Optimal control of a rate-independent system constrained to parameterized balanced viscosity solutions

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

Vorgelegt im

Fachbereich 10
der Universität Kassel
Institut für Mathematik

von

STEPHANIE THOMAS

Datum der Disputation: 25.04.2022

Dissertation

Optimal control of a rate-independent system constrained to parameterized balanced viscosity solutions

Fachbereich Mathematik und Naturwissenschaften der Universität Kassel
Institut für Mathematik

Erstgutachterin: Prof. Dr. Dorothee Knees
Zweitgutachter: Prof. Dr. Christian Meyer

Danksagung

Mein besonderer Dank gilt zunächst meiner Betreuerin Prof. Dr. Dorothee Knees, nicht nur für die Möglichkeit, diese Arbeit zu schreiben, sondern auch für die Begleitung auf dem Weg dorthin, sowie in fachlicher Hinsicht, als auch in Form moralischer Unterstüzung, wenn dies nötig war.

Darüber hinaus möchte ich mich bei allen Kolleginnen und Kollegen des Instituts bedanken. Ich schätze mich glücklich, in einer so freundschaftlichen Atmosphäre arbeiten zu dürfen, die sich sowohl in der Zusammenarbeit und der Wertschätzung der Arbeit der anderen, als auch im persönlichen Austausch zeigt. Namentlich geht hier mein Dank an Thomas Izgin für das gründliche Lesen der ersten beiden Kapitel.

Außerhalb des Instituts danke ich meinen Freunden und Familie für ihre Unterstützung, insbesondere meinen Eltern und Geschwistern, die immer für mich mich da sind und waren.

Zusammenfassung

In dieser Dissertation analysieren wir ein Optimalsteuerungsproblem, welches durch ein ratenunabhängiges System gestellt wird. Dabei bewegen wir uns in einem abstrakten, unendlichdimensionalen Szenario. Das ratenunabhängige System wird durch eine nicht-konvexe Energie charakterisiert, die über eine zeitabhängige äußere Last von der Zeit abhängt, sowie durch ein konvexes, beschränktes, und positiv 1-homogenes Dissipationspotential.

Das Optimalsteuerungsproblem wird durch die äußere Last gesteuert und die zulässigen Zustände werden auf die Menge der normalisierten, parameterisierten balanced viscosity Lösungen (BV Lösungen) des ratenunabhängigen System beschränkt. Lösungen dieser Art werden in dieser Arbeit durch viskose Regularisierung des ratenunabhängigen Systems und anschließenden Übergang zum Grenzwert für verschwindende Viskosität erhalten. Da BV Lösungen in der Regel nicht eindeutig sind, ist ein Meilenstein auf dem Weg zur Existenz einer optimalen Steuerung die Kompaktheit der Menge der BV Lösungen.

Abstract

In this dissertation, we analyze an optimal control problem governed by a rate-independent system in an abstract infinite-dimensional setting. The rate-independent system is characterized by a nonconvex stored energy functional, which depends on time via a time-dependent external loading, and by a convex dissipation potential, which is assumed to be bounded and positively homogeneous of degree one.

The optimal control problem uses the external load as control variable and is constrained to normalized parametrized balanced viscosity solutions (BV solutions) of the rate-independent system. Solutions of this type appear as vanishing viscosity limits of viscously regularized versions of the original rate-independent system. Since BV solutions in general are not unique, as a main ingredient for the existence of optimal solutions we prove the compactness of solution sets for BV solutions.

Contents

List of symbols

$\mathcal{Z}$	state space, reflexive Banach space 50
$\mathcal{V}$	reflexive Banach space 50
$\mathcal{X}$	Banach space 50
$\mathcal{Q}$	general state space 18
$\mathcal{E}:[0,T]\times\mathcal{Q}\to\mathbb{R}$	energy functional in dep. of time and state . . 18
$\mathcal{I}:\mathcal{V}^*\times\mathcal{Z}\to\mathbb{R}$	reduced energy functional in dep. of load and state 51
$\mathcal{P}:[0,T]\times\mathcal{Z}\to\mathbb{R}$	derivative of the scalar function $t\mapsto\mathcal{I}(\ell(t),z)$ 71
$\mathcal{F}:\mathcal{Z}\to[0,\infty)$	non-convex part of the energy functional . . . 51
$\mathcal{R}$, $\mathcal{R}^*$	1-hom. dissipation potential and its convex conjugate 51,22
$\mathcal{R}_\varepsilon=\mathcal{R}+\mathcal{R}_{2,\varepsilon}$, $\mathcal{R}^*_\varepsilon$	augmented dissipation potential and its conv. conjugate 51,54
$\mathcal{R}_{2,\varepsilon}$, $\mathcal{R}^*_{2,\varepsilon}$	viscous augmentation of dissipation potential and its conv. conjugate 51,54
$\ell\in\mathcal{V}^*$	external load 51
$\partial\mathcal{R}$, $\partial\mathcal{R}_\varepsilon$	convex subdifferential 51, 111
$\mathfrak{p}:\mathcal{Z}\times\mathcal{V}^*\to[0,\infty)$	vanishing viscosity contact potential 70
$s_\varepsilon:[0,T]\to[0,S_\varepsilon]$	arclength parameter 70
$\hat{t}_\varepsilon=(s_\varepsilon)^{-1}:[0,S_\varepsilon]\to[0,T]$	parameter change 70
$z_\varepsilon:[0,T]\to\mathcal{Z}$	solution of the viscous problem 59
$\hat{z}_\varepsilon=z_\varepsilon\circ\hat{t}_\varepsilon:[0,S]\to\mathcal{Z}$	reparameterized solution 70
$\mathrm{dist}_{\mathcal{V}^*}(\xi,\partial\mathcal{R}(0))$	distance of $\xi\in\mathcal{V}^*$ to $\partial\mathcal{R}(0)$ w.r.t. $\lVert\cdot\rVert_{\mathcal{V}^*}$ 35
$\mathfrak{e}:[0,T]\times\mathcal{Z}\to[0,\infty)$	distance of driving force to $\partial\mathcal{R}(0)$ in dep. of time and state 71
$\mathrm{m}:\mathcal{V}^*\times\mathcal{Z}\to[0,\infty)$	distance of driving force to $\partial\mathcal{R}(0)$ in dep. of load and state 92
$G=\{s\,\vert\,\mathfrak{e}(\hat{t}(s),\hat{z}(s))>0\}$	subset of $[0,S]$ belonging to a $\mathfrak{p}$-parameterized BV solution $(S,\hat{t},\hat{z})$ 73
$\lambda:[0,S]\to[0,\infty)$	Lagrange parameter in diff. charact. of $\mathfrak{p}$-parameterized BV solution 80
$\mathrm{AC}([0,T];\mathcal{V})$	space of $\mathcal{V}$-valued abs. cont. functions 116
$\mathrm{AC}([0,T];\mathcal{R})$	space of $\mathcal{R}$-abs. cont. functions 72,119

Notation

- For a Banach space B and $p \in [1,\infty)$, we denote with $L^p(0,T;B)$ the linear space that consists of the equivalence classes of Bochner integrable functions $u : [0,T] \to B$ with
$$\int_0^T \|u(t)\|_B^p \,\mathrm{d}t < \infty.$$
Here, we consider two functions equivalent if they coincide almost everwhere. We endow $L^p(0,T;B)$ with the norm
$$\|u\|_{L^p(0,T;B)} := \Big(\int_0^T \|u(t)\|_B^p \,\mathrm{d}t\Big)^{\frac{1}{p}}.$$
With $L^\infty(0,T;B)$, we denote the linear space that consists of the equivalence classes of Bochner measurable functions $u : [0,T] \to B$ that are essentially bounded, i.e., those functions for which there exists $M > 0$ such that for almost all $t \in [0,T]$, it holds that $\|u(t)\|_B \leq M$. The infimum of all these bounds is called the essential supremum $\operatorname{ess\,sup}_{t\in[0,T]} \|u(t)\|_B$. We endow $L^\infty(0,T;B)$ with the norm
$$\|u\|_{L^\infty(0,T;B)} := \operatorname*{ess\,sup}_{t\in[0,T]} \|u(t)\|_B.$$

- We denote by $L^1_{\mathrm{loc}}(0,T;B)$ the space of functions $u : [0,T] \to B$ that are Bochner integrable on every compact subset $K \subset (0,T)$. If $u \in L^1_{\mathrm{loc}}(0,T;B)$, and $v \in L^1_{\mathrm{loc}}(0,T;B)$ is another function such that
$$\text{for all } \phi \in C_0^\infty(0,T) : \int_0^T u(t)\phi'(t)\,\mathrm{d}t = -\int_0^T v(t)\phi(t)\,\mathrm{d}t,$$
then we call v the generalized derivative of u and denote it by $v = \dot{u}$.

- For $p \in [1,\infty]$, we denote with
$$W^{1,p}(0,T;B) := \{u \in L^p(0,T;B) \,|\, \dot{u} \in L^p(0,T;B)\}$$
the space of $L^p(0,T;B)$-functions whose generalized derivative is an element of $L^p(0,T;B)$ as well. We endow $W^{1,p}(0,T;B)$ with the norm
$$\|u\|_{W^{1,p}(0,T;B)} := \|u\|_{L^p(0,T;B)} + \|\dot{u}\|_{L^p(0,T;B)}.$$
For all $p \in [1,\infty]$, the spaces $L^p(0,T;B)$ and $W^{1,p}(0,T;B)$ with their corresponding norms are Banach spaces, see, e.g., [Emm04, Satz 7.1.23, Satz 8.1.6].

- For an interval $I \subseteq \mathbb{R}$ and a Banach space B, we denote with $C(I;B)$ the space of functions $f : I \to B$ that are continuous w.r.t. the norm on B and endow it with the supremum-norm. With $C_{\mathrm{weak}}(I,B)$, we denote the space of functions $f : I \to B$ that are continuous w.r.t. the weak topology on B.

- The symbols C, c are used to denote real, positive constants in estimates. Their value may change from line to line. The notation $C(f_1,\dots,f_n)$ or $c(f_1,\dots,f_n)$ is used to symbolize that the value C or c, respectively, depends only on the entities $f_1,\dots,f_n$.
- The functions $\mathcal{F} : \mathcal{Z} \to [0,\infty)$ defined on p. 51 and $\mathcal{J} : \mathcal{Z} \to \mathbb{R}$ defined on p. 89 depend solely on the state variable $z \in \mathcal{Z}$. Therefore, we denote their Fréchet-derivatives w.r.t. z by $\mathrm{D}\mathcal{F} := \mathrm{D}_z\mathcal{F}$ and $\mathrm{D}\mathcal{J} := \mathrm{D}_z\mathcal{J}$, respectively.
- For a given Banach space B, we denote with $\langle \sigma, u\rangle_B$ the duality pairing of an element $\sigma \in B^*$ with an element $u \in B$, i.e., $\langle \sigma, u\rangle_B := \sigma(u) \in \mathbb{R}$. If no confusion on the ambient space can arise, we may omit the index B.
- $\exp : \mathbb{R} \to \mathbb{R}$ denotes the exponential function.

Chapter 1

Introduction

The aim of this dissertation is the analysis and optimal control of rate-independent systems with non-convex energy. Models of this type arise in a variety of applications, such as for example the brittle damage of a workpiece under the influence of external loads, or in elastoplasticity, see, e.g., [DH08] or [MM09]. The common feature of these dissipative systems is the following: If we apply a time rescaling on the external loads, the solutions of the rescaled system are exactly those obtained by applying the same rescaling to the solutions of the original system. In other words, the internal processes depend on the direction of the rate, but not on its magnitude, which is why these systems are called rate-independent. In the modelling of rate-independent systems, this requirement translates to the condition that the dissipation potential must be positively homogeneous of degree 1. This, however, implies its non-smoothness.

The interplay of a non-smooth dissipation and a non-convex energy causes significant analytical complications: While there exists a variety of different solution concepts, in none of these are the solutions in general unique. What is more, the solutions feature a significant lack of smoothness, up to jumps in time, even if the external loadings are completely smooth. This, of course, poses serious challenges when it comes to the optimal control of these systems.

Let us now go into more detail concerning the modelling of rate-independent systems. On a bounded time interval $[0,T]$ for an end time $T > 0$, we will consider the subdifferential inclusion

$$0 \in \partial\mathcal{R}(\dot{z}(t)) + \mathrm{D}_z\mathcal{J}(t,z(t)) \subset \mathcal{Z}^*, \quad t \in [0,T]; \quad z(0) = z_0. \tag{RIS}$$

Here, $z : [0,T] \to \mathcal{Z}$ constitutes the state variable living in an infinite dimensional state space $\mathcal{Z}$, $\mathcal{J}(\cdot,\cdot) : [0,T]\times\mathcal{Z} \to \mathbb{R}$ denotes the stored energy energy functional, and $\mathcal{R} : \mathcal{Z} \to [0,\infty)$ is the dissipation potential. In order to obtain a rate-independent system, we have to assume that $\mathcal{R}$ is convex and positively 1-homogeneous, i.e., for all $z \in \mathcal{Z}$ and $\alpha \in \mathbb{R}\setminus\{0\}$ we have $\mathcal{R}(\alpha z) = |\alpha|\mathcal{R}(z)$. As mentioned before, this implies that $\mathcal{R}$ is not differentiable, however, it is possible to show that its convex subdifferential $\partial\mathcal{R}(z) \subset \mathcal{Z}^*$ is non-empty. Throughout this work, we assume that for every $t \in [0,T]$, $\mathcal{J}(t,\cdot) : \mathcal{Z} \to \mathbb{R}$ is Fréchet-differentiable and denote by $\mathrm{D}_z\mathcal{J}(t,\cdot) : \mathcal{Z} \to \mathcal{Z}^*$ its Fréchet-derivative.

Given an initial state $z_0 \in \mathcal{Z}$, if we aim to find a curve $z : [0,T] \to \mathcal{Z}$ such that (RIS) is satisfied almost everywhere, the minimal regularity required for z turns out to be $z \in W^{1,1}(0,T;\mathcal{Z})$. Existence results for such curves, called **differential solutions**, are based on strong assumptions on smoothness and convexity of the energy. The now classical weaker concept of **global energetic solutions (GES)** was first introduced in a series of papers by A. Mielke, F. Theil, and V. I. Levitas (see [MT99], [MTL02], [MT04]) and instead relies on a

global stability condition

$$\forall t \in [0,T], \forall \hat{z} \in \mathcal{Z}: \quad \mathcal{J}(t,z(t)) \leq \mathcal{J}(t,\hat{z}) + \mathcal{R}(\hat{z}-z(t)) \tag{S}$$

and an energy balance

$$\forall t \in [0,T]: \mathcal{J}(t,z(t)) + \mathrm{Var}_{\mathcal{R}}(z,[0,T]) = \mathcal{J}(0,z_0) + \int_0^t \partial_s \mathcal{J}(s,z(s))\,\mathrm{d}s. \tag{E}$$

This formulation brings the advantage that it requires significantly less structure of the state space $\mathcal{Z}$ and no differentiability of the energy $\mathcal{J}(t,\cdot): \mathcal{Z} \to \mathbb{R}$, and existence of GES can be shown for non-convex energies. What is more, it lends itself to numerical approximation via a time discretization scheme. Many results on GES have been collected in [MR15], where numerous applications can be found as well.

However, the global nature of the stability condition (S) causes GES to tend to develop discontinuities that could be considered unphysical. This is due to the fact that the state might jump to any $\hat{z} \in \mathcal{Z}$ with the property that the dissipated energy $\mathcal{R}(\hat{z}-z(t))$ does not exceed the release of energy $\mathcal{J}(t,\hat{z}) - \mathcal{J}(t,z(t))$, even if this jump occurs over an energy barrier. In the quest for a weak solution concept that retains some of the advantages of GES, but allows for a better understanding of the non-smoothness, the authors of [EM06] proposed to start instead from a viscously regularized system. Namely, let us assume that $\mathcal{Z}$ is compactly embedded into a second space $\mathcal{V}$, and for $\varepsilon > 0$, we replace the dissipation potential $\mathcal{R}$ by an augmented dissipation potential

$$\mathcal{R}_\varepsilon(z) := \mathcal{R}(z) + \mathcal{R}_{2,\varepsilon}(z) = \mathcal{R}(z) + \tfrac{\varepsilon}{2}\|z\|_{\mathcal{V}}^2.$$

It is then possible to show that for every $\varepsilon > 0$, the resulting rate-dependent system

$$0 \in \partial\mathcal{R}_\varepsilon(\dot{z}_\varepsilon(t)) + \mathrm{D}_z\mathcal{J}(t,z_\varepsilon(t)), \quad t \in [0,T] \tag{1.0.1}$$

has a unique solution z_ε, fulfilling (1.0.1) pointwisely almost everywhere. In order to return to the original system, we would like to pass to the limit $\varepsilon \to 0$. This is done relying on a reparameterization $\hat{z}_\varepsilon := z_\varepsilon \circ \hat{t}_\varepsilon : [0,S] \to \mathcal{Z}$ for each viscous solution z_ε. It is then possible to show that we obtain a limiting pair $(\hat{t}_\varepsilon, \hat{z}_\varepsilon) \xrightarrow{\varepsilon \to 0} (\hat{t}, \hat{z})$.

The exact characterization and properties of the resulting notion of solutions depend on the choice of the artificial time parameter $\hat{t}_\varepsilon$. In this work, we choose to parameterize by means of the **vanishing viscosity contact potential**, given by

$$\mathfrak{p}(v,\xi) := \mathcal{R}(v) + \|v\|_{\mathcal{V}}\,\mathrm{dist}_{\mathcal{V}^*}(\xi, \partial\mathcal{R}(0)) \quad \text{for } v \in \mathcal{V},\ \xi \in \mathcal{V}^*,$$

which was introduced in [MRS12a]. The parameterization is then defined via

$$s_\varepsilon(t) := t + \int_0^t \mathfrak{p}(\dot{z}_\varepsilon(\tau), -\mathrm{D}_z\mathcal{J}(\tau, z_\varepsilon(\tau)))\mathrm{d}\tau \text{ and } \hat{t}_\varepsilon := (s_\varepsilon)^{-1} : [0, S_\varepsilon] \to [0, T],$$

yielding what is known as **$\mathfrak{p}$-parameterized BV solutions** after passage to the vanishing viscosity limit.

One advantage of $\mathfrak{p}$-parameterized BV solutions is the fact that they provide an enhanced resolution of the jumps. Namely, it can be shown that the pair $(\hat{t}, \hat{z})$ satisfies the differential inclusion

$$0 \in \partial\mathcal{R}(\dot{\hat{z}}(s)) + \lambda(s)\partial\mathcal{R}_2(\dot{\hat{z}}(s)) + \mathrm{D}_z\mathcal{J}(\hat{t}(s), \hat{z}(s)) \quad \text{f.a.a. } s \in [0, S] \tag{1.0.2}$$

for a measurable function $\lambda : [0, S] \to [0, \infty)$ such that $\lambda(s)\dot{\hat{t}}(s) = 0$ a.e. in $[0, S]$. Note that $\hat{t} : [0, S] \to [0, T]$ encodes the external time scale. Wherever $\dot{\hat{t}}(s) > 0$ on an interval $(s_1, s_2) \subset [0, S]$, it follows that $\lambda(s) = 0$, and thus, (1.0.2) describes the original rate-independent evolution. On the other hand, whenever $\dot{\hat{t}}(s) = 0$, the external time is frozen, and we obtain from the vanishing viscosity analysis that at the same time, $\dot{\hat{z}}(s) > 0$, so that this is seen as a jump in the external time frame. If $\lambda(s) > 0$, a viscous dissipation is active, which allows for the interpretation of jumps in the rate-independent system as a transition between two end points along a curve following a viscous regime.

The aim of this work is to solve an **optimal control problem governed by** (RIS) **and restricted to $\mathfrak{p}$-parameterized BV solutions**. To be more precise, we will show the existence of a globally optimal solution of an optimal control system of the type

$$\left.\begin{array}{ll} \min & \|\hat{z} - z_{\mathrm{des}}\|_{\mathcal{Z}} + \alpha\|\ell\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} \\ \text{s.t.} & (S, \hat{t}, \hat{z}) \in \widetilde{M}_{ad}, \end{array}\right\} \tag{1.0.3}$$

where the external load ℓ is the control variable, $\alpha > 0$ is a fixed Tikhonov parameter, and z_{des} is a desired state. We restrain the problem to the admissible set $\widetilde{M}_{ad}$ consisting of all $\mathfrak{p}$-parameterized BV solutions to the system (RIS). For this, it is necessary not only to carefully determine the limit equations that characterize $\mathfrak{p}$-parameterized BV solutions, but also to prove compactness of the resulting solution set. The major challenge in this context is to derive an a priori estimate for the driving forces $\mathrm{D}_z\mathcal{J}(\hat{t}(s), \hat{z}(s))$. This estimate will be obtained by another reparameterization argument starting from the subdifferential inclusion (1.0.2). Namely, we will reparameterize (1.0.2) in such a way that the transformed function $\tilde{z}$ satisfies (1.0.1) with $\varepsilon = 1$ and with a constant external load. We are then in a position to apply the a priori estimates previously derived for (1.0.1) to $\tilde{z}$ and finally transfer them back to $\mathfrak{p}$-parameterized BV solutions.

The main results of this dissertation are also the content of [KT18], to which the author made significant and essential contributions. The author would also like to acknowlegde the funding by Deutsche Forschungsgemeinschaft (DFG) through the Priority Programme SPP 1962 *Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization* within Project P09 *Optimal Control of Dissipative Solids: Viscosity Limits and Non-Smooth Algorithms*.

1.1 A glance into the literature on the optimal control of rate-independent systems

Let us now take a glance into the existing literature on the optimal control of rate-independent systems in order to put this work into perspective. For this purpose, we write the optimal control problem (1.0.3) in the more general form

$$\left.\begin{array}{ll}\min & J(z,\ell) := \|z - z_{\text{des}}\| + \alpha\|\ell\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} \\ \text{s.t.} & z \in M_{ad},\end{array}\right\} \tag{1.1.1}$$

where, again, the external load ℓ is the control variable, $\alpha > 0$ is a fixed Tikhonov parameter, and z_{des} is a desired state. In this section however, with a slight abuse of notation, the admissible set M_{ad} consists of all solutions to the system (RIS) in the sense of GES or in the sense of parameterized BV solutions.

In the case that (RIS) has a unique solution, it is possible to define a single-valued solution operator G mapping the control variable ℓ to the corresponding solution z. The operator G is also called the control-to-state map. In this case, (1.1.1) can be formulated as

$$\left.\begin{array}{ll}\min & J(G(\ell),\ell) = \|G(\ell) - z_{\text{des}}\| + \alpha\|\ell\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} \\ \text{s.t.} & \ell \in \mathcal{D},\end{array}\right\} \tag{1.1.2}$$

where $\mathcal{D}$ is the domain of G and the objective function J now depends on the control variable alone. The problem (1.1.2) is sometimes also referred to as the **reduced problem** and can be tackled using what is known as the **implicit programming approach**. In this approach, the control-to-state map G is studied in great detail, in order to be able to apply the implicit function theorem to the map $\ell \mapsto J(G(\ell),\ell)$. The aim is to compute a gradient, or subgradient of some kind, of the objective function. The necessary uniqueness of solutions can be obtained by requiring uniform convexity of the energy functional $\mathcal{E}$, which in turn also implies that all types of solution are practically identical. A classical example of a rate-independent problem with convex energy is the movement of a coin that is trapped underneath a bowl, as it is described in Section 2.1. This example was introduced as **sweeping process** by Jean-Jacques Moreau in 1973 in [Mor73].

Some early research on the optimal control of **convex problems in the scalar valued case** can be found in [Bro87] in the context of hysterises operators. Here, optimality conditions were derived by means of a time discretization approach and a smoothing process. For the treatment of the multi-, but finite dimensional case, the authors of [AC18] and [CP16] used a Moreau-Yosida approximation in order to obtain optimality conditions, whereas [CHHM15] and [CHHM16] use a time discretization approach. A completely different strategy was followed in [BK15], where instead, the directional differentiability of the hysteresis operators is established.

The **convex, but infinite dimensional problem** is treated in the series [Wac12, Wac15, Wac16] for a quasi-static elasto-plastic model, and existence as well as first order optimality conditions are derived via time discretization and subsequent smoothing of the discrete problems. Alternatively, the optimal control of a reduced problem of the form (2.2.12) is treated in [SWW17] by means of a viscous approximation similar to (2.2.13) and smoothing of the dissipation potential, allowing for the establishment of optimality conditions.

When it comes to the non-convex problem, however, not much literature is available. Since we cannot expect unique solutions, it is no longer possible to consider the reduced problem (1.1.2), rendering the implicit programming approach unfeasible. In fact, this turns (1.1.1) into an optimization problem in a function space, rather than an optimal control problem. What is more, in the non-convex case, the different notions of solutions are in general distinct.

There exists some literature on the optimal control of **non-convex rate-independent systems restricted to GES**, which relies on time discretization or regularization approaches. For example, in [Rin08], the existence of a global minimizer of an optimal control problem of the type (1.1.1) is shown by a combination of the direct method of the calculus of variations with Γ-convergence arguments. It is now a natural question to ask whether such a global minimizer of the continuous problem can be obtained as the limit of a sequence of minimizers of the discretized optimization problems. One crucial stepstone to achieve an answer to this question is the construction of recovery sequences for the feasible solutions of the continuous optimization problem, which is often referred to as **reverse approximation**.

The authors of the paper [MR09] showed that under suitable assumptions, reverse approximation of GES of a possibly infinite-dimensional and non-convex rate-independent system by solutions of ε-approximate incremental problems is possible. This ε-approximate incremental scheme was then used in [Rin09] in order to show that global minimizers of optimal control problems of the type (1.1.1) restricted to GES can be approximated by solutions of discretized optimal control problems. Instead of a time discretization, the authors of [MW20] opted for a Yosida regularization of an optimal control problem governed by the equations of quasi-static perfect plasticity at small strain. Under additional assumptions on the smoothness of at least one global minimizer of the unregularized problem, they are then able to show that minimizers of the regularized problem converge to a global minimizer of the unregularized problem. The papers [ELS13, EL14, Ste11] contain existence results for optimal control problems modelling shape-memory alloys and with respect to GES.

If we turn to the **non-convex optimal control problem with parameterized BV solution as the underlying notion of solutions** however, the literature becomes even more scant. In the recent work [KMS21], the authors showed that in a finite dimensional setting, global minimizers of a viscously regularized optimal control problem converge to global minimizers of an optimal control problem governed by a rate-independent system and constricted to parameterized BV solutions, again provided that at least one global minimizer of the limiting problem has additional regularity.

Let us point out that the reverse approximation property is not necessary in order to prove the existence of a global minimizer of (1.1.1), and we will not follow this approach in this work. Instead, we will show the necessary compactness properties of the admissible set in Theorem 4.3.1.

1.2 Outline

The dissertation is structured as follows: In Chapter 2, we give a brief overview on rate-independent systems and present different solution concepts. Their properties and the differences between them are then illustrated by means of a 1-dimensional example. Finally, we establish the standing assumptions on the state space, energy and dissipation functional, as well as the choice of regularization.

In Chapter 3, we study the viscously regularized system and show existence, uniqueness, and regularity of solutions of the viscous system. A special effort is made to improve a priori estimates that already exist in the literature to a degree that allows to later prove compactness of the set of $\mathfrak{p}$-parameterized BV solutions. Subsequently, we carry out the vanishing viscosity analysis. To this end, we define the reparameterization that is appropriate for our needs, and identify the resulting limiting equation. We then show convergence of the reparameterized solutions to $\mathfrak{p}$-parameterized BV solutions. We conclude Chapter 3 by giving equivalent characterizations for $\mathfrak{p}$-parameterized BV solutions.

What follows in Chapter 4 are the crucial a priori estimates for $\mathfrak{p}$-parameterized BV solutions, resulting in the proof of compactness of the solution set.

With these results, we are able to show existence of a solution to the optimal control problem (1.0.3) in Chapter 5.

The main results of this dissertation are summarized in Chapter 6, where we also discuss open questions and possible further research.

In Appendix A, we collect mainly already existing results from convex analysis that are required for the understanding of this work, and in Appendix B, we list some lower semicontinuity results that are used throughout.

The other appendices include arguments that are crucial for the establishment of the main results of this dissertation, but have been moved to the end for the sake of readability. Appendix C contains an overview on Banach space-valued absolutely continuous functions and their properties depending on the properties of the Banach space. We also give some results on the connection to functions of bounded pointwise total variation. The results of this appendix help motivate the definition of the limiting equation for the $\mathfrak{p}$-parameterized BV solutions.

In Appendix D, we prove additional properties of the energy functional and dissipation potential that follow from the assumptions in Section 2.4.1, but are not obvious. We further dedicate Appendix E specifically to proving the convergence of the terms that contain the external loadings. In Appendix F, we establish chain rules for the scalar function $t \mapsto \mathcal{J}(t, z(t))$ under different assumptions on z.

Finally, in Appendix G, we give an explanation for the assumption that the intermediate space $\mathcal{V}$ is uniformly convex with modulus of convexity of power type 2, and sufficient conditions to satisfy this assumption.

Chapter 2

Rate-independent systems and their solutions

2.1 Rate-independent systems

We begin this section by considering one of the simplest examples of rate-independent evolution, which can be found, e.g., in [MR15, Chapter 1.1]. Imagine a coin on a horizontal table with a large glass bowl, which has been turned upside down, on top of it. In the starting position, the center of the bowl is aligned with the center of the coin. If we start to slowly move the bowl, then the coin will not move until the rim of the bowl touches the coin. At this point, the coin will start moving at the same speed as the bowl, in a direction that is perpendicular to the rim of the bowl. As soon as the rim of the bowl stops touching the coin, the coin will stop moving, until both touch again.

If we imagine all movements to be slow enough that inertia can be neglected, this experiment exhibits three properties that are characteristic for rate-independent systems:

- The output (the position of the coin) is driven by an input function (the position of the bowl).
- As soon as the input is constant (the bowl stops moving), the output is constant (the coin stops moving).
- The system has no intrinsic time scale, meaning that a rescaling of the input (change in velocity of the movement of the bowl) results in a corresponding rescaling of the output (i.e., the coin moves at the corresponding velocity, but along the same path).

This experiment can be modelled in the following way: We define an input function $\ell : [0,T] \to \mathbb{R}^2$ to denote the displacement of the center of the bowl, and an output function $q : [0,T] \to \mathbb{R}^2$ to denote the displacement of the center of the coin. We further denote the radius of the bowl by $R > 0$, and that of the coin by $r > 0$. If we assume the centers of the bowl and coin to be aligned at the starting time $t = 0$, i.e., $q(0) = \ell(0) = (0,0)$, then the condition that the coin must remain under the bowl translates into requiring that $|q(t)-\ell(t)| \leq R-r$ (where $|.|$ denotes the euclidian distance in $\mathbb{R}^2$). We also know that $\dot{q}(t)=0$ as long as $|q(t)-\ell(t)| < R-r$, and that, when the coin moves, it does so perpendicularly to the rim of the bowl, i.e.,

$$\text{if } |q(t)-\ell(t)| = R-r, \text{ then } \dot{q}(t) = \lambda(t)(\ell(t)-q(t)) \text{ for some } \lambda(t) > 0.$$

We can determine $\lambda(t) = \dot{\ell}(t)\Big(\ell(t)-q(t)\Big)/(R-r)^2$ by differentiating the constraint on the left hand side and plugging the equation for $\dot{q}$ into the result.

All in all, we arrive at the equation

$$\dot{q}(t) = \begin{cases} 0, & \text{if } |q(t)-\ell(t)| < R-r, \\ \frac{\dot{\ell}(t)(\ell(t)-q(t))}{(R-r)^2}(\ell(t)-q(t)), & \text{else}, \end{cases} \tag{2.1.1}$$

with the initial value $q(0)=(0,0)$. In order to arrive at a more general formulation, we now define $\mathcal{R}:\mathbb{R}^2 \to [0,\infty)$ by $\mathcal{R}(q) := |q|$, and $\mathcal{E}:[0,T]\times\mathbb{R}^2 \to \mathbb{R}$ by $\mathcal{E}(t,q) := \frac{1}{2(R-r)}|q-\ell(t)|^2$. The function $\mathcal{R}$ is 1-homogeneous and therefore not differentiable in $(0,0)$, however, its convex subdifferential (cf. App. A) is given by

$$\partial\mathcal{R}(v) = \begin{cases} \frac{q}{|q|}, & \text{if } q \neq 0, \\ \{\xi \in \mathbb{R}^2 \,|\, |\xi| \leq 1\}, & \text{if } q = 0. \end{cases}$$

The function $\mathcal{E}(t,\cdot):\mathbb{R}^2 \to \mathbb{R}$ is Fréchet differentiable for every $t \in [0,T]$ with derivative $\mathrm{D}_q\mathcal{E}(t,q) = (q-\ell(t))/(R-r)$, and thus, we can reformulate (2.1.1) as the subdifferential inclusion

$$0 \in \partial\mathcal{R}(\dot{q}(t)) + D_q\mathcal{E}(t,q(t)); \quad q(0)=(0,0). \tag{2.1.2}$$

This is a prototype for rate-independent systems that are obtained from differential formulations. While rate-independence can be defined in the context of input-output systems and then applied to a wider variety of problems, for example in control theory, or approached via hysteresis operators, we will not pursue that path and instead refer to [MR15, Chapter 1.2] for that matter.

In this work, if we speak of rate-independent systems, we have the following in mind: A rate-independent system is a triple $(\mathcal{Q},\mathcal{R},\mathcal{E})$ consisting of

- a Banach space $\mathcal{Q}$,
- a dissipation potential $\mathcal{R}(\cdot,\cdot):\mathcal{Q}\times\mathcal{Q} \to [0,\infty]$,
- and an energy functional $\mathcal{E}:[0,T]\times\mathcal{Q} \to \mathbb{R}$,

where we have to require the following from the dissipation potential $\mathcal{R}$ for all $q \in \mathcal{Q}$:

(i) $\mathcal{R}(q,\cdot):\mathcal{Q} \to [0,\infty]$ is convex and lower semicontinuous, and $\mathcal{R}(q,0)=0$

(ii) $\mathcal{R}(q,\cdot)$ is positively 1-homogeneous, i.e., for all $v \in \mathcal{Q}$ and $\lambda > 0$, we have $\mathcal{R}(q,\lambda v) = \lambda \mathcal{R}(q,v)$

We further denote by $\partial_v \mathcal{R}(q,v)$ the partial convex subdifferential, i.e.,

$$\partial_v \mathcal{R}(q,v) := \{\xi \in \mathcal{Q}^* \,|\, \forall w \in \mathcal{Q} : \mathcal{R}(q,w) \geq \mathcal{R}(q,v) + \langle \xi, w - v \rangle\}, \tag{2.1.3}$$

where $\mathcal{Q}^*$ is the dual space of $\mathcal{Q}$, and $\langle \cdot,\cdot \rangle : \mathcal{Q}^* \times \mathcal{Q} \to \mathbb{R}$ denotes the duality pairing. (See Appendix A for more details on convex analysis.) We then seek to solve the inclusion

$$\text{f.a.a. } t \in [0,T] : \quad \partial_{\dot{q}} \mathcal{R}(q(t),\dot{q}(t)) + D_q \mathcal{E}(t,q(t)) \ni 0; \quad q(t_1) = q_1. \tag{RIS}$$

Let us first note that the second requirement on $\mathcal{R}$ is made in order to obtain rate-independence: If we apply a time-rescaling $\alpha : [0,T] \to [0,S]$ to a differentiable solution q, then the rescaled solution $\widetilde{q} := q \circ \alpha$ has the derivative $\dot{\widetilde{q}}(t) = \dot{\alpha}(t)\dot{q}(\alpha(t))$. Therefore, $\widetilde{q}$ is a solution of the rescaled RIS $(\mathcal{Q},\mathcal{R},\mathcal{E}(\alpha(\cdot),\cdot))$ if and only if the subdifferential $\partial_v \mathcal{R}(q,v)$ is 0-homogeneous, that is, if for all $\lambda > 0$, it holds that $\partial_v \mathcal{R}(q,\lambda v) = \partial_v \mathcal{R}(q,v)$. It follows immediately from the definition of the subdifferential that this is equivalent to $\mathcal{R}(q,\cdot)$ being 1-homogeneous.

The second remark concerns the fact that the choice of the state space $\mathcal{Q}$ and consequently its dual space $\mathcal{Q}^*$ determines with respect to which duality the subdifferential (2.1.3) is defined, and ultimately in which space the inclusion (RIS) is considered. This is of particular importance when it comes to solutions obtained by viscosity approximations, where it is usually assumed that the state space is embedded into a second space in which the viscous regularization takes place, see Section 2.2.4. In the following, we will sometimes simply refer to the inclusion (RIS) itself as a rate-independent system, when no confusion on the ambient space $\mathcal{Q}$ can arise.

Thirdly, as we shall see in Section 2.2.2, while we always have the pointwise inclusion (RIS) in mind when speaking of rate-independent systems, actually finding an a.e. pointwise solution is in general not possible and not the goal of this work. Instead, we will concern ourselves with the weaker concept of $\mathfrak{p}$-parameterized BV solutions.

Finally, we note that for the more general considerations that follow, we will allow for the dissipation potential to be state-dependent, that is, to depend on q and not only on $\dot{q}$. However, from Section 2.2.3 onward, we will drop the state-dependence, allowing for significant simplifications.

2.2 Different solution concepts and how they are related

In this section, we provide several distinct concepts for solutions of (RIS), which might coincide or imply each other, depending on the properties of $\mathcal{E}$ and $\mathcal{R}$. The most straight-forward one might be that of **differential solutions** (see Def. 2.2.3), yet it turns out to be too restrictive for most applications, since it requires differentiability of $\mathcal{E}$ and of the solution q. As already mentioned in Section 2.1, in the case of non-convex energies, one cannot expect differentiability of the solution, and it might even have jumps, even if the external loadings are completely smooth.

However, there is a solution concept that is almost completely derivative-free, namely that of **global energetic solutions** (see Def. 2.2.4). This concept comes with several advantages: Since only the derivative of the scalar function $t \mapsto \mathcal{E}(t,q)$ is needed for its definition, it is possible to consider evolutionary systems on a general topological space $\mathcal{Q}$ that is not equipped with a linear structure. What is more, energetic solutions lend themselves to a construction via an **incremental minimization scheme** (see (IMP) on p. 25), which allows for numerical approximation as well as an existence theory based on the direct method of the calculus of variations. Still, energetic solutions come with a major drawback: Since they are based on a **global stability condition** (S), they have a tendency to jump over energy barriers as early as possible, leading to discontinuities that one might consider physically unplausible.

One way to remedy this is to consider instead a viscously regularized system, thereby localizing the stability condition. While we refer to section 2.4.1 for the exact definition of the envolved quantities, for the moment, we consider instead of (RIS) the subdifferential inclusion

$$\text{f.a.a. } t \in [0,T]: \quad \partial\mathcal{R}_\varepsilon(\dot{q}) + D_q\mathcal{E}(t,q) \ni 0; \quad q(t_1) = q_1. \tag{2.2.1}$$

Here, we drop the state dependence of the dissipation potential $\mathcal{R}$ for simplicity, and define $\mathcal{R}_\varepsilon(\dot{q}) := \mathcal{R}(\dot{q}) + \frac{\varepsilon}{2}\|\dot{q}\|$ for a suitable norm $\|\cdot\|$ on $\mathcal{Q}$ and an artificial viscosity parameter $\varepsilon > 0$. Then, (2.2.1) has solutions with higher regularity than those of the original system (RIS). Passing to the limit $\varepsilon \to 0$, we formally arrive back at (RIS), yet, it is necessary to adjust the limit equations. However, we will not pass to the limit directly, but instead reparameterize the viscous solutions by their dissipation arc-length, see (3.2.1) - (3.2.4). In this way, we increase the resolution of temporal jumps, since we obtain in the limit a curve in $[0,T] \times \mathcal{Z}$, connecting the end points of jumps. These curves are then called **parameterized Balanced Viscosity (BV) solutions**, see Def. 3.2.5. There is an intricate connection between the choice of the norm in the definition of $\mathcal{R}_\varepsilon$, the choice of reparameterization of the viscous solution, and the resulting limit equations. For this reason, we will in this chapter only give a brief insight into the train of thought behind parameterized BV solutions and give the exact definitions and results in Chapter 3.2. We will further in Section 2.3 consider in detail an example for an RIS in a non-convex, 1-dimensional setting that highlights the differences between the notions of solutions presented in this section.

Before we go into more detail about these solution concepts, we will give several equivalent formulations of (RIS). Note that some of the following considerations are valid also for dissipation potentials which are not 1-homogeneous. In order to keep the following discussion as broad as possible, we will start with rather weak assumptions on the energy-dissipation framework, and strengthen these assumptions gradually, where it becomes necessary or convenient.

2.2.1 Equivalent formulations of (RIS)

Let us take the subdifferential inclusion (RIS) as a starting point. In a first step, we use mainly results from convex analysis in order to find equivalent formulations of (RIS). The necessary tools for this are collected in Appendix A. One of these tools is the Legendre-Fenchel transform for convex potentials, which is defined as $\mathcal{R}^*(q,\eta) := [\mathcal{R}(q,\cdot)]^*(\eta) = \sup_{v\in\mathcal{Q}}\langle \eta, v\rangle - \mathcal{R}(q,v)$. Now, using the Fenchel equivalences (A.2), we find that for $q \in W^{1,1}(0,T;\mathcal{Q})$, the following three are equivalent for all $\mathcal{R}$ that are convex and lower semicontinuous:

Force balance

$$0 \in \partial_{\dot{q}}\mathcal{R}(q,\dot{q}) + D_q\mathcal{E}(t,q) \subset \mathcal{Q}^* \tag{2.2.2a}$$

Rate equation

$$\dot{q}(t) \in \partial_\xi \mathcal{R}^*(q(t), -D_q\mathcal{E}(t,q(t))) \subset \mathcal{Q} \tag{2.2.2b}$$

Power balance

$$\mathcal{R}(q(t),\dot{q}(t)) + \mathcal{R}^*(q(t),-D_q\mathcal{E}(t,q(t))) = \langle -D_q\mathcal{E}(t,q(t)),\dot{q}(t)\rangle \in \mathbb{R}. \tag{2.2.2c}$$

In the case of a quadratic dissipation potential, (2.2.2b) becomes a **viscous gradient flow**, which is why (2.2.2b) is sometimes called a **generalized gradient flow**, if $\mathcal{R}$ is not quadratic. Now, if the scalar map $t \mapsto \mathcal{E}(t,q(t))$ is regular enough to fulfill the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t,q(t)) = \langle D_q\mathcal{E}(t,q(t)),\dot{q}(t)\rangle + \partial_t\mathcal{E}(t,q(t)), \tag{2.2.3}$$

then via integration over time, the power balance (2.2.2c) becomes the **energy dissipation balance**

$$\begin{aligned}\forall\, 0 \le r < t \le T: \quad \mathcal{E}(t,q(t)) + \int_r^t \mathcal{R}(q(s),\dot{q}(s)) + \mathcal{R}^*(q(s),-D_q\mathcal{E}(s,q(s)))\mathrm{d}s \\ = \mathcal{E}(r,q(r)) + \int_r^t \partial_s\mathcal{E}(s,q(s))\mathrm{d}s.\end{aligned} \tag{2.2.4}$$

In fact, if the chain rule (2.2.3) is valid, it is sufficient to have (2.2.4) as an upper estimate, i.e.,

$$\begin{aligned}\forall\, 0 \le r < t \le T: \quad \mathcal{E}(t,q(t)) + \int_r^t \mathcal{R}(q(s),\dot{q}(s)) + \mathcal{R}^*(q(s),-D_q\mathcal{E}(s,q(s)))\mathrm{d}s \\ \le \mathcal{E}(r,q(r)) + \int_r^t \partial_s\mathcal{E}(s,q(s))\mathrm{d}s\end{aligned} \tag{EDP}$$

to conclude that all three of (2.2.2) hold true almost everywhere on $[0,T]$. This is why (EDP) is also called the **energy dissipation principle** of (2.2.2a). The argument for this is again based on convex analysis, since we find with the Fenchel-Young inequality (A.1) that

$$\langle -D_q\mathcal{E}(t,q(t)),\dot{q}(t)\rangle \le \mathcal{R}(q(t),\dot{q}(t)) + \mathcal{R}^*(q(t),-D_q\mathcal{E}(t,q(t)))$$

holds almost everywhere on $[0,T]$, and the opposite estimate follows from (EDP) with the chain rule, so that we find the power balance (2.2.2c)

In the case of a positively 1-homogeneous dissipation potential $\mathcal{R}$ (that is, if (2.2.2a) is a rate-independent system), simply by exploitation of the definition of the convex subdifferential, we find that (2.2.2a) is equivalent to the evolutionary variational inequality

$$\forall\, w \in \mathcal{Q} : \langle D_q\mathcal{E}(t,q(t)), w - \dot{q}(t)\rangle + \mathcal{R}(q(t),w) - \mathcal{R}(q(t),\dot{q}(t)) \geq 0. \tag{2.2.5}$$

Now, if $\mathcal{R}$ is positively 1-homogeneous, its subdifferential has the following properties:

Lemma 2.2.1. *[MR15, Lemma 1.3.1.]. Let $\mathcal{R} : \mathcal{Q} \to [0,\infty]$ be lower semicontinuous, convex and 1-homogeneous. Then it holds for all $v \in \mathcal{Q}$ that*

$$\partial\mathcal{R}(v) = \{\eta \in \mathcal{R}(0) \,|\, \mathcal{R}(v) = \langle \eta, v\rangle\},$$

as well as, for all $\xi \in \mathcal{Q}^*$:

$$\xi \in \partial\mathcal{R}(v) \Leftrightarrow \Big(\forall w \in \mathcal{Q} : \mathcal{R}(w) - \langle \xi, w\rangle \geq \mathcal{R}(v) - \langle \xi, v\rangle = 0\Big).$$

Therefore, for every $q \in W^{1,1}(0,T;\mathcal{Q})$ and for 1-homogeneous $\mathcal{R}$, the variational inequality (2.2.5) is equivalent to the following two holding true for almost all $t \in [0,T]$:

$$\forall\, w \in \mathcal{Q} : \langle D_q\mathcal{E}(t,q(t)), w\rangle + \mathcal{R}(q(t),w) \geq 0; \tag{S$_{\mathrm{loc}}$}$$

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

The first of these is called the **local stability condition** and is a purely static condition. It is equivalent to requiring that $D_q\mathcal{E}(t,q(t)) + \partial_v\mathcal{R}(q(t),0) \ni 0$, which, in light of the first statement of Lemma 2.2.1, is a relaxation of (RIS). Physically, it can be interpreted as a force balance, requiring that the static frictional forces be strong enough to balance the potential restoring forces. $(\mathrm{E})_{\mathrm{loc}}$ on the other hand is a power balance, establishing a relation between the power associated with the change of state and the dissipation rate. Note that $(\mathrm{E})_{\mathrm{loc}}$ is purely scalar. Again, if the chain rule (2.2.3) holds, we arrive from $(\mathrm{E})_{\mathrm{loc}}$ at the following **energy balance** for all $0 \leq r < t \leq T$

$$\mathcal{E}(t,q(t)) + \int_r^t \mathcal{R}(q(s),\dot{q}(s))\mathrm{d}s = \mathcal{E}(r,q(r)) + \int_r^t \partial_s\mathcal{E}(s,q(s))\mathrm{d}s. \tag{2.2.6}$$

This is in accordance with (EDP), since the 1-homogeneity of $\mathcal{R}$ implies that its convex conjugate is an indicator function, namely

$$\mathcal{R}^*(q,\xi) = \delta_{\partial_v\mathcal{R}(q,0)}(\xi) = \begin{cases} 0, & \text{for } \xi \in \partial_v\mathcal{R}(q,0) \\ \infty, & \text{for } \xi \notin \partial_v\mathcal{R}(q,0), \end{cases}$$

see Lemma 2.4.4. Therefore, (2.2.6) and $(\mathrm{S})_{\mathrm{loc}}$ yield (EDP), and conversely, if (EDP) holds, this implies that $-D_q\mathcal{E}(t,q(t)) \in \partial_v\mathcal{R}(q(t),0)$, hence $(\mathrm{S})_{\mathrm{loc}}$ and $(\mathrm{E})_{\mathrm{loc}}$.

At this point, we also note that in the case of a convex energy $\mathcal{E}(t,\cdot)$ and a 1-homogeneous dissipation potential, the local stability condition $(\mathrm{S})_{\mathrm{loc}}$ is equivalent to the following **global stability condition**

$$\forall\, w \in \mathcal{Q}: \quad \mathcal{E}(t,q(t)) \le \mathcal{E}(t,w) + \mathcal{R}(q(t), w - q(t)). \tag{2.2.7}$$

In fact, we have the following

Lemma 2.2.2. *We define the **local stability set***

$$\mathcal{S}_{loc}(t) := \{q \in \mathcal{Q} \mid \ \mathcal{E}(t,q) < \infty,\ -D_q\mathcal{E}(t,q) \in \partial_v \mathcal{R}(q,0)\}$$

*and the **global stability set***

$$\mathcal{S}_{glob}(t) := \{q \in \mathcal{Q} \mid \ \mathcal{E}(t,q) < \infty,\ \forall w \in \mathcal{Q}: \mathcal{E}(t,q) \le \mathcal{E}(t,w) + \mathcal{R}(q, w-q)\}. \tag{2.2.8}$$

Let $\mathcal{R}: \mathcal{Q}\times\mathcal{Q} \to [0,\infty]$ be a 1-homogeneous dissipation potential. If $\mathcal{E}(t,\cdot) \in C^1(\mathcal{Q},\mathbb{R})$, then $\mathcal{S}_{glob}(t) \subseteq \mathcal{S}_{loc}(t)$ holds for all $t \in [0,T]$. If $\mathcal{E}(t,\cdot)$ is convex, then $\mathcal{S}_{glob}(t) = \mathcal{S}_{loc}(t)$ for all t.

Proof. Let $q \in \mathcal{S}_{\mathrm{glob}}(t)$, i.e., for all $v \in \mathcal{Q}$, we have $\mathcal{E}(t,q) \le \mathcal{E}(t,v) + \mathcal{R}(q, v-q)$. Now let $\varepsilon > 0$ and $w \in \mathcal{Q}$ be arbitrary, and set $v := q + \varepsilon w$. Then we have that

$$\mathcal{E}(t,q) \le \mathcal{E}(t, q+\varepsilon w) + \varepsilon \mathcal{R}(q,w),$$

where we have already used the 1-homogeneity of $w \mapsto \mathcal{R}(q,w)$. Rearrangement of the terms and passage to the limit $\varepsilon \to 0$ then yields $q \in \mathcal{S}_{\mathrm{loc}}(t)$.

Conversely, let $\mathcal{E}(t,\cdot)$ be convex and $q \in \mathcal{S}_{\mathrm{loc}}(t)$. The convexity yields

$$\forall\, v \in \mathcal{Q}: \quad \langle D_q\mathcal{E}(t,q), v-q\rangle \le \mathcal{E}(t,v) - \mathcal{E}(t,q),$$

and local stability means that

$$\forall\, w \in \mathcal{Q}: \quad \langle -D_q\mathcal{E}(t,q), w\rangle \le \mathcal{R}(q,w).$$

Plugging the first condition for $v = w + q$ into the second condition, we arrive at $q \in \mathcal{S}_{\mathrm{glob}}(t)$. ∎

It is often reasonable to assume that the state q decomposes into $q = (y,z)$, where y is a **non-dissipative component** and z is a **dissipative component**. This goes along with a decomposition $\mathcal{Q} = \mathcal{Y} \times \mathcal{Z}$ of the state space into a nondissipative part $\mathcal{Y}$ and a dissipative part $\mathcal{Z}$, where $\mathcal{Y}$ and $\mathcal{Z}$ are Banach spaces. In applications, this can often be interpreted as a splitting into an observable variable y (such as displacement) and an internal variable z that usually is neither directly observable nor controllable from the outside (such as plastic strain, or polarization as in the example in Section 2.4.3).

This distinction comes about since in many applications, the dissipation only depends on the inner variable z, that is,

$$\mathcal{R}(q,\dot{q}) = \mathcal{R}(z,\dot{z}) \quad \text{and} \quad \Big(\mathcal{R}(z,\dot{z}) = 0 \Rightarrow \dot{z} = 0\Big). \tag{2.2.9}$$

This decomposition turns (RIS) into a coupled system, namely

$$D_y\mathcal{E}(t,y,z) = 0, \quad \partial\mathcal{R}(\dot{z}) + \mathrm{D}_z\mathcal{E}(t,y,z) \ni 0. \tag{2.2.10}$$

By minimizing with respect to y first, we obtain the **reduced energy**

$$\mathcal{J}(t,z) := \min\{\mathcal{E}(t,y,z) \,|\, y \in \mathcal{Y}\}. \tag{2.2.11}$$

Having thus satisfied the first of (2.2.10), it remains to solve the **reduced problem**

$$\partial\mathcal{R}(\dot{z}) + \mathrm{D}_z\mathcal{J}(t,z) \ni 0. \tag{2.2.12}$$

Note that we obtain a solution $q = (z,y)$ of the original problem (2.2.10), if we couple a solution $z : [0,T] \to \mathcal{Z}$ of (2.2.12) with a suitable curve $y : [0,T] \to \mathcal{Y}$ such that $y(t) \in \operatorname{Arg\,min} \mathcal{E}(t,\cdot,z(t))$.

2.2.2 Differential solutions

As already mentioned, the notion of differential solutions is the most straightforward one, simply requiring the subdifferential inclusion (RIS) to be fulfilled pointwise almost everywhere. First existence results were shown in [Bré73] for quadratic energies, and more generally in [CV90] for smooth and uniformly convex energies. In order to ensure that all quantities are well-defined and that the reformulations in the previous section are possible, $q \in W^{1,1}(0,T;\mathcal{Q})$ is the minimal regularity that is required.

Definition 2.2.3 (Differential Solutions)**.** *We call $q : [0,T] \to \mathcal{Q}$ a **differential solution** of* (RIS), *if $q \in W^{1,1}(0,T;\mathcal{Q})$ and*

$$\text{f.a.a. } t \in [0,T] : \quad \partial_{\dot{q}}\mathcal{R}(q(t),\dot{q}(t)) + D_q\mathcal{E}(t,q(t)) \ni 0.$$

This is a stronger notion of solution than the energetic one introduced in Section 2.2.3 in the sense that any energetic solution q that fulfills $q \in W^{1,1}(0,T;\mathcal{Q})$ is a differential solution, see Lemma 2.2.6. In this line of reasoning, several existence results can be found e.g. in [MR15], cf. Cor. 3.4.6 or Thm. 3.4.7 therein. There, based on additional assumptions on smoothness and convexity of the energy, energetic solutions are shown to have higher regularity and thus to be differential solutions.

However, in the general nonconvex case, an RIS may have energetic solutions, but not allow for solutions in the differential sense, see the 1-dimensional example in Section 2.3. Even if a non-convex RIS does possess both differential and energetic solutions, this does not guarantee that they coincide, see, e.g. [MR15, Example 1.8.2].

2.2.3 Energetic solutions

Refering back to the decomposition into a dissipative and a non-dissipative component, we will from now on assume reduced state-dependence of the dissipation potential according to (2.2.9). In this work, we will even go as far as to entirely drop the z-dependence of $\mathcal{R}$, that is, we assume hence forth that $\mathcal{R}(z,\dot{z}) = \mathcal{R}(\dot{z})$ depends on $\dot{z}$ alone.

In order to motivate the definition of energetic solutions, we return to the equivalent formulation $(\mathrm{S})_{\mathrm{loc}}$ & $(\mathrm{E})_{\mathrm{loc}}$ from Section 2.2.1. We will define energetic solutions by pairing the global stability condition with the energy balance (2.2.6) that is obtained from $(\mathrm{E})_{\mathrm{loc}}$, which thanks to the new assumption on $\mathcal{R}$ now reads

$$\mathcal{E}(t,q(t)) + \int_r^t \mathcal{R}(\dot{z}(s))\mathrm{d}s = \mathcal{E}(r,q(r)) + \int_r^t \partial_s\mathcal{E}(s,q(s))\mathrm{d}s.$$

In order to obtain a formulation that does not require any differentiability of q, we replace the second term on the left hand side by the **total variation induced by** $\mathcal{R}$, which is definded as

$$\mathrm{Var}_{\mathcal{R}}(z;[r,t]) := \sup\left\{ \sum_{m=1}^{M} \mathcal{R}(z(t_m) - z(t_{m-1})) \,\middle|\, r = t_0 < t_1 < \cdots < t_{M-1} < t_M = t \right\}.$$

In the case that $z \in W^{1,1}(0,T;\mathcal{Z})$, it actually holds that $\int_r^t \mathcal{R}(\dot{z}(s))\mathrm{d}s = \mathrm{Var}_{\mathcal{R}}(z;[r,t])$, cf. Lemma C.18. We thus arrive at the following

Definition 2.2.4. *[Global energetic solutions] A curve $q = (y,z) : [0,T] \to \mathcal{Q} = \mathcal{Y}\times\mathcal{Z}$ is a **(global) energetic solution** of* (RIS), *if $t \mapsto \partial_t\mathcal{E}(t,q(t))$ is integrable and if the global stability* (S) *and the energy balance* (E) *hold for all $t \in [0,T]$:*

$$\forall\, \hat{q} = (\hat{y},\hat{z}) \in \mathcal{Q}: \quad \mathcal{E}(t,q(t)) \le \mathcal{E}(t,\hat{q}) + \mathcal{R}(\hat{z} - z(t)) \tag{S}$$

$$\mathcal{E}(t,q(t)) + \mathrm{Var}_{\mathcal{R}}(z;[0,T]) = \mathcal{E}(0,q(0)) + \int_0^t \partial_s\mathcal{E}(s,q(s))\mathrm{d}s \tag{E}$$

As already mentioned at the beginning of this chapter, one of the main advantages of energetic solutions is the fact that their definition only requires differentiability of the scalar map $t \mapsto \mathcal{E}(t,q)$. If one were to replace the dissipation potential $\mathcal{R}$ by a **dissipation distance** $\mathcal{D} : \mathcal{Z}\times\mathcal{Z} \to [0,\infty]$ and the expression $\mathcal{R}(z_2 - z_1)$ by $\mathcal{D}(z_1,z_2)$, one could thus formulate a theory of rate-independent systems on spaces $\mathcal{Z}$ without linear structure.

What is more, (S) & (E) can be solved via time discretization, where the incremental problems are minimization problems based on the global stability condition (S). To be more precise, let $(0 = t_0, t_1, \ldots, t_N = T)$ be a partition of the interval $[0,T]$, i.e., let $t_0 < t_1 < \cdots < t_N$. We then consider the following **incremental minimization problem**:

$$\text{For given } q_0 \in \mathcal{S}_{\mathrm{glob}}(0) \text{ find } q_k \in \operatorname{Arg\,min}\left\{ \mathcal{E}(t_k,q) + \mathcal{R}(z - z_{k-1}) \,\middle|\, q \in \mathcal{Q} \right\},\ k = 1,\ldots,N \tag{IMP}$$

Now, the discrete minimization problems can be solved in the following way: First, choose an infimizing sequence and show that it has a subsequence that is convergent w.r.t. either the strong or the weak topology on $\mathcal{Q}$. Second, show that the objective functional is lower semicontinuous w.r.t. the same topology. This implies that the limit of the converging subsequence must be a minimizer. In function spaces, this approach is known as the **direct method in the calculus of variations**.

The solutions of the discrete minimization problems are then used to define interpolants. The following additional smallness assumptions (E2) on the energy $\mathcal{E}$ then allow for a priori estimates for these interpolants, and the compatibility of $\mathcal{E}$ and $\mathcal{R}$ as required by (C1) - (C2) ensures that they converge to an energetic solution, see Prop. 2.2.5. To be precise, we need the following assumptions

on the dissipation potential

$$\begin{aligned} &\forall\, z_1, z_2, z_3 \in \mathcal{Z}: \quad \mathcal{R}(z_3 - z_1) \le \mathcal{R}(z_2 - z_1) + \mathcal{R}(z_3 - z_2) \\ &\forall\, z \in \mathcal{Z}: \quad \mathcal{R}(z) = 0 \Leftrightarrow z = 0 \end{aligned} \tag{D1}$$

$$\begin{aligned} &\mathcal{R} : \mathcal{Z} \to [0,\infty] \text{ is lower semicontinuous, convex and} \\ &\text{positively 1-homogeneous} \end{aligned} \tag{D2}$$

on the energy functional

$$\forall\, t \in [0,T]: \quad \mathcal{E}(t,\cdot) : \mathcal{Q} \to \mathbb{R}_\infty \text{ has compact sublevels} \tag{E1}$$

$$\begin{aligned} &\operatorname{Dom} \mathcal{E} = [0,T] \times \operatorname{Dom} \mathcal{E}(0,\cdot), \\ &\exists\, c_{\mathcal{E}} \in \mathbb{R},\ \lambda_{\mathcal{E}} \in L^1(0,T),\ N_{\mathcal{E}} \subset [0,T] \text{ with } \mathcal{L}^1(N_{\mathcal{E}}) = 0 \\ &\forall\, q \in \operatorname{Dom} \mathcal{E}(0,\cdot) : \mathcal{E}(\cdot,q) \in W^{1,1}(0,T), \\ &\qquad\qquad \partial_t \mathcal{E}(t,q) \text{ exists for } t \in [0,T] \setminus N_{\mathcal{E}} \text{ and satisfies} \\ &\qquad\qquad |\partial_t \mathcal{E}(t,q)| \le \lambda_{\mathcal{E}}(t)\Big(\mathcal{E}(t,q) + c_{\mathcal{E}}\Big). \end{aligned} \tag{E2}$$

Calling a sequence $(t_m, q_m)_{m\in\mathbb{N}} \subset [0,T] \times \mathcal{Q}$ a **stable sequence**, if

$$\sup_{m\in\mathbb{N}} \mathcal{E}(t_m, q_m) < \infty \quad \text{and} \quad \forall\, m \in \mathbb{N} : q_m \in \mathcal{S}_{\text{glob}}(t_m),$$

where the global stability set $\mathcal{S}_{\text{glob}}(t)$ was defined in (2.2.8), we further require

the compatibility of $\mathcal{E}$ and $\mathcal{R}$

$$\forall \text{ stable sequences } (t_m, q_m)_{m\in\mathbb{N}} \text{ with } (t_m, q_m) \overset{[0,T]\times\mathcal{Q}}{\longrightarrow} (t,q) \text{ we have:}$$

$$t \in [0,T] \setminus N_{\mathcal{E}} \text{ with } N_{\mathcal{E}} \text{ from (E2)} \Rightarrow \partial_t \mathcal{E}(t,q) = \lim_{m\to\infty} \partial_t \mathcal{E}(t, q_m), \tag{C1}$$

$$q \in \mathcal{S}_{\text{glob}}(t). \tag{C2}$$

It is crucial to realize that the notion of stable sequences intrinsically links the choice of topology in $\mathcal{Q}$ with the properties of $\mathcal{E}$ and $\mathcal{R}$, since the type of convergence in (E1) and in (C1) and (C2) has to coincide. These assumptions now allow for the following existence result:

Proposition 2.2.5. *[MR15, Thm. 2.1.6] Assume that $\mathcal{E}$ and $\mathcal{R}$ fulfill the assumptions* (D1)-(D2), (E1)-(E2), *and the compatibility conditions* (C1)-(C2). *Assume further that the topology of $\mathcal{Q}$ restricted to compact sets is separable and metrizable. Then, for each $q_0 \in \mathcal{S}_{glob}(0)$, there exists an energetic solution $q = (y,z) : [0,T] \to \mathcal{Q}$ to the initial value problem* (RIS), *and q is measurable.*

The proof in [MR15] is based on the incremental minimization problem (IMP) and shows that for the dissipative component z, the right-continuous piecewise constant interpolants converge pointwisely to an energetic solution when the fineness of the partition of $[0,T]$ goes to zero. When it comes to the regularity of the resulting curve, it should be noted that the energy balance (E) in combination with the assumption (E2) are sufficient to show that energetic solutions $q = (y,z) : [0,T] \to \mathcal{Q}$ satisfy a priori estimates for the energies $\mathcal{E}(t,q(t))$ and for the variation $\mathrm{Var}_{\mathcal{R}}(z;[0,T])$. Since in many applications, the energy $\mathcal{E}(t,\cdot)$ is coercive and the dissipation potential $\mathcal{R}$ is a norm on $\mathcal{Q}$, this often translates into L^∞-regularity for q and a BV-estimate for the dissipative component z. However, the *global* stability condition (S) at the core of the definition prompts energetic solutions to change state as soon as the release of the energy $\mathcal{E}(t,q(t)) - \mathcal{E}(t,\hat{q})$ for any possible state $\hat{q} \in \mathcal{Q}$ is no longer compensated by the dissipation $\mathcal{R}(\hat{z} - z(t))$. This can lead to jumps over energy barriers, even if external loadings are completely smooth. Therefore, we cannot expect continuity of energetic solutions. Temporal continuity of energetic solutions can be shown for example under strong convexity assumptions on $\mathcal{E}$, see e.g. [MR15, Section 3.4.2] or [TM10]. As mentioned before, if an energetic solution is absolutely continuous and if the chain rule (2.2.3) holds, it is also a differential solution, as is shown in the next Lemma:

Lemma 2.2.6. *Let $\mathcal{Q} = \mathcal{Y} \times \mathcal{Z}$ be a Banach space and $q = (y,z) : [0,T] \to \mathcal{Q}$ be a global energetic solution of* (RIS). *Let further the chain rule* (2.2.3) *hold true for almost all $t \in [0,T]$. If $z \in W^{1,1}(0,T;\mathcal{Z})$, then q is also a differential solution of* (RIS).

Proof. For the proof, we simply verify that the required regularity is sufficient to execute the reformulations that were done formally in Section 2.2.1. Let q be a global energetic solution, then we know from Lemma 2.2.2 that we have for all $t \in [0,T]$ that $q(t) \in \mathcal{S}_{\mathrm{loc}}(t)$. On the other hand, since $q \in W^{1,1}(0,T;\mathcal{Z})$, we know from Lemma C.18 that $\mathrm{Var}_{\mathcal{R}}(z;[0,T]) = \int_0^T \mathcal{R}(\dot{z}(s))\mathrm{d}s$, which allows us to conclude from (E) that (2.2.6) holds. Applying the chain rule (2.2.3) and localizing the integral, this yields $(\mathrm{E})_{\mathrm{loc}}$. ∎

2.2.4 $\mathfrak{p}$-parameterized Balanced Viscosity (BV) solutions

From now on, we only consider the reduced problem (2.2.12).

While energetic solutions bring considerable advantages, such as the fact that they lend themselves to numerical approximation via time discretization, the lack of smoothness that was mentioned at the end of the previous section remains an issue. This is true in particular since the jumps over energy barriers that might occur could be considered unphysical.

Solution concepts relying on viscosity approximations

One way to prevent these jumps over energy barriers is a regularization approach. Namely, we assume that there is a second Banach space $\mathcal{V}$ into which $\mathcal{Z}$ is compactly embedded, and instead of the rate-independent system (2.2.12), we consider for $\varepsilon > 0$ the viscously regularized system

$$\partial\mathcal{R}_\varepsilon(\dot{z}_\varepsilon) + \mathrm{D}_z\mathcal{J}(t, z_\varepsilon) \ni 0, \tag{2.2.13}$$

where $\mathcal{R}_\varepsilon(z) := \mathcal{R}(z) + \mathcal{R}_{2,\varepsilon}(z) := \mathcal{R}(z) + \frac{\varepsilon}{2}\|z\|_{\mathcal{V}}^2$. We then consider solutions of (2.2.13), which, in contrast to energetic solutions of the rate-independent system (2.2.12), can be shown to be unique under rather mild assumptions and have better regularity. In Section 3.1, we will show the existence of absolutey continuous solutions z_ε of the viscous system, fulfilling the inclusion (2.2.13) almost everywhere, for the setting described in Section 2.4.1, cf. Definition 3.1.1. In this context, it is worth noting that we obtain the curves z_ε via a time discretization scheme similar to (IMP), namely, for a step size $\tau > 0$, we define the partition $(0 = t_0^\tau, t_1^\tau, \dots, t_{N-1}^\tau < T \le t_n^\tau)$ of the interval $[0, T]$ via $t_k^\tau := \tau k$. For a given initial value $z_0 \in \mathcal{Z}$, we then set $z_0^\tau := z_0 \in \mathcal{Z}$ and choose the next iterate

$$z_{k+1}^\tau \in \operatorname{Argmin}\Big\{\mathcal{I}(\ell(t_{k+1}^\tau), z) + \mathcal{R}(z - z_k^\tau) + \tfrac{\varepsilon}{2\tau}\|z - z_k^\tau\|_{\mathcal{V}}^2 \,\Big|\, z \in \mathcal{Z}\Big\}. \tag{2.2.14}$$

cf. (3.1.9). The minimization problem (2.2.14) can be interpreted as a localization of (IMP) in the sense that, when choosing the next iterate, those states z, whose distance $\|z - z_k^\tau\|_{\mathcal{V}}$ to the current iterate is big, are penalized. After passage to the limit in the time discretization scheme, the limiting curves z_ε also fulfill the following energy dissipation balance

$$\begin{aligned}\mathcal{J}(s, z_\varepsilon(s)) + \int_0^s \mathcal{R}_\varepsilon(\dot{z}_\varepsilon(r)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{J}(r, z_\varepsilon(r)))\,\mathrm{d}r \\ = \mathcal{J}(0, z_\varepsilon(0)) + \int_0^s \partial_r\mathcal{J}(r, z(r))\,\mathrm{d}r,\end{aligned} \tag{E$_\varepsilon$}$$

which is in analogy to (E), but now containing the convex conjugate $\mathcal{R}_\varepsilon^*$ of $\mathcal{R}_\varepsilon$, since the augmented dissipation potential $\mathcal{R}_\varepsilon$ is not 1-homogeneous. The natural next step is to pass to the limit $\varepsilon \to 0$ and study the limit of the viscous solutions z_ε.

While we refer to the forthcoming Section 2.4.1 for more details on the regularization in our setting, the intention for now is to give a brief insight into why we will not pass to the limit directly, but instead opt for a reparameterization technique. To this end, we summarize here some of the arguments in [MRS12a, Sections 2.3, 3.1, 4.1], where different viscous approxmation schemes for a rate-independent system in finite dimensions are carried out, since the finite dimensional setting allows for significant simplifications and more straight-forward arguments. In particular, since on finite dimensional spaces all norms are equivalent, the distinction between the spaces $\mathcal{V}$ and $\mathcal{Z}$ becomes obsolete. See also [MRS16, Chaps. 3, 4] for details on ($\mathfrak{p}$-parameterized) BV solutions in infinite dimensions, and Section 3.3 for a discussion in our setting as it is laid out in Section 2.4.1.

Local solutions

As we shall see in Chapter 3.1, the functions z_ε solving (2.2.13) satisfy the a priori bound

$$\int_0^T \mathcal{R}_\varepsilon(\dot{z}_\varepsilon(r)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{J}(r, z_\varepsilon(r)))\,\mathrm{d}r \le C,$$

cf. (3.1.7c), and thus, application of Helly's compactness theorem yields pointwise weak convergence of a subsequence to a curve $z \in \mathrm{BV}([0,T],\mathcal{Z})$. By standard lower semicontinuity arguments, we might then expect from (2.2.13) that the limit z fulfills the local stability condition $\partial\mathcal{R}(0) + \mathrm{D}_z\mathcal{J}(t, z(t)) \ni 0$, i.e., $t \in S_{\mathrm{loc}}$ for almost all $t \in [0,T]$; and passing to the limit inferior in $(\mathrm{E})_\varepsilon$, we arrive at the following

Definition 2.2.7 (Local solutions). *A curve $z \in \mathrm{BV}([0,T];\mathcal{Z})$ is called a **local solution** of the rate-independent system* (2.2.12)*, if it fulfills the local stability condition*

$$\partial\mathcal{R}(0) + \mathrm{D}_z\mathcal{J}(t, z(t)) \ni 0 \textit{ for all } t \in [0,T] \setminus J_z, \tag{S$_{\mathrm{loc}}$}$$

and the energy dissipation inequality

$$\mathcal{J}(s, z(s)) + \mathrm{Var}_{\mathcal{R}}(z;[0,s]) \le \mathcal{J}(0, z(0)) + \int_0^s \partial_r\mathcal{J}(r, z(r))\,\mathrm{d}r \tag{2.2.15}$$

for all $s \in [0,T]$. Here, $J_z \subset [0,T]$ is the at most countable jump set of z.

