Existence of parameterized BV-solutions for rate-independent systems with discontinuous loads.

EXISTENCE OF PARAMETERIZED BV-SOLUTIONS FOR RATE-INDEPENDENT SYSTEMS WITH DISCONTINUOUS LOADS

DOROTHEE KNEES AND CHIARA ZANINI

Abstract. We study a rate-independent system with non-convex energy and in the case of a time-discontinuous loading. We prove existence of the rate-dependent viscous regularization by time-incremental problems, while the existence of the so called parameterized BV-solutions is obtained via vanishing viscosity in a suitable parameterized setting. In addition, we prove that the solution set is compact.

Contents

1. Introduction 1
2. Basic assumptions and estimates for a time-incremental scheme 3
3. Existence and uniqueness of viscous solutions 9
4. The viscosity limit 12
5. Properties of the solution set 18
Appendix A. Properties of $\mathcal{R}$ 22
Appendix B. Kurzweil integrals and convergence 22
Appendix C. Miscellaneous of useful tools 24
C.1. Lower semicontinuity properties 24
C.2. Absolutely continuous functions and BV-functions 24
C.3. A combination of Helly's Theorem and the Ascoli-Arzelà Theorem 25
C.4. Chain rule 25
Acknowledgments 25
References 25

1. Introduction

In this paper the existence of a solution $z : [0,T] \to \mathcal{Z}$ of a doubly nonlinear problem of the type

$$0 \in \partial\mathcal{R}(\partial_t z(t)) + \mathrm{D}\mathcal{I}(z(t)) - \ell(t), \quad z(0) = z_0, \quad t \in [0,T] \tag{1.1}$$

is addressed. The focus is on rate-independent systems and hence we assume that the dissipation functional $\mathcal{R}$ is convex and positively homogeneous of degree one. It is further assumed that the energy functional $\mathcal{I}$ is nonconvex and that the load term ℓ is discontinuous in time. It is well known that even if ℓ is smooth in time, due to the non-convexity of $\mathcal{I}$ the system in general has solutions that are discontinuous in time and that also in general there is no uniqueness (see [MR15] and references therein). In our setting here, a second source for discontinuities is introduced by the discontinuous load term. We prove the existence of (parameterized) balanced viscosity solutions via a vanishing viscosity analysis (Theorem 4.5) and study the compactness of the solution set

2010 *Mathematics Subject Classification.* Primary: 35R05, 49J40 ; Secondary: 74C05, 35Q74, 35D40, 49J45.
Key words and phrases. rate-independent system; discontinuous load; parameterized BV-solution; time-incremental minimum problems; vanishing viscosity limit.

(Proposition 5.3). The analysis is carried out in the semilinear rate-independent setting introduced in [MZ14], compare also [MR15, Example 3.8.4], [Kne18, KT18].

For a more detailed presentation of the arguments let $\mathcal{Z}$, $\mathcal{V}$ be Hilbert spaces and $\mathcal{X}$ a Banach space such that $\mathcal{Z} \Subset \mathcal{V} \subset \mathcal{X}$ (compact and continuous embeddings, respectively). The dissipation functional $\mathcal{R} : \mathcal{X} \to [0, \infty)$ is convex, continuous and positively homogeneous of degree one and it is assumed to be equivalent to the norm on $\mathcal{X}$. The latter assumption simplifies the analysis since then $\partial\mathcal{R}(0)$ is a bounded subset of $\mathcal{X}^*$. However, this assumption rules out the modeling of damage and other unidirectional processes. We work in the semilinear setting where $\mathcal{I} : \mathcal{Z} \to \mathbb{R}$ is of the structure $\mathcal{I}(z) = \frac{1}{2}\langle Az, z\rangle + \mathcal{F}(z)$ with a linear and continuous operator $A \in \mathrm{Lin}(\mathcal{Z}, \mathcal{Z}^*)$ that is bounded and symmetric (we refer to Section 2 for the precise assumptions) and a possibly nonconvex functional $\mathcal{F} : \mathcal{Z} \to [0, \infty)$ that is of lower order with respect to the quadratic term in $\mathcal{I}$. The loads ℓ are taken from $BV([0, T]; \mathcal{V}^*)$. The total energy is given by $\mathcal{E}(t, z) = \mathcal{I}(z) - \langle \ell(t), z\rangle$. As already mentioned, due to the non-convexity of $\mathcal{I}$ solutions to (1.1) are discontinuous in time (even if ℓ is continuous). Several different notions of weak solutions have been introduced in the recent literature (see [MR15] and references therein) allowing for discontinuous solutions, among them the (global) energetic solutions and balanced viscosity solutions (BV-solutions). Let us remark that the solution concepts are not equivalent. Existence of the different solution concepts was obtained for more regular data, while the novelty in this paper is to consider the case of BV-loading. Existence is studied via vanishing viscosity resulting in BV-solutions. For that purpose, we consider the regularized problem

$$0 \in \partial\mathcal{R}(\partial_t z_\varepsilon(t)) + \varepsilon\mathbb{V}\partial_t z_\varepsilon(t) + \mathrm{D}_z\mathcal{E}(t, z_\varepsilon(t)), \quad z_\varepsilon(0) = z_0, \quad t \in [0, T] \tag{1.2}$$

obtained by adding the viscous term $\varepsilon\mathbb{V}\partial_t z(t)$ ($\mathbb{V}$ is a linear operator) to (1.1) with the parameter $\varepsilon > 0$. After having established the existence and uniqueness of solutions to the regularized problem (Proposition 3.3) we study the limit $\varepsilon \to 0$. In order to perform the vanishing viscosity analysis, the inclusion (1.2) is rewritten in a parameterized version, i.e. $t \mapsto z_\varepsilon(t)$ is replaced with $s \mapsto (\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s))$, where $\hat{z}_\varepsilon(s)) = z_\varepsilon(\hat{t}_\varepsilon(s))$. There are different possibilities for choosing the parameterization. We take here the paramterization based on the vanishing viscosity contact potential ([MRS16], see (4.2)). The advantage of this choice is that viscosity limits automatically are normalized in the parameterized picture (see (4.16)). In the convergence proofs we closely follow the arguments in [MRS16] and adapt them to our situation. Due to the semilinear structure of our problem, some stronger statements in particular concerning the regularity of solutions (e.g. $\mathrm{D}\mathcal{E} \in \mathcal{V}^*$ instead of $\mathcal{Z}^*$) compared to those in [MRS16] are possible. Due to the possible discontinuities of the load term ℓ a refined analysis of the power term $\int_0^t \langle \ell(r), \partial_t z_\varepsilon(r)\rangle \,\mathrm{d}r$ and its reparameterized version is necessary. Observe that in the reparameterized version the function $s \mapsto \ell(\hat{t}_\varepsilon(s))$ appears. Interpreting the power term as a Kurzweil integral the limit $\varepsilon \to 0$ can be identified. We refer to [KL09] (and Appendix B) for an overview on the properties of the Kurzweil integral.

In order to perform the vanishing viscosity analysis, estimates for solutions to (1.2) are needed that are uniform with respect to the viscosity parameter ε. Due to the low regularity of the load term ℓ, arguments from the literature are not directly applicable since there it is typically assumed and used that ℓ has temporal H^1 or C^1-smoothness. The new estimates are stated in Propositions 2.3 and 2.5. As a new feature these estimates do not depend on the length of the time interval $[0, T]$ and the constants in the estimates are scaling invariant. This allows for instance to transfer estimates by rescaling arguments to different time intervals without changing the constants. This observation is exploited in the analysis of solution sets to the system (1.1), see Proposition 5.3.

This is not the first paper that investigates solutions to rate-independent systems with discontinuous loads. Let us first mention the article [KL09] that is closest to our investigations. In contrast to our setting, in [KL09] the energy $\mathcal{E}(t, \cdot)$ is assumed to be strictly convex in z and the dissipation potential $\mathcal{R}$ may depend in a discontinuous way on the time. Starting from a time incremental minimization problem (without adding additional viscosity) the authors prove the existence and uniqueness of solutions within their solution class. In addition, if $\mathcal{E}$ is quadratic, they compare this solution with the one obtained from a vanishing viscosity analysis. The analysis is carried out in the physical time and integrals over time intervals are interpreted in the Kurzweil sense. A different approach was followed in [Rec11, Rec16] based on measure theory tools, and originally was developed for the study of the mapping properties of the play operator, solving variational inequalities associated to sweeping processes [Mor77, KL02]. More precisely, in [Rec11, Rec16] the existence results from [Mor77] are re-obtained for discontinuous BV-loadings by using the following steps: reparameterize suitably the problem by "filling in the jumps of the loading ℓ" in order to obtain a Lipschitz-setting, use the better regularity to get existence of a solution, and then parameterize back to the BV-setting via measure theory arguments (instead of time discretization procedure [Mor77]). This approach works thanks to the fact that sweeping processes are rate-independent. The underlying energies in general are convex but the set of admissible forces is allowed to depend on time in a discontinuous way, [RS18]. Translated to our setting this means that $\mathcal{R}$ in addition depends on the time and that $t \mapsto \mathcal{R}(t, z)$ is of bounded variation. It is shown in [Rec11] that the solution z depends on the parameterization chosen, in the sense that, by using segments (geodesics) to fill in the jumps of ℓ, one may get a solution different from the vanishing viscosity one. We refer to [KR14] for a comparison of the different solution concepts. Clearly, a comparison of the parameterized BV-solutions derived in this paper with the above mentioned results would clarify the relations between all these different approaches. This would require to translate back our solutions to the physical time. Due to the length of this paper we postpone this comparison to a future paper.

The paper is organized as follows: in Sec. 2 the precise assumptions are settled and the basic and enhanced estimates are derived in order to do the limiting analysis. In Sec. 3 we pass to the limit in the time incremental viscous problems (expressed as usual in this context via energy balance) and derive existence and uniqueness of solution for $\varepsilon > 0$ fixed. Then in Sec. 4, to perform the vanishing viscosity analysis $\varepsilon \to 0$ we use the reparameterization technique originally introduced in [EM06] and refined in [MRS16], that is we rewrite the problem in a suitable parameterized setting, see (4.2), and pass to the limit as $\varepsilon \to 0$ in this setting. Finally, in Sec. 5 we discuss the regularity properties and compactness of the set of (p)-parameterized solutions. The paper closes with an appendix where basic facts about the Kurzweil integral, about absolutely continuous functions and BV-functions and a chain rule are collected.

2. Basic assumptions and estimates for a time-incremental scheme

Let $\mathcal{X}$ be a Banach space and $\mathcal{Z}, \mathcal{V}$ be separable Hilbert spaces that are densely and compactly, resp. continuously, embedded in the following way:

$$\mathcal{Z} \Subset \mathcal{V} \subset \mathcal{X}. \tag{2.1}$$

Let further $A \in \mathrm{Lin}(\mathcal{Z}, \mathcal{Z}^*)$ and $\mathbb{V} \in \mathrm{Lin}(\mathcal{V}, \mathcal{V}^*)$ be linear symmetric, bounded $\mathcal{Z}$- and $\mathcal{V}$-elliptic operators, i.e. there exist constants $\alpha, \gamma > 0$ such that

$$\forall z \in \mathcal{Z}, \forall v \in \mathcal{V} : \quad \langle Az, z\rangle \geq \alpha \|z\|_{\mathcal{Z}}^2 \ , \quad \langle \mathbb{V}v, v\rangle \geq \gamma \|v\|_{\mathcal{V}}^2 \ , \tag{2.2}$$

and $\langle Az_1, z_2\rangle = \langle Az_2, z_1\rangle$ for all $z_1, z_2 \in \mathcal{Z}$ (and similar for $\mathbb{V}$). Here, $\langle \cdot, \cdot\rangle$ denotes the duality pairings in $\mathcal{Z}$ and $\mathcal{V}$, respectively. We define $\|v\|_{\mathbb{V}} := (\langle \mathbb{V}v, v\rangle)^{\frac{1}{2}}$, which is a norm that is equivalent to the Hilbert space norm $\|\cdot\|_{\mathcal{V}}$. Let further

$$\mathcal{F} \in C^2(\mathcal{Z}; \mathbb{R}) \text{ with } \mathcal{F} \geq 0. \tag{2.3}$$

The functional $\mathcal{F}$ shall play the role of a possibly nonconvex lower order term (cf. [MR15, Section 3.8]). Hence, we assume that

$$\mathrm{D}\mathcal{F} \in C^1(\mathcal{Z};\mathcal{V}^*), \quad \left\|\mathrm{D}^2\mathcal{F}(z)v\right\|_{\mathcal{V}^*} \leq C(1+\|z\|_{\mathcal{Z}}^q)\,\|v\|_{\mathcal{Z}} \tag{2.4}$$

for some $q \geq 1$. For the load we assume

$$\ell \in BV([0,T];\mathcal{V}^*)\,, \tag{2.5}$$

and

$$\mathrm{Var}_{\mathcal{V}^*}(\ell,[a,b]) = \sup_{\text{partitions } (t_k) \text{ of } [a,b]} \sum_k \|\ell(t_k)-\ell(t_{k-1})\|_{\mathcal{V}^*}$$

denotes the total variation of ℓ on $[a,b]$ with respect to $\mathcal{V}^*$.

Energy functionals of the following type are considered

$$\mathcal{I} : \mathcal{Z} \to \mathbb{R}, \quad \mathcal{I}(z) := \frac{1}{2}\langle Az, z\rangle + \mathcal{F}(z), \tag{2.6}$$

$$\mathcal{E} : [0,T]\times\mathcal{Z} \to \mathbb{R}, \quad \mathcal{E}(t,z) = \mathcal{I}(z) - \langle \ell(t), z\rangle\,. \tag{2.7}$$

Clearly, $\mathcal{I} \in C^1(\mathcal{Z};\mathbb{R})$.

