \mathcal{Z} \times \mathcal{Z} \to \mathbb{R}_\infty$, $k \in \mathbb{N}_\infty$, satisfying the following three conditions:

$$\begin{aligned}&\text{For all } k \in \mathbb{N}_\infty \text{ and all } z_1, z_2, z_2 \in \mathcal{Z} :\\ &\mathcal{D}(z_1, z_1) = 0 \text{ and } \mathcal{D}_k(z_1, z_3) \le \mathcal{D}_k(z_1, z_2) + \mathcal{D}(z_2, z_3)\end{aligned} \tag{H1}$$

$$\begin{aligned}&\text{For all compact } K \subseteq \mathcal{Z} \text{ and all } (z_k)_k \subseteq K, z \in \mathcal{Z} :\\ &\text{If } \min\{\mathcal{D}_\infty(z_k, z), \mathcal{D}_\infty(z, z_k)\} \to 0, \text{ then } z_k \to z.\end{aligned} \tag{H2}$$

$$\text{For all } z_{k_j} \to z \text{ and } \tilde{z}_{k_j} \to \tilde{z} \text{ in } \mathcal{Z} : \mathcal{D}_\infty(z, \tilde{z}) \le \liminf_{j\to\infty} \mathcal{D}_{k_j}(z_{k_j}, \tilde{z}_{k_j}). \tag{H3}$$

Theorem 2.13. *Assume* (H1)–(H3). *Then every sequence of functions* $z_k : [0,T] \to \mathcal{Z}$ *with* $\sup_{k\in\mathbb{N}} \mathrm{Diss}_k(z_k; [0,T]) < \infty$ *and* $z_k(t) \subseteq K$ *for a fixed compact set* $K \subseteq \mathcal{Z}$ *and all* $t \in [0,T]$*, admits a subsequence* $(z_{k_j})_j$ *and a limit function* $z \in \mathcal{Z}[0,T]$ *with*

(i) $\lim_{j\to\infty} \mathrm{Diss}_{k_j}(z_{k_j}; [0,t]) =: \delta(t)$ *exists for all* $t \in [0,T]$ *and* $\delta : [0,T] \to [0,\infty)$ *is increasing and* $\delta(0) = 0$,

(ii) $z_{k_j}(t) \to z(t)$ *in* $\mathcal{Z}$ *for all* $t \in [0,T]$,

(iii) $\mathrm{Diss}_\infty(z; [s,t]) \le \delta(t) - \delta(s)$ *for all* $s, t \in [0,T]$, $s \le t$.

Proof. The functions $t \mapsto \mathrm{Diss}_k(z_k; [0,t])$ are increasing and uniformly bounded by assumption. Hence, Lemma 2.12 yields the existence of a sequence of indices k_j and an increasing function $\delta : [0,T] \to [0,\infty)$ such that (i) holds. Additionally, δ has at most countably many discontinuity points by Proposition 2.11, denoted by $E \subseteq [0,T]$.

Let $C \supseteq E$ be countable and dense in $[0,T]$ (for example $C := ([0,T] \cap \mathbb{Q}) \cup E$). For fixed $t \in C$, $(z_{k_j}(t))_j \subseteq K$ is relatively compact and by a diagonal argument, we can select a further subsequence (still denoted by k_j) such that $z_{k_j}|_C \to z$ pointwise for some function $z : C \to \mathcal{Z}$. More precisely, let $C = \{s_1, s_2, \ldots\}$. Then there is a subsequence k_j^1 (of k_j) such that $z_{k_j^1}(s_1)$ converges to some $z(s_1) \in \mathcal{Z}$. Repeating this process, in each step we select a subsequence k_j^i of k_j^{i-1} such that $z_{k_j^i}(s_i)$ converges to $z(s_i)$. Then the diagonal subsequence $\tilde{k}_j := k_j^j$ does the job.

We still have to define z on the points $t_* \in [0,T] \setminus C$, for which we will exploit the continuity of δ in t_*. Fix such a t_* and let $z_* \in \mathcal{Z}$ be an arbitrary accumulation point of $z_{k_j}(t_*)$. It suffices to show that for all sequences $(t_n)_n \subseteq C$ with $t_n \to t_*$ (of which at least one exists by the density of C) it holds that $z(t_n) \to z_*$, because then as $\mathcal{Z}$ is assumed Hausdorff and the choice of accumulation point was arbitrary, there can have been only one accumulation point in the first place.

Let therefore $t_n \to t_*$ and assume (selecting a further subsequence) that $z_{k_j}(t_*) \to z_*$. If $t_n < t_*$, by (H3) we get

$$\begin{aligned}\mathcal{D}_\infty(z(t_n), z_*) &\leq \liminf_{j\to\infty} \mathcal{D}_{k_j}(z_{k_j}(t_n), z_{k_j}(t_*)) \leq \liminf_{j\to\infty} \mathrm{Diss}_{k_j}(z_{k_j}; [t_n, t_*]) \\ &\leq \delta(t_*) - \delta(t_n)\end{aligned}$$

(note that for the last part we employed the additivity of the dissipation, cf. Lemma 2.3) and a similar estimate holds for $t_* < t_n$. Combining, we get

$$\min\{\mathcal{D}_\infty(z(t_n), z_*), \mathcal{D}_\infty(z_*, z(t_n))\} \leq |\delta(t_*) - \delta(t_n)| \to 0 \qquad \text{as } n \to \infty,$$

since δ is continuous in t_*. Thus, (H2) yields $z(t_n) \to z_*$ in $\mathcal{Z}$ and since this must hold for all accumulation points z_* of $z_{k_j}(t_*)$, there is only one such accumulation point and we can unambiguously extend z by defining $z(t_*) := z_*$. In this matter, we can define z on all of $[0, T]$. This establishes (ii).

It remains to show (iii). For any partition $s = \tau_0, \tau_1, \ldots, \tau_N = t$ of the interval $[s, t]$, again using (H3), we have

$$\begin{aligned}\sum_{l=1}^{N} \mathcal{D}_\infty(z(\tau_{l-1}), z(\tau_l)) &\leq \sum_{l=1}^{N} \liminf_{j\to\infty} \mathcal{D}_{k_j}(z_{k_j}(\tau_{l-1}), z_{k_j}(\tau_l)) \\ &\leq \liminf_{j\to\infty} \sum_{l=1}^{N} \mathcal{D}_{k_j}(z_{k_j}(\tau_{l-1}), z_{k_j}(\tau_l)) \\ &\leq \liminf_{j\to\infty} \mathrm{Diss}_{k_j}(z_{k_j}; [s, t]) = \delta(t) - \delta(s).\end{aligned}$$

Taking the supremum over all such partitions, (iii) follows. $\square$

In the special case $\mathcal{Z} = \mathbb{R}$ and the constant sequence $\mathcal{D}(z_1, z_2) := |z_2 - z_1|$ (which obviously fulfills (H1)–(H3)), we recover the classic Helly selection principle:

Corollary 2.14. *A sequence $(f_k)_k$ of uniformly bounded functions $f_k : [a, b] \to \mathbb{R}$ with $M := \sup_{k\in\mathbb{N}} \mathrm{Var}(f_k; [a, b]) < \infty$ admits a subsequence converging pointwise to a limit function $f : [a, b] \to \mathbb{R}$ with $\mathrm{Var}(f; [a, b]) \leq M$.*

2.4.3 Auxiliary results

Hans Hahn showed in [Hah14] that, roughly said, Riemann integrals and Lebesgue integrals on the real line differ only by an exchange of the quantor "∀" to "∃" in the quantification over Riemann sums converging to the integral. More precisely, even for Lebesgue integrals, there always exists at least one Riemann sum converging to the integral (in fact, even more is true, "most" equidistant partitions work).

Lemma 2.15. *Let $u \in \mathrm{L}^1(a,b)$. Then, there exists a sequence of partitions $\Pi_k = (a = \tau_0^k, \dots, \tau_{N(k)}^k = b)$ with $\|\Pi_k\| \to 0$ such that the upper Riemann sums of u associated with the Π_k converge to $\int_a^b u(\tau)\,\mathrm{d}\tau$, i.e.*

$$\sum_{j=1}^{N(k)} u(\tau_j^k)\Delta\tau_j^k \longrightarrow \int_a^b u(\tau)\,\mathrm{d}\tau.$$

A proof can be found either in [Mai04, Theorem A.3.1] or [DFT05, Lemma 4.12] (the latter also contains a generalization to Bochner-integrable functions).

The next proposition shows that for stable processes the "$\geq$"-part in the energy balance (E) is always true and we only need to show the opposite inequality. The proof here is from [MRS08, Proposition 2.4].

Proposition 2.16. *Let $q \in \mathcal{Q}[0,T]$ be a stable process, i.e. $q(t) \in \mathcal{S}_\infty(t)$ for all $t \in [0,T]$, and assume that $\tau \mapsto \partial_t\mathcal{E}_\infty(\tau, q(\tau); u) \in \mathrm{L}^1(0,T)$. Then, for all subintervals $[s,t] \subseteq [0,T]$, the process q satisfies the lower energy estimate*

$$\mathcal{E}_\infty(t, q(t)) + \mathrm{Diss}_\infty(q; [s,t]) \geq \mathcal{E}_\infty(s, q(s)) + \int_s^t \partial_t\mathcal{E}_\infty(\tau, q(\tau))\,\mathrm{d}\tau.$$

Proof. For a partition $s = \tau_0 < \tau_1 < \dots < \tau_N = t$ of the interval $[s,t]$ (which will be chosen later) we test the stability of $q(\tau_{j-1})$ with $q(\tau_j)$ to see

$$\begin{aligned}
\mathcal{E}_\infty(\tau_{j-1}, q(\tau_{j-1})) &\leq \mathcal{E}_\infty(\tau_{j-1}, q(\tau_j)) + \mathcal{D}_\infty(q(\tau_{j-1}), q(\tau_j)) \\
&= \mathcal{E}_\infty(\tau_j, q(\tau_j)) - \int_{\tau_{j-1}}^{\tau_j} \partial_t\mathcal{E}_\infty(\xi, q(\tau_j))\,\mathrm{d}\xi + \mathcal{D}_\infty(q(\tau_{j-1}), q(\tau_j))
\end{aligned}$$

for $j = 1, \dots, N$. From the definition of the dissipation (and using the above inequality), we can deduce

$$\begin{aligned}
\mathcal{E}_\infty(t, q(t)) + \mathrm{Diss}_\infty(q; [s,t]) &\geq \mathcal{E}_\infty(t, q(t)) + \sum_{j=1}^N \mathcal{D}_\infty(q(\tau_{j-1}), q(\tau_j)) \\
&\geq \mathcal{E}_\infty(s, q(s)) + \sum_{j=1}^N \int_{\tau_{j-1}}^{\tau_j} \partial_t\mathcal{E}_\infty(\xi, q(\tau_j))\,\mathrm{d}\xi \\
&\geq \mathcal{E}_\infty(s, q(s)) + \sum_{j=1}^N \partial_t\mathcal{E}_\infty(\tau_j, q(\tau_j))\Delta\tau_j \\
&\quad + \sum_{j=1}^N \int_{\tau_{j-1}}^{\tau_j} \partial_t\mathcal{E}_\infty(\xi, q(\tau_j)) - \partial_t\mathcal{E}_\infty(\tau_j, q(\tau_j))\,\mathrm{d}\xi
\end{aligned}$$

Now, by Lemma 2.15 we can choose a sequence of partitions of the interval $[s,t]$ such that the Riemann sums $\sum_{j=1}^N \partial_t\mathcal{E}_\infty(\tau_j, q(\tau_j))\Delta\tau_j$ converge to $\int_s^t \partial_t\mathcal{E}_\infty(\tau, q(\tau); u)\,\mathrm{d}\tau$. The last sum converges to zero by (A2) and hence the lower estimate is established. $\square$

The next lemma shows an upper semicontinuity property of the stability sets[8].

Lemma 2.17. *For all* $q_k \in \mathcal{S}_k^{\alpha_k}(t_k)$*, where* $\alpha_k \to 0$*, and all* $t_k \to t$*,* $q_k \to q$ *in* $\mathcal{Q}$ *with* $\sup_{k\in\mathbb{N}} \mathcal{E}_k(t_k, q_k) < \infty$*, it holds that* $q \in \mathcal{S}_\infty(t)$*.*

Proof. First observe that for all $M > 0$ by (A1) there exists a constant $L = L(M) > 0$ such that

$$|\mathcal{E}_k(t, \tilde{q}) - \mathcal{E}_k(s, \tilde{q})| \leq L\,|t - s|\,, \tag{2.11}$$

for all $k \in \mathbb{N}_\infty$ and $\tilde{q} \in \mathcal{Q}$ with $\sup_{k\in\mathbb{N}_\infty, t\in[0,T]} \mathcal{E}_k(t, \tilde{q}) \leq M$. This is nothing else than the uniform Lipschitz-continuity of $\mathcal{E}_k(t, q)$.

For better notation define

$$\mathcal{H}_k(s, q', q'') := \mathcal{E}_k(s, q'') + \mathcal{D}_k(q', q'') - \mathcal{E}_k(s, q'),$$

where $q', q'' \in \mathcal{Q}$.

To show $\mathcal{H}_\infty(t, q, \hat{q}) \geq 0$ for all $\hat{q} \in \mathcal{Q}$ (which is equivalent to $q \in \mathcal{S}_\infty(t)$), assume that $\mathcal{E}_\infty(\bullet, \hat{q}) < \infty$ (otherwise, $\mathcal{H}_\infty(t, q, \hat{q}) \geq 0$ is trivial) and let $(\hat{q}_k)_k$ be a recovery sequence for $\hat{q}$ with respect to the Γ-converging sequence $(\mathcal{E}_k(t, \bullet))_k$, cf. ($\Gamma$1). In particular, we have with (2.11)

$$\limsup_{k\to\infty} \mathcal{E}_k(t_k, \hat{q}_k) = \limsup_{k\to\infty} \mathcal{E}_k(t, \hat{q}_k) = \mathcal{E}_\infty(t, \hat{q}).$$

From the lim inf-inequality and again (2.11), we further get

$$\limsup_{k\to\infty} -\mathcal{E}_k(t_k, q_k) = -\liminf_{k\to\infty} \mathcal{E}_k(t, q_k) \leq -\mathcal{E}_\infty(t, q),$$

Combining the last two estimates and using the continuous convergence $\mathcal{D}_k(q_k, \hat{q}_k) \to \mathcal{D}_\infty(q, \hat{q})$ of the dissipation distances, cf. (Γ2), we arrive at

$$\mathcal{H}_\infty(t, q, \hat{q}) \geq \limsup_{k\to\infty} \mathcal{H}_k(t_k, q_k, \hat{q}_k) \geq -\alpha_k \to 0,$$

where the last estimate holds, because $q_k \in \mathcal{S}_k^{\alpha_k}(t_k)$. Hence, $\mathcal{H}_\infty(t, q, \hat{q}) \geq 0$ and the proof is finshed. □

It should be noted that the preceding lemma works just the same if we only consider subsequences of the functionals.

We will also need a little weak-to-strong convergence lemma. The version here is from [FM06, Lemma 3.5].

[8] If a more general existence result is proved where where weaker convergence requirements than (Γ1) and (Γ2) are assumed, the assertion of Lemma 2.17 needs to be assumed additionally to the other assumptions, cf. [MRS08].