Having replaced the global stability condition (S) by $(\mathrm{S})_{\mathrm{loc}}$, and the energy dissipation balance (E) by the inequality (2.2.15), the notion of local solutions is weaker than that of GES. In fact, the inequality (2.2.15) may be strict, that is, local solutions may exhibit a loss of energy. What is more, local solutions lack a description of the behaviour in jumps. To see this, let us first give a precise definition of the jump set J_z. Denoting the left and right limits of $z \in \mathrm{BV}([0,T];\mathcal{Z})$ at an arbitrary time $t \in [0,T]$ by

$$z(t_-) := \lim_{s \nearrow t} z(s); \quad z(t_+) := \lim_{s \searrow t} z(s); \quad z(0_-) := z(0); \quad z(T_+) := z(T), \tag{2.2.16}$$

we define

$$\begin{aligned} J_z &:= \{t \in [0,T] \,|\, z(t_-) \neq z(t) \text{ or } z(t) \neq z(t_+)\} \\ &\supset \mathrm{ess} - J_z := \{t \in [0,T] \,|\, z(t_-) \neq z(t_+)\}, \end{aligned} \tag{2.2.17}$$

where $\mathrm{ess} - J_z$ is the essential jump set of z. If we further denote the energy that is dissipated when changing from a state $z_0 \in \mathcal{Z}$ to a state $z_1 \in \mathcal{Z}$ by

$$\Delta_{\mathcal{R}}(z_0, z_1) := \mathcal{R}(z_1 - z_0); \text{ and } \Delta_{\mathcal{R}}(z_-, z, z_+) := \mathcal{R}(z - z_-) + \mathcal{R}(z_+ - z), \tag{2.2.18}$$

it is possible to deduce the following description of the behaviour in jumps in the finite dimensional setting, see [MRS12a, Props. 2.2, 2.7]:

Proposition 2.2.8. *Let* $z \in \mathrm{BV}([0,T];\mathcal{Z})$ *be*

(i) a global energetic solution of (2.2.12). *Then it holds for every* $t \in J_z$ *that*

$$\begin{aligned}\mathcal{J}(t,z(t)) - \mathcal{J}(t,z(t_-)) &= -\Delta_{\mathcal{R}}(z(t_-),z(t)),\\ \mathcal{J}(t,z(t_+)) - \mathcal{J}(t,z(t)) &= -\Delta_{\mathcal{R}}(z(t),z(t_+)),\\ \mathcal{J}(t,z(t_+)) - \mathcal{J}(t,z(t_-)) &= -\Delta_{\mathcal{R}}(z(t_-),z(t_+)).\end{aligned} \tag{J$_{\mathrm{ener}}$}$$

(ii) a local solution of (2.2.12). *Then it holds for every* $t \in J_z$ *that*

$$\begin{aligned}\mathcal{J}(t,z(t)) - \mathcal{J}(t,z(t_-)) &\le -\Delta_{\mathcal{R}}(z(t_-),z(t)),\\ \mathcal{J}(t,z(t_+)) - \mathcal{J}(t,z(t)) &\le -\Delta_{\mathcal{R}}(z(t),z(t_+)),\\ \mathcal{J}(t,z(t_+)) - \mathcal{J}(t,z(t_-)) &\le -\Delta_{\mathcal{R}}(z(t_-),z(t_+)).\end{aligned} \tag{J$_{\mathrm{local}}$}$$

Together with a suitable subdifferential inclusion (see [MRS12a, Prop. 27]), ($\mathrm{J}_{\mathrm{local}}$) is actually sufficient to conclude that z is a local solution, but not that it fulfills an energy balance. Hence, a finer description of the dissipation in jumps is needed.

BV solutions

To this end, we return to the viscously regularized system (2.2.13). Let us for the moment assume that the external load is constant and that $\mathcal{J}(t,z(t)) = \mathcal{J}(z(t))$ at all times. In this scenario, if we consider two states $z_0, z_1 \in \mathcal{Z}$ that are connected along an almost everywhere differentiable path $\zeta \in \mathrm{AC}([t_0,t_1];\mathcal{Z})$ such that $z(t_0) = z_0$ and $z(t_1) = z_1$ and (2.2.13) is fulfilled, the energy dissipation balance $(\mathrm{E})_\varepsilon$ predicts for the release of energy

$$\mathcal{J}(z_0) - \mathcal{J}(z_1) = \int_{t_0}^{t_1} \mathcal{R}_\varepsilon(\dot{\zeta}(t)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}\mathcal{J}(\zeta(t)))\,\mathrm{d}t.$$

Keeping in mind that we are in need of a lower bound for the right hand side which is independent of $\varepsilon > 0$, it is now natural to consider

$$\mathfrak{p}(v,w) := \inf_{\varepsilon>0}\Big(\mathcal{R}_\varepsilon(v) + \mathcal{R}_\varepsilon^*(w)\Big), \qquad \text{for } v \in \mathcal{V}, w \in \mathcal{V}^*, \tag{2.2.19}$$

satisfying for all $\varepsilon > 0$ the estimate

$$\mathcal{J}(z_0) - \mathcal{J}(z_1) \ge \int_{t_0}^{t_1} \mathfrak{p}(\dot{\zeta}(t)), -\mathrm{D}\mathcal{J}(\zeta(t)))\,\mathrm{d}t.$$

On the other hand, we find

$$\text{for all } v \in \mathcal{V}, w \in \mathcal{V}^* : \langle w, v\rangle_{\mathcal{V}} \le \mathfrak{p}(v,w) \text{ and } \mathcal{R}(v) \le \mathfrak{p}(v,w), \tag{2.2.20}$$

employing the Fenchel-Young inequality (A.1) for the first estimate. Thus, if $\widetilde{\zeta} \in \mathrm{AC}([0,T];\mathcal{Z})$ is an arbitrary curve connecting the states z_0 and z_1, and if we have a chain rule for the scalar map $t \mapsto \mathcal{J}(\widetilde{\zeta}(t))$, we also have the opposite estimate

$$\mathcal{J}(z_0) - \mathcal{J}(z_1) = \int_{t_0}^{t_1} \langle -\mathrm{D}\mathcal{J}(\widetilde{\zeta}(t)), \dot{\widetilde{\zeta}}(t)\rangle_{\mathcal{V}} \,\mathrm{d}t \le \int_{t_0}^{t_1} \mathfrak{p}(\dot{\widetilde{\zeta}}(t), -\mathrm{D}\mathcal{J}(\widetilde{\zeta}(t)))\,\mathrm{d}t.$$

In conclusion,

$$V_{\mathfrak{p}}(\zeta, [t_0, t_1]) := \int_{t_0}^{t_1} \mathfrak{p}(\dot{\zeta}(t), -\mathrm{D}\mathcal{J}(\zeta(t)))\,\mathrm{d}t$$

is always an upper bound for the energy that is released when changing from state $z(t_0)$ to state $z(t_1)$, and this bound is reached along those curves following the viscous regime (2.2.13).

The function $\mathfrak{p}(\cdot,\cdot)$ is called the **vanishing viscosity contact potential** and is studied in great detail in [MRS12a, Section 3].

In order to carry out the vanishing viscosity limit, we again choose a sequence $(z_\varepsilon)_{\varepsilon>0}$ of solutions of the viscously regularized problem (2.2.13) and obtain a weak pointwise limit $z \in \mathrm{BV}([0,T];\mathcal{Z})$. Considering the second estimate in (2.2.20), $\liminf_{\varepsilon\to 0} V_{\mathfrak{p}}(z_\varepsilon, [0,s])$ seems to be a natural candidate to replace $\mathrm{Var}_{\mathcal{R}}(z, [0,s])$ in the energy dissipation inequality (2.2.15) in order to obtain a sharper estimate. However, since the limit function z can only be guaranteed to be of bounded variation, it lacks the necessary differentiability, and $V_{\mathfrak{p}}(z,[0,s])$ is not well-defined in general.

This problem can be circumvented using measure theory in the following way, still following [MRS12a]: The distributional derivative z' of $z \in \mathrm{BV}([0,T];\mathcal{Z})$ defines a Radon vector measure with finite total variation $|z'|$. It has been shown, e.g., in [AFP00, Cor. 3.33], that z' can be decomposed into

$$z' = z'_{\mathscr{L}} + z'_C + z'_J, \quad z'_{\mathrm{co}} := z'_{\mathscr{L}} + z'_C,$$

where $z'_{\mathscr{L}}$ is the absolutely continuous part w.r.t. the Lebesgue measure $\mathscr{L}^1$, and z'_C is the Cantor part, still satisfying $z'_C(\{t\}) = 0$ for all $t \in [0,T]$, whereas z'_J is a discrete measure concentrated on $\mathrm{ess} - J_z$. Therefore, $z'_{\mathrm{co}} := z'_{\mathscr{L}} + z'_C$ is the diffuse part of z'. Introducing the reference measure $\mu := \mathscr{L}^1 + |z'_C|$, the authors of [MRS12a] then define for every $(a,b) \subset (0,T)$

$$\int_a^b \mathrm{d}\mathcal{R}(z'_{\mathrm{co}}) := \int_a^b \mathcal{R}\Big(\frac{\mathrm{d}z'_{\mathrm{co}}}{\mathrm{d}\mu}\Big)\mathrm{d}\mu,$$

This allows for a representation of the pointwise $\mathcal{R}$-variation $\mathrm{Var}_{\mathcal{R}}(z;[a,b])$ of z in terms of its distributional derivate z' as

$$\mathrm{Var}_{\mathcal{R}}(z;[a,b]) = \int_a^b \mathrm{d}\mathcal{R}(z'_{\mathrm{co}}) + \mathrm{Jmp}_{\mathcal{R}}(z;[a,b]) \tag{2.2.21}$$

with the jump contribution (recall the definition of $\Delta_{\mathcal{R}}$ in (2.2.18))

$$\mathrm{Jmp}_{\mathcal{R}}(z;[a,b]) := \Delta_{\mathcal{R}}(z(a), z(a_+)) + \Delta_{\mathcal{R}}(z(b_-), z(b)) + \sum_{t\in J_z\cap(a,b)} \Delta_{\mathcal{R}}(z(t_-), z(t), z(t_+)). \tag{2.2.22}$$

Now, in analogy to (2.2.21)-(2.2.22), the authors of [MRS12a] define the **jump variation** of a curve $z \in \mathrm{BV}([0,T];\mathcal{Z})$ **induced by** $(\mathfrak{p},\mathcal{J})$ **on** $[a,b] \subset [0,T]$ as

$$\mathrm{Jmp}_{\mathfrak{p},\mathcal{J}}(z;[a,b]) := \Delta_{\mathfrak{p},\mathcal{J}}(a;z(a),z(a_+)) + \Delta_{\mathfrak{p},\mathcal{J}}(b;z(b_-),z(b)) + \sum_{t \in J_z \cap (a,b)} \Delta_{\mathfrak{p},\mathcal{J}}(t;z(t_-),z(t),z(t_+)),$$

where for $z_0, z_1 \in \mathcal{V}$,

$$\Delta_{\mathfrak{p},\mathcal{J}}(t;z_0,z_1) := \inf\left\{ \int_{t_0}^{t_1} \mathfrak{p}(\dot{\zeta}(t), -\mathrm{D}_z\mathcal{J}(t,\zeta(t)))\mathrm{d}t \,:\, \zeta \in \mathrm{AC}([t_0,t_1];\mathcal{Z}), \zeta(t_0) = z_0, \zeta(t_1) = z_1 \right\}, \tag{2.2.23}$$

is called the **Finsler cost induced by** $\mathfrak{p}$ **and** $\mathcal{J}$ at time t, and, just as in (2.2.18),

$$\Delta_{\mathfrak{p},\mathcal{J}}(t,z_-,z,z_+) := \Delta_{\mathfrak{p},\mathcal{J}}(t,z_-,z) + \Delta_{\mathfrak{p},\mathcal{J}}(t,z,z_+).$$

Finally, the **(pseudo-)total variation induced by** $(\mathfrak{p},\mathcal{J})$ is defined as

$$\mathrm{Var}_{\mathfrak{p},\mathcal{J}}(z;[a,b]) := \int_a^b \mathrm{d}\mathcal{R}(z'_{\mathrm{co}}) + \mathrm{Jmp}_{\mathfrak{p},\mathcal{J}}(z;[a,b]), \tag{2.2.24}$$

differing from the pointwise $\mathcal{R}$-variation (2.2.21) precisely in the way the jump contribution is measured. Replacing the $\mathcal{R}$-variation by this (pseudo-)total variation and thus obtaining an energy balance in place of an inequality, then yields the following definition of BV solutions:

Definition 2.2.9 (BV solutions)**.** *A curve $z \in \mathrm{BV}([0,T];\mathcal{Z})$ is called a **BV solution** of the rate-independent system* (2.2.12)*, if it fulfills the local stability condition*

$$\partial\mathcal{R}(0) + \mathrm{D}_z\mathcal{J}(t,z(t)) \ni 0 \text{ for all } t \in [0,T] \setminus J_z, \tag{S$_{\mathrm{loc}}$}$$

and the energy dissipation balance

$$\mathcal{J}(s,z(s)) + \mathrm{Var}_{\mathfrak{p},\mathcal{J}}(z;[0,s]) = \mathcal{J}(0,z(0)) + \int_0^s \partial_r\mathcal{J}(r,z(r))\,\mathrm{d}r \tag{2.2.25}$$

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

As expected, the solutions z_ε of the viscously regularized systems (2.2.13) converge pointwisely to a BV solution of (2.2.12) with $\varepsilon \to 0$, cf. [MRS12a, Thm. 4.10]. On the other hand, it is not always the case that every BV solution of an RIS can be obtained as a vanishing viscosity limit. In fact, those curves $z : [0,T] \to \mathcal{Z}$ that can be obtained as pointwise limits of solutions of the viscously regularized systems (2.2.13) are called **approximable solutions** in [MR15, Section 1.8] and in general form a proper subset of the set of BV solutions. We cite here the following example from [MR15, Ex. 1.8.3] to illustrate this fact:

Example 2.2.10. *Let $T > 0$ and let the energy $\mathcal{J} : [0,T] \times \mathbb{R} \to \mathbb{R}$ be given in dependence of the time and state by*

$$\mathcal{J}(t,z) := -\ell(t) \cdot z + \begin{cases} \frac{1}{2}(z+4)^2, & \text{if } z \leq -2 \\ 4 - \frac{1}{2}z^2, & \text{if } -2 \leq z \leq 2 \\ \frac{1}{2}(z-4)^2, & \text{if } z \geq 2, \end{cases} \tag{2.2.26}$$

where the external load is given by $\ell(t) := \min\{t, 6-t\}$. We further assume that the dissipation potential $\mathcal{R} : \mathbb{R} \to [0,\infty)$ is the absolute value function, i.e., $\mathcal{R}(z) := |z|$ for $z \in \mathbb{R}$. We consider the RIS

$$-\mathrm{D}_z\mathcal{J}(t,z(t)) \in \partial\mathcal{R}(\dot{z}(t)), \quad t \in [0,T]; \quad z(0) = -5. \tag{2.2.27}$$

Then there are two different BV solutions

$$z_1(t) := \begin{cases} t-5, & \text{for } t \in [0,3), \\ 6, & \text{for } t \in [3,5], \\ 11-t, & \text{for } t \in [5,9], \\ 3-t, & \text{for } t > 9; \end{cases} \qquad z_2(t) := \begin{cases} t-5, & \text{for } t \in [0,3], \\ -2, & \text{for } t \in [3,5], \\ 3-t, & \text{for } t \geq 5. \end{cases} \tag{2.2.28}$$

For $\varepsilon > 0$, we obtain the curves z_ε as unique solutions of the viscously regularized systems

$$-\mathrm{D}_z\mathcal{J}(t,z_\varepsilon(t)) - \varepsilon\dot{z}_\varepsilon(t) \in \partial\mathcal{R}(\dot{z}_\varepsilon(t)), \quad t \in [0,T]; \quad z_\varepsilon(0) = -5, \tag{2.2.29}$$

and they read

$$z_\varepsilon(t) := \begin{cases} t-5+\varepsilon(\exp(-t/\varepsilon)-1), & \text{for } t \in [0,3], \\ z_\varepsilon^*, & \text{for } t \in [3,t_\varepsilon^*], \\ 3-t+\varepsilon(\exp\Big(\frac{-(t-t_\varepsilon^*)}{\varepsilon}\Big)-1), & \text{for } t \geq t_\varepsilon^*, \end{cases} \tag{2.2.30}$$

where $z_\varepsilon^* = z_\varepsilon(3_-)$ is the left limit of z_ε at $t = 3$, and $t_\varepsilon^* = 3 - z_\varepsilon^*$. Note that $z_\varepsilon^* \to -2$ and $t_\varepsilon^* \to 5$ for $\varepsilon \to 0$. Then it holds for every $t \geq 0$ that $z_\varepsilon(t) \to z_2(t)$ for $\varepsilon \to 0$, whereas the discontinuous BV solution z_1 can not be obtained as a vanishing viscosity limit, i.e., z_2 is an approximable solution, and z_1 is not.

This is in a way the desired result, since we motivated the vanishing viscosity approach by the intention to prevent jumps over energy barriers. Since both $\mathcal{J}$ and and the approximable solution z_2 are continuous, no jumps occour along the graph of $\mathcal{J}(\cdot,z_2(\cdot))$. On the other hand, z_1 has jumps at $t_1 := 3$ and at $t_2 := 9$, and at these points, the energy $\mathcal{J}(\cdot,z_1(\cdot))$ has the left and right limits

$$\mathcal{J}(3_-,z_1(3_-)) = 8 \text{ and } \mathcal{J}(3_+,z_1(3_+)) = 32; \quad \mathcal{J}(9_-,z_1(9_-)) = 8 \text{ and } \mathcal{J}(9_+,z_1(9_+)) = -16,$$

i.e., jumps over energy barriers occur here. The graphs of z_1, z_2 and of the viscous approximations z_ε for $\varepsilon \in \{0.1, 0.2, 0.3\}$ are shown in Figure 2.1

It should be noted that BV solutions fulfill the following jump conditions, compare with Prop. 2.2.8:

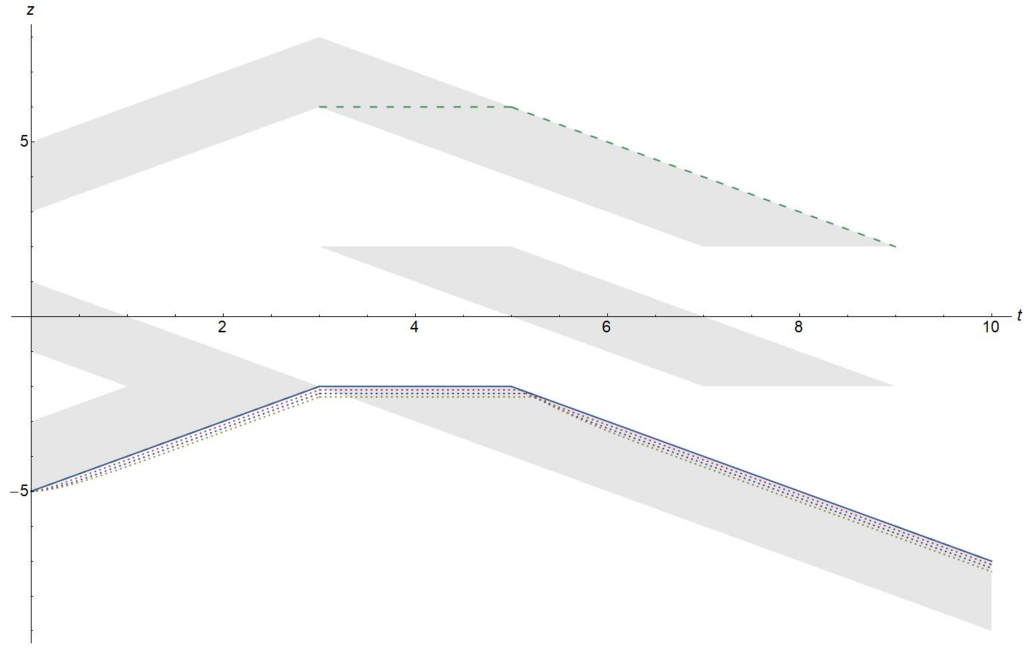


Figure 2.1: Graphs of the two BV solutions z_1 (dashed, green) and z_2 (solid, blue) of the RIS (2.2.27)*. The viscous approximations z_ε for $\varepsilon \in \{0.1, 0.2, 0.3\}$ are dashed; where the graph of z_1 is not visible, it coincides with that of z_2. The admissible set* $\{(t,z) \,|\, -\mathrm{D}_z\mathcal{J}(t,z) \in \partial\mathcal{R}(0)\}$ *is shaded.*

Proposition 2.2.11. *[MRS12a, Props. 2.2, 2.7] Let $z \in \mathrm{BV}([0,T];\mathcal{Z})$ be a BV solution of* (2.2.12)*. Then it holds for every $t \in J_z$ that*

$$\begin{aligned} \mathcal{J}(t,z(t) - \mathcal{J}(t,z(t_-)) &= -\Delta_{\mathfrak{p},\mathcal{J}}(z(t_-),z(t)), \\ \mathcal{J}(t,z(t_+)) - \mathcal{J}(t,z(t)) &= -\Delta_{\mathfrak{p},\mathcal{J}}(z(t),z(t_+)), \\ \mathcal{J}(t,z(t_+)) - \mathcal{J}(t,z(t_-)) &= -\Delta_{\mathfrak{p},\mathcal{J}}(z(t_-),z(t_+)). \end{aligned} \tag{J_{BV}}$$

$\mathfrak{p}$-parameterized Balanced Viscosity solutions

As we have seen on the previous pages, while it is possible to pass to the pointwise limit of a sequence of solutions of the viscously regularized systems, this will in general only yield a limit curve $z \in \mathrm{BV}([0,T];\mathcal{Z})$, exhibiting a substantial lack of differentiability. In this work, we will instead rely on a reparameterization technique that was introduced in [EM06]. The reparameterized solutions then allow for stronger (local) a priori estimates, thus yielding better differentiability after passing to the vanishing viscosity limit. What is more, the resulting notion of solutions has the following advantage: The jumps that might possibly appear in the solutions of the rate-independent system do not shrink down to a singular point in time, but rather, we obtain a jump curve in $[0,T] \times \mathcal{Z}$ that describes the transition between the two end points of the jump.

To this end, we choose an artificial arc-length parameter and transform the viscous system into an artificial time, so that the trajectory $t \mapsto (t, z_\varepsilon(t))$ is rewritten as $s \mapsto (\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s))$. There are several possible choices for the reparameterization, which, together with norm one chooses to define the viscous contribution to the augmented dissipation potential $\mathcal{R}_\varepsilon$, heavily influences the resulting notion of parameterized BV solutions and its properties. In this work, we choose to parameterize by the vanishing viscosity contact potential $\mathfrak{p}(\cdot,\cdot)$ defined in (2.2.19), which in our case of additive viscosity has the explicit representation $\mathfrak{p}(v,\xi) := \mathcal{R}(v) + \|v\|_{\mathcal{V}} \operatorname{dist}_{\mathcal{V}^*}(\xi, \partial\mathcal{R}(0))$, where

$$\operatorname{dist}_{\mathcal{V}^*}(\xi, \partial\mathcal{R}(0)) := \inf\{\|\xi - \eta\|_{\mathcal{V}^*} \,|\, \eta \in \partial\mathcal{R}\}$$

denotes the distance of an element $\xi \in \mathcal{V}^*$ to the set $\partial\mathcal{R}(0)$ w.r.t. the norm on $\mathcal{V}^*$, see [MRS12a, Rem. 3.1]. To be precise, we set

$$\begin{aligned} s_\varepsilon(t) &:= t + \int_0^t \mathfrak{p}(\dot z_\varepsilon(\tau), -\mathrm{D}_z\mathcal{J}(\tau, z_\varepsilon(\tau)))\mathrm{d}\tau, \quad S_\varepsilon := s_\varepsilon(T), \text{ and} \\ \hat t_\varepsilon &:= (s_\varepsilon)^{-1} : [0, S_\varepsilon] \to [0,T]. \end{aligned}$$

We then define $\hat z_\varepsilon := z_\varepsilon \circ \hat t_\varepsilon$ and consider the limit for vanishing viscosity (that is, for $\varepsilon \to 0$). We thus obtain a limit $S \in [0,\infty)$ of the artificial end times S_ε and a limiting pair $(\hat t, \hat z) : [0,S] \to [0,T] \times \mathcal{Z}$. In order to characterize the resulting solution set, we apply the same reparameterization to the energy dissipation balance $(\mathrm{E})_\varepsilon$ associated with the viscous problem (2.2.13) and also pass to the limit $\varepsilon \to 0$ here. In the finite dimensional setting, we arrive at a tuple $(\hat t, \hat z) \in W^{1,\infty}([0,S];\mathbb{R}) \times W^{1,\infty}(0,S;\mathcal{Z})$, fulfilling the limiting energy dissipation balance

$$\begin{aligned} \mathcal{J}(\hat t(s), \hat z(s)) + \int_0^s \mathcal{R}(\dot{\hat z}(r)) + \|\dot{\hat z}(r)\|_{\mathcal{V}} \operatorname{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{J}(\hat t(r), \hat z(r)), \partial\mathcal{R}(0))\,\mathrm{d}r \\ = \mathcal{J}(0, z_0) - \int_0^s \partial_r\mathcal{J}(\hat t(r), \hat z(r))\dot{\hat t}(r)\,\mathrm{d}r. \end{aligned} \tag{2.2.31}$$

In the infinite dimensional setting on the other hand, the curve $\hat z : [0,S] \to \mathcal{Z}$ is in general not differentiable almost everwhere on $[0,S]$. We will show, however, that $\hat z$ is classically differentiable w.r.t. $\|\cdot\|_{\mathcal{V}}$ almost everywhere on the set

$$G := \{s \in [0,S] \,|\, \operatorname{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{J}(\hat t(s), \hat z(s)), \partial\mathcal{R}(0)) > 0\},$$

and that for almost all $s \in [0,S]$, $\hat z$ possesses what is known as a **generalized metric derivative**, denoted by $\mathcal{R}[z'](s)$, see Prop. 3.2.2 for a definition and Appendix C for details. Instead of (2.2.31), for the infinite dimensional setting, we therefore obtain the following energy dissipation balance

$$\begin{aligned} \mathcal{J}(\hat t(s), \hat z(s)) + \int_0^s \mathcal{R}[\hat z'](r) + \|\dot{\hat z}(r)\|_{\mathcal{V}} \operatorname{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{J}(\hat t(r), \hat z(r)), \partial\mathcal{R}(0))\,\mathrm{d}r \\ = \mathcal{J}(0, z_0) - \int_0^s \partial_r\mathcal{J}(\hat t(r), \hat z(r))\dot{\hat t}(r)\,\mathrm{d}r. \end{aligned} \tag{2.2.32}$$

We then define normalized $\mathfrak{p}$-parameterized BV solutions as triples $(S, \hat t, \hat z)$ with a certain regularity such that (2.2.31), or (2.2.32), in infinite dimensions, and a normalization condition is fulfilled, see Definition 3.2.5. One advantage in the choice of the vanishing viscosity contact potential for reparameterization now lies in the fact that the limits $(\hat{t}, \hat{z})$ that are thus obtained are automatically normalized.

It can be shown both in the finite and in the infinite dimensional case (see Lemma 3.3.1), that a normalized $\mathfrak{p}$-parameterized BV solution $(S, \hat{t}, \hat{z})$ satisfies the differential inclusion

$$0 \in \partial\mathcal{R}(\hat{z}'(s)) + \partial\mathcal{R}_2(\lambda(s)\dot{\hat{z}}(s)) + \mathrm{D}_z\mathcal{J}(s, z(s)) \quad \text{f.a.a. } s \in [0, S] \tag{2.2.33}$$

for a measurable function $\lambda : [0, S] \to [0, \infty)$ with $\lambda(s) = 0$ on $(0, S) \setminus G$ and such that $\lambda(s)\dot{\hat{t}}(s) = 0$ a.e. in $[0, S]$. This allows for the following interpretation: Note that $\hat{t} : [0, S] \to [0, T]$ encodes the external time scale. Wherever $\dot{\hat{t}}(s) > 0$ on an interval $(s_1, s_2) \subset [0, S]$, it follows that $\lambda(s) = 0$, so that (2.2.33) describes the original rate-independent evolution. On the other hand, whenever $\dot{\hat{t}}(s) = 0$, the external time is frozen, and we obtain from the normalization condition that at the same time, $\dot{\hat{z}}(s) > 0$. Hence, this is seen as a jump in the external time frame. If $\lambda(s) > 0$, a viscous dissipation is active, which allows for the interpretation of a jump in the rate-independent system as a transition between two end points along a curve following a viscous regime. The vanishing viscosity analysis and the characterization of the resulting $\mathfrak{p}$-parameterized BV solutions in the infinite dimensional setting are carried out in great detail in Section 3.2.

For now, we point out that one has the freedom of choice both when it comes to the norm that characterizes the viscous augmentation $\mathcal{R}_{2,\varepsilon}$ and the reparameterization of the z_ε. Another popular choice of reparameterization is that by the $\|\cdot\|_\mathcal{V}$-arclength, i.e., one sets $\tilde{s}_\varepsilon(t) := t + \int_0^t \|\dot{z}_\varepsilon(\tau)\|_\mathcal{V} \mathrm{d}\tau$, $\tilde{t}_\varepsilon := (\tilde{s}_\varepsilon)^{-1} : [0, \widetilde{S}_\varepsilon] \to [0, T]$, and $\tilde{z}_\varepsilon := z_\varepsilon \circ \tilde{t}_\varepsilon$. This choice then leads to the definition of $\mathcal{V}$-parameterized BV solutions. However, it does not guarantee the normalization condition and not even non-degeneracy of the solutions, see the discussion around [Mie11, Lemma 4.12]. In this context, non-degeneracy means that $\dot{\tilde{t}}(s) + \|\dot{\tilde{z}}(s)\|_\mathcal{V} > 0$ for every $s \in [0, \widetilde{S}]$, and it is a crucial prerequisite to obtain the equivalent characterization via the differential inclusion (2.2.33). Indeed, it is easy to see that, if $-\mathrm{D}_z\mathcal{J}(s, \tilde{z}(s)) \notin \partial\mathcal{R}(0)$, but $\dot{\tilde{t}}(s) + \|\dot{\tilde{z}}(s)\|_\mathcal{V} = 0$, it is not possible to find $\lambda(s)$ such that (2.2.33) is fulfilled. This means that the interpretation and a priori estimates arising from (2.2.33) are not a priori available for $\mathcal{V}$-parameterized BV solutions.

Finally, we mention that there is an intricate interplay between the choice of viscosity norm and of reparameterization, resulting in vastly different regularities, a priori estimates and hence compactness properties of the resulting notion of parameterized BV solutions.

Equivalence between BV and $\mathfrak{p}$-parameterized BV solutions

To summarize, BV solutions are obtained by passing to the limit in the viscously regularized systems immediately, whereas $\mathfrak{p}$-parameterized BV solutions are obtained by applying a reparameterization before passing to the limit. In fact, as the terminology suggests, $\mathfrak{p}$-parameterized BV solutions can be interpreted as parameterized versions of BV solutions. This has been shown in [MRS12a, Thm.

5.8] in finite dimensions, and in [MRS16, Prop. 4.7] in a more general, infinite-dimensional setting. To be more precise, in the finite dimensional case, [MRS12a, Thm. 5.8] reads as follows. We refer to Proposition 3.3.3 for the infinite dimensional setting.

Proposition 2.2.12 (Equivalence of BV and $\mathfrak{p}$-parameterized BV solutions). *If $(S,\hat{t},\hat{z})$ is a $\mathfrak{p}$-parameterized BV solution of* (2.2.12), *then every curve $z:[0,T]\to\mathcal{Z}$ such that*

$$\forall\, t\in[0,T]:\quad z(t)\in\{\hat{z}(s)\,|\,\hat{t}(s)=t\} \tag{2.2.34}$$

is a BV solution of (2.2.12).

Conversely, if $z:[0,T]\to\mathcal{Z}$ is a BV solution of (2.2.12), *then there exists a $\mathfrak{p}$-parameterized BV solution $(S,\hat{t},\hat{z})$ of* (2.2.12) *satisfying* (2.2.34).

This equivalence also shines a light on how the choice of parameterization affects the resulting notion of parameterized BV solution: In fact, in order to obtain a similar result for $\mathcal{V}$-parameterized BV solutions, one has to impose the normalization condition, since it is not automatically obtained from the vanishing viscosity analysis. [Mie11, Cor. 4.22, Prop. 4.24] then assert that normalized $\mathcal{V}$-parameterized BV solutions are equivalent to a subset of BV solutions that exhibit a higher regularity in the sense that at all jump points, the left and right limits can be connected by transitions which are optimal in some sense, see [Mie11, Def. 4.21]. This suggests that choosing the $\mathfrak{p}$-parameterization results in a broader notion on solutions with a closer connection to the notion of BV solutions.

We close this section with the following application of Prop. 2.2.12: While it is one of the main results of this thesis that the reparameterized versions $\hat{z}_\varepsilon$ of the solutions z_ε of the viscously regularized systems converge to a $\mathfrak{p}$-parameterized BV solution with vanishing viscosity (cf. Theorem 3.2.6), the converse is not always true. The particular challenge stems from the fact that the viscously regularized systems allow for unique and smooth solutions, whereas the original (unregularized) system does not. The effect of this was already illustrated in the context of BV solutions in Example 2.2.10. There, we considered the RIS (2.2.27) with two distinct BV solutions, one of which had two jumps, whereas the other was continuous, and argued that the discontinuous one could not be obtained as a vanishing viscosity limit. Using the equivalence between BV and $\mathfrak{p}$-parameterized BV solutions from Prop. 2.2.12, we can now demonstrate that the $\mathfrak{p}$-parameterized BV solution that corresponds to the continuous BV solution of (2.2.27) can be obtained via the vanishing viscosity procedure described in the previous section, whereas the $\mathfrak{p}$-parameterized BV solution corresponding to the discontinuous BV solution of (2.2.27) cannot.

Example 2.2.13 (Continuation of Ex. 2.2.10). *We consider the RIS* (2.2.27) *for the end time $T:=10$, with $\mathcal{J}$ defined in* (2.3.1) *and $\mathcal{R}:\mathbb{R}\to[0,\infty)$ the absolute value function. For $\varepsilon>0$, let $z_\varepsilon:[0,T]\to\mathbb{R}$ defined according to* (2.2.30) *be the solution of* *the viscously regularized system* (2.2.29). *In order to determine*

$$\begin{aligned} s_\varepsilon(t) :&= t + \int_0^t \mathfrak{p}(\dot z_\varepsilon(\tau), -\mathrm{D}_z\mathcal{J}(\tau, z_\varepsilon(\tau)))\mathrm{d}\tau \\ &= t + \int_0^t |\dot z_\varepsilon(\tau)| + |\dot z_\varepsilon(\tau)| \operatorname{dist}(-\mathrm{D}_z\mathcal{J}(\tau, z_\varepsilon(\tau)), \partial\mathcal{R}(0))\mathrm{d}\tau, \end{aligned}$$

we first need to calculate $\operatorname{dist}(-\mathrm{D}_z\mathcal{J}(t, z_\varepsilon(t)), \partial\mathcal{R}(0))$ *for* $t \in [0, T]$. *Note that we have* $\partial\mathcal{R}(0) = [-1, 1]$, *and* $z_\varepsilon(t) < -2$ *for all* $t \in [0, T]$. *Since very similar calculations are carried out in detail in Section 2.3, we will here only give the following results:*

for $t \in [0, 3)$, *it holds that* $-\mathrm{D}_z\mathcal{J}(t, z_\varepsilon(t)) < -1$, *hence*
$\operatorname{dist}(-\mathrm{D}_z\mathcal{J}(t, z_\varepsilon(t)), \partial\mathcal{R}(0)) = -1 - (-\mathrm{D}_z\mathcal{J}(t, z_\varepsilon(t))) = t - z_\varepsilon(t) - 5,$
for $t \in [t_\varepsilon^*, 10]$, *it holds that* $-\mathrm{D}_z\mathcal{J}(t, z_\varepsilon(t)) \in [-1, 1]$, *hence*
$\operatorname{dist}(-\mathrm{D}_z\mathcal{J}(t, z_\varepsilon(t)), \partial\mathcal{R}(0)) = 0,$

and thus for $t \in [0, 3]$:

$$\begin{aligned} s_\varepsilon(t) &= t + \int_0^t |\dot z_\varepsilon(\tau)| + |\dot z_\varepsilon(\tau)| \operatorname{dist}(-\mathrm{D}_z\mathcal{J}(\tau, z_\varepsilon(\tau)), \partial\mathcal{R}(0))\mathrm{d}\tau \\ &= t + \underbrace{\int_0^t \dot z_\varepsilon(\tau) + \dot z_\varepsilon(\tau)\Big(\tau - z_\varepsilon(\tau) - 5\Big)\mathrm{d}\tau}_{\to t \text{ for } \varepsilon \to 0} \\ &\to 2t \text{ for } \varepsilon \to 0, \end{aligned}$$

for $t \in [3, t_\varepsilon^*]$:

$$\begin{aligned} s_\varepsilon(t) &= t + \int_0^3 \dot z_\varepsilon(\tau) + \dot z_\varepsilon(\tau)\Big(z_\varepsilon(\tau) + 5 - \tau\Big)\mathrm{d}\tau + \int_3^t 0\,\mathrm{d}\tau \\ &= t + \underbrace{\int_0^3 \dot z_\varepsilon(\tau) + \dot z_\varepsilon(\tau)\Big(\tau - z_\varepsilon(\tau) - 5\Big)\mathrm{d}\tau}_{\to 3 \text{ for } \varepsilon \to 0} \\ &\to t + 3 \text{ for } \varepsilon \to 0, \end{aligned}$$

for $t \in (t_\varepsilon^*, 10]$:

$$\begin{aligned} s_\varepsilon(t) &= t + \underbrace{\int_0^3 \dot z_\varepsilon(\tau) + \dot z_\varepsilon(\tau)\Big(\tau - z_\varepsilon(\tau) - 5\Big)\mathrm{d}\tau}_{\to 3 \text{ for } \varepsilon \to 0} + \underbrace{\int_{t_\varepsilon^*}^t -\dot z_\varepsilon(\tau)\mathrm{d}\tau}_{\to t-5 \text{ for } \varepsilon \to 0} \\ &\to 2t - 2 \text{ for } \varepsilon \to 0. \end{aligned}$$

In particular, we find that

$$S := \lim_{\varepsilon \to 0} s_\varepsilon(T) = 18. \tag{2.2.35}$$

Therefore, we extend the inverses $\hat{t}_\varepsilon := (s_\varepsilon)^{-1} : [0, s_\varepsilon(10)] \to [0,10]$ *constantly to the interval* $[0,S]$ *and then obtain pointwise convergence to*

$$\hat{t} : [0,S] \to [0,10], \quad \hat{t}(s) := \begin{cases} \frac{s}{2}, & \textit{for } 0 \le s \le 6, \\ s-3, & \textit{for } 6 < s \le 8, \\ \frac{s-2}{2}, & \textit{for } 8 < s \le 18. \end{cases} \tag{2.2.36}$$

The reparameterized curves $\hat{z}_\varepsilon := z_\varepsilon \circ \hat{t}_\varepsilon : [0, s_\varepsilon(T)] \to \mathbb{R}$ *are then constantly continued to* $[0,S]$ *as well, and converge pointwisely to*

$$\hat{z} : [0,S] \to \mathbb{R}, \quad \hat{z}(s) := \begin{cases} \frac{s}{2} - 5, & \textit{for } 0 \le s \le 6, \\ -2, & \textit{for } 6 < s \le 8, \\ 2 - \frac{s}{2}, & \textit{for } 8 < s \le 18. \end{cases} \tag{2.2.37}$$

In summary, the vanishing viscosity procedure described on the previous pages yields the $\mathfrak{p}$*-parameterized BV solution* $(S, \hat{t}, \hat{z})$ *of* (2.2.27) *that is definied via* (2.2.35) - (2.2.37). *In the context of Prop. 2.2.12, we find that the BV solution* z_2 *from* (2.2.28) *fulfills the condition*

$$\forall\, t \in [0,T]: \quad z_2(t) \in \{\hat{z}(s) \,|\, \hat{t}(s) = t\}.$$

Conversely, Prop. 2.2.12 also asserts that there must be a second $\mathfrak{p}$*-parameterized BV solution* $(\widetilde{S}, \widetilde{t}, \widetilde{z})$ *of* (2.2.27) *such that*

$$\forall\, t \in [0,T]: \quad z_1(t) \in \{\widetilde{z}(s) \,|\, \widetilde{t}(s) = t\}$$

holds for the discontinuous BV solution z_1 *from* (2.2.28). *Therefore, the RIS* (2.2.27) *has two distinct* $\mathfrak{p}$*-parameterized BV solutions, one of which cannot be obtained by a vanishing viscosity procedure.*

2.3 A 1-dimensional example

In the previous section, we presented different notions of solutions for rate-independent systems. Let us now discuss an example that highlights the differences between these concepts in a 1-dimensional setting. To this end, we assume that $T = 10$ and that the energy $\mathcal{J} : [0,10] \times \mathbb{R} \to \mathbb{R}$ is given in dependence of the time and state by

$$\mathcal{J}(t,z) := -t \cdot z + \begin{cases} \frac{1}{2}(z+4)^2, & \text{if } z \le -2 \\ 4 - \frac{1}{2}z^2, & \text{if } -2 \le z \le 2 \\ \frac{1}{2}(z-4)^2, & \text{if } z \ge 2, \end{cases} \tag{2.3.1}$$

and the dissipation potential $\mathcal{R} : \mathbb{R} \to [0,\infty)$ is the absolute value function, i.e., $\mathcal{R}(z) := |z|$ for $z \in \mathbb{R}$. Then the convex subdifferential of $\mathcal{R}$ is given by

$$\partial\mathcal{R}(z) = \begin{cases} -1, & \text{if } z < 0, \\ [-1,1], & \text{if } z = 0, \\ 1, & \text{if } z > 0, \end{cases} \tag{2.3.2}$$

and the derivative of $\mathcal{J}$ with respect to the state z by

$$\mathrm{D}_z\mathcal{J}(t,z) := -t + \begin{cases} z+4, & \text{if } z \le -2 \\ -z, & \text{if } -2 \le z \le 2 \\ z-4, & \text{if } z \ge 2. \end{cases}$$

This corresponds to the setting from Example 2.2.10, but with the external loading $\ell(t) := t$. We will now analyze the following RIS

$$-\mathrm{D}_z\mathcal{J}(t,z(t)) \in \partial\mathcal{R}(\dot{z}(t)), \quad t \in [0,10]; \quad z(0) = -5. \tag{2.3.3}$$

Let us first note that the energy $\mathcal{J}(t,\cdot)$ is **nonconvex** in the state variable, with two minima in $z = \pm 4$ and a local maximum in $z = 0$. We should therefore expect that the different notions of solutions of (2.3.3) have different properties.

In fact, let us first show that (2.3.3) **does not possess a solution in the differential sense**. For a proof by contradiction, let us assume that $z_d \in W^{1,1}((0,10);\mathbb{R})$ is a solution of (2.3.3) in the sense of Definition 2.2.3. Since z is differentiable almost everywhere, we can write (2.3.3) equivalently as the system of conditions (2.3.4) - (2.3.6)

$$\text{if } z(t) < -2 \text{ and } \dot{z}(t) > 0: \qquad z(t) = t-5 \tag{2.3.4a}$$
$$\text{if } z(t) < -2 \text{ and } \dot{z}(t) < 0: \qquad z(t) = t-3 \tag{2.3.4b}$$
$$\text{if } z(t) < -2 \text{ and } \dot{z}(t) = 0: \qquad t-5 \le z(t) \le t-3 \tag{2.3.4c}$$

$$\text{if } z(t) \in [-2,2] \text{ and } \dot{z}(t) > 0: \qquad z(t) = 1-t \tag{2.3.5a}$$
$$\text{if } z(t) \in [-2,2] \text{ and } \dot{z}(t) < 0: \qquad z(t) = -1-t \tag{2.3.5b}$$
$$\text{if } z(t) \in [-2,2] \text{ and } \dot{z}(t) = 0: \qquad -1-t \le z(t) \le 1-t \tag{2.3.5c}$$

$$\text{if } z(t) > 2 \text{ and } \dot{z}(t) > 0: \qquad z(t) = t+3 \tag{2.3.6a}$$
$$\text{if } z(t) > 2 \text{ and } \dot{z}(t) < 0: \qquad z(t) = t+5 \tag{2.3.6b}$$
$$\text{if } z(t) > 2 \text{ and } \dot{z}(t) = 0: \qquad t+3 \le z(t) \le t+5 \tag{2.3.6c}$$

Now, since z_d fulfills the initial condition $z_d(0) = -5$, we first turn to (2.3.4). Note that we must have $\dot{z}_d(0) \neq 0$, since z_d cannot constantly take the value -5 on an interval $[0,\delta]$ while still complying with the inequalities in (2.3.4c). Therefore, for some $\delta > 0$, we must have $z_d(t) = t-5$ on $[0,\delta)$, in order to fulfill the initial condition. Let us first assume that $\dot{z}_d(t) \neq 0$ on $[0,3)$. In this case, we must have $z_d(t) = t-5$ on $[0,3)$, and in order to extend z_d continuously to $(3,10]$, we must have that $z_d(3) = -2$.

This new initial condition can only be fulfilled in the cases (2.3.5a) or (2.3.5c). The condition $\dot{z}_d(3) = 0$ can only be fulfilled if z_d has a strict local maximum in $t = 3$, or if there is a $\delta > 0$ such that $z_d(t) \equiv -2$ on $[3,3+\delta]$. In the constant case, for every $t \in (3,3+\delta)$, we find that $z_d(t) = -2 > 1-t$, which is a contradiction to the inequality constraint in (2.3.5c). In the case of a strict local maximum, we must have that $z_d(t) < -2$ and $\dot{z}_d(t) < 0$ in an interval $(3, 3+\delta)$, which is impossible while complying with (2.3.4). Therefore, we must extend z_δ according to (2.3.5a). But then, on an interval $[3, 3+\delta]$ we must have at the same time that $\dot{z}_d(t) > 0$ and $\dot{z}_d(t) = -1 < 0$, which is impossible.

We now assume that there is a $t_* \in (0,3)$ such that $\dot{z}_d(t_*) = 0$. Again, this implies that either t_* is a strict local extremum of z_d, or that there is a $\delta > 0$ such that $z_d \equiv z_* \in (-5,-2)$ is constant on $(t_* - \delta, t_* + \delta)$. If t_* is a strict local extremum of z_d, then there must be a $\delta > 0$ such that $\dot{z}_d(t) < 0$ on $[t_*, t_* + \delta)$, but this is impossible while complying with (2.3.4b). In the constant case, it is impossible to extend z_d continuously beyond $t_* + \delta$, while complying with either (2.3.4a) or (2.3.4b). In conclusion, (2.3.3) does not possess a differential solution.

However, it is possible to solve (2.3.3) in weaker senses: It is noted in [MR15, Ex. 1.8.1] that (2.3.3) has **two distinct global energetic solutions** $z^{\pm}_{\mathrm{GES}}$, which are given by

$$z^{\pm}_{\mathrm{GES}} = \begin{cases} t-5, & \text{if } t \in [0,1), \\ \pm 4, & \text{if } t = 1, \\ t+3, & \text{if } t \in (1,10], \end{cases} \tag{2.3.7}$$

and both **uncountably many distinct local solutions and BV solutions**, since any choice of a value $z_* \in [-2, 6]$ yields a BV solution of (2.3.3) via

$$z_{\mathrm{BV}}(t) = \begin{cases} t-5, & \text{if } t \in [0,3), \\ z_*, & \text{if } t = 3, \\ t+3, & \text{if } t \in (3,10), \end{cases} \tag{2.3.8}$$

whereas for the local solutions, we may choose an arbitrary jump time $t_* \in [1,3]$ and attribute an arbitrary value $z_* \in [3+t_*, 3+t_*+\min\{2, 4\sqrt{t_*-1}\}]$ and thus obtain a local solution

$$z_{\mathrm{loc}}(t) = \begin{cases} t-5, & \text{if } t \in [0,t_*), \\ z_*, & \text{if } t \in [t_*, z_*-3], \\ t+3, & \text{if } t \in (z_*-3, 10]. \end{cases} \tag{2.3.9}$$

The GES and BV solutions are compared in Figure 2.2. For better understanding, the admissible set $\{(t,z) \,|\, -\mathrm{D}_z\mathcal{J}(t,z) \in \partial\mathcal{R}(0)\}$ is included in the graphic. As expected, the GES jump as soon as possible, whereas the BV solutions jump as late as necessary. The figure also illustrates that the non-smoothness of solutions of (2.3.3) is inherent in the definition of $\mathcal{R}$ and $\mathcal{J}$, since it is not possible to smoothly connect the initial point $(0,-5)$ with any end point $(T, z(T))$ without leaving the admissible set. Choosing the initial value $\tilde{z}_0 := 3$ however would yield $-\mathrm{D}_z\mathcal{J}(0,\tilde{z}_0) \in \partial\mathcal{R}(0)$ at the initial time $t = 0$ and allow for the differential solution $z_{\mathrm{diff}} : [0,T] \to \mathbb{R}$, $z_{\mathrm{diff}}(t) := t + 3$.

Figure 2.3 shows the graphs of three distinct local solutions, namely for the choices $(t_*, z_*) = (1.1, 4.3)$, for $(t_*, z_*) = (1.25, 5.5)$ and for $(t_*, z_*) = (2, 7)$. As the

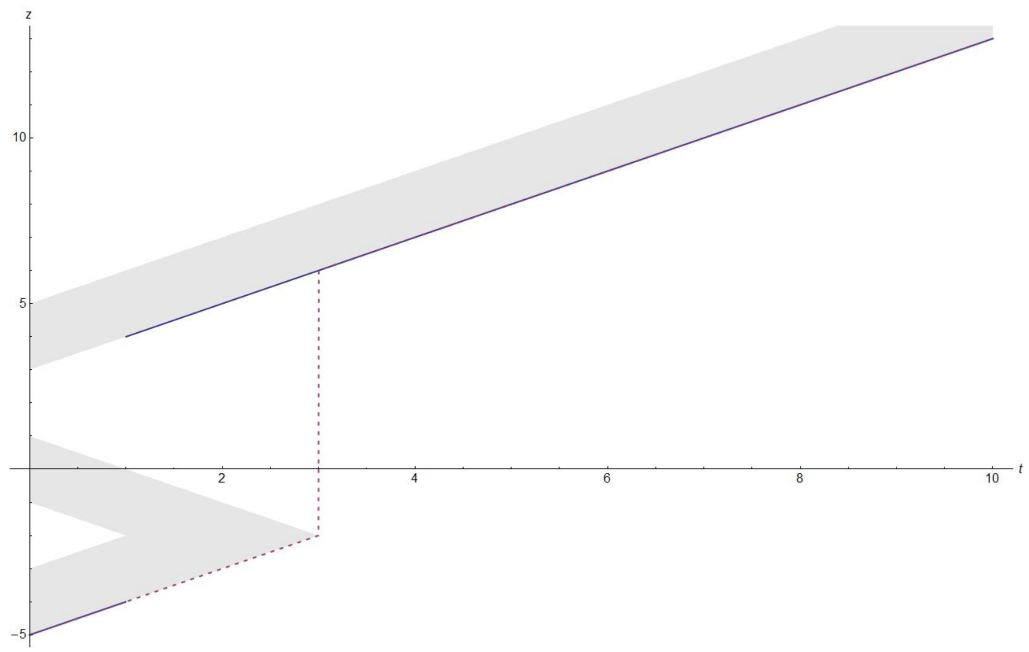


Figure 2.2: *Graphs of the GES (solid) and BV solutions (dashed), different choices for the value assigned at the jump time are possible, cf.* (2.3.7)-(2.3.8). *The admissible set* $\{(t,z) \mid -\mathrm{D}_z\mathcal{J}(t,z) \in \partial\mathcal{R}(0)\}$ *is shaded.*

figure illustrates, while local solutions allow for any choice of the time $t_* \in [1,3]$ at which the jump occurs, the possible choice for the corresponding value z_* in (2.3.9) is restricted in such a way that the graph of z_{loc} is always contained in the admissible set $\{(t,z) \mid -\mathrm{D}_z\mathcal{J}(t,z) \in \partial\mathcal{R}(0)\}$.

We now turn to the **viscous regularization** of (2.3.3), that is, for $0 < \varepsilon$, we consider

$$-\mathrm{D}_z\mathcal{J}(t,z_\varepsilon(t)) - \varepsilon\dot{z}_\varepsilon(t) \in \partial\mathcal{R}(\dot{z}_\varepsilon(t)), \quad t \in [0,10]; \quad z_\varepsilon(0) = -5, \tag{2.3.10}$$

which is equivalent to a system of ordinary differential equations similar to (2.3.4) - (2.3.6), but augmentend by the term $\varepsilon\dot{z}_\varepsilon(t)$. Keeping in mind the initial condition $z_\varepsilon(0) = -5$, we first obtain the the ODE

$$-(z_\varepsilon^1(t) + 4 - t) - \varepsilon\dot{z}_\varepsilon^1(t) = 1; \quad z_\varepsilon^1(0) = -5,$$

which has the solution

$$z_\varepsilon^1(t) := \exp\left(-\frac{t}{\varepsilon}\right)\varepsilon + t - 5 - \varepsilon.$$

This function increases monotonely until a time

$$t_\varepsilon > 0 \text{ for which } z_\varepsilon^1(t_\varepsilon) = -2, \tag{2.3.11}$$

and in order to extend z_ε beyond t_ε, we solve the ODE

$$-(-z_\varepsilon^2(t) - t) - \varepsilon\dot{z}_\varepsilon^2(t) = 1; \quad z_\varepsilon^2(t_\varepsilon) = -2$$

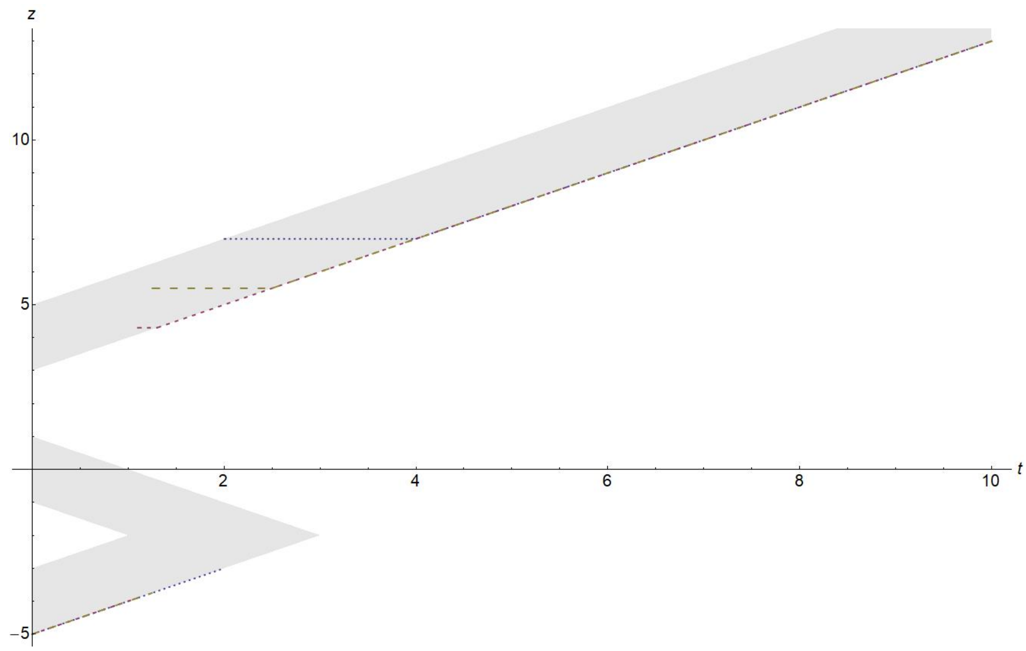


Figure 2.3: *Graphs of local solutions for different choices of jump time t_* and assigned value z_*, cf.* (2.3.9). *The admissible set $\{(t,z) \mid -\mathrm{D}_z\mathcal{J}(t,z) \in \partial\mathcal{R}(0)\}$ is shaded.*

by setting

$$z_\varepsilon^2(t) := \exp\left(\frac{t - t_\varepsilon}{\varepsilon}\right)(\varepsilon + t_\varepsilon - 3) - (\varepsilon + t - 1).$$

Again, this function increases monotonely until a time

$$r_\varepsilon > t_\varepsilon \text{ for which } z_\varepsilon^2(r_\varepsilon) = 2, \tag{2.3.12}$$

and in order to extend z_ε beyond r_ε, we solve the ODE

$$-(z_\varepsilon^3(t) - 4 - t) - \varepsilon \dot{z}_\varepsilon^3(t) = 1; \quad z_\varepsilon^3(r_\varepsilon) = 2$$

by setting

$$z_\varepsilon^3(t) := \exp\left(-\frac{t - r_\varepsilon}{\varepsilon}\right)(\varepsilon - r_\varepsilon - 1) + (t + 3 - \varepsilon).$$

All in all, for $\varepsilon > 0$ the curve

$$z_\varepsilon(t) = \begin{cases} z_\varepsilon^1(t), & \text{if } t \in [0, t_\varepsilon], \\ z_\varepsilon^2(t), & \text{if } t \in (t_\varepsilon, r_\varepsilon), \\ z_\varepsilon^3(t), & \text{if } t \in [r_\varepsilon, 10] \end{cases} \tag{2.3.13}$$

solves the viscously regularized system (2.3.10). Here, t_ε and r_ε are chosen according to (2.3.11) and (2.3.12), respectively, see (2.3.15) for an explicit representation. The graphs of these approximating curves are shown in Figure 2.4 for the values $\varepsilon = 0.01, 0.02, 0.03$. As expected, they converge to a BV solution z_{BV} for vanishing viscosity, i.e., for $\varepsilon \to 0$. To be precise, we have the following pointwise convergence:

Lemma 2.3.1. *Let $\mathcal{J}$ be defined as in* (2.3.1) *and $\mathcal{R}$ as in* (2.3.2). *Let further for $\varepsilon > 0$ the curves z_ε be defined as in* (2.3.13). *Then there exists a BV solution $\widetilde{z}\colon [0,10] \to \mathbb{R}$ such that for every $t \in [0,T]$, it holds that*

$$z_\varepsilon(t) \to \widetilde{z}(t) \text{ for } \varepsilon \to 0.$$

Further, there exists a curve $z : [0,10] \to \mathbb{R}$, which is a local solution and also a BV solution of (2.3.3) *such that for every $t \in [0,T] \setminus \{3\}$, it holds that*

$$z_\varepsilon(t) \to z(t) \text{ for } \varepsilon \to 0.$$

Proof. The crucial ingredient is the convergence of the points t_ε and r_ε that were defined in (2.3.11) and (2.3.12). Therefore, we will first prove that

$$t_\varepsilon \to 3 \quad \text{and} \quad r_\varepsilon \to 3 \quad \text{for } \varepsilon \to 0. \tag{2.3.14}$$

The equations (2.3.11) and (2.3.12) have the unique solutions

$$\begin{aligned} t_\varepsilon &:= 3 + \varepsilon + \varepsilon \cdot W_0(-\exp\Big(\frac{-\varepsilon - 3}{\varepsilon}\Big)), \text{ and} \\ r_\varepsilon &:= -\varepsilon - 1 - \varepsilon \cdot W_{-1}(\frac{-\varepsilon - t_\varepsilon + 3}{\varepsilon} \exp\Big(\frac{-1 - \varepsilon - t_\varepsilon}{\varepsilon}\Big)), \end{aligned} \tag{2.3.15}$$

where W_0 is the principal branch of the Lambert W function, and W_{-1} is its lower branch. Then the first of (2.3.14) follows immediately from the fact that $\lim_{r\to 0} W_0(r) = 0$, see [CGH+96]. For the second convergence, we define

$$g(\varepsilon) := \frac{-\varepsilon - t_\varepsilon + 3}{\varepsilon} \exp\Big(\frac{-1 - \varepsilon - t_\varepsilon}{\varepsilon}\Big)$$

and determine $L := \lim_{\varepsilon\to 0} \varepsilon W_{-1}(g(\varepsilon))$. First note that

$$\frac{-\varepsilon - t_\varepsilon + 3}{\varepsilon} = -1 - \frac{t_\varepsilon - 3}{\varepsilon} = -2 - W_0(-\exp\Big(\frac{-\varepsilon - 3}{\varepsilon}\Big)) \to -2 \text{ for } \varepsilon \to 0, \tag{2.3.16}$$

whereby

$$g(\varepsilon) \nearrow 0 \text{ for } \varepsilon \to 0.$$

For the branch W_{-1}, it holds that $W_{-1}(r) \to -\infty$ for $r \nearrow 0$, [CGH+96], and we apply L'Hôpital's rule several times and thus find that

$$L = \lim_{\varepsilon\to 0} \frac{W_{-1}(g(\varepsilon))}{\varepsilon^{-1}} = \lim_{\varepsilon\to 0} \frac{W_{-1}'(g(\varepsilon))g'(\varepsilon)}{-\varepsilon^{-2}} = \lim_{\varepsilon\to 0} \frac{g'(\varepsilon)}{-\varepsilon^{-2} g(\varepsilon)} \cdot \frac{W_{-1}(g(\varepsilon))}{1 + W_{-1}(g(\varepsilon))}, \tag{2.3.17}$$

where we used the formula $W_{-1}'(z) = \frac{W_{-1}(z)}{z(1+W_{-1}(z))}$ for the derivative of W_{-1}, cf. [CGH+96]. Now, from L'Hôpital's rule, we infer that the second factor in (2.3.17) converges to 1, as well as

$$\begin{aligned} \lim_{\varepsilon\to 0} \frac{g'(\varepsilon)}{-\varepsilon^{-2} g(\varepsilon)} &= \lim_{\varepsilon\to 0} \frac{\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big(\ln(-g(\varepsilon))\big)}{-\varepsilon^{-2}} = \lim_{\varepsilon\to 0} \frac{\ln(-g(\varepsilon))}{\varepsilon^{-1}} \\ &= \lim_{\varepsilon\to 0} \varepsilon \Big(\ln\Big(\frac{\varepsilon + t_\varepsilon - 3}{\varepsilon}\Big) + \frac{-1 - \varepsilon - t_\varepsilon}{\varepsilon}\Big) \\ &= -4, \end{aligned}$$

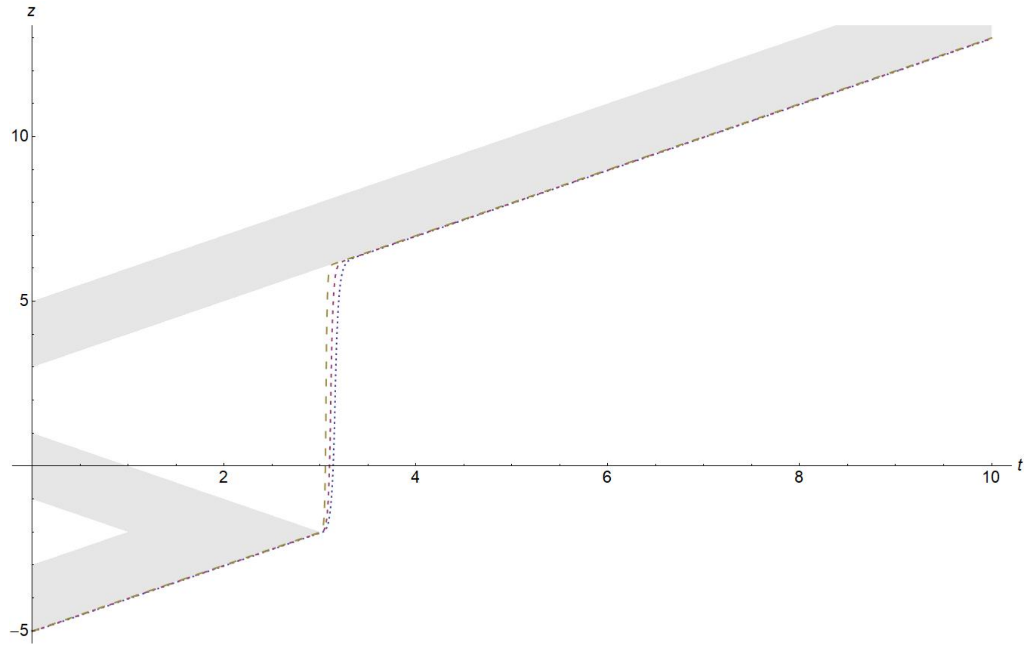


Figure 2.4: *Graphs of viscous approximations z_ε for $\varepsilon \in \{0.01, 0.02, 0.03\}$, cf.* (2.3.13). *The admissible set $\{(t,z) \,|\, -\mathrm{D}_z\mathcal{J}(t,z) \in \partial\mathcal{R}(0)\}$ is shaded.*

using (2.3.16) for the first summand, and the fact that $t_\varepsilon \to 3$ for the second summand. Returning to (2.3.17), we obtain that $L = -4$ and therefore

$$\lim_{\varepsilon\to 0} r_\varepsilon = \lim_{\varepsilon\to 0}\Big(-\varepsilon - 1 - \varepsilon \cdot W_{-1}(g(\varepsilon))\Big) = 3,$$

which is the second of (2.3.14). In fact, since z_ε^1 is strictly increasing in t, and $z_\varepsilon^1(3) < -2$ for $\varepsilon > 0$ small enough, we also find for all ε that are small enough that $3 < t_\varepsilon < r_\varepsilon$. All in all, we obtain for every $t \in [0,10]$

$$\text{if } t \le 3: \quad \lim_{\varepsilon\to 0} z_\varepsilon(t) = \lim_{\varepsilon\to 0} z_\varepsilon^1(t) = t - 5 \tag{2.3.18}$$

$$\text{if } t > 3: \quad \lim_{\varepsilon\to 0} z_\varepsilon(t) = \lim_{\varepsilon\to 0} z_\varepsilon^3(t) = t + 3, \tag{2.3.19}$$

which corresponds to the BV solution with the choice $z_* := 6$ in (2.3.8), and to the local solution with the choices $t_* := 3$ and $z_* := 6$ in (2.3.9). Choosing $z_* := -2$ in (2.3.8), we obtain a BV solution $\widetilde{z}$ for which we have pointwise convergence for all $t \in [0,10]$. ∎

In order to arrive at a $\mathfrak{p}$-**parameterized BV solution** of (2.3.3), we first need to calculate

$$\begin{aligned} s_\varepsilon(t) :&= t + \int_0^t \mathfrak{p}(\dot{z}_\varepsilon(\tau), -\mathrm{D}_z\mathcal{J}(\tau, z_\varepsilon(\tau))\mathrm{d}\tau \\ &= t + \int_0^t |\dot{z}_\varepsilon(\tau)| + |\dot{z}_\varepsilon(\tau)|\,\mathrm{dist}(-\mathrm{D}_z\mathcal{J}(\tau, z_\varepsilon(\tau)), \partial\mathcal{R}(0))\mathrm{d}\tau, \text{ and} \\ S_\varepsilon :&= s_\varepsilon(T), \end{aligned} \tag{2.3.20}$$

cf. p. 35. In this example, we have $\partial\mathcal{R}(0) = [-1,1]$, and we can determine the second integrand in (2.3.20) in the following way: Let first $t \in [0,t_\varepsilon]$. Then we have that $z_\varepsilon(t) \in [-5,-2]$, and

$$-\mathrm{D}_z\mathcal{J}(t,z_\varepsilon(t)) = -z_\varepsilon(t) - 4 + t = -\exp\Big(-\tfrac{t}{\varepsilon}\Big)\varepsilon + 1 - \varepsilon$$

is a strictly mononotely increasing function in t, with $-\mathrm{D}_z\mathcal{J}(0,z_\varepsilon(0)) = 1$. Thus, for $t \in [0,t_\varepsilon]$, we find that

$$\operatorname{dist}(-\mathrm{D}_z\mathcal{J}(t,z_\varepsilon(t)),\partial\mathcal{R}(0)) = -\mathrm{D}_z\mathcal{J}(t,z_\varepsilon(t)) - 1 = -z_\varepsilon^1(t) - 5 + t.$$

With similar arguments, we obtain for all $t \in [0,10]$

$$\operatorname{dist}(-\mathrm{D}_z\mathcal{J}(t,z_\varepsilon(t)),\partial\mathcal{R}(0)) = -\mathrm{D}_z\mathcal{J}(t,z_\varepsilon(t)) - 1 = \begin{cases} -z_\varepsilon^1(t) - 5 + t, & \text{if } t \in [0,t_\varepsilon], \\ z_\varepsilon^2(t) + t - 1, & \text{if } t \in (t_\varepsilon,r_\varepsilon), \\ -z_\varepsilon^3(t) + 3 + t, & \text{if } t \in [r_\varepsilon,10]. \end{cases}$$

We now proceed to determine the limits of $s_\varepsilon(t)$ for $\varepsilon \to 0$ for different values of $t \in [0,10]$. Let first $\varepsilon > 0$ and $t \in [0,3]$. Since all $t_\varepsilon > 3$, and taking into account the monotonicity of z_ε, we find that

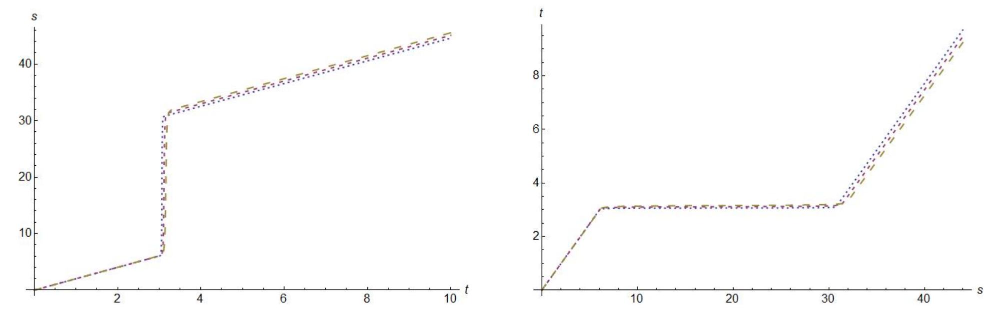
$$\begin{aligned}
s_\varepsilon(t) &= t + \int_0^t \mathfrak{p}(\dot z_\varepsilon^1(\tau), -\mathrm{D}_z\mathcal{J}(\tau,z_\varepsilon^1(\tau))\mathrm{d}\tau \\
&= t + \int_0^t |\dot z_\varepsilon^1(\tau)| + |\dot z_\varepsilon^1(\tau)|(-z_\varepsilon^1(\tau) - 5 + \tau)\mathrm{d}\tau \\
&= t + \int_0^t \dot z_\varepsilon^1(\tau)(-4 - z_\varepsilon^1(\tau) + \tau)\mathrm{d}\tau \\
&= t - 4\int_0^t \dot z_\varepsilon^1(\tau)\mathrm{d}\tau - \tfrac{1}{2}\int_0^t \frac{\mathrm{d}}{\mathrm{d}\tau}\Big(z_\varepsilon^1(\tau)\Big)^2\mathrm{d}\tau + \int_0^t z_\varepsilon^1(\tau)\tau\mathrm{d}\tau \\
&= t - 4(z_\varepsilon^1(t) - z_\varepsilon^1(0)) - \tfrac{1}{2}((z_\varepsilon^1(t))^2 - (z_\varepsilon^1(0))^2) + (z_\varepsilon^1(t)t - z_\varepsilon^1(0)0) - \int_0^t z_\varepsilon^1(\tau)\mathrm{d}\tau \\
&= t - 4(z_\varepsilon^1(t) + 5) - \tfrac{1}{2}((z_\varepsilon^1(t))^2 - 25) + z_\varepsilon^1(t)t - \int_0^t z_\varepsilon^1(\tau)\mathrm{d}\tau.
\end{aligned}$$

We can now apply the convergences from (2.3.14) and (2.3.18) and obtain that

$$s_\varepsilon(t) \to 2t,$$

as well as

$$\begin{aligned}
&\int_0^{t_\varepsilon} \mathfrak{p}(\dot z_\varepsilon(\tau), -\mathrm{D}_z\mathcal{J}(\tau,z_\varepsilon(\tau))\mathrm{d}\tau \\
&\qquad = -4(-2+5) - \tfrac{1}{2}(4-25) - 2t_\varepsilon - \int_0^{t_\varepsilon} z_\varepsilon^1(\tau)\mathrm{d}\tau \to 3.
\end{aligned} \tag{2.3.21}$$



(a) (b)

Figure 2.5: *The graphs of the curves s_ε from* (2.3.20) *(left) and of their inverses $\hat{t}_\varepsilon$ (right) for $\varepsilon \in \{0.01, 0.02, 0.03\}$.*

Let now $t > 3$. Then it holds that $3 < t_\varepsilon < r_\varepsilon < t$ for ε small enough. With similar calculations as for $t \leq 3$, we find that

$$\int_{r_\varepsilon}^{r_\varepsilon} \mathfrak{p}(\dot{z}_\varepsilon(\tau), -\mathrm{D}_z\mathcal{J}(\tau, z_\varepsilon(\tau))\mathrm{d}\tau = \int_{t_\varepsilon}^{r_\varepsilon} \dot{z}_\varepsilon^2(\tau)(z_\varepsilon^2(\tau) + \tau)\mathrm{d}\tau \to 12, \tag{2.3.22}$$

whereby

$$\begin{aligned} s_\varepsilon(t) = t + \int_0^{t_\varepsilon} \dot{z}_\varepsilon^1(\tau)(-4 - z_\varepsilon^1(\tau) + \tau)\mathrm{d}\tau + \int_{t_\varepsilon}^{r_\varepsilon} \dot{z}_\varepsilon^2(\tau)(z_\varepsilon^2(\tau) + \tau)\mathrm{d}\tau \\ + \int_{r_\varepsilon}^{t} \dot{z}_\varepsilon^3(\tau)(4 - z_\varepsilon^3(\tau) + t)\mathrm{d}\tau \\ \to 2t + 24. \end{aligned}$$

In particular, we find that

$$S := \lim_{\varepsilon \to 0} S_\varepsilon = \lim_{\varepsilon \to 0} s_\varepsilon(10) = 44.$$

We now define $\hat{t}_\varepsilon := s_\varepsilon^{-1} : [0, S_\varepsilon] \to [0, 10]$, which we extend constantly to $[0, S]$. We then find the pointwise limit $\hat{t}_\varepsilon \to (\hat{t} : [0, 44] \to [0, 10])$ for

$$\hat{t}(s) := \begin{cases} \frac{s}{2}, & \text{for } 0 \leq s \leq 6, \\ 3, & \text{for } 6 \leq s \leq 30, \\ \frac{s}{2} - 12, & \text{for } 30 \leq s \leq 44. \end{cases} \tag{2.3.23}$$

The graphs of the resulting curves s_ε and of their inverses $\hat{t}_\varepsilon$ are shown in Figure 2.5. Figure 2.6 shows the graphs of the reparameterized curves $\hat{z}_\varepsilon := z_\varepsilon \circ \hat{t}_\varepsilon$.