The dissipation functional $\mathcal{R} : \mathcal{X} \to [0,\infty)$ is assumed to be convex, continuous, positively homogeneous of degree one and

$$\exists c, C > 0\, \forall x \in \mathcal{X}: \quad c\,\|x\|_{\mathcal{X}} \leq \mathcal{R}(x) \leq C\,\|x\|_{\mathcal{X}}\,. \tag{2.8}$$

We refer to Appendix A for the properties of $\mathcal{R}$ which will be used in the following. From (2.4) and (2.8) we deduce the following interpolation estimate, [Kne18, Lemma 1.1]:

Lemma 2.1. *Assume (2.1), (2.3), (2.8) and (2.4). For every $\rho > 0$ and $\kappa > 0$ there exists $C_{\rho,\kappa} > 0$ such that for all $z_1, z_2 \in \mathcal{Z}$ with $\|z_i\|_{\mathcal{Z}} \leq \rho$ we have*

$$\begin{aligned} &|\langle \mathrm{D}\mathcal{F}(z_1) - \mathrm{D}\mathcal{F}(z_2), z_1 - z_2\rangle| \\ &\qquad \leq \kappa\,\|z_1-z_2\|_{\mathcal{Z}}^2 + C_{\rho,\kappa}\min\{\mathcal{R}(z_1-z_2),\mathcal{R}(z_2-z_1)\}\,\|z_1-z_2\|_{\mathbb{V}}\,. \end{aligned} \tag{2.9}$$

As a consequence, $\mathcal{E}$ is λ-convex on sublevels. To be more precise, we have the following estimate: For every $\rho > 0$ there exists $\lambda = \lambda(\rho) > 0$ such that for all $t \in [0,T]$ and all $z_1, z_2 \in \mathcal{Z}$ with $\|z_i\|_{\mathcal{Z}} \leq \rho$ we have

$$\langle \mathrm{D}_z\mathcal{E}(t,z_1) - \mathrm{D}_z\mathcal{E}(t,z_2), z_1-z_2\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \tfrac{\alpha}{2}\,\|z_1-z_2\|_{\mathcal{Z}}^2 - \lambda\,\|z_1-z_2\|_{\mathbb{V}}^2 \tag{2.10}$$

and

$$\mathcal{I}(z_2) - \mathcal{I}(z_1) \geq \langle \mathrm{D}\mathcal{I}(z_1), z_2-z_1\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \tfrac{\alpha}{2}\,\|z_1-z_2\|_{\mathcal{Z}}^2 - \lambda\mathcal{R}(z_2-z_1)\,\|z_2-z_1\|_{\mathcal{V}}\,. \tag{2.11}$$

In the following we replace $\mathrm{D}_z\mathcal{E}(t,z)$ by $\mathrm{D}\mathcal{E}(t,z)$ so that

$$\mathrm{D}\mathcal{E}(t,z) = \mathrm{D}\mathcal{I}(z) - \ell(t) = Az + \mathrm{D}\mathcal{F}(z) - \ell(t).$$

For the proof of the existence theorems we need a further assumption on $\mathcal{F}$:

$$\mathcal{F} : \mathcal{Z} \to \mathbb{R} \text{ and } \mathrm{D}\mathcal{F} : \mathcal{Z} \to \mathcal{Z}^* \text{ are weak-weak continuous.} \tag{2.12}$$

In the next lemma we prove a coercivity estimate for $\mathcal{E}$ and a product estimate which will be used to derive a uniform estimate on $\left\|z_k^N\right\|_{\mathcal{Z}}$, see Proposition 2.3 below. Similar arguments were used in the proof of [KL09, Lemma 3.1].

Lemma 2.2. *Assume (2.1)–(2.5).*

Let $c_0 := \frac{c_{\mathcal{Z}}^2}{\alpha}(1+\|\ell\|_{L^\infty(0,T;\mathcal{V}^*)}^2)$, 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

$$\mathcal{E}(t,v) + c_0 \geq c_{\mathcal{Z}}\,\|v\|_{\mathcal{Z}} \geq \|v\|_{\mathcal{V}}\,. \tag{2.13}$$

A product estimate: Let $\{ a_k \,;\, 1 \leq k \leq N \}$ *with* $a_k \geq 0$ *for every* k*, 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 $c > 0$*,* $\ell \in BV([0,T]; \mathcal{V}^*)$ *and let* $0 \leq t_0 < t_1 < \ldots < t_N \leq T$ *be an arbitrary partition of* $[0,T]$*. Then*

$$\prod_{k=1}^{N}(1 + c\left\|\ell(t_k) - \ell(t_{k-1})\right\|_{\mathcal{V}^*}) \leq \exp\left(c\operatorname{Var}_{\mathcal{V}^*}(\ell, [t_0, t_N])\right) . \tag{2.14}$$

Proof. Let $t \in [0,T]$, $v \in \mathcal{Z}$. By coercivity and Young's inequality

$$\mathcal{E}(t,v) \geq \tfrac{\alpha}{2}\left\|v\right\|_{\mathcal{Z}}^2 - c_{\mathcal{Z}}\left\|\ell(t)\right\|_{\mathcal{V}^*}\left\|v\right\|_{\mathcal{Z}} \geq \tfrac{\alpha}{4}\left\|v\right\|_{\mathcal{Z}}^2 - \tfrac{c_{\mathcal{Z}}^2}{\alpha}\left\|\ell\right\|_{L^\infty(0,T;\mathcal{V}^*)}^2 .$$

Together with $\left\|v\right\|_{\mathcal{V}} \leq c_{\mathcal{Z}}\left\|v\right\|_{\mathcal{Z}} \leq \frac{c_{\mathcal{Z}}^2}{\alpha} + \frac{\alpha}{4}\left\|v\right\|_{\mathcal{Z}}^2$ one obtains (2.13).

Proof of the product estimate: Since for $y \geq 0$ we have $\ln(1+y) \leq y$, it holds

$$\prod_{k=1}^{N}(1 + ca_k) = \exp(\sum_{k=1}^{N}\ln(1 + ca_k)) \leq \exp(c\sum_{k=1}^{N} a_k) .$$

□

We consider viscous regularizations of the rate-independent system $(\mathcal{E}, \mathcal{R}, \mathcal{Z})$ with respect to the intermediate space $\mathcal{V}$. For $\varepsilon \geq 0$ let

$$\mathcal{R}_\varepsilon : \mathcal{V} \to [0,\infty), \quad \mathcal{R}_\varepsilon(v) := \mathcal{R}(v) + \tfrac{\varepsilon}{2}\langle \mathbb{V}v, v\rangle .$$

Properties about $\mathcal{R}_\varepsilon$, $\varepsilon \geq 0$, are collected in the Appendix A.

We start from the usual time-incremental minimization problems: Let $0 = t_0 < t_1 < \ldots < t_N = T$ be an arbitrary partition of $[0,T]$ and let $\tau_k := t_k - t_{k-1}$, for $k = 1, \ldots, N$. With $z_0^N := z_0$, for $k = 1, \ldots, N$ define z_k^N recursively via

$$z_k^N \in \operatorname{Argmin}\{\, \mathcal{E}(t_k, v) + \tau_k\mathcal{R}_\varepsilon\left((v - z_{k-1}^N)/\tau_k\right) \,;\, v \in \mathcal{Z} \,\} . \tag{2.15}$$

Minimizers exist by the direct method in the calculus of variations. In the next proposition we collect the basic estimates for the time-incremental minimization problems.

Proposition 2.3. *Under the above conditions on* $\mathcal{E}$ *and* $\mathcal{R}_\varepsilon$ *there exists a constant* $C > 0$ *such that for all* $\varepsilon \geq 0$*,* $N \in \mathbb{N}$ *and* $1 \leq k \leq N$ *we have, with* c_0 *from Lemma 2.2,*

$$\left\|z_k^N\right\|_{\mathcal{Z}} \leq c_{\mathcal{Z}}^{-1}(\mathcal{E}(0,z_0) + c_0)\exp(\operatorname{Var}_{\mathcal{V}^*}(\ell, [0,t_k])) , \tag{2.16}$$

$$0 \leq c_0 + \mathcal{E}(t_k, z_k^N) \leq (\mathcal{E}(0,z_0) + c_0)\exp(\operatorname{Var}_{\mathcal{V}^*}(\ell, [0,t_k])) , \tag{2.17}$$

$$\sum_{s=1}^{N}\tau_s\mathcal{R}_\varepsilon((z_s^N - z_{s-1}^N)/\tau_s) \leq \widetilde{C} \tag{2.18}$$

with $\widetilde{C} = (\mathcal{E}(0,z_0) + c_0)\Big(1 + \operatorname{Var}_{\mathcal{V}^*}(\ell, [0,T])\exp(\operatorname{Var}_{\mathcal{V}^*}(\ell, [0,T]))\Big)$*. The following energy-dissipation estimates are valid*

$$\mathcal{E}(t_k, z_k^N) + \sum_{s=1}^{k}\tau_s\mathcal{R}_\varepsilon((z_s^N - z_{s-1}^N)/\tau_s) \leq \mathcal{E}(t_0, z_0) + \sum_{s=1}^{k}\langle \ell(t_{s-1}) - \ell(t_s), z_s^N\rangle_{\mathcal{V}^*,\mathcal{V}} , \tag{2.19}$$

$$\mathcal{I}(z_k^N) + \sum_{s=1}^{k}\tau_s\mathcal{R}_\varepsilon((z_s^N - z_{s-1}^N)/\tau_s) \leq \mathcal{I}(z_0) + \sum_{s=0}^{k-1}\langle \ell(t_s), z_{s+1}^N - z_s^N\rangle_{\mathcal{V}^*,\mathcal{V}} . \tag{2.20}$$

Proof. By minimality, we obtain from (2.15) (suppressing the index N) together with (2.13)

$$\begin{aligned}\mathcal{E}(t_k,z_k)+\tau_k\mathcal{R}_\varepsilon((z_k-z_{k-1})/\tau_k) &\le \mathcal{E}(t_{k-1},z_{k-1})+\langle \ell(t_{k-1})-\ell(t_k),z_{k-1}\rangle \\ &\le \mathcal{E}(t_{k-1},z_{k-1})+\|\ell(t_{k-1})-\ell(t_k)\|_{\mathcal{V}^*}\,\|z_{k-1}\|_{\mathcal{V}} \\ &\le \mathcal{E}(t_{k-1},z_{k-1})+\|\ell(t_{k-1})-\ell(t_k)\|_{\mathcal{V}^*}\left(c_0+\mathcal{E}(t_{k-1},z_{k-1})\right).\end{aligned} \tag{2.21}$$

Adding c_0 on both sides yields

$$\mathcal{E}(t_k,z_k)+c_0\le \big(\mathcal{E}(t_{k-1},z_{k-1})+c_0\big)(1+\|\ell(t_{k-1})-\ell(t_k)\|_{\mathcal{V}^*}),$$

and by recursion and (2.14)

$$\begin{aligned}\mathcal{E}(t_k,z_k)+c_0 &\le (\mathcal{E}(t_0,z_0)+c_0)\prod_{s=1}^{k}(1+\|\ell(t_s)-\ell(t_{s-1})\|_{\mathcal{V}^*}) \\ &\le (\mathcal{E}(t_0,z_0)+c_0)\exp(\mathrm{Var}_{\mathcal{V}^*}(\ell,[0,t_k])).\end{aligned}$$

Together with (2.13) we arrive at (2.16) and (2.17). The energy dissipation estimate (2.19) follows from (2.21), again by recursion, while estimate (2.20) is nothing else but a consequence of discrete integration by parts in the power term. Since

$$\left|\sum_{s=1}^{k}\langle \ell(t_{s-1})-\ell(t_s),z_s\rangle_{\mathcal{V}^*,\mathcal{V}}\right|\le c_{\mathcal{Z}}\,\mathrm{Var}_{\mathcal{V}^*}(\ell,[0,T])\sup_k\|z_k\|_{\mathcal{Z}},$$

from (2.19) and (2.13) (i.e. $\mathcal{E}(t_k,z_k)\ge -c_0$) we finally obtain (2.18). $\square$

Remark 2.4. Let $\triangle_N := \max\{\,t_k-t_{k-1}\,;\,1\le k\le N\,\}$ denote the fineness of the partition of $[0,T]$. There exists $m>0$ such that the minimizers z_k^N of (2.15) are unique provided that $\varepsilon>m\triangle_N$. Indeed, by (2.16) the minimizers z_k^N are uniformly bounded with respect to $\varepsilon\ge 0$ and the partitions of $[0,T]$, and they satisfy the inclusion $0\in\partial\mathcal{R}(z_k^N-z_{k-1}^N)+\frac{\varepsilon}{\tau_k}\mathbb{V}(z_k^N-z_{k-1}^N)+\mathrm{D}\mathcal{E}(t_k^N,z_k^N)$. The maximal monotonicity of $\partial\mathcal{R}$ in combination with estimate (2.10) implies uniqueness provided that $\varepsilon/\triangle_N>\lambda$ with λ from (2.10).

In order to carry out the vanishing viscosity analysis we need more refined estimates. In the following $\mathrm{dist}_{\mathbb{V}}(\cdot,\partial\mathcal{R}(0))$ denotes the distance of an element of $\mathcal{V}^*$ to $\partial\mathcal{R}(0)\subset\mathcal{V}^*$, see (A.1).

Proposition 2.5. *Assume* (2.1)–(2.8). *Assume in addition that* $\mathrm{D}\mathcal{E}(0,z_0)\in\mathcal{V}^*$. *Then for all* $\varepsilon\ge 0$, *all* $N\in\mathbb{N}$ *and all partitions* Π_N *of* $[0,T]$ *we have*

$$\sum_{k=1}^{N}\left\|z_k^N-z_{k-1}^N\right\|_{\mathcal{Z}}+\sup_{1\le k\le N}\tfrac{\varepsilon}{\tau_k^N}\left\|z_k^N-z_{k-1}^N\right\|_{\mathbb{V}}\le C_1 \tag{2.22}$$

$$\sup_{1\le k\le N}\left\|\mathrm{D}\mathcal{E}(t_k^N,z_k^N)\right\|_{\mathcal{V}^*}\le \mathrm{diam}_{\mathcal{V}^*}(\partial\mathcal{R}(0))+C_1, \tag{2.23}$$

where $C_1=\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(0,z_0),\partial\mathcal{R}(0))+c_{\mathbb{V}}\,\mathrm{Var}_{\mathcal{V}^*}(\ell,[0,T])+C_I\widetilde{C}$ *with* $\widetilde{C}$ *from* (2.18) *and* $C_I=C_{\rho,\kappa}$ *from* (2.9) *for* $\kappa=\alpha/2$ *and* ρ *is the right hand side of* (2.16). *Finally, for every* $\varepsilon>0$ *there exists a constant* $C_\varepsilon>0$ *such that for all partitions* Π_N *we have*

$$\sum_{k=1}^{N}\tau_k\left\|\frac{z_k^N-z_{k-1}^N}{\tau_k}\right\|_{\mathcal{Z}}^2\le C_\varepsilon. \tag{2.24}$$

Remark 2.6. Observe first that the constants C_1, C_I and $\widetilde{C}$ are independent of the partition Π_N and of $\varepsilon>0$. Observe further that the constants appearing in (2.22)–(2.23) are invariant with respect to a rescaling in time.