Lemma 2.18. *Let $u_k \xrightarrow{*} u$ in $\mathrm{L}^\infty(0,T)$ and assume that also $u = \limsup_{k\to\infty} u_k$ almost everywhere. Then, $u_k \to u$ strongly in $\mathrm{L}^1(0,T)$.*

Proof. Without loss of generality, we may assume $u \equiv 0$. Let $u_k^+ := \max(u_k, 0)$ and define $g_k(t) := \sup\{\, u_j^+(t) \;:\; j \geq k \,\}$. Then, $0 \leq u_k^+ \leq g_k \downarrow \limsup_{j\to\infty} u_j = 0$ and hence, by the monotone convergence theorem (notice the uniform boundedness of $(g_k)_k$ in $\mathrm{L}^\infty(0,T)$),

$$\lim_{k\to\infty} \left\| u_k^+ \right\|_{\mathrm{L}^1(0,T)} \leq \lim_{k\to\infty} \int_0^T g_k(\tau)\,\mathrm{d}\tau = 0.$$

Since $|x| = 2x^+ - x$, where $x^+ := \max(x,0)$,

$$\|u_k\|_{\mathrm{L}^1(0,T)} = 2\left\|u_k^+\right\|_{\mathrm{L}^1(0,T)} - \int_0^T u_k(\tau)\,\mathrm{d}\tau \to 0$$

by the weak*-convergence and the previous estimate. □

2.4.4 Existence theorem

Recall from page 32 the approximate incremental problem for $k \in \mathbb{N}$, $\varepsilon > 0$ and a partition $\Pi = (0 = t_0, \ldots, t_{N(k)} = T)$:

$$\begin{cases} \text{For } j = 1, \ldots, N \text{ inductively find } q_j^k \in \mathcal{Q} \text{ such that} \\ q_j^k \in \operatorname{Argmin}_{\varepsilon\Delta t_j} \left\{\, \mathcal{E}_k(t_j, \hat{q}) + \mathcal{D}_k(q_{j-1}^k, \hat{q}) \;:\; \hat{q} \in \mathcal{Q} \,\right\}. \end{cases} \tag{AIP$_{k,\varepsilon,\Pi}$}$$

Here, q_0^k is some suitable approximation to the initial value q_0, cf. (CC) below.

We now have all the necessary ingredients to prove existence of solutions.

Theorem 2.19. *Let the assumptions from Section 2.4.1 be fulfilled. Assume further that we are given a sequence of partitions $\Pi_k = (0 = t_0^k, \ldots, t_{N(k)}^k = T)$ of the interval $[0,T]$ with $\|\Pi_k\| \to 0$, a falling sequence $(\varepsilon_k)_k$ of positive real numbers with $\varepsilon_k \to 0$, and a sequence of initial values $(q_0^k)_k \subseteq \mathcal{Q}$ satisfying the compatibility condition*

$$q_0^k \in \mathcal{S}_k^{\varepsilon_k\|\Pi_k\|}(0), \qquad q_0^k \to q_0 \text{ in } \mathcal{Q}, \qquad \text{and} \qquad \mathcal{E}_k(0, q_0^k) \to \mathcal{E}_\infty(0, q_0). \tag{CC}$$

Then, the piecewise-constant, right-continuous interpolants $\bar{q}^k$ of discrete solutions $q^k = (q_j^k)_j$ to (AIP$_{k,\varepsilon_k,\Pi_k}$) admit a subsequence q^{k_n} and a solution $q = (\varphi, z) : [0,T] \to \mathcal{F}\times\mathcal{Z} = \mathcal{Q}$ of (S$_\infty$) & (E$_\infty$) to the initial value q_0 with

- *(i)* $\mathcal{E}_{k_n}(t, \bar{q}^{k_n}(t)) \to \mathcal{E}_\infty(t, q(t))$ *for all* $t \in [0,T]$,
- *(ii)* $\mathrm{Diss}_{k_n}(\bar{q}^{k_n}; [0,t]) \to \mathrm{Diss}_\infty(q; [0,t]) < \infty$ *for all* $t \in [0,T]$,
- *(iii)* $z^{k_n}(t) \to z(t)$ *in* $\mathcal{Z}$ *for all* $t \in [0,T]$,
- *(iv)* $\partial_t\mathcal{E}_{k_n}(\bullet, \bar{q}^{k_n}(\bullet)) \to \partial_t\mathcal{E}_\infty(\bullet, q(\bullet))$ *in* $\mathrm{L}^1(0,T)$,
- *(v) For all $t \in [0,T]$, there exists a t-dependent subsequence $(K_n^t)_n$ of k_n such that $\bar{\varphi}^{K_n^t}(t) \to \varphi(t)$ in $\mathcal{F}$.*

Further, any limit of a subsequence obtained in this way solves (S_∞) *&* (E_∞). *If additionally the topology on* $\mathcal{Q}$ *is separable and metrizable on compact sets,* φ *(hence* q*) can be chosen measurable.*

Proof. Step 1: A-priori estimates
Denote by $q^k = (q_0^k, q_1^k, \ldots, q_{N(k)}^k)$ the discrete solution of $(\mathrm{AIP}_{k,\varepsilon_k,\Pi_k})$, which clearly exists. Define

$$e_j^k := \mathcal{E}_k(t_j^k, q_j^k) \qquad \text{and} \qquad \delta_j^k := \mathcal{D}_k(q_{j-1}^k, q_j^k).$$

Since q^k solves the minimum problem in $(\mathrm{AIP}_{k,\varepsilon_k,\Pi_k})$ at time t_j^k up to an error of $\varepsilon_k \Delta t_j^k$, we have for all $\hat{q} \in \mathcal{Q}$ and $j = 1, \ldots, N(k)$

$$e_j^k \le \varepsilon_k \Delta t_j^k + \mathcal{E}_k(t_j^k, \hat{q}) + \mathcal{D}_k(q_{j-1}^k, \hat{q}) - \delta_j^k \le \varepsilon_k \Delta t_j^k + \mathcal{E}_k(t_j^k, \hat{q}) + \mathcal{D}_k(q_j^k, \hat{q}),$$

whence $q_j^k \in \mathcal{S}_k^{\varepsilon_k \Delta t_j^k}(t_j^k)$.

Similarly, testing the $\varepsilon_k \Delta t_j^k$-minimality of q_j^k with q_{j-1}^k and using (2.8) gives, for $j = 1, \ldots, N(k)$,

$$\begin{aligned} e_j^k + \delta_j^k &\le \varepsilon_k \Delta t_j^k + \mathcal{E}_k(t_j^k, q_{j-1}^k) = \varepsilon_k \Delta t_j^k + e_{j-1}^k + \int_{t_{j-1}}^{t_j} \partial_t \mathcal{E}_k(\tau, q_{j-1}^k)\,\mathrm{d}\tau \\ &\le \varepsilon_k \Delta t_j^k + e_{j-1}^k + \int_{t_{j-1}}^{t_j} c_1^E (e_{j-1}^k + c_0^E) \mathrm{e}^{c_1^E(\tau - t_{j-1}^k)}\,\mathrm{d}\tau \\ &= \varepsilon_k \Delta t_j^k + e_{j-1}^k + (e_{j-1}^k + c_0^E)(\mathrm{e}^{c_1^E \Delta t_j^k} - 1). \end{aligned} \tag{2.12}$$

Define $E_j^k := e_j^k + c_0^E + \varepsilon_k / c_1^E$ for $j = 0, \ldots, N(k)$. Then, using the last inequality, we can estimate for $j = 1, \ldots, N(k)$

$$\begin{aligned} E_j^k + \delta_j^k &= e_j^k + c_0^E + \frac{\varepsilon_k}{c_1^E} + \delta_j^k \le \varepsilon_k \Delta t_j^k + (e_{j-1}^k + c_0^E)\mathrm{e}^{c_1^E \Delta t_j^k} + \frac{\varepsilon_k}{c_1^E} \\ &= E_{j-1}^k \mathrm{e}^{c_1^E \Delta t_j^k} + \varepsilon_k \left(\Delta t_j^k + \frac{1 - \mathrm{e}^{c_1^E \Delta t_j^k}}{c_1^E} \right) \\ &= E_{j-1}^k \mathrm{e}^{c_1^E \Delta t_j^k} + \varepsilon_k \int_0^{\Delta t_j^k} 1 - \mathrm{e}^{c_1^E \tau}\,\mathrm{d}\tau \le E_{j-1}^k \mathrm{e}^{c_1^E \Delta t_j^k}, \end{aligned} \tag{2.13}$$

since the last integrand is always less than zero. Let

$$E^* := \sup_{k \in \mathbb{N}} \left(\mathcal{E}_k(0, q_0^k) + c_0^E + \frac{\varepsilon_k}{c_1^E} \right) < \infty,$$

(cf. (CC)). Neglecting $\delta_j^k \ge 0$, by induction it follows from (2.13) that $E_j^k \le \mathrm{e}^{c_1^E t_j^k} E^*$, i.e. the a-priori energy bound

$$e_j^k + c_0^E \le E^* \mathrm{e}^{c_1^E t_j^k} \qquad \text{for } j = 0, \ldots, N(k)$$

holds. Let $\bar{q}^k$ be the right-continuous piecewise-constant interpolant of q^k. By the preceding estimate and (2.7), we get

$$\bar{e}^k(t) := \mathcal{E}_k(t, \bar{q}^k(t)) \leq E^* \mathrm{e}^{c_1^E t} \leq E^* \mathrm{e}^{c_1^E T} \qquad \text{for all } t \in [0,T], k \in \mathbb{N}. \tag{2.14}$$

By (A1), this immediately implies

$$\bar{P}^k(t) := \partial_t \mathcal{E}_k(t, \bar{q}^k(t)) \leq c_1^E (c_0^E + E^* \mathrm{e}^{c_1^E T}) \qquad \text{for all } t \in [0,T], k \in \mathbb{N}. \tag{2.15}$$

Rearranging (2.13), summing up and using $E_{j-1}^k \leq \mathrm{e}^{c_1^E t_{j-1}^k} E^*$ gives

$$\begin{aligned}\sum_{j=1}^{N(k)} \delta_j^k &\leq \sum_{j=1}^{N(k)} \left(E_{j-1}^k \mathrm{e}^{c_1^E \Delta t_j^k} - E_j^k \right) = E_0^k - E_{N(k)}^k + \sum_{j=1}^{N(k)} \left(\mathrm{e}^{c_1^E \Delta t_j^k} - 1 \right) E_{j-1}^k \\ &\leq E_0^k - E_{N(k)}^k + E^* \sum_{j=1}^{N(k)} \left(\mathrm{e}^{c_1^E t_j^k} - \mathrm{e}^{c_1^E t_{j-1}^k} \right) \\ &= E_0^k - E_{N(k)}^k + E^* \mathrm{e}^{c_1^E T} - E^* \leq E^* \mathrm{e}^{c_1^E T},\end{aligned} \tag{2.16}$$

since $E_{N(k)}^k \geq 0$, which is contained implicitly in (A1). It follows that the dissipation of $\bar{q}^k$ stays uniformly bounded, i.e.

$$\mathrm{Diss}_k(\bar{q}^k; [0,T]) = \sum_{j=1}^{N(k)} \delta_j^k \leq E^* \mathrm{e}^{c_1^E T}. \tag{2.17}$$

To apply the classical Helly selection principle in the next step, we also need a uniform bound on the pointwise variation of $\bar{e}_k = \mathcal{E}_k(t, \bar{q}_k(t))$. For $j = 2, \ldots, N(k)$ above it was shown that $q_{j-1}^k \in \mathcal{S}_k^{\varepsilon_k \Delta t_{j-1}^k}(t_{j-1}^k)$ and by testing the stability with q_j^k, and using (2.8), (2.15), one can derive that

$$\begin{aligned} e_{j-1}^k &\leq \varepsilon_k \Delta t_{j-1}^k + \mathcal{E}(t_{j-1}^k, q_j^k) + \mathcal{D}_k(q_{j-1}^k, q_j^k) \\ &\leq \varepsilon_k \Delta t_{j-1}^k + e_j^k + \delta_j^k + \int_{t_{j-1}^k}^{t_j^k} \left| \partial_t \mathcal{E}_k(\tau, q_j^k) \right| \, \mathrm{d}\tau \\ &\leq \varepsilon_k \Delta t_{j-1}^k + c \Delta t_j^k + e_j^k + \delta_j^k, \end{aligned}$$

where here and in the following, c is a *generic* constant, which does not depend on k or j. So, for $j = 2, \ldots, N(k)$, we have shown one half of the following estimate, the other half being contained in (2.12), (2.15):

$$\left| e_j^k + \delta_j^k - e_{j-1}^k \right| \leq c(\Delta t_{j-1}^k + \Delta t_j^k) \tag{2.18}$$

Since $\dot{\bar{e}}_k(t) = \partial_t \mathcal{E}_k(t, q_{j-1}^k)$ for $t \in (t_{j-1}^k, t_j^k)$, estimate (2.15) implies

$$\int_{t_{j-1}^k}^{t_j^k} |\dot{\bar{e}}_k(\tau)| \, \mathrm{d}\tau \leq c\Delta t_j^k. \tag{2.19}$$

Consequently, also using (2.18), the jumps $\Delta e_j^k := \bar{e}^k(t_j^k) - \bar{e}^k(t_j^k - 0)$, $j = 2, \ldots, N(k)$, of $\bar{e}^k$ satisfy

$$\begin{aligned} \left|\Delta e_j^k\right| = \lim_{h \downarrow 0} \left|\bar{e}^k(t_j^k) - \bar{e}^k(t_j^k - h)\right| &\leq \left|e_j^k - e_{j-1}^k\right| + \int_{t_{j-1}^k}^{t_j^k} |\dot{\bar{e}}_k(\tau)| \, \mathrm{d}\tau \\ &\leq c\Delta t_j^k + \varepsilon_k \Delta t_{j-1}^k + \delta_j^k. \end{aligned}$$

By (2.14), the first jump at least satisfies

$$\left|\Delta e_1^k\right| \leq 2E^* \mathrm{e}^{c_1^E T}.$$

The last two estimates on the energy jumps together with (2.16), (2.19) yield the following uniform bound on the pointwise variation:

$$\begin{aligned} \mathrm{Var}(\bar{e}^k; [0,T]) = \sum_{j=1}^{N(k)} \left(\int_{t_{j-1}^k}^{t_j^k} |\dot{\bar{e}}_k(\tau)| \, \mathrm{d}\tau + \left|\Delta e_j^k\right| \right) &\leq cT + \left|\Delta e_1^k\right| + \sum_{j=2}^{N(k)} \left|\Delta e_j^k\right| \\ \leq cT + 2E^* \mathrm{e}^{c_1^E T} + \sum_{j=2}^{N(k)} \delta_j^k &\leq cT + 3E^* \mathrm{e}^{c_1^E T}. \end{aligned} \tag{2.20}$$

Step 2: Selection of subsequences
The t- and k-independent bounds (2.14), (2.17) and (2.20) together with the assumptions (A3), (A5), (Γ2), and (2.10) allow us to invoke the generalized Helly selection principle, Theorem 2.13, and also the classical Helly selection principle, Corollary 2.14, in order to get subsequences (not relabelled) and limit functions $z \in \mathcal{Z}[0,T]$, $e^\infty \in \mathrm{L}^\infty(0,T)$ and $\delta^\infty \in \mathrm{L}^\infty(0,T)$ (increasing and with $\delta^\infty(0) = 0$) such that for all $s, t \in [0,T]$ with $s \leq t$, it holds that

$$\begin{aligned} &\mathrm{Diss}_k(\bar{q}^k; [0,t]) \to \delta^\infty(t), && \mathrm{Diss}_\infty(z; [s,t]) \leq \delta^\infty(t) - \delta^\infty(s), \\ &\bar{e}^k(t) \to e^\infty(t), && \bar{z}^k(t) \to z(t). \end{aligned} \tag{2.21}$$

Now, (2.15) implies that there exists $P^\infty \in \mathrm{L}^\infty(0,T)$ with $\bar{P}^k \overset{*}{\rightharpoonup} P^\infty$ in $\mathrm{L}^\infty(0,T)$. Also, define $P^* := \limsup_{k\to\infty} \bar{P}^k \in \mathrm{L}^\infty(0,T)$ (pointwise). Let $E \subseteq [0,T]$ be an arbitrary Lebesgue-measurable set. Then, Fatou's lemma (for $-\bar{P}^k$, noticing the uniform boundedness) implies

$$\int_E P^*(x) \, \mathrm{d}x = \int_E \limsup_{k\to\infty} \bar{P}^k(x) \, \mathrm{d}x \geq \limsup_{k\to\infty} \int_E \bar{P}^k(x) \, \mathrm{d}x = \int_E P^\infty(x) \, \mathrm{d}x,$$

hence $P^\infty \le P^*$ almost everywhere.