Finally, we aim to find the pointwise limit $\hat{z}$ of the reparameterized curves $\hat{z}_\varepsilon$, in order to obtain the last component of the $\mathfrak{p}$-parameterized BV solution $(S, \hat{t}, \hat{z})$. Let first $0 \leq s < 6$, and let $\tau_\varepsilon \in [0, 10]$ such that $s_\varepsilon(\tau_\varepsilon) = s < 6$. Since we know from (2.3.21) that $s_\varepsilon(t_\varepsilon) \to 6$, it must hold that $\hat{t}_\varepsilon(s) = \tau_\varepsilon < t_\varepsilon$ for ε small enough, and thus

$$\hat{z}_\varepsilon(s) = z_\varepsilon(\hat{t}_\varepsilon(s))) = z_\varepsilon^1(\hat{t}_\varepsilon(s)) \to \frac{s}{2} - 5 =: \hat{z}(s).$$

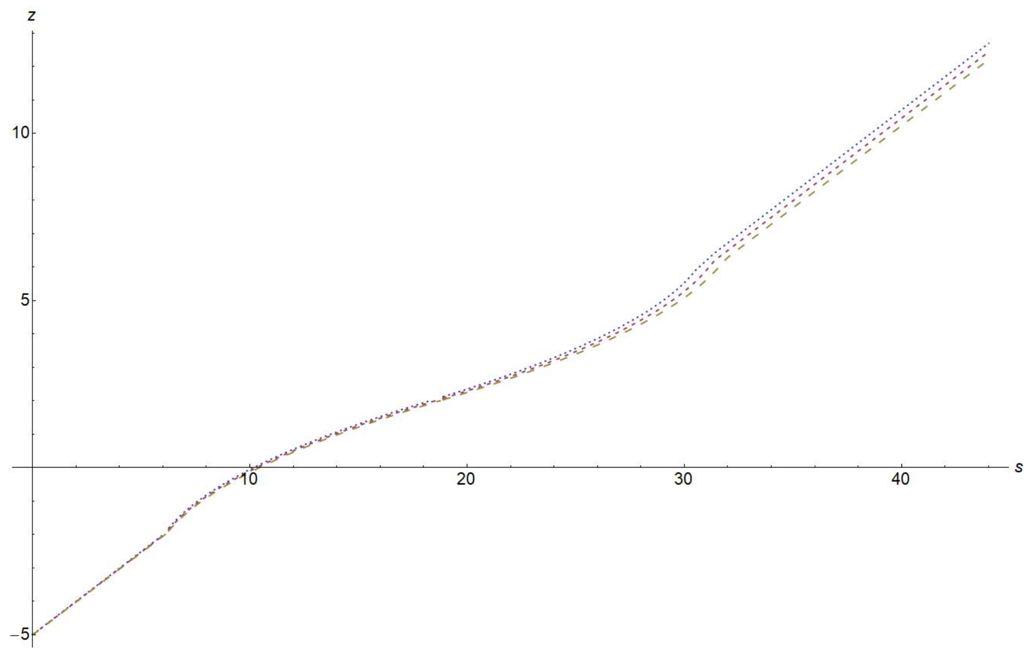


Figure 2.6: *Graphs of reparameterized viscous approximations* $\hat{z}_\varepsilon = z_\varepsilon \circ \hat{t}_\varepsilon$ *for values* $\varepsilon \in \{0.01, 0.02, 0.03\}$.

We proceed in a similar same way for $30 < s \le 44$: Again, let $\tau_\varepsilon \in [0,10]$ such that $s_\varepsilon(\tau_\varepsilon) = s > 30$. Since we know from (2.3.22) that $s_\varepsilon(r_\varepsilon) \to 18$, it must hold that $\hat{t}_\varepsilon(s) = \tau_\varepsilon > r_\varepsilon$ for ε small enough, and thus

$$\hat{z}_\varepsilon(s) = z_\varepsilon(\hat{t}_\varepsilon(s))) = z^3_\varepsilon(\hat{t}_\varepsilon(s)) \to (\frac{s}{2} - 12) + 3 = \frac{s}{2} - 9 =: \hat{z}(s).$$

It remains to determine $\hat{z}(s)$ for $6 \le s \le 30$. We achieve this by using the fact that $\mathfrak{p}$-parameterized BV solutions automatically fulfill the normalization condition (N) from Def. 3.2.5, which reads

$$\dot{\hat{t}}(s) + \mathcal{R}(\dot{\hat{z}}(s)) + |\dot{\hat{z}}(s)|\,\mathrm{dist}(-\mathrm{D}_z\mathcal{J}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)) = 1 \text{ for all } s \in [6,30]. \tag{2.3.24}$$

This allows us to derive an ODE for $\hat{z}$ in the following way: For $s = 6$, we have the initial condition $\hat{z}(6) = -2$, and for $s \in [6,30]$ such that $\hat{z}(s) \in [-2,2]$, we may determine

$$-\mathrm{D}_z\mathcal{J}(\hat{t}(s), \hat{z}(s)) = -\mathrm{D}_z\mathcal{J}(3, \hat{z}(s)) = \hat{z}(s) + 3 \in [1,5],$$

whereby $\mathrm{dist}(-\mathrm{D}_z\mathcal{J}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)) = -\mathrm{D}_z\mathcal{J}(3, \hat{z}(s)) - 1 = \hat{z}(s) + 2$. Thus, taking into account the monotonicity of $\hat{z}$, the normalization condition (2.3.24) reads

$$\dot{\hat{z}}(s) + \dot{\hat{z}}(s)(\hat{z}(s) + 2) = 1. \tag{2.3.25}$$

With the initial condition $\hat{z}(6) = -2$, the ODE (2.3.25) has the unique solution

$$\hat{z}^1(s) := \sqrt{2s - 11} - 3, \quad \text{for } 6 \le s \le 18,$$

Here, the restriction $s \leq 18$ is due to the fact that $\hat{z}^1$ takes the value $\hat{z}^1(18) = 2$, and thus, for $s > 18$, it holds that

$$-\mathrm{D}_z\mathcal{J}(\hat{t}(s), \hat{z}(s)) = -\mathrm{D}_z\mathcal{J}(3, \hat{z}(s)) = -\hat{z}(s) + 7 \in [1, 5].$$

For $s \geq 18$, we therefore obtain in the same way as before from the normalization condition (2.3.24) the ODE

$$\dot{\hat{z}}(s) + \dot{\hat{z}}(s)(-\hat{z}(s) + 6) = 1; \quad \hat{z}(18) = 2,$$

and therefore $\hat{z}$ must be given by

$$\hat{z}^2(s) := 7 - \sqrt{61 - 2s}, \quad \text{for } 18 \leq s \leq 30.$$

In summary, the triples $(S_\varepsilon, \hat{t}_\varepsilon, \hat{z}_\varepsilon)_{\varepsilon>0}$ converge to a $\mathfrak{p}$-parameterized BV solution $(S, \hat{t}, \hat{z})$ defined by

$$S = 44, \quad \hat{t} \text{ from (2.3.23)}, \quad \hat{z}(s) = \begin{cases} \frac{s}{2} - 5, & \text{for } 0 \leq s \leq 6, \\ \sqrt{2s - 11} - 3, & \text{for } 6 \leq s \leq 18, \\ 7 - \sqrt{61 - 2s}, & \text{for } 18 \leq s \leq 30, \\ \frac{s}{2} - 9, & \text{for } 30 \leq s \leq 44. \end{cases} \tag{2.3.26}$$

The graphs of $\hat{t}$ and $\hat{z}$ are included in Figure 2.7. The example illustrates the discussion around the alternative characterization (2.2.33) of $\mathfrak{p}$-parameterized BV solutions by means of a Lagrange parameter $\lambda : [0, S] \to [0, \infty)$: Wherever $\dot{\hat{t}}(s) > 0$, we find that

$$-\mathrm{D}_z\mathcal{J}(\hat{t}(s), \hat{z}(s)) \in \partial\mathcal{R}(\dot{\hat{z}}(s)) \tag{2.3.27}$$

holds true, i.e., it is possible to solve the RIS (2.3.3) locally. On the other hand, wherever $\dot{\hat{t}}(s) = 0$ (that is, for $6 < s < 30$), it holds that $-\mathrm{D}_z\mathcal{J}(\hat{t}(s), \hat{z}(s)) \notin \partial\mathcal{R}(0)$, and the Langrange parameter $\lambda(s) > 0$ is active here. In this interval, the curve $(\hat{t}(s), \hat{z}(s))$ can be interpreted as a transition between the two end points $z_-^* = -2$ and $z_+^* := 6$ of the jump that the BV solutions possess at the time $t_* = 3$. Instead of attributing an arbitrary value $z_* \in [-2, 6]$, the parameterized BV solution smoothly connects these two end points over the length of the interval $[6, 30]$. Conversely, in keeping with Proposition 2.2.12, we can obtain every BV solution of (2.3.3) by setting

$$\hat{s} : [0, 10] \to [0, 44], \quad \hat{s}(t) := \begin{cases} 2t, & \text{for } 0 \leq t \leq 3, \\ 2t + 24, & \text{for } 3 < t \leq 10, \end{cases}$$

choosing an arbitrary value $z_* \in \{\hat{z}(s) \,|\, \hat{t}(s) = 3\} = [\hat{z}(\hat{s}(3_-)), \hat{z}(\hat{s}(3_+))]$, and then defining

$$z(t) := \begin{cases} \hat{z}(\hat{s}(t)) = t - 5, & \text{for } t < 3, \\ z_*, & \text{for } t = 3, \\ \hat{z}(\hat{s}(t)) = t + 3, & \text{for } t > 3. \end{cases}$$

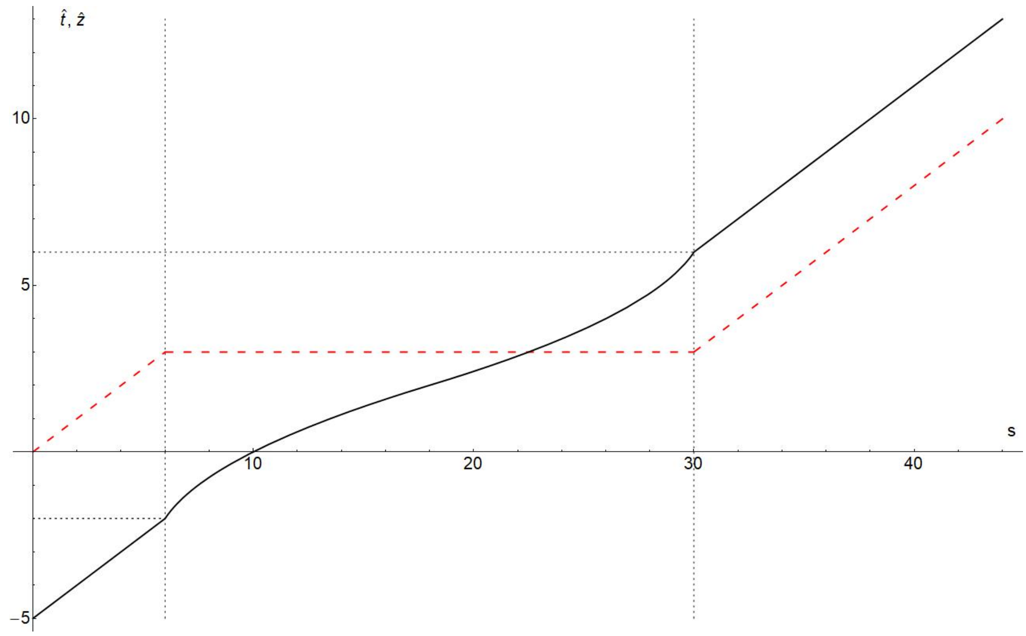


Figure 2.7: *Graphs of $\hat{t}$ (red, dashed) and $\hat{z}$ (black, solid), where $(S, \hat{t}, \hat{z})$ is the $\mathfrak{p}$-parameterized BV solution from* (2.3.26)*. Vertical lines are added at $s = 6$ and at $s = 30$ in order to highlight the transition into and out of the viscous regime when $\hat{t} \equiv 3$ is constant. The horizontal lines through $z^*_- := 2$ and $z^*_+ := 6$ mark the two end points of the jump that the BV solutions possess at the time $t_* = 3$*

2.4 The energy-dissipation framework

In this section, we specify the energy-dissipation framework in which the remainder of this work takes place and collect some basic properties thereof. Since we are interested in parameterized BV solutions, we will first consider a viscously regularized system, for which we show existence and regularity in Section 3.1. For the remaining part of this work, we restrict ourselves to the reduced problem as it was set out in (2.2.9)-(2.2.12). Again, we assume that $\mathcal{R}$ is state-independent, see (2.4.7). Note that the reduced energy $\mathcal{I}$ is given in dependence of an external load ℓ and the state z, that is $\mathcal{I}(\ell(t), z) = \mathcal{J}(t, z)$ for $\mathcal{J}$ from (2.2.11).

2.4.1 Standing assumptions

Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$, $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$ and $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ be Banach spaces, where $\mathcal{Z}$ and $\mathcal{V}$ are assumed to be reflexive, and such that

$$\mathcal{Z} \overset{c}{\hookrightarrow} \mathcal{V} \hookrightarrow \mathcal{X}. \tag{2.4.1}$$

We further assume that the norm $\|\cdot\|_{\mathcal{V}}$ is chosen in such a way that $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ is a uniformly convex Banach space with modulus of convexity of power 2, e.g., a Hilbert space, see Appendix G. We define the **energy functional** $\mathcal{I} : \mathcal{V}^* \times \mathcal{V} \to \mathbb{R}$

in dependence of the external load ℓ and the state z via

$$\mathcal{I}(\ell,z) := \begin{cases} \frac{1}{2}\langle Az,z\rangle_{\mathcal{Z}} + \mathcal{F}(z) - \langle \ell,z\rangle_{\mathcal{V}}, & \text{if } z\in\mathcal{Z} \\ \infty, & \text{if } z\in\mathcal{V}\setminus\mathcal{Z}. \end{cases} \tag{2.4.2}$$

Further, $A : \mathcal{Z}\to\mathcal{Z}^*$ is assumed to be a linear, self-adjoint, bounded and coercive operator from the space $\mathcal{Z}$ into its dual $\mathcal{Z}^*$, so that it holds

$$\forall\, z,w\in\mathcal{Z} : \langle Az,w\rangle_{\mathcal{Z}} = \langle Aw,z\rangle_{\mathcal{Z}}; \quad \exists\,\alpha>0 : \forall\, z\in\mathcal{Z} : \alpha\|z\|_{\mathcal{Z}}^2 \le \langle Az,z\rangle_{\mathcal{Z}}. \tag{2.4.3}$$

We denote by $\langle\cdot,\cdot\rangle_{\mathcal{Z}}$ and $\langle\cdot,\cdot\rangle_{\mathcal{V}}$ the duality pairing between $\mathcal{Z}^*$ and $\mathcal{Z}$ and between $\mathcal{V}^*$ and $\mathcal{V}$, respectively. We allow for a **non-convexity** $\mathcal{F}$ for which we assume that $\mathcal{F}\in C^2(\mathcal{Z},\mathbb{R}_{\ge 0})$ fulfills that

$$\mathrm{D}\mathcal{F} : \mathcal{Z}\to\mathcal{V}^* \text{ is weak-weakly continuous and } \mathrm{D}\mathcal{F}\in C^1(\mathcal{Z},\mathcal{V}^*), \tag{2.4.4a}$$

$$\exists C>0, q\ge 1\,\forall z,v\in\mathcal{Z} : \|\mathrm{D}^2\mathcal{F}(z)[v]\|_{\mathcal{V}^*} \le C(1+\|z\|_{\mathcal{Z}}^q)\|v\|_{\mathcal{Z}}. \tag{2.4.4b}$$

For $T>0$, the function $\ell : [0,T]\to\mathcal{V}^*$ shall characterize the **externally applied loads**. It is assumed that

$$\ell\in W^{1,\infty}(0,T;\mathcal{V}^*). \tag{2.4.5}$$

From now on, we shall always assume that the initial value z_0 and the load ℓ comply with the following compatibility and regularity condition

$$z_0\in\mathcal{Z}, \quad \ell\in W^{1,\infty}(0,T;\mathcal{V}^*), \quad \text{and } \mathrm{D}_z\mathcal{I}(\ell(0),z_0)\in\mathcal{V}^*. \tag{2.4.6}$$

Let further the **dissipation potential** $\mathcal{R} : \mathcal{V}\to[0,\infty)$ be convex, lower semicontinuous, positively homogeneous of degree one and satisfiying

$$\exists C,c>0\forall z\in\mathcal{Z} : c\|z\|_{\mathcal{X}} \le \mathcal{R}(z) \le C\|z\|_{\mathcal{X}}, \tag{2.4.7}$$

and let $\mathcal{R}$ be symmetric, i.e., for all $z\in\mathcal{Z}$ it holds that $\mathcal{R}(z)=\mathcal{R}(-z)$.[1] We will denote the subdifferential of $\mathcal{R}$ with respect to the $\mathcal{Z}-\mathcal{Z}^*$-duality by

$$\partial\mathcal{R}(z) := \{\eta\in\mathcal{Z}^* \,\big|\, \text{for all } w\in\mathcal{Z} : \langle\eta,w-z\rangle_{\mathcal{Z}} \le \mathcal{R}(w)-\mathcal{R}(z)\}. \tag{2.4.8}$$

The aim is to study an optimal control problem governed by the system

$$\text{f.a.a. } t\in[0,T] : \quad 0\in\partial\mathcal{R}(\dot z(t)) + \mathrm{D}_z\mathcal{I}(\ell(t),z(t)), \quad z(0)=z_0. \tag{2.4.9}$$

For $\varepsilon>0$ and given $z_0\in\mathcal{Z}$, we consider the regularized system

$$\text{f.a.a. } t\in[0,T] : \quad 0\in\partial\mathcal{R}_\varepsilon(\dot z_\varepsilon(t)) + \mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t)), \quad z_\varepsilon(0)=z_0, \tag{2.4.10}$$

where the **augmented dissipation potential** $\mathcal{R}_\varepsilon$ is defined as

$$\mathcal{R}_\varepsilon := \mathcal{R} + \mathcal{R}_{2,\varepsilon} \text{ with } \mathcal{R}_{2,\varepsilon}(v) := \frac{\varepsilon}{2}\|v\|_{\mathcal{V}}^2. \tag{2.4.11}$$

[1] The symmetry is only needed in the second step of the proof of Prop. F.1.

2.4.2 Some basic consequences

We first note that we have the following standard estimate for some constants $\lambda, c > 0$ independent of $z \in \mathcal{Z}$ and $\ell \in \mathcal{V}^*$:

$$\forall z \in \mathcal{Z}, \ell \in \mathcal{V}^* : \quad \|\partial_\ell \mathcal{I}(\ell, z)\|_{\mathcal{Z}^*}^2 = \|\ell\|_{\mathcal{Z}^*}^2 \leq \lambda(\mathcal{I}(\ell, z) + c\|\ell\|_{\mathcal{V}^*}^2). \tag{2.4.12}$$

The following property of $\mathcal{I}$ is often called **uniform subdifferentiability**, see, e.g., [MRS16, Section 2.1] or λ-**convexity**, see, e.g., [KZ21, Section 2].

Lemma 2.4.1. *For all $\rho > 0$ there exists a constant $M_\rho \geq 0$ such that $\mathcal{I}$ fulfills for all $z_1, z_2 \in \mathcal{Z}$ with $\|z_i\|_{\mathcal{Z}} \leq \rho$ and for all $\ell \in \mathcal{V}^*$ the estimate*

$$\mathcal{I}(\ell, z_1) \geq \mathcal{I}(\ell, z_2) + \langle \mathrm{D}_z \mathcal{I}(\ell, z_2), z_1 - z_2 \rangle_{\mathcal{Z}} + \frac{\alpha}{4}\|z_1 - z_2\|_{\mathcal{Z}}^2 - M_\rho \|z_1 - z_2\|_{\mathcal{V}} \mathcal{R}(z_1 - z_2), \tag{2.4.13}$$

where $\alpha > 0$ is the constant from (2.4.3).

The term "uniform subdifferentiability" can be explained as follows: The Fréchet subdifferential of $\mathcal{I}(\ell, \cdot) : \mathcal{Z} \to R$ at a point $z \in \mathcal{Z}$ is defined via

$$\partial_z \mathcal{I}(\ell, z) := \{\eta \in \mathcal{Z}^* \,|\, \forall w \xrightarrow{\mathcal{Z}} z : \mathcal{I}(\ell, w) \geq \mathcal{I}(\ell, z) + \langle \eta, w - z \rangle_{\mathcal{Z}} + o(\|w - z\|_{\mathcal{Z}}\}. \tag{2.4.14}$$

Thus, (2.4.13) prescribes a specific form for the remainder in (2.4.14). Also note that, if $M_\rho = 0$ in (2.4.13) for all $\rho > 0$, $\mathcal{I}(\ell, \cdot)$ would be strongly convex. This explains the notion of "λ-convexity" for $\lambda := M_\rho > 0$.

Proof of Lemma 2.4.1. We first show that for every $\rho > 0$ there exists $C_\rho > 0$ such that for all $z_1, z_2 \in \mathcal{Z}$ fulfilling $\|z_i\|_{\mathcal{Z}} \leq \rho$ it holds

$$\mathcal{F}(z_1) \geq \mathcal{F}(z_2) + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2 \rangle_{\mathcal{V}} - C_\rho \|z_1 - z_2\|_{\mathcal{Z}} \|z_1 - z_2\|_{\mathcal{V}}. \tag{2.4.15}$$

Indeed, let $\rho > 0$ and $\|z_1\|_{\mathcal{Z}}, \|z_2\|_{\mathcal{Z}} \leq \rho$. It holds

$$\begin{aligned}
\mathcal{F}(z_1) - \mathcal{F}(z_2) &= \int_0^1 \langle \mathrm{D}\mathcal{F}(z_2 + h(z_1 - z_2)), z_1 - z_2 \rangle_{\mathcal{V}} \,\mathrm{d}h \\
&= \int_0^1 \langle \mathrm{D}\mathcal{F}(z_2 + h(z_1 - z_2)) - \mathrm{D}\mathcal{F}(z_2), z_1 - z_2 \rangle_{\mathcal{V}} \,\mathrm{d}h + \int_0^1 \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2 \rangle_{\mathcal{V}} \,\mathrm{d}h \\
&= \int_0^1 \langle \int_0^1 \mathrm{D}^2\mathcal{F}(z_2 + \sigma h(z_1 - z_2))[h(z_1 - z_2)] \,\mathrm{d}\sigma, z_1 - z_2 \rangle_{\mathcal{V}} \,\mathrm{d}h + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2 \rangle_{\mathcal{V}} \\
&= \int_0^1 \int_0^1 \langle \mathrm{D}^2\mathcal{F}(z_2 + \sigma h(z_1 - z_2))[h(z_1 - z_2)], z_1 - z_2 \rangle_{\mathcal{V}} \,\mathrm{d}\sigma \,\mathrm{d}h + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2 \rangle_{\mathcal{V}} \\
&\geq -\int_0^1 \int_0^1 \|\mathrm{D}^2\mathcal{F}(z_2 + \sigma h(z_1 - z_2))[h(z_1 - z_2)]\|_{\mathcal{V}^*} \|z_1 - z_2\|_{\mathcal{V}} \,\mathrm{d}\sigma \,\mathrm{d}h + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2 \rangle_{\mathcal{V}} \\
&\geq -\int_0^1 \int_0^1 C(1 + \|z_2 + \sigma h(z_1 - z_2)\|_{\mathcal{Z}}^q) \|z_1 - z_2\|_{\mathcal{Z}} h \|z_1 - z_2\|_{\mathcal{V}} \,\mathrm{d}\sigma \,\mathrm{d}h + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2 \rangle_{\mathcal{V}} \\
&\geq -\int_0^1 \int_0^1 C(1 + (\rho + 2\sigma h \rho)^q) \|z_1 - z_2\|_{\mathcal{Z}} h \|z_1 - z_2\|_{\mathcal{V}} \,\mathrm{d}\sigma \,\mathrm{d}h + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2 \rangle_{\mathcal{V}} \\
&\geq -C(1 + (3\rho)^q) \|z_1 - z_2\|_{\mathcal{Z}} \|z_1 - z_2\|_{\mathcal{V}} + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2 \rangle_{\mathcal{V}},
\end{aligned}$$

where the second inequality is justified by (2.4.4b). Thus, (2.4.15) holds true and we proceed with the proof of (2.4.13). By means of Ehrling's Interpolation Lemma, [Wlo87], we obtain for every $\eta > 0$ a constant $C_\eta > 0$ such that

$$\begin{aligned}\mathcal{F}(z_1) - \mathcal{F}(z_2) &\geq -C_\rho \|z_1 - z_2\|_{\mathcal{Z}}(\eta\|z_1 - z_2\|_{\mathcal{Z}} + C_\eta \|z_1 - z_2\|_{\mathcal{X}}) + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2\rangle_{\mathcal{V}} \\ &= -C_\rho \eta \|z_1 - z_2\|_{\mathcal{Z}}^2 - C_\rho C_\eta \|z_1 - z_2\|_{\mathcal{X}} \|z_1 - z_2\|_{\mathcal{Z}} + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2\rangle_{\mathcal{V}} \\ &\geq -C_\rho \eta \|z_1 - z_2\|_{\mathcal{Z}}^2 - C_\rho C_\eta \Big(\frac{C_\eta}{4\eta}\|z_1 - z_2\|_{\mathcal{X}}^2 + \frac{\eta}{C_\eta}\|z_1 - z_2\|_{\mathcal{Z}}^2\Big) + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2\rangle_{\mathcal{V}} \\ &= -2C_\rho \eta \|z_1 - z_2\|_{\mathcal{Z}}^2 - \frac{C_\rho C_\eta^2}{4\eta}\|z_1 - z_2\|_{\mathcal{X}}^2 + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2\rangle_{\mathcal{V}}\end{aligned}$$

Thus, the boundedness of $\mathcal{R}$ by the norm on $\mathcal{X}$ and the boundedness of $\|\cdot\|_{\mathcal{X}}$ by $\|\cdot\|_{\mathcal{V}}$ allow us to conclude that there exist constants $\tilde{C}_{\rho,\eta}, C_{\rho,\eta} > 0$ such that it holds

$$\begin{aligned}\mathcal{F}(z_1) - \mathcal{F}(z_2) &\geq -2C_\rho \eta \|z_1 - z_2\|_{\mathcal{Z}}^2 - \tilde{C}_{\rho,\eta}\|z_1 - z_2\|_{\mathcal{X}}\|z_1 - z_2\|_{\mathcal{V}} + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2\rangle_{\mathcal{V}} \\ &\geq -2C_\rho \eta \|z_1 - z_2\|_{\mathcal{Z}}^2 - C_{\rho,\eta}\mathcal{R}(z_1 - z_2)\|z_1 - z_2\|_{\mathcal{V}} + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2\rangle_{\mathcal{V}}.\end{aligned}$$

Hence, we can now estimate for all $\|z_1\|_{\mathcal{Z}}, \|z_2\|_{\mathcal{Z}} \leq \rho$ as follows:

$$\begin{aligned}\mathcal{I}(\ell, z_1) - \mathcal{I}(\ell, z_2) &= \frac{1}{2}\langle Az_1, z_1\rangle_{\mathcal{Z}} - \frac{1}{2}\langle Az_2, z_2\rangle_{\mathcal{Z}} + \mathcal{F}(z_1) - \mathcal{F}(z_2) - \langle \ell, z_1 - z_2\rangle_{\mathcal{V}} \\ &= \frac{1}{2}\langle A(z_1 + z_2), z_1 - z_2\rangle_{\mathcal{Z}} + \mathcal{F}(z_1) - \mathcal{F}(z_2) - \langle \ell, z_1 - z_2\rangle_{\mathcal{V}} \\ &= \frac{1}{2}\langle A(z_1 - z_2), z_1 - z_2\rangle_{\mathcal{Z}} + \mathcal{F}(z_1) - \mathcal{F}(z_2) + \langle Az_2, z_1 - z_2\rangle_{\mathcal{Z}} - \langle \ell, z_1 - z_2\rangle_{\mathcal{V}} \\ &\geq \frac{\alpha}{2}\|z_1 - z_2\|_{\mathcal{Z}}^2 - C_\rho \eta \|z_1 - z_2\|_{\mathcal{Z}}^2 - C_{\rho,\eta}\mathcal{R}(z_1 - z_2)\|z_1 - z_2\|_{\mathcal{V}} \\ &\qquad + \langle \mathrm{D}\mathcal{F}(z_2), z_1 - z_2\rangle_{\mathcal{V}} + \langle Az_2, z_1 - z_2\rangle_{\mathcal{Z}} - \langle \ell, z_1 - z_2\rangle_{\mathcal{V}}.\end{aligned}$$

Now, choosing $\eta > 0$ so small that $C_\rho \eta \leq \frac{\alpha}{4}$ and identifying the last three terms on the right hand side as $\langle \mathrm{D}_z\mathcal{I}(z_2), z_1 - z_2\rangle_{\mathcal{Z}}$, we arrive at the desired estimate (2.4.13). ■

Another consequence of (2.4.15) is the following

Lemma 2.4.2. *$\mathcal{F}$ is continuous with respect to the weak topology on $\mathcal{Z}$.*

Proof. Let $(z_n)_{n\in\mathbb{N}} \subset \mathcal{Z}$ be a sequence and $z \in \mathcal{Z}$ such that $z_n \rightharpoonup z$ in $\mathcal{Z}$. Then, $z_n \to z$ strongly in $\mathcal{V}$ and there exists $\rho > 0$ such that $\sup_{n\in\mathbb{N}}\|z_n\|_{\mathcal{Z}} \leq \rho$ and, due to the continuity of $\mathrm{D}\mathcal{F}$, $\sup_{n\in\mathbb{N}}\|\mathrm{D}\mathcal{F}(z_n)\|_{\mathcal{V}^*} \leq \rho$. Therefore, (2.4.15) yields

$$\begin{aligned}\mathcal{F}(z) - \mathcal{F}(z_n) &\leq -\langle \mathrm{D}\mathcal{F}(z), z_n - z\rangle_{\mathcal{V}} + C_\rho\|z - z_n\|_{\mathcal{Z}}\|z - z_n\|_{\mathcal{V}} \\ &\leq (\|\mathrm{D}\mathcal{F}(z)\|_{\mathcal{V}^*} + 2\rho C_\rho)\|z - z_n\|_{\mathcal{V}},\end{aligned}$$

as well as

$$\begin{aligned}\mathcal{F}(z_n) - \mathcal{F}(z) &\leq -\langle \mathrm{D}\mathcal{F}(z_n), z - z_n\rangle_{\mathcal{V}} + C_\rho\|z - z_n\|_{\mathcal{Z}}\|z - z_n\|_{\mathcal{V}} \\ &\leq (\rho + 2\rho C_\rho)\|z - z_n\|_{\mathcal{V}},\end{aligned}$$

so that $|\mathcal{F}(z_n) - \mathcal{F}(z)| \to 0$ with $n \to \infty$. ■

From (2.4.4b), we obtain the following estimate, for the proof of which we refer to [Kne19, Lemma 1.1].

Lemma 2.4.3. *[Kne19, Lemma 1.1] For every $\rho > 0$ and $\eta > 0$ there exists $C_{\rho,\eta} > 0$ such that for all $z_1, z_2 \in \mathcal{Z}$ with $\|z_i\|_{\mathcal{Z}} \leq \rho$ it holds*

$$|\langle \mathrm{D}\mathcal{F}(z_1) - \mathrm{D}\mathcal{F}(z_2), z_1 - z_2\rangle_{\mathcal{V}}| \leq \eta \|z_1 - z_2\|_{\mathcal{Z}}^2 + C_{\rho,\eta}\mathcal{R}(z_1 - z_2)\|z_1 - z_2\|_{\mathcal{V}}. \quad (2.4.16)$$

Since the notion of solution for the regularized system (2.4.10) is based on the energy dissipation principle (EDP) (cf. Section 2.2.1), we will at this point also determine the convex conjugate of $\mathcal{R}_\varepsilon$ with respect to the $\mathcal{Z} - \mathcal{Z}^*$-duality which is defined by

$$\mathcal{R}_\varepsilon^*(\eta) := \sup\{\langle \eta, z\rangle_{\mathcal{Z}} - \mathcal{R}_\varepsilon(z) : z \in \mathcal{Z}\} \quad \text{for } \eta \in \mathcal{Z}^*. \quad (2.4.17)$$

The subdifferential and the conjugate of $\mathcal{R}_\varepsilon$ can be identified as follows:

Lemma 2.4.4. *For $\mathcal{R}_\varepsilon$ defined as in 2.4.11, there holds*

(i) $\partial\mathcal{R}_\varepsilon(z) = \partial\mathcal{R}(z) + \partial\mathcal{R}_{2,\varepsilon}(z) \subset \mathcal{V}^*$ *for all* $z \in \mathcal{Z}$, *where*

- $\partial\mathcal{R}(z) = \{\sigma \in \mathcal{V}^* \,|\, \text{for all } w \in \mathcal{V} : \langle \sigma, w - z\rangle_{\mathcal{V}} \leq \mathcal{R}(w) - \mathcal{R}(z)\}$,
- $\partial\mathcal{R}_{2,\varepsilon}(z) = \{\varepsilon\sigma \in \mathcal{V}^* \,|\, \|\sigma\|_{\mathcal{V}^*} = \|z\|_{\mathcal{V}} \text{ and } \langle \sigma, z\rangle_{\mathcal{V}} = \|z\|_{\mathcal{V}}^2\}$;

(ii) $\partial\mathcal{R}(0)$ *is a bounded subset of* $\mathcal{V}^*$;

(iii) for all $z \in \mathcal{Z}$, *it holds that* $\partial\mathcal{R}(z) \subset \partial\mathcal{R}(0)$;

(iv) for all $z \in \mathcal{Z}$, *it holds that, if* $\eta \in \partial\mathcal{R}(z)$, *then* $\langle \eta, z\rangle_{\mathcal{Z}} = \mathcal{R}(z)$;

(v) $\mathcal{R}^*$ *is the indicator function*

$$\mathcal{R}^*(\sigma) = \delta_{\partial\mathcal{R}(0)}(\sigma) = \begin{cases} 0, & \text{for } \sigma \in \partial\mathcal{R}(0) \\ \infty, & \text{for } \sigma \in \mathcal{Z}^* \setminus \partial\mathcal{R}(0); \end{cases}$$

(vi) $\mathcal{R}_\varepsilon^*(\sigma) = \inf_{\eta\in\partial\mathcal{R}(0)} \mathcal{R}_{2,\varepsilon}^*(\sigma - \eta) = \min_{\eta\in\partial\mathcal{R}(0)} \mathcal{R}_{2,\varepsilon}^*(\sigma - \eta)$ *with*

$$\mathcal{R}_{2,\varepsilon}^*(\sigma) = \begin{cases} \frac{1}{2\varepsilon}\|\sigma\|_{\mathcal{V}^*}^2 & \text{if } \sigma \in \mathcal{V}^* \\ \infty & \text{if } \sigma \in \mathcal{Z}^* \setminus \mathcal{V}^*. \end{cases}$$

Proof. We obtain the first assertion in (i) from [AE06, Cor. IV.3.6]. Thanks to Lemma A.1, the subdifferential of $\mathcal{R}_{2,\varepsilon}$ with respect to the $\mathcal{Z} - \mathcal{Z}^*$-duality coincides with that with respect to the $\mathcal{V} - \mathcal{V}^*$-duality which we can determine with help of [AE06, Prop. IV.3.10].

The boundedness asserted in (ii) holds true, since for every $\eta \in \partial\mathcal{R}(0)$, we can find an element $v \in \mathcal{V}$ such that $\|v\|_{\mathcal{V}} = 1$ and $\langle \eta, v\rangle_{\mathcal{V}} = \|\eta\|_{\mathcal{V}^*}$. Hence, we have $\|\eta\|_{\mathcal{V}^*} = \langle \eta, v\rangle_{\mathcal{V}} \leq \mathcal{R}(v) \leq C\|v\|_{\mathcal{X}} \leq CC_{\mathcal{V}\hookrightarrow\mathcal{X}}\|v\|_{\mathcal{V}} = CC_{\mathcal{V}\hookrightarrow\mathcal{X}}$.

For the proof of (iii), let $\eta \in \partial\mathcal{R}(z)$ for some $z \in \mathcal{V}$, i.e., for all $w \in \mathcal{V}$, we have $\langle \eta, w\rangle_{\mathcal{V}} - \mathcal{R}(w) \le \langle \eta, z\rangle_{\mathcal{V}} - \mathcal{R}(z) =: C(z)$. Thus, if there was $w \in \mathcal{V}$ such that $\langle \eta, w\rangle_{\mathcal{V}} - \mathcal{R}(w) > 0$, it would follow for $N \to \infty$ that

$$\langle \eta, Nw\rangle_{\mathcal{V}} - \mathcal{R}(Nw) = N(\langle \eta, w\rangle_{\mathcal{V}} - \mathcal{R}(w)) \to \infty$$

in contradiction to the boundedness of the left hand side by $C(z)$.

To see (iv), let $v \in \mathcal{Z}$ be arbitrary. Then it holds for all $\eta \in \partial\mathcal{R}(v)$ and $\lambda \in (0,1)$ that

$$\mathcal{R}(v) = \frac{\mathcal{R}(v+\lambda v) - \mathcal{R}(v)}{\lambda} \ge \langle \eta, v\rangle_{\mathcal{V}} \ge \frac{\mathcal{R}(v-\lambda v) - \mathcal{R}(v)}{-\lambda} = \mathcal{R}(v),$$

where we have used the 1-homogeneity of $\mathcal{R}$ in the first identity.

The formula (v) holds true since for every $\eta \in \partial\mathcal{R}(0)$, it follows that we have $\mathcal{R}^*(\eta) = \langle \eta, 0\rangle_{\mathcal{Z}} - \mathcal{R}(0) = 0$, whereas for $\eta \in \mathcal{Z}^* \setminus \partial\mathcal{R}(0)$, there exists $z \in \mathcal{Z}$ such that $\langle \eta, z\rangle_{\mathcal{Z}} - \mathcal{R}(z) =: r > 0$. Thus, for $N \to \infty$, we have that $\langle \eta, Nz\rangle_{\mathcal{Z}} - \mathcal{R}(Nz) = Nr \to \infty$, and thus $\mathcal{R}^*(\eta) = \infty$.

In order to prove the formula (vi) for the conjugate functional, we apply [IT79, Thm. 3.3.4.1] and infer that we have $\mathcal{R}_\varepsilon^* = \mathcal{R}^* \oplus \mathcal{R}_{2,\varepsilon}^*$, which is defined by $(f_1 \oplus f_2)(y) := \inf\{f_1(y_1) + f_2(y_2) \,|\, y_1 + y_2 = y\}$. Using (v), it follows for all $\sigma \in \mathcal{Z}^*$ that

$$\mathcal{R}_\varepsilon^*(\sigma) = \inf\{\mathcal{R}^*(\eta) + \mathcal{R}_{2,\varepsilon}^*(\sigma - \eta) \,|\, \eta \in \mathcal{Z}^*\} = \inf\{\mathcal{R}_{2,\varepsilon}^*(\sigma - \eta) \,|\, \eta \in \partial\mathcal{R}(0)\}.$$

Now, we determine $\mathcal{R}_{2,\varepsilon}^*$ by means of Lemma A.1 and obtain

$$\mathcal{R}_{2,\varepsilon}^*(\sigma) = \begin{cases} \mathcal{R}_{2,\varepsilon}^{*,\mathcal{V}}(\sigma), & \text{if } \sigma \in \mathcal{V}^* \\ \infty, & \text{if } \sigma \in \mathcal{Z}^* \setminus \mathcal{V}^*. \end{cases}$$

Furthermore, for arbitrary $\sigma \in \mathcal{V}^*$ and $v \in \mathcal{V}$, it holds

$$\langle \sigma, v\rangle_{\mathcal{V}} - \frac{\varepsilon}{2}\|v\|_{\mathcal{V}}^2 \le \|\sigma\|_{\mathcal{V}^*}\|v\|_{\mathcal{V}} - \frac{\varepsilon}{2}\|v\|_{\mathcal{V}}^2 \le \frac{1}{2\varepsilon}\|\sigma\|_{\mathcal{V}^*}^2.$$

On the other hand, since $\mathcal{V}$ is assumed to be reflexive, we find for every $\sigma \in \mathcal{V}^*$ an element $v_\sigma \in \mathcal{V}$ such that $\langle \sigma, v_\sigma\rangle_{\mathcal{V}} = \|\sigma\|_{\mathcal{V}^*}$ and $\|v_\sigma\|_{\mathcal{V}} = 1$. For $w_\sigma := \frac{\|\sigma\|_{\mathcal{V}^*}}{\varepsilon} v_\sigma$ it follows that

$$\langle \sigma, w_\sigma\rangle_{\mathcal{V}} - \frac{\varepsilon}{2}\|w_\sigma\|_{\mathcal{V}}^2 = \frac{1}{\varepsilon}\|\sigma\|_{\mathcal{V}^*}^2 - \frac{\varepsilon}{2}\frac{\|\sigma\|_{\mathcal{V}^*}^2}{\varepsilon^2} = \frac{1}{2\varepsilon}\|\sigma\|_{\mathcal{V}^*}^2,$$

proving the formula for $\mathcal{R}_{2,\varepsilon}^{*,\mathcal{V}}$. Now, for $\sigma \in \mathcal{V}^*$ let $(\eta_n)_{n\in\mathbb{N}} \subset \partial\mathcal{R}(0)$ be a sequence such that

$$\mathcal{R}_{2,\varepsilon}^*(\sigma - \eta_n) \to \inf_{\eta \in \partial\mathcal{R}(0)} \mathcal{R}_{2,\varepsilon}^*(\sigma - \eta) =: I.$$

If we have $I < \infty$, this implies that $(\eta_n)_{n\in\mathbb{N}} \subset \mathcal{V}^*$ together with the estimate $\sup_{n\in\mathbb{N}} \|\eta_n\|_{\mathcal{V}^*} \le \sup_{n\in\mathbb{N}} \|\eta_n - \sigma\|_{\mathcal{V}^*} + \|\sigma\|_{\mathcal{V}^*} < \infty$. Thus, there exist a subsequence and a limit $\eta \in \mathcal{V}^*$ such that $\eta_{n_k} \overset{*}{\rightharpoonup} \eta$ in $\mathcal{V}^*$, and for all $z \in \mathcal{Z}$, it holds

$$\langle \eta, z\rangle_{\mathcal{Z}} \overset{\text{def.}}{=} \langle \eta, z\rangle_{\mathcal{V}} = \lim_{k\to\infty} \langle \eta_{n_k}, z\rangle_{\mathcal{V}} \overset{\text{def.}}{=} \lim_{k\to\infty} \langle \eta_{n_k}, z\rangle_{\mathcal{Z}} \le \mathcal{R}(z).$$

Therefore, $\eta \in \partial\mathcal{R}(0)$ and

$$I \le \mathcal{R}^*_{2,\varepsilon}(\sigma - \eta) \le \liminf_{k\to\infty} \mathcal{R}^*_{2,\varepsilon}(\sigma - \eta_{n_k}) = I,$$

so that the infimum is indeed a minimum. The same is true if $\mathcal{R}^*_\varepsilon(\sigma) = \infty$, since then we have $\mathcal{R}^*_{2,\varepsilon}(\sigma - \eta) = \infty$ for all $\eta \in \partial\mathcal{R}(0)$. ∎

As a consequence of Lemma 2.4.4, we will no longer specify with respect to which duality we determine the subdifferential and the conjugate functional, since the subdifferentials coincide as subsets of $\mathcal{V}^*$ and the conjugate functional takes finite values only on the subspace $\mathcal{V}^* \subset \mathcal{Z}^*$.

2.4.3 A rate-independent ferrolectric model

The abstract semilinear setting presented above can be applied to a rate-independent ferroelectric model such as in [Kne19, Section 5.3] (which is a simplified version of the model from [SKT+15]). The model uses as variables a displacement field $u : \Omega \to \mathbb{R}^d$, where $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$ is a bounded domain with Lipschitz boundary, the strain field $e(u) := \mathrm{sym}(\nabla u)$, the electric potential $\phi : \Omega \to \mathbb{R}$, the electric field $E = -\nabla\phi$, the electric displacement $D : \Omega \to \mathbb{R}^d$ and the sponatenous polarization $P : \Omega \to \mathbb{R}^d$. The energy functional $\mathcal{E}$ is then given in terms of a free energy density Ψ depending on $e(u)$, D, P and ∇P, as well as an external load ℓ. Assuming vanishing Dirichlet boundary conditions on $\partial\Omega$ for u and ϕ, the function spaces are then chosen as

$$\mathcal{U} := H^1_0(\Omega,\mathbb{R}^d)\times L^2_D(\Omega,\mathbb{R}^d), \quad \mathcal{Z} := H^1(\Omega,\mathbb{R}^d), \quad \mathcal{V} := L^2(\Omega,\mathbb{R}^d), \quad \mathcal{X} := L^1(\Omega,\mathbb{R}^d),$$

where $L^2_D(\Omega,\mathbb{R}^d) := \{D \in L^2(\Omega,\mathbb{R}^d) | \forall\phi \in H^1_0(\Omega,\mathbb{R}^d) : \int_\Omega D\cdot\nabla\phi\,\mathrm{d}x = 0\}$. For displacements $(u,D) \in \mathcal{U}$, $P \in \mathcal{Z}$ and $\ell \in C^1([0,T],(\mathcal{U}^*\times\mathcal{V}^*))$, the energy functional $\mathcal{E} : [0,T]\times\mathcal{U}\times\mathcal{Z} \to \mathbb{R}$ is then given by

$$\mathcal{E}(t,u,D,P) := \int_\Omega \Psi(e(u),D,P,\nabla P)\mathrm{d}x - \langle \ell(t),(u,D,P)^T\rangle,$$

where Ψ is quadratic and convex in $e(u)$, D and ∇P and nonconvex in P. The dissipation potential $\mathcal{R} : \mathcal{X} \to [0,\infty)$ is defined as

$$\mathcal{R}(v) = \gamma\|v\|_{L^1(\Omega)}$$

for a constant $\gamma > 0$. The ferroelectric model then reads: Find $(u,D) : [0,T] \to \mathcal{U}$ and $P : [0,T] \to \mathcal{Z}$ such that $P(0) = P_0$ and

$$\begin{aligned} &0 = \mathrm{D}_u\mathcal{E}(t,u,D,P), \quad 0 = \mathrm{D}_D\mathcal{E}(t,u,D,P) \\ &0 \in \partial\mathcal{R}(\dot{P}(t)) + \mathrm{D}_P\mathcal{E}(t,u,D,P) \end{aligned}$$

Here, P plays the role of z in (2.4.9). As the discussion in [Kne19] shows, under reasonable assumptions on Ψ, this model satisfies all the assumptions made in this section, whereby the results presented in this dissertation allow to solve an optimal control problem in which, e.g., the end time polarization $P(T)$ is prescribed, cf. p. 103f.

2.4.4 Some notes on the structure of existence proofs

As was already pointed out in [MR15, Table 2.1], existence proofs for solutions of rate-independent systems usually consist of similar steps, in which similar kinds of arguments are used. Since the existence proofs can be quite long, what follows is a very schematic overview over these steps. While they are here numbered to allow references to each other, it should be noted that for technical reasons, they will not always occur in that exact order, and will not always have the same number in the following sections. In particular, step 4 may be spaced out throughout a proof in order to supply the specific convergence and regularity results needed to proceed. This being said, the generic steps are:

- **Step 0: Construction of approximating sequences and a priori estimates**

  The approximating sequences may be constructed via a time discretization scheme (see Section 3.1), a viscous regularization (see Section 3.2), or by choosing a minimizing sequence for an objective functional (see Chapter 5). Depending on the intricacy of the construction, this may actually happen outside of the existence proof.

- **Step 1: Selection of convergent subsequences**

  The natural next step is to use these a priori estimates to extract converging subsequences from the approximating sequences and obtain first regularity results.

- **Step 2: Energy dissipation balance - upper bound**

  Now, to show that the limit element of the approximating sequence is actually a solution of the system under consideration, we pass to the limit in the energy dissipation balance. Usually, the convergences acquired in step 2, together with lower semicontinuity principles, are sufficient to pass to the limit inferior and thereby obtain an upper bound for the limiting energy dissipation principle.

- **Step 3: Energy dissipation balance - lower bound**

  In order to show that the energy dissipation estimate is in fact a balance, it is necessary to show the opposite estimate as well. This can be done exploiting the principle that was set out in Section 2.2.1, where we argued that the energy dissipation principle (EDP) (i.e., the upper bound obtained in step 2) together with a chain rule for the scalar function $s \mapsto \mathcal{I}(\ell(t(s)), z(s))$ is actually sufficient to obtain the energy dissipation balance (2.2.4). That is why the crucial ingredient in this step is usually a chain rule for the energy under the given regularity assumptions. In this work, the proofs of the necessary chain rules are mostly collected in Appendix F for the sake of readability.

- **Step 4: Improved convergences and regularity**

  Often, the energy dissipation balance can be used to derive improved convergences or further a priori estimates. In some cases however, additional regularity may be required to prove the chain rule necessary for step 3, which is why the order of these two steps may be switched.

Chapter 3

$\mathfrak{p}$-parameterized BV-solutions: Existence and characterizations

3.1 The viscously regularized problem

We will begin by studying the viscously regularized system as introduced in Chapter 2.4.1. Thanks to the regularization, it is possible to show that (2.4.10) has solutions in the differential sense. The following definition is motivated by the reformulations done at the beginning of Section 2.2.1. In particular, the energy dissipation balance (3.1.3) below can be obtained in the same way as (2.2.4) was in Section 2.2.1.

Let us note that we consider the viscous system on a real interval I, where both a bounded interval $I = (0,a)$ for some $a > 0$, or $I = \mathbb{R}_+ := (0,\infty)$ are permitted. For now, we only need the results obtained for a bounded interval. However, the estimates for the case $I = \mathbb{R}_+$, i.e., for a system that is defined on the positive real half axis, will be crucial in Section 4.3, where we show the compactness of the set of $\mathfrak{p}$-parameterized BV solutions. In what follows, we will use the notations $I_0 := I \cup \{0\}$, and $\overline{I}$ for the closure of I.

Definition 3.1.1 (Solutions of the viscous system)**.** *Let $I = (0,a)$ for $a \in \mathbb{R}_+ \cup \{\infty\}$. For $\varepsilon > 0$, we call a curve $z_\varepsilon \in H^1(\overline{I};\mathcal{Z})$ a solution of the system* (2.4.10)*, if it fulfills the inclusion*

$$-\mathrm{D}_z\mathcal{I}(\ell(t), z_\varepsilon(t)) \in \partial\mathcal{R}_\varepsilon(\dot{z}_\varepsilon(t)) \text{ for a.a. } t \in I, \tag{3.1.1}$$

where $\mathcal{R}_\varepsilon$ is defined in (2.4.11)*. With $\mathcal{L}^\varepsilon(z_0,\ell)$, we denote the set of solutions of* (2.4.10) *associated with the pair (z_0,ℓ).*

In analogy to the discussion in Section 2.2.1, solutions of the differential inclusion (3.1.1) can alternatively be characterized by means of the energy dissipation balance (3.1.3) or even the energy dissipation estimate (3.1.4). The crucial ingredient here is again the chain rule for the scalar function $t \mapsto \mathcal{I}(\ell(t), z_\varepsilon(t))$, which is proven in Lemma D.2. Note that in Section 2.2.1, the energy functional $\mathcal{E} : [0,T] \times \big(\mathcal{U} \times \mathcal{Z}\big) \to [0,\infty)$ is given in dependence of time and state, while we consider here the reduced functional $\mathcal{I} : \mathcal{V}^* \times \mathcal{Z} \to [0,\infty)$ in dependence of the external load $\ell \in \mathcal{V}^*$, and of the dissipative variable $z \in \mathcal{Z}$.

Proposition 3.1.2 (Characterization by energy dissipation balance)**.** *Let $I = (0,a)$ for $a \in \mathbb{R}_+ \cup \{\infty\}$ and let $\varepsilon > 0$ be fixed. For a curve $z_\varepsilon \in H^1(\bar{I};\mathcal{Z})$ with*

$$\sup_{t\in I}|\mathcal{I}(\ell(t),z_\varepsilon(t))| < \infty, \quad \mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t)) \in L^1(I;\mathcal{V}^*),\ \textit{and}$$
$$\int_I \mathcal{R}_\varepsilon(\dot{z}_\varepsilon(t))\,\mathrm{d}t < \infty, \quad \int_I \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t)))\,\mathrm{d}t < \infty, \tag{3.1.2}$$

the following are equivalent:

(i) z_ε *fulfills the* ***differential inclusion*** (3.1.1).

(ii) The ***energy dissipation balance***

$$\int_s^t \mathcal{R}_\varepsilon(\dot{z}_\varepsilon(r)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(r),z_\varepsilon(r)))\,\mathrm{d}r + \mathcal{I}(\ell(t),z_\varepsilon(t))$$
$$= \mathcal{I}(\ell(s),z_\varepsilon(s)) + \int_s^t \langle\dot{\ell}(r),z_\varepsilon(r)\rangle_{\mathcal{V}}\,\mathrm{d}r \tag{3.1.3}$$

holds for every $s,t \in I_0$ such that $0 \le s \le t$.

(iii) The ***energy dissipation estimate***

$$\int_s^t \mathcal{R}_\varepsilon(\dot{z}_\varepsilon(r)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(r),z_\varepsilon(r)))\,\mathrm{d}r + \mathcal{I}(\ell(t),z_\varepsilon(t))$$
$$\le \mathcal{I}(\ell(s),z_\varepsilon(s)) + \int_s^t \langle\dot{\ell}(r),z_\varepsilon(r)\rangle_{\mathcal{V}}\,\mathrm{d}r \tag{3.1.4}$$

holds for every $s,t \in I_0$ such that $0 \le s \le t$.

Proof. The proof is similar to the discussion in Section 2.2.1. Since the arguments where rather formal there, we give a full proof here for completeness.

In order to show that the differential inclusion (3.1.1) implies the energy dissipation balance (3.1.3), we first note that, by the Fenchel equivalence (A.2), (3.1.1) is equivalent to

$$\mathcal{R}_\varepsilon(\dot{z}_\varepsilon(t) + \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t))) = \langle -\mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t)),\dot{z}_\varepsilon(t)\rangle_{\mathcal{V}} \text{ for a.a. } t \in I. \tag{3.1.5}$$

Now, the chain rule $(\mathcal{I}_4)$ from Lemma D.2 yields for almost all $t \in I$

$$\mathcal{R}_\varepsilon(\dot{z}_\varepsilon(t)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t))) = \partial_t\mathcal{I}(\ell(t),z_\varepsilon(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}(\ell(t),z_\varepsilon(t)),$$

and we obtain (3.1.3) by integration with respect to t.

It only remains to show that the energy dissipation estimate (3.1.4) is sufficient to conclude that the differential inclusion (3.1.1) is valid. To this end, we first use again the chain rule $(\mathcal{I}_4)$ to write (3.1.4) as

$$\int_s^t \mathcal{R}_\varepsilon(\dot{z}_\varepsilon(r)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(r),z_\varepsilon(r)))\,\mathrm{d}r$$
$$\le \int_s^t \partial_r\mathcal{I}(\ell(r),z_\varepsilon(r)) - \frac{\mathrm{d}}{\mathrm{d}r}\mathcal{I}(\ell(r),z_\varepsilon(r))\,\mathrm{d}r$$
$$= \int_s^t \langle -\mathrm{D}_z\mathcal{I}(\ell(r),z_\varepsilon(r)),\dot{z}_\varepsilon(r)\rangle_{\mathcal{V}}\mathrm{d}r.$$

Since this inequality holds for all $s,t \in I$ such that $s < t$ and the integrands on both sides are in $L^1(I;\mathbb{R})$, this is sufficient to conclude the pointwise estimate

$$\mathcal{R}_\varepsilon(\dot{z}_\varepsilon(t) + \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t))) \le \langle -\mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t)),\dot{z}_\varepsilon(t)\rangle_{\mathcal{V}} \text{ for a.a. } t \in I. \tag{3.1.6}$$

Since the Fenchel-Young inequality (A.1) also yields the opposite estimate, we find that (3.1.6) in fact holds as an identity, which again is equivalent to the differential inclusion (3.1.1) due to the Fenchel equivalence (A.2). ■

We now proceed to prove existence of solutions for the viscously regularized systems. In order to obtain the desired estimates (3.1.7), one formally takes the time-derivative of the inclusion (3.1.1) and chooses $\dot{z}_\varepsilon$ as a test function in the resulting variational inequality. This can be made rigorous by introducing an additional regularization of the inclusion (3.1.1), or by means of an incremental minimization scheme based on a global stability condition (compare (3.1.9) with (2.2.7)). In this work, we choose the latter approach, and refer to [KRZ11], the preprint version of [KRZ13], for an instalment of the first approach in a model for damage development in elastic materials. See also [KT18] for the simpler case with bounded dissipation and constant load.

Proposition 3.1.3 (Existence for the viscous problem). *Let $a \in \mathbb{R}_+ \cup \{\infty\}$, and let $I = (0,a)$. For every $\varepsilon > 0$, $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ and initial value $z_0 \in \mathcal{Z}$ such that $-\mathrm{D}_z\mathcal{I}(\ell(0),z_0) \in \partial\mathcal{R}(0)$, there exists a unique function $z_\varepsilon \in H^1(\bar{I},\mathcal{Z})$ satisfying* (3.1.1) *and $-\mathrm{D}_z\mathcal{I}(\ell,z_\varepsilon) \in L^\infty(I;\mathcal{V}^*)$.*

Moreover, there exist a function $m(\cdot,\cdot) : [0,\infty)\times[0,\infty) \to [0,\infty)$, mapping bounded sets into bounded sets, and a constant $C > 0$ such that the following estimates are valid for all $\varepsilon > 0$:

$$\forall\, t \in I: \quad \mathcal{I}(\ell(t),z_\varepsilon(t)) \le \Big(\mathcal{I}(\ell(0),z_0) + c_0\Big)\exp\Big(\mathrm{Var}_{\mathcal{V}^*}(\ell,\bar{I})\Big), \tag{3.1.7a}$$

$$\|z_\varepsilon\|_{L^\infty(0,T;\mathcal{Z})} \le C\Big(\mathcal{I}(\ell(0),z_0) + c_0\Big)\exp\Big(\mathrm{Var}_{\mathcal{V}^*}(\ell,\bar{I})\Big), \tag{3.1.7b}$$

$$\int_I \mathcal{R}_\varepsilon(\dot{z}_\varepsilon)\,\mathrm{d}r + \int_I \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(r),z_\varepsilon(r)))\,\mathrm{d}r \le m(\mathcal{I}(\ell(0),z_0),\mathrm{Var}_{\mathcal{V}^*}(\ell,\bar{I})), \tag{3.1.7c}$$

$$\sqrt{\varepsilon}\|z_\varepsilon\|_{H^1(0,T;\mathcal{V})} \le m(\mathcal{I}(\ell(0),z_0),\mathrm{Var}_{\mathcal{V}^*}(\ell,\bar{I})), \tag{3.1.7d}$$

$$\sqrt{\varepsilon}\|z_\varepsilon\|_{H^1(0,T;\mathcal{Z})} \le m(\mathcal{I}(\ell(0),z_0),\mathrm{Var}_{\mathcal{V}^*}(\ell,\bar{I})), \tag{3.1.7e}$$

$$\mathrm{Var}_{\mathcal{V}}(z_\varepsilon,\bar{I}) \le m(\mathcal{I}(\ell(0),z_0),\mathrm{Var}_{\mathcal{V}^*}(\ell,\bar{I})), \tag{3.1.7f}$$

$$\varepsilon\|z_\varepsilon\|_{W^{1,\infty}(0,T;\mathcal{V})} \le m(\mathcal{I}(\ell(0),z_0),\mathrm{Var}_{\mathcal{V}^*}(\ell,\bar{I})), \tag{3.1.7g}$$

$$\|\mathrm{D}_z\mathcal{I}(\ell,z_\varepsilon)\|_{L^\infty(0,T;\mathcal{V}^*)} \le \mathrm{diam}_{\mathcal{V}^*}(\partial\mathcal{R}(0)) + m(\mathcal{I}(\ell(0),z_0),\mathrm{Var}_{\mathcal{V}^*}(\ell,\bar{I})). \tag{3.1.7h}$$

Proof. **Step 0: Uniqueness** Let $\varepsilon > 0$ be fixed, and for $i \in \{1,2\}$ let $z_i \in L^\infty(I;\mathcal{Z})$ with $\dot{z}_i \in L^1(I;\mathcal{V})$ and $\mathrm{D}_z\mathcal{I}(\ell(\cdot),z_i(\cdot)) \in L^\infty(I;\mathcal{V}^*)$ and such that (2.4.10) is satisfied. For the sake of readability, we omit the index ε on the solutions z_i in this step. At points $t \in I$ where both z_i are differentiable, we further choose $\eta_i(t) \in \partial\mathcal{R}_{2,\varepsilon}(\dot{z}_i(t))$, cf. (2.4.11) for the definition of $\mathcal{R}_{2,\varepsilon}$. By the monotonicity of the operator $\partial\mathcal{R}$ we obtain for almost all $t \in I$

$$0 \geq \langle \eta_1(t) - \eta_2(t), \dot{z}_1(t) - \dot{z}_2(t)\rangle_{\mathcal{V}} + \langle \mathrm{D}_z\mathcal{I}(\ell(t), z_1(t)) - \mathrm{D}_z\mathcal{I}(\ell(t), z_2(t)), \dot{z}_1(t) - \dot{z}_2(t)\rangle_{\mathcal{V}},$$

which, after application of the chain rule F.2 reads as

$$\begin{aligned}\langle \eta_1(t) - \eta_2(t), \dot{z}_1(t) - \dot{z}_2(t)\rangle_{\mathcal{V}} &+ \tfrac{1}{2}\tfrac{\mathrm{d}}{\mathrm{d}t}(\langle A(z_1(t) - z_2(t)), z_1(t) - z_2(t)\rangle_{\mathcal{Z}}) \\ &\leq -\langle \mathrm{D}\mathcal{F}(z_1(t)) - \mathrm{D}\mathcal{F}(z_2(t)), \dot{z}_1(t) - \dot{z}_2(t)\rangle_{\mathcal{V}}\end{aligned}$$

Now, since $\mathcal{V}$ is uniformly convex with a modulus of convexity of power 2, we know from Lemma G.1 that the monotone operator $\partial\mathcal{R}_{2,\varepsilon}(\cdot)$ is uniformly convex in the sense of (G.1), i.e., there exists $\gamma > 0$ such that

$$\begin{aligned}\gamma\|\dot{z}_1(t) - \dot{z}_2(t)\|_{\mathcal{V}}^2 &+ \tfrac{1}{2}\tfrac{\mathrm{d}}{\mathrm{d}t}(\langle A(z_1(t) - z_2(t)), z_1(t) - z_2(t)\rangle_{\mathcal{Z}}) \\ &\leq -\langle \mathrm{D}\mathcal{F}(z_1(t)) - \mathrm{D}\mathcal{F}(z_2(t)), \dot{z}_1(t) - \dot{z}_2(t)\rangle_{\mathcal{V}} \\ &\leq \|\mathrm{D}\mathcal{F}(z_1(t)) - \mathrm{D}\mathcal{F}(z_2(t))\|_{\mathcal{V}^*}\|\dot{z}_1(t) - \dot{z}_2(t)\|_{\mathcal{V}} \\ &\leq C\|z_1(t) - z_2(t)\|_{\mathcal{V}}\|\dot{z}_1(t) - \dot{z}_2(t)\|_{\mathcal{V}} \\ &\leq \widetilde{C}\|z_1(t) - z_2(t)\|_{\mathcal{V}}^2 + \frac{\gamma}{2}\|\dot{z}_1(t) - \dot{z}_2(t)\|_{\mathcal{V}}^2, \qquad (3.1.8)\end{aligned}$$

where, in the second to last inequality, we have used (2.4.4a) - (2.4.4b) and the fact that $z_i \in L^\infty(I;\mathcal{Z})$. Using the coercivity of A from (2.4.3) to estimate the left hand side and integrating w.r.t. time, this yields

$$\tfrac{\alpha}{2}\|z_1(s) - z_2(s)\|_{\mathcal{Z}}^2 \leq \widetilde{C}\int_0^s \|z_1(t) - z_2(t)\|_{\mathcal{V}}^2 \mathrm{d}t \leq C\int_0^s \|z_1(t) - z_2(t)\|_{\mathcal{Z}}^2 \mathrm{d}t$$

for almost all $s \in I$. By means of the Gronwall inequality we conclude uniqueness.

Step 1: Time discretization scheme First, let $T > 0$ be arbitrary. We will apply the existence result in Theorem 2.2 in [MRS13] to obtain existence of a solution z_ε on a bounded interval $I = (0,T)$, which is why for steps 1 and 2, we consider the case $a := T \in \mathbb{R}$. We prove in Lemma D.1 and D.2, respectively, that the necessary assumptions $(\mathcal{R}_1)$-$(\mathcal{R}_3)$ on $\mathcal{R}$ and $(\mathcal{I}_1)$-$(\mathcal{I}_5)$ on $\mathcal{I}$ from [MRS13] hold true in the setting described in Section 2.4.1. Sole application of [MRS13, Thm. 2.2] only yields existence of a solution z_ε together with the regularity and convergence results (3.1.27). We therefore return to the time discretization scheme used in the proof of [MRS13, Thm. 2.2] in order to achieve the a priori estimates and improved regularity that are stated in (3.1.7). Namely, we choose a fixed initial value $z_0 \in \mathcal{Z}$ and a time step size $\tau > 0$ to define a partition $\{0 = t_0^\tau < t_1^\tau < \cdots < t_{N-1}^\tau < T \leq t_N^\tau\}$ of $[0,T]$ by $t_n^\tau := n\tau$. We now set $z_0^\tau := z_0$ and obtain $(z_k^\tau)_{k=1,\ldots,N}$ as solutions of

$$z_{k+1}^\tau \in \operatorname{Argmin}\left\{\mathcal{I}(\ell(t_{k+1}^\tau), z) + \tau\mathcal{R}_\varepsilon\left(\frac{z - z_k^\tau}{\tau}\right), z \in \mathcal{Z}\right\}. \qquad (3.1.9)$$

These minimizers exist since we have proven in Lemma D.2 that for every $\tau > 0$ and $k \in \{1,\ldots,N\}$, the objective functional in (3.1.9) has compact sublevels, and is lower semicontinuous, both w.r.t. the strong topology on $\mathcal{V}$ and w.r.t. the weak topology on $\mathcal{Z}$. We now define the following piecewise constant and piecewise linear interpolants:

$$\begin{aligned} \bar{z}_\tau(t) &= z^\tau_{k+1} && \text{for } t \in (t^\tau_k, t^\tau_{k+1}] \\ \hat{z}_\tau(t) &= z^\tau_k + \frac{t - t^\tau_k}{\tau}(z^\tau_{k+1} - z^\tau_k) && \text{for } t \in [t^\tau_k, t^\tau_{k+1}]. \end{aligned}$$

We further use the notation

$$\bar{t}_\tau(t) = t^\tau_{k+1} \text{ for } t \in (t^\tau_k, t^\tau_{k+1}].$$

Note that $(\mathcal{I}_2)$ from Lemma D.2 implies that z^τ_{k+1} chosen according to (3.1.9) fulfills

$$-\mathrm{D}_z\mathcal{I}(\ell(t^\tau_{k+1}), z^\tau_{k+1}) \in \partial\mathcal{R}_\varepsilon\Big(\frac{z^\tau_{k+1} - z^\tau_k}{\tau}\Big) = \partial\mathcal{R}_\varepsilon(\dot{\hat{z}}_\tau(t)) \text{ for } t \in (t^\tau_k, t^\tau_{k+1}). \tag{3.1.10}$$

Step 2: A priori estimates for the approximate solutions We proceed with some a priori estimates. In order to obtain estimates that are independent of the length of the Interval I, we apply a technique from [KZ21] relying on Lemma D.3. First, we show a uniform bound for the energies. To this purpose, note that by the definition of z^τ_k in (3.1.9), we have for all $k \in \mathbb{N}$ that

$$\begin{aligned} \mathcal{I}(\ell(t_k), z^\tau_k) + \tau\mathcal{R}_\varepsilon\left(\frac{z^\tau_k - z^\tau_{k-1}}{\tau}\right) &\le \mathcal{I}(\ell(t_k), z^\tau_{k-1}) \\ &= \mathcal{I}(\ell(t_{k-1}), z^\tau_{k-1}) + \int_{t_{k-1}}^{t_k} \partial_s\mathcal{I}(\ell(s), z^\tau_{k-1})\,\mathrm{d}s \\ &= \mathcal{I}(\ell(t_{k-1}), z^\tau_{k-1}) + \langle \ell(t_{k-1}) - \ell(t_k), z^\tau_{k-1}\rangle_{\mathcal{V}} \\ &\le \mathcal{I}(\ell(t_{k-1}), z^\tau_{k-1}) + \|\ell(t_{k-1}) - \ell(t_k)\|_{\mathcal{V}^*}\|z^\tau_{k-1}\|_{\mathcal{V}} \\ &\le \mathcal{I}(\ell(t_{k-1}), z^\tau_{k-1}) + \|\ell(t_{k-1}) - \ell(t_k)\|_{\mathcal{V}^*}(c_0 + \mathcal{I}(\ell(t_{k-1}), z^\tau_{k-1})), \end{aligned} \tag{3.1.11}$$

where we have used the estimate (D.11) in the last step. Since the second term on the left hand side is non-negative, by adding c_0 on both sides, this estimate reads

$$\mathcal{I}(\ell(t_k), z^\tau_k) + c_0 \le \Big(\mathcal{I}(\ell(t_{k-1}), z^\tau_{k-1}) + c_0\Big)\Big(1 + \|\ell(t_{k-1}) - \ell(t_k)\|_{\mathcal{V}^*}\Big),$$

from which we obtain by recursion and Lemma D.3

$$\begin{aligned} \mathcal{I}(\ell(t_k), z^\tau_k) + c_0 &\le \Big(\mathcal{I}(\ell(0), z_0) + c_0\Big)\prod_{i=1}^{k}\Big(1 + \|\ell(t_{i-1}) - \ell(t_i)\|_{\mathcal{V}^*}\Big) \\ &\le \Big(\mathcal{I}(\ell(0), z_0) + c_0\Big)\exp(\mathrm{Var}_{\mathcal{V}^*}(\ell, [0, t_k])). \end{aligned}$$

Therefore, we have that

$$\forall\, 1 \le k \le N : \sup_{\tau,\varepsilon}\mathcal{I}(\ell(t_k), z^\tau_k) \le \Big(\mathcal{I}(\ell(0), z_0) + c_0\Big)\exp(\mathrm{Var}_{\mathcal{V}^*}(\ell, [0, t_k])) \tag{3.1.12}$$

and consequently by (D.6)

$$\forall\, 1 \le k \le N : \sup_{\tau,\varepsilon} \|z_k^\tau\|_{\mathcal{Z}} \le C\Big(\mathcal{I}(\ell(0), z_0) + c_0\Big)\exp(\mathrm{Var}_{\mathcal{V}^*}(\ell,[0,t_k])) < \infty. \tag{3.1.13}$$

With this estimate, we obtain from (3.1.11) by summation over k that

$$\begin{aligned}\sup_{\varepsilon,\tau>0} \int_I \mathcal{R}_\varepsilon(\dot{\hat{z}}_\tau(t))\,\mathrm{d}t &\le \mathcal{I}(\ell(0), z_0) - \mathcal{I}(\ell(T), z_N^\tau) + \sum_{k=1}^N \langle \ell(t_{k-1}) - \ell(t_k), z_{k-1}^\tau\rangle_{\mathcal{V}} \\ &\le \mathcal{I}(\ell(0), z_0) - \mathcal{I}(\ell(T), z_N^\tau) + \mathrm{Var}_{\mathcal{V}^*}(\ell,[0,T])\|z_{k-1}^\tau\|_{\mathcal{V}} \\ &\le m_1(\mathcal{I}(\ell(0), z_0), \mathrm{Var}_{\mathcal{V}^*}(\ell,[0,T])),\end{aligned} \tag{3.1.14}$$

for a function $m_1(\cdot,\cdot) : [0,\infty)\times[0,\infty) \to [0,\infty)$ that maps bounded sets into bounded sets. Since $\mathcal{R}_\varepsilon$ grows quadratically in $\|z\|_{\mathcal{V}}$ and is coercive w.r.t. $\|\cdot\|_{\mathcal{V}}$, this allows us to conclude that

$$\sup_{\varepsilon,\tau>0} \sqrt{\varepsilon}\|\hat{z}_\tau\|_{H^1(0,T;\mathcal{V})} \le C\cdot m_1(\mathcal{I}(\ell(0), z_0), \mathrm{Var}_{\mathcal{V}^*}(\ell,[0,T])) < \infty. \tag{3.1.15}$$

Furthermore, we infer from (3.1.14) and the superlinear growth of $\mathcal{R}_\varepsilon$ w.r.t. the norm on $\mathcal{V}$ that $(\dot{\hat{z}}_\tau)_{\tau>0} \subset L^1(0,T;\mathcal{V})$ are uniformly integrable, so that $(\hat{z}_\tau)_{\tau>0}$ are equicontinuous w.r.t. the norm on $\mathcal{V}$. Together with the $L^\infty(0,T;\mathcal{Z})$ estimate (3.1.13), this enables us to apply the Arzelá-Ascoli-Theorem [Die69, Thm. 7.5.7] to the sequence $(\hat{z}_\tau)_{\tau>0}$, whereby we find a curve $z_\varepsilon \in C([0,T];\mathcal{V})$ such that the interpolants converge uniformly, up to a subsequence.