Proof of Proposition 2.5. Choose a partition Π_N of $[0,T]$ and $\varepsilon \geq 0$. Let $\{ z_k ; 1 \leq k \leq N \}$ be minimizers according to (2.15) (we omit the index N). Then for all $1 \leq k \leq N$ we have

$$-\xi_k := -\tfrac{\varepsilon}{\tau_k}\mathbb{V}(z_k - z_{k-1}) - \mathrm{D}\mathcal{E}(t_k, z_k) \in \partial\mathcal{R}(z_k - z_{k-1})\,. \tag{2.25}$$

Due to the convexity and one-homogeneity of $\mathcal{R}$ we obtain $-\mathcal{R}(z_k - z_{k-1}) = \langle \xi_k, z_k - z_{k-1}\rangle$ and $\mathcal{R}(z_k - z_{k-1}) \geq \langle -\xi_{k-1}, z_k - z_{k-1}\rangle$, see Appendix A. Hence, after adding these relations and rearranging the terms, for $2 \leq k \leq N$ we arrive at

$$\begin{aligned}\tfrac{\varepsilon}{\tau_k}\left\|z_k - z_{k-1}\right\|_{\mathbb{V}}^2 - \tfrac{\varepsilon}{\tau_{k-1}}\langle \mathbb{V}(z_{k-1} - z_{k-2}), z_k - z_{k-1}\rangle + \langle A(z_k - z_{k-1}), z_k - z_{k-1}\rangle \\ \leq \langle \mathrm{D}\mathcal{F}(z_{k-1}) - \mathrm{D}\mathcal{F}(z_k), z_k - z_{k-1}\rangle + \langle \ell(t_k) - \ell(t_{k-1}), z_k - z_{k-1}\rangle.\end{aligned} \tag{2.26}$$

The left hand side can be estimated as

$$\text{l.h.s.} \geq \Big(\tfrac{\varepsilon}{\tau_k}\left\|z_k - z_{k-1}\right\|_{\mathbb{V}} - \tfrac{\varepsilon}{\tau_{k-1}}\left\|z_{k-1} - z_{k-2}\right\|_{\mathbb{V}}\Big)\left\|z_k - z_{k-1}\right\|_{\mathbb{V}} + \alpha\left\|z_k - z_{k-1}\right\|_{\mathcal{Z}}^2 ,$$

where $\alpha > 0$ is the constant from (2.2). For the right hand side we deduce from Lemma 2.1 (where we choose $\kappa = \frac{\alpha}{2}$ and ρ according to the right hand side in (2.16)) that

$$\text{r.h.s.} \leq \tfrac{\alpha}{2}\left\|z_k - z_{k-1}\right\|_{\mathcal{Z}}^2 + C(\mathcal{R}(z_k - z_{k-1}) + \left\|\ell(t_k) - \ell(t_{k-1})\right\|_{\mathcal{V}^*})\left\|z_k - z_{k-1}\right\|_{\mathbb{V}} .$$

Observe that $C > 0$ is independent of ε and of the partition of $[0,T]$. Joining both inequalities we obtain for all $k \in \{2, \ldots, N\}$

$$\begin{aligned}\tfrac{\varepsilon}{\tau_k}\left\|z_k - z_{k-1}\right\|_{\mathbb{V}} + \tfrac{\alpha}{2c_{\mathcal{Z}}}\left\|z_k - z_{k-1}\right\|_{\mathcal{Z}} \\ \leq \tfrac{\varepsilon}{\tau_{k-1}}\left\|z_{k-1} - z_{k-2}\right\|_{\mathbb{V}} + C(\mathcal{R}(z_k - z_{k-1}) + \left\|\ell(t_k) - \ell(t_{k-1})\right\|_{\mathcal{V}^*}),\end{aligned}$$

where $c_{\mathcal{Z}} > 0$ is the embedding constant for $\mathcal{Z} \subset \mathcal{V}$. Summation with respect to k finally yields (for $2 \leq K \leq N$)

$$\begin{aligned}\tfrac{\varepsilon}{\tau_K}\left\|z_K - z_{K-1}\right\|_{\mathbb{V}} + \tfrac{c\alpha}{2}\sum_{k=2}^{K}\left\|z_k - z_{k-1}\right\|_{\mathcal{Z}} \\ \leq \tfrac{\varepsilon}{\tau_1}\left\|z_1 - z_0\right\|_{\mathbb{V}} + C\sum_{k=2}^{K}(\mathcal{R}(z_k - z_{k-1}) + \left\|\ell(t_k) - \ell(t_{k-1})\right\|_{\mathcal{V}^*})\end{aligned} \tag{2.27}$$

Let now $k = 1$. Choose $\mu \in \partial\mathcal{R}(0)$ such that

$$\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(0, z_0), \partial\mathcal{R}(0)) = \left\|\mu + \mathrm{D}\mathcal{E}(0, z_0)\right\|_{\mathbb{V}} .$$

Together with (2.25) (for $k = 1$) and from the one-homogeneity of $\mathcal{R}$ we obtain

$$\begin{aligned}0 &\geq \langle \mathrm{D}\mathcal{E}(t_1, z_1) + \mu, z_1 - z_0\rangle + \tfrac{\varepsilon}{\tau_1}\langle \mathbb{V}(z_1 - z_0), (z_1 - z_0)\rangle \\ &= \langle \mathrm{D}\mathcal{E}(0, z_0) + \mu, z_1 - z_0\rangle + \langle \mathrm{D}\mathcal{E}(t_1, z_1) - \mathrm{D}\mathcal{E}(0, z_0), z_1 - z_0\rangle + \tfrac{\varepsilon}{\tau_1}\left\|z_1 - z_0\right\|_{\mathbb{V}}^2 .\end{aligned}$$

By the structure of $\mathrm{D}\mathcal{E}$ and after rearranging the terms we obtain

$$\begin{aligned}&\tfrac{\varepsilon}{\tau_1}\left\|z_1 - z_0\right\|_{\mathbb{V}}^2 + \alpha\left\|z_1 - z_0\right\|_{\mathcal{Z}}^2 \\ &\leq -\langle \mathrm{D}\mathcal{E}(0, z_0) + \mu, z_1 - z_0\rangle + \langle (\mathrm{D}\mathcal{F}(z_0) - \mathrm{D}\mathcal{F}(z_1)) + (\ell(t_1) - \ell(t_0)), z_1 - z_0\rangle \\ &\leq \tfrac{\alpha}{2}\left\|z_1 - z_0\right\|_{\mathcal{Z}}^2 + \Big(\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(0, z_0), \partial\mathcal{R}(0)) + \left\|\ell(t_1) - \ell(t_0)\right\|_{\mathbb{V}^*} \\ &\qquad\qquad + C\mathcal{R}(z_1 - z_0)\Big)\left\|z_1 - z_0\right\|_{\mathbb{V}} .\end{aligned} \tag{2.28}$$

For the last estimate we used the definition of μ and similar estimates as for the case $k \geq 2$. Similar to the case $k \geq 2$ we further obtain

$$\begin{aligned}\tfrac{\varepsilon}{\tau_1}\left\|z_1 - z_0\right\|_{\mathbb{V}} + \tfrac{\alpha}{2c_{\mathcal{Z}}}\left\|z_1 - z_0\right\|_{\mathcal{Z}} \\ \leq \mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(0, z_0), \partial\mathcal{R}(0)) + C\big(\left\|\ell(t_1) - \ell(t_0)\right\|_{\mathcal{V}^*} + \mathcal{R}(z_1 - z_0)\big) .\end{aligned}$$

Adding the last estimate to (2.27) finally results in

$$\begin{aligned}\tfrac{\varepsilon}{\tau_K}\left\|z_K-z_{K-1}\right\|_{\mathbb{V}}&+\tfrac{\alpha}{2c_{\mathcal{Z}}}\sum_{k=1}^{K}\left\|z_k-z_{k-1}\right\|_{\mathcal{Z}}\\&\leq \operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(0,z_0),\partial\mathcal{R}(0))+C\operatorname{Var}_{\mathcal{V}^*}(\ell,[0,t_K])+C\sum_{k=1}^{K}\mathcal{R}(z_k-z_{k-1})\,,\end{aligned}\tag{2.29}$$

which is valid for $1\leq K\leq N$. Thanks to Proposition 2.3 the right hand side is uniformly bounded with respect to $\varepsilon\geq 0$ and the partitions of $[0,T]$ and we have shown estimate (2.22).

In order to prove (2.23) observe that $\partial\mathcal{R}(0)\subset\mathcal{Z}^*$ can be identified with a subset of $\mathcal{V}^*$ that is bounded with respect to the $\mathcal{V}^*$-norm, see [Kne18, Lemma A.1]. Hence, for $k\geq 1$ from (2.25) we conclude $-\mathrm{D}\mathcal{E}(t_k,z_k)\in\partial\mathcal{R}(0)+\frac{\varepsilon}{\tau_k}\mathbb{V}(z_k-z_{k-1})\subset\mathcal{V}^*$ and thus by (2.22) we ultimately arrive at (2.23).

For the proof of (2.24) we start again from (2.26). Using $2a(a-b)=a^2-b^2+(a-b)^2$, for the first two terms we obtain after dividing by τ_k for $k\geq 2$

$$\begin{aligned}\tfrac{\varepsilon}{2}\left\|\tfrac{z_k-z_{k-1}}{\tau_k}\right\|_{\mathbb{V}}^2&+\tfrac{\varepsilon}{2}\left\|\left(\tfrac{z_k-z_{k-1}}{\tau_k}\right)-\left(\tfrac{z_{k-1}-z_{k-2}}{\tau_{k-1}}\right)\right\|_{\mathbb{V}}^2+\alpha\tau_k\left\|\tfrac{z_k-z_{k-1}}{\tau_k}\right\|_{\mathcal{Z}}^2\\&\leq\tfrac{\varepsilon}{2}\left\|\tfrac{z_{k-1}-z_{k-2}}{\tau_{k-1}}\right\|_{\mathbb{V}}^2+\langle\mathrm{D}\mathcal{F}(z_{k-1})-\mathrm{D}\mathcal{F}(z_k),\tfrac{z_k-z_{k-1}}{\tau_k}\rangle+\langle\ell(t_k)-\ell(t_{k-1}),\tfrac{z_k-z_{k-1}}{\tau_k}\rangle.\end{aligned}$$

Summation with respect to $2\leq k\leq N$ and adding $(\tau_1^{-1}*(2.28))$ yields

$$\begin{aligned}\tfrac{\varepsilon}{2}\left\|\tfrac{z_N-z_{N-1}}{\tau_N}\right\|_{\mathbb{V}}^2&+\tfrac{\varepsilon}{2}\left\|\tfrac{z_1-z_0}{\tau_1}\right\|_{\mathbb{V}}^2\\&+\tfrac{\varepsilon}{2}\sum_{k=2}^{N}\left\|\left(\tfrac{z_k-z_{k-1}}{\tau_k}\right)-\left(\tfrac{z_{k-1}-z_{k-2}}{\tau_{k-1}}\right)\right\|_{\mathbb{V}}^2+\alpha\sum_{k=1}^{N}\tau_k\left\|\tfrac{z_k-z_{k-1}}{\tau_k}\right\|_{\mathcal{Z}}^2\\&\leq-\langle\mathrm{D}\mathcal{E}(0,z_0)+\mu,\tfrac{z_1-z_0}{\tau_1}\rangle\\&\quad+\sum_{k=1}^{N}\langle\mathrm{D}\mathcal{F}(z_{k-1})-\mathrm{D}\mathcal{F}(z_k),\tfrac{z_k-z_{k-1}}{\tau_k}\rangle+\langle\ell(t_k)-\ell(t_{k-1}),\tfrac{z_k-z_{k-1}}{\tau_k}\rangle\\&=:T_0+T_1+T_2.\end{aligned}\tag{2.30}$$

Clearly, $|T_0|\leq\operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(0,z_0),\partial\mathcal{R}(0))\left\|(z_1-z_0)/\tau_1\right\|_{\mathbb{V}}$. With (2.9) and (2.8), the term T_1 is estimated as

$$|T_1|\leq\tfrac{\alpha}{2}\sum_{k=0}^{N}\tfrac{1}{\tau_k}\left\|z_k-z_{k-1}\right\|_{\mathcal{Z}}^2+C_\alpha\sum_{k=0}^{N}\tfrac{1}{\tau_k}\left\|z_k-z_{k-1}\right\|_{\mathcal{V}}^2.$$

In the term T_2 we shift once more the indices and obtain

$$\begin{aligned}|T_2|&\leq\left|\langle\ell(t_N),\tfrac{z_N-z_{N-1}}{\tau_N}\rangle\right|+\left|\langle\ell(t_1),\tfrac{z_1-z_0}{\tau_1}\rangle\right|+\sum_{k=1}^{N-1}\left|\langle\ell(t_k),\tfrac{z_k-z_{k-1}}{\tau_k}-\tfrac{z_{k+1}-z_k}{\tau_{k+1}}\rangle\right|\\&\leq\tfrac{\varepsilon}{4}\left(\left\|\tfrac{z_N-z_{N-1}}{\tau_N}\right\|_{\mathbb{V}}^2+\left\|\tfrac{z_1-z_0}{\tau_1}\right\|_{\mathbb{V}}^2+\sum_{k=1}^{N}\left\|\left(\tfrac{z_k-z_{k-1}}{\tau_k}\right)-\left(\tfrac{z_{k-1}-z_{k-2}}{\tau_{k-1}}\right)\right\|_{\mathbb{V}}^2\right)+C_\varepsilon\left\|\ell\right\|_{L^\infty(0,T;\mathcal{V}^*)}^2,\end{aligned}$$

where in the last line we applied the Young inequality. Inserting these estimates into (2.30), rearranging the terms and neglecting some nonnegative terms on the left hand side we finally arrive at

$$\frac{\alpha}{2}\sum_{k=0}^{N}\tau_k\left\|\tfrac{z_{k-1}-z_{k-2}}{\tau_k}\right\|_{\mathcal{Z}}^2 \\ \leq C_\varepsilon\left(\operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(0,z_0);\partial\mathcal{R}(0))+\|\ell\|_{L^\infty((0,T);\mathcal{V}^*)}\right)^2+C_\alpha\sum_{k=0}^{N}\tfrac{1}{\tau_k}\|z_k-z_{k-1}\|_{\mathcal{V}}^2 .$$

By (2.18), the last term on the right hand side is bounded by $C\varepsilon^{-1}$, uniformly in N. This proves (2.24). □

3. Existence and uniqueness of viscous solutions

The aim of this section is to prove the existence of solutions to the following system for $\varepsilon > 0$ and given initial value $z_0 \in \mathcal{Z}$:

$$0 \in \partial\mathcal{R}_\varepsilon(\dot z(t)) + \mathrm{D}\mathcal{E}(t, z(t)), \quad z(0) = z_0 . \tag{3.1}$$

Definition 3.1. Let $\varepsilon > 0$, $\ell \in BV([0,T];\mathcal{V}^*)$, $z_0 \in \mathcal{Z}$. A function $z \in H^1([0,T];\mathcal{V}) \cap L^\infty((0,T);\mathcal{Z})$ is a weak solution to (3.1) if $z(0) = z_0$ and if the inclusion (3.1) is satisfied for almost all $t \in [0,T]$.