To construct the limit for the elastic part $\varphi \in \mathcal{F}[0,T]$, we need to select t-dependent subsequences. For this, define for $t \in [0,T]$

$$A(t) := \big\{\, \tilde\varphi \in \mathcal{F} \;:\; \partial_t \mathcal{E}_\infty(t, (\tilde\varphi, z(t))) = P^*(t) \text{ and } \exists (K_n^t)_{n\in\mathbb{N}} \text{ with } \bar\varphi_{K_n^t}(t) \to \tilde\varphi \text{ in } \mathcal{F} \,\big\}.$$

We will see in a moment that $A(t) \neq \emptyset$ for all $t \in [0,T]$. Assuming this, we may choose $\varphi(t) \in A(t)$ arbitrary for all $t \in [0,T]$ (this, of course, uses the Axiom of Choice) and set $q(t) := (\varphi(t), z(t))$ for all $t \in [0,T]$. This q will then be shown to be a solution.

To see that $A(t)$ is nonempty, first choose a subsequence K_n^t such that $\bar P_{K_n^t}(t) \to P^*(t)$. Now, by the uniform energy bound (2.14) and (A3), we can select a further subsequence (not relabelled) such that also $\bar\varphi_{K_n^t}(t) \to \tilde\varphi$ in $\mathcal{F}$ for some $\tilde\varphi \in \mathcal{F}$. Let $m(t,k) \in \{0,\dots,N(k)\}$ be the largest integer such that $t^k_{m(t,k)} \le t$. Then, $\bar q^k(t) = \bar q^k(t^k_{m(t,k)}) = q^k_{m(t,k)} \in \mathcal{S}_k^{\varepsilon_k \Delta t^k_{m(t,k)}}(t^k_{m(t,k)})$, whence assumption (A4) yields

$$\partial_t \mathcal{E}_{K_n^t}(t^k_{m(t,k)}, \bar q^{K_n^t}(t)) \to \partial_t \mathcal{E}_\infty(t, (\tilde\varphi, z(t))) \qquad \text{as } k \to \infty.$$

On the other hand,

$$\lim_{n\to\infty} \partial_t \mathcal{E}_{K_n^t}(t^k_{m(t,k)}, \bar q^{K_n^t}(t)) = \lim_{n\to\infty} \partial_t \mathcal{E}_{K_n^t}(t, \bar q^{K_n^t}(t)) = P^*(t)$$

by the choice of K_n^t. Hence, $\partial_t \mathcal{E}_\infty(t, (\tilde\varphi, z(t))) = P^*(t)$ and $\tilde\varphi \in A(t)$.

Step 3: Stability of the limit process

Fix $t \in [0,T]$ and let again $m(t,k) \in \{0,\dots,N(k)\}$ be the largest integer such that $t^k_{m(t,k)} \le t$. Because $\bar q^k(t) = \bar q^k(t^k_{m(t,k)}) \in \mathcal{S}_k^{\varepsilon_k \|\Pi_k\|}(t^k_{m(t,k)})$ (this was shown in Step 1 for $t > 0$ and assumed in (CC) for $t = 0$), Lemma 2.17 implies $q(t) \in \mathcal{S}_\infty(t)$ for all $t \in [0,T]$.

Step 4: Energy estimates in the limit

By the growth estimate (A1), we see that $\left|\bar e^k(t) - e^k_{j-1}\right| \le c\,\|\Pi_k\|$ for all $t \in [t^k_{j-1}, t_j)$ (recall (2.14)). Let $m(t,k)$ be defined as above. We sum the first line of (2.12) for $j = 1,\dots,m(t,k)$ and get, also employing the uniform bound (2.15) on $\bar P^k = \partial_t \mathcal{E}_k(\bullet, \bar q^k(\bullet))$,

$$\begin{aligned}
\bar e^k(t) + \mathrm{Diss}_k(\bar q^k; [0,t]) &\le \bar e^k_{m(t,k)} + \sum_{j=1}^{m(t,k)} \delta_j^k + c\,\|\Pi_k\| \\
&\le \varepsilon_k t + \bar e^k(0) + \int_0^t \bar P^k(\tau)\,\mathrm{d}\tau + c\,\|\Pi_k\|
\end{aligned}$$

for all $t \in [0,T]$. Passing to the limit $k \to \infty$ in this inequality and using (2.21) as well as $\bar e^k(0) = \mathcal{E}_k(0, q_0^k) \to \mathcal{E}_\infty(0, q_0)$ by (CC), $\bar P^k \overset{*}{\rightharpoonup} P^\infty$ in $\mathrm{L}^\infty(0,T)$ and $P^\infty \le P^*$, we see

$$\begin{aligned}
\mathcal{E}_\infty(t, q(t)) + \mathrm{Diss}_\infty(q; [0,t]) &\le e^\infty(t) + \delta^\infty(t) \le \mathcal{E}_\infty(0, q_0) + \int_0^t P^\infty(\tau)\,\mathrm{d}\tau \\
&\le \mathcal{E}_\infty(0, q_0) + \int_0^t P^*(\tau)\,\mathrm{d}\tau
\end{aligned}$$

for all $t \in [0,T]$.

From Proposition 2.16 ($\partial_t \mathcal{E}_\infty(\bullet, q(\bullet)) \in \mathrm{L}^1(0,T)$ by (2.14) and (A1)) we get the lower energy estimate

$$\begin{aligned}\mathcal{E}_\infty(t,q(t)) + \mathrm{Diss}_\infty(q;[0,t]) &\geq \mathcal{E}_\infty(0,q_0) + \int_0^t \partial_t \mathcal{E}_\infty(\tau,q(\tau))\,\mathrm{d}\tau \\ &= \mathcal{E}_\infty(0,q_0) + \int_0^t P^*(\tau)\,\mathrm{d}\tau\end{aligned}$$

where the last equality follows from the definition of $A(\bullet)$. Combining the last two estimates, we see

$$\begin{aligned}\mathcal{E}_\infty(t,q(t)) + \mathrm{Diss}_\infty(q;[0,t]) &= e^\infty(t) + \delta^\infty(t) = \mathcal{E}_\infty(0,q_0) + \int_0^t P^\infty(\tau)\,\mathrm{d}\tau \\ &= \mathcal{E}_\infty(0,q_0) + \int_0^t P^*(\tau)\,\mathrm{d}\tau = \mathcal{E}_\infty(0,q_0) + \int_0^t \partial_t\mathcal{E}_\infty(\tau,q(\tau))\,\mathrm{d}\tau\end{aligned} \tag{2.22}$$

for all $t \in [0,T]$. In particular, (E_∞) holds.

Step 5: Improved convergence

From (2.22), by $\mathcal{E}_\infty(t,q(t)) \leq e^\infty(t)$ and $\mathrm{Diss}_\infty(q;[0,t]) \leq \delta^\infty(t)$, we immediately have

$$\mathcal{E}_\infty(t,q(t)) = e^\infty(t) \qquad \text{and} \qquad \mathrm{Diss}_\infty(q;[0,t]) = \delta^\infty(t)$$

for all $t \in [0,T]$ and $P^\infty = P^* = \partial_t \mathcal{E}_\infty(\bullet, q(\bullet))$ almost everywhere. Hence, (i) and (ii) hold and an application of Lemma 2.18 yields (iv). The points (iii) and (v) have already been shown to hold in the course of the proof.

Step 6: Measurability

We now assume that $\mathcal{Q}$ is metrizable and separable on compact sets. Then, by (2.14) and (A3), $t \mapsto A(t)$ from Step 2 takes as values subsets of a separable and metrizable space. By (iv) and the definition of $A(\bullet)$ (here, subsequences are not relabelled),

$$\bar{P}^k = \partial_t \mathcal{E}_k(\bullet, \bar{q}^k(\bullet)) \to \partial_t \mathcal{E}_\infty(\bullet, q(\bullet)) = P^* \text{ in } \mathrm{L}^1(0,T),$$

and the right hand side does not depend on the choice of $\tilde{\varphi}$ from $A(t)$ in Step 2. Choosing a further subsequence (again not relabelled), we have $\bar{P}^k \to P^*$ almost everywhere. Hence,

$$\begin{aligned}B(t) &:= \left\{\, \tilde{\varphi} \in \mathcal{F} \;:\; \exists K_n^t \uparrow \infty \text{ (as } n \to \infty) \text{ and } \bar{\varphi}^{K_n^t}(t) \to \tilde{\varphi} \text{ in } \mathcal{F} \,\right\} \\ &= \text{K-}\limsup\{\bar{\varphi}^{k_n}(t)\} \subseteq A(t)\end{aligned}$$

and B is a closed-(set-)valued map with nonempty values. Because it is the Kuratowski upper limit of singletons, Theorem 8.2.5 from [AF90] shows that B is measurable (in the sense of set-valued maps, i.e. $B^{-1}(U) = \{\, t \in [0,T] \;:\; B(t) \cap U \neq \emptyset \,\}$ is measurable for all open sets $U \subseteq \mathcal{F}$). Hence, by Theorem 8.1.3 from [AF90], the map B admits a measurable selection $\varphi \in \mathcal{F}[0,T]$ with $\varphi(t) \in B(t)$ for all $t \in [0,T]$. With this selection, Steps 3 to 5 work just the same. □

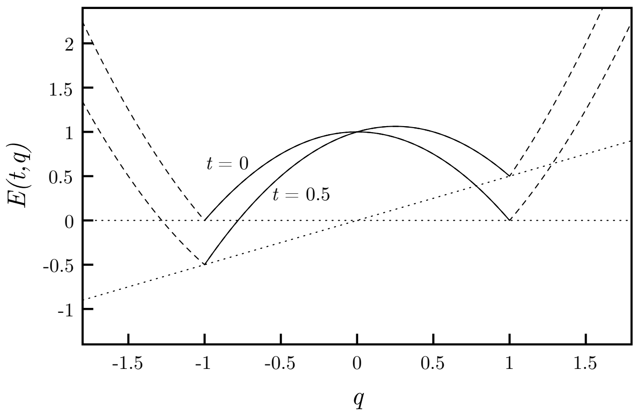

Figure 3: The double-well potential from Counterexample 2.21.

Remark 2.20. If the compatibility condition (CC) is not valid (and this might be difficult to check in applications!), the proof of the preceding existence theorem still works, but stability of the obtained (generalized) solution holds only for $t > 0$. The obtained process immediately jumps from the unstable (incompatible) initial value to a stable value. For many applications, this is enough.

To conclude this section, it should also be mentioned that convergence of solutions of (S_k) & (E_k) to a solution of (S_∞) & (E_∞) can be established, see [MRS08, Theorem 3.1].

2.5 Uniqueness

In this section we consider only single problems. In general, uniqueness cannot be expected from solutions to rate-independent systems since they are highly nonlinear, cf. [BKS04]. In fact, the following counterexample presents a system where solutions to a rate-independent system are not unique.

Counterexample 2.21. On the space $\mathcal{Q} = [-1, 1]$ and in the time interval $[0, T] = [0, 1]$, we consider the potential (see Figure 3)

$$\mathcal{E}(t, q) := (1 + q)(1 - q) + \frac{t}{2}q$$

together with the dissipation distance $\mathcal{D}(u, v) := |v - u|$. It is easily seen that this potential fulfills all the requirements of the theory on the *compact* interval $[-1, 1]$. The functional $\mathcal{E}(t, \bullet)$ has precisely two (strict) local minima, one at -1 and the other at $+1$. There, $\mathcal{E}$ takes the values $\mathcal{E}(t, -1) = -t/2$ and $\mathcal{E}(t, +1) = t/2$, respectively.

At $t = 0$, the set of stable states $\mathcal{S}(0)$ contains 0, because for all $\hat{q} \in \mathcal{Q}$

$$\mathcal{E}(0, \hat{q}) + \mathcal{D}(0, \hat{q}) = 1 - \hat{q}^2 + |\hat{q}| \geq 1 = \mathcal{E}(0, 0),$$

since $\hat{q}^2 \leq |\hat{q}|$ in $[-1,1]$. The following two processes are both solutions to the energetic formulation (S) & (E) with initial value $q_0 = 0 \in \mathcal{S}(0)$:

$$q_-(t) := \begin{cases} 0 & \text{if } t = 0, \\ -1 & \text{if } t \in (0,1] \end{cases} \qquad \text{and} \qquad q_+(t) := \begin{cases} 0 & \text{if } t = 0, \\ +1 & \text{if } t \in (0,1]. \end{cases}$$

We have $-1, +1 \in \mathcal{S}(t)$ for all $t \in [0,1]$. For -1 this is clear since -1 is the global minimum of $\mathcal{E}(t,\bullet)$ for all $t \geq 0$. For $+1$, the hyperplane $q \mapsto tq/2$ supports the graph of $\mathcal{E}(t,\bullet)$. This hyperplane has at most slope $1/2$; therefore, if we add the linear map $\hat{q} \mapsto \mathcal{D}(+1,\hat{q}) = 1 - \hat{q}$ (since $\hat{q} \in [-1,1]$) to $\mathcal{E}(t,\bullet)$, we still have a hyperplane with negative slope $-1/2$ supporting $\mathcal{E}(t,\bullet) + \mathcal{D}(+1,\bullet)$ and going through $(+1, \mathcal{E}(t,+1))$. Hence also $+1 \in \mathcal{S}(t)$. We have thus established the validity of (S) for q_- and q_+. Further, for $t \in (0,1]$, the energy balance (E) holds as well:

$$\begin{aligned} \mathcal{E}(t, q_\pm(t)) + \mathrm{Diss}_{\mathcal{D}}(q_\pm; [0,t]) = \mathcal{E}(t, \pm 1) + \mathcal{D}(0, \pm 1) &= \frac{\pm t}{2} + 1 \\ &= \mathcal{E}(0, q_0) + \int_0^t \partial_t \mathcal{E}(\tau, q_\pm(\tau)) \,\mathrm{d}\tau \end{aligned}$$

since $\partial_t \mathcal{E}(\bullet, q_\pm(\bullet)) = \pm 1/2$ almost everywhere. For $t = 0$, the energy balance is trivial.
Hence, both q_- and q_+ solve (S) & (E) and there is no uniqueness.

Note that $\mathcal{E}(t,\bullet)$ from the preceding counterexample is the restriction of the double-well potential (see Figure 3)

$$(t,q) \mapsto |1-q|\,|1+q| + \frac{t}{2}q$$

to the interval $[-1,1]$ (we refrained from carrying out the example on a bigger space for ease of notation only). Hence, the last counterexample physically represents a phase transition problem, where the energies of the two phases change in the course of time due to a prescribed loading ($q \mapsto tq/2$). We will return to this example in the next chapter.

The nonuniqueness of solutions is due to nonconvexity (or degenerate convexity) of the stored energy functional (as seen in the previous example). Therefore, if we strengthen the convexity assumptions, also uniqueness is to be expected, and indeed this is the case if additionally the sets of stable states are convex and the functionals are of a special form (cf. [Mie05, Theorem 4.2]):

Theorem 2.22. *Let $\mathcal{Q}$ be a Banach space and assume that $\mathcal{E}$ takes the form $\mathcal{E}(t,q) = \mathcal{W}(q) + \langle l(t), q\rangle$, where $l(t) \in \mathcal{Q}^*$ for all $t \in [0,T]$ and $\mathcal{W} : \mathcal{Q} \to \mathbb{R}_\infty$ is strictly convex. Further, assume that $\mathcal{D}(q_1,q_2) = \mathcal{R}(q_2 - q_1)$ for a convex, 1-homogeneous dissipation potential $\mathcal{R} : \mathcal{Q} \to [0,\infty]$ and let $\mathcal{S}(t)$ be convex for all $t \in [0,T]$. Then, for each initial value $q_0 \in \mathcal{S}(0)$, there is at most one process solving* (S) *&* (E).