Next, we are going to derive further estimates for the interpolants' derivatives. First, we infer from the subdifferential inclusion (3.1.10) and the characterization of the subdifferentials in Lemma 2.4.4 that for all $k \ge 0$, there exists $\varepsilon\sigma_{k+1} \in \partial\mathcal{R}_{2,\varepsilon}\left(\frac{z_{k+1}^\tau - z_k^\tau}{\tau}\right)$, i.e. $\sigma_{k+1} \in \mathcal{V}^*$ such that

$$\|\sigma_{k+1}\|_{\mathcal{V}^*} = \|\frac{z_{k+1}^\tau - z_k^\tau}{\tau}\|_{\mathcal{V}} \text{ and } \langle \sigma_{k+1}, \frac{z_{k+1}^\tau - z_k^\tau}{\tau}\rangle_{\mathcal{V}} = \|\frac{z_{k+1}^\tau - z_k^\tau}{\tau}\|_{\mathcal{V}}^2,$$

which fulfills

$$\underbrace{-\mathrm{D}_z\mathcal{I}(\ell(t_{k+1}), z_{k+1}^\tau) - \varepsilon\sigma_{k+1}}_{=:\xi_{k+1}} \in \partial\mathcal{R}\left(\frac{z_{k+1}^\tau - z_k^\tau}{\tau}\right) \subset \partial\mathcal{R}(0), \tag{3.1.16}$$

where the last inclusion follows from Lemma 2.4.4. Also due to this Lemma, we find that for all $k \ge 0$, it follows from $\xi_{k+1} \in \partial\mathcal{R}\left(\frac{z_{k+1}^\tau - z_k^\tau}{\tau}\right)$ that

$$\langle \xi_{k+1}, \frac{z_{k+1}^\tau - z_k^\tau}{\tau}\rangle_{\mathcal{V}} = \mathcal{R}\left(\frac{z_{k+1}^\tau - z_k^\tau}{\tau}\right) \Leftrightarrow \langle \xi_{k+1}, z_{k+1}^\tau - z_k^\tau\rangle_{\mathcal{V}} = \mathcal{R}\left(z_{k+1}^\tau - z_k^\tau\right). \tag{3.1.17}$$

For all $k \ge 1$, it follows from $\xi_k \in \partial\mathcal{R}(0)$ that

$$\forall v \in \mathcal{V} : \mathcal{R}(v) \ge \langle \xi_k, v\rangle_{\mathcal{V}}, \tag{3.1.18}$$

and setting $\sigma_0 := 0$, we obtain the same for ξ_0 from our assumption on the initial value z_0. Now, the identity (3.1.17) and the inequality (3.1.18) for $v = z^\tau_{k+1} - z^\tau_k$ yield

$$\begin{aligned} 0 &\geq \langle \xi_k, z^\tau_{k+1} - z^\tau_k\rangle_{\mathcal{V}} - \langle \xi_{k+1}, z^\tau_{k+1} - z^\tau_k\rangle_{\mathcal{V}} \\ &= \langle \xi_k - \xi_{k+1}, z^\tau_{k+1} - z^\tau_k\rangle_{\mathcal{V}} \\ &= \langle \mathrm{D}_z\mathcal{I}(\ell(t_{k+1}), z^\tau_{k+1}) - \mathrm{D}_z\mathcal{I}(\ell(t_k), z_k), z^\tau_{k+1} - z^\tau_k\rangle_{\mathcal{V}} + \varepsilon\underbrace{\langle \sigma_{k+1}, z^\tau_{k+1} - z^\tau_k\rangle_{\mathcal{V}}}_{=1/\tau\|z^\tau_{k+1}-z^\tau_k\|^2_{\mathcal{V}}} - \varepsilon\langle \sigma_k, z^\tau_{k+1} - z^\tau_k\rangle_{\mathcal{V}} \end{aligned} \tag{3.1.19}$$

and thus, using estimate (2.4.16) for $\eta = \frac{\alpha}{2}$,

$$\begin{aligned} \alpha\|z^\tau_{k+1} - z^\tau_k\|^2_{\mathcal{Z}} + \frac{\varepsilon}{\tau}\|z^\tau_{k+1} - z^\tau_k\|^2_{\mathcal{V}} &\leq \langle A(z^\tau_{k+1} - z^\tau_k), z^\tau_{k+1} - z^\tau_k\rangle_{\mathcal{V}} + \frac{\varepsilon}{\tau}\|z^\tau_{k+1} - z^\tau_k\|^2_{\mathcal{V}} \\ &\leq \langle \ell(t_{k+1}) - \ell(t_k), z^\tau_{k+1} - z^\tau_k\rangle_{\mathcal{V}} - \langle \mathrm{D}\mathcal{F}(z^\tau_{k+1}) - \mathrm{D}\mathcal{F}(z^\tau_k), z^\tau_{k+1} - z^\tau_k\rangle_{\mathcal{V}} + \varepsilon\langle \sigma_k, z^\tau_{k+1} - z^\tau_k\rangle_{\mathcal{V}} \\ &\leq \|\ell(t_{k+1}) - \ell(t_k)\|_{\mathcal{V}^*}\|z^\tau_{k+1} - z^\tau_k\|_{\mathcal{V}} + \frac{\alpha}{2}\|z^\tau_{k+1} - z^\tau_k\|^2_{\mathcal{Z}} + C_{\rho,\eta}\mathcal{R}\left(z^\tau_{k+1} - z^\tau_k\right)\|z^\tau_{k+1} - z^\tau_k\|_{\mathcal{V}} \\ &\qquad + \varepsilon\|\sigma_k\|_{\mathcal{V}^*}\|z^\tau_{k+1} - z^\tau_k\|_{\mathcal{V}}. \end{aligned}$$

By absorbing the $\mathcal{Z}$-norm of $z^\tau_{k+1} - z^\tau_k$ into the term on the left hand side and using the embedding of $\mathcal{Z}$ into $\mathcal{V}$, this estimate reads

$$\frac{\alpha}{2}\|z^\tau_{k+1} - z^\tau_k\|_{\mathcal{Z}} + \frac{\varepsilon}{\tau}\|z^\tau_{k+1} - z^\tau_k\|_{\mathcal{V}} \leq \|\ell(t_{k+1}) - \ell(t_k)\|_{\mathcal{V}^*} + C_{\rho,\eta}\mathcal{R}\left(z^\tau_{k+1} - z^\tau_k\right) + \varepsilon\|\sigma_k\|_{\mathcal{V}^*}. \tag{3.1.20}$$

For $k \geq 1$, we can estimate the last term by $\varepsilon/\tau\|z^\tau_k - z^\tau_{k-1}\|_{\mathcal{V}}$, whereas for $k = 0$, we have to use the assumption $-\mathrm{D}_z\mathcal{I}(0, z_0) \in \partial\mathcal{R}(0)$ in order to conclude that estimate (3.1.20) is true for $\sigma_0 = 0$. Summation of (3.1.20) for $k = 0, \ldots, N$ gives

$$\frac{\varepsilon}{\tau}\|z^\tau_N - z^\tau_{N-1}\|_{\mathcal{V}} + \frac{\alpha}{2}\sum_{k=0}^{N}\|z^\tau_{k+1} - z^\tau_k\|_{\mathcal{Z}} \leq \sum_{k=0}^{N}\Big(\|\ell(t_{k+1}) - \ell(t_k)\|_{\mathcal{V}^*} + C_{\rho,\eta}\mathcal{R}\left(z^\tau_{k+1} - z^\tau_k\right)\Big),$$

i.e.

$$\frac{\varepsilon}{\tau}\|z^\tau_N - z^\tau_{N-1}\|_{\mathcal{V}} + \frac{\alpha}{2}\int_0^T \|\dot{\hat{z}}_\tau(t)\|_{\mathcal{Z}}\,\mathrm{d}t \leq \sum_{k=0}^{N}\|\ell(t_{k+1}) - \ell(t_k)\|_{\mathcal{V}^*} + C_{\rho,\eta}\int_0^T \mathcal{R}\left(\dot{\hat{z}}_\tau(t)\right)\mathrm{d}t$$

and since the right hand side is uniformly bounded according to (3.1.14), this implies that

$$\sup_{\varepsilon,\tau>0}\|\dot{\hat{z}}_\tau\|_{L^1(0,T;\mathcal{Z})} \leq \mathrm{Var}_{\mathcal{V}^*}(\ell, [0,T]) + m_1(\mathcal{I}(\ell(0), z_0), \mathrm{Var}_{\mathcal{V}^*}(\ell, [0,T])) < \infty. \tag{3.1.21}$$

With similiar arguments, we obtain an estimate for the $\mathcal{V}$-norms of the interpolants' derivatives from (3.1.20). Namely, by summation up to an arbitrary index $k \in \{0, \ldots, N\}$, it follows that it holds for all $\tau > 0$ that

$$\frac{\varepsilon}{\tau}\|z^\tau_{k+1} - z^\tau_k\|_{\mathcal{V}} = \varepsilon\|\dot{\hat{z}}_\tau(t)\|_{\mathcal{V}} \leq \mathrm{Var}_{\mathcal{V}^*}(\ell, [0, t_{k+1}]) + \int_0^{t_{k+1}} C_{\rho,\eta}\mathcal{R}(\dot{\hat{z}}_\tau(t))\,\mathrm{d}t,$$

and together with (3.1.14), this implies

$$\sup_{\varepsilon>0,\tau>0} \varepsilon\|\hat{z}_\tau\|_{W^{1,\infty}(0,T;\mathcal{V})} \leq \mathrm{Var}_{\mathcal{V}^*}(\ell,[0,T]) + m_1(\mathcal{I}(\ell(0),z_0),\mathrm{Var}_{\mathcal{V}^*}(\ell,[0,T])) < \infty. \tag{3.1.22}$$

Having this estimate in mind, we reconsider the subdifferential inclusion (3.1.16), which reads

$$-\mathrm{D}_z\mathcal{I}(\ell(t_{k+1}),z_{k+1}^\tau) \in \partial\mathcal{R}(0) - \varepsilon\sigma_{k+1},$$

where $\sigma_{k+1} \in \mathcal{V}^*$ with $\varepsilon\|\sigma_{k+1}\|_{\mathcal{V}^*} = \varepsilon\|\frac{z_{k+1}^\tau - z_k^\tau}{\tau}\|_{\mathcal{V}}$. Thus, the right hand side of the above inclusion is uniformly bounded and it holds that

$$\sup_{\varepsilon,\tau>0}\|\mathrm{D}_z\mathcal{I}(\ell(\bar{t}_\tau),\bar{z}_\tau)\|_{L^\infty(0,T;\mathcal{V}^*)} \leq \mathrm{diam}_{\mathcal{V}^*}(\partial\mathcal{R}(0)) + m(\mathcal{I}(\ell(0),z_0),\mathrm{Var}_{\mathcal{V}^*}(\ell,[0,T])). \tag{3.1.23}$$

Let us now return to the variational inequality (3.1.19) in order to obtain an estimate for $\|\dot{\hat{z}}_\tau\|_{L^2(0,T;\mathcal{Z})}$. We first divide (3.1.19) by $\tau > 0$ and obtain for $t \in (t_k,t_{k+1})$, similarly to (3.1.20), the estimate

$$\varepsilon\langle\sigma_{k+1}-\sigma_k,\dot{\hat{z}}_\tau(t)\rangle_{\mathcal{V}} + \tau\alpha\|\dot{\hat{z}}_\tau(t)\|_{\mathcal{Z}}^2 \leq \langle\ell(t_{k+1})-\ell(t_k),\dot{\hat{z}}(t)\rangle_{\mathcal{V}} + \tau\tfrac{\alpha}{2}\|\dot{\hat{z}}_\tau(t)\|_{\mathcal{Z}}^2 + \tau\tilde{C}_{\varrho,\eta}\|\dot{\hat{z}}_\tau(t)\|_{\mathcal{V}}^2, \tag{3.1.24}$$

where we also exploited the embedding $\mathcal{V} \hookrightarrow \mathcal{X}$. We now proceed as in [KRZ13, Prop. 4.1] in order to estimate the first term on the left for $s \in (t_{k-1},t_k)$ by

$$\begin{aligned}\langle\sigma_{k+1}-\sigma_k,\dot{\hat{z}}_\tau(t)\rangle_{\mathcal{V}} &= \frac{1}{2}\Big(\|\dot{\hat{z}}_\tau(t)\|_{\mathcal{V}}^2 - \|\dot{\hat{z}}_\tau(s)\|_{\mathcal{V}}^2 + \langle\sigma_{k+1},\dot{\hat{z}}_\tau\rangle_{\mathcal{V}} + \langle\sigma_k,\dot{\hat{z}}(s)\rangle_{\mathcal{V}} - 2\langle\sigma_k,\dot{\hat{z}}_\tau(t)\rangle_{\mathcal{V}}\Big)\\ &\geq \frac{1}{2}\Big(\|\dot{\hat{z}}_\tau(t)\|_{\mathcal{V}}^2 - \|\dot{\hat{z}}_\tau(s)\|_{\mathcal{V}}^2 + \Big(\|\dot{\hat{z}}_\tau(t)\|_{\mathcal{V}} - \|\dot{\hat{z}}_\tau(s)\|_{\mathcal{V}}\Big)^2\Big)\\ &\geq \tfrac{1}{2}\|\dot{\hat{z}}_\tau(t)\|_{\mathcal{V}}^2 - \tfrac{1}{2}\|\dot{\hat{z}}_\tau(s)\|_{\mathcal{V}}^2.\end{aligned}$$

With this estimate, (3.1.24) now reads

$$\tfrac{\varepsilon}{2}\|\dot{\hat{z}}_\tau(t)\|_{\mathcal{V}}^2 - \tfrac{\varepsilon}{2}\|\dot{\hat{z}}_\tau(s)\|_{\mathcal{V}}^2 + \tau\tfrac{\alpha}{2}\|\dot{\hat{z}}_\tau(t)\|_{\mathcal{Z}}^2 \leq \langle\ell(t_{k+1})-\ell(t_k),\dot{\hat{z}}(t)\rangle_{\mathcal{V}} + \tau\tilde{C}_{\varrho,\eta}\|\dot{\hat{z}}_\tau(t)\|_{\mathcal{V}}^2$$

for all $k \geq 0$, where we may set $\dot{\hat{z}}_\tau(s) := 0$ for $k = 0$ thanks to our assumption that $-\mathrm{D}\mathcal{I}_z(\ell(0),z_0) \in \partial\mathcal{R}(0)$. Summation for $k = 0,\dots,N$ now yields

$$\begin{aligned}\tfrac{\varepsilon}{2}\|\dot{\hat{z}}_\tau(T)\|_{\mathcal{V}}^2 + \tfrac{\alpha}{2}\int_0^{t_N}\|\dot{\hat{z}}_\tau(r)\|_{\mathcal{Z}}^2\mathrm{d}r &\leq \sum_{k=0}^{N}\langle\ell(t_{k+1})-\ell(t_k),\dot{\hat{z}}(t)\rangle_{\mathcal{V}} + \tilde{C}_{\varrho,\eta}\int_0^{t_N}\|\dot{\hat{z}}_\tau(r)\|_{\mathcal{V}}^2\mathrm{d}r\\ &\leq \|\dot{\ell}\|_{L^\infty(0,T;\mathcal{V}^*}\|\dot{\hat{z}}_\tau\|_{L^1(0,T;\mathcal{V})} + \tilde{C}_{\varrho,\eta}\|\dot{\hat{z}}_\tau\|_{L^2(0,T;\mathcal{V})}^2,\end{aligned} \tag{3.1.25}$$

where we used the fact that the partition $\{t_0,\dots,t_n\}$ is equidistant in the second inequality. We can now estimate the right hand side of (3.1.25) by means of

(3.1.15) and (3.1.21), and also estimate $\hat{z}_\tau$ according to (3.1.13), and thus find the existence of a function $m_2 : [0,\infty) \times [0,\infty) \to [0,\infty)$ that maps bounded sets into bounded sets such that

$$\sup_{\tau>0} \|\hat{z}_\tau\|_{H^1(0,T;\mathcal{Z})} \leq m_2(\mathrm{Var}_{\mathcal{V}^*}(\ell,[0,T]),\mathcal{I}(\ell(0),z_0))\left(1+\frac{1}{\sqrt{\varepsilon}}\right) \tag{3.1.26}$$

Step 3: Convergence and energy dissipation balance Now, applying [MRS13, Theorem 2.2], we find that there exist a solution z_ε of (3.1.1) on the interval $(0,T)$ in the sense of Definition 3.1.1 such that $\mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t)) \in L^1(0,T;\mathcal{V}^*)$ and a sequence $\tau_n \searrow 0$ such that

$$\hat{z}_{\tau_n}, \overline{z}_{\tau_n} \to z_\varepsilon \text{ in } L^\infty(0,T;\mathcal{V}), \tag{3.1.27a}$$

$$\hat{z}_{\tau_n} \rightharpoonup z_\varepsilon \text{ in } W^{1,1}(0,T;\mathcal{V}), \tag{3.1.27b}$$

$$\mathcal{I}(\ell(t),\overline{z}_{\tau_n}(t)) \to \mathcal{I}(\ell(t),z_\varepsilon(t)) \text{ for all } t \in [0,T], \tag{3.1.27c}$$

$$\int_s^t \mathcal{R}_\varepsilon(\dot{\hat{z}}_{\tau_n}(r))\,\mathrm{d}r \to \int_s^t \mathcal{R}_\varepsilon(\dot{z}_\varepsilon(r))\,\mathrm{d}r \text{ for all } 0 \leq s < t \leq T. \tag{3.1.27d}$$

We will use the a priori estimates from Step 2 in order to obtain the improved estimates (3.1.7) in the next step. Before that, let us convince ourselves that (2.4.10) can be solved not only on bounded intervals $(0,T)$, but also has global in time solutions in the case that $I = (0,\infty)$. To this end, we first note that (3.1.1) can be solved on the closed interval $[0,T]$. Indeed, since the solution z_ε of (3.1.1) on $(0,T)$ is an element of $\mathrm{AC}([0,T];\mathcal{V})$ (cf. Lemma C.6), it can be extended to $[0,T]$ with a value $z_\varepsilon(T)$. What is more, (3.1.1) has a solution $\tilde{z}_\varepsilon \in W^{1,1}(0,T+1;\mathcal{V})$ on the interval $(0,T+1)$, and due to the uniqueness, it holds that $\tilde{z}_\varepsilon(t) = z_\varepsilon(t)$ for almost all $0 < t < T$. We thereby infer that

$$z_\varepsilon(T) = \lim_{t \nearrow T} z_\varepsilon(t) = \lim_{t \nearrow T} \tilde{z}_\varepsilon(t) = \tilde{z}_\varepsilon(T),$$

so that z_ε must be a solution of (3.1.1) on $[0,T]$. We now argue by contradiction: Assume that there is a time $T_* > 0$ such that the solution on $[0,T_*]$ cannot be extended beyond T_*. Now, applying [MRS13, Theorem 2.2] to the system (2.4.10) on an interval $[T_*,S]$ for some $S > T_*$ with the new initial state $z_\varepsilon(T_*)$ yields a contradiction.

Step 4: A priori estimates for viscous solutions We now proceed to transfer the a priori estimates for the interpolants that were derived in step 2 onto the viscous solution z_ε in order to improve the estimates obtained from [MRS13, Thm. 2.2]. From the estimate (3.1.26), we conclude that there exists a curve $\tilde{z}_\varepsilon \in H^1(I;\mathcal{Z})$ such that

$$\hat{z}_\tau \rightharpoonup \tilde{z}_\varepsilon \text{ in } H^1(I;\mathcal{Z}) \subset H^1(I;\mathcal{V}),$$

and we infer that $z_\varepsilon \in H^1(I;\mathcal{Z})$ together with the estimate (3.1.7e). From the estimate (3.1.21), we further obtain (3.1.7f), and estimate (3.1.22) implies (3.1.7g). Finally, in order to show the estimate (3.1.7h), we first show that

$$\mathrm{D}_z\mathcal{I}(\ell(\overline{t}_{\tau_n}(t)),\overline{z}_{\tau_n}(t)) \rightharpoonup \mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t)) \text{ weakly in } \mathcal{V}^* \text{ a.e. in } I. \tag{3.1.28}$$

Indeed, let $t \in I$ be such that $\dot{\ell}(t)$ exists. Then we have that

$$\|\mathrm{D}_z\mathcal{I}(\ell(t),\overline{z}_{\tau_n}(t)) - \mathrm{D}_z\mathcal{I}(\ell(\overline{t}_{\tau_n}(t)),\overline{z}_{\tau_n}(t))\|_{\mathcal{V}^*} \le \|\ell(t)-\ell(\overline{t}_{\tau_n}(t))\|_{\mathcal{V}^*} \le \tau_n\|\dot{\ell}\|_{L^\infty(I;\mathcal{V}^*)}, \tag{3.1.29}$$

whereby we may conclude from the a priori estimate (3.1.23) that

$$\sup_{n\in\mathbb{N}}\|\mathrm{D}_z\mathcal{I}(\ell(t),\overline{z}_{\tau_n}(t))\|_{\mathcal{V}^*} < \infty,$$

so that there exists $\xi(t) \in \mathcal{V}^*$ such that

$$\mathrm{D}_z\mathcal{I}(\ell(t),\overline{z}_{\tau_n}(t)) \rightharpoonup \xi(t) \text{ weakly in } \mathcal{V}^*$$

along a (not re-labelled) subsequence that might depend on t. Now, taking into account (3.1.27a) and (3.1.27c), we infer that we have for almost all $t \in I$ the convergences that are necessary to apply property $(\mathcal{I}_5)$ of the energy functional (cf. Lemma D.2), and obtain that

$$\mathrm{D}_z\mathcal{I}(\ell(t),\overline{z}_{\tau_n}(t)) \rightharpoonup \mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t)) \text{ weakly in } \mathcal{V}^*.$$

Using (3.1.29) again, we obtain (3.1.28) and therefore for all $t \in I$ the estimate

$$\begin{aligned}\|\mathrm{D}_z\mathcal{I}(\ell(t),z_\varepsilon(t))\|_{\mathcal{V}^*} &\le \liminf_{n\to\infty}\|\mathrm{D}_z\mathcal{I}(\ell(\overline{t}_{\tau_n}(t)),\overline{z}_{\tau_n}(t))\|_{\mathcal{V}^*}\\ &\le \sup_{\tau>0}\|\mathrm{D}_z\mathcal{I}(\ell(\overline{t}_\tau),\overline{z}_\tau)\|_{L^\infty(I;\mathcal{V}^*)}\\ &\le \operatorname{diam}_{\mathcal{V}^*}(\partial\mathcal{R}(0)) + m(\mathcal{I}(\ell(0),z_0),\operatorname{Var}_{\mathcal{V}^*}(\ell,I))\end{aligned}$$

and (3.1.7h) ensues.

It remains to prove the estimates (3.1.7a) - (3.1.7d). The estimate (3.1.7a) for the energies follows from (3.1.12) by lower semicontinuity, which then implies the boundedness of $\|z_\varepsilon\|_{L^\infty(I;\mathcal{Z})}$ in (3.1.7b) thanks to the coercivity estimate (D.6). Now, the energy dissipation balance (3.1.3) and the boundedness of the energies (3.1.7a) yield for every $0 < T < a$

$$\begin{aligned}\int_0^T \mathcal{R}_\varepsilon(\dot{z}_\varepsilon)\,\mathrm{d}r &+ \int_0^T \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(r),z_\varepsilon(r)))\,\mathrm{d}r\\ &\le \mathcal{I}(\ell(0),z_0) - \mathcal{I}(\ell(T),z_\varepsilon(T)) + \int_I \langle\dot{\ell}(r),z_\varepsilon(r)\rangle_{\mathcal{V}}\mathrm{d}r\\ &\le \mathcal{I}(\ell(0),z_0) - \mathcal{I}(\ell(T),z_\varepsilon(T)) + \operatorname{Var}_{\mathcal{V}^*}(\ell,I)\|z_\varepsilon\|_{L^\infty(I;\mathcal{V})}\end{aligned}$$

which implies that the dissipation terms are bounded as asserted in (3.1.7c). Finally, the estimate (3.1.7d) for the $H^1(I;\mathcal{V})$-norms of z_ε follows from (3.1.7c) and the definition of $\mathcal{R}_\varepsilon$ via the $\mathcal{V}$-norm on z_ε in (2.4.11).

■

As mentioned in the beginning of this section, a version of the results from Theorem 3.1.3 will be needed again in Section 4.3. To be more precise, we will need regularity properties and estimates for solutions of the system

$$0 \in \partial\mathcal{R}_1(\dot{z}(t)) + \mathrm{D}\mathcal{J}(z(t)) - \ell_*,\ t \in I$$

for a constant load $\ell_* \in \mathcal{V}^*$, where $\mathcal{J} : \mathcal{Z} \to \mathbb{R}$ is defined by

$$\mathcal{J}(z) := \tfrac{1}{2}\langle Az, z\rangle + \mathcal{F}(z) = \mathcal{I}(\ell, z) + \ell$$

and $\mathcal{R}_1(z) = \mathcal{R}(z) + \frac{1}{2}\|z\|_{\mathcal{V}}^2$. For this reason, we now give the following corollary for later use.

Corollary 3.1.4 (Solutions of the autonomous system (3.1.30))**.**

(i) Existence of solutions and regularity: For every $\ell_* \in \mathcal{V}^*$ and $z_0 \in \mathcal{Z}$ such that $\mathrm{D}\mathcal{J}(z_0) \in \mathcal{V}^*$, there exists a function $z \in L^\infty(I;\mathcal{Z})$ with $\dot{z} \in L^2(I;\mathcal{V})$ that satisfies $z(0) = z_0$ and the inclusion

$$0 \in \partial\mathcal{R}_1(\dot{z}(t)) + \mathrm{D}\mathcal{J}(z(t)) - \ell_*,\ t \in I \tag{3.1.30}$$

for almost all $t \in I$. Moreover, this solution belongs to $W^{1,\infty}(I;\mathcal{V})$, and it holds that $\mathrm{Var}_{\mathcal{Z}}(z;\overline{I})) < \infty$ and $\mathrm{D}\mathcal{J}(z(\cdot)) \in L^\infty(I;\mathcal{V}^*)$.

(ii) Uniqueness of solutions: For every $\ell_* \in \mathcal{V}^*$ and $z_0 \in \mathcal{Z}$ there exists at most one function $z \in L^\infty(I;\mathcal{Z})$ with $\dot{z} \in L^1(I;\mathcal{V})$ and $\mathrm{D}\mathcal{J}(z(\cdot)) \in L^\infty(I;\mathcal{V}^*)$ that satisfies $z(0) = z_0$ and the inclusion (3.1.30) *for almost all $t \in I$.*

(iii) Uniform estimates: There exist functions $m_1, m_2 : \mathcal{Z} \times \mathcal{V}^* \to [0,\infty)$ that map bounded sets on bounded sets such that for all $\ell_* \in \mathcal{V}$ and all $z_0 \in \mathcal{Z}$ with $\mathrm{D}\mathcal{J}(z_0) \in \mathcal{V}^*$ it holds: Let z be the solution of (3.1.30) *corresponding to (z_0, ℓ_*). Then*

$$\|z\|_{L^\infty(I;\mathcal{Z})} \le m_1(z_0, \ell_*), \tag{3.1.31}$$

$$\|\dot{z}\|_{L^\infty(I;\mathcal{V})} + \mathrm{Var}_{\mathcal{Z}}(z;\overline{I}) \le m_2(z_0,\ell_*)\Big(\mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}\mathcal{J}(z_0) + \ell_*, \partial\mathcal{R}(0)) + m_1(z_0,\ell_*)\Big), \tag{3.1.32}$$

$$\|\mathrm{D}\mathcal{J}(z(\cdot))\|_{L^\infty(I;\mathcal{V}^*)} \le \mathrm{diam}_{\mathcal{V}^*}(\partial\mathcal{R}(0)) + \|\ell_*\|_{\mathcal{V}^*} + C\|\dot{z}\|_{L^\infty(I;\mathcal{V})}. \tag{3.1.33}$$

Remark 3.1.5. *Let $z_0 \in \mathcal{Z}$, $\ell_* \in \mathcal{V}^*$ and assume that $-\mathrm{D}\mathcal{J}(z_0) + \ell_* \in \partial\mathcal{R}(0)$. Then the constant function $z(t) = z_0$, $t \in \overline{I}$, is the unique solution of* (3.1.30).

3.2 Vanishing viscosity analysis

The aim of this section is to perform the limit $\varepsilon \to 0$ in the regularized system (2.4.10). Let therefore z_ε be solutions of the viscous problem (2.4.10) in the sense of Def. 3.1.1. Following the discussion in Section 2.2.4, before extracting convergent subsequences from $(z_\varepsilon)_{\varepsilon>0}$, we parameterize the graph of each viscous solution by its respective arclength measured w.r.t. the dissipation associated with their paths. To be precise, we choose the following parameterization of the functions $(z_\varepsilon)_\varepsilon$: For $v \in \mathcal{Z}$ and $w \in \mathcal{V}^*$ let

$$\mathfrak{p}(v, w) := \mathcal{R}(v) + \|v\|_{\mathcal{V}}\mathrm{dist}_{\mathcal{V}^*}(w, \partial\mathcal{R}(0)), \tag{3.2.1a}$$

$$s_\varepsilon(t) := t + \int_0^t \mathfrak{p}(\dot z_\varepsilon(r), -\mathrm{D}_z\mathcal{I}(\ell(r), z_\varepsilon(r)))\,\mathrm{d}r, \tag{3.2.1b}$$

$$S_\varepsilon := s_\varepsilon(T), \tag{3.2.1c}$$

where $\mathfrak{p}(\cdot,\cdot)$ is the **vanishing viscosity contact potential** and fulfills for all $v \in \mathcal{Z}$ and $w \in \mathcal{V}^*$ (see [MRS12a, Rem. 3.1])

$$\mathfrak{p}(v,w) = \inf_{\varepsilon>0}\Big\{\mathcal{R}_\varepsilon(v) + \mathcal{R}_\varepsilon^*(w)\Big\} = \mathcal{R}(v) + \inf_{\varepsilon>0}\Big\{\frac{\varepsilon}{2}\|v\|_{\mathcal{V}}^2 + \frac{1}{2\varepsilon}\mathrm{dist}_{\mathcal{V}^*}(w, \partial\mathcal{R}(0))^2\Big\}.$$

In particular, with the Fenchel-Young inequality (A.1), it holds for all $v \in \mathcal{Z}$ and $w \in \mathcal{V}^*$ that

$$\langle w, v\rangle_{\mathcal{V}} \le \mathfrak{p}(v,w). \tag{3.2.2}$$

Thanks to Young's inequality, the formula for $\mathcal{R}_\varepsilon^*$ in Lemma 2.4.4, and the estimate (3.1.7c) for the dissipation terms, S_ε is uniformly bounded in ε, so that for $\varepsilon \to 0$, there exists a converging subsequence, whose limit we denote by S. Moreover, s_ε is a strictly increasing function in t. Hence, its inverse function

$$\hat t_\varepsilon := (s_\varepsilon)^{-1} : [0, S_\varepsilon] \to [0,T] \tag{3.2.3}$$

exists and we define

$$\hat z_\varepsilon := z_\varepsilon \circ \hat t_\varepsilon : [0,S_\varepsilon] \to \mathcal{Z} \quad \text{ by } \hat z_\varepsilon(s) := z_\varepsilon(\hat t_\varepsilon(s)). \tag{3.2.4}$$

Observe that $\hat t_\varepsilon$ are uniformly Lipschitz continuous, since for $s = s_\varepsilon(t)$ such that $\dot s_\varepsilon(t) \neq 0$, which are all $t > 0$, it holds

$$\dot{\hat t}_\varepsilon(s) = \frac{1}{\dot s_\varepsilon(\hat t_\varepsilon(s))} = \frac{1}{\dot s_\varepsilon(t)} \le 1. \tag{3.2.5}$$

In particular, $\dot{\hat t}_\varepsilon(s) \neq 0$ for all $s > 0$. After constantly continuating every $\hat t_\varepsilon$ to $[0,S]$, we assume for simplicity that $S_\varepsilon = S$. Now, the absolute continuity of the z_ε and Corollary C.5 imply that for all $\varepsilon > 0$, $\hat z_\varepsilon \in \mathrm{AC}([0,S];\mathcal{Z})$ with the following change of variable: Let $\varepsilon > 0$. Then it holds for all $\sigma \in [0,S]$ such that $\hat t_\varepsilon$ is differentiable in σ and z_ε is differentiable in $\hat t_\varepsilon(\sigma)$, that $\hat z_\varepsilon$ is differentiable in σ and $\dot{\hat z}_\varepsilon(\sigma) = \dot z_\varepsilon(\hat t_\varepsilon(\sigma))\dot{\hat t}_\varepsilon(\sigma)$. Thus, for all $0 \le r < s \le S$, we have:

$$\begin{aligned}\hat z_\varepsilon(s) - \hat z_\varepsilon(r) = z_\varepsilon(\hat t_\varepsilon(s)) - z_\varepsilon(\hat t_\varepsilon(r)) &= \int_{\hat t_\varepsilon(r)}^{\hat t_\varepsilon(s)} \dot z_\varepsilon(\tau)\,\mathrm{d}\tau = \int_{\hat t_\varepsilon(r)}^{\hat t_\varepsilon(s)} \dot z_\varepsilon(\hat t_\varepsilon(s_\varepsilon(\tau)))\,\mathrm{d}\tau \\ &= \int_r^s \dot z_\varepsilon(\hat t_\varepsilon(\sigma))\dot{\hat t}_\varepsilon(\sigma)\,\mathrm{d}\sigma = \int_r^s \dot{\hat z}_\varepsilon(\sigma)\,\mathrm{d}\sigma.\end{aligned}$$

Moreover, it holds for all $s \in [0, S_\varepsilon]$

$$
\begin{aligned}
\hat{t}_\varepsilon(s) + \int_0^s \mathfrak{p}(\dot{\hat{z}}_\varepsilon(r), -\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_\varepsilon(r)), \hat{z}_\varepsilon(r)))\,\mathrm{d}r \\
&= \hat{t}_\varepsilon(s) + \int_0^s \dot{\hat{t}}_\varepsilon(r)\mathfrak{p}(\dot{z}_\varepsilon(\hat{t}_\varepsilon(r)), -\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_\varepsilon(r)), z_\varepsilon(\hat{t}_\varepsilon(r))))\,\mathrm{d}r \\
&= \hat{t}_\varepsilon(s) + \int_0^{\hat{t}_\varepsilon(s)} \mathfrak{p}(\dot{z}_\varepsilon(r), -\mathrm{D}_z\mathcal{I}(\ell(r), z_\varepsilon(r)))\,\mathrm{d}r \\
&= s_\varepsilon(\hat{t}_\varepsilon(s)) \\
&= s,
\end{aligned}
$$

so that differentiating w.r.t. s on both sides yields

$$
\dot{\hat{t}}_\varepsilon(s) + \mathfrak{p}(\dot{\hat{z}}_\varepsilon(s), -\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_\varepsilon(s)), \hat{z}_\varepsilon(s))) = 1 \tag{3.2.6}
$$

for almost all $s \in [0, S_\varepsilon]$. We further use the abbreviation

$$
\mathfrak{e}(t, z) := \mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(t), z), \partial\mathcal{R}(0)), \tag{3.2.7}
$$

with the convention that $\|z\|_\mathcal{V}\mathfrak{e}(t,z) = 0$, if $\mathrm{D}_z\mathcal{I}(\ell(t), z) \in \mathcal{Z}^* \setminus \mathcal{V}^*$, so that we have

$$
\mathfrak{p}(v, -\mathrm{D}_z\mathcal{I}(\ell(t), z)) = \mathcal{R}(v) + \|v\|_\mathcal{V}\mathfrak{e}(t, z).
$$

We further denote by $\mathcal{P}(t,z) := \partial_t\mathcal{I}(\ell(t), z)$ the derivative of the scalar function $t \mapsto \mathcal{I}(\ell(t), z)$. Evaluating the energy dissipation balance (3.1.3) for z_ε at the times $t = \hat{t}_\varepsilon(s_2)$ and $s = \hat{t}_\varepsilon(s_1)$ now yields

$$
\begin{aligned}
\mathcal{I}(\ell(\hat{t}_\varepsilon(s_1)), &\hat{z}_\varepsilon(s_1)) - \mathcal{I}(\ell(\hat{t}_\varepsilon(s_2)), z_\varepsilon(\hat{t}_\varepsilon(s_2))) \\
&= \int_{\hat{t}_\varepsilon(s_1)}^{\hat{t}_\varepsilon(s_2)} \mathcal{R}_\varepsilon(\dot{z}_\varepsilon(\tau)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(\tau), z_\varepsilon(\tau)))\,\mathrm{d}\tau - \int_{\hat{t}_\varepsilon(s_1)}^{\hat{t}_\varepsilon(s_2)} \mathcal{P}(\tau, z_\varepsilon(\tau))\,\mathrm{d}\tau \\
&= \int_{s_1}^{s_2} \dot{\hat{t}}_\varepsilon(r)\Big[\mathcal{R}_\varepsilon(\dot{z}_\varepsilon(\hat{t}_\varepsilon(r))) + \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_\varepsilon(r)), z_\varepsilon(\hat{t}_\varepsilon(r))))\Big]\mathrm{d}r \\
&\qquad - \int_{s_1}^{s_2} \dot{\hat{t}}_\varepsilon(r)\mathcal{P}(\hat{t}_\varepsilon(r), z_\varepsilon(\hat{t}_\varepsilon(r)))\,\mathrm{d}r \\
&= \int_{s_1}^{s_2} \mathcal{R}(\dot{\hat{z}}_\varepsilon(r)) + \frac{\varepsilon}{\dot{\hat{t}}_\varepsilon(r)}\|\dot{\hat{z}}_\varepsilon(r)\|_\mathcal{V}^2 + \dot{\hat{t}}_\varepsilon(r)\mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_\varepsilon(r)), \hat{z}_\varepsilon(r)))\,\mathrm{d}r \\
&\qquad - \int_{s_1}^{s_2} \dot{\hat{t}}_\varepsilon(r)\mathcal{P}(\hat{t}_\varepsilon(r), \hat{z}_\varepsilon(r))\,\mathrm{d}r,
\end{aligned} \tag{3.2.8}
$$

which we write as

$$
\begin{aligned}
\mathcal{I}(\ell(\hat{t}_\varepsilon(s_1)), \hat{z}_\varepsilon(s_1)) = \mathcal{I}(\ell(\hat{t}_\varepsilon(s_2)), \hat{z}_\varepsilon(s_2)) &+ \int_{s_1}^{s_2} M_\varepsilon(\hat{t}_\varepsilon(s), \dot{\hat{z}}_\varepsilon(s), \mathfrak{e}(\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s))\,\mathrm{d}s \\
&- \int_{s_1}^{s_2} \partial_\ell\mathcal{I}(\ell(\hat{t}_\varepsilon(s)), \hat{z}_\varepsilon(s))\dot{\ell}(\hat{t}_\varepsilon(s))\dot{\hat{t}}_\varepsilon(s)\,\mathrm{d}s
\end{aligned}
$$

for

$$M_\varepsilon : \begin{cases} (0,\infty)\times\mathcal{V}\times[0,\infty) & \to [0,\infty] \\ (\alpha,v,\zeta) & \mapsto \mathcal{R}(v)+\frac{\varepsilon}{2\alpha}\|v\|_{\mathcal{V}}^2+\frac{\alpha}{2\varepsilon}\zeta^2 \end{cases}$$

Thus, a hypothetical limit $(\hat{t},\hat{z})$ of the sequence $(\hat{t}_\varepsilon,\hat{z}_\varepsilon)_\varepsilon$ should fulfill an estimate that is obtained by lower semicontinuity arguments, so that one might expect a functional M_0 such that

$$\begin{aligned}\mathcal{I}(\ell(\hat{t}(0)),z_0)\geq \mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s))+&\int_0^s M_0(\hat{t}(r),\dot{\hat{z}}(r),\mathfrak{c}(\hat{t}(r),\hat{z}(r)))\,\mathrm{d}r\\ -&\int_0^s \partial_\ell\mathcal{I}(\ell(\hat{t}(r)),\hat{z}(r))\dot{\ell}(\hat{t}(r))\dot{\hat{t}}(r)\,\mathrm{d}s.\end{aligned}$$

However, we are not able to show that the limiting state $\hat{z}$ is still differentiable w.r.t. the new variable s, so that we cannot give M_0 in dependence of $\dot{\hat{z}}(r)$, cf. the limiting energy dissipation principle (EDB) in Definition 3.2.5. This is due to the fact that we are only able to obtain an estimate for the differences $\hat{z}(s)-\hat{z}(r)$ measured w.r.t. the dissipation potential $\mathcal{R}$ (cf. (3.2.17)), which does not induce a reflexive topology on $\mathcal{V}$ in general. While we refer to Appendix C for a more detailed discussion of how this affects differentiability, at this point, we only give the necessary definitions.

Definition 3.2.1 ($\mathcal{R}$-absolutely continuous functions). *For a subset $K\subseteq\mathcal{V}$ and a subinterval $[a,b]\subseteq[0,T]$, we say that a curve $v:[a,b]\to K$ is $\mathcal{R}$-absolutely continuous, if there exists a nonnegative function $m\in L^1(a,b)$ such that*

$$\mathcal{R}(v(t)-v(s))\leq\int_s^t m(r)\,\mathrm{d}r \quad \text{for every } a\leq s<t\leq b \tag{3.2.9}$$

and denote by $\mathrm{AC}([a,b];K,\mathcal{R})$ the set of all $\mathcal{R}$-absolutely continuous curves $[a,b]\to K$ and by $\mathrm{AC}([a,b];\mathcal{R})$ the set of all $\mathcal{R}$-absolutely continuous curves $[a,b]\to\mathcal{V}$.

$\mathcal{R}$-absolutely continuous curves fulfill the following notion of differentiability:

Proposition 3.2.2 (Generalized metric derivatives). *[RMS08, Prop.2.2] For every curve $v\in\mathrm{AC}([a,b];K,\mathcal{R})$, the limit*

$$\mathcal{R}[v'](t):=\lim_{h\searrow 0}\mathcal{R}\left(\frac{v(t+h)-v(t)}{h}\right)=\lim_{h\searrow 0}\mathcal{R}\left(\frac{v(t)-v(t-h)}{h}\right)$$

exists almost everywhere and is called the ***generalized metric derivative*** *of v. Moreover, the function $t\mapsto\mathcal{R}[v'](t)$ belongs to $L^1(0,T)$, it is an admissible integrand in (3.2.9) and is minimal with this property, i.e., if m is another function satisfying (3.2.9), then $\mathcal{R}[v'](t)\leq m(t)$ almost everywhere.*

Definition 3.2.3. *For a subinterval $[a,b]\subseteq[0,T]$, we denote by $\mathrm{AC}^\infty([a,b];\mathcal{X})$ the set of all $\mathcal{R}$-absolutely continuous curves $v:[a,b]\to\mathcal{V}$ whose generalized metric derivative $\mathcal{R}[v']$ is an element of $L^\infty(a,b)$.*

Remark 3.2.4. *For a curve $v : [a,b] \to \mathcal{V}$, we use the notation*

$$\dot{v}(t) := \lim_{s\to t} \frac{v(t)-v(s)}{t-s} \in \mathcal{V} \qquad \textit{for the classical time-derivative.}$$

$$\mathcal{R}[v'](t) := \lim_{h\searrow 0} \mathcal{R}\left(\frac{v(t+h)-v(t)}{h}\right) \in \mathbb{R} \qquad \textit{for the generalized metric derivative.}$$

If v is classicaly differentiable in $t \in [0,T]$, then $\mathcal{R}(\dot{v}(t)) \geq \mathcal{R}[v'](t)$.

Absolutely continuous curves w.r.t the norm on $\mathcal{V}$ are $\mathcal{R}$-absolutely continuous.

The following Definition 3.2.5 shows that the function $\mathfrak{e}(\cdot,\cdot)$ defined in (3.2.7) is crucial for the characterization of the limit $(\hat{t},\hat{z})$. Namely, defining the set

$$G := \{s \in [0,S] \,|\, \mathfrak{e}(\hat{t}(s),\hat{z}(s)) > 0\},$$

we find that, on its complement $[0,S] \setminus G$, $\hat{z}$ can be understood as a solution of the relaxated problem

$$\partial\mathcal{R}(0) + \mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)) \ni 0, \tag{S$_{\text{loc}}$}$$

resulting in a somewhat weaker description of $\hat{z}$ here than on the set G. Formally, this is foreshadowed by the arclength (3.2.6) of the reparameterized solutions $\hat{z}_\varepsilon$, which yields the estimate

$$\|\dot{\hat{z}}_\varepsilon(s)\|_{\mathcal{V}}\mathfrak{e}(\dot{\hat{t}}_\varepsilon(s),\hat{z}_\varepsilon(s)) \leq 1 \quad \text{a.e. on } [0,S]. \tag{3.2.10}$$

On connected components $[a,b] \subset G$, we will later use (3.2.10) to derive local a priori estimates for $\|\dot{\hat{z}}_\varepsilon\|_{L^\infty([a,b];\mathcal{V})}$, which allow us to prove the differentiability of $\hat{z}$ on $[a,b]$, see the discussion around (3.2.16).

In anticipation of Theorem 3.2.6, we now give the following definition of $\mathfrak{p}$-parameterized BV solutions of the system (2.4.9).

Definition 3.2.5 ($\mathfrak{p}$-parameterized Balanced Viscosity solutions)**.** *Let $z_0 \in \mathcal{Z}$ and $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$. A triple $(S,\hat{t},\hat{z})$ with*

$$S > 0, \quad \hat{t} \in W^{1,\infty}((0,S),\mathbb{R}), \quad \textit{and } \hat{z} \in \mathrm{AC}([0,S];\mathcal{R}) \cap L^\infty(0,S;\mathcal{Z})$$

*is a $\mathfrak{p}$-**parameterized, normalized BV solution** of the rate-independent system* (2.4.9) *with data (z_0,ℓ), if the set*

$$\begin{aligned} G :&= \{s \in [0,S] \,|\, \mathfrak{e}(\hat{t}(s),\hat{z}(s)) > 0\} \qquad (3.2.11a)\\ &= \{s \in [0,S] \,|\, -\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)) \notin \partial\mathcal{R}(0)\} \end{aligned}$$

is a relatively open subset of $[0,S]$ such that

$$\hat{z} \in W^{1,1}_{loc}(G;\mathcal{V}), \tag{3.2.11b}$$

$$\mathrm{D}_z\mathcal{I}(\ell\circ\hat{t},\hat{z}) \in L^\infty_{loc}(G;\mathcal{V}^*). \tag{3.2.11c}$$

Furthermore, the following conditions shall be satisfied:

Complementarity and normalization condition: *For almost all $s \in [0,S]$, it holds*

$$\dot{\hat{t}}(s) \geq 0, \quad \hat{t}(0) = 0, \quad \hat{t}(S) = T, \quad \hat{z}(0) = z_0 \tag{3.2.12a}$$

$$\dot{\hat{t}}(s)\,\mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\partial\mathcal{R}(0)) = 0 \tag{3.2.12b}$$

$$1 = \begin{cases} \dot{\hat{t}}(s) + \mathcal{R}[\hat{z}'](s) + \|\dot{\hat{z}}(s)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(s),\hat{z}(s)), & \text{if } s \in G \\ \dot{\hat{t}}(s) + \mathcal{R}[\hat{z}'](s), & \text{if } s \in [0,S]\setminus G \end{cases} \tag{N}$$

Energy-dissipation balance: *For all $s \in [0,S]$, it holds*

$$\begin{aligned}\mathcal{I}(\ell(0),z_0) = \mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)) + \int_0^s \mathcal{R}[\hat{z}'](r) + \|\dot{\hat{z}}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(s),\hat{z}(s))\,\mathrm{d}r \\ - \int_0^s \partial_\ell\mathcal{I}(\ell(\hat{t}(r)),\hat{z}(r))\dot{\ell}(\hat{t}(r))\dot{\hat{t}}(r)\,\mathrm{d}r,\end{aligned} \tag{EDB}$$

where we adopt the convention that $\|\dot{\hat{z}}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(s),\hat{z}(s)) = 0$ for $s \in [0,S]\setminus G$.

With $\mathcal{L}(z_0,\ell)$ we denote the set of normalized, $\mathfrak{p}$-parameterized BV solutions associated with the pair (z_0,ℓ).

We will now prove that $\mathfrak{p}$-parameterized BV solutions exist. Apart from the estimate (3.2.13c), the arguments in the proof of following theorem are in essence similar to those in the proof of [MRS16, Thm. 4.3]. However, our additional assumptions on the energy functional allow for much more direct arguments and a notation with much higher readability.

Theorem 3.2.6 (Main existence result)**.** *Let (z_0,ℓ) comply with* (2.4.6)*. Let $(z_\varepsilon)_{\varepsilon>0}$ be a family of solutions of the regularized problem* (2.4.10) *satisfying the estimates* (3.1.7)*, and let $\hat{t}_\varepsilon$, $\hat{z}_\varepsilon$ be defined according to* (3.2.3)-(3.2.4)*. Then there exist a subsequence $\varepsilon_n \to 0$ and curves $\hat{t} \in W^{1,\infty}(0,S)$ and*

$$\hat{z} \in \mathrm{BV}([0,S];\mathcal{Z}) \cap \mathrm{AC}([0,S];\mathcal{R}) \cap C([0,S];\mathcal{V}) \cap C_{\mathrm{weak}}([0,S];\mathcal{Z})$$

such that $(S,\hat{t},\hat{z}) \in \mathcal{L}(z_0,\ell)$ and for $G = \{s \in [0,S] \,|\, \mathfrak{e}(\hat{t}(s),\hat{z}(s)) > 0\}$ and every connected component $[a,b] \subset G$, it holds:

$$\hat{t}_{\varepsilon_n} \overset{*}{\rightharpoonup} \hat{t} \text{ in } W^{1,\infty}(0,S), \tag{3.2.13a}$$

$$\hat{z}_{\varepsilon_n}(s) \rightharpoonup \hat{z}(s) \text{ in } \mathcal{Z} \text{ for all } s \in [0,S], \tag{3.2.13b}$$

$$\|\mathrm{D}_z\mathcal{I}(\ell \circ \hat{t},\hat{z})\|_{L^\infty(0,S;\mathcal{V}^*)} < \infty, \tag{3.2.13c}$$

$$\hat{z}_{\varepsilon_n} \overset{*}{\rightharpoonup} \hat{z} \text{ in } W^{1,\infty}(a,b;\mathcal{V}), \tag{3.2.13d}$$

$$\hat{z}_{\varepsilon_n} \to \hat{z} \text{ in } L^1(0,S;\mathcal{V}). \tag{3.2.13e}$$

Furthermore, it holds for all $0 \le r < s \le S$ *that*

$$\mathcal{I}(\ell(\hat{t}_{\varepsilon_n}(s)), \hat{z}_{\varepsilon_n}(s)) \to \mathcal{I}(\ell(\hat{t}(s)), \hat{z}(s)), \tag{3.2.13f}$$

$$\begin{aligned}\int_r^s \mathcal{R}(\dot{\hat{z}}_{\varepsilon_n}(\sigma)) + \frac{\varepsilon_n}{\dot{\hat{t}}_{\varepsilon_n}(\sigma)} \|\dot{\hat{z}}_{\varepsilon_n}(\sigma)\|_{\mathcal{V}}^2 + \dot{\hat{t}}_{\varepsilon_n}(\sigma) \mathcal{R}^*_{\varepsilon_n}(-\mathrm{D}_z \mathcal{I}(\ell(\hat{t}_{\varepsilon_n}(\sigma)), \hat{z}_{\varepsilon_n}(\sigma))) \,\mathrm{d}\sigma \\ \to \int_r^s \mathcal{R}[\hat{z}'](\sigma) + \|\dot{\hat{z}}(\sigma)\|_{\mathcal{V}} \mathrm{dist}(-\mathrm{D}_z \mathcal{I}(\ell(\hat{t}(\sigma)), \hat{z}(\sigma)), \partial\mathcal{R}(0)) \,\mathrm{d}\sigma\,.\end{aligned} \tag{3.2.13g}$$

Proof. **Step 1: Compactness of rescaled solutions** (3.2.13a) is clear thanks to the uniform Lipschitz estimate (3.2.5) for the $\hat{t}_\varepsilon$. As a first step, we apply Proposition C.13 to the sequence $(\hat{z}_n)_{n\in\mathbb{N}}$ and thus obtain the existence of an element $\hat{z} \in C([0,S];\mathcal{V}) \cap C_{\mathrm{weak}}([0,S];\mathcal{Z})$ such that $\hat{z}_n \to \hat{z}$ uniformly in $C([0,S];\mathcal{V})$ and (3.2.13b) is fulfilled. Thanks to Lemma C.16 and Remark C.17, the absolute continuity of the $\hat{z}_\varepsilon$ w.r.t. the norm on $\mathcal{Z}$ and the pointwise weak convergence (3.2.13b) are sufficient to conclude that $\hat{z} \in \mathrm{BV}([0,S],\mathcal{Z})$.

Next, we want to show the following implication:

$$\begin{array}{ll}\underline{\text{If}} & t_n \to t \text{ in } \mathbb{R},\, z_n \rightharpoonup z \text{ in } \mathcal{Z} \text{ and } \sup\limits_{n\in\mathbb{N}} \|\mathrm{D}_z\mathcal{I}(\ell(t_n), z_n)\|_{\mathcal{V}^*} < \infty \\ \underline{\text{then}} & \mathrm{D}_z\mathcal{I}(\ell(t_n), z_n) \rightharpoonup \mathrm{D}_z\mathcal{I}(\ell(t), z) \text{ in } \mathcal{V}^* \text{ and } \mathfrak{e}(t,z) \le \liminf\limits_{n\to\infty} \mathfrak{e}(t_n, z_n).\end{array} \tag{3.2.14}$$

To this end, we note that the map

$$d : \begin{cases} \mathcal{V}^* \to \mathbb{R} \\ \xi \mapsto \mathrm{dist}_{\mathcal{V}^*}(\xi, \partial\mathcal{R}(0)) \end{cases}$$

is convex and continuous w.r.t. the norm on $\mathcal{V}^*$, hence lower semicontinuous w.r.t. the weak topology on $\mathcal{V}^*$. Thus, it remains to show the weak convergence $\mathrm{D}_z\mathcal{I}(\ell(t_n), z_n) \rightharpoonup \mathrm{D}_z\mathcal{I}(\ell(t), z)$ in $\mathcal{V}^*$. From the convergence $z_n \rightharpoonup z$ in $\mathcal{Z}^*$ and the assumptions on A and $\mathcal{F}$, it follows immediately that $Az_n + \mathrm{D}\mathcal{F}(z_n) \rightharpoonup Az + \mathrm{D}\mathcal{F}(z)$ in $\mathcal{Z}^*$. Since ℓ is continuous w.r.t. the norm on $\mathcal{V}^*$, $\mathrm{D}_z\mathcal{I}(\ell(t_n), z_n) \rightharpoonup \mathrm{D}_z\mathcal{I}(\ell(t), z)$ in $\mathcal{Z}^*$ ensues. Now, the boundedness of the sequence $\mathrm{D}_z\mathcal{I}(\ell(t_n), z_n)$ w.r.t. the norm on $\mathcal{V}^*$ implies the existence of a subsequence converging weakly in $\mathcal{V}^*$, whose limit has to coincide with that in $\mathcal{Z}^*$. Therefore, the entire sequence weakly converges to $\mathrm{D}_z\mathcal{I}(\ell(t), z)$ in $\mathcal{V}^*$ and (3.2.14) is proven.

Now, the a priori estimate (3.1.7h) implies that

$$\sup_{\varepsilon>0, s\in[0,S]} \|\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_\varepsilon(s)), \hat{z}_\varepsilon(s))\|_{\mathcal{V}^*} < \infty$$

holds true. By the first part of (3.2.14), we conclude that

$$\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_\varepsilon(s)), \hat{z}_\varepsilon(s)) \rightharpoonup \mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)), \hat{z}(s)) \text{ in } \mathcal{V}^* \text{ for all } s \in [0,S],$$

so that the limit element has to be an element of $\mathcal{V}^*$ as well, fulfilling (3.2.13c).

We now choose a connected component $[a,b] \subset G = \{s \in [0,S] \,|\, \mathfrak{e}(\hat{t}(s),\hat{z}(s)) > 0\}$ and a sequence $(s_n)_{n\in\mathbb{N}} \subset [a,b]$ such that

$$c := \inf_{s\in[a,b]} \mathfrak{e}(\hat{t}(s),\hat{z}(s)) = \lim_{n\to\infty} \mathfrak{e}(\hat{t}(s_n),\hat{z}(s_n))$$

and convince ourselves that $c > 0$: For every subsequence $(s_{n_k})_{k\in\mathbb{N}}$ that converges to a point σ, we conclude that σ is an element of the closed interval $[a,b]$ and apply (3.2.14) to $t_k := \hat{t}(s_{n_k})$ and $z_k := \hat{z}(s_{n_k})$ and obtain

$$\begin{aligned} c &= \inf_{s\in[a,b]} \mathfrak{e}(\hat{t}(s),\hat{z}(s)) = \lim_{n\to\infty} \mathfrak{e}(\hat{t}(s_n),\hat{z}(s_n)) \\ &= \lim_{k\to\infty} \mathfrak{e}(\hat{t}(s_{n_k}),\hat{z}(s_{n_k})) = \liminf_{k\to\infty} \mathfrak{e}(\hat{t}(s_{n_k}),\hat{z}(s_{n_k})) \\ &\geq \mathfrak{e}(\hat{t}(\sigma),\hat{z}(\sigma)) > 0. \end{aligned}$$

Using (3.2.14) again, we infer that

$$\liminf_{n\to\infty} \mathfrak{e}(\hat{t}_{\varepsilon_n}(s),\hat{z}_{\varepsilon_n}(s)) \geq \mathfrak{e}(\hat{t}(s),\hat{z}(s)) \geq c, \tag{3.2.15}$$

which implies that

$$\forall\, s \in [a,b] \,\exists\, N \in \mathbb{N} \,\forall\, n \geq N : \mathfrak{e}(\hat{t}_{\varepsilon_n}(s),\hat{z}_{\varepsilon_n}(s)) \geq \frac{c}{2}.$$

Plugging this into (3.2.6) gives

$$\begin{aligned} &\forall\, s \in [a,b] \,\exists\, N \in \mathbb{N} \,\forall\, n \geq N : \\ &\qquad \frac{c}{2}\|\dot{\hat{z}}_{\varepsilon_n}(s)\|_{\mathcal{V}} \leq \dot{\hat{t}}_{\varepsilon_n}(s) + \mathfrak{e}(\hat{t}_{\varepsilon_n}(s),\hat{z}_{\varepsilon_n}(s))\|\dot{\hat{z}}_{\varepsilon_n}(s)\|_{\mathcal{V}} + \mathcal{R}(\dot{\hat{z}}_{\varepsilon_n}(s)) = 1, \end{aligned} \tag{3.2.16}$$

and thus

$$\forall\, s \in [a,b] : \limsup_{n\to\infty} \|\dot{\hat{z}}_{\varepsilon_n}(s)\|_{\mathcal{V}} \leq \frac{2}{c},$$

implying that $\hat{z}_{\varepsilon_n}$ are uniformly bounded on $[a,b]$ in $W^{1,\infty}(a,b;\mathcal{V})$ and we obtain (3.2.13d). The strong convergence (3.2.13e) follows from (3.2.13b) and the uniform bound (3.1.7b) in $L^\infty(0,T;\mathcal{Z})$ by means of Vitali's convergence theorem, see Lemma B.1 in Appendix B. Lastly, the estimate

$$\mathcal{R}(\hat{z}(s)-\hat{z}(r)) \leq \lim_{n\to\infty} \mathcal{R}(\hat{z}_{\varepsilon_n}(s)-\hat{z}_{\varepsilon_n}(r)) \leq \limsup_{n\to\infty} \int_r^s \underbrace{\mathcal{R}(\dot{\hat{z}}_{\varepsilon_n}(\sigma))}_{\leq 1} \mathrm{d}\sigma \leq s-r \tag{3.2.17}$$

for all $0 \leq r < s \leq S$ proves that $z \in \mathrm{AC}([0,S];\mathcal{R})$ with $\mathcal{R}[\hat{z}'] \leq 1$ almost everywhere.

Step 2: Energy dissipation balance - upper bound Young's inequality implies that

$$\|\dot{\hat{z}}_{\varepsilon_n}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}_{\varepsilon_n}(r),\hat{z}_{\varepsilon_n}(r)) \leq \frac{\varepsilon_n}{\dot{\hat{t}}_{\varepsilon_n}(r)}\|\dot{\hat{z}}_{\varepsilon_n}(r)\|_{\mathcal{V}}^2 + \dot{\hat{t}}_{\varepsilon_n}(r)\mathcal{R}^*_{\varepsilon_n}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_{\varepsilon_n}(r)),\hat{z}_{\varepsilon_n}(r))) \tag{3.2.18}$$

for all $r \in [0,S]$. Since weak*-convergence in $W^{1,\infty}(a,b;\mathcal{V})$ implies weak convergence in $L^1(a,b;\mathcal{V})$, we conclude by means of Lemma B.2 from (3.2.15) and (3.2.13d) that for all $[a,b] \subset G$ it holds

$$\int_a^b \|\dot{\hat{z}}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(r),\hat{z}(r))\,\mathrm{d}r \le \liminf_{n\to\infty} \int_a^b \|\dot{\hat{z}}_{\varepsilon_n}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}_{\varepsilon_n}(r),\hat{z}_{\varepsilon_n}(r))\,\mathrm{d}r.$$

Using the convention that $\|\dot{\hat{z}}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(r),\hat{z}(r)) = 0$ if $r \in [0,S]\setminus G$, the same inequality holds for the integral on arbitrary intervals in $[0,S]$. Due to Lemma C.16 we also know that

$$\int_0^s \mathcal{R}[\hat{z}'](\sigma)\,\mathrm{d}\sigma = \mathrm{Var}_{\mathcal{R}}(\hat{z},[0,s]) \le \liminf_{n\to\infty} \mathrm{Var}_{\mathcal{R}}(\hat{z}_{\varepsilon_n},[0,s]) = \liminf_{n\to\infty} \int_0^s \mathcal{R}[\hat{z}'_{\varepsilon_n}](\sigma)\,\mathrm{d}\sigma,$$

allowing us to conclude that

$$\begin{aligned}
&\int_0^s \mathcal{R}[\hat{z}'](r) + \|\dot{z}(r)\|_{\mathcal{V}}\mathfrak{e}(t(r),z(r))\,\mathrm{d}r \\
&\quad\le \liminf_{n\to\infty} \int_0^s \mathcal{R}(\dot{\hat{z}}_{\varepsilon_n}(r)) + \|\dot{\hat{z}}_{\varepsilon_n}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}_{\varepsilon_n}(r),\hat{z}_{\varepsilon_n}(r))\,\mathrm{d}r \qquad (3.2.19)\\
&\quad\le \liminf_{n\to\infty} \int_0^s \mathcal{R}(\dot{\hat{z}}_{\varepsilon_n}(r)) + \frac{\varepsilon_n}{\dot{\hat{t}}_{\varepsilon_n}(r)}\|\dot{\hat{z}}_{\varepsilon_n}(r)\|_{\mathcal{V}}^2 + \dot{\hat{t}}_{\varepsilon_n}(r)\mathcal{R}^*_{\varepsilon_n}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_{\varepsilon_n}(r)),\hat{z}_{\varepsilon_n}(r)))\,\mathrm{d}r.
\end{aligned}$$

Now, in order to show the convergence of the power term, we calculate for fixed $z \in \mathcal{Z}$ and point of differentiability $r \in [0,S]$

$$-\dot{\hat{t}}(r)\mathcal{P}(\hat{t}(r),z) = -\partial_\ell\mathcal{I}(\ell(\hat{t}(r)),z)\dot{\ell}(\hat{t}(r))\dot{\hat{t}}(r) = \langle\dot{\ell}(\hat{t}(r)),z\rangle_{\mathcal{V}}\dot{\hat{t}}(r).$$

Now, Lemma E.1 implies that

$$\dot{\hat{t}}_{\varepsilon_n}\cdot(\dot{\ell}\circ\hat{t}_{\varepsilon_n}) = (\ell\circ\hat{t}_{\varepsilon_n})' \overset{*}{\rightharpoonup} (\ell\circ\hat{t})' = \dot{\hat{t}}\cdot(\dot{\ell}\circ\hat{t}) \text{ in } L^\infty(0,S;\mathcal{V}^*),$$

which together with the strong convergence of $\hat{z}_{\varepsilon_n}$ to $\hat{z}$ in $L^1(0,S;\mathcal{V})$ according to (3.2.13e) is sufficient to conclude

$$\int_0^s \langle\dot{\ell}(\hat{t}(r)),\hat{z}(r)\rangle_{\mathcal{V}}\dot{t}(r)\,\mathrm{d}r = \lim_{n\to\infty}\int_0^s \langle\dot{\ell}(\hat{t}_{\varepsilon_n}(r)),\hat{z}_{\varepsilon_n}(r)\rangle_{\mathcal{V}}\dot{\hat{t}}_{\varepsilon_n}(r)\,\mathrm{d}r. \qquad (3.2.20)$$

Finally, we use the fact that the nonconvex part $\mathcal{F}$ of the energy $\mathcal{I}$ is weakly continuous according to Lemma 2.4.2 and obtain

$$\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)) \le \liminf_{n\to\infty}\mathcal{I}(\ell(\hat{t}_{\varepsilon_n}(s)),\hat{z}_{\varepsilon_n}(s)), \qquad (3.2.21)$$

so that (t,z) fulfills the energy dissipation estimate

$$\begin{aligned}
\mathcal{I}(\ell(0),z_0) \ge \mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)) &+ \int_0^s \mathcal{R}[\hat{z}'](r) + \|\dot{\hat{z}}(r)\|_{\mathcal{V}}\mathrm{dist}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(r)),\hat{z}(r)),\partial\mathcal{R}(0))\,\mathrm{d}r \\
&- \int_0^s \partial_\ell\mathcal{I}(\ell(\hat{t}(r)),\hat{z}(r))\dot{\ell}(\hat{t}(r))\dot{\hat{t}}(r)\,\mathrm{d}r.
\end{aligned}$$

Step 3: The limit fulfills the complementarity condition A change of variable in the a priori estimate (3.1.7c) for the dissipation terms shows that

$$0 \leq \int_0^{S_{\varepsilon_n}} \dot{\hat{t}}_{\varepsilon_n}(s)\mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_{\varepsilon_n}(s)),\hat{z}_{\varepsilon_n}(s)),\partial\mathcal{R}(0))^2\,\mathrm{d}s \leq \varepsilon_n C,$$

where C is independent of ε_n, whereby the above integral converges to 0 with $n \to \infty$. Now, thanks to the liminf-estimate (3.2.15) for $\mathfrak{e}$, Lemma B.2 implies that

$$\begin{aligned} 0 &= \lim_{n\to\infty}\int_0^{S_{\varepsilon_n}} \dot{\hat{t}}_{\varepsilon_n}(s)\mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_{\varepsilon_n}(s)),\hat{z}_{\varepsilon_n}(s)),\partial\mathcal{R}(0))^2\,\mathrm{d}s \\ &\geq \int_0^{S} \dot{\hat{t}}(s)\mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\partial\mathcal{R}(0))^2\,\mathrm{d}s \geq 0, \end{aligned}$$

and the complementarity condition (3.2.12b) ensues.

Step 4: Energy dissipation balance - lower bound For the opposite estimate in the energy dissipation balance, we use the chain rule inequality from Proposition F.1 in the appendix, which implies for all $s \in [0,S]$ the estimate

$$-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)) + \mathcal{P}(\hat{t}(s),\hat{z}(s))\dot{\hat{t}}(s) \leq \mathcal{R}[\hat{z}'](s) + \|\dot{\hat{z}}(s)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(s),\hat{z}(s)),$$

and we infer the lower bound by integration w.r.t. time.

Step 5: Improved convergences Next, we want to show the convergences (3.2.13f) and (3.2.13g) of the energies and the dissipation terms, since so far, we have only established liminf-inequalities in (3.2.19) and (3.2.21). However, this follows easily from the liminf-inequalities and the fact that the energy dissipation balance (EDB) implies that

$$\begin{aligned} \lim_{n\to\infty}\int_0^s &\mathcal{R}(\dot{\hat{z}}_{\varepsilon_n}(r)) + \frac{\varepsilon_n}{\dot{\hat{t}}_{\varepsilon_n}(r)}\|\dot{\hat{z}}_{\varepsilon_n}(r)\|_{\mathcal{V}}^2 + \dot{\hat{t}}_{\varepsilon_n}(r)\mathcal{R}^*_{\varepsilon_n}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_{\varepsilon_n}(r)),\hat{z}_{\varepsilon_n}(r)))\,\mathrm{d}r + \mathcal{I}(\ell(\hat{t}_{\varepsilon_n}(s)),\hat{z}_{\varepsilon_n}(s)) \\ &= \mathcal{I}(\ell(0),z_0) + \int_0^s \partial_\ell\mathcal{I}(\ell(\hat{t}(r)),\hat{z}(r))\dot{\ell}(\hat{t}(r))\dot{\hat{t}}(r)\,\mathrm{d}r \\ &= \int_0^s \mathcal{R}[\hat{z}'](r) + \|\dot{\hat{z}}(r)\|_{\mathcal{V}}\mathrm{dist}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(r)),\hat{z}(r)),\partial\mathcal{R}(0))\,\mathrm{d}r + \mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)). \end{aligned}$$

Step 6: The limit is normalized It remains to show that the normalization condition (N) is fulfilled. On the one hand, the liminf-inequality (3.2.19) and the normalization condition (3.2.6) for the parameterized viscous solutions $\hat{z}_\varepsilon$ yield for all $0 \leq a < b \leq S$:

$$\begin{aligned} \int_a^b &\mathcal{R}[\hat{z}'](r) + \|\dot{\hat{z}}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(r),\hat{z}(r))\,\mathrm{d}r \\ &\leq \liminf_{n\to\infty}\int_a^b \mathcal{R}[\hat{z}'_{\varepsilon_n}](r) + \|\dot{\hat{z}}_{\varepsilon_n}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}_{\varepsilon_n}(r),\hat{z}_{\varepsilon_n}(r))\,\mathrm{d}r \\ &= \liminf_{n\to\infty}\int_a^b 1 - \dot{\hat{t}}_{\varepsilon_n}(r)\,\mathrm{d}r = \int_a^b 1 - \dot{\hat{t}}(r)\,\mathrm{d}r. \end{aligned}$$

Since the integrands on both the left hand and the right hand side are in $L^1(0,S)$, we conclude for almost all $s \in [0,S]$:

$$\begin{aligned}\mathcal{R}[\hat{z}'](s) + \|\dot{\hat{z}}(s)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(s),\hat{z}(s)) &= \lim_{h\to 0}\frac{1}{h}\int_s^{s+h}\mathcal{R}[\hat{z}'](r) + \|\dot{\hat{z}}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(r),\hat{z}(r))\,\mathrm{d}r\\ &\le \lim_{h\to 0}\frac{1}{h}\int_s^{s+h} 1-\dot{\hat{t}}(r)\,\mathrm{d}r = 1-\dot{\hat{t}}(s). \end{aligned}\tag{3.2.22}$$

On the other hand, Young's inequality as well as the convergence (3.2.13g) of the dissipation terms implies for all $0 \le a < b \le S$:

$$\begin{aligned}\int_a^b 1-\dot{\hat{t}}(r)\,\mathrm{d}r &= \lim_{n\to\infty}\int_a^b 1-\dot{\hat{t}}_{\varepsilon_n}(r)\,\mathrm{d}r\\ &= \lim_{n\to\infty}\int_a^b \mathcal{R}[\hat{z}'_{\varepsilon_n}](r) + \|\dot{\hat{z}}_{\varepsilon_n}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}_{\varepsilon_n}(r),\hat{z}_{\varepsilon_n}(r))\,\mathrm{d}r\\ &\le \lim_{n\to\infty}\int_a^b \mathcal{R}[\hat{z}'_{\varepsilon_n}](r) + \frac{\varepsilon_n}{\dot{\hat{t}}_{\varepsilon_n}(r)}\|\dot{\hat{z}}_{\varepsilon_n}(r)\|_{\mathcal{V}}^2 + \dot{\hat{t}}_{\varepsilon_n}(r)\mathcal{R}^*_{\varepsilon_n}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}_{\varepsilon_n}(r)),\hat{z}_{\varepsilon_n}(r)))\,\mathrm{d}r\\ &= \int_a^b \mathcal{R}[\hat{z}'](r) + \|\dot{\hat{z}}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(r),\hat{z}(r))\,\mathrm{d}r,\end{aligned}$$

and the same localization argument as in (3.2.22) gives the opposite pointwise estimate.