As is common in the study of rate independent systems it is more convenient to work with an equivalent formulation, namely De Giorgi's energy dissipation principle.

Lemma 3.2. *Let $z \in H^1([0,T];\mathcal{V}) \cap L^\infty((0,T);\mathcal{Z})$ with $z(0) = z_0 \in \mathcal{Z}$. The following properties are equivalent:*

(a) z is a weak solution to (3.1) in the sense of Definition 3.1.
(b) For all $t \in [0,T]$ we have

$$\mathcal{I}(z(t)) + \int_0^t \mathcal{R}_\varepsilon(\dot z(s)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}\mathcal{E}(s,z(s)))\,\mathrm{d}s = \mathcal{I}(z_0) + \int_0^t \langle \ell(s), \dot z(s)\rangle\,\mathrm{d}s. \tag{3.2}$$

(c) For all $t \in [0,T]$ we have

$$\mathcal{I}(z(t)) + \int_0^t \mathcal{R}_\varepsilon(\dot z(s)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}\mathcal{E}(s,z(s)))\,\mathrm{d}s \leq \mathcal{I}(z_0) + \int_0^t \langle \ell(s), \dot z(s)\rangle\,\mathrm{d}s. \tag{3.3}$$

If z satisfies any of these properties then $Az \in L^\infty((0,T);\mathcal{Z}^*) \cap L^2((0,T);\mathcal{V}^*)$.

Proof. The proof follows standard arguments relying on convex analysis and the chain rules provided in Proposition C.6, see e.g. [KT18, Proposition E.1].

Indeed, let z be a weak solution to (3.1). The fact that $\partial\mathcal{R}(0)$ can be identified with a subset of $\mathcal{V}^*$ that is bounded with respect to the norm in $\mathcal{V}^*$, and the assumptions on $\mathcal{F}$ and ℓ imply that $Az \in L^\infty((0,T);\mathcal{Z}^*) \cap L^2((0,T);\mathcal{V}^*)$. Convex analysis arguments and the chain rule provided in Proposition C.6 yield the identity

$$\mathcal{R}_\varepsilon(\dot z(t)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}\mathcal{E}(t,z(t))) \\ = \langle -\mathrm{D}\mathcal{I}(z(t)), \dot z(t)\rangle_{\mathcal{V}^*,\mathcal{V}} + \langle \ell(t), \dot z(t)\rangle_{\mathcal{V}^*,\mathcal{V}} = -\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}(z(t)) + \langle \ell(t), \dot z(t)\rangle_{\mathcal{V}^*,\mathcal{V}}$$

that is valid for almost all t. Integration with respect to t implies (3.2). From this, (3.3) is an obvious consequence.

Assume now that z satisfies (3.3). Since $\int_0^T \mathcal{R}_\varepsilon^*(-\mathrm{D}\mathcal{E}(r,z(r)))\,\mathrm{d}r < \infty$, it follows that $\mathrm{D}\mathcal{E}(\cdot, z(\cdot)) \in L^2(0,T;\mathcal{V}^*)$ and in particular that $Az \in L^\infty((0,T);\mathcal{Z}^*) \cap L^2((0,T);\mathcal{V}^*)$. By the Fenchel inequality and the chain rule we deduce

$$\int_0^t \langle -\mathrm{D}\mathcal{E}(s,z(s)),\dot z(s)\rangle\,\mathrm{d}s \le \int_0^t \mathcal{R}_\varepsilon(\dot z(s)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}\mathcal{E}(s,z(s)))\,\mathrm{d}s$$
$$\overset{(3.3)}{\le} \mathcal{I}(z_0) - \mathcal{I}(z(t)) + \int_0^t \langle \ell(s),\dot z(s)\rangle\,\mathrm{d}s = \int_0^t \big(-\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}(z(s))\big) + \langle \ell(s),\dot z(s)\rangle\,\mathrm{d}s.$$

Hence, (3.2) is valid. Localizing the integral identity and using once more the tools from convex analysis finally shows that z is a weak solution. □

For $\ell \in BV([0,T];\mathcal{V}^*)$ let ℓ_- and ℓ_+ denote the left and the right continuous representative. The identity (3.2) reveals that the weak solutions of (3.1) for ℓ are also weak solutions for ℓ_+ and ℓ_-.

Proposition 3.3. *Assume (2.1)–(2.8). For every $\ell \in BV([0,T];\mathcal{V}^*)$, $z_0 \in \mathcal{Z}$ and $\varepsilon > 0$ there exists a unique weak solution z_ε of (3.1). This solution coincides with the weak solutions for ℓ_+ and ℓ_-. Moreover, $\sup_{\varepsilon>0}\|z_\varepsilon\|_{L^\infty((0,T);\mathcal{Z})} < \infty$.*

If in addition we assume that $\mathrm{D}\mathcal{E}(0,z_0) \in \mathcal{V}^*$, then the weak solution belongs to $H^1((0,T);\mathcal{Z})$ and there exists a constant $C > 0$ such that for all $\varepsilon > 0$ the corresponding weak solution satisfies

$$\|z_\varepsilon\|_{W^{1,1}((0,T);\mathcal{Z})} + \varepsilon\,\|\dot z_\varepsilon\|_{L^\infty(0,T;\mathcal{V})} + \|\mathrm{D}\mathcal{E}(\cdot,z_\varepsilon)\|_{L^\infty((0,T);\mathcal{V}^*)} \le C. \tag{3.4}$$

Remark 3.4. The constant in (3.4) has the same structure as the constants in (2.22)–(2.23).

Proof of Proposition 3.3. Uniqueness of weak solutions:

For $i \in \{1,2\}$ let $\ell_i \in \{\ell,\ell_+,\ell_-\}$ and let z_i be a weak solution for (3.1) corresponding to ℓ_i with $z_i(0) = z_0$. Since $\partial\mathcal{R}$ is maximal monotone, the inclusion (3.1) implies

$$\langle A(z_1(t)-z_2(t)),\dot z_1(t)-\dot z_2(t)\rangle_{\mathcal{V}^*,\mathcal{V}} + \varepsilon\,\|\dot z_1(t)-\dot z_2(t)\|_{\mathcal{V}}^2$$
$$\le \langle \mathrm{D}\mathcal{F}(z_2(t)) - \mathrm{D}\mathcal{F}(z_1(t)) + (\ell_1(t)-\ell_2(t)),\dot z_1(t)-\dot z_2(t)\rangle_{\mathcal{V}^*,\mathcal{V}},$$

which is valid for almost all $t \in [0,T]$. Integration with respect to t yields

$$\tfrac{\alpha}{2}\,\|z_1(t)-z_2(t)\|_{\mathcal{Z}}^2 + \varepsilon\int_0^t \|\dot z_1(s)-\dot z_2(s)\|_{\mathcal{V}}^2\,\mathrm{d}s$$
$$\le \tfrac{\alpha}{2}\,\|z_1(0)-z_2(0)\|_{\mathcal{Z}}^2 + \int_0^t \langle \mathrm{D}\mathcal{F}(z_2(s)) - \mathrm{D}\mathcal{F}(z_1(s)),\dot z_1(s)-\dot z_2(s)\rangle\,\mathrm{d}s$$
$$+ \int_0^t \|\ell_1(s)-\ell_2(s)\|_{\mathcal{V}^*}\,\|\dot z_1(s)-\dot z_2(s)\|_{\mathcal{V}}\,\mathrm{d}s.$$

Observe that the first and the last term on the right hand side are zero since ℓ_1 and ℓ_2 differ at most on a countable set. Thanks to (2.4) and Young's inequality the integral on the right hand side can be estimated as

$$\int_0^t \langle \mathrm{D}\mathcal{F}(z_2(s)) - \mathrm{D}\mathcal{F}(z_1(s)),\dot z_1(s)-\dot z_2(s)\rangle\,\mathrm{d}s$$
$$\le C\int_0^t \|z_1(s)-z_2(s)\|_{\mathcal{Z}}\,\|\dot z_1(s)-\dot z_2(s)\|_{\mathcal{V}}\,\mathrm{d}s$$
$$\le \int_0^t \tfrac{\varepsilon}{2}\,\|\dot z_1(s)-\dot z_2(s)\|_{\mathcal{V}}^2\,\mathrm{d}s + C_\varepsilon\int_0^t \|z_1(s)-z_2(s)\|_{\mathcal{Z}}^2\,\mathrm{d}s.$$

Joining these inequalities and applying the Gronwall Lemma finishes the proof of uniqueness.

Existence of weak solutions:

Let $\varepsilon > 0$ be fixed. Let $(\Pi_N)_{N\in\mathbb{N}}$ be a sequence of partitions of $[0,T]$ with fineness $\triangle_N \searrow 0$ and let $(z_k^N)_{k\le N}$ be minimizers of (2.15). We introduce the following piecewise affine and piecewise linear interpolants:

$$\widetilde{z}_N(t) := z_{k-1}^N + \tfrac{t-t_{k-1}}{\tau_k}(z_k^N - z_{k-1}^N), \quad t \in [t_{k-1}^N, t_k^N],$$
$$\underline{z}_N(t) := z_{k-1}^N, \quad t \in [t_{k-1}^N, t_k^N); \quad \overline{z}_N(t) := z_k^N,\ \bar{t}_N(t) := t_k^N, \quad t \in (t_{k-1}^N, t_k^N].$$

By Proposition 2.3 the functions $\widetilde{z}_N$, $\bar{z}_N$, $\underline{z}_N$ are uniformly bounded (w.r. to N and ε) in the space $L^\infty((0,T);\mathcal{Z})$. Moreover, we have

$$\|\widetilde{z}_N\|_{H^1((0,T);\mathcal{V})} \le C/\sqrt{\varepsilon} \tag{3.5}$$

with a constant $C > 0$ that is independent of the partition Π_N. Thus, there exists $z \in L^\infty((0,T);\mathcal{Z}) \cap H^1((0,T);\mathcal{V})$ and a (not relabeled) subsequence such that

$$\widetilde{z}_N, \overline{z}_N, \underline{z}_N \overset{*}{\rightharpoonup} z \text{ weakly}* \text{ in } L^\infty((0,T);\mathcal{Z}), \tag{3.6}$$
$$\widetilde{z}_N \rightharpoonup z \text{ weakly in } H^1((0,T);\mathcal{V}), \tag{3.7}$$
$$\widetilde{z}_N(t), \overline{z}_N(t), \underline{z}_N(t) \rightharpoonup z(t) \text{ weakly in } \mathcal{Z} \text{ for all } t \in [0,T], \tag{3.8}$$

where the last line is a consequence of (3.6) and (3.7). Thanks to (3.5) the limits of the different interpolants coincide. All accumulation points obtained in this way are uniformly bounded in $L^\infty((0,T);\mathcal{Z})$ with respect to $\varepsilon > 0$ and the chosen sequence of partitions. With the above definitions, for $t > 0$ the inclusion (2.25) can be rewritten as $-\mathrm{D}\mathcal{E}(\bar{t}_N(t), \overline{z}_N(t)) \in \partial\mathcal{R}_\varepsilon(\dot{\widetilde{z}}_N(t))$, and by convex analysis and the chain rule we obtain

$$\mathcal{R}_\varepsilon(\dot{\widetilde{z}}_N(t)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}\mathcal{E}(\bar{t}_N(t), \overline{z}_N(t)))$$
$$= -\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}(\widetilde{z}(t)) + \langle \ell(\bar{t}_N(t)), \dot{\widetilde{z}}_N(t)\rangle + \langle \mathrm{D}\mathcal{I}(\widetilde{z}(t)) - \mathrm{D}\mathcal{I}(\overline{z}(t)), \dot{\widetilde{z}}_N(t)\rangle.$$

Integration with respect to t results in a discrete version of the energy dissipation estimate (3.3) with an additional error term: For all $t \in [0,T]$

$$\mathcal{I}(\widetilde{z}_N(t)) + \int_0^t \mathcal{R}_\varepsilon(\dot{\widetilde{z}}_N(s)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}\mathcal{E}(\bar{t}_N(s), \overline{z}_N(s)))\,\mathrm{d}s$$
$$\le \mathcal{I}(z_0) + \int_0^t \langle \ell(\bar{t}_N(s)), \dot{\widetilde{z}}_N(s)\rangle\,\mathrm{d}s + \int_0^t r_N(s)\,\mathrm{d}s, \tag{3.9}$$

where $r_N(t) = \langle \mathrm{D}\mathcal{I}(\widetilde{z}_N(t)) - \mathrm{D}\mathcal{I}(\overline{z}_N(t)), \dot{\widetilde{z}}_N(t)\rangle$. Next we pass to the limit $N \to \infty$ in (3.9). Since $\widetilde{z}_N(t) - \overline{z}_N(t) = \dot{\widetilde{z}}_N(t)(t - \bar{t}_N(t))$, with (2.10) we find

$$r_N(t) = -(\bar{t}_N(t) - t)^{-1}\langle \mathrm{D}\mathcal{I}(\widetilde{z}_N(t)) - \mathrm{D}\mathcal{I}(\overline{z}_N(t)), \widetilde{z}_N(t) - \overline{z}_N(t)\rangle \le \lambda\tau_k \big\|\dot{\widetilde{z}}_N(t)\big\|_{\mathcal{V}}^2,$$

and $\lambda > 0$ is independent of $\varepsilon > 0$ and the partition Π_N. Hence, relying on estimate (3.5) we obtain

$$\limsup_{N\to\infty} \int_0^t r_N(s)\,\mathrm{d}s \le \lambda \limsup_{N\to\infty} \triangle_N \,\|\widetilde{z}_N\|_{H^1((0,T);\mathcal{V})}^2 = 0,$$

as $\limsup_{N\to\infty} \triangle_N = 0$. Concerning the power term observe first that $\bar{t}_N(t) \searrow t$ for $N \to \infty$, and hence, $\ell(\bar{t}_N(t)) \to \ell(t+) = \ell_+(t)$ strongly in $\mathcal{V}^*$ (for all $t \in [0,T]$). Since $\ell \in L^\infty((0,T);\mathcal{V}^*)$ this implies in particular that $\ell \circ \bar{t}_N \to \ell_+$ strongly in $L^2((0,T);\mathcal{V}^*)$. Taking into account the weak convergence of $(\dot{\widetilde{z}}_N)_N$ in $L^2((0,T);\mathcal{V})$ we obtain

$$\int_0^t \langle \ell \circ \bar{t}_N, \dot{\widetilde{z}}_N\rangle\,\mathrm{d}s \to \int_0^t \langle \ell_+, \dot{z}\rangle\,\mathrm{d}s.$$

The discrete energy dissipation estimate (3.9) in particular implies that

$$\sup_N \int_0^T \mathcal{R}_\varepsilon^*(-\mathrm{D}\mathcal{E}(\bar{t}_N, \overline{z}_N))\,\mathrm{d}s < \infty$$

and hence $\mathrm{D}\mathcal{E}(\bar{t}_N, \overline{z}_N)$ is uniformly bounded (with respect to N) in $L^2((0,T);\mathcal{V}^*)$. Thanks to (3.8) we also have pointwise weak convergence in $\mathcal{Z}^*$ of $\mathrm{D}\mathcal{E}(\bar{t}_N(t), \overline{z}_N(t))$ to $\mathrm{D}\mathcal{E}(t+, z(t))$ so that altogether $\mathrm{D}\mathcal{E}(\bar{t}_N, \overline{z}_N) \rightharpoonup \mathrm{D}\mathcal{E}(\cdot+, z(\cdot))$ weakly in $L^2((0,T);\mathcal{V}^*)$. By lower semicontinuity we therefore obtain for the left hand side in (3.9)

$$\liminf_N \,(\text{l.h.s}) \geq \mathcal{I}(z(t)) + \int_0^t \mathcal{R}_\varepsilon(\dot{z}(s)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}\mathcal{E}(s+, z(s)))\,\mathrm{d}s.$$

In summary we have shown that z satisfies (3.3) with ℓ_+ and therefore also with ℓ. Hence, by Lemma 3.2 z is a weak solution to (3.1) for ℓ.