Proof. Assume the existence of two processes $q, \tilde{q} : [0,T] \to \mathcal{Q}$ both fulfilling (S) & (E) and satisfying $q(0) = \tilde{q}(0) = q_0$. Then, by convexity of $\mathcal{S}(t)$, the combined process $y(t) := (q(t) + \tilde{q}(t))/2$ is also stable. Thus, Proposition 2.16 implies the lower energy estimate

$$\mathcal{E}(t, y(t)) + \mathrm{Diss}_{\mathcal{D}}(y; [0,t]) \geq \mathcal{E}(0, q_0) + \int_0^t \partial_t \mathcal{E}(\tau, y(\tau))\, \mathrm{d}\tau \tag{2.23}$$

For all $t \in [0,T]$. If for some $t_0 \in [0,T]$, $q(t_0) \neq \tilde{q}(t_0)$, then

$$\begin{aligned}
&\mathcal{E}(t_0, y(t_0)) + \mathrm{Diss}_{\mathcal{D}}(y; [0,t_0]) \\
&\qquad < \frac{1}{2}\Big[\mathcal{E}(t_0, q(t_0)) + \mathcal{E}(t_0, \tilde{q}(t_0)) + \mathrm{Diss}_{\mathcal{D}}(q; [0,t_0]) + \mathrm{Diss}_{\mathcal{D}}(\tilde{q}; [0,t_0])\Big]
\end{aligned}$$

by the strict convexity of $\mathcal{E}(t_0, \bullet)$ and the convexity of $\mathrm{Diss}_{\mathcal{D}}(\bullet; [0,t_0])$, which follows from the convexity and 1-homogeneity of $\mathcal{R}$. As both q and $\tilde{q}$ fulfill (E), the right hand side of the preceding inequality is equal to

$$\begin{aligned}
&\mathcal{E}(0, q_0) + \frac{1}{2}\int_0^{t_0} \big(\partial_t \mathcal{E}(\tau, q(\tau)) + \partial_t \mathcal{E}(\tau, \tilde{q}(\tau))\big)\, \mathrm{d}\tau \\
&\qquad = \mathcal{E}(0, q_0) + \int_0^{t_0} \Big\langle \dot{l}(\tau), \frac{q(\tau) + \tilde{q}(\tau)}{2} \Big\rangle\, \mathrm{d}\tau = \mathcal{E}(0, q_0) + \int_0^{t_0} \partial_t \mathcal{E}(\tau, y(\tau))\, \mathrm{d}\tau,
\end{aligned}$$

where we employed that $\partial_t \mathcal{E}(t, q) = \langle \dot{l}(t), q \rangle$ is linear in q. Therefore, the strict upper energy estimate

$$\mathcal{E}(t_0, y(t_0)) + \mathrm{Diss}_{\mathcal{D}}(y; [0,t_0]) < \mathcal{E}(0, q_0) + \int_0^{t_0} \partial_t \mathcal{E}(\tau, y(\tau))\, \mathrm{d}\tau$$

holds, contradicting (2.23). □

In the case of *uniform* convexity, uniqueness of a solution can directly be established without any assumptions on the stability sets, cf. [Mie05, Proposition 4.1]. In this case, solutions are even Lipschitz-continuous and depend Lipschitz-continuously in the initial value, see [Mie05, Theorem 4.3]. Recent development on uniqueness can be found in [MR07b].

3 Reverse approximation

The leading question of this section is whether every solution of (S) & (E) is representable as limit of discrete solutions to the *strict* incremental problem (IP). This, however, is not true in general as can be seen by the following counterexample:

Counterexample 3.1. Consider the situation of Counterexample 2.21. There it was shown that on the space $\mathcal{Q} = [-1, 1]$ the processes q_- and q_+ with

$$q_-(t) = \begin{cases} 0 & \text{if } t = 0, \\ -1 & \text{if } t \in (0, 1] \end{cases} \qquad \text{and} \qquad q_+(t) = \begin{cases} 0 & \text{if } t = 0, \\ +1 & \text{if } t \in (0, 1]. \end{cases}$$

both solve (S) & (E) for the functionals

$$\mathcal{E}(t, q) := (1 + q)(1 - q) + \frac{t}{2}q \qquad \text{and} \qquad \mathcal{D}(u, v) := |v - u|\,.$$

The strict incremental problem (IP), however, will always select q_-: In the first step, at time $t_1 > 0$, we seek the global minimizer of $\mathcal{E}(t_1, \bullet) + \mathcal{D}(0, \bullet)$. But this global minimizer clearly is -1 (see Figure 3). So, the discrete solution will jump to $q_1 = -1$ and, because $-1 \in \mathcal{S}(t)$ for all t, stay there forever. Passing to the limit, we get the solution process q_- from above. The other solution q_+, however, is not selected. This shows that not all solutions of (S) & (E) correspond to discrete solutions if we only allow strict minimizers in the incremental problem.

The preceding result reflects the interesting observation that while there might be a "preferred" solution (q_- in the counterexample), another solution can also occur if we allow for small (in fact, arbitrarily small) perturbations. This is nothing else but the instability of the evolution problem and seems to be well in line with physical intuition.

For many purposes, it suffices to show that we can find a solution to an *approximate* incremental problem such as (AIP); to prove this "reverse approximation" is the aim of this chapter. First, we investigate single problems in Section 3.1 and afterwards, building on these results, investigate the corresponding question for sequences of problems in Section 3.2. In a certain sense, we construct recovery sequences for the limit solution, but not for ordinary Γ-limits of functionals, but in the more complex situation of convergent sequences of time-dependent energy and dissipation functionals. Finally, in Section 3.3 we investigate contraction properties of the minimization problems in (AIP) in the case of strictly convex functionals (based on the results from Section 1.4) and we use reverse approximation to derive a stability estimate on the difference of energetic solutions to the same initial value.

3.1 Single problems

Throughout this section, we silently assume (A1') and (A5') from page 33 to hold (they are standard in the theory and are also needed to establish existence, see Sections 2.3 and 2.4).

By the Gronwall lemma, (A1') implies

$$\mathcal{E}(t,q) + c_0^E \le (\mathcal{E}(s,q) + c_0^E)\mathrm{e}^{c_1^E|t-s|} \qquad \text{for all } t,s \in [0,T], q \in \mathcal{Q}.$$

Using (A1') again, we also get

$$|\partial_t \mathcal{E}(t,q)| \le c_1^E(\mathcal{E}(s,q) + c_0^E)\mathrm{e}^{c_1^E|t-s|} \qquad \text{for all } t,s \in [0,T], q \in \mathcal{Q}. \tag{3.1}$$

Similarly, for an energetic solution $q \in \mathcal{Q}[0,T]$ and for all $t \in [0,T]$, we have the a-priori estimates

$$\mathcal{E}(t,q(t)) + c_0^E \le (\mathcal{E}(0,q(0)) + c_0^E)\mathrm{e}^{c_1^E t}, \tag{3.2}$$

$$|\partial_t \mathcal{E}(t,q(t))| \le c_1^E(\mathcal{E}(0,q(0)) + c_0^E)\mathrm{e}^{c_1^E t}. \tag{3.3}$$

To see this, notice that with (E) and (A1'), we have for all $t \in [0,T]$

$$\begin{aligned}\mathcal{E}(t,q(t)) + c_0^E &\le \mathcal{E}(0,q(0)) + c_0^E + \int_0^t \partial_t\mathcal{E}(\tau,q(\tau))\,\mathrm{d}\tau \\ &\le \mathcal{E}(0,q(0)) + c_0^E + \int_0^t c_1^E(\mathcal{E}(\tau,q(\tau)) + c_0^E)\,\mathrm{d}\tau\end{aligned}$$

and hence (3.2) follows by applying Gronwall's inequality (note $\mathcal{E}(t,q(t)) + c_0^E \ge 0$). Invoking again (A1'), we also have (3.3).

We commence with a lemma which allows us to estimate the energy of approximate minimizers.

Lemma 3.2. *Let $q \in \mathcal{Q}[0,T]$ be a solution of* (S) *&* (E) *with initial value $q_0 = q(0) \in \mathcal{S}(0)$ and let $\Pi = (0 = t_0, \dots, t_N = T)$ be a partition of the interval $[0,T]$. Let*

$$q_j := q(t_j) \qquad \textit{for } j = 1,\dots,N-1.$$

Then, for all $\delta \ge 0$ there exists $M = M(q_0,\delta) \in \mathbb{R}$ such that for all $j = 1,\dots,N$ and for all

$$q_j^* \in \mathrm{Argmin}_\delta \left\{ \mathcal{E}(t_j,\hat{q}) + \mathcal{D}(q_{j-1},\hat{q}) \;:\; \hat{q} \in \mathcal{Q} \right\},$$

it holds that $\mathcal{E}(s,q_j^*) \le M$ for all $s \in [0,T]$. Further, the quantity $M(q_0,\delta)$ can be chosen increasing in δ.

Proof. First, we use the δ–minimality of q_j^* to derive

$$\begin{aligned}\mathcal{E}(t_j,q_j^*)+\mathcal{D}(q_{j-1},q_j^*)&\le \mathcal{E}(t_j,q_{j-1})+\mathcal{D}(q_{j-1},q_{j-1})+\delta\\ &\le \mathcal{E}(t_{j-1},q_{j-1})+\int_{t_{j-1}}^{t_j}\partial_t\mathcal{E}(\tau,q_{j-1})\,\mathrm{d}\tau+\delta,\end{aligned}\tag{3.4}$$

where we exploited $\mathcal{D}(q_{j-1},q_{j-1})=0$. We can now use the growth estimate (3.1) to deduce

$$\begin{aligned}\int_{t_{j-1}}^{t_j}\partial_t\mathcal{E}(\tau,q_{j-1})\,\mathrm{d}\tau&\le \int_{t_{j-1}}^{t_j} c_1^E\left(\mathcal{E}(t_{j-1},q_{j-1})+c_0^E\right)\mathrm{e}^{c_1^E(\tau-t_{j-1})}\,\mathrm{d}\tau\\ &=\left(\mathcal{E}(t_{j-1},q_{j-1})+c_0^E\right)\left(\mathrm{e}^{c_1^E\Delta t_j}-1\right).\end{aligned}\tag{3.5}$$

The a-priori bound (3.2) on the energy of the continuous solution provides the necessary information to estimate the term $\mathcal{E}(t_{j-1},q_{j-1})$. Indeed,

$$\mathcal{E}(t_{j-1},q_{j-1})+c_0^E\le\left(\mathcal{E}(0,q_0)+c_0^E\right)\mathrm{e}^{c_1^E t_{j-1}}\le\left(\mathcal{E}(0,q_0)+c_0^E\right)\mathrm{e}^{c_1^E T}=:L$$

where $L=L(q_0)$ only depends on q_0 (and on c_0^E, c_1^E, but this dependency is henceforth omitted for all constants). We combine this with the previous estimates (3.4), (3.5) to get

$$\begin{aligned}\mathcal{E}(t_j,q_j^*)&\le \mathcal{E}(t_j,q_j^*)+\mathcal{D}(q_{j-1},q_j^*)\\ &\le \mathcal{E}(t_{j-1},q_{j-1})+\left(\mathcal{E}(t_{j-1},q_{j-1})+c_0^E\right)\left(\mathrm{e}^{c_1^E\Delta t_j}-1\right)+\delta\\ &\le L+L\left(\mathrm{e}^{c_1^E T}-1\right)+\delta=:L_1=L_1(q_0,\delta).\end{aligned}$$

Using the analogue of (2.7) for single functionals with $s=t_j$, the result follows with $M=(L_1(q_0,\delta)+c_0^E)\mathrm{e}^{c_1^E T}-c_0$. The monotonicity claim is clear. □

We are now in a position to prove that every solution of (S) & (E) gives rise to a solution of (cf. page 32)

$$\begin{cases}\text{For } j=1,\dots,N \text{ inductively find } q_j\in\mathcal{Q} \text{ such that}\\ q_j\in\operatorname{Argmin}_{\varepsilon\Delta t_j}\left\{\,\mathcal{E}(t_j,\hat q)+\mathcal{D}(q_{j-1},\hat q)\ :\ \hat q\in\mathcal{Q}\,\right\}\end{cases}\tag{AIP$_{\varepsilon,\Pi}$}$$

for any partition $\Pi=(0=t_0,\dots,t_N=T)$ and a suitably chosen constant $\varepsilon=c_R$.

Theorem 3.3. *Let $q\in\mathcal{Q}[0,T]$ be a solution of* (S) *&* (E) *with initial value $q_0=q(0)\in\mathcal{S}(0)$. Then, there exits a constant $c_R=c_R(q_0)>0$ such that for any partition $\Pi=(0=t_0,\dots,t_N=T)$ of the interval $[0,T]$, the values $q_j:=q(t_j)$, $j=1,\dots,N$, solve* (AIP$_{\varepsilon,\Pi}$) *with $\varepsilon=c_R$ (fixed), i.e.*

$$q_j\in\operatorname{Argmin}_{c_R\Delta t_j}\left\{\,\mathcal{E}(t_j,\hat q)+\mathcal{D}(q_{j-1},\hat q)\ :\ \hat q\in\mathcal{Q}\,\right\}\qquad \textit{for } j=1,\dots,N.$$

The quantity c_R is called the **reverse approximation constant** *of the system to the initial value q_0.*

Proof. The energy balance (E) implies

$$\mathcal{E}(t_j, q_j) + \mathrm{Diss}_{\mathcal{D}}(q; [t_{j-1}, t_j]) = \mathcal{E}(t_{j-1}, q_{j-1}) + \int_{t_{j-1}}^{t_j} \partial_t \mathcal{E}(\tau, q(\tau))\,\mathrm{d}\tau.$$

The stability $q_{j-1} \in \mathcal{S}(t_{j-1})$ here reads as

$$\mathcal{E}(t_{j-1}, q_{j-1}) \leq \mathcal{E}(t_{j-1}, \hat{q}) + \mathcal{D}(q_{j-1}, \hat{q}) \qquad \text{for all } \hat{q} \in \mathcal{Q}.$$

Together with $\mathcal{D}(q_{j-1}, q_j) \leq \mathrm{Diss}_{\mathcal{D}}(q; [t_{j-1}, t_j])$ this gives

$$\mathcal{E}(t_j, q_j) + \mathcal{D}(q_{j-1}, q_j) \leq \mathcal{E}(t_{j-1}, \hat{q}) + \mathcal{D}(q_{j-1}, \hat{q}) + \int_{t_{j-1}}^{t_j} \partial_t \mathcal{E}(\tau, q(\tau))\,\mathrm{d}\tau \tag{3.6}$$

and we continue by estimating the integral term using the growth estimate (3.3) to find

$$\begin{aligned} \int_{t_{j-1}}^{t_j} \partial_t \mathcal{E}(\tau, q(\tau))\,\mathrm{d}\tau &\leq \int_{t_{j-1}}^{t_j} c_1^E \left(\mathcal{E}(0, q_0) + c_0^E\right) \mathrm{e}^{c_1^E \tau}\,\mathrm{d}\tau \\ &\leq c_1^E \left(\mathcal{E}(0, q_0) + c_0^E\right) \mathrm{e}^{c_1^E T} \Delta t_j. \end{aligned} \tag{3.7}$$

We assume $\mathcal{E}(t_{j-1}, \hat{q}) < \infty$ (this will be justified later). Then, the quantity $\mathcal{E}(t_{j-1}, \hat{q})$ can be transformed and estimated using (3.1) as follows:

$$\begin{aligned} \mathcal{E}(t_{j-1}, \hat{q}) &= \mathcal{E}(t_j, \hat{q}) - \int_{t_{j-1}}^{t_j} \partial_t \mathcal{E}(\tau, \hat{q})\,\mathrm{d}\tau \\ &\leq \mathcal{E}(t_j, \hat{q}) + \int_{t_{j-1}}^{t_j} c_1^E \left(\mathcal{E}(0, \hat{q}) + c_0^E\right) \mathrm{e}^{c_1^E \tau}\,\mathrm{d}\tau \end{aligned} \tag{3.8}$$

Now choose $\hat{q} := q_j^*$ with

$$q_j^* \in \mathrm{Argmin}_\delta \left\{ \mathcal{E}(t_j, \hat{q}) + \mathcal{D}(q_{j-1}, \hat{q}) \ : \ \hat{q} \in \mathcal{Q} \right\}$$

for some $0 < \delta \leq \min\{1, \Delta t_j\}$. Such a q_j^* always exists. By Lemma 3.2, we can bound $\mathcal{E}(0, q_j^*)$ in (3.8) by a constant $M = M(q_0, \delta) \leq M(q_0, 1) =: M_1$ (note the monotonicity of M in δ), which does not depend on Π (or any other quantities except q_0). This gives

$$\mathcal{E}(t_{j-1}, q_j^*) \leq \mathcal{E}(t_j, q_j^*) + c_1^E (M_1 + c_0^E) \mathrm{e}^{c_1^E T} \Delta t_j. \tag{3.9}$$

Plugging (3.7) and (3.9) into (3.6), we see

$$\begin{aligned} \mathcal{E}(t_j, q_j) + \mathcal{D}(q_{j-1}, q_j) &\leq \mathcal{E}(t_j, q_j^*) + \mathcal{D}(q_{j-1}, q_j^*) + c_1^E(\mathcal{E}(0, q_0) + M_1 + 2c_0^E) \mathrm{e}^{c_1^E T} \Delta t_j \\ &\leq \inf \left\{ \mathcal{E}(t_j, \hat{q}) + \mathcal{D}(q_{j-1}, \hat{q}) \ : \ \hat{q} \in \mathcal{Q} \right\} + \delta + \tilde{c} \Delta t_j, \\ &\leq \inf \left\{ \mathcal{E}(t_j, \hat{q}) + \mathcal{D}(q_{j-1}, \hat{q}) \ : \ \hat{q} \in \mathcal{Q} \right\} + c_R \Delta t_j \end{aligned}$$

where we have set $\tilde{c} := c_1^E(\mathcal{E}(0, q_0) + M_1 + 2c_0^E) \mathrm{e}^{c_1^E T}$ and $c_R := 1 + \tilde{c}$. As M_1 only depends on q_0, so does c_R and the proof is complete. □

Remark 3.4. Counterexample 3.1 showed that the error order $\mathrm{O}(\|\Pi\|)$ is optimal: If the solution q_+ is to be selected, the discrete solution must jump from 0 to +1 at time t_1. The difference between $\mathcal{E}(t_1, -1)$ and $\mathcal{E}(t_1, +1)$ is t_1, hence the error in the minimization of $\mathcal{E}(t_1, \bullet) + \mathcal{D}(0, \bullet)$ is $t_1 = \mathrm{O}(\|\Pi\|)$ and nothing better than linear order can be archieved.