Step 7: G is a relatively open set Finally, we verify that G as defined in (3.2.11a) is indeed an open set. To this end, we choose $s \in G$ and a sequence $(s_n)_{n\in\mathbb{N}} \subset [0,S]$ such that $s_n \to s$. We can now use the implication (3.2.14), since $\hat{z} \in C_{\mathrm{weak}}([0,S],\mathcal{Z})$ and we already have proven the estimate (3.2.13c), and conclude $\liminf_{n\to\infty}\mathfrak{e}(s_n,\hat{z}(s_n)) \ge \mathfrak{e}(s,\hat{z}(s))$. In other words, there exists $N \in \mathbb{N}$ such that for all $n \ge N$ it holds that $\mathfrak{e}(s_n,\hat{z}(s_n)) \ge \mathfrak{e}(s,\hat{z}(s))$.

Now, for every $\sigma \in G$ an argument by contradiction shows that there exists a radius $r > 0$ small enough, such that for each $s \in B_r(\sigma)\cap[0,S]$, we have that $\mathfrak{e}(s,\hat{z}(s)) \ge \frac{1}{2}\mathfrak{e}(\sigma,\hat{z}(\sigma)) > 0$, i.e., G is relatively open in $[0,S]$. ■

3.3 Equivalent characterizations

The following differential characterization of $\mathfrak{p}$-parameterized BV solutions is classical, cf. [Mie11] or [MRS16, Prop. 4.6] in a more general setting. It is crucial to obtain the a priori estimates necessary for the proof of compactness of the solution set, and therefore ultimately for proving existence of a solution of the optimal control problem. Hence, we will give a proof here for completeness.

Lemma 3.3.1 (Differential characterization)**.** *If $(S,\hat{t},\hat{z})$ is a normalized $\mathfrak{p}$-parameterized BV solution of the system (2.4.9), then $\hat{t}$ is constant on each connected component of G, and there exists a measurable function $\lambda : (0,S) \to [0,\infty)$ with $\lambda(s) = 0$ on $[0,S]\setminus G$ and such that on each connected component $(a,b) \subset G$ the differential* *inclusion*

$$0 \in \partial\mathcal{R}(\dot{\hat{z}}(s)) + \partial\mathcal{R}_2(\lambda(s)\dot{\hat{z}}(s)) + \mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)) \quad \textit{and} \quad \dot{t}(s)\lambda(s) = 0 \tag{3.3.1}$$

holds true for almost all $s \in (a,b)$, *where* $\mathcal{R}_2(v) := \frac{1}{2}\|v\|_{\mathcal{V}}^2$. *Furthermore, for almost all* $s \in G$, *in holds that*

$$\lambda(s) = \mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\partial\mathcal{R}(0)) \Big/ \left\|\dot{\hat{z}}(s)\right\|_{\mathcal{V}}.$$

Conversely, if an absolutely continuous curve $(\hat{t},\hat{z}) : [0,S] \to [0,T]\times\mathcal{Z}$ *satisfies* (3.3.1) *for almost all* $s \in [0,S]$ *for a measurable function* $\lambda : (0,S) \to [0,\infty)$, *and the map* $s \mapsto \mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s))$ *is absolutely continuous, and if further* $(\hat{t},\hat{z})$ *are non-degenerate in the sense that*

$$\dot{\hat{t}}(s) + \mathcal{R}(\dot{\hat{z}}(r)) + \|\dot{\hat{z}}(r)\|_{\mathcal{V}}\mathfrak{c}(\hat{t}(s),\hat{z}(s)) > 0 \textit{ for all } s \in [0,S], \tag{3.3.2}$$

then $(S,\hat{t},\hat{z})$ *is a* $\mathfrak{p}$*-parameterized BV solution to the system* (2.4.9).

Proof. Let us first assume that we are given a normalized $\mathfrak{p}$-parameterized BV solution $(S,\hat{t},\hat{z})$ of (2.4.9). Since $\mathfrak{c}(\hat{t}(s),\hat{z}(s)) > 0$ on G, from the complementarity condition (3.2.12b) we deduce that $\hat{t}$ is constant on each connected component of G. In order to verify (3.3.1), let $[a,b]$ be such a connected component. Since by assumption $\hat{z} \in W^{1,1}((a,b);\mathcal{V})$, we have $\mathcal{R}[\hat{z}'](s) = \mathcal{R}(\dot{\hat{z}}(s))$ for almost all $s \in (a,b)$, cf. [AGS05, Remark 1.1.3]. Thus, localizing the energy dissipation identity (EDB) (where we apply the chain rule formulated in Proposition F.2) yields

$$\mathcal{R}(\dot{\hat{z}}(s)) + \langle \mathrm{D}_z\mathcal{I}(\hat{t}(s),\hat{z}(s)),\dot{\hat{z}}(s)\rangle_{\mathcal{V}} + \left\|\dot{\hat{z}}(s)\right\|_{\mathcal{V}} \mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\partial\mathcal{R}(0)) = 0, \tag{3.3.3}$$

which is valid for almost all $s \in (a,b)$. Since $\hat{t}$ is constant on (a,b), from the normalization condition (N) (cf. Def. 3.2.5) it follows that $\dot{\hat{z}}(s) \neq 0$ almost everywhere on (a,b). Hence, with

$$\lambda(s) = \begin{cases} \mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\partial\mathcal{R}(0))/\left\|\dot{\hat{z}}(s)\right\|_{\mathcal{V}}, & \text{if } \dot{\hat{z}}(s) \neq 0, \\ 0, & \text{otherwise,} \end{cases}$$

we have $\langle \lambda(s)\mu(s),\dot{\hat{z}}(s)\rangle_{\mathcal{V}} = \left\|\dot{\hat{z}}(s)\right\|_{\mathcal{V}} \mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\partial\mathcal{R}(0))$ for every element $\mu(s) \in \partial\mathcal{R}_2(\dot{z}(s))$. Multiplying (3.3.3) with $\lambda(s)$ and taking the one-homogeneity of $\mathcal{R}$, as well as the characterization of the convex dual $\left(\mathcal{R}+\mathcal{R}_2\right)^*$ from Lemma 2.4.4 into account, we obtain

$$\begin{aligned} &\langle -\mathrm{D}_z\mathcal{I}(\hat{t}(s),\hat{z}(s)),\lambda(s)\dot{\hat{z}}(s)\rangle_{\mathcal{V}} \\ &\qquad = \mathcal{R}(\lambda(s)\dot{\hat{z}}(s)) + \mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\partial\mathcal{R}(0))^2 \\ &\qquad = \mathcal{R}(\lambda(s)\dot{\hat{z}}(s)) + \tfrac{1}{2}\|\lambda(s)\dot{\hat{z}}(s)\|_{\mathcal{V}}^2 + \tfrac{1}{2}\,\mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\partial\mathcal{R}(0))^2 \\ &\qquad = \left(\mathcal{R}+\mathcal{R}_2\right)\left(\lambda(s)\dot{\hat{z}}(s)\right) + \left(\mathcal{R}+\mathcal{R}_2\right)^*\left(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s))\right). \end{aligned}$$

By means of the Fenchel equivalence (A.2), we conclude

$$-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)) \in \partial\Big(\mathcal{R}+\mathcal{R}_2\Big)\Big(\lambda(s)\dot{\hat{z}}(s)\Big)$$

which implies (3.3.1), since $\partial\mathcal{R}(\cdot)$ is 0-homogeneous.

For the proof of the converse, following the proof of [MRS16, Prop. 4.6], we first observe that under the given assumptions, if $s \in [0,S]$ is chosen such that $\lambda(s) = 0$ and $\hat{z}$ is differentiable, then $\mathfrak{e}(\hat{t}(s),\hat{z}(s)) = 0$, and since this implies $\langle -\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\dot{\hat{z}}(s)\rangle_{\mathcal{V}} = \mathcal{R}(\dot{\hat{z}}(s))$ according to Lemma 2.4.4, we infer (3.3.3). On the other hand, if $\lambda(s) > 0$, then $\dot{\hat{t}}(s) = 0$, and if $\hat{z}$ is differentiable at s, we infer that $\dot{\hat{z}}(s) \neq 0$ from the non-degeneracy (3.3.2). What is more, applying first the Fenchel-equivalence (A.2) and then the inequality (3.2.2), we obtain

$$\begin{aligned}\langle -\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\dot{\hat{z}}(s)\rangle_{\mathcal{V}} &= \mathcal{R}(\dot{\hat{z}}(s)) + \frac{\lambda(s)}{2}\|\dot{\hat{z}}(s)\|_{\mathcal{V}}^2 + \frac{1}{2\lambda(s)}\mathfrak{e}(\hat{t}(s),\hat{z}(s))^2 \\ &\geq \mathcal{R}(\dot{\hat{z}}(s)) + \|\dot{\hat{z}}(s)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(s),\hat{z}(s)) \\ &\geq \langle -\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\dot{\hat{z}}(s)\rangle_{\mathcal{V}},\end{aligned}$$

and consequently, all inequalities must indeed have been equalities. Therefore, we must have $\mathfrak{e}(\hat{t}(s),\hat{z}(s)) > 0$, since otherwise, the second identity above would not be fulfilled. We can now infer that (3.3.3) holds almost everywhere on $[0,S]$. Since $s \mapsto \mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s))$ is presupposed to be absolutely continuous, it is differentiable almost everywhere and its derivative

$$\tfrac{\mathrm{d}}{\mathrm{d}s}\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)) = \langle \mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\dot{\hat{z}}(s)\rangle_{\mathcal{V}} + \partial_\ell\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s))\dot{\ell}(\hat{t}(s))\dot{\hat{t}}(s)$$

is an element of $L^1(0,S)$, so that we may infer the energy dissipation balance (EDB). ■

As was already pointed at the end of Section 2.2.4, this differential representation allows for the interpretation of jumps in the rate-independent system as transitions between two end points along a curve following a viscous regime, with an explicit description of this transition given by the curve $(\hat{t},\hat{z})$.

Let us now return to the connection between BV and $\mathfrak{p}$-parameterized BV solutions that was hinted at at the end of Section 2.2.4. First, we need to update the definition of BV solutions to our infinite dimensional setting. Just as in Section 2.2.4, BV solutions are characterized by a local stability condition $(\mathrm{S})_{\mathrm{loc}}$ and an energy dissipation balance (3.3.5), which is given in terms of the (pseudo-) total variation $\mathrm{Var}_{\mathfrak{p},\mathcal{I}}$ induced by $\mathfrak{p}$ and $\mathcal{I}$. However, the discussion around the proof of (3.2.11) suggests that we need to adjust the definition (2.2.23) of the Finsler cost in the following way: The first issue concerns the fact that in the infinite dimensional setting, we can only expect BV solutions to be of bounded variation w.r.t. the dissipation potential $\mathcal{R}$, see Definition 3.3.2. While it is well-known, see, e.g., [Fed69, 2.5.16], that functions in $\mathrm{BV}([a,b];\mathcal{Z})$ have left and right limits at every time $t \in (a,b)$, it is not a priori clear that the same is true for functions in $\mathrm{BV}([a,b];\mathcal{R})$. In Lemma C.19, we show that these functions in fact do have left and right limits w.r.t. the norm on $\mathcal{V}$, and not only w.r.t. $\mathcal{R}$ and give a definition of their jump set J_z. The second issue concerns the transition between the left and right limit at such a jump point: Instead of allowing only for $\mathcal{V}$-absolutely continuous transitions between two states z_0 and z_1, we consider for $t \in [0,T]$ the richer set of **admissible transition curves** $\mathcal{T}(t;z_0,z_1)$, consisting of all curves $\zeta \in \mathrm{AC}([0,1];\mathcal{R})$ whose restriction to the relatively open set

$$G(t,\zeta) := \{r \in [0,1] \,|\, \mathfrak{e}(t,\zeta(r)) > 0\}$$

belongs to $\mathrm{AC}_{\mathrm{loc}}(G(t,\zeta);\mathcal{V})$ and such that $\zeta(0) = z_0$ and $\zeta(1) = z_1$. For such admissible transition curves, the expression

$$\mathfrak{p}_t[\zeta,\zeta'](r) := \begin{cases} \mathcal{R}(\dot\zeta(r)) + \|\dot\zeta(r)\|_{\mathcal{V}}\mathfrak{e}(t,\zeta(r)) & \text{if } r \in G(t,\zeta) \\ \mathcal{R}[\zeta'](r) & \text{if } r \in [0,1] \setminus G(t,\zeta) \end{cases}$$

is well-defined, and we define the **Finsler dissipation cost** induced by $\mathfrak{p}$ at time t as (see [MRS16, Def. 3.5])

$$\Delta_{\mathfrak{p}}(t;z_0,z_1) := \inf\Big\{\int_0^1 \mathfrak{p}_t[\zeta,\zeta'](r)\mathrm{d}r \,|\, \zeta \in \mathcal{T}(t;z_0,z_1)\Big\}. \tag{3.3.4}$$

In analogy to (2.2.24), we now define for $z \in \mathrm{BV}([0,T];K,\mathcal{R})$

$$\begin{aligned}\mathrm{Jmp}_{\mathfrak{p}}(z;[a,b]) :=& \Delta_{\mathfrak{p}}(a;z(a),z(a_+)) + \Delta_{\mathfrak{p}}(b;z(b_-),z(b)) \\ &+ \sum_{t\in J_z\cap(a,b)} \Big(\Delta_{\mathfrak{p}}(t;z(t_-),z(t)) + \Delta_{\mathfrak{p}}(t;z(t),z(t_+))\Big)\end{aligned}$$

and

$$\mathrm{Var}_{\mathfrak{p}}(z;[a,b]) := \mathrm{Var}_{\mathcal{R}}(z;[a,b]) - \mathrm{Jmp}_{\mathcal{R}}(z;[a,b]) + \mathrm{Jmp}_{\mathfrak{p}}(z;[a,b]).$$

Note that for all z_0, $z_1 \in \mathcal{V}$ and $\zeta \in \mathcal{T}(t;z_0,z_1)$, it holds that

$$\Delta_{\mathcal{R}}(z_0,z_1) = \mathcal{R}(z_1 - z_0) \le \mathrm{Var}_{\mathcal{R}}(\zeta;[0,1]) = \int_0^1 \mathcal{R}[\zeta'](r)\mathrm{d}r \le \int_0^1 \mathfrak{p}_t[\zeta,\zeta'](r)\mathrm{d}r,$$

so that $\Delta_{\mathcal{R}}(z_0,z_1) \le \Delta_{\mathfrak{p}}(z_0,z_1)$ and therefore also $\mathrm{Var}_{\mathcal{R}}(z;[a,b]) \le \mathrm{Var}_{\mathfrak{p}}(z;[a,b])$.

We finally arrive at following definition of BV solutions, which is taken from [MRS16, Def. 3.10]:

Definition 3.3.2 (BV solutions)**.** *A curve $z \in \mathrm{BV}([0,T];\mathcal{R})$ is called a **BV solution** of the rate-independent system* (2.4.9), *if it fulfills the local stability condition*

$$\partial\mathcal{R}(0) + \mathrm{D}_z\mathcal{I}(\ell(t),z(t)) \ni 0 \textit{ for all } t \in [0,T] \setminus J_z, \tag{S$_{\mathrm{loc}}$}$$

and the energy dissipation balance

$$\mathcal{I}(\ell(t),z(t)) + \mathrm{Var}_{\mathfrak{p}}(z;[0,t]) = \mathcal{I}(\ell(0),z(0)) + \int_0^t \partial_\ell\mathcal{I}(\ell(r),z(r))\dot\ell(r)\,\mathrm{d}r \tag{3.3.5}$$

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

In analogy to in the finite-dimensional case presented in Section 2.2.4, it is shown in [MRS16, Thm. 3.11] that solutions z_ε of the viscously regularized system (2.4.10) in the sense of Definition 3.1.1 converge to a BV solution with vanishing viscosity, that is, when $\varepsilon \to 0$. What is more, in [MRS16, Prop. 4.7], the authors prove that BV solutions and $\mathfrak{p}$-parameterized BV solutions are equivalent in the following sense: Every BV solution can be parameterized in such a way that a $\mathfrak{p}$-parameterized BV solution is obtained, and conversely, $\mathfrak{p}$-parameterized BV solutions can be interpreted as graphs of BV solutions, see (3.3.6). What follows is a simplified version of [MRS16, Prop. 4.7], together with a more detailed version of the proof given therein, adjusted to our notation.

Proposition 3.3.3 (Equivalence between BV and $\mathfrak{p}$-parameterized BV solutions). *If* $(S,\hat{t},\hat{z})$ *is a* $\mathfrak{p}$*-parameterized BV solution of* (2.4.9), *then every curve* $z:[0,T]\to\mathcal{Z}$ *such that*

$$\forall\, t\in[0,T]:\quad z(t)\in\{\hat{z}(s)\,|\,\hat{t}(s)=t\} \tag{3.3.6}$$

is a BV solution of (2.4.9) *in the sense of Definition 3.3.2.*

Conversely, if $z\in \mathrm{BV}([0,T];\mathcal{R})$ *is a BV solution of* (2.4.9), *then there exists a triple* $(S,\hat{t},\hat{z})\in\mathbb{R}_+\times W^{1,\infty}([0,S])\times \mathrm{AC}([0,S];\mathcal{R})$ *satisfying* (3.3.6) *which is a* $\mathfrak{p}$*-parameterized BV solution of* (2.4.9).

Proof. Let us first assume that we are given a $\mathfrak{p}$-parameterized BV solution $(S,\hat{t},\hat{z})$ of (2.4.9), and a curve $z:[0,T]\to\mathcal{Z}$ such that (3.3.6) holds true. We now define an inverse $\hat{s}:[0,T]\to[0,S]$ of $\hat{t}$ through (3.3.6), that is, for every $t\in[0,T]$, we choose a fixed $s\in[0,S]$ such that $(t,z(t))=(\hat{t}(s),\hat{z}(s))$ and set $\hat{s}(t):=s$. Thus, we have that $z=\hat{z}\circ\hat{s}:[0,T]\to\mathcal{Z}$, and that $t\in J_z$ if and only if $t\in J_{\hat{s}}$ and $\hat{t}(s)\equiv t$ for all $s\in[\hat{s}(t_-),\hat{s}(t)]$. Here, J_z and $J_{\hat{s}}$ denote the jump sets of z and $\hat{s}$, respectively, and $\hat{s}(t_-)$ is the left limit of $\hat{s}$ in t, cf. Lemma C.19 and (2.2.16) - (2.2.17), respectively. If necessary, for $t\in J_{\hat{s}}$, we alter $\hat{s}$ and choose $\hat{s}(t)\in[\hat{s}(t_-),\hat{s}(t_+)]$. We can now verify that $z\in\mathrm{BV}([0,T];\mathcal{R})$, since we find for all $0\le t_0<t_1\le T$ that

$$\begin{aligned}
\mathrm{Var}_{\mathcal{R}}(z,[t_0,t_1]) &= \inf\{\sum_{i=0}^{N}\mathcal{R}(z(\tau_i)-z(\tau_{i-1}))\,|\,t_0=\tau_0<\dots<\tau_N=t_1\}\\
&= \inf\{\sum_{i=0}^{N}\mathcal{R}(\hat{z}(\hat{s}(\tau_i))-\hat{z}(\hat{s}(\tau_{i-1})))\,|\,t_0=\tau_0<\dots<\tau_N=t_1\}\\
&= \mathrm{Var}_{\mathcal{R}}(\hat{z},[\hat{s}(t_0),\hat{s}(t_1)])=\int_{t_0}^{t_1}\mathcal{R}[\hat{z}'](\tau)\mathrm{d}\tau,
\end{aligned}$$

where we have used Lemma C.16 in the last equation.

In order to verify $(\mathrm{S})_{\mathrm{loc}}$, note that for all $t\in[0,T]\setminus J_z$, we have the equivalence

$$0\in\partial\mathcal{R}(0)+\mathrm{D}_z\mathcal{I}(\ell(t),z(t))\Leftrightarrow \mathfrak{e}(\hat{t}(\hat{s}(t)),\hat{z}(\hat{s}(t)))=0\Leftrightarrow t\in[0,T]\setminus\hat{t}(G), \tag{3.3.7}$$

where G was defined in (3.2.11a). Now, from the complementarity condition (3.2.12b), we infer that $\dot{\hat{t}}\equiv 0$ on G, and thus, since $\hat{t}$ is absolutely continuous, the

Lebesgue-measure of $\hat{t}(G)$ is given by

$$\mathscr{L}^1(\hat{t}(G)) = \int_{\hat{t}(G)} 1\,\mathrm{d}r = \int_G \dot{\hat{t}}(r)\mathrm{d}r = 0.$$

Therefore, the local stability condition $(\mathrm{S})_{\mathrm{loc}}$ holds on $[0,T] \setminus (\hat{t}(G) \cup J_z)$, which is dense subset of $[0,T] \setminus J_z$. With the lower semicontinuity property (3.2.14) of $\mathfrak{e}$, this allows us to conclude that $(\mathrm{S})_{\mathrm{loc}}$ holds everwhere on $[0,T] \setminus J_z$. (We refer to the forthcoming Theorem 4.2.1 for a proof that we have the necessary a priori estimate for $\mathrm{D}_z\mathcal{I}(\ell(t), z(t))$.) Returning to the equivalence (3.3.7), this implies that

$$J_z = \hat{t}(G). \tag{3.3.8}$$

For the proof of the energy dissipation balance (3.3.5), note that for a given $t \in J_z$, the restriction $(\hat{z} : [\hat{s}(t_-), \hat{s}(t_+)] \to \mathcal{Z}) \in \mathcal{T}(t; z_0, z_1)$ is an admissible transition curve, which yields that

$$\Delta_{\mathfrak{p}}(t, z(t_-), z(t)) \le \int_{s(t_-)}^{s(t)} \mathfrak{p}_t[z, z'](r)\mathrm{d}r \text{ and } \Delta_{\mathfrak{p}}(t, z(t), z(t_+)) \le \int_{s(t)}^{s(t_+)} \mathfrak{p}_t[z, z'](r)\mathrm{d}r,$$

and therefore we have for all $t \in [0,T]$ that

$$\mathrm{Var}_{\mathfrak{p}}(z; [0,t]) \le \int_{\hat{s}(0)}^{\hat{s}(t)} \mathcal{R}[\hat{z}'](r) + \|\dot{\hat{z}}(r)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(r), \hat{z}(r))\mathrm{d}r.$$

Evaluating the energy dissipation balance (EDB) fulfilled by $(\hat{t}, \hat{z})$ at times $\hat{s}(0)$ and $\hat{s}(t)$ and a change of variable yield for z the energy dissipation inequality

$$\mathcal{I}(\ell(t), z(t)) + \mathrm{Var}_{\mathfrak{p}}(z; [0,t]) \le \mathcal{I}(\ell(0), z(0)) + \int_0^t \partial_\ell \mathcal{I}(\ell(r), z(r))\dot{\ell}(r)\,\mathrm{d}r \tag{3.3.9}$$

for all $t \in [0,T]$. In order to show that the estimate (3.3.9) together with $(\mathrm{S})_{\mathrm{loc}}$ is sufficient to obtain the balance (3.3.5), we need a chain rule inequality for BV solutions, which is provided by [MRS16, Thm. 3.13].

Let us now conversely assume that we are given a BV solution of (2.4.9) in the sense of Definition 3.3.2. We choose the following parameterization:

$$\text{for } t \in [0,T], \text{ set } \hat{s}(t) := t + \mathrm{Var}_{\mathfrak{p}}(z, [0,t]), \quad S := \hat{s}(T).$$

Then the jump set $J_{\hat{s}} = J_u = (t_n)_{n \in \mathbb{N}}$ is at most countable, and we may define for $I_n := (\hat{s}(t_{n-}), \hat{s}(t_{n+}))$ and $I := \bigcup I_n$

$$\hat{t} := \hat{s}^{-1} : [0,S] \setminus I \to [0,T], \quad \text{and } \hat{z} := z \circ \hat{t} : [0,S] \setminus I \to \mathcal{Z}.$$

We further extend $\hat{t}$ and $\hat{z}$ to I in the following way: For $n \in \mathbb{N}$, we denote by $r_n : \overline{I_n} \to [0,1]$ the unique affine and strictly increasing function mapping $\overline{I_n}$ onto

$[0,1]$. Now, according to [MRS16, Thm. 3.7], for $s \in I_n$, there exists a unique admissible transition

$$\left.\begin{aligned} &\zeta_n \in \mathcal{T}(t_n; z(t_{n-}), z(t_{n+})) \text{ such that} \\ &\qquad \zeta_n(r_n(\hat{s}(t_n))) = z(t_n) \text{ and} \\ &\qquad \textstyle\int_0^1 \mathfrak{p}_{t_n}[\zeta_n, \zeta_n'](r)\mathrm{d}r = \Delta_{\mathfrak{p}}(t_n; z(t_{n-}), z(t_n)) + \Delta_{\mathfrak{p}}(t_n; z(t_n), z(t_{n+})) \end{aligned}\right\} \tag{3.3.10}$$

hold true. The definition of $\hat{t}$ and $\hat{z}$ on I is then given by

$$\text{for } s \in I_n: \quad \hat{t}(s) :\equiv t_n, \quad \hat{z}(s) := \zeta_n(r_n(s)). \tag{3.3.11}$$

Obviously, (3.3.6) is fulfilled, since $z = \hat{z} \circ \hat{s}$. Just as in (3.2.5), we find that we have $\hat{t} \in W^{1,\infty}([0,S];\mathbb{R})$. Furthermore, $\hat{z} \in \mathrm{AC}([0,S];\mathcal{R})$: Let us first note that by a change of variable, $\mathrm{Var}_{\mathcal{R}}(\hat{z};[0,S]) = \mathrm{Var}_{\mathcal{R}}(z;[0,T]) < \infty$. Therefore, using the fact that $\mathrm{Var}_{\mathfrak{p}}$ is additive according to [MRS16, Rem. 3.9], it holds for all $0 \le s_1 < s_2 \le S$ that

$$\begin{aligned} \hat{t}(s_2) - \hat{t}(s_1) + \mathrm{Var}_{\mathcal{R}}(\hat{z};[s_1,s_2]) &= \hat{t}(s_2) - \hat{t}(s_1) + \mathrm{Var}_{\mathcal{R}}(z;[\hat{t}(s_1),\hat{t}(s_2)]) \\ &\le \hat{t}(s_2) - \hat{t}(s_1) + \mathrm{Var}_{\mathfrak{p}}(z;[\hat{t}(s_1),\hat{t}(s_2)]) \\ &= \hat{t}(s_2) + \mathrm{Var}_{\mathfrak{p}}(z;[0,\hat{t}(s_2)]) - \big(\hat{t}(s_1) + \mathrm{Var}_{\mathfrak{p}}(z;[0,\hat{t}(s_1)])\big) \\ &= \hat{s}(\hat{t}(s_2)) - \hat{s}(\hat{t}(s_1)) = s_2 - s_1. \end{aligned}$$

Taking into account the monotonicity of $\hat{t}$, this yields

$$\mathcal{R}(\hat{z}(s_2) - \hat{z}(s_1)) \le \mathrm{Var}_{\mathcal{R}}(\hat{z};[s_1,s_2]) \le s_2 - s_1,$$

and thus $\hat{z} \in \mathrm{AC}([0,S];\mathcal{R})$. What is more, arguing as in (3.3.7) (using again Theorem 4.2.1 for the regularity of $\mathrm{D}_z\mathcal{I}$), we find that the set $G = I$ is relatively open. The complementarity condition (3.2.12b) and the improved local regularity $\hat{z} \in \mathrm{AC}_{\mathrm{loc}}(G;\mathcal{V})$ are now a direct consequence of the definition in (3.3.11). We find the upper estimate for the energy dissipation balance (EDB) as follows: For the dissipation terms, it holds that (with ζ_n defined in (3.3.10))

$$\begin{aligned} &\int_0^S \mathcal{R}[\hat{z}'](s)\mathrm{d}s + \int_G \|\dot{z}(s)\|_{\mathcal{V}} \mathfrak{e}(\hat{t}(s),\hat{z}(s))\mathrm{d}s \\ &\quad = \mathrm{Var}_{\mathcal{R}}(\hat{z};[0,S]) + \int_G \|\dot{z}(s)\|_{\mathcal{V}} \mathfrak{e}(\hat{t}(s),\hat{z}(s))\mathrm{d}s \\ &\quad = \mathrm{Var}_{\mathcal{R}}(z;[0,T]) + \sum_{n\in\mathbb{N}} \int_0^1 \|\dot{\zeta}_n(s)\|_{\mathcal{V}} \mathfrak{e}(t_n,\zeta_n(s))\mathrm{d}s \\ &\quad = \mathrm{Var}_{\mathcal{R}}(z;[0,T]) + \sum_{n\in\mathbb{N}} \int_0^1 \mathfrak{p}_{t_n}[\zeta_n,\zeta_n'](s)\mathrm{d}s - \sum_{n\in\mathbb{N}} \int_0^1 \mathcal{R}[\zeta_n'](s)\mathrm{d}s \\ &\quad \le \mathrm{Var}_{\mathcal{R}}(z;[0,T]) + \sum_{n\in\mathbb{N}} \int_0^1 \mathfrak{p}_{t_n}[\zeta_n,\zeta_n'](s)\mathrm{d}s - \sum_{n\in\mathbb{N}} \mathcal{R}(\zeta_n(1) - \zeta_n(0)) \\ &\quad = \mathrm{Var}_{\mathcal{R}}(z;[0,T]) + \mathrm{Jmp}_{\mathfrak{p}}(z;[0,T]) - \mathrm{Jmp}_{\mathcal{R}}(z;[0,T]) \\ &\quad = \mathrm{Var}_{\mathfrak{p}}(z;[0,T]). \end{aligned}$$

Since z complies with (3.3.5), this yields

$$\int_0^S \mathcal{R}[\hat{z}'](s)\mathrm{d}s + \int_G \|\dot{z}(s)\|_{\mathcal{V}}\mathfrak{e}(\hat{t}(s),\hat{z}(s))\mathrm{d}s \\ \leq \mathcal{I}(\ell(0),z_0) - \mathcal{I}(\ell(S),\hat{z}(S)) + \int_0^S \partial_\ell \mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s))\dot{\ell}(\hat{t}(s))\dot{\hat{t}}(s)\mathrm{d}s.$$

Just as in step 4 in the proof of Theorem 3.2.6, we use the chain rule F.1 to conclude that the opposite estimate holds true as well, and obtain (EDB). ■

Remark 3.3.4. *Note that the identity* (3.3.8) *implies that, for a given* $\mathfrak{p}$*-parameterized BV solution* $(S,\hat{t},\hat{z})$*, all corresponding BV solutions* $z:[0,T]\to\mathcal{Z}$ *have jumps at exactly the same points* $t\in\hat{t}(G)$*, and only their value at the time of the jump can be chosen anywhere on the curve* $s\mapsto(\hat{t}(s),\hat{z}(s))$ *for* $s\in[\hat{s}(t_-),\hat{s}(t_+)]$*, where* $\hat{s}$ *is an inverse of* $\hat{t}$ *as described at the very beginning of the previous proof.*

In conclusion, the previous discussion shows that $\mathfrak{p}$-parameterized BV solutions are indeed parameterized versions of BV solutions, yet with higher regularity with respect to the arclength parameter s and supplying a more detailed characterization of the behaviour at the jump points.

Chapter 4

A priori estimates for $\mathfrak{p}$-parameterized BV solutions and compactness of the solution set

4.1 A priori estimates

Let us first collect the following classical a priori estimates for normalized $\mathfrak{p}$-parameterized BV solutions, which are derived from the energy-dissipation identity (EDB) and the coercivity of $\mathcal{I}$:

Lemma 4.1.1. *For all $z_0 \in \mathcal{Z}, \ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ and all normalized $\mathfrak{p}$-parameterized BV solutions $(S,t,z) \in \mathcal{L}(z_0,\ell)$, it holds with λ, c from (2.4.12) for all $s \in [0,S]$:*

$$\begin{aligned}\mathcal{I}(\ell(t(s)),z(s)) &+ \int_0^s \mathcal{R}[z'](r) + \|\dot z(r)\|_{\mathcal{V}}\mathfrak{e}(t(s),z(s))\,\mathrm{d}r \\ &\le \Big(\mathcal{I}(\ell(0),z_0) + \frac{1}{2}\|\ell\|^2_{H^1(0,T;\mathcal{V}^*)}(1+\lambda c)\Big)\Big(1 + \frac{\lambda T}{2}\exp\Big(\frac{\lambda T}{2}\Big)\Big) \qquad (4.1.1)\end{aligned}$$

Proof. The estimate for the energies is a consequence of the Gronwall inequality applied to the energy-dissipation identity (EDB). Indeed, the power integral in (EDB) can be estimated as follows (using (2.4.12)):

$$\begin{aligned}\int_0^s \partial_\ell \mathcal{I}(\ell(t(r)),z(r))\dot\ell(t(r))\dot t(r)\,\mathrm{d}r &\le \frac{1}{2}\int_0^s \|\partial_\ell\mathcal{I}(\ell(t),z)\|^2_{\mathcal{Z}^*}\dot t\,\mathrm{d}r + \frac{1}{2}\int_0^s \|\dot\ell(t)\|^2_{\mathcal{V}^*}\dot t\,\mathrm{d}r \\ &\le \int_0^s \frac{\lambda}{2}(\mathcal{I}(\ell(t),z) + c\|\ell(t)\|^2_{\mathcal{V}^*})\dot t\,\mathrm{d}r + \frac{1}{2}\|\ell\|^2_{H^1(0,T;\mathcal{V}^*)},\end{aligned}$$

and thus

$$\begin{aligned}&\mathcal{I}(\ell(t(s)),z(s)) \\ &\quad\le \mathcal{I}(\ell(t(0)),z_0) + \frac{\lambda c}{2}\int_0^s \|\ell(t(r))\|^2_{\mathcal{V}^*}\dot t(r)\,\mathrm{d}r + \int_0^s \frac{\lambda}{2}\dot t\mathcal{I}(\ell(t),z)\,\mathrm{d}r + \frac{1}{2}\|\ell\|^2_{H^1(0,T;\mathcal{V}^*)} \\ &\quad= \underbrace{\mathcal{I}(\ell(t(0)),z_0) + \frac{1}{2}\|\ell\|^2_{H^1(0,T;\mathcal{V}^*)} + \frac{\lambda c}{2}\int_0^t \|\ell(\tau)\|^2_{\mathcal{V}^*}\,\mathrm{d}\tau}_{=:\alpha(s)\le\alpha(S)\le\mathcal{I}(\ell(0),z_0)+\frac{1}{2}\|\ell\|^2_{H^1(0,T;\mathcal{V}^*)}(1+\lambda c)} + \int_0^s \underbrace{\dot t(r)\frac{\lambda}{2}}_{\beta(r)}\mathcal{I}(\ell(t(r)),z(r))\,\mathrm{d}r,\end{aligned}$$

which leads to the estimate

$$\mathcal{I}(\ell(t(s)), z(s)) \le \alpha(s)\exp\Big(\int_0^s \beta(r)\,\mathrm{d}r\Big) \le \alpha(S)\exp\Big(\frac{\lambda T}{2}\Big). \tag{4.1.2}$$

Plugging this estimate into the energy dissipation balance (EDB) again yields

$$\begin{aligned}
&\mathcal{I}(\ell(t(s)), z(s)) + \int_0^s \mathcal{R}[z'](r) + \|\dot z(r)\|_{\mathcal{V}}\mathfrak{c}(t(s), z(s))\,\mathrm{d}r \\
&\quad\le \mathcal{I}(\ell(t(0)), z_0) + \int_0^s \frac{\lambda}{2}(\mathcal{I}(\ell(t), z) + c\|\ell(t)\|_{\mathcal{V}^*}^2)\dot t\,\mathrm{d}r + \frac{1}{2}\|\ell\|_{H^1(0,T;\mathcal{V}^*)}^2 \\
&\quad\le \mathcal{I}(\ell(t(0)), z_0) + \frac{1}{2}\|\ell\|_{H^1(0,T;\mathcal{V}^*)}^2 + \int_0^s \frac{\lambda}{2}(\alpha(S)\exp\Big(\frac{\lambda T}{2}\Big)) + c\|\ell(t(r))\|_{\mathcal{V}^*}^2)\dot t(r)\,\mathrm{d}r \\
&\quad= \alpha(s) + \frac{\lambda T}{2}\alpha(S)\exp\Big(\frac{\lambda T}{2}\Big) \\
&\quad\le \Big(\mathcal{I}(\ell(0), z_0) + \frac{1}{2}\|\ell\|_{H^1(0,T;\mathcal{V}^*)}^2(1 + \lambda c)\Big)\Big(1 + \frac{\lambda T}{2}\exp\Big(\frac{\lambda T}{2}\Big)\Big).
\end{aligned}$$

■

Corollary 4.1.2. *For every $L > 0$, there exists a constant $C_L > 0$ such that for all $z_0 \in \mathcal{Z}$, $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ with $\|z_0\|_{\mathcal{Z}} + \|\ell\|_{H^1(0,T;\mathcal{V}^*)} \le L$ and all normalized $\mathfrak{p}$-parameterized BV solutions $(S,t,z) \in \mathcal{L}(z_0,\ell)$ it holds*

$$S + \|z\|_{L^\infty(0,S;\mathcal{Z})} + \mathcal{I}(\ell(t(S)), z(S)) < C_L$$

Proof. Let us first note that (4.1.1) from the preceding Lemma implies the existence of an upper bound R for the energies. Now, the coercivity estimate (D.6) in the first step of the proof of Proposition 3.1.3 yields the desired bound for $\|z\|_{L^\infty(0,S;\mathcal{Z})}$. Finally, integrating the normalization condition (N) for (S,t,z) with respect to s yields

$$S = \int_0^S \dot t(s) + \mathcal{R}[z'](r) + \|\dot z(r)\|_{\mathcal{V}}\mathfrak{c}(t(s), z(s))\,\mathrm{d}s = T + \int_0^S \mathcal{R}[z'](r) + \|\dot z(r)\|_{\mathcal{V}}\mathfrak{c}(t(s), z(s))\,\mathrm{d}s,$$

where the right hand side of the above equation is uniformly bounded according to (4.1.1). ■

4.2 Uniform estimates for the driving forces

Since the aim of the next Section 4.3 is to show compactness of the set of $\mathfrak{p}$-parameterized BV solutions, we will ultimately choose a sequence of $\mathfrak{p}$-parameterized BV solutions and show that it contains subsequence that converges to a $\mathfrak{p}$-parameterized BV solution, see Theorem 4.3.1. In order to pass to the limit in the energy dissipation balance, we require estimates on the driving forces $\mathrm{D}_z\mathcal{I}(\ell(\hat t(\cdot)), \hat z(\cdot))$ that are uniform for all $\mathfrak{p}$-parameterized BV solutions $(\hat S, \hat t, \hat z)$. From the complementarity condition (3.2.12b), we already know that we have $\mathfrak{c}(\hat t(s), \hat z(s)) = 0$ on $[0,S] \setminus G$, meaning that $\mathrm{D}_z\mathcal{I}(\ell(\hat t(s)), \hat z(s)) \in \partial\mathcal{R}(0)$ for almost all

$s \in [0,S] \setminus G$. Since $\mathcal{R}(0)$ is a bounded set in $\mathcal{V}^*$, this implies a uniform estimate in $L^\infty([0,S] \setminus G; \mathcal{V}^*)$. What is more, the inclusion $\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(\cdot)), \hat{z}(\cdot)) \in L^\infty_{\mathrm{loc}}(G, \mathcal{V}^*)$ is always satisfied by definition, but no a priori estimate is obvious here. In this section, we want to show that $\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(\cdot)), \hat{z}(\cdot)) \in L^\infty(0,S;\mathcal{V}^*)$ together with uniform estimates on sets

$$M_\rho := \{(S,\hat{t},\hat{z}); (S,\hat{t},\hat{z}) \in \mathcal{L}(z_0,\ell) \text{ for } (z_0,\ell) \text{ satisfying (2.4.6)} \\ \text{and } \|z_0\|_{\mathcal{Z}} + \|\ell\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} \le \rho\}.$$

To this end, we start with the inclusion (3.3.1) and choose a reparameterization in such a way that the transformed function $\widetilde{z}$ satisfies

$$0 \in \partial\mathcal{R}_1(\widetilde{z}'(r)) + D\mathcal{J}(\widetilde{z}'(r)) - \ell_*, \quad r > 0$$

for a constant load $\ell_* \in \mathcal{V}^*$, where $\mathcal{J} : \mathcal{Z} \to \mathbb{R}$ is defined by

$$\mathcal{J}(z) := \tfrac{1}{2}\langle Az, z\rangle + \mathcal{F}(z) = \mathcal{I}(\ell, z) + \ell, \tag{4.2.1}$$

see (3.1.30). We already derived the essential estimates on this system in Corollary 3.1.4 and it now remains to transfer them to the original one.

Theorem 4.2.1 (Bound for the driving forces)**.**
There exists a function $m : \mathcal{Z} \times W^{1,\infty}(0,T;\mathcal{V}^*) \to [0,\infty)$ mapping bounded sets to bounded sets such that for all $z_0 \in \mathcal{Z}$, $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ satisfying (2.4.6) *and all $(S,\hat{t},\hat{z}) \in \mathcal{L}(z_0,\ell)$ we have*

$$\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(\cdot)), \hat{z}(\cdot)) \in L^\infty(0,S;\mathcal{V}^*), \\ \mathrm{D}\mathcal{J}(\hat{z}(\cdot)) \in C_{weak}([0,S];\mathcal{V}^*),$$

with $\mathcal{J}$ from (4.2.1)*. Furthermore, for all measurable choices $\mu : [0,S] \to \mathcal{V}^*$ such that $\mu(s) \in \partial\mathcal{R}_2(\dot{z}(s))$ almost everwhere, and for the measurable function $\lambda : (0,S) \to [0,\infty)$ from Lemma 3.3.1, it holds*

$$\left\|\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(\cdot)), \hat{z}(\cdot))\right\|_{L^\infty(0,S;\mathcal{V}^*)} + \left\|\lambda\mu\right\|_{L^\infty(G;\mathcal{V}^*)} \le m(z_0,\ell).$$

Remark 4.2.2. *As a byproduct, in the proof of Theorem 4.2.1 we show that the function λ from* (3.3.1) *is positive almost everywhere on G, and that the function $s \mapsto 1/\lambda(s)$ belongs to $L^1_{loc}(G)$ but that it is not integrable on any connected component of G.*

Proof. As already stated at the beginning of this section, outside of the set G, we have $\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(\cdot)), \hat{z}(\cdot)) \in L^\infty((0,S)\setminus G;\mathcal{V}^*)$, together with the a priori estimate $\left\|\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(\cdot)), \hat{z}(\cdot))\right\|_{L^\infty((0,S)\setminus G;\mathcal{V}^*)} \le \mathrm{diam}_{\mathcal{V}^*}(\partial\mathcal{R}(0))$, and it remains to study the behavior on the set G. For that purpose we start from the differential inclusion (3.3.1). We recall that by the definition of parameterized solutions, the set G is a relatively open subset of $[0,S]$. Let $(a,b) \subset G$ be a maximal connected component of G. By Lemma 3.3.1, $\hat{t}$ is constant on (a,b). Hence, $\ell \circ \hat{t}$ is constant on (a,b) as well and we denote its value with ℓ_*. Now, for each compact subset

$K \subset (a,b)$ we have $\hat{z} \in W^{1,1}(K;\mathcal{V})$ and $\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(\cdot)),\hat{z}(\cdot)) \in L^\infty(K;\mathcal{V}^*)$, which implies that $\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(\cdot)),\hat{z}(\cdot)) \in C_{\mathrm{weak}}(K;\mathcal{V}^*)$. Here, we used the fact that the continuous representative of $\hat{z}$ takes values in the space $\mathcal{Z}$ according to Lemma C.14. Thus, by lower semicontinuity, there exists $c_K > 0$ such that

$$\mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\partial\mathcal{R}(0)) \geq c_K > 0 \text{ for all } s \in K.$$

The normalization condition (N) from Def. 3.2.5 now implies that $\|\hat{z}'(s)\|_{\mathcal{V}} \leq c_K^{-1}$ almost everywhere on K and hence $\lambda(s) \geq c_K^2 > 0$ almost everywhere on K, where we used the representation of λ from Lemma 3.3.1. This observation was already made in [MRS16].

The next aim is to perform a change of variables $s \mapsto r$ and $(a,b) \to (0,\Lambda)$ such that (3.3.1) rewritten in the new variable is of the form (3.1.30).

Assume first that there is $s_* \in (a,b)$ such that $1/\lambda \notin L^1((a,s_*))$. The above considerations imply that for every $\varepsilon > 0$ there exists a constant $c_\varepsilon > 0$ such that $\lambda^{-1}\big|_{(a+\varepsilon,s_*)} \leq c_\varepsilon$. Hence, since λ^{-1} is not integrable on (a,s_*), λ^{-1} is unbounded in a neighborhood of a. To be more precise, for every $n \in \mathbb{N}$ the set $\Sigma_n := \{s \in (a,a+\frac{1}{n});\, 1/\lambda(s) \geq n\}$ has positive Lebesgue measure. From the normalization property and the structure of λ we therefore deduce that

$$\text{for all } n \in \mathbb{N} \text{ and almost all } s \in \Sigma_n: \qquad \mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\partial\mathcal{R}(0)) \leq \tfrac{1}{\sqrt{n}}.$$

For $n \in \mathbb{N}$, let now $s_n \in \Sigma_n$ be points with $\lambda(s_n)^{-1} \geq n$ and such that it holds that $\mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)),\partial\mathcal{R}(0)) \leq \frac{1}{\sqrt{n}}$. Clearly, $\lim_n s_n = a$ and without loss of generality we may assume that the sequence $(s_n)_n$ is decreasing. We next study the system (3.3.1) on the intervalls (s_n,b). For $s \in (s_n,b)$, let $\Lambda_n(s) := \int_{s_n}^s \frac{1}{\lambda(\sigma)}\mathrm{d}\sigma$. The above considerations show that Λ_n is well defined on (s_n,b). Moreover, Λ_n is strictly increasing and the inverse function $\Lambda_n^{-1} : [0,\Lambda_n(b)) \to [s_n,b)$ exists. We remark that $\Lambda_n(b) = \infty$ is not excluded. For $r \in [0,\Lambda_n(b))$, let $\widetilde{z}_n(r) := \hat{z}(\Lambda_n^{-1}(r))$. Observe that $\widetilde{z}_n \in W^{1,1}(0,\Lambda_n(b-\delta);\mathcal{V})$ for every $\delta > 0$ according to [Sie20, Lemma A.2.7.], and that the function $\widetilde{z}_n$ solves the Cauchy problem

$$0 \in \partial\mathcal{R}(\dot{\widetilde{z}}_n(r)) + \partial\mathcal{R}_2(\dot{\widetilde{z}}_n(s)) + \mathrm{D}\mathcal{J}(\widetilde{z}_n(r)) - \ell_*, \quad r \in (0,\Lambda(b)), \tag{4.2.2}$$
$$\widetilde{z}_n(0) = \hat{z}(s_n). \tag{4.2.3}$$

Hence, Corollary 3.1.4 is applicable and implies in particular that it holds that $\mathrm{D}\mathcal{J}(\widetilde{z}_n(\cdot)) \in L^\infty(0,\Lambda_n(b);\mathcal{V}^*)$ and that $\widetilde{z}_n$ is the unique solution of (4.2.2) - (4.2.3).

We now apply the third part of Corollary 3.1.4 for $I = \mathbb{R}_+$, and obtain the estimate

$$\|\mathrm{D}\mathcal{J}(\widetilde{z}_n(\cdot))\|_{L^\infty(0,\infty;\mathcal{V}^*)} \leq m_2(\hat{z}(s_n),\ell_*)\Big(\mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s_n)),\hat{z}(s_n)),\partial\mathcal{R}(0)) + m_1(\hat{z}(s_n),\ell_*)\Big),$$

where $m_1,m_2 : \mathcal{Z}\times\mathcal{V}^* \to [0,\infty)$ are functions that map bounded sets on bounded sets and that do not depend on n. This immediately translates into the inclusion $\mathrm{D}\mathcal{J}(\hat{z}(\cdot)) \in L^\infty(s_n,b;\mathcal{V}^*)$ along with the estimate

$$\begin{aligned}\|\mathrm{D}\mathcal{J}(\hat{z}(\cdot))\|_{L^\infty(s_n,b;\mathcal{V}^*)} &\leq m_2(\hat{z}(s_n),\ell_*)\Big(\mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s_n)),\hat{z}(s_n)),\partial\mathcal{R}(0)) + m_1(\hat{z}(s_n),\ell_*)\Big) \qquad (4.2.4)\\ &\leq \widetilde{m}_2(z_0,\ell)\Big(\tfrac{1}{\sqrt{n}} + \widetilde{m}_1(z_0,\ell)\Big),\end{aligned}$$

where $\widetilde{m}_1, \widetilde{m}_2 : \mathcal{Z} \times W^{1,\infty}(0,T;\mathcal{V}^*) \to [0,\infty)$ are functions that map bounded sets on bounded sets and depend on $\mathcal{I},\mathcal{R}$ and embedding constants, only. The previous estimate is of the structure $\alpha_n \le \beta_n$ with an increasing sequence $(\alpha_n)_n$ and a decreasing sequence $(\beta_n)_n$. Hence, for $s_n \searrow 0$ we obtain $\mathrm{D}\mathcal{J}(\hat{z}(\cdot)) \in L^\infty(a,b;\mathcal{V}^*)$ along with a bound that ultimately depends on $\|z_0\|_{\mathcal{Z}}$ and $\|\ell\|_{W^{1,\infty}(0,T;\mathcal{V}^*)}$, only.

Assume next that $\lambda^{-1} \in L^1(a,s_*)$ for every $s_* < b$. In this case, we use the transformation $\Lambda(s) := \int_a^s \frac{1}{\lambda(\sigma)}\mathrm{d}\sigma$, and the transformed function $\widetilde{z}$ satisfies (3.1.30) on $(0,\Lambda(b))$ with the initial condition $\widetilde{z}(0) = \hat{z}(a)$. But again, since G is open, a does not belong to G and hence, $-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(a)),\hat{z}(a)) \in \partial\mathcal{R}(0)$. Arguing with Remark 3.1.5 as above, this leads to a contradiction to the normalization condition. As a consequence, λ^{-1} is not bounded close to a. The same arguments yield a contradiction for the case that $0 \notin G$ and $\lambda^{-1} \in L^1(0,s_*)$, since $0 \notin G$ is equivalent to the inclusion $-\mathrm{D}_z\mathcal{I}(\ell(0),\hat{z}(0)) \in \partial\mathcal{R}(0)$.

If $0 \notin G$, then the proof of Theorem 4.2.1 is finished. Otherwise let $[0,b)$ be a maximal connected component of G. But now we can argue exactly in the same way as before with 0 instead of s_n in (4.2.4). ∎

4.3 Compactness of solution sets

The aim of this section is to derive compactness properties of the sets

$$M_\rho := \{(S,\hat{t},\hat{z});\, (S,\hat{t},\hat{z}) \in \mathcal{L}(z_0,\ell) \text{ for } (z_0,\ell) \text{ with (2.4.6)} \\ \text{and } \|z_0\|_{\mathcal{Z}} + \|\ell\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} \le \rho\}. \tag{4.3.1}$$

for arbitrary $\rho \ge 0$. These properties will be based on the uniform estimates derived in the previous two sections.

Theorem 4.3.1 (Properties of the solution set)**.** *Let $\rho > 0$ and $z_0 \in \mathcal{Z}$. Then the set M_ρ is compact in the following sense: For every sequence $(S_n,\hat{t}_n,\hat{z}_n)_{n\in\mathbb{N}} \subseteq M_\rho$ with $(S_n,\hat{t}_n,\hat{z}_n) \in \mathcal{L}(z_0,\ell_n)$ and such that (z_0,ℓ_n) satisfy* (2.4.6)*, there exist a subsequence (denoted by the same symbols for simplicity) and limit elements $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ and $(S,\hat{t},\hat{z}) \in \mathcal{L}(z_0,\ell)$ such that (z_0,ℓ) comply with* (2.4.6) *and*

$$S_n \to S \text{ in } \mathbb{R}, \quad \hat{t}_n \overset{*}{\rightharpoonup} \hat{t} \text{ in } W^{1,\infty}(0,S), \quad \hat{t}(S) = T, \quad \ell_n \overset{*}{\rightharpoonup} \ell \text{ in } W^{1,\infty}(0,T;\mathcal{V}^*), \tag{4.3.2}$$

$$\hat{z}_n \overset{*}{\rightharpoonup} \hat{z} \text{ in } L^\infty(0,S;\mathcal{Z}) \text{ and } \hat{z}_n \to \hat{z} \text{ uniformly in } C([0,S],\mathcal{V}), \tag{4.3.3}$$

$$\hat{z}_n(S_n) \to \hat{z}(S) \text{ strongly in } \mathcal{V}, \tag{4.3.4}$$

$$\mathrm{D}\mathcal{J}(\hat{z}_n) \overset{*}{\rightharpoonup} \mathrm{D}\mathcal{J}(\hat{z}) \text{ in } L^\infty(0,S;\mathcal{V}^*), \tag{4.3.5}$$

and for every $s \in [0,S]$, it holds that

$$\hat{t}_n(s) \to \hat{t}(s), \quad \hat{z}_n(s) \rightharpoonup \hat{z}(s) \text{ in } \mathcal{Z}, \quad \mathrm{D}\mathcal{J}(\hat{z}_n(s)) \rightharpoonup \mathrm{D}\mathcal{J}(\hat{z}(s)) \text{ in } \mathcal{V}^*, \tag{4.3.6}$$

$$\hat{z}_n(s) \to \hat{z}(s) \text{ strongly in } \mathcal{Z}, \quad \mathcal{J}(\hat{z}_n(s))) \to \mathcal{J}(\hat{z}(s)) \text{ in } \mathbb{R}. \tag{4.3.7}$$

Furthermore, the map $s \mapsto \mathrm{D}\mathcal{J}(\hat{z}(s))$ is continuous w.r.t. the weak topology on $\mathcal{V}^*$.

Proof. **Step 1: Extraction of convergent subsequences** Let $(S_n, \hat{t}_n, \hat{z}_n)_{n\in\mathbb{N}} \subseteq M_\rho$ be a sequence as in the theorem and for $n \in \mathbb{N}$ let $G_n \subset [0,S]$ be the corresponding open sets according to Definition 3.2.5. Thanks to Corollary 4.1.2, we infer the first of (4.3.2). If $S > S_n$, we extend all functions $\hat{z}_n$ and $\hat{t}_n$ constantly to $[0,S]$ by their value at S_n and thus obtain the first of (4.3.3). Due to (3.2.12a) and the normalization condition (N), the second of (4.3.2) ensues, and since $W^{1,\infty}(0,S)$ is compactly embedded into $C([0,S])$, also the third of (4.3.2) as well as the first of (4.3.6). Combining the a priori estimate for $\|\hat{z}_n\|_{L^\infty(0,S;\mathcal{Z})}$ from Cor. 4.1.2 and the normalization condition (N), we conclude uniform convergence of $\hat{z}_n$ to $\hat{z}$ in $\mathcal{V}$ and pointwise weak convergence in $\mathcal{Z}$ along a subsequence by means of Proposition C.13. The same proposition also yields that $\hat{z} \in \mathrm{AC}^\infty([0,S];\mathcal{X})$. We also obtain (4.3.4) with the following estimate:

$$\|\hat{z}_n(S_n) - \hat{z}(S)\|_{\mathcal{V}} \le \|\hat{z}_n(S_n) - \hat{z}_n(S)\|_{\mathcal{V}} + \|\hat{z}_n(S) - \hat{z}(S)\|_{\mathcal{V}} \to 0,$$

where for the convergence of the first term, we exploit the equicontinuity of the sequence $(\hat{z}_n)_n$ (cf. the proof of Proposition C.13) and the second summand tends to zero due to the uniform convergence (4.3.3).

In order to show (4.3.5), we first note that thanks to the a priori estimate in Theorem 4.2.1, there are an element $\xi \in \mathcal{V}^*$ such that $\mathrm{D}\mathcal{J}(\hat{z}_n) \overset{*}{\rightharpoonup} \xi$ in $L^\infty(0,S;\mathcal{V}^*)$ as well as pointwise limits such that $\mathrm{D}\mathcal{J}(\hat{z}_n(s)) \rightharpoonup \mu(s)$ in $\mathcal{V}^*$ for all $s \in [0,S]$ along a subsubsequence. Now, since we also have $\hat{z}_n(s) \rightharpoonup \hat{z}(s)$ in $\mathcal{Z}$, and $\mathrm{D}\mathcal{F}$ is supposed to be weakly continuous (cf. (2.4.4a)), we also know that $\mathrm{D}\mathcal{J}(\hat{z}_n(s)) \rightharpoonup \mathrm{D}\mathcal{J}(\hat{z}(s))$ in $\mathcal{Z}^*$, whereby (4.3.5) and the third of (4.3.6) ensue along a subsequence. A standard argument by contradiction shows convergence along the entire sequence. By the same arguments, we obtain the weak continuity of $s \mapsto \mathrm{D}\mathcal{J}(\hat{z}(s))$.

It remains to show that $(S, \hat{t}, \hat{z}) \in \mathcal{L}(z_0, \ell)$. As a first step, we show that the complementarity identity (3.2.12b) is valid.

Step 2: Complementarity condition To this end, we want to apply Lemma B.2 and first note that $\dot{\hat{t}}_n \overset{*}{\rightharpoonup} \dot{\hat{t}}$ in $L^\infty(0,S)$ implies weak convergence also in $L^1(0,S)$. Furthermore, we have $\ell_n(\hat{t}_n(s)) \rightharpoonup \ell(\hat{t}(s))$ in $\mathcal{V}^*$ for all $s \in [0,S]$ according to Lemma E.1. Together with the weak convergence of $\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)), \hat{z}_n(s))$ according to (4.3.6) and the weak lower semicontinuity of $\mathrm{dist}_{\mathcal{V}^*}(\cdot, \partial\mathcal{R}(0))$, this implies

$$\mathfrak{c}(\hat{t}(s), \hat{z}(s)) \le \liminf_{n\to\infty} \mathfrak{c}(\hat{t}_n(s), \hat{z}_n(s)) \text{ for all } s \in [0,S]. \tag{4.3.8}$$

This allows us to conclude by means of Lemma B.2 that we have

$$0 \le \int_0^S \hat{t}'(s)\mathfrak{c}(\hat{t}(s), \hat{z}(s))\,\mathrm{d}s \le \liminf_{n\to\infty} \int_0^S \hat{t}_n'(s)\mathfrak{c}(\hat{t}_n(s), \hat{z}_n(s))\,\mathrm{d}s = 0,$$

and since the integrand is nonnegative, (3.2.12b) ensues.

Step 3: Energy dissipation balance - upper bound Next, we want to show that (EDB) is valid with $\le$ instead of $=$. For every $n \in \mathbb{N}$ and $s \in [0,S]$, it holds with the abbreviations

$$\mathrm{m}(\ell, z) := \mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell, z), \partial\mathcal{R}(0))$$

$$\hat{\mathcal{E}}_n(s, v) := \mathcal{I}(\ell_n(\hat{t}_n(s)), v) = \mathcal{J}(v) - \langle \hat{\ell}_n(s), v\rangle, \text{ where } \hat{\ell}_n := \ell_n \circ \hat{t}_n, \text{ and}$$

$$\hat{\mathcal{E}}(s, v) := \mathcal{I}(\ell(\hat{t}(s)), v) = \mathcal{J}(v) - \langle \hat{\ell}(s), v\rangle, \text{ where } \hat{\ell} := \ell \circ \hat{t}$$

that (note that $\mathrm{m}(\hat{\ell}_n(r), z) = \mathfrak{e}(\hat{t}_n(r), z)$)

$$\hat{\mathcal{E}}_n(s, \hat{z}_n(s)) + \int_0^s \mathcal{R}[\hat{z}_n'](r)\,\mathrm{d}r + \int_{[0,s]\cap G_n} \|\dot{\hat{z}}_n(r)\|_{\mathcal{V}} \mathrm{m}(\hat{\ell}_n(r), \hat{z}_n(r))\,\mathrm{d}r = \hat{\mathcal{E}}_n(0, z_0) - \int_0^s \langle \dot{\hat{\ell}}_n(r), \hat{z}_n(r)\rangle\,\mathrm{d}r. \tag{4.3.9}$$

Now, the second of (4.3.6) together with the lower semicontinuity of $v \mapsto \mathcal{J}(v)$ w.r.t. the weak topology on $\mathcal{Z}$, as well as Lemma E.1 imply for all $s \in [0, S]$ that

$$\liminf_{n\in\mathbb{N}} \hat{\mathcal{E}}_n(s, \hat{z}_n(s)) \geq \hat{\mathcal{E}}(s, \hat{z}(s)) \text{ and } \lim_{n\to\infty} \hat{\mathcal{E}}_n(0, z_0) = \hat{\mathcal{E}}(0, z_0). \tag{4.3.10}$$

For the first dissipation integral, it follows by means of Helly's selection principle, [MM05, Theorem 3.2], for all $s \in [0, S]$ that

$$\liminf_{n\to\infty} \int_0^s \mathcal{R}[\hat{z}_n'](r)\,\mathrm{d}r \geq \int_0^s \mathcal{R}[\hat{z}'](r)\,\mathrm{d}r. \tag{4.3.11}$$

According to Lemma E.1 in combination with the second of (4.3.3), the load term fulfills the convergence

$$\int_0^s \langle \dot{\hat{\ell}}(r), \hat{z}(r)\rangle\,\mathrm{d}r = \lim_{n\to\infty} \int_0^s \langle \dot{\hat{\ell}}_n(r), \hat{z}_n(r)\rangle\,\mathrm{d}r, \tag{4.3.12}$$

and it remains to study the second dissipation term. First, we show that the set $G := \{s \in [0, S] : \mathrm{m}(\hat{\ell}(s), \hat{z}(s)) > 0\}$ is a relatively open subset of $[0, S]$. To this end, let $(s_k)_{k\in\mathbb{N}} \subset [0, S] \setminus G$ be a sequence converging to an element $s \in [0, S]$. By the weak continuity of $s \mapsto \mathrm{D}\mathcal{J}(\hat{z}(s))$, we obtain

$$0 = \liminf_{n\to\infty} \mathrm{m}(\hat{\ell}(s_n), \hat{z}(s_n)) \geq \mathrm{m}(\hat{\ell}(s), \hat{z}(s)) = 0, \tag{4.3.13}$$

so that $s \in [0, S] \setminus G$ and G is indeed relatively open. Next, we are going to show the improved regularity of $\hat{z}$ on G. Let $K \subset G$ be compact. By the same arguments as above, we conclude that $c := \liminf_{s\in K} \mathrm{m}(\hat{\ell}(s), \hat{z}(s)) > 0$. Thus, for every $s \in K$, there exists $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$ we have $\mathrm{m}(\ell_n(\hat{t}_n(s)), \hat{z}_n(s)) \geq \frac{c}{2}$, and a proof by contradiction shows that N_0 can be chosen independently of $s \in K$. Therefore, the normalization condition (N) implies that $\sup_{n\geq N_0} \|\hat{z}_n'\|_{L^\infty(K;\mathcal{V})} \leq \frac{2}{c}$, whence it follows in combination with (4.3.3) that $\hat{z}_n \overset{*}{\rightharpoonup} \hat{z} \in W^{1,\infty}(K; \mathcal{V})$. Now, by means of Proposition B.3 and having in mind (4.3.13), we may conclude that

$$\liminf_{n\to\infty} \int_K \|\dot{\hat{z}}_n(r)\|_{\mathcal{V}} \mathrm{m}(\hat{\ell}_n(r), \hat{z}_n(r))\,\mathrm{d}r \geq \int_K \|\dot{\hat{z}}(r)\|_{\mathcal{V}} \mathrm{m}(\hat{\ell}(r), \hat{z}(r))\,\mathrm{d}r, \tag{4.3.14}$$

What is more, from (4.3.13), we infer that for every $s \in G$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, it holds that $s \in G_n$, so that the characteristic functions $\chi_{G_n\cap G}$ converge pointwisely to χ_G on $[0, S]$. Since for every $n \in N$, the function $\chi_{G_n\cap G}$ is dominated by the integrable function χ_G, we in fact have strong convergence in $L^1(0,S)$. We now consider for every $s \in [0,S]$

$$\begin{aligned}
&\liminf_{n\to\infty}\Big(\int_{[0,s]\cap(G_n\cap G)} \|\dot{\hat z}_n(r)\|_{\mathcal V}\mathrm{m}(\hat\ell_n(r),\hat z_n(r))\,\mathrm{d}r - \int_{[0,s]\cap G}\|\dot{\hat z}(r)\|_{\mathcal V}\mathrm{m}(\hat\ell(r),\hat z(r))\,\mathrm{d}r\Big)\\
&\quad=\liminf_{n\to\infty}\int_{[0,s]}\chi_{G_n\cap G}\cdot\|\dot{\hat z}_n(r)\|_{\mathcal V}\mathrm{m}(\hat\ell_n(r),\hat z_n(r)) - \chi_G\cdot\|\dot{\hat z}(r)\|_{\mathcal V}\mathrm{m}(\hat\ell(r),\hat z(r))\,\mathrm{d}r\\
&\quad=\liminf_{n\to\infty}\Bigg(\int_{[0,s]}\|\dot{\hat z}_n(r)\|_{\mathcal V}\mathrm{m}(\hat\ell_n(r),\hat z_n(r))\cdot\Big(\chi_{G_n\cap G}-\chi_G\Big)\mathrm{d}r \qquad (4.3.15)\\
&\qquad\qquad+\int_{[0,s]}\chi_G\cdot\Big(\|\dot{\hat z}_n(r)\|_{\mathcal V}\mathrm{m}(\hat\ell_n(r),\hat z_n(r)) - \|\dot{\hat z}(r)\|_{\mathcal V}\mathrm{m}(\hat\ell(r),\hat z(r))\Big)\mathrm{d}r\Bigg).
\end{aligned}$$

Now, for the first integral, we use the strong convergence $\chi_{G_n\cap G}\to\chi_G$ in $L^1(0,S)$, together with the uniform boundedness of $(\|\dot{\hat z}_n\|_{\mathcal V}\mathrm{m}(\hat\ell_n,\hat z_n))_{n\in\mathbb N}$ in $L^\infty(0,S)$ according to the normalization condition (N) from Def. 3.2.5, to infer that

$$\lim_{n\to\infty}\int_{[0,s]}\|\dot{\hat z}_n(r)\|_{\mathcal V}\mathrm{m}(\hat\ell_n(r),\hat z_n(r))\cdot\Big(\chi_{G_n\cap G}-\chi_G\Big)\mathrm{d}r = 0. \qquad (4.3.16)$$

If we are able to show that

$$\liminf_{n\to\infty}\int_{[0,s]}\chi_G\cdot\Big(\underbrace{\|\dot{\hat z}_n(r)\|_{\mathcal V}\mathrm{m}(\hat\ell_n(r),\hat z_n(r))}_{=:g_n(r)} - \underbrace{\|\dot{\hat z}(r)\|_{\mathcal V}\mathrm{m}(\hat\ell(r),\hat z(r))}_{=:g(r)}\Big)\mathrm{d}r\ge 0, \qquad (4.3.17)$$

then (4.3.15) implies

$$\begin{aligned}
\liminf_{n\to\infty}\int_{[0,s]\cap G_n}\|\dot{\hat z}_n(r)\|_{\mathcal V}\mathrm{m}(\hat\ell_n(r),\hat z_n(r))\,\mathrm{d}r &\ge \liminf_{n\to\infty}\int_{[0,s]\cap(G_n\cap G)}\|\dot{\hat z}_n(r)\|_{\mathcal V}\mathrm{m}(\hat\ell_n(r),\hat z_n(r))\,\mathrm{d}r\\
&\ge\int_{[0,s]\cap G}\|\dot{\hat z}(r)\|_{\mathcal V}\mathrm{m}(\hat\ell(r),\hat z(r))\,\mathrm{d}r, \qquad (4.3.18)
\end{aligned}$$

where the left-hand side of (4.3.18) is finite according to (4.1.1).

For the proof of (4.3.17), let $(K_j)_{j\in\mathbb N}$ be an exhaustion of G by compact sets. Then we find for every $j\in\mathbb N$ that

$$\begin{aligned}
\liminf_{n\to\infty}\int_{[0,S]}\chi_G(g_n-g)\,\mathrm{d}r &\ge \liminf_{n\to\infty}\int_{[0,S]}\Big(\chi_G-\chi_{K_j}\Big)g_n\,\mathrm{d}r\\
&\quad+\liminf_{n\to\infty}\int_{[0,S]}\chi_{K_j}(g_n-g)\,\mathrm{d}r+\liminf_{n\to\infty}\int_{[0,S]}\Big(\chi_G-\chi_{K_j}\Big)g\,\mathrm{d}r.
\end{aligned}$$

Now, the second summand is non-negative according to (4.3.14), whereas for the first and the last summand, we can again argue just as in (4.3.16) in order to find for arbitrary $\delta>0$ an index $J\in\mathbb N$ that is big enough, so that

$$\begin{aligned}
&\liminf_{n\to\infty}\int_{[0,S]}|\chi_G-\chi_{K_J}|\cdot|g_n|\,\mathrm{d}r\le\liminf_{n\to\infty}\int_{[0,S]}|\chi_G-\chi_{K_J}|\,\mathrm{d}r<\frac{\delta}{2},\ \text{and}\\
&\liminf_{n\to\infty}\int_{[0,S]}|\chi_G-\chi_{K_J}|\cdot|g|\,\mathrm{d}r\le\liminf_{n\to\infty}\int_{[0,S]}|\chi_G-\chi_{K_J}|\,\mathrm{d}r<\frac{\delta}{2}.
\end{aligned}$$

Thus, we find for every $\delta > 0$ that $\liminf_{n\to\infty} \int_{[0,S]} \chi_G(g_n - g)\,\mathrm{d}r \geq -\delta$, which is (4.3.17)

We are now in the position to pass to the limit inferior in (4.3.9), and using (4.3.10) - (4.3.12), as well as (4.3.18), we find that (EDB) is valid with $\leq$ instead of $=$. Similarly, we pass to the limit inferior in the normalization condition and find that $\hat{z}$ satisfies (N) with $\geq$ instead of $=$, which reads

$$\begin{aligned} 1 &= \liminf_{n\to\infty}\Big(\int_0^s \mathcal{R}[\hat{z}_n'](r)\,\mathrm{d}r + \int_{[0,s]\cap G_n} \|\dot{\hat{z}}_n(r)\|_{\mathcal{V}} \mathrm{m}(\hat{\ell}_n(r), \hat{z}_n(r))\,\mathrm{d}r\Big) \\ &\geq \int_0^s \mathcal{R}[\hat{z}'](r)\,\mathrm{d}r + \int_{[0,s]\cap G} \|\dot{\hat{z}}(r)\|_{\mathcal{V}} \mathrm{m}(\hat{\ell}(r), \hat{z}(r))\,\mathrm{d}r \end{aligned} \tag{4.3.19}$$

Step 4: Energy dissipation balance - lower bound In order to show the opposite estimates, we follow the ideas from [KZ21].

We first show that $s \mapsto \mathcal{J}(\hat{z}(s))$ is continuous on $[0,S]$ and hence uniformly continuous. From $\hat{z} \in C([0,S];\mathcal{V}) \cap L^\infty(0,S;\mathcal{Z})$ we obtain $\hat{z} \in C_{\mathrm{weak}}([0,S];\mathcal{Z})$. Hence, thanks to the assumptions (2.4.4a), $\mathcal{F}(\hat{z}(\cdot))$ is continuous on $[0,S]$ and $\mathrm{D}\mathcal{F}(\hat{z}(\cdot))$ belongs to $C_{\mathrm{weak}}([0,S];\mathcal{V}^*)$. Since the same is true for $\mathrm{D}\mathcal{J}(\hat{z}(\cdot))$, we conclude that $A\hat{z}(\cdot)$ is continuous with respect to the weak topology in $\mathcal{V}^*$, as well. But this ensures the continuity of the term $s \mapsto \langle A\hat{z}(s), \hat{z}(s)\rangle_{\mathcal{V}^*,\mathcal{V}}$ and ultimately the continuity of $\mathcal{J}(\hat{z}(\cdot))$.