Improved estimates: Assume in addition that $\mathrm{D}\mathcal{E}(0, z_0) \in \mathcal{V}^*$. Then from Proposition 2.5 we obtain

$$\|\widetilde{z}_N\|_{W^{1,1}((0,T);\mathcal{Z})} + \varepsilon \left\|\dot{\widetilde{z}}_N\right\|_{L^\infty(0,T;\mathcal{V})} + \|\mathrm{D}\mathcal{E}(\bar{t}_N, \overline{z}_N)\|_{L^\infty((0,T);\mathcal{V}^*)} \leq C,$$

and $C > 0$ is independent of ε and Π_N. Moreover, $\|\widetilde{z}_N\|_{H^1((0,T),\mathcal{Z})} \leq C_\varepsilon$, uniformly in N. Hence, by weak compactness and lower semicontinuity, for $N \to \infty$ we obtain the improved regularity of z as well as (3.4). $\square$

4. The viscosity limit

In order to study the limit $\varepsilon \to 0$ we use the reparameterization technique originally introduced in [EM06] and refined in [MRS16], among others. In this section we assume

$$\text{(2.1)–(2.8) and that } \mathrm{D}\mathcal{E}(0, z_0) \in \mathcal{V}^*. \tag{4.1}$$

Let

$$\mathrm{p} : \mathcal{V} \times \mathcal{V}^* \to \mathbb{R}, \quad \mathrm{p}(v, w) := \mathcal{R}(v) + \|v\|_{\mathbb{V}} \operatorname{dist}_{\mathbb{V}}(w, \partial\mathcal{R}(0))$$

denote the so called *vanishing viscosity contact potential*, [MRS16]. Observe that by Young's inequality, for all $\varepsilon > 0$ we have $\mathrm{p}(v, w) \leq \mathcal{R}_\varepsilon(v) + \mathcal{R}_\varepsilon^*(w)$. Let $\varepsilon > 0$ and let z_ε be a weak solution of the viscous problem (3.1). As in [MRS16], we define

$$s_\varepsilon(t) := t + \int_0^t \mathrm{p}(\dot{z}_\varepsilon(r), -\mathrm{D}\mathcal{E}(r, z_\varepsilon(r))\,\mathrm{d}r, \quad S_\varepsilon := s_\varepsilon(T)\,. \tag{4.2}$$

By definition, $s_\varepsilon : [0,T] \to [0, S_\varepsilon]$ is strictly monotone and hence invertible. We denote with $\hat{t}_\varepsilon : [0, S_\varepsilon] \to [0,T]$ the inverse of s_ε. Furthermore, let

$$\hat{z}_\varepsilon(s) := z_\varepsilon(\hat{t}_\varepsilon(s)), \quad \hat{\ell}_\varepsilon(s) := \ell(\hat{t}_\varepsilon(s))\,. \tag{4.3}$$

Clearly, $\hat{t}_\varepsilon \in W^{1,\infty}((0, S_\varepsilon))$ and for almost all s we have

$$\hat{t}_\varepsilon'(s) + \mathrm{p}(\hat{z}_\varepsilon'(s), -\mathrm{D}\mathcal{E}(\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s))) = 1. \tag{4.4}$$

In the next proposition we collect regularity properties and (uniform) estimates that are valid for the transformed quantities.

Proposition 4.1. *Assume* (4.1). *Then* $\sup_{\varepsilon>0} S_\varepsilon < \infty$, $\hat{z}_\varepsilon$ *belongs to the space* $H^1((0,S_\varepsilon);\mathcal{Z}) \cap W^{1,\infty}((0,S_\varepsilon);\mathcal{V})$ *and there is a constant* $C > 0$ *such that for all* $\varepsilon > 0$ *and with* $I_\varepsilon := (0, S_\varepsilon)$ *we have*

$$\|\hat{z}_\varepsilon\|_{W^{1,1}(I_\varepsilon;\mathcal{Z})} + \|\hat{z}_\varepsilon'\|_{L^\infty(I_\varepsilon;\mathcal{X})} + \varepsilon \left\|(\hat{t}_\varepsilon')^{-1}\hat{z}_\varepsilon'\right\|_{L^\infty(I_\varepsilon;\mathcal{V})} + \left\|\mathrm{D}\mathcal{E}(\hat{t}_\varepsilon, \hat{z}_\varepsilon)\right\|_{L^\infty(I_\varepsilon;\mathcal{V}^*)} < C. \tag{4.5}$$

Moreover, $\hat{\ell}_\varepsilon \in BV([0, S_\varepsilon];\mathcal{V}^*)$ *with* $\operatorname{Var}_{\mathcal{V}^*}(\hat{\ell}_\varepsilon, [0, S_\varepsilon]) = \operatorname{Var}_{\mathcal{V}^*}(\ell, [0,T])$.

Proof. Observe that $S_\varepsilon \le T + \int_0^T \mathcal{R}_\varepsilon(\dot z_\varepsilon(r)) + \mathcal{R}_\varepsilon^*(-\mathrm{D}\mathcal{E}(r, z_\varepsilon(r)))\,\mathrm{d}r$. From the identity (3.2) and estimate (3.4) we deduce the uniform bound for $(S_\varepsilon)_\varepsilon$. Since $\hat t_\varepsilon$ is Lipschitz continuous, the regularity of $\hat z_\varepsilon$ and estimate (4.5) immediately follow from Proposition 3.3. Observe finally that thanks to the strict monotonicity of s_ε we have $\mathrm{Var}_{\mathcal{V}^*}(\hat\ell_\varepsilon, [a,b]) = \mathrm{Var}_{\mathcal{V}^*}(\ell, [\hat t_\varepsilon(a), \hat t_\varepsilon(b)])$. □

As a consequence, by compactness we obtain

Proposition 4.2. *Assume* (4.1).

Let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a sequence with $\varepsilon_n \searrow 0$ for $n \to \infty$. Then there exist $S > 0$, a triple $(\hat t, \hat z, \hat\ell)$ with $\hat t \in W^{1,\infty}(0, S; \mathbb{R})$, $\hat z \in AC^\infty([0, S]; \mathcal{X}) \cap C([0, S]; \mathcal{V}) \cap BV([0, S]; \mathcal{Z}) \cap C_{weak}([0, S]; \mathcal{Z})$ and $\hat\ell \in BV([0, S]; \mathcal{V}^*)$ and a subsequence of $(\varepsilon_n)_n$ such that for $n \to \infty$ (we suppress the index n)

$$S_\varepsilon \to S; \quad \hat t_\varepsilon \overset{*}{\rightharpoonup} \hat t \text{ weakly* in } W^{1,\infty}(0, S),\ \hat t(S) = T, \tag{4.6}$$

$$\hat z_\varepsilon \rightharpoonup \hat z \text{ weakly* in } L^\infty(0, S; \mathcal{Z}) \text{ and uniformly in } C([0, S]; \mathcal{V}), \tag{4.7}$$

$$\hat\ell_\varepsilon \overset{*}{\rightharpoonup} \hat\ell,\ \mathrm{D}\mathcal{I}(\hat z_\varepsilon) \overset{*}{\rightharpoonup} \mathrm{D}\mathcal{I}(\hat z) \text{ weakly* in } L^\infty(0, S; \mathcal{V}^*), \tag{4.8}$$

and for every $s \in [0, S]$

$$\hat t_\varepsilon(s) \to \hat t(s), \quad \hat z_\varepsilon(s) \rightharpoonup \hat z(s) \text{ weakly in } \mathcal{Z}\ , \tag{4.9}$$

$$\mathrm{D}\mathcal{I}(\hat z_\varepsilon(s)) \rightharpoonup \mathrm{D}\mathcal{I}(\hat z(s)) \text{ weakly in } \mathcal{V}^*, \quad \hat\ell_\varepsilon(s) \to \hat\ell(s) \text{ strongly in } \mathcal{V}^*. \tag{4.10}$$

The function $s \mapsto \mathcal{I}(\hat z(s))$ is uniformly continuous on $[0, S]$, the function $s \mapsto \mathrm{D}\mathcal{I}(\hat z(s))$ belongs to $C_{weak}([0, S]; \mathcal{V}^*)$ and $\hat t'(s) \ge 0$ for almost all s. Moreover, $\hat\ell$ can be characterized as follows: For every $t_* \in [0, T]$ there exists $s_* \in \hat t^{-1}(t_*)$ such that for all $s \in [0, S]$ with $\hat t(s) = t_*$ we have

$$\hat\ell(s) = \begin{cases} \ell(\hat t(s)-) & s < s_* \\ \ell(\hat t(s)+) & s > s_* \end{cases} \quad \text{and} \quad \hat\ell(s_*) \in \{\ell(t_*), \ell(t_*+), \ell(t_*-)\}. \tag{4.11}$$

Remark 4.3. In the previous proposition we tacitly extend all functions by their constant value in S_ε, if $S_\varepsilon < S$.

Proof of Proposition 4.2. The uniform bounds provided in Proposition 4.1 in combination with Proposition C.5 yield the convergence properties of the sequence $(\hat z_\varepsilon)_\varepsilon$ and the regularity of the limit function $\hat z$. The first assertion in (4.10) is a consequence of the weak continuity of $\mathrm{D}\mathcal{I} : \mathcal{Z} \to \mathcal{Z}^*$, (4.9) and the uniform estimate (4.5). From this we also obtain the second part of (4.8). By the very same argument the weak continuity of $s \mapsto \mathrm{D}\mathcal{I}(\hat z(s))$ in $\mathcal{V}^*$ ensues.

Let us next show that $s \mapsto \mathcal{I}(\hat z(s))$ is continuous and thus uniformly continuous on $[0, S]$. As stated above, we have $\mathrm{D}\mathcal{I}(\hat z(\cdot)) \in C_{\mathrm{weak}}([0, S]; \mathcal{V}^*)$. But this is also separately valid for the mappings $s \mapsto A\hat z(s)$ and $z \mapsto \mathrm{D}\mathcal{F}(\hat z(s))$. Indeed, since $\hat z \in L^\infty(0, S; \mathcal{Z})$ the assumed bound in (2.4) yields $\mathrm{D}\mathcal{F}(\hat z(\cdot)) \in L^\infty(0, S; \mathcal{V}^*)$. Combining this with assumption (2.12) and the fact that $\hat z \in C_{\mathrm{weak}}([0, S]; \mathcal{Z})$, we obtain $\mathrm{D}\mathcal{F}(\hat z(\cdot)) \in C_{\mathrm{weak}}([0, S]; \mathcal{V}^*)$, and hence also $A\hat z(\cdot) \in C_{\mathrm{weak}}([0, S]; \mathcal{V}^*)$. By standard arguments we ultimately obtain the continuity of $s \mapsto \mathcal{I}(\hat z(s))$.

It remains to discuss the sequence $(\hat\ell_\varepsilon)_\varepsilon$. The Banach space valued version of Helly's selection principle, [BP86], applied to the sequence $(\hat\ell_\varepsilon)_\varepsilon$ yields (4.8) and weak convergence in (4.10). Since ℓ possesses (strong) left and right limits in $\mathcal{V}^*$ and since $(\ell(\hat t_\varepsilon(s)))_\varepsilon$ converges weakly for all s, it follows that $\hat\ell(s)$ belongs to the set $\{\ell(\hat t(s)), \ell(\hat t(s)+), \ell(\hat t(s)-)\}$ and that $\hat\ell_\varepsilon(s) \to \hat\ell(s)$ strongly in $\mathcal{V}^*$. Let $t_* \in [0, T]$. If t_* is a point of continuity of ℓ, the proof of the representation formula for $\hat\ell$ is finished. Assume now that t_* is a jump point of ℓ with $\ell(t_*-) \ne \ell(t_*+)$ (the arguments here below can easily be adapted to the case $\ell(t_*-) = \ell(t_*+) \ne \ell(t_*)$). By monotonicity and continuity of $\hat t$ we have $\hat t^{-1}(t_*) = [a, b]$ for some $a < b$. Let $s \in [a, b]$ with $\hat\ell(s) = \ell(t_*+)$. This implies that there is $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$ we have $\hat t_\varepsilon(s) \ge t_*$. Again by monotonicity this implies that $\hat t_\varepsilon(\sigma) \ge t_*$ for every $\sigma \in [s, b]$ and every $\varepsilon < \varepsilon_0$. Hence, for all these σ we have $\hat\ell(\sigma) = \ell(t_*+)$. Let $s_+ := \inf\{\, s \in [a, b]\,;\, \hat\ell(s) = \ell(t_*+)\,\}$. Then $\hat\ell(s) = \ell(t_*+)$ for all $s \in (s_+, b]$.