In summary, the results of the previous section suggest that ($\mathrm{AIP}_{\varepsilon,\Pi}$) is better suited than (IP) as a time discretization of (S) & (E). There is also another interesting conclusion to be drawn: While for the existence proof in Theorem 2.9 (see Theorem 2.19 for the proof) we need a local approximation grade of $\varepsilon\Delta t_j$ with $\varepsilon \to 0$, i.e. of order $\mathrm{o}(\Delta t_j)$, when going backwards from a solution, we can only expect an approximation grade of $c_R\Delta t_j = \mathrm{O}(\Delta t_j)$ by the preceding counterexample.

Remark 3.5. An inspection of the proof of Theorem 3.3 shows that the reverse approximation constant c_R does depend on the initial value q_0 only through the energy $\mathcal{E}(0, q_0)$, so c_R can be made independent of q_0 if an additional energy bound on all admissible initial values is imposed.

3.2 Sequences of problems

In this section we will show that also in the case of a sequence of systems, solutions to (S_∞) & (E_∞) can be approximated by solutions to ($\mathrm{AIP}_{k,\varepsilon,\Pi}$). The result is based on the reverse approximation result of the previous section and Section 1.3. Applications include relaxation and smooth approximation, see Chapter 4.

In all of the following, we assume (A1), (A3), (A5), (Γ1) and (Γ2) from Section 2.4.1 to hold.

Similarly to the previous section, the following a-priori estimates follow for all energetic solutions $q \in \mathcal{Q}[0,T]$ and for all $t \in [0,T]$ (by the Gronwall lemma):

$$\begin{aligned} \mathcal{E}_\infty(t, q(t)) + c_0^E &\le (\mathcal{E}_\infty(0, q(0)) + c_0^E)\mathrm{e}^{c_1^E t}, \\ |\partial_t \mathcal{E}_\infty(t, q(t))| &\le c_1^E(\mathcal{E}_\infty(0, q(0)) + c_0^E)\mathrm{e}^{c_1^E t}, \end{aligned} \tag{3.10}$$

cf. [Mie05, Section 3.1].

The next theorem shows that solutions to (S_∞) & (E_∞) all originate from solutions to the the usual approximate time-incremental problem (cf. page 32)

$$\begin{cases} \text{For } j = 1, \dots, N \text{ inductively find } q_j^k \in \mathcal{Q} \text{ such that} \\ q_j^k \in \mathrm{Argmin}_{\varepsilon\Delta t_j} \left\{ \mathcal{E}_k(t_j, \hat{q}) + \mathcal{D}_k(q_{j-1}^k, \hat{q}) \ : \ \hat{q} \in \mathcal{Q} \right\} \end{cases} \tag{$\mathrm{AIP}_{k,\varepsilon,\Pi}$}$$

for all $\varepsilon > 0$ and all sufficiently fine partitions $\Pi = (0 = t_0, \dots, t_N = T)$ of $[0,T]$.

Theorem 3.6. *Let $q_\infty : [0,T] \to \mathcal{Q}$ be a solution to* (S_∞) *&* (E_∞) *with initial value $q_0 = q_\infty(0) \in \mathcal{S}_\infty(0)$. Then, for all $\varepsilon > 0$, for all partitions $\Pi = (0 = t_0, \dots, t_N = T)$*

of $[0,T]$ *with* $\|\Pi\| \leq \varepsilon/(2c_R)$ ($c_R = c_R(q_0) > 0$ *is the reverse approximation constant from Theorem 3.3 applied to* $\mathcal{E}_\infty$ *and* $\mathcal{D}_\infty$), *and for all* k, *there exist a discrete solution* $q_k^\Pi := (q_0^k, q_1^k, \dots, q_N^k)$, *defined on the partition* Π, *of the* ε*-approximate incremental problem* $(\mathrm{AIP}_{k,\varepsilon,\Pi})$ *associated with* $\mathcal{E}_k$ *and* $\mathcal{D}_k$, *such that* $q_j^k \to q(t_j)$ *as* $k \to \infty$.

Proof. The main idea of the proof is to first construct a discrete solution to $(\mathrm{AIP}_{\infty,\varepsilon/2,\Pi})$ and then show how this discrete solution can be changed to yield a solution of $(\mathrm{AIP}_{k,\varepsilon,\Pi})$ for k sufficiently large. In detail, however, some further technicalities are needed. In particular, one has to be careful in choosing the q_j^k, because the approximate incremental problem at step j depends on the choice of the previous q_{j-1}^k.

As $\mathcal{E}_\infty$ and $\mathcal{D}_\infty$ fulfill all the prerequisites of Theorem 3.3, for a partition Π sufficiently fine, i.e. $\|\Pi\| \leq \varepsilon/(2c_R)$, we can find a discrete $\varepsilon/2$-solution $\tilde{q}^\Pi = (\tilde{q}_0, \tilde{q}_1, \dots, \tilde{q}_N)$ for $(\mathrm{AIP}_{\infty,\varepsilon/2,\Pi})$, i.e.

$$\tilde{q}_j \in \mathrm{Argmin}_{\varepsilon/2} \big(\mathcal{E}_\infty(t_j,\bullet) + \mathcal{D}_\infty(\tilde{q}_{j-1},\bullet)\big) \qquad \text{for } j = 1,\dots,N.$$

Note that Theorem 3.3 uses the choice $\tilde{q}_j = q(t_j)$.

By assumptions (Γ1) and (Γ2), we have the joint Γ-convergence

$$\mathcal{E}_\infty(t_j,\bullet) + \mathcal{D}_\infty(\tilde{q}_{j-1},\bullet) = \Gamma\text{-}\lim_k(\mathcal{E}_k(t_j,\bullet) + \mathcal{D}_k(\tilde{q}_{j-1},\bullet)),$$

cf. (2.9). Condition (A3) provides the equi-mild coercivity (1.1) (in fact, take $E := \mathcal{E}_\infty(t_j,\tilde{q}_{j-1}) + 1$ and observe that for a recovery sequence $q_{j-1}^k \to \tilde{q}_{j-1}$, $\inf_\mathcal{Q} \big(\mathcal{E}_k(t_j,.) + \mathcal{D}_k(\tilde{q}_{j-1},.)\big) \leq E$ for all k large enough). Now, the invocation of Proposition 1.10 (with $\delta = \varepsilon/6$) on each $\tilde{q}_j$ provides us with recovery sequences $(q_j^k)_k$ with $q_j^k \to \tilde{q}_j$ as $k \to \infty$ and

$$\mathcal{E}_k(t_j,q_j^k) + \mathcal{D}_k(\tilde{q}_{j-1},q_j^k) \leq \inf_\mathcal{Q} \big(\mathcal{E}_k(t_j,\bullet) + \mathcal{D}_k(\tilde{q}_{j-1},\bullet)\big) + \frac{\varepsilon}{2} + \frac{\varepsilon}{6} \tag{3.11}$$

for all $j = 1,\dots,N$ and k sufficiently large (note that $q_0^k = \tilde{q}_0 = q_0$).

Because $q_j^k \to \tilde{q}_j$ as $k \to \infty$ and the energies are bounded for k sufficiently large (cf. (3.11)), the continuous convergence assumption (Γ2) shows $\mathcal{D}_k(q_j^k,\tilde{q}_j) \to 0$ and $\mathcal{D}_k(\tilde{q}_j,q_j^k) \to 0$ as $k \to \infty$, i.e.

$$\max\big\{\mathcal{D}_k(q_j^k,\tilde{q}_j), \mathcal{D}_k(\tilde{q}_j,q_j^k)\big\} \leq \frac{\varepsilon}{6} \tag{3.12}$$

for all $j = 1,\dots,N$ and k sufficiently large.

So far we have constructed sequences and selected some $k_0 = k_0(\varepsilon)$ large enough such that (3.11) and (3.12) are fulfilled for all q_j^k with $k \geq k_0$. We still need to show that these q_j^k form a discrete solution to $(\mathrm{AIP}_{k,\varepsilon,\Pi})$.

For all $k \geq k_0$ and all $j = 2,\dots,N$, we find by the triangle inequality and (3.12)

$$\begin{aligned}\inf_\mathcal{Q} \big(\mathcal{E}_k(t_j,\bullet) + \mathcal{D}_k(\tilde{q}_{j-1},\bullet)\big) &\leq \inf_\mathcal{Q} \big(\mathcal{E}_k(t_j,\bullet) + \mathcal{D}_k(q_{j-1}^k,\bullet)\big) + \mathcal{D}_k(\tilde{q}_{j-1},q_{j-1}^k) \\ &\leq \inf_\mathcal{Q} \big(\mathcal{E}_k(t_j,\bullet) + \mathcal{D}_k(q_{j-1}^k,\bullet)\big) + \frac{\varepsilon}{6}.\end{aligned} \tag{3.13}$$

In the case $j = 1$, we have $q_0^k = \tilde{q}_0 = q_0$ for all k and hence (3.13) also holds for $j = 1$.

Now, using first the triangle inequality, then (3.11) and (3.12), and finally (3.13), we deduce

$$\begin{aligned}\mathcal{E}_k(t_j, q_j^k) + \mathcal{D}_k(q_{j-1}^k, q_j^k) &\le \mathcal{E}_k(t_j, q_j^k) + \mathcal{D}_k(\tilde{q}_{j-1}, q_j^k) + \mathcal{D}_k(q_{j-1}^k, \tilde{q}_{j-1}) \\ &\le \inf_{\mathcal{Q}} \big(\mathcal{E}_k(t_j, \bullet) + \mathcal{D}_k(\tilde{q}_{j-1}, \bullet)\big) + \frac{\varepsilon}{2} + \frac{\varepsilon}{6} + \frac{\varepsilon}{6} \\ &\le \inf_{\mathcal{Q}} \big(\mathcal{E}_k(t_j, \bullet) + \mathcal{D}_k(q_{j-1}^k, \bullet)\big) + \frac{\varepsilon}{2} + \frac{\varepsilon}{6} + \frac{\varepsilon}{6} + \frac{\varepsilon}{6} \\ &= \inf_{\mathcal{Q}} \big(\mathcal{E}_k(t_j, \bullet) + \mathcal{D}_k(q_{j-1}^k, \bullet)\big) + \varepsilon.\end{aligned}$$

But this is just

$$q_j^k \in \operatorname{Argmin}_\varepsilon \big(\mathcal{E}_k(t_j, \bullet) + \mathcal{D}_k(q_{j-1}^k, \bullet)\big) \qquad \text{for } j = 1, \ldots, N.$$

This shows the existence of solutions to $(\mathrm{AIP}_{k,\varepsilon,\Pi})$ for $k \ge k_0 = k_0(\varepsilon)$. Trivially, we can fill up this sequence for $k < k_0$ with arbitrary solutions to $(\mathrm{AIP}_{k,\varepsilon,\Pi})$. The claim $q_j^k \to \tilde{q}_j = q(t_j)$ is clear by the choice of the q_j^k. □

Next, we see that we can even find a sequence of partitions such that the discrete approximations converge on a dense set if the topology is metrizable on compact sets:

Theorem 3.7. *Besides the usual assumptions, assume that $\mathcal{Q}$ is metrizable on compact subsets. Let $q_\infty : [0,T] \to \mathcal{Q}$ be a solution to (S_∞) & (E_∞) with initial value $q_0 = q_\infty(0) \in \mathcal{S}_\infty(0)$. Then, there exists a sequence of (nested) partitions $\Pi_n = (0 = t_0^n, \ldots, t_{N(n)}^n = T)$ of $[0,T]$ with $\|\Pi_n\| \to 0$, a sequence $(k_n)_n$ of problem indices ($k_n \uparrow \infty$), and discrete solution $q_{k_n}^{\Pi_n} = (q_0^{k_n}, q_1^{k_n}, \ldots, q_{N(n)}^{k_n})$, defined on the partition Π_n, of the approximate incremental problem $(\mathrm{AIP}_{k_n,\varepsilon_n,\Pi_n})$ such that the piecewise constant interpolants $\bar{q}_{k_n}^{\Pi_n} : [0,T] \to \mathcal{Q}$ of these discrete solutions converge on a dense subset of $[0,T]$ to the solution q_∞.*

Proof. Let $\Pi_n \subseteq \Pi_{n+1}$ be a sequence of nested partitions with $\|\Pi_n\| \le 1/(2c_R n)$, where $c_R = c_R(q_0) > 0$ is the reverse approximation constant from Theorem 3.3. Then, $A := \bigcup_{n\in\mathbb{N}} \Pi_n$ is dense in $[0,T]$. Applying Theorem 3.6, we find for each $n \in \mathbb{N}$ a sequence $q^{n,k} = (q_j^{n,k})_{j=0,\ldots,N(n)}$, $k \in \mathbb{N}$, of solutions to $(\mathrm{AIP}_{k,\varepsilon_n,\Pi_n})$ such that

$$q_j^{n,k} \to q(t_j^n), \qquad \mathcal{E}_k(t_j^n, q_j^{n,k}) \to \mathcal{E}(t_j^n, q(t_j^n)) \qquad \text{as } k \to \infty$$

for all $n \in \mathbb{N}$, $j = 0, ..., N(n)$.

With $E := \sup\{\, \mathcal{E}_\infty(t, q(t)) \;:\; t \in [0,T] \,\} < \infty$ (cf. (3.10)), we find $K(n) \in \mathbb{N}$ such that $\max\{\, \mathcal{E}_k(t_j^n, q_j^{n,k}) \;:\; j = 0, 1, ..., N(n) \,\} \le E + 1$ for all $k \ge K(n)$. Since the sublevels are uniformly contained in a compact set by (A3), we obtain, for fixed $n \in \mathbb{N}$,

$$D_n(k) = \max\big\{\, d(q_j^{n,k}, q(t_j^n)) \;:\; j = 0, 1, ..., N(n) \,\big\} \to 0 \qquad \text{as } k \to \infty,$$

where $d(\bullet,\bullet)$ is the metric on the sublevel corresponding to the energy bound $E+1$. Choose the subsequence $(k_n)_n$, $k_n \uparrow \infty$, such that $k_n \geq K(n)$ and $D_n(k_n) \leq 1/n$, and define the solutions $q_{k_n}^{\Pi_n} := (q_0^{k_n}, \ldots, q_{N(n)}^{k_n})$ via $q_j^{k_n} = q_j^{n,k_n}$. Since the sequence of partitions is nested, for each $t \in A$ there exists an index $m(t)$ such that $t \in \Pi_n$ for $n \geq m(t)$, i.e., $t = t_{J(t,n)}^n$. Thus,

$$d(\bar{q}_{k_n}^{\Pi_n}(t), q(t)) = d(q_{J(t,n)}^{n,k_n}, q(t_{J(t,n)}^n)) \leq D_n(k_n) \leq \frac{1}{n} \quad \text{for } n \geq m(t).$$

This is the desired convergence result and the theorem is established. □

In certain situations it is possible to strengthen the preceding result to convergence everywhere by using similar ideas as in the proof of the Helly selection Theorem 2.13, cf. [MR07a, Remark 4.4].