For $s \in [0,S]$, let $\mu(s) \in \partial\mathcal{R}(0)$ such that $\| -\mathrm{D}_z\hat{\mathcal{E}}(s,\hat{z}(s)) - \mu(s)\|_{\mathcal{V}^*} = \mathrm{m}(\hat{\ell}(s), \hat{z}(s))$. An application of (2.4.13) yields for every $s \in [0,S)$ and $0 < h < S - s$

$$\begin{aligned} \mathcal{J}(\hat{z}(s+h)) - \mathcal{J}(\hat{z}(s)) \geq& \langle \mathrm{D}_z\hat{\mathcal{E}}(s,\hat{z}(s)), \Delta_h\hat{z}(s)\rangle + \langle \hat{\ell}(s), \Delta_h\hat{z}(s)\rangle - \lambda\mathcal{R}(\Delta_h\hat{z}(s))\|\Delta_h\hat{z}(s)\|_{\mathcal{V}} \\ =& \langle \mathrm{D}_z\hat{\mathcal{E}}(s,\hat{z}(s)) + \mu(s), \Delta_h\hat{z}(s)\rangle + \langle \hat{\ell}(s), \Delta_h\hat{z}(s)\rangle - \langle \mu(s), \Delta_h\hat{z}(s)\rangle \\ & - \lambda\mathcal{R}(\Delta_h\hat{z}(s))\|\Delta_h\hat{z}(s)\|_{\mathcal{V}}, \end{aligned}$$

where we abbreviate $\Delta_h\hat{z}(s) := \hat{z}(s+h) - \hat{z}(s)$. Now, thanks to the choice of $\mu(s)$, we can estimate the first term on the right hand side by

$$\begin{aligned} -\langle \mathrm{D}_z\hat{\mathcal{E}}(s,\hat{z}(s)) + \mu(s), \Delta_h\hat{z}(s)\rangle &\leq \|\mathrm{D}_z\hat{\mathcal{E}}(s,\hat{z}(s)) + \mu(s)\|_{\mathcal{V}^*}\|\Delta_h\hat{z}(s)\|_{\mathcal{V}} \\ &= \mathrm{m}(\hat{\ell}(s), \hat{z}(s))\|\Delta_h\hat{z}(s)\|_{\mathcal{V}}, \end{aligned}$$

and the third term by $\langle \mu(s), \Delta_h\hat{z}(s)\rangle \leq \mathcal{R}(\Delta_h\hat{z}(s))$. Therefore, rearrangement of the terms leads to the estimate

$$\begin{aligned} \mathcal{J}(\hat{z}(s+h)) - \mathcal{J}(\hat{z}(s)) + \mathrm{m}(\hat{\ell}(s), \hat{z}(s))\|\Delta_h\hat{z}(s)\|_{\mathcal{V}} + (1 + \lambda\|\Delta_h\hat{z}(s)\|_{\mathcal{V}})\mathcal{R}(\Delta_h\hat{z}(s)) \\ \geq \langle \hat{\ell}(s), \Delta_h\hat{z}(s)\rangle, \end{aligned}$$

which we divide by $h > 0$ and integrate with respect to s to obtain for every $0 \leq \sigma_1 < \sigma_2 \leq S - h$

$$\begin{aligned} \int_{\sigma_1}^{\sigma_2} \tfrac{1}{h}\Big(\mathcal{J}(\hat{z}(s+h)) - \mathcal{J}(\hat{z}(s))\Big)\mathrm{d}s + \int_{\sigma_1}^{\sigma_2} \mathrm{m}(\hat{\ell}(s), \hat{z}(s))\|\tfrac{1}{h}\Delta_h\hat{z}(s)\|_{\mathcal{V}}\,\mathrm{d}s& \\ + \int_{\sigma_1}^{\sigma_2} (1 + \lambda\|\Delta_h\hat{z}(s)\|_{\mathcal{V}})\mathcal{R}(\tfrac{1}{h}\Delta_h\hat{z}(s))\,\mathrm{d}s& \\ \geq \int_{\sigma_1}^{\sigma_2} \langle \hat{\ell}(s), \tfrac{1}{h}\Delta_h\hat{z}(s)\rangle\,\mathrm{d}s.& \end{aligned} \tag{4.3.20}$$

Now, since $s \mapsto \mathcal{J}(\hat{z}(s))$ is uniformly continuous (as shown above), the first integral converges to $\mathcal{J}(\hat{z}(\sigma_2)) - \mathcal{J}(\hat{z}(\sigma_1))$ with $h \to 0$. For the second integral, we have to distinguish the cases $s \in [0,S] \setminus G$, where we have $\mathrm{m}(\hat{\ell}(s), \hat{z}(s)) = 0$, and $s \in G$, where we can argue as follows: Since $\hat{z} \in W^{1,\infty}_{\mathrm{loc}}(G;\mathcal{V})$, we find by the Dominated Convergence Theorem for every $K \Subset G$

$$\lim_{h\to 0} \int_{(\sigma_1,\sigma_2)\cap K} \mathrm{m}(\hat{\ell}(s), \hat{z}(s)) \|\tfrac{1}{h}\Delta_h \hat{z}(s)\|_{\mathcal{V}} \,\mathrm{d}s = \int_{(\sigma_1,\sigma_2)\cap K} \mathrm{m}(\hat{\ell}(s), \hat{z}(s)) \|\hat{z}'(s)\|_{\mathcal{V}} \,\mathrm{d}s.$$

Furthermore, since we have $\hat{z} \in C([0,S];\mathcal{V})$, it follows that $\Delta_h \hat{z}(s) \to 0$ strongly in $\mathcal{V}$ and uniformly in s, and since $\hat{z} \in \mathrm{AC}^\infty([0,S];\mathcal{X})$, we infer by means of the results in Appendix C that $\mathcal{R}(\frac{1}{h}\Delta_h \hat{z}(s)) \to \mathcal{R}[\hat{z}'](s)$ for almost every $s \in [0,S]$. Keeping in mind that $\mathcal{R}[\hat{z}'](\sigma) \leq 1$ for almost all $\sigma \in [0,S]$ due to the normalization inequality (4.3.19), as well as the fact that $\mathcal{R}[z']$ is an admissible integrand in (C.4), we obtain the estimate

$$\mathcal{R}(\tfrac{1}{h}\Delta_h \hat{z}(s)) = \frac{1}{h}\mathcal{R}(\hat{z}(s+h) - \hat{z}(s)) \leq \frac{1}{h}\int_s^{s+h} \mathcal{R}[\hat{z}'](\sigma)\mathrm{d}\sigma \leq 1.$$

All in all, we find a constant $C > 0$ independent of h such that it holds for almost every $s \in [0,S]$

$$(1 + \lambda\|\Delta_h \hat{z}(s)\|_{\mathcal{V}})\mathcal{R}(\tfrac{1}{h}\Delta_h \hat{z}(s)) \leq 1 + C\|\hat{z}\|_{L^\infty(0,S;\mathcal{V})},$$

so that the Dominated Convergence Theorem implies that

$$\lim_{h\to 0} \int_{\sigma_1}^{\sigma_2} (1 + \lambda\|\Delta_h \hat{z}(s)\|_{\mathcal{V}})\mathcal{R}(\tfrac{1}{h}\Delta_h \hat{z}(s)) \,\mathrm{d}s = \int_{\sigma_1}^{\sigma_2} \mathcal{R}[\hat{z}'](s)\,\mathrm{d}s.$$

Finally, for the term on the right hand side of (4.3.20), we find

$$\begin{aligned}
&\int_{\sigma_1}^{\sigma_2} \langle \hat{\ell}(s), \tfrac{1}{h}\Delta_h \hat{z}(s)\rangle \,\mathrm{d}s \\
&\quad = \frac{1}{h}\Big(\int_{\sigma_1+h}^{\sigma_2+h} \langle \hat{\ell}(s-h) - \hat{\ell}(s), \hat{z}(s)\rangle + \langle \hat{\ell}(s), \hat{z}(s)\rangle - \langle \hat{\ell}(s-h), \hat{z}(s-h)\rangle \,\mathrm{d}s\Big) \\
&\quad = -\int_{\sigma_1+h}^{\sigma_2+h} \langle \tfrac{\hat{\ell}(s)-\hat{\ell}(s-h)}{h}, \hat{z}(s)\rangle \,\mathrm{d}s + \frac{1}{h}\int_{\sigma_2}^{\sigma_2+h} \langle \hat{\ell}(s), \hat{z}(s)\rangle \,\mathrm{d}s - \frac{1}{h}\int_{\sigma_1}^{\sigma_1+h} \langle \hat{\ell}(s), \hat{z}(s)\rangle \,\mathrm{d}s \\
&\quad = -\int_{\sigma_1}^{\sigma_2} \langle \tfrac{\hat{\ell}(s+h)-\hat{\ell}(s)}{h}, \hat{z}(s+h)\rangle \,\mathrm{d}s + \frac{1}{h}\int_{\sigma_2}^{\sigma_2+h} \langle \hat{\ell}(s), \hat{z}(s)\rangle \,\mathrm{d}s - \frac{1}{h}\int_{\sigma_1}^{\sigma_1+h} \langle \hat{\ell}(s), \hat{z}(s)\rangle \,\mathrm{d}s.
\end{aligned} \tag{4.3.21}$$

In order to show convergence, we apply (2) in Lemma E.2 to $v = \hat{\ell}$ and (3) in the same Lemma to $v = \hat{z}$ and obtain for the first term on the right hand side of (4.3.21) the convergence

$$\lim_{h\to 0} \int_{\sigma_1}^{\sigma_2} \langle \tfrac{\hat{\ell}(s+h)-\hat{\ell}(s)}{h}, \hat{z}(s+h)\rangle \,\mathrm{d}s = \lim_{h\to 0} \int_{\sigma_1}^{\sigma_2} \langle L_h \hat{\ell}(s), S_h \hat{z}(s)\rangle \,\mathrm{d}s = \int_{\sigma_1}^{\sigma_2} \langle \hat{\ell}'(s), \hat{z}(s)\rangle \,\mathrm{d}s,$$

while the second and third term converge to $\langle\hat{\ell}(\sigma_2),\hat{z}(\sigma_2)\rangle$ and $\langle\hat{\ell}(\sigma_1),\hat{z}(\sigma_1)\rangle$, respectively, for almost all σ_1,σ_2. What is more, since both $\hat{\ell}\in W^{1,\infty}(0,S;\mathcal{V}^*)$ and $\hat{z}\in C([0,S],\mathcal{V})$ are continuous, the product $s\mapsto\langle\hat{\ell}(s),\hat{z}(s)\rangle$ is uniformly continuous on $[0,S]$, and we have convergence for all $\sigma_1,\sigma_2\in[0,S]$. Alltogether, we can now pass to the limit in (4.3.20) and obtain the opposite estimate in (EDB), which is therefore valid as an identity.

Step 5: Improved convergences We can now procede to show that the estimates (4.3.11) and (4.3.14) can be improved to equalities by standard arguments. We first employ Lemma E.1 and write the energy dissipation balance in the following way,

$$\begin{aligned}
\lim_{n\to\infty}&\Big(\int_0^\sigma \mathcal{R}[\hat{z}_n'](s)\,\mathrm{d}s+\int_{(0,\sigma)\cap G_n}\mathrm{m}(\hat{\ell}_n(s),\hat{z}_n(s))\|\hat{z}_n'(s)\|_{\mathcal{V}}\,\mathrm{d}s+\hat{\mathcal{E}}_n(\sigma,\hat{z}_n(\sigma))\Big)\\
&=\hat{\mathcal{E}}(0,\hat{z}(0))+\int_0^\sigma\langle\hat{\ell}'(s),\hat{z}(s)\rangle\,\mathrm{d}s\\
&=\int_0^\sigma\mathcal{R}[\hat{z}'](s)\,\mathrm{d}s+\int_{(0,\sigma)\cap G}\mathrm{m}(\hat{\ell}(s),\hat{z}(s))\|\hat{z}'(s)\|_{\mathcal{V}}\,\mathrm{d}s+\hat{\mathcal{E}}(\sigma,\hat{z}(\sigma))
\end{aligned}$$

where we also made use of the strong convergence of $\hat{z}_n$ to $\hat{z}$ according to (4.3.3). We may now conclude that in fact

$$\lim_{n\to\infty}\int_0^\sigma\mathcal{R}[\hat{z}_n'](s)\,\mathrm{d}s=\int_0^\sigma\mathcal{R}[\hat{z}'](s)\,\mathrm{d}s$$

and

$$\lim_{n\to\infty}\int_{(0,\sigma)\cap G_n}\mathrm{m}(\hat{\ell}_n(s),\hat{z}_n(s))\|\hat{z}_n'(s)\|_{\mathcal{V}}\,\mathrm{d}s=\int_{(0,\sigma)\cap G}\mathrm{m}(\hat{\ell}(s),\hat{z}(s))\|\hat{z}'(s)\|_{\mathcal{V}}\,\mathrm{d}s$$

are valid for all $\sigma\in[0,S]$. Now, writing

$$\int_0^\sigma\mathcal{R}[\hat{z}_n'](s)\,\mathrm{d}s+\int_{(0,\sigma)\cap G}\mathrm{m}(\hat{\ell}_n(s),\hat{z}_n(s))\|\hat{z}_n'(s)\|_{\mathcal{V}}\,\mathrm{d}s=\int_0^\sigma\Big(1-\hat{t}_n'(s)\Big)\,\mathrm{d}s,$$

the above convergences yield the normalization condition (N).

Finally, we exploit the uniform subdifferentiability (2.4.13) and the fact that all $\mathrm{D}\mathcal{J}(z_n)$ map into the space $\mathcal{V}^*$ to deduce for every $s\in[0,S]$ the estimate

$$\mathcal{J}(\hat{z}(s))\geq\mathcal{J}(\hat{z}_n(s))+\tfrac{\alpha}{4}\|\hat{z}(s)-\hat{z}_n(s)\|_{\mathcal{Z}}^2+\langle\mathrm{D}\mathcal{J}(\hat{z}_n(s)),\hat{z}(s)-\hat{z}_n(s)\rangle_{\mathcal{V}}.$$

Now, the last term on the right hand side converges to zero as the pairing of a weakly and a strongly convergent sequence, cf. (4.3.3) and (4.3.6), respectively. Since $\mathcal{J}$ is weakly lower semicontinuous w.r.t. the norm on $\mathcal{Z}$, we find for every $s\in[0,S]$ that

$$\begin{aligned}
\mathcal{J}(\hat{z}(s))&\geq\liminf_{n\to\infty}\Big(\mathcal{J}(\hat{z}_n(s))+\tfrac{\alpha}{4}\|\hat{z}(s)-\hat{z}_n(s)\|_{\mathcal{Z}}^2+\langle\mathrm{D}\mathcal{J}(\hat{z}_n(s)),\hat{z}(s)-\hat{z}_n(s)\rangle_{\mathcal{V}}\Big)\\
&\geq\liminf_{n\to\infty}\Big(\mathcal{J}(\hat{z}_n(s))\Big)\geq\mathcal{J}(\hat{z}(s)),
\end{aligned}\tag{4.3.22}$$

so that all the inequalities in (4.3.22) must have been equalities , and thus (4.3.7). ■

Compactness of the solution set for the viscously regularized system

Showing the corresponding result for the viscously regularized systems (3.1.1) is much more straight-forward, since we have already established the crucial a priori estimate for the driving forces $-\mathrm{D}_z\mathcal{I}(\ell, z_\varepsilon)$ in (3.1.7h). We can therefore show that, for fixed $\varepsilon > 0$, and arbitrary $\rho > 0$, the set

$$M_\rho^\varepsilon := \{z \in H^1(0,T;\mathcal{Z});\, z_\varepsilon \in \mathcal{L}^\varepsilon(z_0,\ell) \text{ for } (z_0,\ell) \text{ with (2.4.6)} \\ \text{and } \|z_0\|_{\mathcal{Z}} + \|\ell\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} \le \rho\}. \tag{4.3.23}$$

is sequentially compact.

Proposition 4.3.2 (Properties of the solution set for the viscous problems)**.**
Let $\varepsilon > 0$ be fixed. Let $\rho > 0$ and $z_0 \in \mathcal{Z}$. Then the set M_ρ^ε is compact in the following sense: For every sequence $(z_n)_{n\in\mathbb{N}} \subseteq M_\rho^\varepsilon$ with $z_n \in \mathcal{L}^\varepsilon(z_0,\ell_n)$ and such that (z_0,ℓ_n) satisfy (2.4.6)*, there exists a subsequence (denoted by the same symbols for simplicity) and limit elements $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ and $z \in H^1(0,T;\mathcal{Z})$ such that $z \in \mathcal{L}^\varepsilon(z_0,\ell)$ and (z_0,ℓ) comply with* (2.4.6) *and*

$$\ell_n \overset{*}{\rightharpoonup} \ell \text{ in } W^{1,\infty}(0,T;\mathcal{V}^*), \quad z_n \rightharpoonup z \text{ in } H^1(0,T;\mathcal{Z}) \tag{4.3.24}$$

$$z_n \overset{*}{\rightharpoonup} z \text{ in } L^\infty(0,T;\mathcal{Z}) \text{ and } z_n \to z \text{ uniformly in } C([0,T],\mathcal{V}), \tag{4.3.25}$$

$$\mathrm{D}\mathcal{J}(z_n) \overset{*}{\rightharpoonup} \mathrm{D}\mathcal{J}(z) \text{ in } L^\infty(0,T;\mathcal{V}^*), \quad \mathrm{D}_z\mathcal{I}(\ell_n,z_n) \overset{*}{\rightharpoonup} \mathrm{D}_z\mathcal{I}(\ell,z) \text{ in } L^\infty(0,T;\mathcal{V}^*) \tag{4.3.26}$$

and for every $t \in [0,T]$, it holds that

$$z_n(t) \rightharpoonup z(t) \text{ in } \mathcal{Z}, \quad \mathrm{D}\mathcal{J}(z_n(t)) \rightharpoonup \mathrm{D}\mathcal{J}(z(t)) \text{ in } \mathcal{V}^*, \tag{4.3.27}$$

$$z_n(t) \to z(t) \text{ strongly in } \mathcal{Z}, \quad \mathcal{J}(z_n(t)) \to \mathcal{J}(z(t)) \text{ in } \mathbb{R}. \tag{4.3.28}$$

Furthermore, the map $t \mapsto \mathrm{D}\mathcal{J}(z(t))$ is continuous w.r.t. the weak topology on $\mathcal{V}^*$.

Proof. **Step 1: Extraction of convergent subsequences** Let $(z_n)_{n\in\mathbb{N}} \subset M_\rho^\varepsilon$ be a sequence such that $z_n \in \mathcal{L}^\varepsilon(z_0,\ell_n)$ and (z_0,ℓ_n) comply with (2.4.6). Since the sequence $(\ell_n)_{n\in\mathbb{N}} \subset W^{1,\infty}(0,T;\mathcal{V}^*)$ is bounded by definition of M_ρ^ε, we also find that the sequence $(z_n)_{n\in\mathbb{N}} \subset H^1(0,T;\mathcal{Z})$ is bounded according to (3.1.7e), and we infer the existence of $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ and $z \in H^1(0,T;\mathcal{Z})$ such that (4.3.24) is valid along a subsequence.

The first convergence in (4.3.25) follows from the uniform $L^\infty(0,T;\mathcal{Z})$- estimate (3.1.7b), and we also may infer the first of (4.3.27). Since $\varepsilon > 0$ is fixed, we also find that the sequence $(z_n)_{n\in\mathbb{N}}$ is bounded in $W^{1,\infty}(0,T;\mathcal{V})$ and therefore equicontinuous w.r.t. the norm on $\mathcal{V}$. Thus, by application of the Arzelá-Ascoli-Theorem [Die69, Thm. 7.5.7], we also infer the second of (4.3.25).

What is more, we use the estimate (3.1.7h) in order to obtain

$$\sup_{n\in\mathbb{N}} \|\mathrm{D}_z\mathcal{I}(\ell_n,z_n)\|_{L^\infty(0,T;\mathcal{V}^*)} \le \mathrm{diam}_{\mathcal{V}^*}(\partial\mathcal{R}(0)) + \widetilde{m}(\mathcal{I}(\ell(0),z_0), \mathrm{Var}_{\mathcal{V}^*}(\ell,[0,T]))$$

for a function $\widetilde{m}(\cdot,\cdot):[0,\infty)\times[0,\infty)\to[0,\infty)$ that maps bounded sets into bounded sets, which is why we find an element $\xi\in L^\infty(0,T;\mathcal{V}^*)$ such that $\mathrm{D}_z\mathcal{I}(\ell_n,z_n)\overset{*}{\rightharpoonup}\xi$ in $L^\infty(0,T;\mathcal{V}^*)$. Taking into account the first of (4.3.24), this implies the weak convergence $\mathrm{D}\mathcal{J}(z_n)\overset{*}{\rightharpoonup}\xi+\ell_n$ in $L^\infty(0,T;\mathcal{V}^*)$. By exactly the same arguments as in the first step of the proof of Theorem 4.3.1, we find (4.3.26) and the second of (4.3.27), as well as the continuity of $t\mapsto\mathrm{D}\mathcal{J}(z(t))$ w.r.t. the weak topology on $\mathcal{V}^*$. Finally, the convergences (4.3.28) follow by the same arguments as in (4.3.22).

Step 2: The limit is a solution of the viscous problem In order to show the the limiting elements z and ℓ thus obtained indeed fulfill the relation $z\in\mathcal{L}(z_0,\ell)$, we prove that z complies with the energy dissipation estimate (3.1.4) with lower semicontinuity arguments. In fact, for every $n\in\mathbb{N}$, the curve z_n does fulfill (3.1.4), which reads for $0\le s<t\le T$ as

$$\begin{aligned}\mathcal{I}(\ell_n(s),z_n(s))-\mathcal{I}(\ell_n(t),z_n(t))+\int_s^t\langle\dot{\ell}_n(r),z_n(r)\rangle_{\mathcal{V}}\mathrm{d}r\\ \ge\int_s^t\mathcal{R}_\varepsilon(\dot{z}_n(r))+\mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell_n(r),z_n(r)))\mathrm{d}r.\end{aligned}\tag{4.3.29}$$

Now, for the left hand side of (4.3.29), it holds that

$$\mathcal{I}(\ell_n(s),z_n(s))=\mathcal{J}(z_n(s))-\langle\ell_n(s),z_n(s)\rangle_{\mathcal{V}}\to\mathcal{J}(z(s))-\langle\ell(s),z(s)\rangle_{\mathcal{V}}=\mathcal{I}(\ell(s),z(s)),$$

where the first term converges according to (4.3.28), and the second as the product of a weakly and a strongly convergent sequence according to (4.3.24) and the second of (4.3.25). By the same argument, $\mathcal{I}(\ell_n(t),z_n(t))$ converges to $\mathcal{I}(\ell(t),z(t))$. As for the last term on the left hand side of (4.3.29), we infer from (4.3.24) that $\dot{\ell}_n\overset{*}{\rightharpoonup}\dot{\ell}$ in $L^\infty(0,T;\mathcal{V}^*)$ and from (4.3.25) that $z_n\to z$ strongly in $L^\infty(0,T;\mathcal{V})$, so that

$$\int_s^t\langle\dot{\ell}_n(r),z_n(r)\rangle_{\mathcal{V}}\mathrm{d}r\to\int_s^t\langle\dot{\ell}(r),z(r)\rangle_{\mathcal{V}}\mathrm{d}r.$$

We now turn to the right hand side of (4.3.29). Note that both $\mathcal{R}_\varepsilon:\mathcal{V}\to[0,\infty)$ and $\mathcal{R}_\varepsilon^*:\mathcal{V}^*\to[0,\infty)$ are weakly lower semicontinuous w.r.t. the norm on $\mathcal{V}$ and $\mathcal{V}^*$, respectively (cf. Appendix A for the properties of the convex dual $\mathcal{R}_\varepsilon^*$). Now, from the second of (4.3.26), we we infer that

$$\mathrm{D}_z\mathcal{I}(\ell_n(t),z_n(t))\rightharpoonup\mathrm{D}_z\mathcal{I}(\ell(t),z(t))\text{ in }\mathcal{V}^*\text{ f.a.a. }t\in[0,T].$$

All in all, passing to the limit inferior in (4.3.29), we find

$$\begin{aligned}&\mathcal{I}(\ell(s),z(s))-\mathcal{I}(\ell(t),z(t))+\int_s^t\langle\dot{\ell}(r),z(r)\rangle_{\mathcal{V}}\mathrm{d}r\\ &\qquad=\liminf_{n\to\infty}\left(\int_s^t\mathcal{R}_\varepsilon(\dot{z}_n(r))+\mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell_n(r),z_n(r)))\mathrm{d}r\right)\\ &\qquad\ge\int_s^t\liminf_{n\to\infty}\Big(\mathcal{R}_\varepsilon(\dot{z}_n(r))+\mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell_n(r),z_n(r)))\Big)\mathrm{d}r\\ &\qquad\ge\int_s^t\mathcal{R}_\varepsilon(\dot{z}(r))+\mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(r),z(r)))\mathrm{d}r,\end{aligned}$$

which is (3.1.4). ■

Chapter 5

Existence of globally optimal controls

We now turn to the optimal control problem governed by (2.4.9). Our control variable is $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ and the admissible set M_{ad} consists of all normalized, $\mathfrak{p}$-parametrized solutions of the system (2.4.9) with data z_0 and ℓ. To be more precise, we define

$$M_{ad} := \Big\{(S,\hat{t},\hat{z},\ell) \in \mathbb{R}_+ \times W^{1,\infty}(0,S) \times \mathrm{AC}(0,S;\mathcal{X}) \times W^{1,\infty}(0,T;\mathcal{V}^*) \,|\, (z_0,\ell) \text{ comply with (2.4.6), and } (S,\hat{t},\hat{z}) \in \mathcal{L}(z_0,\ell)\Big\}.$$

Then, the optimal control problem under consideration reads as follows:

$$\left.\begin{array}{ll} \min & J(S,\hat{z},\ell) := j(\hat{z}(S)) + \alpha\|\ell\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} \\ \text{s.t.} & (S,\hat{t},\hat{z},\ell) \in M_{ad}. \end{array}\right\} \tag{OCP}$$

Herein, $\alpha > 0$ is a fixed Tikhonov parameter and $j : \mathcal{V} \to \mathbb{R}$ is bounded from below and continuous, e.g. $j(z) := \|z - z_{\mathrm{des}}\|_{\mathcal{V}}$ for a desired end state $z_{\mathrm{des}} \in \mathcal{V}$.

We now have the following existence result:

Theorem 5.0.1 (Main existence result)**.** *Let $\alpha > 0$ be a fixed Tikhonov parameter, $z_0 \in \mathcal{Z}$ be chosen such that there exists $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ such that (z_0,ℓ) complies with* (2.4.6) *and let $j : \mathcal{V} \to \mathbb{R}$ be bounded from below and continuous. Then, the optimal control problem* (OCP) *has a globally optimal solution.*

Proof. Since j is prerequisited to be continuous and bounded from below, we find that $I := \inf\{J(S,\hat{z},\ell) \,|\, (S,\hat{t},\hat{z},\ell) \in M_{ad}\} > -\infty$. We choose an infimizing sequence $((S_n,\hat{t}_n,\hat{z}_n,\ell_n))_{n\in\mathbb{N}} \subset M_{ad}$, i.e.

$$I = \lim_{n\to\infty} J(S_n,\hat{z}_n,\ell_n).$$

Due to the boundedness assumption on j, we find that

$$R := \sup_{n\in\mathbb{N}} \|\ell_n\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} < \infty,$$

and $((S_n,\hat{t}_n,\hat{z}_n))_{n\in\mathbb{N}} \subset M_{\|z_0\|_{\mathcal{Z}}+R}$ with M_ρ as in (4.3.1) for $\rho > 0$. According to Theorem 4.3.1, this is a sequentially compact set. Thus, there exist limit elements $\ell_* \in W^{1,\infty}(0,T;\mathcal{V}^*)$ and $(S_*,t_*,z_*) \in \mathcal{L}(z_0,\ell_*)$ and a subsequence (not relabelled for simplicity) such that (z_0,ℓ_*) comply with (2.4.6) and we have in particular the convergences (cf. (4.3.4))

$$\ell_n \overset{*}{\rightharpoonup} \ell_* \text{ in } W^{1,\infty}(0,T;\mathcal{V}^*) \text{ and } \hat{z}_n(S_n) \to z_*(S_*) \text{ in } \mathcal{V}.$$

Therefore, $(S_*,t_*,z_*,\ell_*) \in M_{ad}$, and since j is assumed to be continuous, we infer that

$$I \le J(S_*,z_*,\ell_*) \le \liminf_{n\to\infty}\Big(j(\hat{z}_n(S_n)) + \alpha\|\ell_n\|_{W^{1,\infty}(0,T;\mathcal{V}^*)}\Big) = \lim_{n\to\infty} J(S_n,\hat{z}_n,\ell_n) = I,$$

so that (S_*,t_*,z_*,ℓ_*) is indeed a minimizer of J on the admissible set M_{ad}. ■

Remark 5.0.2. *For given data (z_0,ℓ) complying with (2.4.6) and a $\mathfrak{p}$-parameterized BV solution $(S,\hat{t},\hat{z}) \in \mathcal{L}(z_0,\ell)$, it cannot be guaranteed that the driving forces at the end time S are contained in the admissible set, i.e., that it holds that*

$$-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(S)),\hat{z}(S)) \in \partial\mathcal{R}(0). \tag{5.0.1}$$

In fact, according to Remark 3.3.4, for $s \in [0,S]$, it holds that

$$-\mathrm{D}_z\mathcal{I}(\ell(\hat{t}(s)),\hat{z}(s)) \in \partial\mathcal{R}(0) \Leftrightarrow \hat{t}(s) \in [0,T]\setminus J_z,$$

where $z : [0,T] \to \mathcal{Z}$ is any BV solution of (2.4.9) chosen such that for all $t \in [0,T]$, it holds that $z(t) \in \{\hat{z}(s)\,|\,\hat{t}(s) = t\}$. Hence, if z has a jump in $T = \hat{t}(S)$, the inclusion (5.0.1) is not satisfied. Let us note that in order to make sense of these pointwise evaluations of $-\mathrm{D}_z\mathcal{I}$ in $\hat{z}$, we have to use the fact that the $\mathcal{V}$-continuous representative of $\hat{z}$ takes values in the space $\mathcal{Z}$ according to Lemma C.14. If one is interested only in those $\mathfrak{p}$-parameterized BV solutions that comply with (5.0.1), this can be achieved by adding this requirement to the admissible set, i.e., instead of (OCP), one considers

$$\left.\begin{array}{ll} \min & J(S,\hat{z},\ell) := j(\hat{z}(S)) + \alpha\|\ell\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} \\ \text{s.t.} & (S,\hat{t},\hat{z},\ell) \in \widetilde{M}_{ad}. \end{array}\right\}, \tag{OCP$_{ad}$}$$

where

$$\widetilde{M}_{ad} := \Big\{(S,\hat{t},\hat{z},\ell) \in \mathbb{R}_+ \times W^{1,\infty}(0,S) \times \mathrm{AC}(0,S;\mathcal{X}) \times W^{1,\infty}(0,T;\mathcal{V}^*)\,| \\ (z_0,\ell) \textit{ comply with } (2.4.6),\ (S,\hat{t},\hat{z}) \in \mathcal{L}(z_0,\ell), \textit{ and } (5.0.1)\Big\}.$$

This smaller admissible set is now defined in such a way that, among all the possible controls ℓ, only those that yield a $\mathfrak{p}$-parameterized BV solution satisfying (5.0.1), are considered. Note that $\widetilde{M}_{ad}$ is non-empty only if there exists a BV solution of (2.4.9) that does not possess a jump in T.

The proof of Theorem 5.0.1 can then be adjusted in the following way: We now choose an infimizing sequence $((S_n,t_n,z_n,\ell_n))_{n\in\mathbb{N}} \subset \widetilde{M}_{ad}$ for the objective functional, *and it only remains to show that its limit (S_*, t_*, z_*, ℓ_*) is still contained in the admissible set $\widetilde{M}_{ad}$, i.e., that*

$$-\mathrm{D}_z\mathcal{I}(\ell_*(t_*(S_*)), z_*(S_*)) = -\mathrm{D}_z\mathcal{I}(\ell_*(T), z_*(S_*)) \in \partial\mathcal{R}(0). \tag{5.0.2}$$

To this end, we note that it holds for all $n \in \mathbb{N}$ that

$$-\mathrm{D}_z\mathcal{I}(\ell_n(t_n(S_n)), z_n(S_n)) = -\mathrm{D}\mathcal{J}(z_n(S_n)) + \ell_n(T) \in \partial\mathcal{R}(0), \tag{5.0.3}$$

and we already have the convergences $\mathrm{D}\mathcal{J}(z_n(S_n)) \rightharpoonup \mathrm{D}\mathcal{J}(z_*(S_*))$ in $\mathcal{V}^*$ (cf. (4.3.6)), as well as $\ell_n(T) \rightharpoonup \ell_*(T)$ weakly in $\mathcal{V}^*$ according to Lemma E.1.(iii). Since $\partial\mathcal{R}(0)$ is closed w.r.t. the weak topology on $\mathcal{V}^*$ (cf. Appendix A), this implies (5.0.2).

Application to the ferroelectric model

Returning to the rate-independent ferroelectric model that was presented in Section 2.4.3, we can now apply Theorem 5.0.1 in order to achieve a prescribed end time polarization. Remember that the model reads as follows: For a bounded domain $\Omega \subset \mathbb{R}^d$ with Lipschitz boundary, $d \in \{2,3\}$, we choose the function spaces

$$\mathcal{U} := H_0^1(\Omega, \mathbb{R}^d) \times L_D^2(\Omega, \mathbb{R}^d), \quad \mathcal{Z} := H^1(\Omega, \mathbb{R}^d), \quad \mathcal{V} := L^2(\Omega, \mathbb{R}^d), \quad \mathcal{X} := L^1(\Omega, \mathbb{R}^d),$$

where $L_D^2(\Omega, \mathbb{R}^d) := \{D \in L^2(\Omega, \mathbb{R}^d) | \forall \phi \in H_0^1(\Omega, \mathbb{R}^d) : \int_\Omega D \cdot \nabla\phi \,\mathrm{d}x = 0\}$. The energy functional $\mathcal{E} : (\mathcal{U}^* \times \mathcal{V}^*) \times \mathcal{U} \times \mathcal{Z} \to \mathbb{R}$ is then given in dependence of the external load $\ell = (\ell_{\mathcal{U}^*}, \ell_{\mathcal{V}^*}) \in \mathcal{U}^* \times \mathcal{V}^*$, a displacement field $u \in H_0^1(\Omega, \mathbb{R}^d)$, the electric displacement $D \in L_D^2(\Omega, \mathbb{R}^d)$ and the sponatenous polarization $P \in \mathcal{Z}$, where P plays the role of z in (2.4.9). The dissipation potential $\mathcal{R} : \mathcal{X} \to [0,\infty)$ is defined as

$$\mathcal{R}(v) = \gamma \|v\|_{L^1(\Omega)}$$

for a constant $\gamma > 0$. For given end time $T > 0$ and initial value $P_0 \in \mathcal{Z}$, it is now the task to find $(u, D) : [0,T] \to \mathcal{U}$ and $P : [0,T] \to \mathcal{Z}$ such that $P(0) = P_0$ and

$$\left.\begin{array}{l} 0 = \mathrm{D}_u\mathcal{E}(\ell(t), u, D, P), \quad 0 = \mathrm{D}_D\mathcal{E}(\ell(t), u, D, P) \\ 0 \in \partial\mathcal{R}(\dot{P}(t)) + \mathrm{D}_P\mathcal{E}(\ell(t), u, D, P) \end{array}\right. \quad \text{for all } t \in [0,T]. \tag{5.0.4}$$

In order to apply Theorem 5.0.1, we now reduce the problem according to the procedure described in (2.2.10) - (2.2.12), that is, for fixed $\ell_{\mathcal{U}^*} \in \mathcal{U}^*$, we define the reduced energy

$$\mathcal{J} : \mathcal{V}^* \times \mathcal{Z} \to \mathbb{R} \text{ via } \mathcal{J}(\ell_{\mathcal{V}^*}, P) := \min\Big\{\mathcal{E}(\ell_{\mathcal{V}^*}, \ell_{\mathcal{U}^*}, u, D, P) \,\Big|\, (u, D) \in \mathcal{U}\Big\},$$

Under reasonable assumptions on the energy $\mathcal{E}$ (cf. [Kne19]), we can now apply the results from Chapters 3 - 5. Thus, if we choose any data $(P_0, \ell_{\mathcal{V}^*})$ in compliance with (2.4.6), we obtain from Theorem 3.2.6 that the set $\mathcal{L}(P_0, \ell_{\mathcal{V}^*})$ of normalized, $\mathfrak{p}$-parameterized BV solutions of

$$0 \in \mathrm{D}_P\mathcal{J}(\ell_{\mathcal{V}^*}(t), P(t)) + \partial\mathcal{R}(\dot{P}(t)) \tag{5.0.5}$$

is not empty. What is more, from Theorem 4.3.1, we know that for arbitrary $\rho > 0$, the set

$$M_\rho := \{(S,\hat{t},\hat{P});\, (S,\hat{t},\hat{P}) \in \mathcal{L}(P_0,\ell_{\mathcal{V}^*}) \text{ for } (P_0,\ell_{\mathcal{V}^*}) \text{ with (2.4.6)} \\ \text{and } \|P_0\|_{\mathcal{Z}} + \|\ell_{\mathcal{V}^*}\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} \leq \rho\}$$

is sequentially compact, and thus, with Theorem 5.0.1, we infer that for a given desired end time polarization $P_{\text{des}} \in \mathcal{Z}$, the optimal control problem

$$\left.\begin{array}{ll} \min & J(S,\hat{P},\ell_{\mathcal{V}^*}) := \|\hat{P}(S) - P_{\text{des}}\|_{\mathcal{Z}} + \alpha\|\ell_{\mathcal{V}^*}\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} \\ \text{s.t.} & (S,\hat{t},\hat{P},\ell_{\mathcal{V}^*}) \in M_{ad}. \end{array}\right\},$$

where

$$M_{ad} := \Big\{(S,\hat{t},\hat{P},\ell_{\mathcal{V}^*}) \in \mathbb{R}_+ \times W^{1,\infty}(0,S) \times \mathrm{AC}(0,S;\mathcal{X}) \times W^{1,\infty}(0,T;\mathcal{V}^*) \,| \\ (P_0,\ell_{\mathcal{V}^*}) \text{ comply with (2.4.6), and } (S,\hat{t},\hat{P}) \in \mathcal{L}(P_0,\ell_{\mathcal{V}^*})\Big\},$$

admits a globally optimal solution (S_*,t_*,P_*,ℓ_*) such that $(S_*,t_*,P_*) \in \mathcal{L}(P_0,\ell_*)$.

It remains to transfer this result back to the original problem (5.0.4). To this end, for a fixed $\ell_{\mathcal{U}^*} \in C^1([0,T];\mathcal{U}^*)$, couple P_* with a curve $(\widetilde{u},\widetilde{D}) : [0,T] \to \mathcal{U}$ such that

$$(\widetilde{u},\widetilde{D})(t_*(s)) \in \operatorname{Argmin}\Big\{\mathcal{E}((\ell_*,\ell_{\mathcal{U}^*})(t_*(s)),u,D,P_*(s)) \,\Big|\, (u,D) \in \mathcal{U}\Big\}.$$

It is also possible to first translate the parameterized solution $(t_*,P_*) : [0,S_*] \to \mathcal{Z}$ back into the physical time. This can be done applying Proposition 3.3.3 in order to obtain a BV solution $P_{\text{BV}} : [0,T] \to \mathcal{Z}$ by choosing

$$\text{for every } t \in [0,T] \text{ a value } P_{\text{BV}}(t) \in \{P_*(s) \,|\, t_*(s) = t\}.$$

We then couple P_{BV} with a curve $(\widehat{u},\widehat{D}) : [0,T] \to \mathcal{U}$ such that

$$(\widehat{u},\widehat{D})(t) \in \operatorname{Argmin}\Big\{\mathcal{E}((\ell_*,\ell_{\mathcal{U}^*})(t),u,D,P_{\text{BV}}(t)) \,\Big|\, (u,D) \in \mathcal{U}\Big\}.$$

Note that the values of $(\widetilde{u},\widetilde{D})$ and of $(\widehat{u},\widehat{D})$ can be chosen in such a way that it holds

$$\text{for every } t \in [0,T] : (\widehat{u},\widehat{D},P_{\text{BV}})(t) \in \Big\{(\widetilde{u} \circ t_*, \widetilde{D} \circ t_*, P_*)(s) \,\Big|\, t_*(s) = t\Big\}.$$

The viscously regularized optimal control problem

It is also possible to show the existence of optimal controls for the viscously regularized problem, which takes the following form: For fixed $\varepsilon > 0$, we define the admissible set

$$M_{ad}^\varepsilon := \Big\{(z,\ell) \in H^1(0,T;\mathcal{Z}) \times W^{1,\infty}(0,T;\mathcal{V}^*) \,| \\ (z_0,\ell) \text{ comply with (2.4.6), and } z \in \mathcal{L}^\varepsilon(z_0,\ell)\Big\},$$

which contains all solutions of the viscously regularized system (2.4.10) that correspond to compatible data z_0 and ℓ.

In analogy to (OCP), the regularized optimal control problem reads as:

$$\left.\begin{array}{ll} \min & J(z,\ell) := j(z(T)) + \alpha\|\ell\|_{W^{1,\infty}(0,T;\mathcal{V}^*)} \\ \text{s.t.} & (z,\ell) \in M_{ad}^{\varepsilon}. \end{array}\right\}, \tag{OCP$_\varepsilon$}$$

and again, $\alpha > 0$ is a fixed Tikhonov parameter and $j : \mathcal{V} \to \mathbb{R}$ is bounded from below and continuous. (OCP$_\varepsilon$) has a globally optimal solution:

Proposition 5.0.3. *Let $\alpha > 0$ be a fixed Tikhonov parameter, $z_0 \in \mathcal{Z}$ be chosen such that there exists $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ such that (z_0,ℓ) complies with* (2.4.6) *and let the function $j : \mathcal{V} \to \mathbb{R}$ be bounded from below and continuous. Then, the optimal control problem* (OCP$_\varepsilon$) *has a globally optimal solution.*

Proof. The proof is completely analogous to that of Theorem 5.0.1. We choose an infimizing sequence $(z_n,\ell_n)_{n\in\mathbb{N}} \subset M_{ad}^{\varepsilon}$ for J. Setting

$$\rho := \|z_0\|_{\mathcal{Z}} + \sup_{n\in\mathbb{N}} \|\ell_n\|_{W^{1,\infty}(0,T;\mathcal{V}^*)},$$

we find that the sequence $(z_n,\ell_n)_{n\in\mathbb{N}}$ is contained in the set M_ρ^{ε}, which is sequentially compact according to Prop. 4.3.2. Thus, (z_n,ℓ_n) converges to a minimizer

$$(z^{\varepsilon},\ell^{\varepsilon}) \in \operatorname{argmin}\{J(z,\ell) \,|\, (z,\ell) \in M_{ad}^{\varepsilon}\}.$$

■

As for now, it remains an open question whether minimizers of the viscously regularized optimal control problems (OCP$_\varepsilon$) converge to a minimizer of (OCP) with vanishing viscosity.

Chapter 6

Conclusion and outlook

In this dissertation, we obtained an existence result for an optimal control system governed by a rate-independent system and constrained to $\mathfrak{p}$-parameterized BV solutions. To the best of the author's knowledge, together with [KT18], this is the first work dealing with this type of problem in a non-convex and infinite-dimensional setting.

The interplay of a non-convex energy functional and a non-smooth dissipation potential causes inherent non-smoothness of the rate-independent system and thus also of the corresponding solutions. In Section 2.2, we saw how different notions of solution characterize these temporal discontinuities. In particular, by means of a viscous regularization and subsequent reparameterization, we arrived at the notion of $\mathfrak{p}$-parameterized BV solutions which allows for a precise description of the behaviour at jump points as a transition between two semi-stable states along a curve following a viscous regime. We also noted how the choice of parameterization affects the resulting notion of parameterized BV solution, leading us to choose the $\mathfrak{p}$-parameterization.

We then proceeded to analyze the viscously regularized system by means of a time discretization scheme along the lines of [MRS13]. The main difference consists in a more detailed a priori analysis, yielding the crucial a priori estimate (3.1.7h) for the driving forces $\mathrm{D}_z\mathcal{I}(\ell, z_\varepsilon)$. The subsequent vanishing viscosity analysis then motivated the definition of $\mathfrak{p}$-parameterized BV solutions, leading to the existence result in Theorem 3.2.6. The arguments in this section were closely related to those in [MRS16], however, the additional assumptions in our setting allowed for simplifications in the notation as well as in some arguments, allowing for more straight-forward proofs.

The crucial ingredient for solving the optimal control problem was the sequential compactness of the set of $\mathfrak{p}$-parameterized BV solutions as it is formulated in Theorem 4.3.1. Here, the key argument relied on an equivalent, differential characterization of $\mathfrak{p}$-parameterized BV solutions. While this characterization is well-known in the literature, it was used here to obtain the critical a priori estimate for the driving forces. This was done by means of a time reparameterization on those time intervals where the driving forces are not contained in the bounded set $\partial\mathcal{R}(0)$ and the external load is constant. The reparameterized solutions could thus be interpreted as solutions of an autonomous viscously regulated rate-independent system on $\mathbb{R}_+$, for which we had derived the necassary estimates in the previous section.

There remains, of course, a number of open questions:

- In this work, we opted to tackle the optimal control problem (OCP) by showing sequential compactness of the set of $\mathfrak{p}$-parameterized BV solutions. The next natural question is whether the minimizers of the viscously regularized optimal control problems (OCP_ε) can be used to approximate a minimizer of (OCP). A crucial step for this would be to investigate whether a given $\mathfrak{p}$-parameterized BV solution of the rate-independent system can be obtained as the limit of reparameterized solutions of the viscously regularized system. This is often referred to as **reverse approximation property** and is still an open question for parameterized BV solutions. The particular challenge when it comes to rate-independent systems stems from the fact that the regularized systems allow for unique and smooth solutions, whereas the original (unregularized) system does not. It seems therefore unlikely that it is possible to obtain *every* parameterized BV solution as a limit of solutions of the regularized systems, and there are, in fact, examples that demonstrate the contrary, such as Example 2.2.13 in a 1-dimensional setting. Examples in the context of perfect plasticity can be found, e.g., in [Suq81] or [MW20, Ex. 3.10]. In the context of optimal control, however, the control ℓ can be used as an additional variable, and it seems promising to find a sequence of tuples (z_n, ℓ_n) which are feasible for the regularized system such that the corresponding values of the objective function converge to the globally optimal value in the unregularized problem. This approach was followed in [MW20] for perfect elasto-plasticity and in [KMS21] in an abstract finite-dimensional setting, and it remains an open question how it can be applied to the present abstract infinite-dimensional setting.

- For our arguments, the boundedness of the subdifferential $\partial\mathcal{R}(0)$ in $\mathcal{V}^*$ was critical on many occasions as a means to derive a priori estimates for the driving forces $\mathrm{D}_z\mathcal{I}$. However, the subdifferential is bounded only if $\mathcal{R}$ is, meaning that the results in this work cannot be applied to study **damage**. In a typical model for damage, one could choose the state space $\mathcal{Z} = H^1(\Omega)$, the viscosity space $\mathcal{V} = L^2(\Omega)$, and $\mathcal{X} = L^1(\Omega)$ for a domain $\Omega \subset \mathbb{R}^3$. In this model, $z : [0,T] \times \Omega \to [0,1]$ is the damage variable, attributing for every time $t \in [0,T]$ to every point $x \in \Omega$ a value indicating the damage present at this point. Here, $z(t,x) = 1$ means no damage, and $z(t,x) = 0$ means complete damage. In order to prevent healing, one defines a **unidirectional dissipation potential** $\mathcal{R}$ via

$$\mathcal{R}(\dot z(t)) := \begin{cases} \int_\Omega |\dot z(t,x)|\,\mathrm{d}x = \|\dot z(t)\|_{\mathcal{X}}, & \text{if } \dot z(t,x) \le 0 \text{ for all } x \in \Omega, \\ \infty, & \text{else.} \end{cases}$$

  Since $\mathcal{R}$ is positively 1-homogeneous, this yields a rate-independent system, but with unbounded dissipation. Systems of this type have been studied, e.g., in the series [KRZ13, KRZ15, KRZ19] and numerically in [Sie21]. While [KRZ13, KRZ15] give existence results for weak notions of parameterized solutions, [KRZ19] guarantees existence of a BV solution in a stronger sense and with higher regularity, though under rather resctrictive assumptions on the domain Ω. It would be interesting to study the optimal control of such damage models.

- Is this work, we consider a semi-linear and state-independent setting. To be more precise, we assumed in Section 2.4.1 that the state-derivative $D_z\mathcal{I}$ of the potential energy is semilinear, and that the dissipation potential depends solely on $\dot{z}$ and not on the state z. These are of course rather restrictive assumptions, allowing for significant simplifications. There exists some literature discussing **more general energy functionals** such as [MRS16], where existence of $\mathfrak{p}$-parameterized BV solutions is shown for an abstract, quite general, infinite-dimensional setting. Here, the energy is supposed to be lower semi-continuous, Fréchet-subdifferentiable, uniformly subdifferentiable in the sense of Lemma 2.4.1, coercive and fulfilling certain estimates for the power $\partial_t\mathcal{I}(\ell(t), z)$. However, when it comes to state-dependent dissipation potentials, the literatur becomes rather scant. In [MR07], existence of energetic solutions is shown, whereas [MRS09] introduces the notion of **parameterized metric solutions**, which is obtained by a vanishing viscosity approach. The authors of [BFM12] show existence of BV solutions for a model for non-associative elasto-plasticity. Beyond these papers, further mathematical research on problems with **state-dependent dissipation** is needed.

Appendix A

Elements of convex analysis

We collect here some basic definitions and results from convex analysis that are fundamental to the energy-dissipation framework. The proofs can be found, e.g., in [ET76, Chapter 1.4-1.5]. We begin with the definition of a function's convex conjugate.

Let S be a Banach space, S^* its dual space, and let $\langle \cdot, \cdot \rangle_S$ denote the duality pairing between S and S^*. Let further $F : S \to \overline{\mathbb{R}}$ be a function. For $\sigma \in S^*$, we define

$$F^{*,S}(\sigma) := \sup_{s \in S} \{\langle \sigma, s \rangle_S - F(s) \,|\, s \in S\},$$

and call the function $F^{*,S} : S^* \to \overline{\mathbb{R}}$ the **convex conjugate** of F. Independently of the properties of F, its convex conjugate $F^{*,S}$ is convex and lower semicontinuous. If $F : S \to [0, \infty]$ itself is a convex and lower semicontinuous function, we can identify $F^{**} := (F^*)^* = F$.

We define the **subdifferential** of F at a point $s \in S$ by

$$\partial_S F(s) := \{\sigma \in S^* \,|\, \forall t \in S : \langle \sigma, t - s \rangle_S \leq F(t) - F(s)\}.$$

The subdifferential is always a closed, convex subset of S^*. If F is convex, and finite and continuous at a point $s_0 \in S$, then $\partial_S F(s) \neq \emptyset$ for all $s \in (\mathrm{Dom}\ F)^\circ$, where $(\mathrm{Dom}\ F)^\circ$ denotes the interior of the domain of F. Consequently, $\partial_S F(s)$ is closed w.r.t. the weak topology on S^*. An element $\sigma \in \partial_S F(s)$ is called a **subgradient** of F in s. By the very definition of $F^{*,S}$, we have the **Fenchel-Young inequality**

$$\text{for all } s \in S \text{ and } \sigma \in S^* : \quad F(s) + F^{*,S}(\sigma) \geq \langle \sigma, s \rangle_S. \tag{A.1}$$

The opposite inequality holds true if and only if σ is a subgradient of F in s. To be precise, we have the **Fenchel equivalences**

$$\sigma \in \partial_S F(s) \quad \Leftrightarrow \quad s \in F^{*,S}(\sigma) \quad \Leftrightarrow \quad F(s) + F^{*,S}(\sigma) = \langle \sigma, s \rangle_S. \tag{A.2}$$

While our energy-dissipation framework takes place in a triple of Banach spaces $\mathcal{Z}$, $\mathcal{V}$, and $\mathcal{X}$ such that $\mathcal{Z} \hookrightarrow \mathcal{V} \hookrightarrow \mathcal{X}$, it is a priori not obvious how the subdifferential and the convex conjugate of the dissipation potential with respect to the $\mathcal{V}-\mathcal{V}^*$-duality relate to those with respect to the $\mathcal{Z}-\mathcal{Z}^*$-duality. The following Lemma helps establish that connection.

Lemma A.1. *Let $i : \mathcal{Z} \to \mathcal{V}$ denote the embedding from $\mathcal{Z}$ into $\mathcal{V}$ and $i^* : \mathcal{V}^* \to \mathcal{Z}^*$ its dual, i.e.*

$$\text{for all } \sigma \in \mathcal{V}^* \text{ and } z \in \mathcal{Z} : \langle i^*\sigma, z\rangle_{\mathcal{Z}} = \langle \sigma, iz\rangle_{\mathcal{V}}.$$

Let further $F : \mathcal{V} \to [0,\infty)$ be convex and lower semicontinuous. Within this lemma, whenever no space is indicated in the notation of ∂F or F^*, we consider the subdifferential or the convex conjugate of F with respect to the $\mathcal{Z} - \mathcal{Z}^*$-duality.

Further, for a linear and continuous functional $B : \mathcal{V}^* \to \mathcal{Z}^*$ and a convex and lower semicontinuous map $h : \mathcal{V}^* \to [0,\infty)$, we define a map $Bh : \mathcal{Z}^* \to [0,\infty]$ by

$$Bh(\eta) := \inf\{h(\sigma) \,|\, \sigma \in \mathcal{V}^*, B\sigma = \eta\}.$$

Here, we set $\inf\emptyset := \infty$. *Then the following identities are valid:*

(i) *For all $z \in \mathcal{Z}$: $\partial(F \circ i)(z) = i^*\partial_{\mathcal{V}} F(iz)$.*

(ii) $(F \circ i)^* = i^* F^{*,\mathcal{V}}$.

In particular, it holds:

$$\text{for all } z \in \mathcal{Z} : \partial F(z) = \partial_{\mathcal{V}} F(z) \subset \mathcal{V}^* \subset \mathcal{Z}^*$$

$$\text{for all } \eta \in \mathcal{Z}^* : F^*(\eta) = \begin{cases} F^{*,\mathcal{V}}(\eta), & \text{if } \eta \in \mathcal{V}^* \\ \infty, & \text{if } \eta \in \mathcal{Z}^* \setminus \mathcal{V}^*. \end{cases}$$

Proof. For the proof of (i), we cite [ET76, Chapter I, Prop.5.7], which asserts the following: Let $A : \mathcal{Z} \to \mathcal{V}$ be a linear and continuous operator for which there exists $z_0 \in \mathcal{Z}$ such that F is continuous in Az_0. Then it holds for all elements $z \in \mathcal{Z}$ that $\partial(F \circ A)(z) = A^*\partial_{\mathcal{V}}(Az)$, where the subdifferential on the left hand side is determined with respect to the $\mathcal{Z} - \mathcal{Z}^*$-duality.

For the proof of (ii), we follow closely the arguments in the proof of [HUL01, Thm.2.2.1] and show first that for B and h given as in the lemma, we have $(Bh)^{*,\mathcal{Z}^*} = h^{*,\mathcal{V}^*} \circ B^*$. To this end, choose an arbitrary element $z \in (\mathcal{Z}^*)^* = \mathcal{Z}$ and determine

$$\begin{aligned}
(Bh)^{*,\mathcal{Z}^*}(z) &= \sup\{\langle \eta, z\rangle_{\mathcal{Z}} - \inf\{h(\sigma) \,|\, \sigma \in \mathcal{V}^*, B\sigma = \eta\} | \eta \in \mathcal{Z}^*\} \\
&= \sup\{\langle \eta, z\rangle_{\mathcal{Z}} - h(\sigma) \,|\, \eta \in \mathcal{Z}^*, \sigma \in \mathcal{V}^*, B\sigma = \eta\} \\
&= \sup\{\langle B\sigma, z\rangle_{\mathcal{Z}} - h(\sigma) \,|\, \sigma \in \mathcal{V}^*\} \\
&= \sup\{\langle \sigma, B^*z\rangle_{\mathcal{V}} - h(\sigma) \,|\, \sigma \in \mathcal{V}^*\} \\
&= h^{*,\mathcal{V}^*}(B^*(z)).
\end{aligned}$$

Thus, for $h = F^{*,\mathcal{V}}$ and $B = i^*$, we obtain that $F \circ i = (i^*F^{*,\mathcal{V}})^{*,\mathcal{Z}^*}$ and consequently $(F \circ i)^* = ((i^*F^{*,\mathcal{V}})^{*,\mathcal{Z}^*})^*$. Now, we have for all $\sigma \in \mathcal{V}^*$ that $(i^*F^{*,\mathcal{V}})(\sigma) = F^{*,\mathcal{V}}(\sigma)$ and for all $\sigma \in \mathcal{Z}^* \setminus \mathcal{V}^*$ that $(i^*F^{*,\mathcal{V}})(\sigma) = \infty$ and thus $(F \circ i)^* = ((i^*F^{*,\mathcal{V}})^{*,\mathcal{Z}^*})^* = i^*F^{*,\mathcal{V}}$. ∎

Appendix B

Lower semicontinuity results

Lemma B.1 (Vitali's convergence theorem). *Let $(z_n)_{n\in\mathbb{N}} \subset L^1(0,S;\mathcal{V})$ be a sequence and $z:[0,T]\to\mathcal{V}$ be measurable such that*

(i) $z_n(s) \to z(s)$ *for almost all* $s \in [0,S]$, *strongly in* $\mathcal{V}$,

(ii) $(z_n)_{n\in\mathbb{N}}$ *is uniformly integrable, i.e.,*

  (a) $\forall\, \varepsilon > 0\, \exists\, K \subset [0,S]: \sup_{n\in\mathbb{N}} \int_{[0,S]\setminus K} \|z_n(s)\|_{\mathcal{V}}\, \mathrm{d}s < \varepsilon$, *and*

  (b) $\forall\, \varepsilon > 0\, \exists\, \delta > 0\, \forall A \subset [0,S]: \lambda(A) \le \delta \Rightarrow \sup_{n\in\mathbb{N}} \int_A \|z_n(s)\|_{\mathcal{V}}\, \mathrm{d}s \le \varepsilon.$

Then $z_n \to z$ strongly in $L^1(0,S;\mathcal{V})$. (Here, λ denotes the one-dimensional Lebesgue-measure.)

Proof. The proof is completely analoguous to the proof of the theorem for scalar-valued functions, see e.g. [Els05, Satz 5.6]. ∎

A proof of the following Lemma can be found in [MRS12b, Lemma 4.3]

Lemma B.2. *Let I be a bounded real interval, and let $h_n, h,\ m_n, m : I \to [0,\infty)$ for $n \in \mathbb{N}$ be measurable functions satisfying*

$$h(s) \le \liminf_{n\to\infty} h_n(s) \text{ for a.a. } s \in I, \quad m_n \rightharpoonup m \in L^1(I).$$

Then

$$\int_I h(s)m(s)ds \le \liminf_{n\to\infty} \int_I h_n(s)m_n(s)ds.$$

For a proof of the following Proposition we refer to [Kne19, Lemma B.1].

Proposition B.3. *Let $v_n, v \in L^\infty(0,S;\mathcal{V})$ with $v_n \overset{*}{\rightharpoonup} v$ in $L^\infty(0,S;\mathcal{V})$ and let further $\delta_n, \delta \in L^1(0,S;[0,\infty))$ with $\liminf_{n\to\infty} \delta_n(s) \ge \delta(s)$ for almost all s. Then*

$$\liminf_{n\to\infty} \int_0^S \|v_n(s)\|_{\mathcal{V}} \delta_n(s)\, \mathrm{d}s \ge \int_0^S \|v(s)\|_{\mathcal{V}} \delta(s)\, \mathrm{d}s. \tag{B.1}$$

Appendix C

Absolutely continuous and BV-functions

In this section, we collect several concepts in the context of absolute continuity and the total variation of functions. This is motivated by the fact that in an infinite-dimensional setting, we can only expect $\mathfrak{p}$-parameterized BV solutions to be absolutely continuous w.r.t. the norm on $\mathcal{X}$. Since $\mathcal{X}$ is a non-reflexive Banach space in general, absolute continuity w.r.t. $\|\cdot\|_{\mathcal{X}}$ is not sufficient to obtain differentiability. In this appendix, we aim to achieve an understanding of how the properties of the ambient space influence the differentiability of absolutely continuous functions. We shall see that it is possible to prove that absolutely continuous functions with values in non-reflexive Banach spaces fulfill a generalized notion of differentiability, see Def. C.7. This lays the foundation for the definition of the **generalized metric derivative** (Prop. C.10), which is crucial for identifying the limiting energy dissipation balance in the vanishing viscosity analysis (cf. Section 3.2). We will also obtain an insight into the connections between absolutely continuous functions and functions of bounded variation.

Absolutely continuous functions

For a real-valued function $f : [0,T] \to \mathbb{R}$, absolute continuity is most frequently defined as follows: For every $\varepsilon > 0$, there exists $\delta > 0$ such that for every sequence of pairwise disjoint intervals $(s_k, t_k) \subset [0,T]$ with $\sum_k t_k - s_k < \delta$ it holds

$$\sum_k |f(t_k) - f(s_k)| < \varepsilon.$$

This notion can easily be generalized for the case that f maps into a metric space, cf. Definition C.1. For real-valued functions, this criterion is equivalent to f being almost everywhere differentiable with a summable derivative $\dot{f}$ such that the fundamental theorem of calculus is valid, i.e.

$$\text{for all } 0 \le t \le T : f(t) = f(0) + \int_0^t \dot{f}(s)\,\mathrm{d}s, \tag{C.1}$$

see, e.g. [Rud99]. The identity (C.1) can be weakened into requiring the purely metric condition

$$\exists m \in L^1(0,T) : \forall 0 \le s \le t \le T : |f(t) - f(s)| \le \int_s^t m(r)\,\mathrm{d}r. \tag{C.2}$$

Our first objective is to show that absolute continuity and the validity of (C.1) or (C.2) are mutually equivalent in the case that f maps into a reflexive Banach space, cf. Corollary C.5.

Definition C.1 (Absolutely continuous functions)**.** *For a subset K of a Banach space B and a subinterval $[a,b] \subseteq [0,T]$, we say that a curve $f : [a,b] \to K$ is absolutely continuous if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every sequence of pairwise disjoint intervals $(s_k, t_k) \subset [a,b]$ with $\sum_k t_k - s_k < \delta$ it holds that*

$$\sum_k \|f(t_k) - f(s_k)\|_B < \varepsilon.$$

We denote by $\mathrm{AC}([a,b],K;B)$ the space of all absolutely continuous curves $[a,b] \to K$ and by $\mathrm{AC}([a,b];B)$ the space of all absolutely continuous curves $[a,b] \to B$.

The following Lemma states that antiderivatives of Bochner integrable functions are examples for absolutely continuous functions.

Lemma C.2. *[HvVW16, Lemma 2.5.8][Emm04, Satz 7.1.19] Let B be a Banach space (not necessarily reflexive) and $f \in L^1(0,T;B)$. The function $g : [0,T] \to B$ defined by*

$$g(t) := \int_0^t f(s)\,\mathrm{d}s$$

is absolutely continuous and both the weak derivative ∂g and the classical derivative g' exist, the latter almost everywhere. Further, we have that

$$\partial g = g' = f \in L^1(0,T;B).$$

In fact, the following proposition asserts that for reflexive Banach spaces, antiderivatives of Bochner integrable functions are the only examples for absolutely continuous functions. To be more precise, the above definition of absolute continuity is sufficient to conclude the validity of the fundamental theorem of calculus, which in turn implies the validity of an estimate like in (C.2).

Proposition C.3. *[Emm04, Satz 7.1.20] If B is a reflexive Banach space and the map $f : [0,T] \to B$ is absolutely continuous, then the classical derivative $\dot f(t)$ exists for almost every $t \in [0,T]$ and $\dot f$ is Bochner integrable and for any fixed $t_0 \in [0,T]$, it holds that*

$$f(t) = f(t_0) + \int_{t_0}^t \dot f(s)\,\mathrm{d}s.$$

Next, we want to show that the analogue to (C.2) in Banach spaces implies absolute continuity.

Lemma C.4. *Let B be a (not necessarily reflexive) Banach space and $f : [0,T] \to B$ a curve such that*

$$\exists m \in L^1(0,T) : \forall 0 \le s \le t \le T : \|f(t) - f(s)\|_B \le \int_s^t m(r)\,\mathrm{d}r. \tag{C.3}$$

Then f is absolutely continuous.

Proof. Let f and m be as in (C.3). Set

$$M(t) := \int_0^t m(r)\,\mathrm{d}r \text{ for } t \in [0,T].$$

Then Lemma C.2 allows us to conclude that the real function $M : [0,T] \to \mathbb{R}$ is absolutely continuous. Furthermore, for every $0 \le s < t \le T$ it holds

$$0 \le \int_s^t m(r)\,\mathrm{d}r = \int_0^t m(r)\,\mathrm{d}r - \int_0^s m(r)\,\mathrm{d}r = M(t) - M(s).$$

For arbitrary $\varepsilon > 0$, choose $\delta > 0$ such that for every sequence of pairwise disjoint intervals $(s_k, t_k) \subset [0,T]$ with $\sum_k t_k - s_k < \delta$ it holds

$$\sum_k M(t_k) - M(s_k) = \sum_k |M(t_k) - M(s_k)| < \varepsilon.$$

We infer

$$\sum_k \|f(t_k) - f(s_k)\|_B \le \sum_k \int_{s_k}^{t_k} m(r)\,\mathrm{d}r = \sum_k M(t_k) - M(s_k) < \varepsilon.$$

■

This implies the following corollaries:

Corollary C.5. *Let B be a reflexive Banach space, then for a curve $f : [0,T] \to B$, the following are equivalent:*

(i) *f is absolutely continuous.*

(ii) $\exists m \in L^1(0,T) : \forall 0 \le s \le t \le T : \|f(t) - f(s)\|_B \le \int_s^t m(r)\,\mathrm{d}r$.

(iii) *f is differentiable almost everywhere, its derivative $\dot f$ is Bochner integrable, and for any fixed $t_0 \in [0,T]$, it holds that*

$$f(t) = f(t_0) + \int_{t_0}^t \dot f(s)\,\mathrm{d}s.$$

Proof. The implication $(i) \Rightarrow (ii)$ follows easily from Proposition C.3, and the implication $(i) \Rightarrow (iii)$ was asserted just there. $(ii) \Rightarrow (i)$ was the content of Lemma C.4, and $(iii) \Rightarrow (i)$ that of Lemma C.2. ■

Corollary C.6. *Let B be a reflexive Banach space. Then*

$$W^{1,1}(0,T;B) = \mathrm{AC}([0,T];B).$$

If B is not reflexive, the inclusion $W^{1,1}(0,T;B) \subseteq \mathrm{AC}([0,T];B)$ still holds true.

Proof. Let $f \in W^{1,1}(0,T;B)$, i.e., the weak derivative ∂f exists in $L^1(0,T;B)$. Therefore, Lemma C.2 implies that the antiderivative $g : [0,T] \to B$, which is defined by $g(t) := f(0) + \int_0^t \partial f(s)\,\mathrm{d}s$ is absolutely continuous and weakly differentiable with $\partial g = \partial f$ in $L^1(0,T;B)$. Thus, we have that $\partial(g-f) = 0$ and we conclude by means of [HvVW16, Prop. 2.5.3] that $f-g$ is a constant almost everywhere. Comparison of the values in $t=0$ yields that $f=g$ is absolutely continuous.

Now, let B be reflexive and $f \in \mathrm{AC}([0,T];B)$. Point (iii) in Corollary C.5 tells us that the pointwise derivative $\dot{f}$ exists almost everywhere and defines a curve in $L^1(0,T;B)$ such that $f(t) = f(0) + \int_0^t \dot{f}(s)\,\mathrm{d}s$ for almost all $t \in [0,T]$. Now, according to Lemma C.2, f is weakly differentiable with weak derivative $\partial f = \dot{f} \in L^1(0,T;B)$, and the estimate

$$\|f(t)\|_B \leq \|f(0)\|_B + \int_0^T \|\dot{f}(s)\|_B\,\mathrm{d}s = \|f(0)\|_B + \|\dot{f}\|_{L^1(0,T;B)}$$

implies that $f \in L^\infty(0,T;B) \subset L^1(0,T;B)$, and thus $f \in W^{1,1}(0,T;B)$. ∎

In the case that B is not reflexive, absolute continuity does no longer imply differentiability, i.e., we cannot expect that for any sequence $(s_n)_{n\in\mathbb{N}} \subset (0,T)$ converging to t, the sequence of difference quotients $(\frac{f(s_n)-f(t)}{s_n-t})_{n\in\mathbb{N}} \subset B$ converges in B. However, it is possible to show that the sequence $(\frac{\|f(s_n)-f(t)\|_B}{s_n-t})_{n\in\mathbb{N}} \subset \mathbb{R}$ of real numbers is convergent, yielding the following generalized notion of differentiability.

Proposition C.7 (Metric derivatives)**.** *[AGS05, Theorem 1.1.2] Let B be a Banach space (not necessarily reflexive). If f fulfills* (C.3)*, the limit*

$$\|f'\|_B(t) := \lim_{s\to t} \frac{\|f(s)-f(t)\|_B}{|s-t|}$$

exists almost everywhere. Moreover, the function $t \mapsto \|f'\|_B(t)$ belongs to $L^1(0,T)$, it is an admissible integrand in (C.3) *and is minimal with this property, i.e., if m is another function satisfying* (C.3)*, then $\|f'\|_B(t) \leq m(t)$ almost everywhere. $\|f'\|_B$ is called the **metric derivative** of f.*

In Lemma C.4, we have already seen that (C.3) implies absolute continuity, even if B is not reflexive. In order to show that the converse is true as well, we argue similary to the proof of [AGS05, Theorem 1.1.2] cited above.

Lemma C.8. *Let B be a Banach space (not necessarily reflexive). If $f : [0,T] \to B$ is absolutely continuous, then* (C.3) *is valid. Thus,* (i) *and* (ii) *of Corollary C.5 are equivalent, even if B is not reflexive.*

Proof. Since $[0,T]$ is separable and f is continuous, there exists a sequence $(b_n)_{n\in\mathbb{N}}$ which is dense in the image $f([0,T])$. For $n \in \mathbb{N}$ and $t \in [0,T]$, we set

$$d_n(t) := \|b_n - f(t)\|_B.$$

Since for all $0 \le s < t \le T$ it holds $|d_n(t) - d_n(s)| \le \|f(t) - f(s)\|_B$, the real functions $d_n : [0,T] \to \mathbb{R}$ are absolutely continuous, hence almost everywhere differentiable with derivatives $\dot{d}_n \in L^1(0,T)$ according to Proposition C.3. What is more, the derivatives fulfill the estimate

$$|\dot{d}_n(t)| = \lim_{s\to t} \frac{|d_n(s) - d_n(t)|}{|s-t|} \le \lim_{s\to t} \frac{\|f(s) - f(t)\|_B}{|s-t|} = \|f'\|_B(t) \text{for a.a. } t \in [0,T]$$

Thus, the sequence $g_n : [0,T] \to \mathbb{R}$ defined by

$$g_n(t) := \sup_{1\le k\le n} |\dot{d}_k(t)|$$

is bounded by the integrable function $\|f'\|_B$ and monotonely convergent to

$$d(t) := \sup_{n\in\mathbb{N}} |\dot{d}_n(t)| = \lim_{n\to\infty} g_n(t).$$

Therefore, the Monotone Convergence Theorem assures that $d \in L^1((0,T))$. We conclude for every $0 \le s < t \le T$

$$\|f(t) - f(s)\|_B = \sup_{n\in\mathbb{N}} |d_n(t) - d_n(s)| \le \int_s^t d(r)\,\mathrm{d}r.$$

■

For examples of absolutely continuous curves into not reflexive Banach spaces that are not differentiable, see e.g. [RMS08, Section 7] and [Emm04, Beispiel 7.1.21]. Indeed, [HvVW16, Thm. 2.5.12] asserts that a.e. differentiability of locally absolutely continuous functions with values in a Banach space B is equivalent to said Banach space having the Radon-Nikodỳm property. Reflexive spaces always have the Radon-Nikodỳm property, see [HvVW16, Thm. 1.3.21].

It is now possible to generalize the notion of metric derivatives to the class of $\mathcal{R}$-absolutely continuous curves, where $\mathcal{R}$ is a dissipation potential and not necessarily a norm. The following definition was already given in Section 3.2, but is repeated here for convenience.

Definition C.9. *[$\mathcal{R}$-absolutely continuous functions] Let $\mathcal{V}$ be a Banach space, and let $\mathcal{R} : \mathcal{V} \to [0,\infty)$ be convex, lower semicontinuous and positively 1-homogeneous. For a subset $K \subseteq \mathcal{V}$ and a subinterval $[a,b] \subseteq [0,T]$, we say that a curve $v : [a,b] \to K$ is $\mathcal{R}$-absolutely continuous, if there exists a non-negative function $m \in L^1(a,b)$ such that*

$$\mathcal{R}(v(t) - v(s)) \le \int_s^t m(r)\,\mathrm{d}r \quad \textit{for every } a \le s < t \le b, \tag{C.4}$$

and denote by $\mathrm{AC}([a,b];K,\mathcal{R})$ the set of all $\mathcal{R}$-absolutely continuous curves $[a,b] \to K$, and by $\mathrm{AC}([a,b];\mathcal{R})$ the set of all $\mathcal{R}$-absolutely continuous curves $[a,b] \to \mathcal{V}$.

Even if $\mathcal{R}$ does not induce a norm on $\mathcal{V}$, $\mathcal{R}$-absolutely continuous curves still have a metric derivative:

Proposition C.10. *[Generalized metric derivatives][RMS08, Prop.2.2] Let $\mathcal{R}$ be given as in Def. C.9. If $v \in \mathrm{AC}([a,b];K,\mathcal{R})$, the limit*

$$\mathcal{R}[v'](t) := \lim_{h\searrow 0} \mathcal{R}\left(\frac{v(t+h)-v(t)}{h}\right) = \lim_{h\searrow 0} \mathcal{R}\left(\frac{v(t)-v(t-h)}{h}\right)$$

exists almost everywhere. Moreover, the function $t \mapsto \mathcal{R}[v'](t)$ belongs to $L^1(0,T)$, it is an admissible integrand in (C.4) *and is minimal with this property, i.e., if m is another function satisfying* (C.4)*, then $\mathcal{R}[v'](t) \leq m(t)$ almost everywhere.*

Remark C.11. *In the setting of the energy-dissipation framework as described in Section 2.4.1, let $v \in \mathrm{AC}([a,b];\mathcal{V})$. Since $v : [a,b] \to \mathcal{V}$ fulfills* (C.3) *with respect to the norm on $\mathcal{V}$ and the space $\mathcal{V}$ is continuously embedded into $\mathcal{X}$, the boundedness of $\mathcal{R}$ by the norm on $\mathcal{X}$ (cf.* (2.4.7)*) implies that absolutely continuous curves w.r.t the norm on $\mathcal{V}$ are $\mathcal{R}$-absolutely continuous.*

The following useful observation on $\mathcal{R}$-absolutely continuous functions relies on a combination of a generalized version of Helly's Theorem and the Arzelà-Ascoli Theorem. The arguments are closely related to those in [MRS16, AGS05]. Let us first introduce the following notation:

Definition C.12. *For $[a,b] \subseteq [0,T]$ and $p \in [1,\infty]$, we denote by $\mathrm{AC}^p([a,b];\mathcal{X})$ the set of all $\mathcal{R}$-absolutely continuous curves $v : [a,b] \to \mathcal{V}$ whose generalized metric derivative $\mathcal{R}[v']$ is an element of $L^p(a,b)$.*

Proposition C.13. *Let $\mathcal{Z}$ be a reflexive Banach space, $\mathcal{V}, \mathcal{X}$ further Banach spaces such that* (2.4.1) *is satisfied and assume that $\mathcal{R} : \mathcal{X} \to [0,\infty)$ complies with* (2.4.7).