In a similar way we define $s_- := \sup\{\, s \in [a,b]\,;\, \hat\ell(s) = \ell(t_*-)\,\}$ and obtain $\hat\ell(s) = \ell(t_*-)$ for all $s \in [a, s_-)$. Observe that $s_- \le s_+$. Assume now that $s_- < s_+$ and let $s_1 < s_2 \in (s_-, s_+)$ which implies $\hat\ell(s_1) = \hat\ell(s_2) = \ell(t_*)$. But this is only possible if there exists $\varepsilon_1 > 0$ such that for all $\varepsilon < \varepsilon_1$ we have $\hat t_\varepsilon(s_1) = t_* = \hat t_\varepsilon(s_2)$, which is a contradiction to the strict monotonicity of $\hat t_\varepsilon$. Hence, $s_- = s_+ =: s_*$ and the proof is finished. $\square$

Next we rewrite the energy dissipation estimate (3.3) in the new variables and investigate the limit $\varepsilon \to 0$. For that purpose we need to introduce some more notation. For a curve $z : [0,S] \to \mathcal{X}$ we define

$$\mathrm{Var}_{\mathcal{R}}(z,[a,b]) := \sup_{\text{partitions } (t_i)_i \text{ of } [a,b]} \sum_{i=1}^m \mathcal{R}(z(t_i) - z(t_{i-1}))$$

as the $\mathcal{R}$ dissipation ($\mathcal{R}$ variation) along the curve z. Thanks to the assumptions on $\mathcal{R}$ we have $\mathrm{Var}_{\mathcal{R}}(z;[a,b]) < \infty$ if and only if $\mathrm{Var}_{\mathcal{X}}(z;[a,b]) < \infty$.

Let $\hat{\mathcal{E}}(s,v) := \mathcal{I}(v) - \langle \hat\ell(s), v\rangle$. In order to shorten the notation let

$$\mathrm{e}(f,v) := \mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{I}(v) + f, \partial\mathcal{R}(0)). \tag{4.12}$$

With this, $\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s,\hat z(s)), \partial\mathcal{R}(0)) = \mathrm{e}(\hat\ell(s), \hat z(s))$. For $f \in BV([0,S];\mathcal{V}^*)$ and $v \in \mathcal{Z}$ let

$$\mathrm{m}(f(s),v) := \min\{\mathrm{e}(f(s),v), \mathrm{e}(f(s-),v), \mathrm{e}(f(s+),v)\}. \tag{4.13}$$

The next lemma shows that $\mathrm{m}(\cdot,\cdot)$ is lower semicontinuous.

Lemma 4.4. *Let $f \in BV([0,S];\mathcal{V}^*)$, $(v_n)_n \subset \mathcal{Z}$ with $\mathrm{D}\mathcal{I}(v_n) \rightharpoonup \mathrm{D}\mathcal{I}(v)$ weakly in $\mathcal{V}^*$ and $(s_n)_n, s \subset [0,S]$ with $s_n \to s$. Then $\liminf_n \mathrm{m}(f(s_n), v_n) \ge m(f(s),v)$.*

Proof. Observe that the accumulation points of the sequences $(f(s_n+))_n$, $(f(s_n))_n$, $(f(s_n-))_n$ belong to the set $\{f(s), f(s+), f(s-)\}$. Hence, by the lower semicontinuity of the functional $\mathrm{dist}_{\mathbb{V}}$ we conclude. $\square$

Theorem 4.5. *Assume (4.1). Then there exist $S > 0$, $\hat t \in W^{1,\infty}(0,S;\mathbb{R})$, $\hat z \in AC^\infty([0,S];\mathcal{X}) \cap C([0,S];\mathcal{V}) \cap BV([0,S];\mathcal{Z}) \cap C_{weak}([0,S];\mathcal{Z})$ and $\hat\ell \in BV([0,S];\mathcal{V}^*)$ as in (4.11) such that $\mathcal{I}(\hat z) \in C([0,S])$, $\mathrm{D}\mathcal{I}(\hat z) \in L^\infty(0,S;\mathcal{V}^*) \cap C_{weak}([0,S];\mathcal{V}^*)$. Let $G := \{\, s \in [0,S]\,;\, \mathrm{m}(\hat\ell(s), \hat z(s)) > 0\,\}$. The set G is open and $\hat z \in W^{1,\infty}_{loc}(G;\mathcal{V})$. Moreover, for almost every $s \in [0,S]$*

$$\hat t'(s) \ge 0,\ \hat t(S) = T,\ \hat z(0) = z_0, \tag{4.14}$$

$$\hat t'(s)\,\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s,\hat z(s)), \partial\mathcal{R}(0)) = 0\,, \tag{4.15}$$

$$1 = \begin{cases} \hat t'(s) + \mathcal{R}[\hat z'](s) & \text{if } s \notin G \\ \hat t'(s) + \mathcal{R}[\hat z'](s) + \|\hat z'(s)\|_{\mathcal{V}}\,\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s,\hat z(s)), \partial\mathcal{R}(0)) & \text{if } s \in G \end{cases}. \tag{4.16}$$

Furthermore, $\hat t' = 0$ almost everywhere on G. Finally, for every $s \in [0,S]$

$$\mathcal{I}(\hat z(s)) + \int_0^s \mathcal{R}[\hat z'](r)\,\mathrm{d}r + \int_{(0,s)\cap G} \|\hat z'(r)\|_{\mathbb{V}}\,\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(r,\hat z(s)), \partial\mathcal{R}(0))\,\mathrm{d}r$$
$$= \mathcal{I}(z_0) + \int_0^s \langle \hat\ell(r), \mathrm{d}\hat z(r)\rangle\,. \tag{4.17}$$

Every tuple $(S, \hat t, \hat z, \hat\ell)$ obtained as a limit as in Proposition 4.2 satisfies the above conditions.

The integral on the right hand side in (4.17) is understood as a Kurzweil integral, see Appendix B.

Proof. For $\varepsilon > 0$ let z_ε be a solution to (3.1) and let $(S_\varepsilon, \hat{t}_\varepsilon, \hat{z}_\varepsilon, \hat{\ell}_\varepsilon)_{\varepsilon>0}$ be a sequence constructed from $(z_\varepsilon)_\varepsilon$ that converges to $(S, \hat{t}, \hat{z}, \hat{\ell})$ as stated Proposition 4.2. The aim is to show that $(S, \hat{t}, \hat{z}, \hat{\ell})$ has the properties formulated in Theorem 4.5.

Complementarity identity (4.15): Since $\partial\mathcal{R}(\dot{z}_\varepsilon(t)) \subset \partial\mathcal{R}(0)$, from (3.1) we deduce

$$-\mathrm{D}\mathcal{E}(\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s)) \in \partial\mathcal{R}(0) + \tfrac{\varepsilon}{\hat{t}_\varepsilon'(s)} \mathbb{V}\hat{z}_\varepsilon'(s), \tag{4.18}$$

which implies that $\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s)), \partial\mathcal{R}(0)) \le \frac{\varepsilon}{\hat{t}_\varepsilon'(s)} \|\hat{z}_\varepsilon'(s)\|_{\mathbb{V}}$. Since $\partial\mathcal{R}(0)$ is bounded in $\mathcal{V}^*$, by lower semicontinuity and in combination with (4.5) and (4.10) it follows that $\mathrm{D}\hat{\mathcal{E}}(\cdot, \hat{z}(\cdot)) \in L^\infty((0, S); \mathcal{V}^*)$. Moreover, since $\varepsilon \|\dot{z}_\varepsilon\|^2_{L^2((0,T);\mathcal{V})}$ is uniformly bounded (cf. (3.3) and Proposition 3.3), we obtain

$$\sup_\varepsilon \varepsilon \left\| (\hat{t}_\varepsilon')^{-\frac{1}{2}} \hat{z}_\varepsilon' \right\|^2_{L^2((0,S);\mathcal{V})} = \sup_\varepsilon \varepsilon \|\dot{z}_\varepsilon\|^2_{L^2((0,T);\mathcal{V})} =: C < \infty.$$

Since $\hat{t}_\varepsilon'(s) \le 1$, we therefore arrive at $\int_0^S \left(\hat{t}_\varepsilon' \, \mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(\hat{t}_\varepsilon, \hat{z}_\varepsilon), \partial\mathcal{R}(0))\right)^2 \mathrm{d}s \le \varepsilon C$. Thanks to (4.10), for almost every s we have $\liminf_\varepsilon \mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s)), \partial\mathcal{R}(0)) \ge \mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s, \hat{z}(s)), \partial\mathcal{R}(0))$. Hence, Proposition C.1 implies

$$\begin{aligned} 0 &\ge \liminf_\varepsilon \int_0^S (\hat{t}_\varepsilon')^2 \,\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(\hat{t}_\varepsilon, \hat{z}_\varepsilon), \partial\mathcal{R}(0))^2 \,\mathrm{d}s \\ &\ge \int_0^S (\hat{t}'(s))^2 \,\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s, \hat{z}(s)), \partial\mathcal{R}(0))^2 \,\mathrm{d}s \ge 0 \end{aligned}$$

from which (4.15) is an immediate consequence.

Energy dissipation estimate (4.17), $\le$: For every $\varepsilon > 0$ and $s \in [0, S]$ we have the energy dissipation estimate

$$\mathcal{I}(\hat{z}_\varepsilon(s)) + \int_0^s \mathrm{p}(\hat{z}_\varepsilon'(r), -\mathrm{D}\mathcal{E}(\hat{t}_\varepsilon(r), \hat{z}_\varepsilon(r))\,\mathrm{d}r \le \mathcal{I}(z_0) + \int_0^s \langle \hat{\ell}_\varepsilon(r), \hat{z}_\varepsilon'(r)\rangle \,\mathrm{d}r, \tag{4.19}$$

which is a reparameterized version of (3.3) in combination with the estimate for $\mathrm{p}(\cdot,\cdot)$.

Thanks to Proposition B.1 we have $\int_0^s \langle \hat{\ell}_\varepsilon, \hat{z}_\varepsilon'\rangle \,\mathrm{d}r \to \int_0^s \langle \hat{\ell}(r), \mathrm{d}\hat{z}(r)\rangle$, where the last term is to be interpreted as a Kurzweil integral. By lower semicontinuity, for every s it holds $\liminf_\varepsilon \mathcal{I}(\hat{z}_\varepsilon(s)) \ge \mathcal{I}(\hat{z}(s))$ and it remains to pass to the limit inferior in the dissipation integral. Again by Helly, [MM05, Theorem 3.2], we obtain

$$\liminf_{\varepsilon\to 0} \int_0^s \mathcal{R}(\hat{z}_\varepsilon'(r))\,\mathrm{d}r \ge \mathrm{Var}_{\mathcal{R}}(\hat{z}, [0, s]) = \int_0^s \mathcal{R}[\hat{z}'](r)\,\mathrm{d}r,$$

where for the last identity we have applied Lemma C.3 with $p = \infty$.

The remaining term $\int_0^s \|\hat{z}_\varepsilon'(r)\|_{\mathbb{V}} \,\mathrm{e}(\hat{\ell}_\varepsilon(r), \hat{z}_\varepsilon(r))\,\mathrm{d}r$ is more delicate and we follow the arguments in [MRS16] exploiting in addition the uniform bound $\mathrm{D}\mathcal{I}(\hat{z}_\varepsilon) \in L^\infty((0, S_\varepsilon); \mathcal{V}^*)$. We recall the definition of $\mathrm{m}(\cdot,\cdot)$ in (4.13). The set

$$G = \{\, s \in [0, S]\,;\, \mathrm{m}(\hat{\ell}(s), \hat{z}(s)) > 0 \,\}$$

is relatively open (w.r. to $[0, S]$). Indeed, let $(s_n)_n \subset [0, S]\backslash G$ with $s_n \to s$. By Proposition 4.2 we have $\mathrm{D}\mathcal{I}(\hat{z}(s_n)) \rightharpoonup \mathrm{D}\mathcal{I}(\hat{z}(s))$ weakly in $\mathcal{V}^*$. Hence, with Lemma 4.4 we obtain $0 = \liminf_n m(\hat{\ell}(s_n), \hat{z}(s_n)) \ge m(\hat{\ell}(s), \hat{z}(s)) = 0$, consequently $s \notin G$.

Next, as in [MRS16], we derive an improved uniform regularity estimate for $(\hat{z}_\varepsilon)_\varepsilon$ that is valid on compact subsets of G and that allows us to give a meaning to $\hat{z}'$ on G. Let $K \subset G$ be compact. By lower semicontinuity it follows that $c := \inf_K \mathrm{m}(\hat{\ell}(s), \hat{z}(s))$ is positive. Again by lower semicontinuity for every $s \in K$ it holds

$$\liminf_\varepsilon \mathrm{e}(\hat{\ell}_\varepsilon(s), \hat{z}_\varepsilon(s)) \ge \mathrm{m}(\hat{\ell}(s), \hat{z}(s)) \ge c.$$

Hence, for every $s \in K$ there exists $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$ we have $\mathrm{e}(\hat{\ell}_\varepsilon(s), \hat{z}_\varepsilon(s)) \geq c/2$. A proof by contradiction shows that ε_0 in fact can be chosen independently of $s \in K$. From the normalization property (4.4) we therefore deduce that $\sup_{\varepsilon<\varepsilon_0} \|\hat{z}_\varepsilon'\|_{L^\infty(K;\mathcal{V})} \leq 2/c$ and hence $(\hat{z}_\varepsilon)_\varepsilon$ converges weakly$*$ in $W^{1,\infty}(K;\mathcal{V})$ to $\hat{z}$. Now we are in the position to apply Proposition C.1 to conclude that

$$\liminf_{\varepsilon} \int_K \|\hat{z}_\varepsilon'(s)\|_{\mathcal{V}} \operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s)), \partial\mathcal{R}(0))\,\mathrm{d}s \\ \geq \int_K \|\hat{z}'(s)\|_{\mathcal{V}} \operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s, \hat{z}(s)), \partial\mathcal{R}(0))\,\mathrm{d}s. \quad (4.20)$$

In summary we have proved (4.17) with $\leq$ instead of equality. By similar arguments we obtain (4.16) with $\geq$ instead of equality.