Clearly, just like in Section 3.1, we cannot expect strict approximability of solutions to (S_∞) & (E_∞) by discrete solutions of $(\mathrm{IP}_{k,\Pi})$ instead of $(\mathrm{AIP}_{k,\varepsilon,\Pi})$. This, in fact, has been settled already in Counterexample 3.1, because the latter shows that even for a *constant* sequence of functionals, we cannot get strict approximability.

To conclude this section, we show that one also cannot expect approximability of solutions to (S_∞) & (E_∞) by *time-continuous* solutions to (S_k) & (E_k) instead of approximate *time-incremental* solutions to $(\mathrm{AIP}_{k,\varepsilon,\Pi})$.

Counterexample 3.8. Consider the state-space $\mathcal{Q} = [0,1]$, the time interval $[0,T] = [0,2]$ and the stationary energy functionals

$$\begin{aligned} \mathcal{E}_k(t,q) &:= \frac{q^2}{2k} - q, \qquad \text{for all } k \in \mathbb{N} \\ \mathcal{E}_\infty(t,q) &:= -q \end{aligned}$$

for $t \in [0,2]$ and $q \in \mathbb{R}$. Also, choose $\mathcal{D}_k(q_1,q_2) = \mathcal{D}_\infty(q_1,q_2) := \mathcal{D}(q_1,q_2) := |q_2 - q_1|$. As initial value we select $q_0 = 0$. This setting can be seen as a degenerately convex problem in the limit $k = \infty$ with strictly convex approximations for $k \in \mathbb{N}$.

The process

$$q_\infty(t) := \begin{cases} 0 & \text{if } t \in [0,1), \\ 1 & \text{if } t \in [1,2], \end{cases}$$

is one of the many solutions of the rate-independent formulation associated with $\mathcal{E}_\infty$ and $\mathcal{D}$. The stable states $\mathcal{S}_\infty(t)$ are easily seen to be the whole space, i.e. $\mathcal{S}_\infty(t) = \mathcal{S}_\infty(0) = [0,1]$, thus the stability condition is trivially fulfilled. For $t \in [0,1)$, the energy balance is trivial and for $t \in [1,2]$ we have

$$\begin{aligned} \mathcal{E}_\infty(t, q_\infty(t)) + \mathrm{Diss}_{\mathcal{D}}(q_\infty; [0,t]) &= \mathcal{E}_\infty(t,1) + \mathcal{D}(0,1) = -1 + 1 = 0 \\ &= \mathcal{E}_\infty(0, q_0) + \int_0^t \partial_t \mathcal{E}_\infty(\tau, q_\infty(\tau))\, \mathrm{d}\tau. \end{aligned}$$

Hence, q_∞ is an energetic solution of (S_∞) & (E_∞). We now show that q_∞ cannot be approximated by solutions to (S_k) & (E_k).

For all $k \in \mathbb{N}$, the stability sets $\mathcal{S}_k(t) = \mathcal{S}_k(0)$ again are the whole space $[0,1]$, since it holds for all $q, \hat{q} \in [0,1]$ that

$$\begin{aligned}\mathcal{E}_k(t,\hat{q}) + \mathcal{D}(q,\hat{q}) - \mathcal{E}_k(t,q) &= \frac{\hat{q}^2 - q^2}{2k} + (q-\hat{q}) + |\hat{q} - q| \\ &= \left.\begin{cases} (\hat{q}^2 - q^2)/(2k) & \text{if } \hat{q} \geq q \\ (q-\hat{q})(2 - (\hat{q}+q)/(2k)) \geq (q-\hat{q})(2-k^{-1}) & \text{if } \hat{q} < q \end{cases}\right\} \geq 0,\end{aligned}$$

i.e. $q \in \mathcal{S}_k(t)$. Now, the zero-process $q_k \equiv 0$ trivially is a solution of (S_k) & (E_k) and because the problem is strictly convex and the stability sets are convex, we immediately get the uniqueness of this solution from Theorem 2.22. But, obviously, the zero-process does not approximate q_∞ in any reasonable sense.

The results of the preceding section have shown that ($\mathrm{AIP}_{k,\varepsilon,\Pi}$) is the “right” problem to work with for reverse approximation. This complements the existence results from Sections 2.3 and 2.4, where also some variant of (AIP) was used instead of (IP) (which is by the way also possible for some, but not all, existence theorems).

It should finally be noted that the preceding reverse approximation results can also be made quantitative, which is for example useful when considering numerical approximations. See Section 5 of [MR07a] for details.

3.3 Contraction properties of the incremental problems and stability

In this section, we investigate the contraction properties of the minimization problems in (AIP) and, as a consequence of the reverse approximation results from the previous section, prove a theorem for the distance between two solutions to the same initial value.

In all of the following, we assume (A1), (A3), (A5), (Γ1), (Γ2) and additionally

Compact metrizability of $\mathcal{Q}$:
All compact sets $K \subseteq \mathcal{Q}$ are metrizable. (CM)

We can now prove the following proposition, which gives information about the behavior of the sets of ε-minimizers of the incremental problem (AIP) in the strictly convex case. This could, for example, be used for error estimates in numerical schemes.

Proposition 3.9. *Let $K \subseteq \mathcal{Q}$ be compact such that*

$$\bigcup_{t\in[0,T], q\in K, k\in\mathbb{N}_\infty} \mathrm{Argmin}_\delta \big(\mathcal{E}_k(t,\bullet) + \mathcal{D}_k(q,\bullet)\big) \tag{3.14}$$

is relatively compact for some $\delta > 0$ and $\sup_{q\in K, k\in\mathbb{N}_\infty} \mathcal{E}_k(0,q) < \infty$. Also assume that $\mathcal{E}_\infty(t,\bullet) + \mathcal{D}_\infty(q,\bullet)$ is strictly convex for all $t \in [0,T]$ and all $q \in K$. Then, for all $\eta > 0$ there exist $\varepsilon_0 = \varepsilon_0(\eta) > 0$ and $k_0 = k_0(\eta) \in \mathbb{N}$ such that for all $\varepsilon \in [0,\varepsilon_0]$ and $k \geq k_0$ it holds that

$$\operatorname{diam}(\operatorname{Argmin}_\varepsilon(\mathcal{E}_k(t,\bullet) + \mathcal{D}_k(q,\bullet))) \leq \eta$$

for all $t \in [0,T]$, $q \in K$

Of course, by (A1), it suffices to check (3.14) for fixed $t = t_0$.

Proof. Assume to the contrary that there exists an $\eta > 0$ such that we can find a strictly increasing sequence $k_j \uparrow \infty$, a falling sequence $\varepsilon_j \downarrow 0$ and corresponding sequences $(t_j)_j \subseteq [0,T]$, $(q_j)_j \subseteq K$ with

$$\operatorname{diam}(\operatorname{Argmin}_{\varepsilon_j}(\mathcal{E}_{k_j}(t_j,\bullet) + \mathcal{D}_{k_j}(q_j,\bullet))) > \eta.$$

As K is compact we can find subsequences (not relabelled) with $t_j \to t \in [0,T]$ and $q_j \to q \in K$. From the definition of the diameter, we can choose two sequences $(x_j)_j \subseteq K$ and $(y_j)_j \subseteq K$ with $d(x_j,y_j) \geq \eta$ and with

$$x_j, y_j \in \operatorname{Argmin}_{\varepsilon_j}(\mathcal{E}_{k_j}(t_j,\bullet) + \mathcal{D}_{k_j}(q_j,\bullet)) \subseteq \operatorname{Argmin}_\mu(\mathcal{E}_{k_j}(t_j,\bullet) + \mathcal{D}_{k_j}(q_j,\bullet)),$$

where the last inclusion holds for all $\mu > 0$ and j so large that $\varepsilon_j \leq \mu$. Selecting a subsequence by assumption (3.14), we can assume without loss of generality that $x_j \to x$ and $y_j \to y$ with $d(x,y) \geq \eta$ (as in (1.7)).

From the definition of the Kuratowski–upper limit we know that

$$x, y \in \text{K-}\limsup_j \operatorname{Argmin}_\mu(\mathcal{E}_{k_j}(t_j,\bullet) + \mathcal{D}_{k_j}(q_j,\bullet)).$$

It remains to show that

$$\text{K-}\limsup_j \operatorname{Argmin}_\mu(\mathcal{E}_{k_j}(t_j,\bullet) + \mathcal{D}_{k_j}(q_j,\bullet)) \subseteq \operatorname{Argmin}_\mu(\mathcal{E}_\infty(t,\bullet) + \mathcal{D}_\infty(q,\bullet)). \quad (3.15)$$

Then we have $x, y \in \operatorname{Argmin}_\mu(\mathcal{E}_\infty(t,\bullet) + \mathcal{D}_\infty(q,\bullet))$ and hence, for μ was arbitrary, also $x, y \in \operatorname{Argmin}(\mathcal{E}_\infty(t,\bullet) + \mathcal{D}_\infty(q,\bullet))$. But because of strict convexity, this set is (at most) a singleton, which contradicts $d(x,y) \geq \eta$.

To show (3.15), we note that $\mathcal{E}_{k_j}(t_j,\bullet) + \mathcal{D}_{k_j}(q_j,\bullet)$ Γ-converges to $\mathcal{E}_\infty(t,\bullet) + \mathcal{D}_\infty(q,\bullet)$, this follows from (A1), (Γ1) and (Γ2) (subsequences of Γ-converging sequences also Γ-converge and we always stay in a subset of bounded energy, also cf. the method of proof in Lemma 2.17). Now apply Theorem 1.15 to get

$$\begin{aligned}
&\text{K-}\limsup_j \operatorname{Argmin}_\mu(\mathcal{E}_{k_j}(t_j,\bullet) + \mathcal{D}_{k_j}(q_j,\bullet)) \\
&\quad \subseteq \bigcap_{\delta>0} \text{K-}\limsup_j \operatorname{Argmin}_{\mu+\delta}(\mathcal{E}_{k_j}(t_j,\bullet) + \mathcal{D}_{k_j}(q_j,\bullet)) \\
&\quad = \operatorname{Argmin}_\mu(\mathcal{E}_\infty(t,\bullet) + \mathcal{D}_\infty(q,\bullet)).
\end{aligned}$$

This shows (3.15) and therefore completes the proof. $\square$

Note that (3.14) is not restrictive since the usual a-priori estimates, for example (3.10), together with (A3) allow to restrict attention to a compact subset of $\mathcal{Q}$. Also, in the preceding proposition we only assumed strict convexity for the limit problem and not for the approximating problems.

We end this chapter by showing how the shape of the sets of ε-minimizers determines the **reachability region**, i.e. the region of $[0,T]\times\mathcal{Q}$ where energetic solutions to the same initial value can proceed. This in particulat provides a "stability estimate" on the distance between different solutions of (S_∞) & (E_∞) to the same initial value (recall from Section 2.5 that in general uniqueness of solutions cannot be expected).

Proposition 3.10. *For a fixed initial value $q_0 \in \mathcal{S}_\infty(0)$, let*

$$M_0 := c_R + 2c_1^E(\mathcal{E}_\infty(0,q_0) + c_0^E)\mathrm{e}^{2c_1^E T},$$

where c_R is the reverse approximation constant, and assume that $d(\bullet,\bullet)$ is a metric on the compact set $\mathrm{cl}(\bigcup_{t\in[0,T]}\{\, q\in\mathcal{Q} \,:\, \mathcal{E}_\infty(t,q) \le M_0 T\,\})$. Define the increasing **modulus of contraction**

$$\omega_0(t) := \mathrm{diam}(\mathrm{Argmin}_{M_0 t}(\mathcal{E}_\infty(0,\bullet) + \mathcal{D}_\infty(q_0,\bullet))).$$

Then, two solution processes $q,\tilde q : [0,T]\to\mathcal{Q}$ of (S_∞) & (E_∞) to the same initial value $q_0\in\mathcal{S}_\infty(0)$ satisfy

$$d(q(t),\tilde q(t)) \le \omega_0(t).$$

The function $\omega_0 : [0,T]\to\mathbb{R}_\infty$ is right-continuous and if additionally $\mathcal{E}_\infty(0,\bullet)$ is convex[1]*, it is even continuous.*

Proof. The case $t=0$ is trivial. Let $t\in(0,T]$. We apply the Reverse Approximation Theorem 3.3 to the problem on the shorter time interval $[0,t]$ and with the trivial partition $\Pi=(0,t)$ to get

$$q(t),\tilde q(t) \in \mathrm{Argmin}_{c_R t}(\mathcal{E}_\infty(t,\bullet) + \mathcal{D}_\infty(q_0,\bullet))$$

and it follows by (3.1) and (3.2) that

$$\begin{aligned}
\mathcal{E}_\infty(0,q(t)) &\le \mathcal{E}_\infty(t,q(t)) + c_1^E(\mathcal{E}_\infty(t,q(t)) + c_0^E)\mathrm{e}^{c_1^E T} t\\
&\le \mathcal{E}_\infty(t,q_0) + c_R t + c_1^E(\mathcal{E}_\infty(0,q_0) + c_0^E)\mathrm{e}^{2c_1^E T} t\\
&\le \mathcal{E}_\infty(0,q_0) + c_R t + 2c_1^E(\mathcal{E}_\infty(0,q_0) + c_0^E)\mathrm{e}^{2c_1^E T} t,
\end{aligned}$$

likewise for $\tilde q(t)$. Hence, because $q_0\in\mathrm{Argmin}(\mathcal{E}_\infty(0,\bullet)+\mathcal{D}_\infty(q_0,\bullet))$ by stability,

$$q(t),\tilde q(t) \in \mathrm{Argmin}_{M_0 t}(\mathcal{E}_\infty(0,\bullet) + \mathcal{D}_\infty(q_0,\bullet))$$

[1]This, of course, assumes that $\mathcal{Q}$ is a linear space.

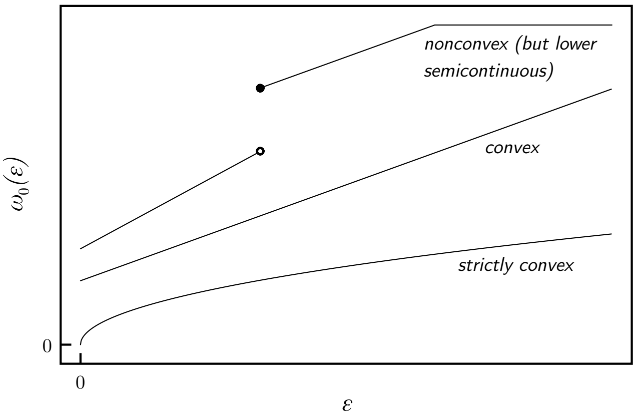

Figure 4: Several possible functions ω_0 (the notes underneath the graphs indicate additional qualifications on $\mathcal{E}$).

and the claim follows.

The properties of ω_0 follow from Proposition 1.16 (note that coercivity follows from (3.14) and lower semicontinuity is clear since we are working with Γ-limits, cf. Proposition 1.5). □

If additionally $\mathcal{E}_\infty(t, \bullet) + \mathcal{D}_\infty(q_0, \bullet)$ is strictly convex, then $\omega_0(0) = 0$. Some examples for functions ω_0 are given in Figure 4. If one thinks of these graphs mirrored by the ε-axis and with the ω_0-axis replaced by the space $\mathcal{Q}$ (visualized as a one-dimensional space where q_0 takes its place at the origin), one sees the shape of the reachability region of the process.

Of course, the preceding proposition can also be adapted to the case of only one system (just erase all "$_\infty$").

Remark 3.11. Except for the continuity properties of ω_0, all the results of the preceding proposition also hold true if the metric $d(\bullet, \bullet)$ *does not generate the topology* on $\mathrm{cl}(\bigcup_{t\in[0,T]}\{\, q \in \mathcal{Q} \;:\; \mathcal{E}_\infty(t, q) \le M_0 T \,\})$. For example, one can use an arbitrary norm to measure the distance between two energetic solutions. Of course, one has to be able to estimate ω_0, i.e. diameters of sublevels with respect to $d(\bullet, \bullet)$, but this might be possible in applications (for example, it could be possible to use the $\mathcal{Q}$-norm if $\mathcal{Q}$ is a Banach space).