(a) The set $\mathrm{AC}^1([a,b];\mathcal{X}) \cap L^\infty((a,b);\mathcal{Z})$ is contained in $C([a,b];\mathcal{V})$ and there exists $C > 0$ such that for all $z \in \mathrm{AC}^1([a,b];\mathcal{X}) \cap L^\infty((a,b);\mathcal{Z})$ we have

$$\|z\|_{C([a,b];\mathcal{V})} \leq C\|z\|_{L^\infty((a,b);\mathcal{Z})}.$$

(b) Let $(z_n)_n \subset \mathrm{AC}^\infty([a,b];\mathcal{X}) \cap L^\infty((a,b);\mathcal{Z})$ be uniformly bounded in the sense that $A := \sup_n \|z_n\|_{L^\infty((a,b);\mathcal{Z})} < \infty$ and $B := \sup_n \|\mathcal{R}[z']\|_{L^\infty((a,b))} < \infty$.

Then there exists $z \in \mathrm{AC}^\infty([a,b];\mathcal{X}) \cap L^\infty((a,b);\mathcal{Z})$ and a (not relabelled) subsequence $(z_n)_n$ such that

$$z_n \to z \text{ uniformly in } C([a,b];\mathcal{V}), \tag{C.5}$$

$$\forall t \in [a,b] \quad z_n(t) \rightharpoonup z(t) \text{ weakly in } \mathcal{Z}. \tag{C.6}$$

(c) It is $L^\infty((a,b);\mathcal{Z}) \cap C([a,b];\mathcal{V}) \subset C_{weak}([a,b];\mathcal{Z})$.

Proof. In order to verify (a), let $z \in \mathrm{AC}^1([a,b];\mathcal{X}) \cap L^\infty((a,b);\mathcal{Z})$. By the Ehrling Lemma, [Wlo87], for every $\mu > 0$ there exists $C_\mu > 0$ such that for all $t,s \in [a,b]$ we have

$$\begin{aligned}\|z(t)-z(s)\|_{\mathcal{V}} &\le \mu\|z(t)-z(s)\|_{\mathcal{Z}} + C_\mu\|z(t)-z(s)\|_{\mathcal{X}} \\ &\le 2\mu\|z\|_{L^\infty(a,b;\mathcal{Z})} + \widetilde{C}_\mu \int_s^t \mathcal{R}[z'](r)\,\mathrm{d}r.\end{aligned}$$

This implies that $z \in C([a,b];\mathcal{V})$. The norm estimate follows from the embedding $\mathcal{Z} \hookrightarrow \mathcal{V}$ and (a) is proven.

For (b) let $(z_n)_n \subset \mathrm{AC}^\infty([a,b];\mathcal{X}) \cap L^\infty((a,b);\mathcal{Z})$ as in part (b) of the Proposition. Again by Ehrling's Lemma, for every $\mu > 0$ there exists $C_\mu > 0$ such that for all $t > s \in [a,b]$ and $n \in \mathbb{N}$ we have

$$\begin{aligned}\|z_n(t)-z_n(s)\|_{\mathcal{V}} &\le \mu\|z_n(t)-z_n(s)\|_{\mathcal{Z}} + C_\mu\|z_n(t)-z_n(s)\|_{\mathcal{X}} \\ &\le 2\mu A + \widetilde{C}_\mu \int_s^t \mathcal{R}[z_n'](r)\,\mathrm{d}r \le 2\mu A + C_\mu B|t-s|.\end{aligned}$$

This implies the equicontinuity of the sequence $(z_n)_n$ in $C([a,b];\mathcal{V})$. Indeed, for $\varepsilon > 0$ choose $\mu < \varepsilon/(4A)$ and $\delta < \varepsilon/(2BC_\mu)$. Then for all $n \in \mathbb{N}$, $s,t \in [a,b]$ with $|s-t| < \delta$ we have $\|z_n(s)-z_n(t)\|_{\mathcal{V}} \le \varepsilon$. Together with $z_n(t) \in \mathcal{K}$ for all t and n for a set $\mathcal{K}$ that is sequentially compact w.r.t. to the norm on $\mathcal{V}$, by the classical version of the Arzelà-Ascoli Theorem, see e.g. [Die69, Thm. 7.5.7], we obtain (C.5) for a subsequence. After possibly extracting a further subsequence, the generalized version of Helly's Theorem, see e.g. [MM05, Theorem 3.2] guarantees (C.6). By lower semicontinuity we conclude that for every $s < t \in [a,b]$

$$\mathcal{R}(z(t)-z(s)) \le \liminf_n \mathcal{R}(z_n(t)-z_n(s)) \le \int_s^t B\,\mathrm{d}s,$$

and thus $z \in \mathrm{AC}^\infty([a,b];\mathcal{X})$. Standard arguments finally imply that

$$L^\infty((a,b);\mathcal{Z}) \cap C([a,b];\mathcal{V}) \subset C_{\mathrm{weak}}([a,b];\mathcal{Z}).$$

■

The following statement is onely loosely connected to the previous ones in that it allows to interpret the continuous representative of a function that is continuous w.r.t. the norm on $\mathcal{V}$ and bounded w.r.t. the norm on $\mathcal{Z}$ as a $\mathcal{Z}$-valued function:

Lemma C.14. *Let $z \in C(0,S;\mathcal{V}) \cap L^\infty(0,S;\mathcal{Z})$, then it holds for all $s \in [0,S]$ that $\|z(s)\|_{\mathcal{Z}} \le \|z\|_{L^\infty(0,S;\mathcal{Z})}$.*

Proof. The Lebesgue differentiation theorem tells us that it holds for almost all $s \in [0,S]$ that

$$z(s) = \lim_{h\to 0} \tfrac{1}{2h} \int_{s-h}^{s+h} z(\sigma)\mathrm{d}\sigma \text{ in } \mathcal{V}. \tag{C.7}$$

Since z is continuous w.r.t. $\|\cdot\|_{\mathcal{V}}$, this holds true for all $s \in [0, S]$. To see this, let $s \in [0, S]$ be arbitrary and let $(s_n)_{n\in\mathbb{N}} \subset (0, S)$ be a sequence with $s_n \nearrow s$ and such that (C.7) holds true for every $n \in \mathbb{N}$. Let now $\delta > 0$ be arbitrary, and $h > 0$ be arbitrary, but fixed, then it holds for every $n \in \mathbb{N}$ that

$$
\begin{aligned}
z(s)-&\frac{1}{2h}\int_{s-h}^{s+h} z(\sigma)\mathrm{d}\sigma \\
&= (z(s)-z(s_n)) + \left(z(s_n) - \frac{1}{2h}\int_{s_n-h}^{s_n+h} z(\sigma)\mathrm{d}\sigma\right) + \frac{1}{2h}\left(\int_{s_n-h}^{s_n+h} z(\sigma)\mathrm{d}\sigma - \int_{s-h}^{s+h} z(\sigma)\mathrm{d}\sigma\right) \\
&= (z(s)-z(s_n)) + \left(z(s_n) - \frac{1}{2h}\int_{s_n-h}^{s_n+h} z(\sigma)\mathrm{d}\sigma\right) + \frac{1}{2h}\int_{s_n-h}^{s_n+h} z(\sigma-(s-s_n)) - z(\sigma)\mathrm{d}\sigma,
\end{aligned}
$$

where we consider the constant continuation of z to $[-h, S+h]$, if necessary. Now, since z is uniformly continuous on $[0, S]$, we find $N \in \mathbb{N}$ big enough, such that

$$
\|z(s)-z(s_N)\|_{\mathcal{V}} < \frac{\delta}{3} \quad \text{and} \quad \|z(\cdot-(s-s_N)) - z(\cdot)\|_{L^\infty(0,S;\mathcal{V})} < \frac{\delta}{3}.
$$

Lastly, there is $h_0 > 0$ so small that we find for all $0 < h < h_0$ that

$$
\|z(s_N) - \frac{1}{2h}\int_{s_N-h}^{s_N+h} z(\sigma)\mathrm{d}\sigma\|_{\mathcal{V}} < \frac{\delta}{3}.
$$

All in all, for every $\delta > 0$, there is $h_0 > 0$ such that for all $0 < h < h_0$, it holds that

$$
\|z(s) - \frac{1}{2h}\int_{s-h}^{s+h} z(\sigma)\mathrm{d}\sigma\|_{\mathcal{V}} < \delta,
$$

which is (C.7). Since it also holds for all $s \in [0, S]$ and all $h > 0$ that

$$
\|\frac{1}{2h}\int_{s-h}^{s+h} z(\sigma)\mathrm{d}\sigma\|_{\mathcal{Z}} \le \|z\|_{L^\infty(0,S;\mathcal{Z})},
$$

the continuous representative of z is in fact $\mathcal{Z}$-valued. ■

Functions of bounded variation

We will also make use of the space of functions of bounded variation. For the remainder of this section, we use the notations and assumptions introduced in Section 2.4.1

Definition C.15 (Functions of bounded variation). *(i) The pointwise total variation of a function $f : [a, b] \to B$ into a Banach space B is defined as*

$$
\mathrm{Var}_B(f;[a,b]) := \sup\Big\{\sum_{m=1}^{M} \|f(t_m) - f(t_{m-1})\|_B \;\Big|\; a = t_0 < t_1 < \cdots < t_{M-1} < t_M = b\Big\},
$$

and for $K \subseteq B$, we denote by $\mathrm{BV}([a,b];K,B)$ *and* $\mathrm{BV}([a,b];B)$ *the sets of all functions $f : [a,b] \to K$ or, respectively, $[a,b] \to B$ with finite pointwise total variation.*

(ii) *The pointwise total $\mathcal{R}$-variation of a function $v : [0,T] \to \mathcal{V}$ on the interval $[a,b] \subseteq [0,T]$ is defined via*

$$\mathrm{Var}_{\mathcal{R}}(v;[a,b]) := \sup\Big\{ \sum_{m=1}^{M} \mathcal{R}(v(t_m) - v(t_{m-1})) \quad \Big| \quad a = t_0 < t_1 < \cdots < t_{M-1} < t_M = b \Big\},$$

and for $K \subseteq \mathcal{V}$, we denote by $\mathrm{BV}([a,b];K,\mathcal{R})$ and $\mathrm{BV}([a,b];\mathcal{R})$ the set of all functions $v : [a,b] \to K$, or, respectively, $[a,b] \to \mathcal{V}$ with finite pointwise total $\mathcal{R}$-variation.

The pointwise total variation has the following representation formula and lower semicontinuity property.

Lemma C.16. (i) *If $v \in \mathrm{AC}([0,T];K,\mathcal{R})$, then $v \in \mathrm{BV}([0,T];K,\mathcal{R})$ and*

$$\mathrm{Var}_{\mathcal{R}}(v,[a,b]) = \int_a^b \mathcal{R}[v'](s)\,\mathrm{d}s \text{ for all } 0 \le a < b \le T. \tag{C.8}$$

(ii) *If $(v_n)_{n\in\mathbb{N}} \subset \mathrm{BV}([a,b];\mathcal{R})$ is a sequence of functions of bounded $\mathcal{R}$-variation and $v : [a,b] \to \mathcal{V}$ is a curve such that*

$$\infty > \liminf_{n\to\infty} \mathcal{R}(v_n(t) - v_n(s)) \ge \mathcal{R}(v(t) - v(s)) \text{ for all } s,t \in [0,T],$$

then $v \in \mathrm{BV}([a,b];\mathcal{R})$ and

$$\mathrm{Var}_{\mathcal{R}}(v,[a,b]) \le \liminf_{n\to\infty} \mathrm{Var}_{\mathcal{R}}(v_n,[a,b]) \text{ for all } 0 \le a < b \le T.$$

Proof. ad (i): It follows from the definitions that

$$\begin{aligned}\mathrm{Var}_{\mathcal{R}}(v,[a,b]) &\le \sup\Big\{ \sum_{m=1}^{M} \int_{t_{m-1}}^{t_m} \mathcal{R}[v'](s)\,\mathrm{d}s \quad \Big| \quad a = t_0 < t_1 < \cdots < t_{M-1} < t_M = b \Big\} \\ &= \int_a^b \mathcal{R}[v'](s)\,\mathrm{d}s. \end{aligned} \tag{C.9}$$

On the other hand, let $0 \le a < b \le T$, $t \in [a,b]$ and $h > 0$. If $t + h > T$, we constantly continue v to the interval $[T, t+h]$ with the value $v(T)$. Then it holds that

$$\frac{\mathcal{R}(v(t+h) - v(t))}{h} \le \frac{1}{h} \mathrm{Var}_{\mathcal{R}}(v,[t,t+h]),$$

where the left hand side converges to $\mathcal{R}[v'](t)$ with $h \to 0$ for almost all $t \in [0,T]$. By means of Fatou's Lemma, we obtain

$$\begin{aligned}\int_a^b \mathcal{R}[v'](t)\,\mathrm{d}t &\le \liminf_{h\to 0} \int_a^b \frac{\mathcal{R}(v(t+h) - v(t))}{h}\,\mathrm{d}t \le \liminf_{h\to 0} \int_a^b \frac{1}{h} \mathrm{Var}_{\mathcal{R}}(v,[t,t+h])\,\mathrm{d}t \\ &= \liminf_{h\to 0} \int_a^b \frac{1}{h}\Big(\mathrm{Var}_{\mathcal{R}}(v,[0,t+h]) - \mathrm{Var}_{\mathcal{R}}(v,[0,t]) \Big)\mathrm{d}t \\ &= \liminf_{h\to 0} \frac{1}{h}\Big(\int_{a+h}^{b+h} \mathrm{Var}_{\mathcal{R}}(v,[0,t])\,\mathrm{d}t - \int_a^b \mathrm{Var}_{\mathcal{R}}(v,[0,t])\,\mathrm{d}t \Big) \\ &= \liminf_{h\to 0} \Big(\frac{1}{h} \int_b^{b+h} \mathrm{Var}_{\mathcal{R}}(v,[0,t])\,\mathrm{d}t - \frac{1}{h} \int_a^{a+h} \mathrm{Var}_{\mathcal{R}}(v,[0,t])\,\mathrm{d}t \Big). \end{aligned} \tag{C.10}$$

We now consider the bounded real-valued function

$$V : [0,T] \to \mathbb{R}, \quad V(t) := \mathrm{Var}_{\mathcal{R}}(v,[0,t]) \text{ for } t \in (0,T], \quad V(0) := 0,$$

which is monotone and thus an element of $L^1(0,T)$. It therefore holds for almost all $s \in [0,T]$ that

$$\lim_{h\to 0} \frac{1}{h} \int_s^{s+h} \mathrm{Var}_{\mathcal{R}}(v,[0,t])\,\mathrm{d}t = \lim_{h\to 0} \frac{1}{h} \int_s^{s+h} V(t)\,\mathrm{d}t = V(s).$$

Returning to (C.10), we thus find for almost all $0 \le a < b \le T$ that

$$\int_a^b \mathcal{R}[v'](t)\,\mathrm{d}t \le V(b) - V(a) = \mathrm{Var}_{\mathcal{R}}(v,[a,b]).$$

In order to show that (C.8) holds for all choices $0 \le a < b \le T$, let first $b \in (0,T]$ and $(b_n)_{n\in\mathbb{N}} \subset (0,b)$ be a sequence such that $b_n \nearrow b$ with $n \to \infty$, and

$$\text{for all } n \in \mathbb{N} : \quad \int_0^{b_n} \mathcal{R}[v'](t)\,\mathrm{d}t = \mathrm{Var}_{\mathcal{R}}(v,[0,b_n]). \tag{C.11}$$

Then the functions $\chi_{[0,b_n]}\mathcal{R}[v']$, where χ_I denotes the characteristic function of an interval $I \subset [0,T]$, converge pointwisely to $\chi_{[0,b]}\mathcal{R}[v']$ and are uniformly bounded by $\chi_{[0,T]}\mathcal{R}[v']$, so that

$$\lim_{n\to\infty} \int_0^T \chi_{[0,b_n]}(s)\mathcal{R}[v'](s)\,\mathrm{d}s = \int_0^T \chi_{[0,b]}(s)\mathcal{R}[v'](s)\,\mathrm{d}s, \tag{C.12}$$

i.e., the left hand side of (C.11) converges. For the right hand side, we find with (C.9) that

$$\begin{aligned}
\mathrm{Var}_{\mathcal{R}}(v,[0,b]) &= \lim_{n\to\infty}\Big(\mathrm{Var}_{\mathcal{R}}(v,[0,b_n]) + \mathrm{Var}_{\mathcal{R}}(v,[b_n,b])\Big) \\
&\le \lim_{n\to\infty}\Big(\mathrm{Var}_{\mathcal{R}}(v,[0,b_n]) + \int_{b_n}^b \mathcal{R}[v'](s)\,\mathrm{d}s\Big) \\
&= \lim_{n\to\infty} \mathrm{Var}_{\mathcal{R}}(v,[0,b_n]),
\end{aligned} \tag{C.13}$$

where we argue as in (C.11) in the last step. Note that converse of (C.13) holds trivially, which is why we have in fact an identity instead of an estimate. This means that the right-hand side of (C.11) converges as well, and we find for arbitrary $b \in (0,T]$ that

$$\int_0^b \mathcal{R}[v'](t)\,\mathrm{d}t = \mathrm{Var}_{\mathcal{R}}(v,[0,b]).$$

For arbitrary $a \in [0,T]$, using that $\mathrm{Var}_{\mathcal{R}}(v,[a,b]) = \mathrm{Var}_{\mathcal{R}}(v,[0,b]) - \mathrm{Var}_{\mathcal{R}}(v,[0,a])$, we finally find (C.8).

ad (ii): Let Z be the set of all finite partitions of the interval $[a,b]$ and for $\mathfrak{z} := (a = t_0 < t_1 < \cdots < t_M = b) \in Z$, let $f_{\mathfrak{z}} : \mathrm{BV}([0,T];\mathcal{R}) \to \mathbb{R}$ be the function defined by

$$f_{\mathfrak{z}}(v) := \sum_{m=1}^{M} \mathcal{R}(v(t_m) - v(t_{m-1})),$$

so that we have for all $v \in \mathrm{BV}([a,b];\mathcal{R})$

$$\mathrm{Var}_{\mathcal{R}}(v,[a,b]) = \sup_{\mathfrak{z}\in Z} f_{\mathfrak{z}}(v).$$

Now, if $(v_n)_{n\in\mathbb{N}} \subset \mathrm{BV}([a,b];\mathcal{R})$ is a sequence of functions such that the inequality $\liminf_{n\to\infty} \mathcal{R}(v_n(t) - v_n(s)) \geq \mathcal{R}(v(t) - v(s))$ holds for all $s,t \in [0,T]$, we have for all $\mathfrak{z} \in Z$ that

$$\liminf_{n\to\infty} f_{\mathfrak{z}}(v_n) \geq f_{\mathfrak{z}}(v).$$

Let $(v_{n_k})_{k\in\mathbb{N}}$ be a subsequence such that $\mathrm{Var}_{\mathcal{R}}(v_{n_k},[a,b]) \to \liminf_{n\to\infty} \mathrm{Var}_{\mathcal{R}}(v_n,[a,b])$ for $k \to \infty$. Then we have for all $\mathfrak{z} \in Z$

$$\liminf_{n\to\infty} \mathrm{Var}_{\mathcal{R}}(v_n,[a,b]) = \lim_{k\to\infty} \mathrm{Var}_{\mathcal{R}}(v_{n_k},[a,b]) \geq \liminf_{k\to\infty} f_{\mathfrak{z}}(v_{n_k}) \geq f_{\mathfrak{z}}(v),$$

and thus

$$\mathrm{Var}_{\mathcal{R}}(v,[a,b]) = \sup_{\mathfrak{z}\in Z} f_{\mathfrak{z}}(v) \leq \liminf_{n\to\infty} \mathrm{Var}_{\mathcal{R}}(v_n,[a,b]) < \infty.$$

■

Remark C.17. *The statements that were made in Lemma C.16 are also true for absolutely continuous functions or functions of bounded variation w.r.t. a norm. In this case, the condition in statement (ii) is fulfilled in case of pointwise weak convergence of v_n to v.*

In the next lemma, we prove that if $v \in W^{1,1}(0,T;\mathcal{X})$, its total $\mathcal{R}$-variation can be represented by means of its derivative.

Lemma C.18. *Let $v \in W^{1,1}(0,T;\mathcal{X})$. Then it holds for all $0 \leq a < b \leq T$ that*

$$\mathrm{Var}_{\mathcal{R}}(v,[a,b]) = \int_a^b \mathcal{R}(\dot{v}(s))\mathrm{d}s.$$

Proof. Since $\mathcal{R}$ is lower semicontinuous, it follows immediately that at any point $s \in [a,b]$ of differentiability, we have

$$\mathcal{R}[v'](s) = \liminf_{h\to 0} \mathcal{R}\Big(\frac{v(s+h)-v(s)}{h}\Big) \geq \mathcal{R}(\dot{v}(s)),$$

so that by Lemma C.16.(i), it follows that

$$\mathrm{Var}_{\mathcal{R}}(v,[a,b]) = \int_a^b \mathcal{R}[v'](s)\mathrm{d}s \geq \int_a^b \mathcal{R}(\dot{v}(s))\mathrm{d}s.$$

On the other hand, Jensens' inequality (see [HvVW16, Lemma 1.2.11]) yields for any partition $a = t_0 < t_1 < \cdots < t_N = b$ of the interval $[a,b]$ and any $i \in \{0,\dots,N-1\}$ that

$$\mathcal{R}\Big(\frac{1}{t_{i+1}-t_i}\int_{t_i}^{t_{i+1}} \dot{v}(s)\mathrm{d}s\Big) \leq \frac{1}{t_{i+1}-t_i}\int_{t_i}^{t_{i+1}} \mathcal{R}(\dot{v}(s))\mathrm{d}s.$$

Since the Fundamental Theorem of Calculus is valid in $W^{1,1}(0,T;\mathcal{X})$, even if $\mathcal{X}$ is not reflexive, see [HvVW16, Prop. 2.5.9], we can use the 1-homogeneity of $\mathcal{R}$ to find that

$$\mathcal{R}(v(t_{i+1}) - v(t_i)) = \mathcal{R}\Big(\int_{t_i}^{t_{i+1}} \dot{v}(s)\mathrm{d}s\Big) \leq \int_{t_i}^{t_{i+1}} \mathcal{R}(\dot{v}(s))\mathrm{d}s.$$

Summation over $i = 0,\dots,N-1$ yields

$$\sum_{i=0}^{N-1} \mathcal{R}(v(t_{i+1}) - v(t_i)) \leq \int_a^b \mathcal{R}(\dot{v}(s))\mathrm{d}s,$$

whence

$$\mathrm{Var}_{\mathcal{R}}(v,[a,b]) = \int_a^b \mathcal{R}[v'](s)\mathrm{d}s \leq \int_a^b \mathcal{R}(\dot{v}(s))\mathrm{d}s.$$

■

It is well-known, see, e.g., [Fed69, 2.5.16], that functions in $\mathrm{BV}([a,b];\mathcal{V})$ have left and right limits at every time $t \in (a,b)$, and a right or left limit at a or b, respectively. This fact allows for the definition of a jump set, see (2.2.17). In the following lemma, we show that the same is true for functions in $\mathrm{BV}([a,b];K,\mathcal{R})$, if $K \subset \mathcal{V}$ is a compact subset of $\mathcal{V}$. In particular, these functions have have left and right limits w.r.t. the norm on $\mathcal{V}$, and not only w.r.t. $\mathcal{R}$.

Lemma C.19. *Let $K \subset \mathcal{Z}$ be a subset such that $\sup_{z\in K} \|z\|_{\mathcal{Z}} < \infty$.*

(i) There exists a continuous from the right, subadditive, monotonely increasing function $\Omega_K : [0,\infty) \to [0,\infty)$ such that for all $z_1, z_2 \in K$, it holds

$$\|z_1 - z_2\|_{\mathcal{V}} \leq \Omega_K(\mathcal{R}(z_1 - z_2)).$$

In particular, it holds that $\Omega_K(r) \xrightarrow{r\to 0} 0$.

(ii) Let now $z \in \mathrm{BV}([a,b];K,\mathcal{R})$. For every $t \in [a,b]$, the limits

$$z(t_-) := \lim_{s\nearrow t} z(s);\ z(t_+) := \lim_{s\searrow t} z(s), \quad \text{where } z(a_-) := z(a);\ z(b_+) := z(b), \tag{C.14}$$

exist w.r.t. the norm $\|\cdot\|_{\mathcal{V}}$, and we define

$$J_z := \{t \in [0,T] \,|\, z(t_-) \neq z(t) \text{ or } z(t) \neq z(t_+)\}.$$

Proof. ad (i): The arguments are inspired by [MRS16, Section 2.2]. First, an application of Ehrling's Interpolation Lemma yields for every $\delta > 0$ a constant $M_\delta > 0$ such that for all $z \in \mathcal{Z}$, it holds

$$\|z\|_{\mathcal{V}} \leq \delta\|z\|_{\mathcal{Z}} + M_\delta\|z\|_{\mathcal{X}}.$$

Now, since K is bounded from above with respect to $\|\cdot\|_{\mathcal{Z}}$ and $\|\cdot\|_{\mathcal{X}}$ is bounded from above by $\mathcal{R}$, this implies that for every $\delta > 0$, there exists $N_\delta > 0$ such that for all $z \in K$, it holds

$$\|z\|_{\mathcal{V}} \leq \delta + N_\delta \mathcal{R}(z).$$

We now set for $r \geq 0$

$$\Omega_K(r) := \inf_{\delta>0}\{\delta + N_\delta r\} \text{ and } \Omega_K(0) = 0.$$

It is obvious that Ω_K is monotonely increasing, and since for all $s, r \in [0,\infty)$, it holds that

$$\frac{r}{r+s}\Omega_K(r+s) = \inf_{\delta>0}\{\frac{r}{r+s}\delta + N_\delta r\} \leq \Omega_K(r),$$

as well as

$$\frac{s}{r+s}\Omega_K(r+s) = \inf_{\delta>0}\{\frac{s}{r+s}\delta + N_\delta s\} \leq \Omega_K(s),$$

it is also subadditive. Now let $r \in [0,\infty)$ and $(r_n)_{n\in\mathbb{N}}$ be a sequence such that $r_n \to r$ and $r_n \geq r$ for all $n \in \mathbb{N}$. For arbitrary $\varepsilon > 0$, choose $\delta_0 > 0$ such that $\delta_0 + N_{\delta_0} r \leq \inf_{\delta>0}\{\delta + N_\delta r\} + \frac{\varepsilon}{2}$, and $L \in \mathbb{N}$ such that for all $n \geq L$, it holds $r_n - r < \frac{\varepsilon}{2N_{\delta_0}}$. It follows for all $n \geq L$ that

$$0 \leq \Omega_K(r_n) - \Omega_K(r) \leq \delta_0 + N_{\delta_0} r_n - (\delta_0 + N_{\delta_0} r) + \frac{\varepsilon}{2} = N_{\delta_0}(r_n - r) + \frac{\varepsilon}{2} < \varepsilon,$$

and thus $\Omega_K(r_n) \to \Omega_K(r)$.

ad (ii): Let $z \in \mathrm{BV}([a,b];K,\mathcal{R})$. We procede as in [Fed69, 2.5.16] and define the monotonely increasing scalar function

$$V : (a,b] \to \mathbb{R}, \quad V(t) := \mathrm{Var}_{\mathcal{R}}(z,[a,t]).$$

Since V is also a bounded function, its left and right limits, defined as in (C.14) exist everywhere. Let now $t \in (a,b)$, and $(s_n)_{n\in\mathbb{N}} \subset (a,t)$ be a (w.l.o.g. monotonely increasing) sequence with $\lim_{n\to\infty} s_n = t$. With (i), it follows that

$$\begin{aligned}\|z(s_m) - z(s_n)\|_{\mathcal{V}} &\leq \Omega_K\Big(\mathcal{R}(z(s_m) - z(s_n))\Big) \leq \Omega_K\Big(\mathrm{Var}_{\mathcal{R}}(z;[s_n,s_m])\Big)\\ &= \Omega_K\Big(\mathrm{Var}_{\mathcal{R}}(z;[a,s_m]) - \mathrm{Var}_{\mathcal{R}}(z;[a,s_n])\Big) \leq \Omega_K\Big(V(t_-) - V(s_n)\Big),\end{aligned}$$

whence $\big(z(s_n)\big)_{n\in\mathbb{N}}$ converges w.r.t. $\|\cdot\|_{\mathcal{V}}$ to a limit $z_* \in \mathcal{V}$. If $(\sigma_n)_{n\in\mathbb{N}}$ is a further sequence with $\sigma_n \nearrow t$, it follows from (changing the roles of σ_n and s_n if necessary)

$$\begin{aligned}\|z(\sigma_n) - z(s_n)\|_{\mathcal{V}} &\leq \Omega_K(\mathcal{R}(z(\sigma_n) - z(s_n))) \leq \Omega_K\Big(\mathrm{Var}_{\mathcal{R}}(z;[s_n,\sigma_n])\Big)\\ &= \Omega_K\Big(\mathrm{Var}_{\mathcal{R}}(z;[a,\sigma_n]) - \mathrm{Var}_{\mathcal{R}}(z;[a,s_n])\Big) \leq \Omega_K\Big(V(t_-) - V(s_n)\Big),\end{aligned}$$

that the limit z_* is unique, and we denote it by $z(t_-)$. The existence of $z(t_+)$ can be shown in the same way. ∎

Appendix D

Additional properties of energy and dissipation functional

The following Lemmas enable us use the existence result in Theorem 2.2 in [MRS13] for the proof of Theorem 3.1.3.

Lemma D.1. *For every $\varepsilon > 0$, the following assertions hold true:*

$$\mathcal{R}_\varepsilon : \mathcal{V} \to [0,\infty) \text{ is lower semicontinuous and convex,} \tag{$\mathcal{R}_1$}$$

$$\mathcal{R}(0) = 0, \qquad \lim_{\|v\|_{\mathcal{V}}\to\infty} \frac{\mathcal{R}_\varepsilon(v)}{\|v\|_{\mathcal{V}}} = \infty, \qquad \lim_{\|\xi\|_{\mathcal{V}^*}\to\infty} \frac{\mathcal{R}^*_\varepsilon(\xi)}{\|\xi\|_{\mathcal{V}^*}} = \infty, \tag{$\mathcal{R}_2$}$$

$$\text{for all } v \in \mathcal{V} \text{ and } \eta_1, \eta_2 \in \partial\mathcal{R}_\varepsilon(v): \quad \mathcal{R}^*_\varepsilon(\eta_1) = \mathcal{R}^*_\varepsilon(\eta_2). \tag{$\mathcal{R}_3$}$$

Here, $\mathcal{R}_\varepsilon := \mathcal{R} + \mathcal{R}_{2,\varepsilon}$, with $\mathcal{R}_{2,\varepsilon}(v) := \frac{\varepsilon}{2}\|v\|^2_{\mathcal{V}}$, see (2.4.11)*.*

Proof. ($\mathcal{R}_1$) follows from the lower semicontinuity and convexity of the norm on $\mathcal{V}$ and the first and second of ($\mathcal{R}_2$) are obvious. In order to obtain the third assertion, we use the reflexivity of the space $\mathcal{V}$ to choose for every $\xi \in \mathcal{V}^*$ an element $\tilde{v} \in \mathcal{V}^{**} = \mathcal{V}$ with $\|\tilde{v}\|_{\mathcal{V}} = 1$ and $\langle \xi, \tilde{v}\rangle_{\mathcal{V}} = \|\xi\|_{\mathcal{V}^*}$ and then set $v := \frac{\|\xi\|_{\mathcal{V}^*}}{\varepsilon}\tilde{v}$. We can now estimate

$$\begin{aligned} \mathcal{R}^*_\varepsilon(\xi) &\geq \langle \xi, v\rangle - C\|v\|_{\mathcal{V}} - \frac{\varepsilon}{2}\|v\|^2_{\mathcal{V}} \\ &\geq \frac{1}{2\varepsilon}\|\xi\|^2_{\mathcal{V}^*} - \frac{C}{\varepsilon}\|\xi\|_{\mathcal{V}^*} \end{aligned}$$

for a constant $C > 0$ depending only on the embedding $\mathcal{V} \hookrightarrow \mathcal{X}$, which implies the third assertion of ($\mathcal{R}_2$). Next, in order to show ($\mathcal{R}_3$), we want to show that for all $v \in \mathcal{V}$, the mapping $\lambda \mapsto \mathcal{R}_\varepsilon(\lambda v)$ is differentiable in $\lambda = 1$. Indeed, we have

$$\begin{aligned} \lim_{\lambda\to 0} \frac{\mathcal{R}_\varepsilon(v+\lambda v) - \mathcal{R}_\varepsilon(v)}{\lambda} &= \lim_{\lambda\to 0}\left(\mathcal{R}(v) + \frac{\frac{\varepsilon}{2}\|(1+\lambda)v\|^2_{\mathcal{V}} - \frac{\varepsilon}{2}\|v\|^2_{\mathcal{V}}}{\lambda}\right) \\ &= \mathcal{R}(v) + \lim_{\lambda\to 0}\frac{\frac{\varepsilon}{2}(2\lambda+\lambda^2)\|v\|^2_{\mathcal{V}}}{\lambda} \\ &= \mathcal{R}(v) + \varepsilon\|v\|^2_{\mathcal{V}}. \end{aligned}$$

Let now $\eta \in \partial\mathcal{R}_\varepsilon(v)$. Since we also have for all $\lambda \in (0,1)$

$$\frac{\mathcal{R}_\varepsilon(v+\lambda v)-\mathcal{R}_\varepsilon(v)}{\lambda} \geq \langle \eta, v\rangle_{\mathcal{V}} \geq \frac{\mathcal{R}_\varepsilon(v-\lambda v)-\mathcal{R}_\varepsilon(v)}{-\lambda},$$

and the limits for $\lambda \to 0$ on the right hand side and on the left hand side exist and coincide, it follows that for all $\eta_1, \eta_2 \in \partial\mathcal{R}_\varepsilon(v)$ the identity

$$\langle \eta_1, v\rangle_{\mathcal{V}} = \langle \eta_2, v\rangle_{\mathcal{V}} = \mathcal{R}(v) + \varepsilon\|v\|_{\mathcal{V}}^2$$

must hold, which implies $(\mathcal{R}_3)$. ■

Lemma D.2. *For every $\varepsilon > 0$, the energy functional $\mathcal{I}$ and the dissipation potential $\mathcal{R}_\varepsilon$ have the following properties:*
Lower semicontinuity and boundedness from below

$$\begin{aligned} &z \mapsto \mathcal{I}(\ell(t), z) \textit{ is lower semicontinuous w.r.t. the strong topology on } \mathcal{V}, \\ &\exists C_0 \in \mathbb{R} : \forall (t,z) \in [0,T]\times\mathcal{Z} : \mathcal{I}(\ell(t),z) \geq C_0. \end{aligned} \tag{$\mathcal{I}_0$}$$

Coercivity

$$\begin{aligned} &\forall \tau_0 > 0,\ t \in [0,T]: \\ &z \mapsto \mathcal{I}(\ell(t),z) + \tau_0\mathcal{R}_\varepsilon(\frac{z}{\tau_0}) \textit{ has compact sublevels w.r.t. the strong topology on } \mathcal{V}. \end{aligned} \tag{$\mathcal{I}_1$}$$

Variational sum rule *If for some $v_0 \in \mathcal{V}$ and $\tau > 0$, the point $\overline{v}$ is a minimizer of the map $v \mapsto \mathcal{I}(\ell(t), v) + \tau\mathcal{R}_\varepsilon((v - v_0)/\tau)$, then $\overline{v}$ satisfies the Euler-Lagrange equation*

$$0 \in \mathrm{D}_z\mathcal{I}(\ell(t),\overline{v}) + \partial\mathcal{R}_\varepsilon((\overline{v}-v_0)/\tau) \textit{ in } \mathcal{V}^*. \tag{$\mathcal{I}_2$}$$

Time-dependence

$$\left.\begin{aligned} &\forall z \in \mathcal{Z}:\ t \mapsto \mathcal{I}(\ell(t),z) \textit{ is in } W^{1,\infty}(0,T), \\ &\textit{differentiable a.e. with derivative } \mathcal{P}(t,z) := \partial_t\mathcal{I}(\ell(t),z) = \langle\dot{\ell}(t), z\rangle_{\mathcal{V}}; \\ &\exists C_1 > 0:\ \textit{f.a.a. } t \in [0,T], \forall\, z \in \mathcal{Z} : |\partial_t\mathcal{I}(\ell(t),z)| \leq C_1\mathcal{I}(\ell(t),z). \end{aligned}\right\} \tag{$\mathcal{I}_3$}$$

Chain rule *For every $v \in \mathrm{AC}([0,T];\mathcal{V})$ with*

$$\begin{aligned} &\sup_{t\in[0,T]}|\mathcal{I}(\ell(t),v(t))| < \infty, \quad \mathrm{D}_z\mathcal{I}(\ell(\cdot),v(\cdot)) \in L^1(0,T;\mathcal{V}^*), \textit{ and} \\ &\int_0^T \mathcal{R}_\varepsilon(\dot{v}(t))\,\mathrm{d}t < \infty, \qquad \int_0^T \mathcal{R}_\varepsilon^*(-\mathrm{D}_z\mathcal{I}(\ell(t),v(t)))\,\mathrm{d}t < \infty, \end{aligned} \tag{D.1}$$

the map $t \mapsto \mathcal{I}(\ell(t),v(t))$ is absolutely continuous and

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}(\ell(t),v(t)) = \langle \mathrm{D}_z\mathcal{I}(\ell(t),v(t)), \dot{v}(t)\rangle_{\mathcal{V}} + \partial_t\mathcal{I}(\ell(t),v(t)) \textit{ for a.a. } t \in (0,T). \tag{$\mathcal{I}_4$}$$

Weak closedness of $(\mathcal{I}, \mathrm{D}_z\mathcal{I})$ *For all* $t \in [0,T]$ *and for all sequences* $(v_n)_{n\in\mathbb{N}} \subset \mathcal{V}$ *we have the following property:*

$$\begin{aligned}&\underline{\textit{if}}\\ &\quad v_n \rightharpoonup v \text{ in } \mathcal{V}, \mathrm{D}_z\mathcal{I}(\ell(t),v_n) \rightharpoonup \xi \text{ in } \mathcal{V}^*, \partial_t\mathcal{I}(\ell(t),v_n) \to p, \mathcal{I}(\ell(t),v_n) \to I \text{ in } \mathbb{R},\\ &\underline{\textit{then}}\\ &\quad v_n \to v \text{ strongly in } \mathcal{Z} \text{ and } \xi = \mathrm{D}_z\mathcal{I}(\ell(t),v) \text{ and } p = \partial_t\mathcal{I}(\ell(t),v) \text{ and } I = \mathcal{I}(\ell(t),v).\end{aligned} \tag{$\mathcal{I}_5$}$$

It further holds for C_1 *from* ($\mathcal{I}_3$) *for all* $s < t \in [0,T]$ *and* $z \in \mathcal{Z}$*:*

$$\mathcal{I}(\ell(t),z) \le \mathcal{I}(\ell(s),z)e^{C_1(t-s)}. \tag{D.2}$$

Proof. Proof of ($\mathcal{I}_0$): We start by showing that $\mathcal{I}(\ell(t),\cdot)$ is lower semicontinuous with respect to the topology on $\mathcal{V}$. It is clear that the restriction of $\mathcal{I}(\ell(t),\cdot)$ to $\mathcal{Z}$ is lower semicontinuous with respect to the weak topology on $\mathcal{Z}$, since the linear term $\langle\ell(t),\cdot\rangle_{\mathcal{V}}$ is obviously continuous with respect to the weak topology on $\mathcal{Z}$ and the same is true for $\mathcal{F}$ according to Lemma 2.4.2. Now, the continuity of $\langle A\cdot,\cdot\rangle_{\mathcal{Z}}$ with respect to the strong topology on $\mathcal{Z}$ together with its convexity yield the weak lower semicontinuity of $\langle A\cdot,\cdot\rangle_{\mathcal{Z}}$ and hence of $\mathcal{I}(\ell(t),\cdot)$ on $\mathcal{Z}$. Next, we want to show that $\mathcal{I}(\ell(t),\cdot)$ is lower semicontinuous on $\mathcal{Z}$ with respect to the topology on $\mathcal{V}$. To this end, let $(z_n)_{n\in\mathbb{N}} \subset \mathcal{Z}$ be a sequence and $z \in \mathcal{V}$ such that $z_n \to z$ in $\mathcal{V}$. We have to distinguish two cases: If $z \in \mathcal{V}\setminus\mathcal{Z}$, it suffices to show that $\liminf_{n\to\infty}\mathcal{I}(\ell(t),z_n) = \infty$. Indeed, assume that $\liminf_{n\to\infty}\mathcal{I}(\ell(t),z_n) =: C < \infty$ and choose a subsequence $(z_{n_k})_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty}\mathcal{I}(\ell(t),z_{n_k}) = C$. As we shall see in (D.6), this implies that $\sup_{k\in\mathbb{N}}\|z_{n_k}\|_{\mathcal{Z}} < \infty$, so that there exist a $\mathcal{Z}$-weakly converging sub-subsequence and a weak limit $z_0 \in \mathcal{Z}$. Since this implies strong convergence to z_0 in $\mathcal{V}$, we conclude that $z = z_0 \in \mathcal{Z}$, which contradicts our assumption $z \in \mathcal{V}\setminus\mathcal{Z}$. Thus, we have $\liminf_{n\to\infty}\mathcal{I}(\ell(t),z_n) = \infty = \mathcal{I}(\ell(t),z)$. In the second case, we assume that $z \in \mathcal{Z}$. If $\liminf_{n\to\infty}\mathcal{I}(\ell(t),z_n) = \infty$, then the inequality $\liminf_{n\to\infty}\mathcal{I}(\ell(t),z_n) \ge \mathcal{I}(\ell(t),z)$ ensues immediately. Now, let $\liminf_{n\to\infty}\mathcal{I}(\ell(t),z_n) =: C < \infty$. Suppose that $C < \mathcal{I}(\ell(t),z)$. As in the first case, we choose a subsequence $(z_{n_k})_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty}\mathcal{I}(\ell(t),z_{n_k}) = C$ and a $\mathcal{Z}$-weakly converging sub-subsequence $(z_{n_{k_l}})_{l\in\mathbb{N}}$ whose weak limit is z. Now, the weak lower semicontinuity of $\mathcal{I}(\ell(t),\cdot)$ on $\mathcal{Z}$ yields

$$\mathcal{I}(\ell(t),z) \le \liminf_{l\to\infty}\mathcal{I}(\ell(t),z_{n_{k_l}}) = \lim_{k\to\infty}\mathcal{I}(\ell(t),z_{n_k}) = C < \mathcal{I}(\ell(t),z),$$

a contradiction. Thus, if $\liminf_{n\to\infty}\mathcal{I}(\ell(t),z_n) < \infty$, it has to be greater or equal to $\mathcal{I}(\ell(t),z)$. Note that it would have been sufficient to assume that $z_n \rightharpoonup z$ weakly in $\mathcal{V}$, thus $\mathcal{I}(\ell(t),\cdot)$ is even weakly lower semicontinuous on $\mathcal{Z}$ with respect to the topology on $\mathcal{V}$. Hence, the same is true for its continuation to $\mathcal{V}$.

Next, we determine the constant C_0 such that the second of ($\mathcal{I}_0$) is met. To this end, we estimate for all $z \in \mathcal{Z}$:

$$\begin{aligned}\tfrac{1}{2}\langle Az, z\rangle_{\mathcal{Z}} + \mathcal{F}(z) - \langle \ell(t), z\rangle_{\mathcal{V}} &\geq \tfrac{\alpha}{2}\|z\|_{\mathcal{Z}}^2 - \|\ell(t)\|_{\mathcal{V}^*}\|z\|_{\mathcal{V}} \\ &\geq \tfrac{\alpha}{2}C_{\mathcal{Z}\hookrightarrow\mathcal{V}}^{-2}\|z\|_{\mathcal{V}}^2 - \frac{1}{2\delta}\|\ell(t)\|_{\mathcal{V}^*}^2 - \frac{\delta}{2}\|z\|_{\mathcal{V}}^2 \\ &\geq C_c\|z\|_{\mathcal{V}}^2 - \frac{1}{2\delta_0}\|\ell\|_{L^\infty(0,T;\mathcal{V}^*)}^2 \qquad \text{(D.3)} \\ &\geq 2C_c\|z\|_{\mathcal{V}} - C_c - \frac{1}{2\delta_0}\|\ell\|_{L^\infty(0,T;\mathcal{V}^*)}^2, \qquad \text{(D.4)}\end{aligned}$$

where we chose

$$0 < \delta_0 < \alpha C_{\mathcal{Z}\hookrightarrow\mathcal{V}}^{-2} \qquad \text{(D.5)}$$

for the embedding constant $C_{\mathcal{Z}\hookrightarrow\mathcal{V}}$, so that $C_c := \frac{\alpha}{2}C_{\mathcal{Z}\hookrightarrow\mathcal{V}}^{-2} - \frac{\delta_0}{2} > 0$. We conclude $(\mathcal{I}_0)$ from (D.3) for the constant $C_0 := -\frac{1}{2\delta_0}\|\ell\|_{L^\infty(0,T;\mathcal{V}^*)}^2$. In fact, as is noted in [MRS13], if the functionals $\mathcal{I}(\ell(t),\cdot)$ are bounded from below by some constant independent of t, up to a translation, it is not restrictive to assume such a constant to be strictly positive. For this reason, for the remainder of this proof, we consider instead of $\mathcal{I}$ the functional that is translated by

$$C_\ell := C_c - C_0 = C_c + \frac{1}{2\delta_0}\|\ell\|_{L^\infty(0,T;\mathcal{V}^*)}^2 > 0.$$

Proof of $(\mathcal{I}_3)$: Let us first note that the estimate (D.4) now reads

$$\|z\|_{\mathcal{V}} \leq C\mathcal{I}(\ell(t), z)$$

for $C := 1/(2C_c) > 0$. We can now verify $(\mathcal{I}_3)$: For all $z \in \mathcal{Z}$ and almost all $t \in [0,T]$, it holds that

$$\lim_{s\to t}\frac{\mathcal{I}(\ell(t),z) - \mathcal{I}(\ell(s),z)}{t-s} = \lim_{s\to t}\langle\frac{\ell(t)-\ell(s)}{t-s}, z\rangle_{\mathcal{V}} = \langle\dot{\ell}(t), z\rangle_{\mathcal{V}}.$$

Thus, we can estimate the time derivative $\partial_t\mathcal{I}(\ell(t),z) = \langle\dot{\ell}(t),z\rangle_{\mathcal{V}}$ by

$$|\partial_t\mathcal{I}(\ell(t),z)| \leq \|\dot{\ell}(t)\|_{\mathcal{V}^*}\|z\|_{\mathcal{V}} \leq \|\dot{\ell}\|_{L^\infty(0,T;\mathcal{V}^*)}C\mathcal{I}(\ell(t),z)$$

for all $z \in \mathcal{Z}$ and almost all $t \in [0,T]$.
Proof of $(\mathcal{I}_1)$: We now show the compactness of the energy sublevels with respect to the topology on $\mathcal{V}$ as follows: For all $t \in [0,T]$ and $E > 0$, $\mathcal{I}(\ell(t),z) + \tau_0\mathcal{R}_\varepsilon(\frac{z}{\tau_0}) \leq E$ implies that

$$\begin{aligned}E \geq \mathcal{I}(\ell(t),z) &\geq \tfrac{\alpha}{2}\|z\|_{\mathcal{Z}}^2 - \frac{\delta_0}{2}\|z\|_{\mathcal{V}}^2 \\ &\geq \frac{\alpha - \delta_0 C_{\mathcal{Z}\hookrightarrow\mathcal{V}}^2}{2}\|z\|_{\mathcal{Z}}^2, \qquad \text{(D.6)}\end{aligned}$$

where the last constant is greater than zero thanks to the choice of δ_0 in (D.5). Thus, the sublevels of $\mathcal{I}(\ell(t),\cdot) + \tau_0\mathcal{R}_\varepsilon(\frac{\cdot}{\tau_0})$ are bounded in $\mathcal{Z}$. Since $\mathcal{Z}$ is reflexive,

this implies the existence of a weakly convergent subsequence in $\mathcal{Z}$, which converges strongly in $\mathcal{V}$, and we infer $(\mathcal{I}_1)$.
Proof of $(\mathcal{I}_2)$: Let for some $v_0 \in \mathcal{V}$ and $\tau > 0$ the point $\overline{v}$ be a minimizer of the map $v \mapsto \mathcal{I}(\ell(t), v) + \tau\mathcal{R}_\varepsilon((v - v_0)/\tau)$. Since $\mathcal{I}$ takes finite values only in $\mathcal{Z}$, we infer that $\overline{v} \in \mathcal{Z}$ and that we have the following inclusion in $\mathcal{Z}^*$:

$$0 \in \mathrm{D}_z\mathcal{I}(\ell(t), \overline{v}) + \partial\mathcal{R}_\varepsilon(\frac{\overline{v} - v_0}{\tau}),$$

which is why there exists $\xi \in \partial\mathcal{R}_{2,\varepsilon}(\frac{\overline{v}-v_0}{\tau})$ such that

$$-\mathrm{D}_z\mathcal{I}(\ell(t), \overline{v}) - \xi \in \partial\mathcal{R}(\frac{\overline{v} - v_0}{\tau}). \tag{D.7}$$

From Lemma 2.4.4, we know that $\partial\mathcal{R}(\frac{\overline{v}-v_0}{\tau}) \subset \partial\mathcal{R}(0) \subset \mathcal{V}^*$ and that $\partial\mathcal{R}_{2,\varepsilon}(z) \subset \mathcal{V}^*$ for all $z \in \mathcal{V}$, so that $\xi \in \mathcal{V}^*$, and we conclude that (D.7) is valid in $\mathcal{V}^*$ as well.
Proof of $(\mathcal{I}_4)$: Let $v \in \mathrm{AC}([0, T]; \mathcal{V})$ fulfill (D.1). By (D.6), we infer that we have $\|v\|_{L^\infty(0,T;\mathcal{Z})} =: \rho < \infty$. In order to show absolute continuity, we decompose the scalar function $t \mapsto \mathcal{I}(\ell(t), v(t))$ into three summands and show absolute continuity of each of those. First, consider the map $H : [0, T] \to \mathbb{R}$, $H(t) := \langle \ell(t), v(t)\rangle_\mathcal{V}$. Let $s, t \in [0, T]$. Then we have the following estimate:

$$\begin{aligned}
|H(t)-&H(s)| \\
&= |\langle \ell(t), v(t)\rangle_\mathcal{V} - \langle \ell(s), v(s)\rangle_\mathcal{V}| \\
&= |\langle \ell(t) - \ell(s), v(t) - v(s)\rangle_\mathcal{V} + \langle \ell(t) - \ell(s), v(s)\rangle_\mathcal{V} + \langle \ell(s), v(t) - v(s)\rangle_\mathcal{V}| \\
&\le \|\ell(t) - \ell(s)\|_{\mathcal{V}^*}\|v(t) - v(s)\|_\mathcal{V} + \|\ell(t) - \ell(s)\|_{\mathcal{V}^*}\|v(s)\|_\mathcal{V} + \|\ell(s)\|_{\mathcal{V}^*}\|v(t) - v(s)\|_\mathcal{V} \\
&\le 3\|\ell\|_{L^\infty(0,T;\mathcal{V}^*)}\|v(t) - v(s)\|_\mathcal{V} + \|v\|_{L^\infty(0,T;\mathcal{V})}\|\dot{\ell}\|_{L^\infty(0,T;\mathcal{V}^*)}|t - s|,
\end{aligned}$$

and since v is absolutely continuous w.r.t. the norm on $\mathcal{V}$, this implies absolute continuity of H. Next, let $I : [0, T] \to \mathbb{R}$ be defined by $I(t) := \mathcal{F}(v(t))$ and let $s, t \in [0, T]$ be arbitrary. Since $\mathcal{F} \in C^1(\mathcal{Z}, \mathcal{V}^*)$, it holds that

$$\begin{aligned}
|I(s) - I(t)| &= |\int_0^1 \frac{\mathrm{d}}{\mathrm{d}h}\mathcal{F}(v(t) + h(v(s) - v(t)))\,\mathrm{d}h| \\
&= |\int_0^1 \mathrm{D}\mathcal{F}(v(t) + h(v(s) - v(t)))[v(s) - v(t)]\,\mathrm{d}h \\
&\le \int_0^1 \|\mathrm{D}\mathcal{F}(v(t) + h(v(s) - v(t)))\|_{\mathcal{V}^*}\,\mathrm{d}h\,\|v(s) - v(t)\|_\mathcal{V}.
\end{aligned} \tag{D.8}$$

We now estimate $\mathrm{D}\mathcal{F}(w)$ for arbitrary $w \in \mathcal{Z}$ by means of (2.4.4b) as follows:

$$\begin{aligned}
\|\mathrm{D}\mathcal{F}(w)\|_{\mathcal{V}^*} &\le \|\mathrm{D}\mathcal{F}(0)\|_{\mathcal{V}^*} + \int_0^1 \|\mathrm{D}_z^2\mathcal{F}(hw)[w]\|_{\mathcal{V}^*}\,\mathrm{d}h \\
&\le \|\mathrm{D}\mathcal{F}(0)\|_{\mathcal{V}^*} + \int_0^1 C(1 + |h|^q\|w\|_\mathcal{Z}^q)\,\mathrm{d}h\,\|w\|_\mathcal{V} \\
&\le \|\mathrm{D}\mathcal{F}(0)\|_{\mathcal{V}^*} + C(1 + \|w\|_\mathcal{Z}^q)\|w\|_\mathcal{V}
\end{aligned}$$

Plugging this estimate into (D.8), we obtain

$$\begin{aligned}|I(s)-I(t)| \leq& \|\mathrm{D}\mathcal{F}(0)\|_{\mathcal{V}^*}\\ &+ C(1+(\|v(t)\|_{\mathcal{Z}}+\|v(s)-v(t)\|_{\mathcal{Z}})^q)(\|v(t)\|_{\mathcal{V}}+\|v(s)-v(t)\|_{\mathcal{V}})\|v(s)-v(t)\|_{\mathcal{V}}\\ \leq& C(1+\|v\|^{q+1}_{L^\infty(0,T;\mathcal{Z})})\|v(s)-v(t)\|_{\mathcal{V}},\end{aligned}$$

so that I is absolutely continuous as well. Finally, we define $J:[0,T]\to\mathbb{R}$ by $J(t):=\frac{1}{2}\langle Av(t),v(t)\rangle_{\mathcal{V}}$. Let us first note that J defines an element in $L^1(0,T;\mathbb{R})$, since we may conclude that $v\in H^1(0,T;\mathcal{V})$ and $Av\in L^2(0,T;\mathcal{V}^*)$ from (D.1). Indeed, the first follows from the boundedness of $\int_0^T \mathcal{R}_\varepsilon(\dot v(t))\,\mathrm{d}t$. In order to show the second, for $t\in[0,T]$, we choose $\eta(t)\in\partial\mathcal{R}(0)\subset\mathcal{V}^*$ such that

$$\begin{aligned}\|-\mathrm{D}_z\mathcal{I}(\ell(t),v(t))-\eta(t)\|^2_{\mathcal{V}^*} &= \mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(t),v(t)),\partial\mathcal{R}(0))\\ &= 2\varepsilon\mathcal{R}^*_\varepsilon(-\mathrm{D}_z\mathcal{I}(\ell(t),v(t))),\end{aligned}$$

where the second identity holds true thanks to the characterization of $\mathcal{R}^*_\varepsilon$ from Lemma 2.4.4. Now, the boundedness of $\partial\mathcal{R}(0)$ in $\mathcal{V}^*$ allows us to estimate

$$\begin{aligned}\int_0^T \|-\mathrm{D}_z\mathcal{I}(\ell(t),v(t))\|^2_{\mathcal{V}^*}\,\mathrm{d}t &\leq C\int_0^T \|-\mathrm{D}_z\mathcal{I}(\ell(t),v(t))-\eta(t)\|^2_{\mathcal{V}^*}+\|\eta(t)\|^2_{\mathcal{V}^*}\,\mathrm{d}t\\ &\leq 2C\varepsilon\int_0^T \mathcal{R}^*_\varepsilon(-\mathrm{D}_z\mathcal{I}(\ell(t),v(t)))\,\mathrm{d}t + CT\sup\{\|\eta\|^2_{\mathcal{V}^*}\,|\,\eta\in\partial\mathcal{R}(0)\}\\ &< \infty.\end{aligned}$$

The estimates for $\mathrm{D}_z\mathcal{I}(\ell,v)$ can be translated into estimates for Av as follows: Since $\mathrm{D}\mathcal{F}:\mathcal{Z}\to\mathcal{V}^*$ is presupposed to be weakly continuous, it is also bounded with respect to the norm on $\mathcal{V}^*$ on bounded subsets of $\mathcal{Z}$. Combining this with $\|\ell\|_{L^\infty(0,T;\mathcal{V}^*)}<\infty$, we find for all $t\in[0,T]$:

$$\begin{aligned}\|Av(t)\|_{\mathcal{V}^*} &\leq \|Av(t)-\mathrm{D}_z\mathcal{I}(\ell(t),v(t))\|_{\mathcal{V}^*}+\|\mathrm{D}_z\mathcal{I}(\ell(t),v(t))\|_{\mathcal{V}^*}\\ &\leq \|\mathrm{D}\mathcal{F}(v(t))+\ell(t)\|_{\mathcal{V}^*}+\|\mathrm{D}_z\mathcal{I}(\ell(t),v(t))\|_{\mathcal{V}^*}\\ &\leq \|\mathrm{D}_z\mathcal{I}(\ell(t),v(t))\|_{\mathcal{V}^*}+C.\end{aligned}$$

Thus, $J\in L^1((0,T);\mathbb{R})$. Next, we show that $J\in\mathrm{AC}([0,T];\mathbb{R})=W^{1,1}((0,T);\mathbb{R})$ with $\frac{\mathrm{d}}{\mathrm{d}t}J(t)=\langle Av(t),\dot v(t)\rangle_{\mathcal{V}}$ by proving for test functions $\phi\in C_0^\infty((0,T);\mathbb{R})$ that

$$-\int_0^T J(t)\dot\phi(t)\,\mathrm{d}t=\int_0^T\langle Av(t),\dot v(t)\rangle_{\mathcal{V}}\phi(t)\,\mathrm{d}t.$$

To this end, for fixed $\phi\in C_0^\infty((0,T);\mathbb{R})$ and $h<h_0:=\mathrm{dist}(\mathrm{supp}(\phi),\{0,T\})$, we define

$$D_hJ(t):=\frac{1}{2h}\Big(\langle Av(t+h),v(t+h)\rangle_{\mathcal{V}}-\langle Av(t),v(t)\rangle_{\mathcal{V}}\Big)\quad\text{and}\quad DJ(t):=\langle Av(t),\dot v(t)\rangle_{\mathcal{V}},$$

where we consider the constant continuation of v to the interval $[0,T+h]$, and claim that

$$-\int_0^T J(t)\dot\phi(t)\,\mathrm{d}t \overset{(i)}{=} \lim_{h\to0}\int_0^T D_hJ(t)\phi(t)\,\mathrm{d}t \overset{(ii)}{=} \int_0^T DJ(t)\phi(t)\,\mathrm{d}t. \tag{D.9}$$

Proof of (i): Using the constant continuation of ϕ to the interval $[-h, T+h]$, it holds that

$$\begin{aligned}
\lim_{h\to0}\int_0^T D_hJ(t)\phi(t)\,\mathrm{d}t &= \lim_{h\to0}\int_0^T \frac{1}{h}(J(t+h)\phi(t) - J(t)\phi(t))\,\mathrm{d}t \\
&= \lim_{h\to0}\int_h^{T+h} \frac{1}{h}(J(\tau)\phi(\tau-h) - J(\tau-h)\phi(\tau-h))\,\mathrm{d}\tau \\
&= \lim_{h\to0}\int_h^{T+h} \frac{1}{h}\Big(J(\tau)(\phi(\tau-h)-\phi(\tau)) + J(\tau)\phi(\tau) - J(\tau-h)\phi(\tau-h)\Big)\,\mathrm{d}\tau \\
&= \lim_{h\to0}\Big(-\int_h^{T+h} J(\tau)\frac{\phi(\tau)-\phi(\tau-h)}{h}\,\mathrm{d}\tau + \frac{1}{h}\int_h^{T+h} J(\tau)\phi(\tau)\,\mathrm{d}\tau - \frac{1}{h}\int_0^T J(t)\phi(t)\,\mathrm{d}t\Big) \\
&= \lim_{h\to0}\Big(-\int_h^{T+h} J(\tau)\frac{\phi(\tau)-\phi(\tau-h)}{h}\,\mathrm{d}\tau + \frac{1}{h}\int_0^{h} J(\tau)\phi(\tau)\,\mathrm{d}\tau - \frac{1}{h}\int_T^{T+h} J(t)\phi(t)\,\mathrm{d}t\Big) \\
&= \lim_{h\to0}\Big(-\int_0^{T} J(\tau)\frac{\phi(\tau)-\phi(\tau-h)}{h}\,\mathrm{d}\tau\Big),
\end{aligned}$$

where the last equality holds due to the fact that h is chosen so small that $\phi(t) = 0$ for $t \in (-h, h) \cup (T-h, T+h)$. Now, the remaining integrand on the right hand side converges pointwisely to $J(t)\dot{\phi}(t)$ and is uniformly bounded by the integrable function $J(\cdot)C\|\dot{\phi}\|_{L^\infty((0,T);\mathbb{R})}$ for a constant $C > 0$. Thus, we infer the validity of (i) by means of the Theorem of Dominated Convergence.

Proof of (ii): Application of assertions (2) and (3) from Lemma E.2 yields

$$\begin{aligned}
\lim_{h\to0}\int_0^T D_hJ(t)\phi(t)\,\mathrm{d}t &= \lim_{h\to0}\int_0^T \frac{1}{h}\Big(\langle Av(t+h), v(t+h)\rangle_{\mathcal{V}} - \langle Av(t), v(t)\rangle_{\mathcal{V}}\Big)\phi(t)\,\mathrm{d}t \\
= \lim_{h\to0}\Big(&\int_0^T \langle Av(t+h), \phi(t)\frac{v(t+h)-v(t)}{h}\rangle_{\mathcal{V}}\,\mathrm{d}t + \int_0^T \langle \frac{A(v(t+h)-v(t))}{h}, \phi(t)v(t)\rangle_{\mathcal{V}}\,\mathrm{d}t\Big) \\
= \lim_{h\to0}\Big(&\int_0^T \langle \underbrace{Av(t+h)}_{\xrightarrow{(3)} Av}, \underbrace{L_hv(t)}_{\xrightarrow{(2)} \phi\dot{v}}\rangle_{\mathcal{V}}\,\mathrm{d}t + \int_0^T \langle Av(t), \underbrace{L_hv(t)}_{\xrightarrow{(2)} \phi\dot{v}}\rangle_{\mathcal{V}}\,\mathrm{d}t\Big) = \int_0^T DJ(t)\phi(t)\,\mathrm{d}t,
\end{aligned}$$

where we have used the assumption that A is self-adjoint in the third identity, and L_h is the difference quotient operator from Lemma E.2.

Having thus verified that (D.9) holds true, the map $t \mapsto \mathcal{I}(\ell(t), v(t))$ can be written as the sum

$$\mathcal{I}(\ell(t), v(t)) = J(t) + I(t) - H(t)$$

of three absolutely continuous functions and is therefore absolutely continuous itself.

In order to show the chain rule, let $t \in [0, T]$ be a point such that the derivatives $\dot{\ell}(t)$, $\dot{v}(t)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}(\ell(t), v(t))$ exist. First, we note that due to the continuity of $\mathrm{D}\mathcal{F} : \mathcal{Z} \to \mathcal{V}^*$ and using (3) of Lemma E.2, we have that

$$\begin{aligned}
\mathrm{D}_z\mathcal{I}(\ell(t+h), v(t+h)) &= Av(t+h) + \mathrm{D}\mathcal{F}(v(t+h)) + \ell(t+h) \\
&\xrightarrow{h\to0} Av(t) + \mathrm{D}\mathcal{F}(v(t)) + \ell(t) = \mathrm{D}_z\mathcal{I}(\ell(t), v(t)) \text{ in } \mathcal{V}^*,
\end{aligned}$$

and by the same arguments, we obtain that

$$\begin{aligned}\mathrm{D}_z\mathcal{I}(\ell(t-h),v(t-h)) &= Av(t-h)+\mathrm{D}\mathcal{F}(v(t-h))+\ell(t-h)\\ &\xrightarrow{h\to 0} Av(t)+\mathrm{D}\mathcal{F}(v(t))+\ell(t)=\mathrm{D}_z\mathcal{I}(\ell(t),v(t)) \text{ in } \mathcal{V}^*.\end{aligned}$$

Furthermore, since $v(t\pm h)\xrightarrow{h\to 0} v(t)$ strongly in $\mathcal{X}$, the boundedness (2.4.7) of $\mathcal{R}$ implies that

$$0=\lim_{h\to 0} c\|v(t)-v(t\pm h)\|_{\mathcal{X}} \le \lim_{h\to 0}\mathcal{R}(v(t)-v(t\pm h)) \le \lim_{h\to 0} C\|v(t)-v(t\pm h)\|_{\mathcal{X}}=0.$$

Therefore, using (2.4.13), we obtain for every $h>0$ the estimate

$$\begin{aligned}&\frac{1}{h}\Big(\mathcal{I}(\ell(t),v(t))-\mathcal{I}(\ell(t-h),v(t-h))\Big)\\ &\quad=\frac{1}{h}\Big(\mathcal{I}(\ell(t),v(t))-\mathcal{I}(\ell(t-h),v(t))\Big)+\frac{1}{h}\Big(\mathcal{I}(\ell(t-h),v(t))-\mathcal{I}(\ell(t-h),v(t-h))\Big)\\ &\quad\ge\frac{1}{h}\langle\ell(t)-\ell(t-h),v(t)\rangle_{\mathcal{V}}+\langle\mathrm{D}_z\mathcal{I}(\ell(t-h),v(t-h)),\frac{v(t)-v(t-h)}{h}\rangle_{\mathcal{V}}\\ &\qquad\qquad -M_\rho\|\frac{v(t)-v(t-h)}{h}\|_{\mathcal{V}}\mathcal{R}(v(t)-v(t-h))\\ &\quad\to\mathcal{P}(t,v(t))+\langle\mathrm{D}_z\mathcal{I}(\ell(t),v(t)),\dot v(t)\rangle_{\mathcal{V}}.\end{aligned}\tag{D.10}$$

Analoguously, an approximation of t from above by a sequence $t+h_n$, where $h_n\searrow 0$, gives the opposite estimate: We estimate as in (D.10) but now divide by $-h_n<0$

$$\begin{aligned}&\frac{1}{-h}\Big(\mathcal{I}(\ell(t),v(t))-\mathcal{I}(\ell(t+h),v(t+h))\Big)\\ &\quad=\frac{1}{-h}\Big(\mathcal{I}(\ell(t),v(t))-\mathcal{I}(\ell(t+h),v(t))\Big)+\frac{1}{-h}\Big(\mathcal{I}(\ell(t+h),v(t))-\mathcal{I}(\ell(t+h),v(t+h))\Big)\\ &\quad=\frac{1}{h}\Big(\mathcal{I}(\ell(t+h),v(t))-\mathcal{I}(\ell(t),v(t))\Big)+\frac{1}{-h}\Big(\mathcal{I}(\ell(t+h),v(t))-\mathcal{I}(\ell(t+h),v(t+h))\Big)\\ &\quad\le\frac{1}{h}\langle\ell(t+h)-\ell(t),v(t)\rangle_{\mathcal{V}}+\langle\mathrm{D}_z\mathcal{I}(\ell(t+h),v(t+h)),\frac{v(t)-v(t+h)}{-h}\rangle_{\mathcal{V}}\\ &\qquad\qquad +M_\rho\|\frac{v(t)-v(t+h)}{h}\|_{\mathcal{V}}\mathcal{R}(v(t)-v(t+h))\\ &\quad\to\mathcal{P}(t,v(t))+\langle\mathrm{D}_z\mathcal{I}(\ell(t),v(t)),\dot v(t)\rangle_{\mathcal{V}}.\end{aligned}$$

Proof of ($\mathcal{I}_5$): Let $t\in[0,T]$ and $(v_n)_{n\in\mathbb{N}}\subset\mathcal{V}$ be a sequence as in ($\mathcal{I}_5$). From the convergence $\mathcal{I}(\ell(t),v_n)\to\mathcal{I}$, we infer that $\sup_{n\in\mathbb{N}}|\mathcal{I}(\ell(t),v_n)|<\infty$, so that $\sup_{n\in\mathbb{N}}\|v_n\|_{\mathcal{Z}}<\infty$ by (D.6). Thus, there exist a subsequence and a weak limit in $\mathcal{Z}$, which must coincide with v, i.e. $v_{n_k}\xrightarrow{k\to\infty} v$ in $\mathcal{Z}$ and $v_{n_k}\xrightarrow{k\to\infty} v$ in $\mathcal{V}$. By repeating the argument for every subsequence, we conclude that the whole sequence converges weakly in $\mathcal{Z}$ and strongly in $\mathcal{V}$. We show the convergence of the powers by

$$p=\lim_{n\to\infty}\partial_t\mathcal{I}(\ell(t),v_n)=\lim_{n\to\infty}-\langle\dot\ell(t),v_n\rangle_{\mathcal{V}}=-\langle\dot\ell(t),v\rangle_{\mathcal{V}}=\partial_t\mathcal{I}(\ell(t),v).$$

Due to the weak convergence $v_n \rightharpoonup v$ in $\mathcal{Z}$, we have that

$$\mathrm{D}_z\mathcal{I}(\ell(t),v_n) = Av_n + \mathrm{D}\mathcal{F}(v_n) + \ell(t) \rightharpoonup Av + \mathrm{D}\mathcal{F}(v) + \ell(t) = \mathrm{D}_z\mathcal{I}(\ell(t),v) \text{ in } \mathcal{Z}^*,$$

and since the left hand side converges weakly to ξ in $\mathcal{V}^*$, we conclude that we have $\xi = \mathrm{D}_z\mathcal{I}(\ell(t),v)$ as claimed. In order to show the convergence of the energies, we observe that $v_n \to v$ strongly in $\mathcal{V}$, and that the boundedness of the sequence $(v_n)_{n\in\mathbb{N}}$ in $\mathcal{Z}$ makes the estimate (2.4.13) available, which leads to

$$\mathcal{I}(\ell(t),v) \geq \mathcal{I}(\ell(t),v_n) + \langle \underbrace{\mathrm{D}_z\mathcal{I}(\ell(t),v_n)}_{\rightharpoonup \xi \text{ in } \mathcal{V}^*}, \underbrace{v - v_n}_{\to 0 \text{ in } \mathcal{V}} \rangle_{\mathcal{V}} + \frac{\alpha}{4}\|v - v_n\|_{\mathcal{Z}}^2 + M \underbrace{\|v_n - v\|_{\mathcal{V}}\mathcal{R}(v_n - v)}_{\to 0}$$

and thus, since we already established lower semicontinuity of $\mathcal{I}$ w.r.t. the norm topology on $\mathcal{V}$,

$$\begin{aligned}
\mathcal{I}(\ell(t),v) &\geq \liminf_{n\to\infty}\Big(\mathcal{I}(\ell(t),v_n) + \frac{\alpha}{4}\|v_n - v\|_{\mathcal{Z}}^2\Big) \\
&\geq \liminf_{n\to\infty}\mathcal{I}(\ell(t),v_n) \\
&\geq \mathcal{I}(\ell(t),v),
\end{aligned}$$

whereby we conclude $\mathcal{I}(\ell(t),v) = \liminf_{n\to\infty}\mathcal{I}(\ell(t),v_n) = \lim_{n\to\infty}\mathcal{I}(\ell(t),v_n)$ as well as $\|v_n - v\|_{\mathcal{Z}} \to 0$.