In order to prove that in fact an identity is valid in (4.17) and (4.16) we follow ideas from [MRS12a]. For $s \in [0, S]$ let $\mu(s) \in \partial\mathcal{R}(0)$ with $\left\|-\mathrm{D}\hat{\mathcal{E}}(s, \hat{z}(s)) - \mu(s)\right\|_{\mathcal{V}^*} = \operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s, \hat{z}(s)), \partial\mathcal{R}(0))$. Then from (2.11) for every $s \in [0, S)$ and $h > 0$ (such that $s + h \in [0, S]$ and with $\triangle_h \hat{z}(s) = \hat{z}(s+h) - \hat{z}(s)$) we obtain

$$\mathcal{I}(\hat{z}(s+h)) - \mathcal{I}(\hat{z}(s)) \\ \geq \langle \mathrm{D}\hat{\mathcal{E}}(s, \hat{z}(s)) + \mu(s), \triangle_h \hat{z}(s)\rangle + \langle \hat{\ell}(s), \triangle_h \hat{z}(s)\rangle - \langle \mu(s), \triangle_h \hat{z}(s)\rangle \\ - C\,\|\triangle_h \hat{z}(s)\|_{\mathcal{V}}\, \mathcal{R}(\triangle_h \hat{z}(s))\,. \quad (4.21)$$

Thanks to the definition of μ we have the estimates

$$\langle \mathrm{D}\hat{\mathcal{E}}(s, \hat{z}(s)) + \mu(s), \triangle_h \hat{z}(s)\rangle \\ \leq \left\|\mathrm{D}\hat{\mathcal{E}}(s, \hat{z}(s)) + \mu(s)\right\|_{\mathbb{V}^*} \|\triangle_h \hat{z}(s)\|_{\mathbb{V}} = \operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s, \hat{z}(s)), \partial\mathcal{R}(0))\, \|\triangle_h \hat{z}(s)\|_{\mathbb{V}}$$

and $\mathcal{R}(v) \geq \langle \mu(s), v\rangle$ for all $v \in \mathcal{Z}$. Hence, after rearranging the terms in (4.21) and integration with respect to s, for $\sigma_1 < \sigma_2 \leq S - h$ we find

$$\int_{\sigma_1}^{\sigma_2} h^{-1}(\mathcal{I}(\hat{z}(s+h)) - \mathcal{I}(\hat{z}(s)))\,\mathrm{d}s \\ + \int_{\sigma_1}^{\sigma_2} (1 + C\,\|\triangle_h \hat{z}(s)\|_{\mathcal{V}})\mathcal{R}(h^{-1}\triangle_h \hat{z}(s)) + \operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s, \hat{z}(s)), \partial\mathcal{R}(0))\,\left\|h^{-1}\triangle_h \hat{z}(s)\right\|_{\mathcal{V}}\,\mathrm{d}s \\ \geq \int_{\sigma_1}^{\sigma_2} \langle \hat{\ell}(s), h^{-1}\triangle_h \hat{z}(s)\rangle\,\mathrm{d}s\,.$$

The next aim is to pass to the limit $h \searrow 0$ in this energy dissipation estimate. Lemma B.2 implies that $\lim_{h\searrow 0}\int_{\sigma_1}^{\sigma_2} \langle \hat{\ell}(s), h^{-1}\triangle_h \hat{z}(s)\rangle\,\mathrm{d}s = \int_{\sigma_1}^{\sigma_2} \langle \hat{\ell}(s), \mathrm{d}\hat{z}(s)\rangle$. Moreover, since $s \mapsto \mathcal{I}(\hat{z}(s))$ is uniformly continuous (cf. Proposition 4.2), for the first term on the left hand side we obtain

$$\lim_{h\searrow 0} \int_{\sigma_1}^{\sigma_2} h^{-1}(\mathcal{I}(\hat{z}(s+h)) - \mathcal{I}(\hat{z}(s)))\,\mathrm{d}s = \mathcal{I}(\hat{z}(\sigma_2)) - \mathcal{I}(\hat{z}(\sigma_1)).$$

Since $\hat{z} \in C([0, S]; \mathcal{V})$, we obtain $\triangle_h \hat{z}(s) \to 0$ strongly in $\mathcal{V}$ and uniformly in s. Furthermore, since $z \in AC^\infty([0, S]; \mathcal{X})$, the limit $\lim_{h\searrow 0} \mathcal{R}(h^{-1}\triangle_h z_h)$ exists for almost all s and equals to $\mathcal{R}[\hat{z}'](s)$, cf. Appendix C.2. By the Lebesgue Theorem we thus obtain

$$\lim_{h\searrow 0} \int_{\sigma_1}^{\sigma_2} (1 + C\,\|\triangle_h \hat{z}(s)\|_{\mathcal{V}})\mathcal{R}(h^{-1}\triangle_h \hat{z}(s))\,\mathrm{d}s = \int_{\sigma_1}^{\sigma_2} \mathcal{R}[\hat{z}'](s)\,\mathrm{d}s.$$

The definition of G and that fact that $\mathrm{e}(\hat\ell(s),\hat z(s))$ and $\mathrm{e}(\hat\ell(s\pm),\hat z(s))$ differ in at most countably many points imply that $\mathrm{e}(\hat\ell(s),\hat z(s))=0$ for almost all $s\in[0,S]\backslash G$. Thus,

$$\int_{\sigma_1}^{\sigma_2}\mathrm{dist}_{\mathbb V}(-\mathrm{D}\hat{\mathcal E}(s,\hat z(s)),\partial\mathcal R(0))\big\|\tfrac1h\triangle_h\hat z(s)\big\|_{\mathcal V}\,\mathrm{d}s=\int_{(\sigma_1,\sigma_2)\cap G}\mathrm{e}(\hat\ell(s),\hat z(s))\,\big\|\tfrac1h\triangle_h\hat z(s)\big\|_{\mathcal V}\,\mathrm{d}s.$$

Since $\hat z\in W^{1,\infty}_{\mathrm{loc}}(G;\mathcal V)$, by Lebesgue's theorem we deduce for each $K\Subset G$

$$\lim_{h\searrow0}\int_{(\sigma_1,\sigma_2)\cap K}\mathrm{e}(\hat\ell(s),\hat z(s))\,\big\|h^{-1}\triangle_h\hat z(s)\big\|_{\mathcal V}\,\mathrm{d}s=\int_{(\sigma_1,\sigma_2)\cap K}\mathrm{e}(\hat\ell(s),\hat z(s))\,\|\hat z'(s)\|_{\mathcal V}\,\mathrm{d}s.$$

To summarize, we have shown the following: By continuity of $\mathcal I(\hat z(\cdot))$ and taking into account Proposition B.1, for all $(a,b)\subset G$ we have

$$\mathcal I(\hat z(b))-\mathcal I(\hat z(a))+\int_a^b\mathcal R[\hat z'(s)]+\mathrm{e}(\hat\ell(s),\hat z(s))\,\|\hat z'(s)\|_{\mathbb V}\,\mathrm{d}s\ge\int_a^b\langle\hat\ell(s),\mathrm{d}\hat z(s)\rangle,\tag{4.22}$$

while for every $[\alpha,\beta]\subset[0,S]\backslash G$

$$\mathcal I(\hat z(\beta))-\mathcal I(\hat z(\alpha))+\int_\alpha^\beta\mathcal R[\hat z'(s)]\,\mathrm{d}s\ge\int_\alpha^\beta\langle\hat\ell(s),\mathrm{d}\hat z(s)\rangle.\tag{4.23}$$

Since G is the disjoint union of at most countably many (relatively) open intervals and keeping in mind [KL09, Proposition 1.4], a telescopic sum argument finally implies that for all $\sigma_1<\sigma_2\in[0,S]$ the energy dissipation estimate

$$\mathcal I(\hat z(\sigma_2))-\mathcal I(\hat z(\sigma_1))+\int_{\sigma_1}^{\sigma_2}\mathcal R[\hat z'(s)]\,\mathrm{d}s+\int_{(\sigma_1,\sigma_2)\cap G}\mathrm{e}(\hat\ell(s),\hat z(s))\,\|\hat z'(s)\|_{\mathbb V}\,\mathrm{d}s\\ \ge\int_{\sigma_1}^{\sigma_2}\langle\hat\ell(s),\mathrm{d}\hat z(s)\rangle$$

is valid. Together with the opposite estimate (i.e. (4.17) with $\le$) we finally obtain (4.17) with an equality.

Improved convergences: By standard arguments it follows that in fact for all $s\in[0,S]$ it holds

$$\begin{aligned}\lim_{\varepsilon\to0}\mathcal I(\hat z_\varepsilon(s))&=\mathcal I(\hat z(s)),\\ \lim_{\varepsilon\to0}\int_0^s\mathcal R(\hat z_\varepsilon'(r))\,\mathrm{d}r&=\int_0^s\mathcal R[\hat z'](r)\,\mathrm{d}r,\\ \lim_{\varepsilon\to0}\int_0^s\|\hat z_\varepsilon'(r)\|_{\mathbb V}\,\mathrm{dist}_{\mathbb V}(-\mathrm{D}\mathcal E(\hat t_\varepsilon(r),\hat z_\varepsilon(r)),\partial\mathcal R(0))\,\mathrm{d}r&\\ =\int_{(0,s)\cap G}\|\hat z'(r)\|_{\mathbb V}\,\mathrm{dist}_{\mathbb V}(-\mathrm{D}\hat{\mathcal E}(r,\hat z(r)),\partial\mathcal R(0))\,\mathrm{d}r\,.&\end{aligned}$$

In order to prove that the limit solution is normalized, i.e. in order to verify (4.16), we rewrite $\int_0^s\mathcal R(\hat z_\varepsilon'(r))+\|\hat z_\varepsilon'(r)\|_{\mathbb V}\,\mathrm{dist}_{\mathbb V}(-\mathrm{D}\mathcal E(\hat t_\varepsilon(r),\hat z_\varepsilon(r)),\partial\mathcal R(0))\,\mathrm{d}r=\int_0^s(1-\hat t_\varepsilon'(r))\,\mathrm{d}r$ and use the above convergences to conclude. ☐

Definition 4.6. Assume (4.1). A tuple $(S,\hat t,\hat z,\hat\ell)$ with $S>0$, $\hat t\in W^{1,\infty}((0,S);\mathbb R)$, $\hat z\in AC^\infty([0,S];\mathcal X)\cap L^\infty((0,S);\mathcal Z)$ and $\hat\ell\in BV([0,S];\mathcal V^*)$ is a normalized, p-parameterized balanced viscosity solution of the rate-independent system associated with $(\mathcal I,\mathcal R,\ell,z_0)$ if $\hat\ell$ is of the form (4.11), if there exists an open set $G\subset[0,S]$ such that $\hat z\in W^{1,1}_{\mathrm{loc}}(G;\mathcal V)$, $\mathrm{D}\hat{\mathcal E}(\cdot,\hat z(\cdot))\in L^\infty_{\mathrm{loc}}(G;\mathcal V^*)$ and such that $\mathrm{m}(\hat\ell(s),\hat z(s))>0$ for all $s\in G$ and $\mathrm{m}(\hat\ell(s),\hat z(s))=0$ for all $s\in[0,S]\backslash G$, and if (4.14)–(4.17) are satisfied.

With $\mathcal L(\ell,z_0)$ we denote the set of normalized, p-parameterized balanced viscosity solutions associated with $(\mathcal I,\mathcal R,\ell,z_0)$.

If (4.1) is satisfied then by Theorem 4.5 the set $\mathcal L(\ell,z_0)$ is not empty.

5. Properties of the solution set

The next lemma shows that all elements of $\mathcal{L}(\ell, z_0)$ enjoy the same regularity properties as the limit functions obtained in Proposition 4.2 (except possibly the $BV([0,S];\mathcal{Z})$ regularity) with bounds that are uniform with respect to the set $\mathcal{L}(\ell, z_0)$. While estimates (5.2)–(5.3) here below are immediate consequences of the energy dissipation balance (4.17) and the normalization property (4.16), the uniform L^∞-bound for $\mathrm{D}\hat{\mathcal{E}}$, i.e. (5.4), requires a more refined analysis.

Lemma 5.1. *Assume* (4.1).

Every normalized, p*-parameterized balanced viscosity solution* $(S, \hat{t}, \hat{z}, \hat{\ell}) \in \mathcal{L}(\ell, z_0)$ *of the rate-independent system associated with* $(\mathcal{I}, \mathcal{R}, \ell, z_0)$ *(according to Definition 4.6) satisfies*

(1) $\mathcal{I}(\hat{z}(\cdot))$ *belongs to* $C([0,S];\mathbb{R})$.

(2) $\hat{t}$ *is constant on the closure of each connected component of* G *and there exists a measurable function* $\lambda : (0,S) \to [0,\infty)$ *with* $\lambda(s) = 0$ *on* $(0,S)\backslash G$ *such that on each connected component* $(a,b) \subset G$ *the differential inclusion*

$$0 \in \partial\mathcal{R}(\hat{z}'(s)) + \lambda(s)\mathbb{V}\hat{z}'(s) + \mathrm{D}\hat{\mathcal{E}}(s, \hat{z}(s)) \tag{5.1}$$

is satisfied, for almost all $s \in (a,b)$.

For almost all $s \in G$ *we have* $\lambda(s) = \mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s,\hat{z}(s)), \partial\mathcal{R}(0)) / \|\hat{z}'(s)\|_{\mathbb{V}}$.

(3) Estimates:

There exists a constant $c > 0$ *(depending on* $\|z_0\|_{\mathcal{Z}}$, $\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(0,z_0), \partial\mathcal{R}(0))$, $\|\ell\|_{L^\infty(0,T;\mathcal{V}^*)}$, $\mathrm{Var}_{\mathcal{V}^*}(\ell,[0,T])$, *and* $\mathrm{diam}_{\mathcal{V}^*}(\partial\mathcal{R}(0))$, *only) such that for all normalized,* p*-parameterized balanced viscosity solutions associated with* (z_0, ℓ) *it holds* $\mathrm{D}\hat{\mathcal{E}}(\cdot, \hat{z}(\cdot)) \in L^\infty((0,S);\mathcal{V}^*)$ *and*

$$\|\hat{z}\|_{L^\infty((0,S);\mathcal{Z})} \le c, \;\; S \le c, \tag{5.2}$$

$$\int_0^S \mathcal{R}[\hat{z}'](s)\,\mathrm{d}s + \int_{(0,S)\cap G} \|\hat{z}'(s)\|_{\mathbb{V}}\,\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s,\hat{z}(s)), \partial\mathcal{R}(0))\,\mathrm{d}s \le c. \tag{5.3}$$

$$\left\|\mathrm{D}\hat{\mathcal{E}}(\cdot, \hat{z}(\cdot))\right\|_{L^\infty((0,S);\mathcal{V}^*)} \le c, \tag{5.4}$$

$$\|\lambda\mathbb{V}\hat{z}'\|_{L^\infty(G;\mathcal{V}^*)} \le c. \tag{5.5}$$

Finally, $\mathrm{D}\mathcal{I}(\hat{z}(\cdot)) \in C_{weak}([0,S];\mathcal{V}^*)$.