4 Applications

This section applies the theoretical results developed in the previous chapters to concrete (integral) functionals. First, we investigate reverse approximation for a single system in Section 4.1. Then we see how this functional can be interpreted as the relaxation (i.e. the weakly lower semicontinuous envelope) of a nonconvex functional in Section 4.2 and finally consider approximation by regularized functionals in Section 4.3.

Not mentioned here are applications in numerics, where reverse approximation can yield a backward analysis of a numerical scheme, cf. [MR07a]. Other applications of reverse approximation, in particular applications to the optimal control of rate-independent systems, are currently under investigation by the author.

4.1 An example for reverse approximation

Let the state-space $\mathcal{Q}$ $(= \mathcal{Z})$ be the Sobolev space $\mathrm{H}^1(0,1)$ equipped with its weak topology. Consider the functional

$$\mathcal{E}(t,v) := \int_0^1 W^{**}(v'(x)) + v(x)^2 - f(t,x)v(x)\,\mathrm{d}x \qquad \text{for } t \in [0,T], v \in \mathrm{H}^1(0,T).$$

with a bounded function $f \in \mathrm{C}^1([0,T]\times[0,1])$, which represents a prescribed loading, and $W^{**}: \mathbb{R} \to [0,\infty)$ given by

$$W^{**}(s) = \begin{cases} (s+1)^2 & \text{if } s < -1, \\ 0 & \text{if } s \in [-1,1], \\ (s-1)^2 & \text{if } s > 1, \end{cases}$$

which is the convexification of $W(s) := \min\{(s-1)^2, (s+1)^2\}$, see Figure 5. This functional and its variants discussed below have applications in solid-solid phase transformations, cf. [Mül93], which also contains many references on other applications. As dissipation distance, we use the $\mathrm{L}^1(0,1)$–norm, i.e.

$$\mathcal{D}(u,v) := \|v-u\|_1\,.$$

One can now estimate using Young's inequality

$$\mathcal{E}(t,v) \geq \int_0^1 W^{**}(v'(x)) + \frac{v(x)^2}{2} - \frac{f^*}{2}\,\mathrm{d}x, \tag{4.1}$$

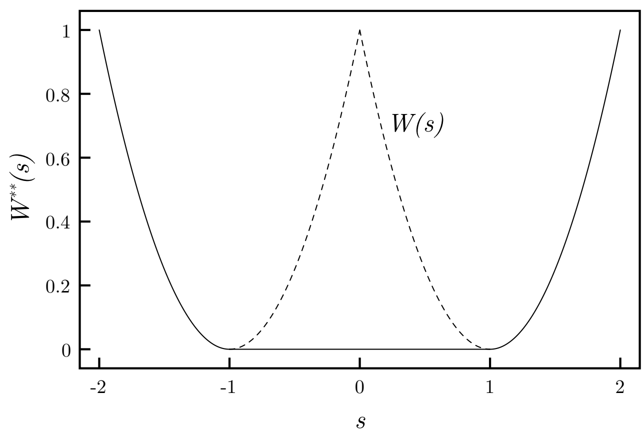

Figure 5: The energy densities W and W^{**}

where $f^* := \max_{[0,T]\times[0,1]} f$. Now, W^{**} satisfies

$$W^{**}(s) \geq \frac{s^2}{2} \qquad \text{for } |s| \geq s_0 \qquad \left(s_0 = \frac{\sqrt{2}}{\sqrt{2}-1} \right)$$

as can be seen from $W^{**}(s)/s^2 = (1 \pm s^{-1})^2 \geq 1/2$ for $|s| \geq s_0$. Hence, we have the coercivity estimate

$$2W^{**}(s) + s_0^2 \geq s^2 \qquad \text{for all } \; s \in \mathbb{R}.$$

Using this and (4.1), $\mathcal{E}(t,v) \leq E$ for $E \in \mathbb{R}$ implies

$$\|v\|^2_{\mathrm{L}^2(0,1)} \leq 2E + f^* \qquad \text{and} \qquad \|v'\|^2_{\mathrm{L}^2(0,1)} \leq 2E + s_0^2 + f^*,$$

which shows that all sublevels are uniformly bounded in $\mathrm{H}^1(0,1)$. Also, $\mathcal{E}(t,\bullet)$ is convex and strongly continuous (one can select an almost everywhere converging subsequence and use Fatou's lemma to pass to the limit in the nonnegative first part of the integrand; the second part with f is clearly strongly continuous). Hence, the sublevels are strongly closed and convex, thus weakly closed and $\mathcal{E}(t,\bullet)$ has compact sublevels.

With the preceding considerations in mind, we have that under suitable assumptions on the loading f, assumptions (A1')–(A4') hold. The assumptions (A5')–(A6') on the dissipation distance hold by Example 2.8. Hence, Theorem 2.9 shows the existence of a solution $q : [0,T] \to \mathrm{H}^1(0,1)$ to the energetic formulation (S) & (E) associated with the above functionals $\mathcal{E}$ and $\mathcal{D}$ and a stable initial value $q_0 \in \mathcal{S}(0)$. In particular, solutions to the approximate time-incremental problem

$$\begin{cases} \text{For } j = 1, \dots, N \text{ inductively find } q_j \in \mathcal{Q} \text{ such that} \\ q_j \in \operatorname{Argmin}_{\varepsilon\Delta t_j} \Big\{ \displaystyle\int_0^1 W^{**}(\hat{q}') + \hat{q}^2 - f(t)\hat{q} + |\hat{q} - q_{j-1}| \;\mathrm{d}x \;:\; \hat{q} \in \mathrm{H}^1(0,1) \Big\}. \end{cases}$$

on a partition $\Pi = (0 = t_0, \dots, t_N = T)$ of the interval $[0,T]$ admit a subsequence, which converges pointwise to a solution $q : [0,T] \to \mathrm{H}^1(0,1)$ of (S) & (E) (with bounded $\mathrm{L}^1(0,1)$-variation) if we let $\|\Pi\| \to 0$ and $\varepsilon \to 0$. The Reverse Approximation Theorem 3.3 shows that also all solutions originate from this time-incremental problem (with $\varepsilon = c_R$ fixed).

4.2 Relaxation

We might encounter energy functionals $\mathcal{E} : \mathcal{Q} \to \mathbb{R}_\infty$ for which an infimizing sequence converges, but the limit is no minimizer of $\mathcal{E}$. Such functionals $\mathcal{E}$ cannot have closed sublevels, i.e. they are not lower semicontinuous. In applications, this situation may be caused by the development of microstructure [Mül93, Mül99]. In order to still analyse such functionals, we can "relax" $\mathcal{E}$ to its lower semicontinuous envelope $\mathcal{E}^{**} := \operatorname{lsc}\mathcal{E}$ and study the problem associated with the new functional $\mathcal{E}^{**}$. The framework of Γ-convergence is designed in such a way that if we take the Γ-limit of the constant sequence $(\mathcal{E})_k$, we arrive at the relaxation $\mathcal{E}^{**}$ of $\mathcal{E}$, this was the content of Proposition 1.6. Thus, we can apply the reverse approximation methods for sequences developed above in order to understand the connection between the original and the relaxed problem.

Concretely, one can examine the energy functional (now named $\mathcal{E}^{**}$) from the last section again and define

$$\begin{aligned} \mathcal{E}(t,v) &:= \int_0^1 W(v'(x)) + v(x)^2 - f(t,x)v(x)\,\mathrm{d}x, \\ \mathcal{E}^{**}(t,v) &:= \int_0^1 W^{**}(v'(x)) + v(x)^2 - f(t,x)v(x)\,\mathrm{d}x \end{aligned}$$

for $t \in [0,T]$ and $v \in \mathrm{H}^1(0,T)$ (recall that $W(s) := \min\{(s-1)^2, (s+1)^2\}$, see Figure 5).

To see that $\mathcal{E}^{**}$ is the relaxation of $\mathcal{E}$, it suffices to notice that the integrand of $\mathcal{E}^{**}$ is the convexification with respect to $v'(x)$ of the integrand of $\mathcal{E}$. This implies that $\mathcal{E}^{**} = \operatorname{lsc}\mathcal{E}$ by standard results, see for example Theorem 2.18 in [Bra02][1].

Again, the Reverse Approximation Theorems 3.6 and 3.7 are applicable and we can conclude that every energetic solution to the rate-independent system associated with $\mathcal{E}^{**}$ and $\mathcal{D}$ originates from time-discrete solutions to the approximate time-incremental problems for the energy functional $\mathcal{E}$. This shows an important point of reverse approximation: Originally, we do not know whether the relaxed problem still carries any information on the original problem. Of course, the relaxed and the original problem

[1]The main argument goes as follows: $\mathcal{E}^{**} \le \operatorname{lsc}\mathcal{E}$ is clear by $\mathcal{E}^{**} \le \mathcal{E}$ and the lower semicontinuity of $\mathcal{E}^{**}$. For the other direction, one considers for any $z_1, z_2 \in \mathbb{R}$ and $\lambda \in (0,1)$ a suitable oscillatory function whose derivative takes only the values z_1 and z_2 in the ratio λ and $1-\lambda$, respectively. This derivative converges weakly to the constant $\lambda z_1 + (1-\lambda)z_2$ by the Riemann–Lebesgue lemma. One can conclude that the integrand of the lower semicontinuous envelope must be convex in the derivative of the function to which the functional is applied, hence $\operatorname{lsc}\mathcal{E} \le \mathcal{E}^{**}$.

should be somehow related in order for relaxation to be useful. The reverse approximation results from the preceding chapter give some insight into this relationship: The main Existence Theorem 2.19 shows that the relaxed problem is not "too small", i.e. a sequence of solutions to the approximate incremental problem ($\mathrm{AIP}_{k,\varepsilon,\Pi}$) for the original energy functional $\mathcal{E}$ admits a limit point, which is an energetic solution to (S^{**}) & (E^{**}) for the relaxed functional $\mathcal{E}^{**}$. In this work, we have shown that the approximate incremental problems ($\mathrm{AIP}_{k,\varepsilon,\Pi}$) also are not "too big", i.e. for every solution of (S^{**}) & (E^{**}) we can find an associated sequence of solutions to ($\mathrm{AIP}_{k,\varepsilon,\Pi}$).

In the terminology of the relaxation theory for rate-independent problems as introduced in [Mie03, Mie04], we have shown the *lower incremental relaxation* condition. Such a conditions has previously been seen to hold in the special case of the theory of phase transitions in elastic solids [The02].

4.3 Regular approximation

Sometimes it is useful to approximate functionals with regularized versions

Let again the state-space $\mathcal{Q}$ be the Sobolev space $\mathrm{H}^1(0,1)$ equipped with its weak topology. Consider the functionals ($k \in \mathbb{N}$)

$$\mathcal{E}_k(t,v) := \begin{cases} \displaystyle\int_0^1 \frac{1}{k}(v''(x))^2 + W(v'(x)) + v(x)^2 - f(t,x)v(x)\,\mathrm{d}x & \text{if } v \in \mathrm{H}^2(0,1), \\ +\infty & \text{otherwise}, \end{cases}$$
$$\mathcal{E}_\infty(t,v) := \int_0^1 W^{**}(v'(x)) + v(x)^2 - f(t,x)v(x)\,\mathrm{d}x \qquad \text{for } v \in \mathrm{H}^1(0,T),$$

where f, W and W^{**} are as above. Note that in order to apply $\mathcal{E}_k$, we need twice (weak) differentiability, which is only given in the subspace $\mathrm{H}^2(0,1)$. Therefore, we have set $\mathcal{E}_k := +\infty$ on $\mathrm{H}^1(0,1) \setminus \mathrm{H}^2(0,1)$. See [Mül93] for deeper investigations into these functionals. As before, for all k we use the $\mathrm{L}^1(0,1)$–norm as dissipation distance.

All the $\mathcal{E}_k$ have closed and bounded sublevels in $\mathrm{H}^1(0,1)$ (can be shown similarly to the previous sections). The coercivity bounds are uniform in k (for the $\mathrm{H}^1(0,1)$–norm) since we can simply disregard the first term of the integrand and proceed as before. Thus, solutions to the rate-independent problems associated with $\mathcal{E}_k$ and $\mathcal{D}$ exist by Theorem 2.19 and can be seen as generalized solutions to the evolutionary differential inclusion (which is obtained by formally differentiating the functionals)

$$0 \in \mathrm{Sign}(\partial_t q) + \frac{2}{k}\partial_x^4 q - \partial_x(\mathrm{D}W(\partial_x q)) + 2q - f(t,\bullet) \qquad \text{a.e. in } (t,x) \in [0,T] \times \Omega,$$

together with a smooth stable initial condition $q(0,\bullet) = q_0 \in \mathrm{H}^2(0,1)$. In the previous formula, $\mathrm{Sign} : \mathbb{R} \rightrightarrows \mathbb{R}$ stands for the multi-valued signum function

$$\mathrm{Sign}(x) := \partial(|\bullet|)(x) = \begin{cases} \{-1\} & \text{if } x < 0, \\ [-1,1] & \text{if } x = 0, \\ \{1\} & \text{if } x > 0. \end{cases}$$

Similarly, solutions to the problem associated with $\mathcal{E}_\infty$ and $\mathcal{D}$ satisfy

$$0 \in \mathrm{Sign}(\partial_t q) - \partial_x(\mathrm{D}W^{**}(\partial_x q)) + 2q - f(t,\bullet) \qquad \text{a.e. in } (t,x) \in [0,T] \times \Omega$$

and this time the equivalence of the energetic formulation with the preceding differential inclusion is even rigorously true by Proposition 2.6.

The $\mathcal{E}_k$ Γ-converge to $\mathcal{E}_\infty$, which is seen as follows: The lim inf-inequality is clear since $(v'')^2/k \geq 0$ and we already saw the Γ-convergence of the functional $\mathcal{E}$ without the k-dependent term to $\mathcal{E}^{**}(= \mathcal{E}_\infty)$ in Section 4.2. For the lim sup-inequality at $t \in [0,T]$ and $u \in \mathrm{H}^1(0,1)$, choose a sequence $(\tilde{v}_j)_j \subseteq \mathrm{H}^1(0,1)$ with $\tilde{v}_j \rightharpoonup u$ in $\mathrm{H}^1(0,1)$ and $\mathcal{E}(t,\tilde{v}_j) \to \mathcal{E}_\infty(t,u)$, which exists since $\mathcal{E}_\infty = \mathcal{E}^{**}$ is the weakly lower semicontinuous envelope of $\mathcal{E}$ (cf. the considerations in Section 4.2). Now, choose a smooth approximation $v_j \in \mathrm{C}^\infty[0,1]$ to $\tilde{v}_j$ with

$$\|v_j - \tilde{v}_j\|_{\mathrm{H}^1(0,1)} \leq \frac{1}{j} \qquad \text{and} \qquad |\mathcal{E}(t,v_j) - \mathcal{E}(t,\tilde{v}_j)| \leq \frac{1}{j}$$

by using the usual smoothing method with a mollifier (and an extension to some interval strictly containing $(0,1)$); the second condition can be fulfilled, because $\mathcal{E}(t,\bullet)$ is strongly continuous. Since $v_j \in \mathrm{C}^\infty[0,1]$, $\|v_j\|_{\mathrm{H}^2(0,1)} < \infty$ and we can inductively find for each j an $m(j) \geq \max\{m(j-1), j\}$ with

$$\|v_j\|_{\mathrm{H}^2(0,1)} \leq m(j)^{1/4} \, \|v_1\|_{\mathrm{H}^2(0,1)} \, .$$

(for $j = 1$ this holds with $m(1) = 1$). Now, we assemble the "spaced" sequence $(u_k)_k$ as follows:

$$(u_k)_k = (\underbrace{v_1, \ldots, v_1}_{m(2)-1 \text{ times}}, \underbrace{v_2, \ldots, v_2}_{m(3)-m(2) \text{ times}}, \ldots, \underbrace{v_i, \ldots, v_i}_{m(i+1)-m(i) \text{ times}}, \ldots)$$

so that $\|u_k\|_{\mathrm{H}^2(0,1)} \leq k^{1/4} \, \|v_1\|_{\mathrm{H}^2(0,1)}$ holds for all $k \in \mathbb{N}$. Since $\|v_j - \tilde{v}_j\|_{\mathrm{H}^1(0,1)} \leq 1/j$ and $\tilde{v}_j \rightharpoonup u$ in $\mathrm{H}^1(0,1)$, also $u_k \rightharpoonup u$ in $\mathrm{H}^1(0,1)$ and by construction

$$\int_0^1 \frac{1}{k}(u_k'')^2 \, \mathrm{d}x \leq \frac{1}{k} \, \|u_k\|_{\mathrm{H}^2(0,1)}^2 \leq \frac{1}{\sqrt{k}} \, \|v_1\|_{\mathrm{H}^2(0,1)}^2 \longrightarrow 0 \qquad \text{as } k \to \infty.$$

Hence, $\mathcal{E}_k(t,u_k) \to \mathcal{E}_\infty(t,u)$ and $(u_k)_k$ is the sought recovery sequence for u.