Finally, (D.2) follows from $(\mathcal{I}_3)$ by means of the Gronwall inequality via

$$\mathcal{I}(\ell(t),z) \leq \mathcal{I}(\ell(s),z) + \int_s^t \partial_t\mathcal{I}(\ell(\sigma),z)\,\mathrm{d}\sigma \leq \mathcal{I}(\ell(s),z) + \int_s^t C_1\mathcal{I}(\ell(\sigma),z)\,\mathrm{d}\sigma,$$

and thus

$$\mathcal{I}(\ell(t),z) \leq \mathcal{I}(\ell(s),z)e^{C_1(t-s)}.$$

■

The following Lemma is taken directly from [KZ21, Lemma 2.2]

Lemma D.3. *Let $c_0 := \frac{c_{\mathcal{Z}}^2}{2}(1 + \|\ell\|_{L^\infty(0,T;\mathcal{V}^*)}^2)$, where $c_{\mathcal{Z}}$ is the embedding constant for $\mathcal{Z} \subset \mathcal{V}$. Then for every $t \in [0,T]$ and $v \in \mathcal{Z}$ we have*

$$\mathcal{I}(\ell(t),v) + c_0 \geq c_{\mathcal{Z}}\|z\|_{\mathcal{Z}} \geq \|v\|_{\mathcal{V}}. \tag{D.11}$$

Furthermore, the following product estimate is valid: Let $\{a_k; 1 \leq k \leq N\} \subset \mathbb{R}_{\geq 0}$ and $c > 0$. Then

$$\prod_{k=1}^{N}(1 + ca_k) \leq \exp\left(c\sum_{k=1}^{N} a_k\right).$$

As a consequence, let $c > 0$, $\ell \in \mathrm{BV}([0,T];\mathcal{V}^*)$ and let $0 \leq t_0 < t_1 < \cdots < t_N \leq T$ be an arbitrary partition of $[0,T]$. Then

$$\prod_{k=1}^{N}(1 + c\|\ell(t_k) - \ell(t_{k-1})\|_{\mathcal{V}^*}) \leq \exp(c\,\mathrm{Var}_{\mathcal{V}^*}(\ell,[t_0,t_N])). \tag{D.12}$$

Appendix E

Convergence of the load term

Lemma E.1.

(i) *Let $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ and $\hat{t} \in W^{1,\infty}(0,S)$ with $\hat{t}(0)=0, \hat{t}(S)=T$, and $\hat{t}'(s) \geq 0$ f.a.a. $s \in (0,S)$ be given. Then $\hat{\ell} := \ell \circ \hat{t} \in W^{1,\infty}(0,S;\mathcal{V}^*)$ with*

$$\frac{\mathrm{d}}{\mathrm{d}s}(\ell \circ \hat{t})(s) = \begin{cases} \dot{\ell}(\hat{t}(s))\dot{\hat{t}}(s), & \text{if } \ell \text{ is differentiable in } \hat{t}(s) \\ 0, & \text{if } \ell \text{ is not differentiable in } \hat{t}(s) \end{cases} \quad \text{f.a.a. } s \in [0,S], \tag{E.1}$$

and $\|\dot{\hat{\ell}}\|_{L^\infty(0,S;\mathcal{V}^*)} \leq \|\dot{\hat{t}}\|_{L^\infty(0,S)}\|\dot{\ell}\|_{L^\infty(0,T;\mathcal{V}^*)}$.

(ii) *Let further $(\hat{t}_n)_{n\in\mathbb{N}} \subset W^{1,\infty}(0,S;\mathbb{R})$ be a sequence with $\hat{t}_n \overset{*}{\rightharpoonup} \hat{t}$ in $W^{1,\infty}(0,S;\mathbb{R})$ and with $\hat{t}_n(0)=0$, and $\hat{t}_n'(s) \geq 0$ f.a.a. $s \in (0,S)$ and $n \in \mathbb{N}$. Then*

$$\ell \circ \hat{t}_n \overset{*}{\rightharpoonup} \ell \circ \hat{t} \text{ in } W^{1,\infty}(0,S;\mathcal{V}^*).$$

(iii) *If $(\ell_n)_{n\in\mathbb{N}} \subset W^{1,\infty}(0,T;\mathcal{V}^*)$ with $\ell_n \overset{*}{\rightharpoonup} \ell$ in $W^{1,\infty}(0,T;\mathcal{V}^*)$, then $\ell_n(t) \rightharpoonup \ell(t)$ weakly in $\mathcal{V}^*$ for all $t \in [0,T]$.*

(iv) *Let now $(\ell_n)_{n\in\mathbb{N}} \subset W^{1,\infty}(0,T;\mathcal{V}^*)$ and $(\hat{t}_n)_{n\in\mathbb{N}} \subset W^{1,\infty}(0,S;\mathbb{R})$ be sequences such that $\ell_n \overset{*}{\rightharpoonup} \ell$ in $W^{1,\infty}(0,T;\mathcal{V}^*)$ and $\hat{t}_n \overset{*}{\rightharpoonup} \hat{t}$ in $W^{1,\infty}(0,S;\mathbb{R})$. Then*

$$\ell_n \circ \hat{t}_n \overset{*}{\rightharpoonup} \ell \circ \hat{t} \text{ in } W^{1,\infty}(0,S;\mathcal{V}^*) \tag{E.2}$$

Proof. *(i)* Let us first prove the chain rule (E.1) in analogy to the finite dimensional case. Let $s_0 \in [0,S]$ be such that $\hat{t}$ is differentiable in s_0. According to [CH98, Thm. 1.4.35], this is the case almost everywhere in $[0,S]$. First assume that ℓ is differentiable in $\hat{t}(s_0)$. For $t, t_0 \in [0,T]$, we define

$$D(t,t_0) := \begin{cases} \frac{\ell(t)-\ell(t_0)}{t-t_0}, & \text{if } t \neq t_0, \\ \dot{\ell}(t_0), & \text{if } t = t_0. \end{cases}$$

Then, for those points t_0 in which ℓ is differentiable, the map $D(\cdot, t_0) : [0,T] \to \mathcal{V}^*$ is norm-continuous in t_0. Therefore, it holds in $\mathcal{V}^*$ that

$$\begin{aligned}
\lim_{s\to s_0} \frac{\ell(\hat{t}(s)) - \ell(\hat{t}(s_0))}{s - s_0} &= \lim_{s\to s_0}\left(\frac{D(\hat{t}(s), \hat{t}(s_0)) \cdot (\hat{t}(s) - \hat{t}(s_0))}{s - s_0}\right) \\
&= \lim_{s\to s_0}\Big(D(\hat{t}(s), \hat{t}(s_0))\Big) \cdot \lim_{s\to s_0}\left(\frac{\hat{t}(s) - \hat{t}(s_0)}{s - s_0}\right) \\
&= D(\hat{t}(s_0), \hat{t}(s_0))\dot{\hat{t}}(s_0) \\
&= \dot{\ell}(\hat{t}(s_0))\dot{\hat{t}}(s_0),
\end{aligned}$$

where the second to last equation follows from the fact that $\hat{t}$ is continuous, and we infer the first of (E.1).

We now turn to the case that ℓ is not differentiable in $\hat{t}(s_0)$. We denote by $M \subset [0,S]$ the set of all points $s \in [0,S]$ such that $\hat{t}$ is differentiable in s, but ℓ is not differentiable in $\hat{t}(s)$. Again from [CH98, Thm. 1.4.35], we infer that there exists a set $N \subset [0,T]$ of $\mathscr{L}^1$-measure 0 with $\hat{t}(M) \subseteq N$, so that $\mathscr{L}^1(\hat{t}(M)) = 0$. Let us now assume that there exists an interval $I \subseteq M$ such that $\dot{\hat{t}}(s) > 0$ on I. A change of coordinates then yields

$$0 = \mathscr{L}^1(\hat{t}(I)) = \int_{\hat{t}(I)} 1 \,\mathrm{d}t = \int_I \dot{\hat{t}}(s)\mathrm{d}s > 0,$$

which is a contradiction. Analoguously, there cannot exist an interval $I \subseteq M$ such that $\dot{\hat{t}}(s) < 0$ on I. Thus, $\hat{t}$ is constant on all connected components of M, and we infer the second of (E.1). Then, for almost all $s \in (0,S)$, we have

$$\|\dot{\hat{\ell}}(s)\|_{\mathcal{V}^*} = \|\dot{\ell}(\hat{t}(s))\dot{\hat{t}}(s)\|_{\mathcal{V}^*} \le \|\dot{\hat{t}}\|_{L^\infty(0,S)}\|\dot{\ell}\|_{L^\infty(0,T;\mathcal{V}^*)}.$$

(ii) We first use the compact embedding $W^{1,\infty}(0,S) \overset{c}{\hookrightarrow} C([0,S])$ and obtain that $\hat{t}_n \to \hat{t}$ uniformly on $[0,S]$, so that we also find

$$\sup_{s\in[0,S]} \|\ell(t_n(s)) - \ell(t(s))\|_{\mathcal{V}^*} \le \|\dot{\ell}\|_{L^\infty(0,S;\mathcal{V}^*)} \sup_{s\in[0,S]} |t_n(s) - t(s)| \to 0,$$

i.e., $\ell \circ t_n \to \ell \circ t$ uniformly w.r.t. the norm on $\mathcal{V}^*$. Since $(\ell \circ t_n)_{n\in\mathbb{N}}$ is a bounded sequence in $W^{1,\infty}(0,S;\mathcal{V}^*)$, we also know that there exist $\tilde{\ell} \in W^{1,\infty}(0,S;\mathcal{V}^*)$ and a not relabeled subsequence such that $\ell \circ t_n \overset{*}{\rightharpoonup} \tilde{\ell}$ weakly in $W^{1,\infty}(0,S;\mathcal{V}^*)$ and the weak limit coincides with $\ell \circ \hat{t}$, so that we have $\ell \circ t_n \overset{*}{\rightharpoonup} \ell \circ \hat{t}$ in $W^{1,\infty}(0,S;\mathcal{V}^*)$ as claimed.

(iii) We in fact show that the following stronger implication holds true for all sequences $(f_n)_{n\in\mathbb{N}} \subset W^{1,\infty}(0,T;\mathcal{V}^*)$:

$$\begin{aligned}
&\underline{\text{if}} \quad f_n \overset{*}{\rightharpoonup} f \text{ in } W^{1,\infty}(0,T;\mathcal{V}^*),\\
&\underline{\text{then}} \quad f_n(t) \to f(t) \text{ strongly in } \mathcal{Z}^* \text{ for all } t \in [0,T].
\end{aligned} \tag{E.3}$$

To this end, we argue as follows: Let $f_n \overset{*}{\rightharpoonup} f$ in $W^{1,\infty}(0,T;\mathcal{V}^*)$, then in particular $f_n \overset{*}{\rightharpoonup} f$ in $L^\infty(0,T;\mathcal{V}^*)$, i.e.,

$$\begin{aligned}
\forall \varphi \in L^1(0,T;\mathcal{V}): &\quad \int_0^T \langle f_n(t) - f(t), \varphi(t)\rangle_{\mathcal{V}}\,\mathrm{d}t \to 0, \textit{ i.e.,}\\
\forall \varphi \in L^\infty(0,T;\mathcal{V}): &\quad \int_0^T \langle f_n(t) - f(t), \varphi(t)\rangle_{\mathcal{V}}\,\mathrm{d}t \to 0, \textit{ i.e.,}\\
\forall \varphi \in L^\infty(0,T;\mathcal{Z}): &\quad \int_0^T \langle f_n(t) - f(t), \varphi(t)\rangle_{\mathcal{V}}\,\mathrm{d}t \to 0,
\end{aligned}$$

and thus, $f_n \rightharpoonup f$ weakly in $L^1(0,T;\mathcal{Z}^*)$. What is more, since $\sup_{n\in\mathbb{N}} \|f_n(t)\|_{\mathcal{V}^*} < \infty$ almost everwhere, there is a function, denoted by $\tilde{f}$, such that $f_n(t) \rightharpoonup \tilde{f}(t)$ weakly in $\mathcal{V}^*$ and $f_n(t) \to \tilde{f}(t)$ strongly in $\mathcal{Z}^*$ almost everwhere. Now, the boundedness of the sequence $(f_n)_{n\in\mathbb{N}}$ in $L^\infty(0,T;\mathcal{V}^*)$ implies the boundedness of the sequence $(f_n)_{n\in\mathbb{N}}$ in $L^\infty(0,T;\mathcal{Z}^*)$. Thus, the Theorem of Dominated Convergence is applicable, and we infer that $f_n \to \tilde{f}$ strongly in $L^1(0,T;\mathcal{V}^*)$, which is why f and $\tilde{f}$ have to coincide almost everwhere, and we find that $f_n(t) \rightharpoonup f(t)$ weakly in $\mathcal{V}^*$ almost everywhere.

Since both f_n and f are continuous w.r.t. $\|\cdot\|_{\mathcal{V}^*}$, this holds everywhere in $[0,T]$: Let $t \in [0,T]$ be arbitrary and $(t_k)_{k\in\mathbb{N}} \subset (0,T)$ be a sequence with $t_k \nearrow t$ and such that we have for all $k \in \mathbb{N}$ that $f_n(t_k) \rightharpoonup f(t_k)$ weakly in $\mathcal{V}^*$. Let us denote $\rho := \sup_{n\in\mathbb{N}} \|f_n\|_{W^{1,\infty}(0,T;\mathcal{V}^*)}$. Let further $v \in \mathcal{V}$ and $\delta > 0$ be arbitrary, and $n \in \mathbb{N}$ be arbitrary, but fixed, then it holds for all $k \in \mathbb{N}$

$$\begin{aligned}
|\langle f(t) - f_n(t), v\rangle_{\mathcal{V}}| &\le |\langle f(t) - f(t_k), v\rangle_{\mathcal{V}}| + |\langle f(t_k) - f_n(t_k), v\rangle_{\mathcal{V}}| + |\langle f_n(t_k) - f_n(t), v\rangle_{\mathcal{V}}|\\
&\le |\langle f(t) - f(t_k), v\rangle_{\mathcal{V}}| + |\langle f(t_k) - f_n(t_k), v\rangle_{\mathcal{V}}| + \int_{t_k}^t \|\dot{f}_n(\tau)\|_{\mathcal{V}^*}\mathrm{d}\tau \cdot \|v\|_{\mathcal{V}}\\
&\le |\langle f(t) - f(t_k), v\rangle_{\mathcal{V}}| + |\langle f(t_k) - f_n(t_k), v\rangle_{\mathcal{V}}| + (t - t_k)\rho\|v\|_{\mathcal{V}}.
\end{aligned}$$

We can now choose $K \in \mathbb{N}$ so big that

$$|\langle f(t) - f(t_K), v\rangle_{\mathcal{V}}| < \tfrac{\delta}{3} \quad \text{and} \quad (t - t_K)\rho\|v\|_{\mathcal{V}} < \tfrac{\delta}{3}.$$

Finally, we find $N \in \mathbb{N}$ such that for all $n \ge N$, it holds $|\langle f(t_K) - f_n(t_K), v\rangle_{\mathcal{V}}| < \frac{\delta}{3}$, which implies that $f_n(t) \rightharpoonup f(t)$ weakly in $\mathcal{V}^*$, and (E.3) is proven.

We can now prove *(iv)* as follows: From the estimate

$$\sup_{n\in\mathbb{N}} \|\ell_n \circ \hat{t}_n\|_{W^{1,\infty}(0,S;\mathcal{V}^*)} \le \sup_{n\in\mathbb{N}} \{\|\ell_n\|_{L^\infty(0,T;\mathcal{V}^*)}, \|\dot{\ell}_n\|_{L^\infty(0,T;\mathcal{V}^*)} \|\dot{\hat{t}}_n\|_{L^\infty(0,S)}\} < \infty,$$

we infer the existence of an element $\tilde{\ell} \in W^{1,\infty}(0,S^*;\mathcal{V}^*)$ such that $\ell_n \circ \hat{t}_n \overset{*}{\rightharpoonup} \tilde{\ell}$ in $W^{1,\infty}(0,S;\mathcal{V}^*)$ and it remains to show that $\tilde{\ell} = \ell \circ \hat{t}$. Indeed, from $\ell_n \overset{*}{\rightharpoonup} \ell$ in $W^{1,\infty}(0,T;\mathcal{V}^*)$, it follows with (E.3) that $\ell_n(t) \to \ell(t)$ strongly in $\mathcal{Z}^*$ almost everwhere and from $\hat{t}_n \overset{*}{\rightharpoonup} \hat{t}$ in $W^{1,\infty}(0,S)$, it follows that $\hat{t}_n \to \hat{t}$ in $C([0,S])$. Now, let $s \in [0,S]$, let $\varepsilon > 0$ be arbitrary and $N \in \mathbb{N}$ sufficiently big, such that for all $n \ge N$, it holds that

$$\|\ell_n(\hat{t}(s)) - \ell(\hat{t}(s))\|_{\mathcal{Z}^*} < \frac{\varepsilon}{2} \quad \text{and} \quad \|\hat{t}_n - \hat{t}\|_{L^\infty(0,S)} < \frac{\varepsilon}{2\sup_{n\in\mathbb{N}} \|\dot{\ell}_n\|_{L^\infty(0,T;\mathcal{Z}^*)}}.$$

Then, for all $n \geq N$ and $s \in [0, S]$, it holds that

$$\begin{aligned}\|\ell_n(\hat{t}_n(s)) - \ell(\hat{t}(s))\|_{\mathcal{Z}^*} &\leq \|\ell_n(\hat{t}_n(s)) - \ell_n(\hat{t}(s))\|_{\mathcal{Z}^*} + \|\ell_n(\hat{t}(s)) - \ell(\hat{t}(s))\|_{\mathcal{Z}^*} \\ &\leq \|\dot{\ell}_n\|_{L^\infty(0,T;\mathcal{Z}^*)}|\hat{t}_n(s) - \hat{t}(s)| + \|\ell_n(\hat{t}(s)) - \ell(\hat{t}(s))\|_{\mathcal{Z}^*} \\ &< \varepsilon,\end{aligned}$$

which implies the pointwise convergence $\ell_n(\hat{t}_n(s)) \to \ell(\hat{t}(s))$ in $\mathcal{Z}^*$. However, from $\ell_n \circ t_n \overset{*}{\rightharpoonup} \tilde{\ell}$ in $W^{1,\infty}(0, S; \mathcal{V}^*)$ it also follows with (E.3) that $\ell_n(\hat{t}_n(s)) \to \tilde{\ell}(s)$ strongly in $\mathcal{Z}^*$, so that $\tilde{\ell}(s) = \ell(\hat{t}(s))$ almost everywhere and (E.2) ensues. ∎

Lemma E.2. *Let $\phi \in C_0^\infty(0, T; \mathbb{R})$ be arbitrary, but fixed. We define the function $L_h : H^1(0, T; \mathcal{V}) \to L^2(0, T; \mathcal{V})$ by*

$$L_h v(t) := \phi(t)\frac{1}{h}(v(t + h) - v(t)),$$

using the constant continuation of v to $[0, T + h]$. Then it holds that

(1) L_h is a well defined, continuous, linear operator such that

$$\forall\, v \in H^1(0, T; \mathcal{V}) : \|L_h v\|_{L^2(0,T;\mathcal{V})} \leq \|\phi\|_{L^\infty(0,T)}\|\dot{v}\|_{L^2(0,T;\mathcal{V})}.$$

(2) For all $v \in H^1(0, T; \mathcal{V})$, it holds that $L_h(v) \xrightarrow{h\to 0} \phi\dot{v}$ strongly in $L^2(0, T; \mathcal{V})$.

(3) For all $w \in L^2(0, T; \mathcal{V}^*)$ and their continuation $\widetilde{w} : \mathbb{R} \to \mathcal{V}^*$ of w to $\mathbb{R}$ by zero, it holds that $\widetilde{w}(\cdot + h) \xrightarrow{h\to 0} \widetilde{w}$ in $L^2(0, T; \mathcal{V}^*)$.

Proof. In order to prove (1), we estimate for $v \in C^\infty([0, T]; \mathcal{V})$

$$\begin{aligned}\|L_h v\|^2_{L^2(0,T;\mathcal{V})} &= \int_0^T |\phi(t)|^2 \|\int_0^1 \dot{v}(t + sh)\,\mathrm{d}s\|^2_{\mathcal{V}}\,\mathrm{d}t \\ &\leq \int_0^T |\phi(t)|^2 \Big(\int_0^1 \|\dot{v}(t + sh)\|_{\mathcal{V}}\,\mathrm{d}s\Big)^2\,\mathrm{d}t \\ &\leq \int_0^T |\phi(t)|^2 \int_0^1 \|\dot{v}(t + sh)\|^2_{\mathcal{V}}\,\mathrm{d}s\,\mathrm{d}t \\ &\leq \int_0^1 \int_{\mathrm{supp}(\phi)} \|\dot{v}(t + sh)\|^2_{\mathcal{V}}\,\mathrm{d}t\,\mathrm{d}s\,\|\phi\|^2_{L^\infty(0,T)} \\ &\leq \|\phi\|^2_{L^\infty(0,T)}\|\dot{v}\|^2_{L^2(0,T;\mathcal{V})}.\end{aligned}$$

Since $C^\infty([0, T]; \mathcal{V})$ is dense in $H^1(0, T; \mathcal{V})$ according to [Emm04, Satz 8.1.9], we infer (1). For the proof of (2), we again assume that $v \in C^\infty([0, T]; \mathcal{V})$ and obtain

$$\|L_h v - \phi\dot{v}\|^2_{L^2(0,T;\mathcal{V})} = \int_0^T |\phi(t)|^2 \|\frac{v(t + h) - v(t)}{h} - \dot{v}\|^2_{\mathcal{V}}\,\mathrm{d}t \xrightarrow{h\to 0} 0,$$

due to the Dominated Convergence Theorem, since the integrand is uniformly bounded by $C\|\phi\|^2_{L^\infty(0,T)}\|\dot{v}\|^2_{L^\infty(0,T;\mathcal{V})}$. Now, let $v \in H^1(0, T; \mathcal{V})$ and $\eta > 0$. Then we choose $\tilde{v} \in C^\infty([0,T];\mathcal{V})$ such that $\|\tilde{v} - v\|_{H^1(0,T;\mathcal{V})} < \frac{\eta}{3\|\phi\|_{L^\infty(0,T)}}$ and $h_1 < h_0$ so small that for all $0 < h < h_1$, it holds that $\|L_h\tilde{v} - \phi\dot{\tilde{v}}\|_{L^2(0,T;\mathcal{V})} < \frac{\eta}{3}$. It follows for all $0 < h < h_1$:

$$
\begin{aligned}
&\|L_h v - \phi\dot{v}\|_{L^2(0,T;\mathcal{V})} \\
&\qquad \le \|L_h v - L_h\tilde{v}\|_{L^2(0,T;\mathcal{V})} + \|L_h\tilde{v} - \phi\dot{\tilde{v}}\|_{L^2(0,T;\mathcal{V})} + \|\phi\dot{\tilde{v}} - \phi\dot{v}\|_{L^2(0,T;\mathcal{V})} \\
&\qquad \le \|\phi\|_{L^\infty(0,T)}\|\dot{\tilde{v}} - \dot{v}\|_{L^2(0,T;\mathcal{V})} + \|L_h\tilde{v} - \phi\dot{\tilde{v}}\|_{L^2(0,T;\mathcal{V})} + \|\phi\|_{L^\infty(0,T)}\|\dot{\tilde{v}} - \dot{v}\|_{L^2(0,T;\mathcal{V})} \\
&\qquad < \eta.
\end{aligned}
$$

For the proof of (3), let $w \in C^\infty([0,T];\mathcal{V}^*)$, then it holds that

$$
\begin{aligned}
&\|\widetilde{w}(\cdot + h) - \widetilde{w}\|^2_{L^2(0,T;\mathcal{V}^*)} \\
&\qquad = \int_0^T \|\widetilde{w}(t+h) - \widetilde{w}(t)\|^2_{\mathcal{V}^*}\,\mathrm{d}t \\
&\qquad \le C\Big(\int_0^{T-h} \|w(t+h) - w(t)\|^2_{\mathcal{V}^*}\,\mathrm{d}t + \int_{T-h}^T \|w(T) - w(t)\|^2_{\mathcal{V}^*}\,\mathrm{d}t\Big) \\
&\qquad \le C\Big(\int_0^{T-h} \|w(t+h) - w(t)\|^2_{\mathcal{V}^*}\,\mathrm{d}t + h\|w\|^2_{L^\infty(0,T;\mathcal{V}^*)}\,\mathrm{d}t\Big) \\
&\qquad \xrightarrow{h\to 0} 0,
\end{aligned}
$$

since the first integrand converges to 0 and is bounded by $2\|w\|^2_{L^\infty(0,T;\mathcal{V}^*)}$. Now, let $w \in L^2(0,T;\mathcal{V}^*)$ and $\eta > 0$. Since $C^\infty([0,T],\mathcal{V}^*)$ is dense in $L^2(0,T;\mathcal{V}^*)$ according to [GP06, Remark 2.2.4], we can choose $w^* \in C^\infty([0,T];\mathcal{V}^*)$ such that $\|w - w^*\|^2_{L^2(0,T;\mathcal{V}^*)} < \frac{\eta}{3}$. We further find $0 < h_1 < h_0$ so small that for all $0 < h < h_1$, it holds $\|\widetilde{w}^*(\cdot + h) - \widetilde{w}^*\|^2_{L^2(0,T;\mathcal{V}^*)} < \frac{\eta}{3}$. This implies for all $0 < h < h_1$ that

$$
\begin{aligned}
&\|\widetilde{w}(\cdot + h) - \widetilde{w}\|^2_{L^2(0,T;\mathcal{V}^*)} \\
&\qquad \le \|\widetilde{w}(\cdot + h) - \widetilde{w}^*(\cdot + h)\|^2_{L^2(0,T;\mathcal{V}^*)} + \|\widetilde{w}^*(\cdot + h) - \widetilde{w}^*\|^2_{L^2(0,T;\mathcal{V}^*)} + \|\widetilde{w}^* - \widetilde{w}\|^2_{L^2(0,T;\mathcal{V}^*)} \\
&\qquad \le \|w - w^*\|^2_{L^2(0,T;\mathcal{V}^*)} + \|\widetilde{w}^*(\cdot + h) - \widetilde{w}^*\|^2_{L^2(0,T;\mathcal{V}^*)} + \|w^* - w\|^2_{L^2(0,T;\mathcal{V}^*)} \\
&\qquad < \eta,
\end{aligned}
$$

and (3) is proven. ∎

Appendix F

Chain rules

In the proof of Theorem 3.2.6, we used the following chain rule, which is a simplified version of [MRS16, Theorem 4.4]. Since we use different notation here and the additional assumptions significantly simplify some arguments, we give a full proof here for convenience.

Proposition F.1 (Parameterized chain rule)**.**
Let $(t,z) \in W^{1,\infty}(0,S) \times \big(\mathrm{AC}([0,S];\mathcal{R}) \cap L^\infty(0,S;\mathcal{Z})\big)$ *comply with* (3.2.11)-(3.2.12) *and the normalization condition* (N) *from Def. 3.2.5. Then the map* $s \mapsto \mathcal{I}(\ell(t(s)), z(s))$ *is absolutely continuous and its derivative fulfills almost everywhere in* $[0,S]$

$$|\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{I}(\ell(t(s)),z(s)) - \mathcal{P}(t(s),z(s))\dot{t}(s)| \le \mathcal{R}[z'](s) + \|\dot{z}(s)\|_{\mathcal{V}}\mathfrak{c}(t(s),z(s)). \tag{F.1}$$

Proof. We have for the constants c, C from (2.4.7) that for all $z \in \mathcal{Z}$

$$\mathcal{R}(z) \le C\|z\|_{\mathcal{X}} = \frac{C}{c}c\|-z\|_{\mathcal{X}} \le \frac{C}{c}\mathcal{R}(-z).$$

Step 1: Absolute continuity Let $0 \le r < s \le S$. In order to show absolute continuity, we have to distinguish three cases: Let $G \subset [0,S]$ be defined according to (3.2.11a).

First case: $r, s \in [0,S] \setminus G$. Application of (2.4.13) gives

$$\begin{aligned}
&\mathcal{I}(\ell(t(s)),z(s)) - \mathcal{I}(\ell(t(r)),z(r)) \\
&\quad = \mathcal{I}(\ell(t(s)),z(s)) - \mathcal{I}(\ell(t(s)),z(r)) + \langle \ell(t(r)) - \ell(t(s)), z(r)\rangle_{\mathcal{V}} \\
&\quad \le M_R\|z(r)-z(s)\|_{\mathcal{V}}\mathcal{R}(z(r)-z(s)) - \frac{\alpha}{4}\|z(r)-z(s)\|_{\mathcal{Z}}^2 \\
&\qquad\qquad + \langle -\mathrm{D}_z\mathcal{I}(\ell(t(s)),z(s)), z(r)-z(s)\rangle_{\mathcal{V}} + \|\dot{\ell}\|_{L^\infty(0,T;\mathcal{V}^*)}\|\dot{t}\|_{L^\infty(0,S)}R(s-t) \\
&\quad \le 2M_R R\frac{C}{c}\mathcal{R}(z(s)-z(r)) + \frac{C}{c}\mathcal{R}(z(s)-z(r)) + C_2(s-t) \\
&\quad \le \int_r^s \frac{C}{c}(2M_R R + 1)\mathcal{R}[z'](\tau) + C_2\,\mathrm{d}\tau.
\end{aligned}$$

Second case: $r, s \in [\alpha, \beta] \subseteq G$. Since z is $\mathcal{V}$-absolutely continuous on $[\alpha,\beta]$ according to (3.2.11b), we can argue as follows: Since we already know that $\mathcal{I}$

complies with the chain rule ($\mathcal{I}_4$) (cf. Lemma D.2) and only the linear term depends on time, it is sufficient to show that

$$s \mapsto \langle \ell(t(s)), v(s)\rangle_{\mathcal{V}}$$

is absolutely continuous. To this end, we choose arbitrary $0 \le r < s \le S$ and estimate

$$\begin{aligned}
&|\langle \ell(t(s)), v(s)\rangle_{\mathcal{V}} - \langle \ell(t(r)), v(r)\rangle_{\mathcal{V}}| \\
&\qquad \le |\langle \ell(t(s)) - \ell(t(r)), v(s)\rangle_{\mathcal{V}}| + |\langle \ell(t(r)), v(s) - v(r)\rangle_{\mathcal{V}}| \\
&\qquad \le \|\dot{\ell}\|_{L^\infty(0,T;\mathcal{V}^*)}(s-r)\|v(s)\|_{\mathcal{V}} + \|\ell\|_{L^\infty(0,T;\mathcal{V}^*)}\|v(s)-v(r)\|_{\mathcal{V}},
\end{aligned}$$

as well as

$$\begin{aligned}
&|\langle \ell(t(s)), v(s)\rangle_{\mathcal{V}} - \langle \ell(t(r)), v(r)\rangle_{\mathcal{V}}| \\
&\qquad \le \|\dot{\ell}\|_{L^\infty(0,T;\mathcal{V}^*)}(s-r)\|v(r)\|_{\mathcal{V}} + \|\ell\|_{L^\infty(0,T;\mathcal{V}^*)}\|v(s)-v(r)\|_{\mathcal{V}}.
\end{aligned}$$

Thus, $g(r) := \|\ell\|_{W^{1,\infty}(0,T;\mathcal{V}^*)}\|v(r)\|_{\mathcal{V}}$ defines a function in $L^1(0,S)$ such that

$$|\langle \ell(t(s)), v(s)\rangle_{\mathcal{V}} - \langle \ell(t(r)), v(r)\rangle_{\mathcal{V}}| \le \max\{g(r), g(s)\}|s-r| + \|\dot{\ell}\|_{L^\infty(0,T;\mathcal{V}^*)}\int_r^s \|\dot{v}(t)\|_{\mathcal{V}}\,\mathrm{d}t$$

for all $r, s \in [0,S]$, and we may proceed as in the proof of [AGS05, Theorem 1.2.5] to obtain absolute continuity of $s \mapsto \langle \ell(t(s)), v(s)\rangle_{\mathcal{V}}$, and consequently of $s \mapsto \mathcal{I}(\ell(t(s)), z(s))$ and the following chain rule holds for almost all $s \in [\alpha, \beta]$:

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{I}(\ell(t(s)), v(s)) = \langle \mathrm{D}_z\mathcal{I}(\ell(t(s)), v(s)), \dot{v}(s)\rangle_{\mathcal{V}} + \partial_\ell \mathcal{I}(\ell(t(s)), v(s))\dot{\ell}(t(s))\dot{t}(s). \tag{F.2}$$

Since $\dot{t}(s) = 0$ according to (3.2.12b), this implies the inequality

$$|-\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{I}(\ell(t(s)), z(s))| = |\langle -\mathrm{D}_z\mathcal{I}(\ell(t(s)), z(s)), \dot{z}(s)\rangle_{\mathcal{V}}| \le \mathcal{R}[z'](s) + \|\dot{z}(s)\|_{\mathcal{V}}\mathfrak{e}(t(s), z(s))$$

by the same arguments as in the proof of [KRZ13, Lemma 5.2]: We choose for every $s \in [0,S]$ an element $\mu(s) \in \partial\mathcal{R}(0)$ such that

$$\|-\mathrm{D}_z\mathcal{I}(\ell(t(s)), z(s)) - \mu(s)\|_{\mathcal{V}^*} = \mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(t(s)), z(s)), \partial\mathcal{R}(0)).$$

Observe that the distance in $\mathcal{V}^*$ is indeed realized by some element in $\partial\mathcal{R}(0)$ according to Lemma 2.4.4. Now, we use the parameterized chain rule (F.2) above and obtain

$$\begin{aligned}
|-\frac{\mathrm{d}}{\mathrm{d}r}\mathcal{I}(\ell(t(r)), v(r))| &= |\langle -\mathrm{D}_z\mathcal{I}(\ell(t(r)), z(r)) - \mu(r), \dot{z}(r)\rangle_{\mathcal{V}} + \langle \mu(r), \dot{z}(r)\rangle_{\mathcal{V}}| \\
&\le \|-\mathrm{D}_z\mathcal{I}(\ell(t(r)), z(r)) - \mu(r)\|_{\mathcal{V}^*}\|\dot{z}(r)\|_{\mathcal{V}} + \mathcal{R}(\dot{z}(r)) \\
&= \mathrm{dist}_{\mathcal{V}^*}(-\mathrm{D}_z\mathcal{I}(\ell(t(s)), z(s)), \partial\mathcal{R}(0))\|\dot{z}(r)\|_{\mathcal{V}} + \mathcal{R}(\dot{z}(r)).
\end{aligned}$$

Third case: $r \in G$ and $s \in [0,S] \setminus G$. Choose σ as the right boundary point of the connected component of G containing r. We then combine the first and second case and obtain

$$
\begin{aligned}
&|\mathcal{I}(\ell(t(s)),z(s))-\mathcal{I}(\ell(t(r)),z(r))\|\\
&\qquad\le \int_\sigma^s \frac{C}{c}(2M_R R+1)\mathcal{R}[z'](\tau)+C_2\,\mathrm{d}\tau + \int_r^\sigma \mathcal{R}[z'](\tau)+\|\dot z(\tau)\|_{\mathcal{V}}\mathfrak{c}(t(\tau),z(\tau))\,\mathrm{d}\tau\\
&\qquad\le \int_r^s h(\tau)\,\mathrm{d}\tau \qquad \text{for some } h\in L^1(0,S) \text{ that is independent of } s \text{ and } r.
\end{aligned}
$$

In all three cases, we conclude by help of Lemma C.4 that $s \mapsto \mathcal{I}(\ell(t(s)),z(s))$ is indeed absolutely continuous.

Step 2: Chain rule inequality If s lies in the closure of a connected component contained in G, the chain rule inequality (F.1) follows directly from the consideration in the second case of the first step. Else, we choose a sequence $r_n = s-h_n \in [0,S]\setminus G$ such that $h_n \searrow 0$. This implies $-\mathrm{D}_z\mathcal{I}(\ell(t(r_n)),z(r_n)) \in \partial\mathcal{R}(0)$ as well as $\|z(s)-z(r_n)\|_{\mathcal{V}} \to 0$, since $z \in C([0,S];\mathcal{V})$ according to Proposition C.13. Thus, we can estimate the difference quotient as follows:

$$
\begin{aligned}
&\tfrac{1}{h_n}\Big(\mathcal{I}(\ell(t(s)),z(s))-\mathcal{I}(\ell(t(r_n)),z(r_n))\Big)\\
&\quad= \tfrac{1}{h_n}\Big(\mathcal{I}(\ell(t(s)),z(s))-\mathcal{I}(\ell(t(r_n)),z(s))\Big)+\tfrac{1}{h_n}\Big(\mathcal{I}(\ell(t(r_n)),z(s))-\mathcal{I}(\ell(t(r_n)),z(r_n))\Big)\\
&\quad\ge \tfrac{1}{h_n}\int_{r_n}^s \mathcal{P}(t(\tau),z(s))\dot t(\tau)\,\mathrm{d}\tau - \tfrac{1}{h_n}M_R\|z(s)-z(r_n)\|_{\mathcal{V}}\mathcal{R}(z(s)-z(r_n))+\tfrac{\alpha}{4h_n}\|z(s)-z(r_n)\|_{\mathcal{Z}}^2\\
&\qquad\qquad + \tfrac{1}{h_n}\langle \mathrm{D}_z\mathcal{I}(\ell(t(r_n)),z(r_n)),z(s)-z(r_n)\rangle_{\mathcal{V}}\\
&\quad\ge \tfrac{1}{h_n}\int_{r_n}^s \mathcal{P}(t(\tau),z(s))\dot t(\tau)\,\mathrm{d}\tau - M_R\underbrace{\|z(s)-z(r_n)\|_{\mathcal{V}}}_{\to 0}\underbrace{\mathcal{R}\left(\frac{z(s)-z(r_n)}{h_n}\right)}_{\to\mathcal{R}[z'](s)}-\tfrac{1}{h_n}\mathcal{R}(z(s)-z(r_n)) \qquad \text{(F.3)}\\
&\quad\to \mathcal{P}(t(s),z(s)\dot t(s)-\mathcal{R}[z'](s),
\end{aligned}
$$

where the first term converges in every Lebesgue point s of $\dot t$, and the last term converges almost everwhere according to Prop. C.10. An approximation of s from above by a sequence $r_n = s+h_n \in [0,S]\setminus G$, where $h_n \searrow 0$, gives the opposite estimate: [1] We estimate as in (F.3) but now divide by $s-r_n = -h_n < 0$

$$
\begin{aligned}
&\tfrac{1}{-h_n}\Big(\mathcal{I}(\ell(t(s)),z(s))-\mathcal{I}(\ell(t(r_n)),z(r_n))\Big)\\
&\quad\le -\tfrac{1}{h_n}\int_{r_n}^s \mathcal{P}(t(\tau),z(s))\dot t(\tau)\,\mathrm{d}\tau + \tfrac{1}{h_n}M_R\|z(s)-z(r_n)\|_{\mathcal{V}}\mathcal{R}\,(z(s)-z(r_n))+\tfrac{1}{h_n}\mathcal{R}(z(s)-z(r_n))\\
&\quad\le \tfrac{1}{h_n}\int_s^{r_n} \mathcal{P}(t(\tau),z(s))\dot t(\tau)\,\mathrm{d}\tau + M_R\tfrac{C}{c}\underbrace{\|z(s)-z(r_n)\|_{\mathcal{V}}}_{\to 0}\underbrace{\mathcal{R}\left(\frac{z(r_n)-z(s)}{h_n}\right)}_{\to\mathcal{R}[z'](s)}+\mathcal{R}\left(\frac{z(r_n)-z(s)}{h_n}\right)\\
&\quad\to \mathcal{P}(t(s),z(s)\dot t(s)+\mathcal{R}[z'](s).
\end{aligned}
$$

■

[1] At this point, the symmetry of $\mathcal{R}$ is needed, since otherwise, we only obtain an estimate against $\mathcal{R}[-z'](s)$.

To prove the uniform estimate for the driving forces, we need the following chain rule:

Proposition F.2. *Let $z \in H^1((0,T);\mathcal{V}) \cap L^\infty((0,T);\mathcal{Z})$ and $\mathrm{D}\mathcal{J}(z(\cdot)) \in L^\infty((0,T);\mathcal{V}^*)$. Then for almost all t, the mapping $t \mapsto \mathcal{J}(z(t))$ is differentiable and we have the identity*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{J}(z(t)) = \langle Az(t), \dot{z}(t)\rangle_{\mathcal{V}^*,\mathcal{V}} + \langle \mathrm{D}\mathcal{F}(z(t)), \dot{z}(t)\rangle_{\mathcal{V}^*,\mathcal{V}}. \tag{F.4}$$

Integrated version of the chain rule: Let $z \in W^{1,1}((0,T);\mathcal{V}) \cap L^\infty((0,T);\mathcal{Z})$ with $\mathrm{D}\mathcal{J}(z(\cdot)) \in L^\infty((0,T);\mathcal{V}^*)$ and assume that $t \mapsto \mathcal{J}(z(t))$ is continuous on $[0,T]$. Then for all $t_1 < t_2 \in [0,T]$

$$\mathcal{J}(z(t_2)) - \mathcal{J}(z(t_1)) = \int_{t_1}^{t_2} \langle \mathrm{D}\mathcal{J}(z(r)), \dot{z}(t)\rangle_{\mathcal{V}^*,\mathcal{V}}\,\mathrm{d}r. \tag{F.5}$$

Proof. For the proof of (F.4), we start from the λ-convexity of $\mathcal{I}$ from (2.4.13), yielding that for $\rho := \|z\|_{L^\infty((0,T);\mathcal{Z})}$, there exists $M_\rho > 0$ such that for all $h > 0$ we have the estimate

$$\begin{aligned} &h^{-1}\Big(\mathcal{J}(z(t{+}h)) - \mathcal{J}(z(t))\Big) \\ &\qquad \geq \langle \mathrm{D}\mathcal{J}(z(t)), h^{-1}(z(t+h) - z(t))\rangle_{\mathcal{V},\mathcal{V}^*} - M_\rho \mathcal{R}(z(t+h) - z(t))\|h^{-1}(z(t+h) - z(t))\|_{\mathcal{V}}. \end{aligned}$$

Note that we may consider the duality pairing in $\mathcal{V}^*$ on the right-hand side, since we already know that $\mathrm{D}\mathcal{J}(z(\cdot))$ maps into $\mathcal{V}^*$. Using [CH98, Thm. 1.4.35], we find that $h^{-1}(z(t+h) - z(t)) \to \dot{z}(t)$ strongly in $\mathcal{V}$ almost everwhere, so that the right-hand side converges to $\langle \mathrm{D}\mathcal{J}(z(t)), \dot{z}(t)\rangle_{\mathcal{V},\mathcal{V}^*}$ for $h \searrow 0$. On the other hand, using the same estimate, but changing the roles of $z(t)$ and $z(t+h)$, we also find the opposite estimate

$$\begin{aligned} &h^{-1}\Big(\mathcal{J}(z(t)) - \mathcal{J}(z(t+h))\Big) \\ &\qquad \geq \langle \mathrm{D}\mathcal{J}(z(t+h)), h^{-1}(z(t) - z(t+h))\rangle_{\mathcal{V},\mathcal{V}^*} - M_\rho \mathcal{R}(z(t) - z(t+h))\|h^{-1}(z(t) - z(t+h))\|_{\mathcal{V}} \\ &\qquad \to -\langle \mathrm{D}\mathcal{J}(z(t)), \dot{z}(t)\rangle_{\mathcal{V},\mathcal{V}^*} \end{aligned}$$

for almost all t. In order to prove the last convergence, note that $z(t+h) \to z(t)$ strongly in $\mathcal{V}$, and together with $\|z\|_{L^\infty((0,T);\mathcal{Z})} < \infty$, this implies $z(t+h) \rightharpoonup z(t)$ weakly in $\mathcal{Z}$. Taking into account assumption (2.4.4a), this implies the weak convergence $\mathrm{D}\mathcal{J}(z(t+h)) \rightharpoonup \mathrm{D}\mathcal{J}(z(t))$ in $\mathcal{Z}^*$. Since $\mathrm{D}\mathcal{J}(\cdot) \in L^\infty((0,T);\mathcal{V}^*)$, this implies weak convergence also in $\mathcal{V}^*$, and thus convergence of the duality pairing. Thus, we have shown that

$$\lim_{h\searrow 0} h^{-1}\Big(\mathcal{J}(z(t+h)) - \mathcal{J}(z(t))\Big) = \langle \mathrm{D}\mathcal{J}(z(t)), \dot{z}(t)\rangle_{\mathcal{V},\mathcal{V}^*} = \langle Az(t), \dot{z}(t)\rangle_{\mathcal{V}^*,\mathcal{V}} + \langle \mathrm{D}\mathcal{F}(z(t)), \dot{z}(t)\rangle_{\mathcal{V}^*,\mathcal{V}}.$$

A similar argument for $h < 0$ proves (F.4)

For the proof of the integrated version of the chain rule, let $t_1 < t_2 \in [0,T)$ and $h_0 > 0$ such that $t_2 + h_0 \leq T$. Then for all $0 < h \leq h_0$, the uniform subdifferentiability (2.4.13) implies

$$
\begin{aligned}
h^{-1}\int_{t_1}^{t_2} &\mathcal{J}(z(t+h)) - \mathcal{J}(z(t))\,\mathrm{d}t \\
&\geq \int_{t_1}^{t_2} \langle \mathrm{D}\mathcal{J}(z(t)), h^{-1}(z(t+h)-z(t))\rangle\,\mathrm{d}t - \tfrac{\lambda}{h}\int_{t_1}^{t_2} \|z(t+h)-z(t)\|_{\mathcal{V}}^2\,\mathrm{d}t,
\end{aligned}
$$

where $\lambda > 0$ depends on $\|z\|_{L^\infty(0,T;\mathcal{Z})}$. Thanks to the continuity of $\mathcal{J}(z(\cdot))$, for the left hand side we obtain $\lim_{h\to 0} h^{-1}\int_{t_1}^{t_2} \mathcal{J}(z(t+h)) - \mathcal{J}(z(t))\,\mathrm{d}t = \mathcal{J}(z(t_2)) - \mathcal{J}(z(t_1))$. Since $z \in W^{1,1}((0,T);\mathcal{V})$, on each $(t_1,t_2) \Subset (0,T)$ the difference quotients converge strongly in the following sense: $h^{-1}(z(\cdot+h)-z(\cdot)) \to \dot{z}(\cdot)$ strongly in $L^1((t_1,t_2);\mathcal{V})$, [CH98, Cor. 1.4.39]. Thus the first integral on the right hand side converges to $\int_{t_1}^{t_2}\langle \mathrm{D}\mathcal{J}(z(r)), \dot{z}(t)\rangle\,\mathrm{d}r$, while the second integral on the right hand side converges to zero. A similar argument for $h < 0$ finally proves (F.5). ∎

Appendix G

On the assumption that $\mathcal{V}$ is uniformly convex

It should be noted that, for almost every argument in this dissertation, we only need to assume that $\mathcal{V}$ is a reflexive Banach space. If $\mathcal{V}$ is uniformly convex, then the Milman-Pettis-Theorem, [Pet39], asserts that $\mathcal{V}$ is also reflexive. The converse is not true in general, see, e.g., [Day41]. The requirement that $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ be uniformly convex with a modulus of convexity of power 2 is needed only in the proof of uniqueness of solutions of the viscously regularized system, that is in Step 0 of the proof of Prop. (3.1.3). To be more precise, we require the uniform convexity of $\partial\mathcal{R}_{2,1}(\cdot)$ in order to obtain the crucial estimate (3.1.8). However, according to [PR86, Prop. 2.11], both requirements are equivalent on smooth Banach spaces, and Lemma G.1 here below asserts that one of the implications is true for arbitrary Banach spaces.

Uniform convexity of a normed space $(X, \|\cdot\|_X)$ is defined in [BGHV09] as follows: The **modulus of convexity** is the function $\delta_{\|\cdot\|_X} : [0,2) \to [0,\infty]$ defined by

$$\delta_{\|\cdot\|_X}(t) := \inf\Big\{ 1 - \frac{1}{2}\|x+y\|_X \quad \Big| \quad \|x\|_X, \|y\|_X \le 1, \quad \|x-y\|_X = t \Big\},$$

where the infimum over the empty set is ∞. The function $\delta_{\|\cdot\|_X}(\cdot)$ is monotonely increasing, cf. [Die84]. The space $(X, \|\cdot\|_X)$ is called **uniformly convex** if it holds that $\delta_{\|\cdot\|_X}(t) > 0$ for all $t > 0$; and it is called **uniformly convex with modulus of convexity of power type** p if there exists $C > 0$ such that $\delta_{\|\cdot\|_X}(t) \ge Ct^p$ for all $t > 0$. It should be stressed here that uniform convexity is a property of the norm, and not of the topology, that is, choosing an equivalent norm on X need not preserve uniform convexity or the power type of the modulus of convexity. We now have the following implication

Lemma G.1. *Let $(X, \|\cdot\|_X)$ be a Banach space. If $(X, \|\cdot\|_X)$ is uniformly convex with modulus of convexity of power type 2, then there exists a constant $C > 0$, depending on $(X, \|\cdot\|_X)$ alone, such that for the map $R : X \to \mathbb{R}$, $R(x) := \frac{1}{2}\|x\|_X^2$, it holds that*

$$\begin{aligned} &\forall\, x, y \in X \setminus \{0\}, \forall\, \eta_x, \in \partial R(x), \eta_y \in \partial R(y) : \\ &\qquad\qquad \langle \eta_x - \eta_y, x - y \rangle_X \ge C\|x-y\|_X^2. \end{aligned} \tag{G.1}$$

The proof is completely analoguous to that of [PR86, Prop. 2.11]. There, the requirement that X be smooth has the only effect of rendering $\partial R(\cdot)$ a single-valued function $X \to X^*$. We give the proof here for completeness.

Proof. Let $\delta : [0,2) \to [0,\infty]$ be the modulus of convexity of X. As a first step, we show that for all $x, y \in X \setminus \{0\}$, it holds that

$$\frac{\|x+y\|_X^2}{2(\|x\|_X^2 + \|y\|_X^2)} \leq 1 - \delta\Big(\frac{\|x-y\|_X}{(2(\|x\|_X^2 + \|y\|_X^2))^{\frac{1}{2}}}\Big). \tag{G.2}$$

To this end, first note that it holds that

$$\begin{aligned}\forall\, x,y \in X \text{ such that } \|x\|_X,\|y\|_X \leq 1,\ \|x-y\|_X \leq 1 :&\\ \|x+y\|_X \leq 1 - \delta(2\|x-y\|_X) \leq 1 - \delta(\|x-y\|_X),&\end{aligned}$$

were the last estimate is justified by the monotonicity of δ. This implies that

$$\begin{aligned}\forall\, x,y \in X \text{ such that } \|x\|_X,\|y\|_X \leq 1,\ \|x \pm y\|_X \leq 1 :&\\ \|x+y\|_X^2 \leq 1 - \delta(\|x-y\|_X).&\end{aligned} \tag{G.3}$$

Now, for arbitrary $x, y \in X \setminus \{0\}$, if we set

$$\tilde{x} := \frac{x}{(2(\|x\|_X^2 + \|y\|_X^2)^{\frac{1}{2}}} \quad \text{and} \quad \tilde{y} := \frac{y}{(2(\|x\|_X^2 + \|y\|_X^2))^{\frac{1}{2}}},$$

it holds that $\|\tilde{x}\|_X, \|\tilde{y}\|_X \leq \frac{1}{\sqrt{2}}$, and infer from

$$\|x \pm y\|_X^2 \leq \|x\|_X^2 + \|y\|_X^2 + 2\|x\|_X \cdot \|y\|_X \leq 2(\|x\|_X^2 + \|y\|_X^2)$$

that $\|\tilde{x} \pm \tilde{y}\|_X \leq 1$, allowing us to apply (G.3) to $\tilde{x}$ and $\tilde{y}$ and thereby obtain (G.2).

Let now $x, y \in X \setminus \{0\}$ be arbitrary, and $\eta_x, \in \partial R(x), \eta_y \in \partial R(y)$. Then it holds for all $z \in X$ that

$$\langle \eta_x, z - x\rangle_X + \tfrac{1}{2}\|x\|_X^2 \leq \tfrac{1}{2}\|z\|_X^2,$$

which, for $z = \frac{x+y}{2}$, reads

$$\langle \eta_x, x - y\rangle_X - \|x\|_X^2 \geq -\|\frac{x+y}{2}\|_X^2, \tag{G.4}$$

and in the same way, we obtain the estimate

$$-\langle \eta_y, x - y\rangle_X - \|y\|_X^2 \geq -\|\frac{x+y}{2}\|_X^2. \tag{G.5}$$

Adding (G.4) and (G.5) yields

$$\langle \eta_x - \eta_y, x - y\rangle_X \geq \|x\|_X^2 + \|y\|_X^2 - \tfrac{1}{2}\|x+y\|_X^2. \tag{G.6}$$

Now, using the estimate (G.2), (G.6) reads

$$\begin{aligned}\langle \eta_x - \eta_y, x-y\rangle_X &\geq (\|x\|_X^2 + \|y\|_X^2)\cdot \delta\Big(\frac{\|x-y\|_X}{(2(\|x\|_X^2+\|y\|_X^2))^{\frac{1}{2}}}\Big)\\ &\geq C\cdot(\|x\|_X^2+\|y\|_X^2)\Big(\frac{\|x-y\|_X}{(2(\|x\|_X^2+\|y\|_X^2))^{\frac{1}{2}}}\Big)^2\\ &= \widetilde{C}\|x-y\|_X^2,\end{aligned}$$

where $C>0$ is a constant that depends on X and $\|\cdot\|_X$ alone.

■

An obvious way to obtain a space that is uniformly convex with modulus of convexity of power type 2 is to choose a Hilbert space $\mathcal{V}$, such as the Sobolev space $H^1(\Omega)$. For a Banach space $\mathcal{V}$, it would also be sufficient to require the existence of an operator

$$\left.\begin{array}{c}\mathbb{V}\in \mathrm{Lin}(\mathcal{V},\mathcal{V}^*) \text{ that is linear, self-adjoint, bounded and elliptic, i.e.,}\\ \forall\, w,v\in\mathcal{V}: \langle \mathbb{V}w,v\rangle_{\mathcal{V}} = \langle \mathbb{V}v,w\rangle_{\mathcal{V}}\\ \exists\,\Gamma,\gamma>0: \forall v\in\mathcal{V}:\ \gamma\|v\|_{\mathcal{V}}^2 \leq \langle\mathbb{V}v,v\rangle_{\mathcal{V}} \leq \Gamma\|v\|_{\mathcal{V}}^2\end{array}\right\} \tag{G.7}$$

and then define an equivalent norm $|\!|\!|\cdot|\!|\!|$ on $\mathcal{V}$ via

$$|\!|\!|v|\!|\!| := \sqrt{\langle \mathbb{V}v,v\rangle_{\mathcal{V}}}.$$

In this case, the viscous augmentation $\mathcal{R}_{2,1}(v) := \frac{1}{2}|\!|\!|v|\!|\!|^2$ is Fréchet-differentiable with Fréchet-derivative $\mathrm{D}\mathcal{R}_{2,1}(v) = \mathbb{V}v$, since it holds that

$$\begin{aligned}&\lim_{\|h\|_{\mathcal{V}}\to 0}\frac{1}{\|h\|_{\mathcal{V}}}\Big(\mathcal{R}_{2,1}(v+h)-\mathcal{R}_{2,1}(v)-\langle\mathbb{V}v,h\rangle_{\mathcal{V}}\Big)\\ &\quad=\lim_{\|h\|_{\mathcal{V}}\to 0}\frac{1}{2\|h\|_{\mathcal{V}}}\Big(\langle\mathbb{V}h,h\rangle_{\mathcal{V}}+\langle\mathbb{V}v,h\rangle_{\mathcal{V}}+\langle\mathbb{V}h,v\rangle_{\mathcal{V}}-2\langle\mathbb{V}v,h\rangle_{\mathcal{V}}\Big)\\ &\quad=\lim_{\|h\|_{\mathcal{V}}\to 0}\frac{|\!|\!|h|\!|\!|^2}{2\|h\|_{\mathcal{V}}}=0,\end{aligned}$$

and thus has the single-valued convex subdifferential $\partial\mathcal{R}_{2,1}(v) = \{\mathrm{D}\mathcal{R}_{2,1}(v)\}$. From here, the uniform convexity of $\partial\mathcal{R}_{2,1}(\cdot)$ is a direct consequence of the ellipticity of $\mathbb{V}$. However, requiring the existence of an operator $\mathbb{V}\in\mathrm{Lin}(\mathcal{V},\mathcal{V}^*)$ that satisfies (G.7) effectively turns $\mathcal{V}$ into an inner product space. Now, since $\mathcal{V}$ is complete w.r.t. the norm $\|\cdot\|_{\mathcal{V}}$, the same is true for the equivalent norm $|\!|\!|\cdot|\!|\!|$, so that we must have been dealing with a Hilbert space from the beginning.

Other possible choices for $\mathcal{V}$ that are not Hilbert spaces, would be Sobolev spaces $W^{k,p}(\Omega)$. If endowed with the right norm, these are uniformly convex:

Lemma G.2. *Let $\Omega\subset\mathbb{R}^d$ be a domain and for $p>1$, $k\in\mathbb{N}$ we consider*

$$W^{k,p}(\Omega) := \{f\in L^p(\Omega)\,|\,\forall|\alpha|\leq k : \mathrm{D}^\alpha f\in L^p(\Omega)\}$$

together with the norm

$$\|f\|_{k,p} := \sum_{|\alpha|\le k} \|\mathrm{D}^\alpha f\|_{L^p(\Omega)}$$

Then $W^{k,p}(\Omega)$ is uniformly convex.

Proof. This follows from the fact that $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ is uniformly convex with modulus of convexity of power p, see [Han56]. Let δ_p be the modulus of convexity of $\|\cdot\|_{L^p(\Omega)}$. Let further $t > 0$ and $f, g \in W^{k,p}(\Omega)$ such that $\|f-g\|_{k,p} = t$. In particular, we have that $t_0 := \|f-g\|_{L^p(\Omega)} > 0$. For $|\alpha| \le k$, let $t_\alpha := \|\mathrm{D}^\alpha(f-g)\|_{L^p(\Omega)} \ge 0$, then

$$\frac{1}{2}\|f\|_{k,p} + \frac{1}{2}\|g\|_{k,p} - \frac{1}{2}\|f-g\|_{k,p} \ge \sum_{|\alpha|\le k} \delta_p(t_\alpha) \ge \delta_p(t_0) > 0.$$

∎

Now, every uniformly convex normed space can be endowed with an equivalent norm (that is, with a norm that induces the same topology) whose modulus of convexity is of power type 2:

Lemma G.3. *Let $(X, \|\cdot\|_{\mathcal{X}})$ be a uniformly convex space. Then there is an equivalent norm $|\!|\!|\cdot|\!|\!|$ on X that has a modulus of convexity of power type 2.*

Proof. We use [BGHV09, Thm 4.3] according to which it is sufficient to show existence of a function $f : X \to \mathbb{R}$ that is continuous, uniformly convex, and satisfies $f(x) \le \|x\|_X^2$ for all $x \in X$. Simple calculations show that $f(x) := \frac{1}{2}\|x\|_X^2$ is such a function. ∎

In conclusion, there exists an equivalent norm $|\!|\!|\cdot|\!|\!|_{k,p}$ on $W^{k,p}(\Omega)$ such that $(W^{k,p}(\Omega), |\!|\!|\cdot|\!|\!|_{k,p})$ is uniformly convex with modulus of convexity of power type 2.

Remark G.4. *If $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ is uniformly convex with a modulus of convexity of power type $p > 2$, we can still follow the arguments in the proof of Lemma G.1. Now, instead of* (3.1.8), *we obtain that*

$$\begin{aligned} \gamma\|\dot z_1(t) - \dot z_2(t)\|_{\mathcal{V}}^p + \tfrac{1}{2}\tfrac{\mathrm{d}}{\mathrm{d}t}(\langle A(z_1(t)-z_2(t)), z_1(t)-z_2(t)\rangle_{\mathcal{Z}}) \\ \le \widetilde{C}\|z_1(t)-z_2(t)\|_{\mathcal{V}}^{\frac{p}{p-1}} + \frac{\gamma}{2}\|\dot z_1(t) - \dot z_2(t)\|_{\mathcal{V}}^p, \end{aligned}$$

i.e.,

$$\tfrac{\alpha}{2}\|z_1(s)-z_2(s)\|_{\mathcal{Z}}^2 \le \widetilde{C}\int_0^s \|z_1(t)-z_2(t)\|_{\mathcal{V}}^{\frac{p}{p-1}}\,\mathrm{d}t \le C\int_0^s \|z_1(t)-z_2(t)\|_{\mathcal{Z}}^{\frac{p}{p-1}}\,\mathrm{d}t.$$

In order to turn this into a useful estimate for $\|z_1(s)-z_2(s)\|_{\mathcal{Z}}^2$ however, it seems that some generalized Gronwall inequalities are needed.

Index

Bibliography

[AC18] Chems Eddine Arroud and Giovanni Colombo. A maximum principle for the controlled sweeping process. *Set-Valued Var. Anal.*, 26(3):607–629, 2018.

[AE06] Jean-Pierre Aubin and Ivar Ekeland. *Applied nonlinear analysis*. Dover Publications, Inc., Mineola, NY, 2006. Reprint of the 1984 original.

[AFP00] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[AGS05] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Basel: Birkhäuser, 2005.

[BFM12] Jean-F. Babadjian, Gilles A. Francfort, and Maria G. Mora. Quasi-static evolution in nonassociative plasticity: the cap model. *SIAM J. Math. Anal.*, 44(1):245–292, 2012.

[BGHV09] Jonathan Borwein, Antonio J. Guirao, Petr Hájek, and Jon Vanderwerff. Uniformly convex functions on Banach spaces. *Proc. Amer. Math. Soc.*, 137(3):1081–1091, 2009.

[BK15] Martin Brokate and Pavel Krejčí. Weak differentiability of scalar hysteresis operators. *Discrete Contin. Dyn. Syst.*, 35(6):2405–2421, 2015.

[Bré73] Haïm Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Mathematics Studies, No. 5. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973.

[Bro87] Martin Brokate. *Optimale Steuerung von gewöhnlichen Differentialgleichungen mit Nichtlinearitäten vom Hysteresis-Typ*, volume 35 of *Methoden und Verfahren der Mathematischen Physik [Methods and Procedures in Mathematical Physics]*. Verlag Peter D. Lang, Frankfurt am Main, 1987.

[CGH+96] Robert M. Corless, Gaston H. Gonnet, David E. G. Hare, David J. Jeffrey, and Donald E. Knuth. On the Lambert W function. *Adv. Comput. Math.*, 5(4):329–359, 1996.

[CH98] Thierry Cazenave and Alain Haraux. *An Introduction to Semilinear Evolution Equations*, volume 13 of *Oxford Lecture Ser. Math. Appl.* Clarendon Press, 1998.

[CHHM15] Giovanni Colombo, René Henrion, Nguyen D. Hoang, and Boris S. Mordukhovich. Discrete approximations of a controlled sweeping process. *Set-Valued Var. Anal.*, 23(1):69–86, 2015.

[CHHM16] Giovanni Colombo, René Henrion, Nguyen D. Hoang, and Boris S. Mordukhovich. Optimal control of the sweeping process over polyhedral controlled sets. *J. Differential Equations*, 260(4):3397–3447, 2016.

[CP16] Giovanni Colombo and Michele Palladino. The minimum time function for the controlled Moreau's sweeping process. *SIAM J. Control Optim.*, 54(4):2036–2062, 2016.

[CV90] Pierluigi Colli and Augusto Visintin. On a class of doubly nonlinear evolution equations. *Comm. Partial Differential Equations*, 15(5):737–756, 1990.

[Day41] Mahlon M. Day. Reflexive Banach spaces not isomorphic to uniformly convex spaces. *Bull. Amer. Math. Soc.*, 47:313–317, 1941.

[DH08] Bojan J. Dimitrijevic and Klaus Hackl. A method for gradient enhancement of continuum damage models. *Technische Mechanik*, 28:43–52, 2008.

[Die69] Jean Dieudonné. Foundations of modern analysis. Enlarged and corrected printing. New York-London: Academic Press. XV, 387 p., 1969.

[Die84] Joseph Diestel. *Sequences and series in Banach spaces*, volume 92 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1984.

[EL14] Michela Eleuteri and Luca Lussardi. Thermal control of a rate-independent model for permanent inelastic effects in shape memory materials. *Evol. Equ. Control Theory*, 3(3):411–427, 2014.

[Els05] Jürgen Elstrodt. *Maß- und Integrationstheorie*. Springer-Lehrbuch. [Springer Textbook]. Springer-Verlag, Berlin, fourth edition, 2005. Grundwissen Mathematik. [Basic Knowledge in Mathematics].

[ELS13] Michela Eleuteri, Luca Lussardi, and Ulisse Stefanelli. Thermal control of the Souza-Auricchio model for shape memory alloys. *Discrete Contin. Dyn. Syst. Ser. S*, 6(2):369–386, 2013.

[EM06] Messoud A. Efendiev and Alexander Mielke. On the rate-independent limit of systems with dry friction and small viscosity. *J. Convex Anal.*, 13(1):151–167, 2006.

[Emm04] Etienne Emmrich. *Gewöhnliche und Operator-Differentialgleichungen*. Vieweg & Sohn Verlag, Wiesbaden, 2004.

[ET76] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*. Studies in Mathematics and its Applications, Vol. 1. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1976. Translated from the French.

[Fed69] Herbert Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[GP06] Leszek Gasiński and Nikolaos Papageorgiou. *Nonlinear Analysis*, volume 9 of *Mathematical Analysis and Applications*. Chapman & Hall/CRC, 2006.

[Han56] Olof Hanner. On the uniform convexity of L^p and l^p. *Ark. Mat.*, 3:239–244, 1956.

[HUL01] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of convex analysis*. Grundlehren Text Editions. Springer-Verlag, Berlin, 2001. Abridged version of *Convex analysis and minimization algorithms. I*, Springer, Berlin, 1993.

[HvVW16] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis. *Analysis in Banach spaces. Volume I. Martingales and Littlewood-Paley theory.* Cham: Springer, 2016.

[IT79] Alexander D. Ioffe and Vladimir M. Tihomirov. *Theory of extremal problems*, volume 6 of *Studies of Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1979.

[KMS21] Dorothee Knees, Christian Meyer, and Michael Sievers. Optimal control of non-convex rate-independent systems via vanishing viscosity – the finite dimensional case. Preprint SPP1962-171, 6 2021.

[Kne19] Dorothee Knees. Convergence analysis of time-discretisation schemes for rate-independent systems. *ESAIM Control Optim. Calc. Var.*, 25:Paper No. 65, 38, 2019.

[KRZ11] Dorothee Knees, Riccarda Rossi, and Chiara Zanini. A vanishing viscosity approach to a rate-independent damage model. WIAS Preprint 1633, Weierstrass Institute, Berlin, 2011.

[KRZ13] Dorothee Knees, Ricarda Rossi, and Chiara Zanini. A vanishing viscosity approach to a rate-independent damage model. *Math. Models Methods Appl. Sci.*, 23(4):565–616, 2013.

[KRZ15] Dorothee Knees, Riccarda Rossi, and Chiara Zanini. A quasilinear differential inclusion for viscous and rate-independent damage systems in non-smooth domains. *Nonlinear Anal. Real World Appl.*, 24:126–162, 2015.

[KRZ19] Dorothee Knees, Riccarda Rossi, and Chiara Zanini. Balanced viscosity solutions to a rate-independent system for damage. *European J. Appl. Math.*, 30(1):117–175, 2019.

[KT18] Dorothee Knees and Stephanie Thomas. Optimal control for a rate-independent system constrained to parameterized balanced viscosity solutions. Technical report, University of Kassel, 2018.

[KZ21] Dorothee Knees and Chiara Zanini. Existence of parameterized BV-solutions for rate-independent systems with discontinuous loads. *Discrete Contin. Dyn. Syst., Ser. S*, 14(1):121–149, 2021.

[Mie11] Alexander Mielke. Differential, energetic, and metric formulations for rate-independent processes. In *Nonlinear PDE's and applications. C.I.M.E. summer school, Cetraro, Italy, June 23–28, 2008*, pages 87–170. Berlin: Springer, 2011.

[MM05] Andreas Mainik and Alexander Mielke. Existence results for energetic models for rate-independent systems. *Calc. Var. Partial Differ. Equ.*, 22(1):73–99, 2005.

[MM09] Andreas Mainik and Alexander Mielke. Global existence for rate-independent gradient plasticity at finite strain. *J. Nonlinear Sci.*, 19(3):221–248, 2009.

[Mor73] Jean-Jacques Moreau. Probléme d'évolution associé á un convexe mobile d'un espace hilbertien. *C. R. Acad. Sci. Paris Sér. A-B*, 276:A791–A794, 1973.

[MR07] Alexander Mielke and Riccarda Rossi. Existence and uniqueness results for a class of rate-independent hysteresis problems. *Math. Models Methods Appl. Sci.*, 17(1):81–123, 2007.

[MR09] Alexander Mielke and Filip Rindler. Reverse approximation of energetic solutions to rate-independent processes. *NoDEA Nonlinear Differential Equations Appl.*, 16(1):17–40, 2009.

[MR15] Alexander Mielke and Tomáš Roubíček. *Rate-independent systems. Theory and application.* New York, NY: Springer, 2015.

[MRS09] Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. Modeling solutions with jumps for rate-independent systems on metric spaces. *Discrete Contin. Dyn. Syst.*, 25(2):585–615, 2009.

[MRS12a] Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. BV solutions and viscosity approximations of rate-independent systems. *ESAIM Control Optim. Calc. Var.*, 18(1):36–80, 2012.

[MRS12b] Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. Variational convergence of gradient flows and rate-independent evolutions in metric spaces. *Milan J. Math.*, 80(2):381–410, 2012.

[MRS13] Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. Nonsmooth analysis of doubly nonlinear evolution equations. *Calc. Var. Partial Differential Equations*, 46(1-2):253–310, 2013.

[MRS16] Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. Balanced viscosity (BV) solutions to infinite-dimensional rate-independent systems. *J. Eur. Math. Soc. (JEMS)*, 18(9):2107–2165, 2016.

[MT99] Alexander Mielke and Florian Theil. A mathematical model for rate-independent phase transformations with hysteresis. Proceedings of the Workshop on "Models of Continuum Mechanics in Analysis and Engineering, pages 117–129, Aachen, 1999. Shaker-Verlag.

[MT04] Alexander Mielke and Florian Theil. On rate-independent hysteresis models. *NoDEA Nonlinear Differential Equations Appl.*, 11(2):151–189, 2004.

[MTL02] Alexander Mielke, Florian Theil, and Valery I. Levitas. A variational formulation of rate-independent phase transformations using an extremum principle. *Arch. Ration. Mech. Anal.*, 162(2):137–177, 2002.

[MW20] Christian Meyer and Stephan Walther. Optimal control of perfect plasticity part ii: Displacement tracking. Preprint SPP1962-136, 3 2020.

[Pet39] Billy James Pettis. A proof that every uniformly convex space is reflexive. *Duke Math. J.*, 5(2):249–253, 1939.

[PR86] Esteban I. Poffald and Simeon Reich. An incomplete Cauchy problem. *J. Math. Anal. Appl.*, 113(2):514–543, 1986.

[Rin08] Filip Rindler. Optimal control for nonconvex rate-independent evolution processes. *SIAM J. Control Optim.*, 47(6):2773–2794, 2008.

[Rin09] Filip Rindler. Approximation of rate-independent optimal control problems. *SIAM J. Numer. Anal.*, 47(5):3884–3909, 2009.

[RMS08] Riccarda Rossi, Alexander Mielke, and Giuseppe Savaré. A metric approach to a class of doubly nonlinear evolution equations and applications. *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (5)*, 7(1):97–169, 2008.

[Rud99] Walter Rudin. *Reelle und komplexe Analysis*. R. Oldenbourg Verlag, Munich, 1999. Translated from the third English (1987) edition by Uwe Krieg.

[Sie20] Michael Sievers. *A Numerical Scheme for Rate-Independent Systems*. PhD thesis, TU Dortmund, 2020. unpublished thesis.

[Sie21] Michael Sievers. Convergence analysis of a local stationarity scheme for rate-independent systems and application to damage. Preprint SPP1962-168, 4 2021.

[SKT+15] David Schrade, Marc-André Keip, Huy Thai, Jörg Schröder, Bob Svendsen, Ralf Müller, and Dietmar Gross. Coordinate-invariant phase field modeling of ferroelectrics, part I: Model formulation and single-crystal simulations. *GAMM-Mitt.*, 38(1):102–114, 2015.

[Ste11] Ulisse Stefanelli. Magnetic control of magnetic shape-memory single crystals. *Physica B: Condensed Matter*, 407:9–11, 01 2011.

[Suq81] Pierre-M. Suquet. Sur les équations de la plasticité: existence et régularité des solutions. *J. Mécanique*, 20(1):3–39, 1981.

[SWW17] Ulisse Stefanelli, Daniel Wachsmuth, and Gerd Wachsmuth. Optimal control of a rate-independent evolution equation via viscous regularization. *Discrete Contin. Dyn. Syst. Ser. S*, 10(6):1467–1485, 2017.

[TM10] Marita Thomas and Alexander Mielke. Damage of nonlinearly elastic materials at small strain—existence and regularity results. *ZAMM Z. Angew. Math. Mech.*, 90(2):88–112, 2010.

[Wac12] Gerd Wachsmuth. Optimal control of quasi-static plasticity with linear kinematic hardening, Part I: Existence and discretization in time. *SIAM J. Control Optim.*, 50(5):2836–2861 + loose erratum, 2012.

[Wac15] Gerd Wachsmuth. Optimal control of quasistatic plasticity with linear kinematic hardening II: Regularization and differentiability. *Z. Anal. Anwend.*, 34(4):391–418, 2015.

[Wac16] Gerd Wachsmuth. Optimal control of quasistatic plasticity with linear kinematic hardening III: Optimality conditions. *Z. Anal. Anwend.*, 35(1):81–118, 2016.

[Wlo87] Joseph Wloka. *Partial differential equations*. Cambridge etc.: Cambridge University Press, 1987.