As before, for adequately chosen loadings f, we have existence and we know that solutions to the incremental problem (AIP) for the kth problem admit a subsequence converging to a solution of (S_∞) & (E_∞). The results of Section 3.2 imply that *every* solution to (S_∞) & (E_∞) can be approximated by solutions to $(\mathrm{AIP}_{k,\varepsilon,\Pi})$.

Symbols

The following table contains some of the used symbols:

$\mathbb{N}$	the natural numbers (without 0)
$\mathbb{R}_\infty$	$= \mathbb{R} \cup \{+\infty\}$
$\mathbb{N}_\infty$	$= \mathbb{N} \cup \{+\infty\}$
Ω	a generic open, bounded domain in $\mathbb{R}^d$ with Lipschitz-boundary
$\overline{\Omega}, \operatorname{cl}\Omega$	the closure of Ω (in the ambient space)
$\Subset$	$E \Subset \Omega$ if and only if $\overline{E} \subseteq \Omega$ and $\overline{E}$ compact ($E, \Omega \subseteq \mathbb{R}^d$ open)
$\operatorname{diam} A$	the diameter of a set, see page 16
$\mathcal{X}[a,b]$	the space of all functions $f : [a,b] \to \mathcal{X}$
$f(a, \bullet)$	a function with an open argument (interprete as $x \mapsto f(a,x)$)
$\operatorname{supp} f$	$= \operatorname{cl}\{\, x \in \Omega \;:\; f(x) \neq 0 \,\}$ for $f : \Omega \to \mathbb{R}$
$\operatorname{lsc} f$	the lower semicontinuous envelope of $f : \mathcal{X} \to \mathbb{R}_\infty$, see page 8
δ_x	$\delta_x(y) = 1$ if and only if $x = y$ (the Kronecker delta)
∂_t, ∂_x	partial t, x-derivatives
∂ (no index)	(convex) subdifferential, see page 24
$\mathrm{D}F$	Gâteaux-derivative of the functional $F : \mathcal{X} \to \mathbb{R}_\infty$, see page 24
$\mathrm{O}(f)$	at most growth of order f (Landau symbol)
$\mathrm{o}(f)$	growth of order less than f (Landau symbol)
$F : \mathcal{X} \rightrightarrows \mathcal{Y}$	a multi-valued map with $F(x) \subseteq \mathcal{Y}$ for all $x \in \mathcal{X}$
Sign	$= \partial(\lvert\bullet\rvert)$ (the multi-valued signum function), see page 69
$\rightharpoonup$	weak convergence
$\overset{*}{\rightharpoonup}$	weak* convergence
$\hookrightarrow$	a continuous embedding
$\overset{c}{\hookrightarrow}$	a compact embedding
$\langle x^*, x \rangle$	the duality product between $x \in \mathcal{X}^*$ and $x \in \mathcal{X}$ ($\mathcal{X}$ a Banach space)
$\lVert \Pi \rVert$	the fineness of the partition Π, see page 31
$\mathrm{C}(\mathcal{X})$	$= \{\, u : \mathcal{X} \to \mathbb{R} \;:\; u \text{ is continuous} \,\}$
$\mathrm{C}^k(\mathcal{X})$	$= \{\, u : \mathcal{X} \to \mathbb{R} \;:\; u \text{ is } k\text{-times continuously differentiable} \,\}$, $k \in \mathbb{N}_\infty$
$\mathrm{C}^k_c(\mathcal{X})$	$= \left\{\, u \in \mathrm{C}^k(\mathcal{X}) \;:\; \operatorname{supp} u \Subset \mathcal{X} \,\right\}$

$\mathrm{C}^k_{(c)}(\mathcal{X};\mathcal{Y})$	like $\mathrm{C}^k_{(c)}(\mathcal{X})$, but functions with values in the space $\mathcal{Y}$
$\mathrm{L}^p(\Omega)$	the space of p-Lebesgue integrable functions modulo equality a.e.
$\mathrm{L}^p(\Omega;\mathcal{Y})$	the space of p-Bochner integrable functions with values in the Banach space $\mathcal{Y}$
$\mathrm{W}^{1,p}(\Omega)$	$=\{\, u \in \mathrm{L}^p(\Omega) \;:\; Du \in \mathrm{L}^p(\Omega;\mathbb{R}^d) \,\}$ where Du is the distributional or generalized derivative (Sobolev space)
$\mathrm{W}^{1,p}(\Omega;\mathcal{Y})$	$=\{\, u \in \mathrm{L}^p(\Omega;\mathcal{Y}) \;:\; Du \in \mathrm{L}^p(\Omega;\mathcal{Y}) \,\}$ (Sobolev–Bochner space)
$\mathrm{L}^p(a,b;\mathcal{Y})$	$=\mathrm{L}^p((a,b);\mathcal{Y})$
$\mathrm{W}^{1,p}(a,b;\mathcal{Y})$	$=\mathrm{W}^{1,p}((a,b);\mathcal{Y})$
$\operatorname{Argmin} F$	the set of minimizers (if any) of $F:\mathcal{X}\to\mathbb{R}_\infty$, see page 12
$\operatorname{Argmin}_\varepsilon F$	the set of ε-minimizers of $F:\mathcal{X}\to\mathbb{R}_\infty$, see page 12
$\Gamma\text{-}\lim_k F_k$	the Γ-limit of the sequence $(F_k:\mathcal{X}\to\mathbb{R}_\infty)_k$, see page 6
$\Gamma\text{-}\liminf_k F_k$	the Γ-limes inferior of the sequence $(F_k:\mathcal{X}\to\mathbb{R}_\infty)_k$, see page 6
$\Gamma\text{-}\limsup_k F_k$	the Γ-limes superior of the sequence $(F_k:\mathcal{X}\to\mathbb{R}_\infty)_k$, see page 6
$\mathrm{K}\text{-}\liminf_k A_k$	the Kuratowski lower limit of the sequence of sets $(A_k)_k$, see page 14
$\mathrm{K}\text{-}\limsup_k A_k$	the Kuratowski upper limit of the sequence of sets $(A_k)_k$, see page 14
$\mathrm{Var}(f;[a,b])$	the pointwise variation of $f:[a,b]\to\mathbb{R}$, see page 37
$\mathrm{Diss}_{\mathcal{D}}(q;[s,t])$	the total dissipation of the process $q\in\mathcal{Q}[s,t]$ with respect to the dissipation distance $\mathcal{D}:\mathcal{Q}\times\mathcal{Q}\to\mathbb{R}_\infty$, see page 26
S_H	the solution functor for the system H, see page 22
INT	the category of all intervals, see page 22
REP	the category of all orientation-preserving C^1-reparametrizations of intervals, see page 23
SOP	the categoty of all solution operators, see page 22
$\bowtie$	the concatenation of adjacent intervals or solution operators, see page 22
α_*	the transformation of an interval or of a solution operator induced by the reparametrization $\alpha\in\mathsf{REP}$, see page 23

Bibliography

[AF90] J. P. Aubin and H. Frankowska, *Set-Valued Analysis*, Systems & Control: Foundations & Applications, vol. 2, Birkhäuser, 1990.

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free-Discontinuity Problems*, Oxford Mathematical Monographs, Oxford University Press, 2000.

[AP02] F. Auricchio and L. Petrini, *Improvements and algorithmical considerations on a recent three-dimensional model describing stress-induced solid phase transformations*, Internat. J. Numer. Methods Engrg. **55** (2002), 1255–1284.

[AS05] F. Auricchio and U. Stefanelli, *Well-posedness and approximation for a one-dimensional model for shape memory alloys*, Math. Models Methods Appl. Sci. **15** (2005), 1301–1327.

[BKS04] M. Brokate, P. Krejčí, and H. Schnabel, *On uniqueness in evolution quasivariational inequalities*, J. Convex Anal. **11** (2004), 111–130.

[BP86] V. Barbu and T. Precupanu, *Convexity and optimization in Banach spaces*, third ed., D. Reidel Publishing Co., Dordrecht, 1986.

[Bra02] A. Braides, *Γ-convergence for Beginners*, Oxford Lecture Series in Mathematics and its Applications, vol. 22, Oxford University Press, 2002.

[CHM02] C. Carstensen, K. Hackl, and A. Mielke, *Non-convex potentials and microstructures in finite-strain plasticity*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **458** (2002), 299–317.

[Con90] J. B. Conway, *A Course in Functional Analysis*, 2. ed., Graduate Texts in Mathematics, vol. 96, Springer, 1990.

[CV90] P. Colli and A. Visintin, *On a class of doubly nonlinear evolution equations*, Comm. Partial Differential Equations **15** (1990), 737–756.

[Dac89] B. Dacorogna, *Direct Methods in the Calculus of Variations*, Applied Mathematical Sciences, vol. 78, Springer-Verlag, 1989.

[Dal93] G. Dal Maso, *An Introduction to Γ-Convergence*, Progress in Nonlinear Differential Equations and Their Applications, vol. 8, Birkhäuser, 1993.

[DFT05] G. Dal Maso, G. A. Francfort, and R. Toader, *Quasistatic crack growth in nonlinear elasticity*, Arch. Ration. Mech. Anal. **176** (2005), 165–225.

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

[FM06] G. A. Francfort and A. Mielke, *Existence results for a class of rate-independent material models with nonconvex elastic energies*, J. Reine Angew. Math. **595** (2006), 55–91.

[Fré02] M. Frémond, *Non-smooth thermomechanics*, Springer, Berlin, 2002.

[GGZ74] H. Gajewski, K. Gröger, and K. Zacharias, *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, Akademie Verlag, 1974.

[GT98] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Grundlehren der mathematischen Wissenschaften, vol. 224, Springer, 1998.

[Hah14] H. Hahn, *Über Annäherungen an Lebesgue'sche Integrale durch Riemann'sche Summen*, Sitzungsber. Math. Phys. Kl. K. Akad. Wiss. Wien **123** (1914), 713–743.

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

[HR99] W. Han and B. D. Reddy, *Plasticity (mathematical theory and numerical analysis)*, Interdisciplinary Applied Mathematics, vol. 9, Springer, 1999.

[Kač85] J. Kačur, *Method of Rothe in Evolution Equations*, Teubner-Texte zur Mathematik, vol. 80, Teubner, 1985.

[KMR05] M. Kružík, A. Mielke, and T. Roubíček, *Modelling of microstructure and its evolution in shape-memory-alloy single-crystals, in particular in CuAlNi*, Meccanica **40** (2005), 389–418.

[KMR06] M. Kočvara, A. Mielke, and T. Roubíček, *A rate-independent approach to the delamination problem*, Math. Mech. Solids **11** (2006), 423–447.

[Kre98] R. Kress, *Numerical Analysis*, Graduate Texts in Mathematics, vol. 181, Springer, 1998.

[KW03] M. Kamlah and Z. Wang, *A thermodynamically and microscopically motivated constitutive model for piezoceramics*, Comput. Materials Science **28** (2003), 409–418.

[Mac98] S. Mac Lane, *Categories for the Working Mathematician*, 2. ed., Graduate Texts in Mathematics, vol. 5, Springer, 1998.

[Mai04] A. Mainik, *A rate-independent model for phase transformations in shape-memory alloys*, Ph.D. thesis, Universität Stuttgart, 2004.

[MB89] G. A. Maugin and E. Bassiouny, *Continuum thermodynamics of electromechanical hysteresis in ceramics*, Continuum mechanics and its applications (Burnaby, BC, 1988), Hemisphere, New York, 1989, pp. 225–235.

[Mie03] A. Mielke, *Evolution of rate-independent inelasticity with microstructure using relaxation and Young measures*, IUTAM Symposium on Computational Mechanics of Solid Materials at Large Strains (Stuttgart, 2001), Solid Mech. Appl., vol. 108, Kluwer Acad. Publ., Dordrecht, 2003, pp. 33–44.

[Mie04] ———, *Deriving new evolution equations for microstructures via relaxation of variational incremental problems*, Comput. Methods Appl. Mech. Engrg. **193** (2004), 5095–5127.

[Mie05] ———, *Evolution of rate-independent systems*, Evolutionary equations. Vol. II, Handb. Differ. Equ., Elsevier/North-Holland, Amsterdam, 2005, pp. 461–559.

[Mie07] ———, *Modelling and analysis of rate-independent processes*, Lipschitz lecture notes, January 2007.

[Mie08] ———, *Weak-convergence methods for Hamiltonian multiscale problems*, Discrete Contin. Dyn. Syst. **20** (2008), 53–79.

[MM05] A. Mainik and A. Mielke, *Existence results for energetic models for rate-independent systems*, Calc. Var. Partial Differential Equations **22** (2005), 73–99.

[MR06] A. Mielke and T. Roubíček, *Numerical approaches to rate-independent processes and applications in inelasticity*, ESAIM Math. Mod. Num. Anal. (2006), Submitted. WIAS Preprint 1169.

[MR07a] A. Mielke and F. Rindler, *Reverse approximation of energetic solutions to rate-independent processes*, submitted, WIAS-Preprint 1271, November 2007.

[MR07b] A. Mielke and R. Rossi, *Existence and uniqueness results for a class of rate-independent hysteresis problems*, Math. Models Methods Appl. Sci. **17** (2007), 81–123.

[MRS08] A. Mielke, T. Roubíček, and U. Stefanelli, Γ*-limits and relaxations for rate-independent evolutionary problems*, Calc. Var. Partial Differential Equations **31** (2008), 387–416.

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

[MT04] ______, *On rate-independent hysteresis models*, NoDEA Nonlinear Differential Equations Appl. **11** (2004), 151–189.

[MT06] A. Mielke and A. M. Timofte, *An energetic material model for time-dependent ferroelectric behaviour: existence and uniqueness*, Math. Methods Appl. Sci. **29** (2006), 1393–1410.

[MTL02] A. Mielke, F. Theil, and V. I. Levitas, *A variational formulation of rate-independent phase transformations using an extremum principle*, Arch. Ration. Mech. Anal. **162** (2002), 137–177.

[Mül93] S. Müller, *Singular perturbations as a selection criterion for periodic minimizing sequences*, Calc. Var. Partial Differential Equations **1** (1993), 169–204.

[Mül99] ______, *Variational models for microstructure and phase transitions*, Calculus of variations and geometric evolution problems (Cetraro, 1996), Lecture Notes in Math., vol. 1713, Springer, Berlin, 1999, pp. 85–210.

[Nat75] I. P. Natanson, *Theorie der Funktionen einer reellen Veränderlichen*, Akademie Verlag, 1975.

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

[QV94] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer Series in Computational Mathematics, vol. 23, Springer, 1994.

[Rin08a] F. Rindler, *Approximation of rate-independent optimal control problems*, in preparation, 2008.

[Rin08b] ______, *Optimal control for nonconvex rate-independent evolution processes*, SIAM J. Control Optim. (to appear) (2008).

[Roc70] R. T. Rockafellar, *Convex Analysis*, Princeton Mathematical Series, vol. 28, Princeton University Press, 1970.

[Rou05] T. Roubíček, *Nonlinear Partial Differential Equations with Applications*, International Series of Numerical Mathematics, vol. 153, Birkhäuser, 2005.

[Sch70] H. Schubert, *Kategorien I*, Heidelberger Taschenbücher, vol. 65, Springer, 1970.

[SS04] E. Sandier and S. Serfaty, *Gamma-convergence of gradient flows with applications to Ginzburg-Landau*, Comm. Pure Appl. Math. **57** (2004), 1627–1672.

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

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

Index

Selbstständigkeitserklärung

Die selbstständige und eigenhändige Anfertigung dieser Arbeit versichere ich an Eides statt.