Proof. Continuity of $\mathcal{I}(\hat{z}(\cdot))$ (claim (1)): The energy dissipation identity (4.17) and the normalization property (4.16) imply that for all $a, b \in (0,S)$ we have

$$|\mathcal{I}(\hat{z}(b)) - \mathcal{I}(\hat{z}(a))| \le |b-a| + \left|\hat{t}(b) - \hat{t}(a)\right| + \left|\int_a^b \langle \hat{\ell}(s), \mathrm{d}\hat{z}(s)\rangle\right|.$$

Since $\hat{z} \in C([0,S];\mathcal{V})$ (cf. Proposition C.5) and taking into account estimate (B.1), the latter integral can be estimated as

$$\begin{aligned}\left|\int_a^b \langle \hat{\ell}(s), \mathrm{d}\hat{z}(s)\rangle\right| &= \left|\int_a^b \langle \hat{\ell}(s), \mathrm{d}(\hat{z}(s) - \hat{z}(a))\rangle\right| \\ &\le \left(\left\|\hat{\ell}\right\|_{L^\infty((0,S);\mathcal{V}^*} + \mathrm{Var}_{\mathcal{V}^*}(\hat{\ell},[0,S])\right) \|\hat{z}(\cdot) - \hat{z}(a)\|_{C([a,b];\mathcal{V})} =: f(b).\end{aligned}$$

Since $\lim_{b\to a} f(b) = 0$, the continuity of the mapping $s \mapsto \mathcal{I}(\hat{z}(s))$ ensues.

Proof of claim (2): Since $m(\hat{\ell}(s), \hat{z}(s)) > 0$ on G, from the complementarity condition (4.15) we deduce that $\hat{t}$ is constant on each connected component of G. In order to verify (5.1), let $[a,b] \Subset G$. Since by assumption $\hat{z} \in W^{1,1}((a,b);\mathcal{V})$, the identities $\mathcal{R}[\hat{z}'](s) = \mathcal{R}(\hat{z}'(s))$ and $\int_\alpha^\beta \langle \hat{\ell}(s), \mathrm{d}\hat{z}(s)\rangle = \int_\alpha^\beta \langle \hat{\ell}(s), \hat{z}'(s)\rangle\,\mathrm{d}s$ are valid for almost all $s \in (a,b)$ and all $\alpha < \beta \in (a,b)$, cf. [KL09, Proposition 1.10]. Thus, localizing the energy dissipation identity (4.17) (we apply the integrated version of the chain rule (C.7) and exploit the continuity of $\mathcal{I}(\hat{z}(\cdot))$ provided in the first part of the proposition) yields

$$\mathcal{R}(\hat{z}'(s)) + \langle \mathrm{D}\mathcal{I}(\hat{z}(s)) - \hat{\ell}(s), \hat{z}'(s)\rangle_{\mathcal{V}^*,\mathcal{V}} + \|\hat{z}'(s)\|_{\mathbb{V}} \operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s,\hat{z}(s)), \partial\mathcal{R}(0)) = 0 \tag{5.6}$$

which is valid for almost all $s \in (a,b)$. Since $\hat{t}$ is constant on (a,b), from (4.16) it follows that $\hat{z}'(s) \neq 0$ almost everywhere on (a,b). Hence, with

$$\lambda(s) := \begin{cases} \operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s,\hat{z}(s)), \partial\mathcal{R}(0)) / \|\hat{z}'(s)\|_{\mathbb{V}} & \text{if } \hat{z}'(s) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

we have $\|\hat{z}'(s)\|_{\mathbb{V}} \operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s,\hat{z}(s)), \partial\mathcal{R}(0)) = \langle \lambda(s)\mathbb{V}\hat{z}'(s), \hat{z}'(s)\rangle$ and (5.1) follows from (5.6) and the one-homogeneity of $\mathcal{R}$. This finishes the proof of claim (2) in Lemma 5.1.

Proof of the estimates (claim (3)): The verification of (5.2)–(5.3) starts from the energy dissipation identity (4.17). Indeed, for all $b \in [0,S]$ we deduce relying on the coercivity estimate for $\mathcal{I}$ and on [KL09, Theorem 1.9] (cf. (B.1) in the Appendix)

$$\tfrac{\alpha}{2} \|\hat{z}(b)\|_{\mathcal{Z}}^2 \leq \mathcal{I}(z_0) + (\|\ell\|_{L^\infty(0,T;\mathcal{V}^*)} + \operatorname{Var}_{\mathbb{V}^*}(\ell,[0,T])\, \|\hat{z}\|_{L^\infty(0,S;\mathcal{Z})} \,.$$

Here, we also used that $\operatorname{Var}_{\mathbb{V}^*}(\ell,[0,T]) = \operatorname{Var}_{\mathbb{V}^*}(\hat{\ell},[0,S])$. From this the claimed uniform bounds in (5.2)–(5.3) are an immediate consequence taking into account the normalization condition (4.16).

Let us finally show the higher regularity of $\mathrm{D}\hat{\mathcal{E}}(\cdot,\hat{z}(\cdot))$ along with estimate (5.4). Observe that $m(\hat{\ell}(s),\hat{z}(s)) = 0$ for all $s \in (0,S)\backslash G$. Since $\hat{\ell}(s), \hat{\ell}(s+), \hat{\ell}(s-)$ differ in at most countably many points, this implies that $\operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s,\hat{z}(s)), \partial\mathcal{R}(0)) = 0$ almost everywhere on $(0,S)\backslash G$. Since $\partial\mathcal{R}(0)$ is a bounded subset of $\mathcal{V}^*$, for almost all $s \in [0,S]\backslash G$ we obtain $\left\|\mathrm{D}\hat{\mathcal{E}}(s,\hat{z}(s))\right\|_{\mathcal{V}^*} \leq \operatorname{diam}_{\mathcal{V}^*}(\partial\mathcal{R}(0))$, which is (5.4) restricted to the set $(0,S)\backslash G$.

The regularity and the estimate with respect to the set G will be deduced by a rescaling argument relying on the differential inclusion (5.1), Proposition 3.3 and Remark 3.4. Let $(a,b) \subset G$ be a nonempty maximal connected component of G. A proof by contradiction relying on the lower semi-continuity property of $\mathrm{m}(\cdot,\cdot)$ stated in Lemma 4.4 shows that for every compact $K \Subset (a,b)$ there exists $c_K > 0$ such that $\mathrm{m}(\hat{\ell}(s),\hat{z}(s)) \geq c_K$ for all $s \in K$. From the normalization condition we thus obtain $\|\hat{z}'(s)\|_{\mathcal{V}} \leq c_K^{-1}$ almost everywhere on K and hence $\lambda(s) \geq c_K^2 > 0$ on K. Thus $\lambda^{-1} \in L^\infty_{\mathrm{loc}}(a,b)$.

We now distinguish two cases, namely case (a), where there exists $s_* \in (a,b)$ such that $\lambda^{-1} \notin L^1((a,s_*))$ and the simpler case (b), where we assume that for all $s_* \in (a,b)$ the function λ^{-1} belongs to $L^1((a,s_*))$.

Case (a): Assume that $\lambda^{-1} \notin L^1((a,s_*))$. Since $\lambda^{-1} \in L^\infty_{\mathrm{loc}}(a,b)$, for every $\varepsilon > 0$ there exists $c_\varepsilon > 0$ such that $\lambda^{-1}\big|_{(a+\varepsilon,s_*)} \leq c_\varepsilon$. Since by assumption λ^{-1} is not integrable on (a,s_*), λ^{-1} is unbounded towards the point a. To be more precise, for every $n \in \mathbb{N}$ the set

$$S_n := \{\, s \in (a, a+\tfrac{1}{n})\,;\, \tfrac{1}{\lambda(s)} \geq n \,\} = \{\, s \in (a, a+\tfrac{1}{n})\,;\, \lambda(s) \leq \tfrac{1}{n} \,\}$$

has positive Lebesgue measure. Moreover, taking into account the normalization property (4.16) and the structure of λ, we deduce

$$\text{for all } n \in \mathbb{N} \text{ and almost all } s \in S_n \quad \operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s,\hat{z}(s)), \partial\mathcal{R}(0)) \leq \tfrac{1}{\sqrt{n}} \,. \tag{5.7}$$

Let now $s_n \in S_n$ such that $\operatorname{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s_n,\hat{z}(s_n)), \partial\mathcal{R}(0)) \leq \frac{1}{\sqrt{n}}$. Without loss of generality we assume that the sequence $(s_n)_{n\in\mathbb{N}}$ is decreasing and converging to a. Observe that $-\mathrm{D}\hat{\mathcal{E}}(s_n,\hat{z}(s_n)) \notin \partial\mathcal{R}(0)$ for all n since $\mathrm{m}(\hat{\ell}(s_n),\hat{z}(s_n)) > 0$ on G. Observe further that $\hat{z}$ satisfies the following initial value problem with $z_{0,n} := \hat{z}(s_n)$

$$0 \in \partial\mathcal{R}(\hat{z}'(s)) + \lambda(s)\mathbb{V}\hat{z}'(s) + \mathrm{D}\hat{\mathcal{E}}(s,\hat{z}(s)), \quad s \in (s_n,b),$$

$$\hat{z}(s_n) = z_{0,n}, \quad \mathrm{D}\hat{\mathcal{E}}(s_n,z_{0,n}) \in \mathcal{V}^*.$$

We next rescale this system as follows: For $s \in [s_n, b)$ let $\Lambda(s) := \int_{s_n}^s \frac{1}{\lambda(r)}\,\mathrm{d}r$. The above considerations show that Λ is well defined for all $s \in [s_n, b)$. However, for $s \nearrow b$ one might have $\Lambda(s) \to \infty$. Moreover, Λ is strictly increasing, continuous and the inverse function $\sigma := \Lambda^{-1} : [0, \Lambda(b)) \to [s_n, b)$ exists. For $r \in [0, \Lambda(b))$ let $\widetilde{z}(r) := \hat{z}(\sigma(r))$, $\widetilde{\ell}(r) = \hat{\ell}(\sigma(r))$ and $\widetilde{\mathcal{E}}(r, v) = \hat{\mathcal{E}}(\sigma(r), v) = \mathcal{I}(v) - \langle \widetilde{\ell}(r), v\rangle$. The function $\widetilde{z}$ solves the Cauchy problem

$$0 \in \partial\mathcal{R}(\widetilde{z}'(r)) + \mathbb{V}\widetilde{z}'(r) + \mathrm{D}\widetilde{\mathcal{E}}(r, \widetilde{z}(r)), \quad r \in (0, \Lambda(b)),$$
$$\widetilde{z}(0) = z_{0,n}, \quad \mathrm{D}\widetilde{\mathcal{E}}(0, \widetilde{z}(0)) \in \mathcal{V}^*.$$

Thus, Proposition 3.3 and Remark 3.4 are applicable and imply in particular that $\mathrm{D}\mathcal{I}(\widetilde{z}) \in L^\infty((0, \Lambda(b)); \mathcal{V}^*)$ with a bound that depends on $\|\widetilde{z}(0)\|_{\mathcal{Z}}$, $\mathrm{Var}_{\mathcal{V}^*}(\widetilde{\ell}; [0, \Lambda(b)])$, $\|\widetilde{\ell}\|_{L^\infty(0,\Lambda(b);\mathcal{V}^*)}$ and $\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\widetilde{\mathcal{E}}(0, \widetilde{z}(0)), \partial\mathcal{R}(0))$, only. This immediately translates into $\mathrm{D}\mathcal{I}(\hat{z}) \in L^\infty((s_n, b); \mathcal{V}^*)$ with

$$\begin{aligned}
&\|\mathrm{D}\mathcal{I}(\hat{z})\|_{L^\infty((s_n,b);\mathcal{V}^*)} \\
&\le c\Big(\|\hat{z}(s_n)\|_{\mathcal{Z}} + \mathrm{Var}_{\mathcal{V}^*}(\hat{\ell}, [s_n, b]) + \big\|\hat{\ell}\big\|_{L^\infty(s_n,b;\mathcal{V}^*)} + \mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(s_n, \hat{z}(s_n)), \partial\mathcal{R}(0))\Big) \\
&\le c\big(\|\hat{z}\|_{L^\infty((0,S);\mathcal{Z})} + \mathrm{Var}_{\mathcal{V}^*}(\ell; [0, T]) + \|\ell\|_{L^\infty(0,T;\mathcal{V}^*)} + \frac{1}{\sqrt{n}}\big),
\end{aligned}$$

and the constant c is independent of the chosen solution $\hat{z}$ and of s_n. For $n \to \infty$ we ultimately obtain $\mathrm{D}\mathcal{I}(\hat{z}) \in L^\infty((a, b); \mathcal{V}^*)$ with a bound that depends on the data z_0, ℓ, only.

Case (b): Now we assume that $\lambda^{-1} \in L^1((a, s_*))$ for every $s_* \in (a, b)$. Since G is open and since (by assumption) (a, b) is a maximal connected component of G, we have $a \notin G$ and hence, $\mathrm{m}(\hat{\ell}(a), \hat{z}(a)) = 0$. As above, we rescale the equation by applying the following transformation: Let $\Lambda(s) := \int_a^s \frac{1}{\lambda(r)}\,\mathrm{d}r$ and $\sigma := \Lambda^{-1}$ its inverse function. For $r \in (0, \Lambda(b))$ we define $\widetilde{z}(r) := \hat{z}(\sigma(r))$ and $\widetilde{\ell}(r) := \hat{\ell}(\sigma(r))$. The function $\widetilde{z}$ satisfies the initial value problem

$$\widetilde{z}(0) = \hat{z}(a), \quad 0 \in \partial\mathcal{R}(\widetilde{z}'(r)) + \mathbb{V}\widetilde{z}'(r) + \mathrm{D}\mathcal{I}(\widetilde{z}(r)) - \widetilde{\ell}(r) \text{ for a.a. } r \in (0, \Lambda(b))$$

with $\mathrm{D}\mathcal{I}(\widetilde{z}(0)) - \widetilde{\ell}(0) \in \mathcal{V}^*$. By Proposition 3.3 we have $\mathrm{D}\mathcal{I}(\widetilde{z}) \in L^\infty((0, \Lambda(b)); \mathcal{V}^*)$ with a bound depending only on $\|\hat{z}(a)\|_{\mathcal{Z}}$, on $\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(a, \hat{z}(a)), \partial\mathcal{R}(0))$ and on $\mathrm{Var}_{\mathcal{V}^*}(\widetilde{\ell}, [0, \Lambda(b)])$. This immediately carries over to $\mathrm{D}\mathcal{I}(\hat{z}) \in L^\infty(a, b; \mathcal{V}^*)$ with the same bound. Observe that there exists $\ell_* \in \{\hat{\ell}(a), \hat{\ell}(a+), \hat{\ell}(a-)\}$ with $-\mathrm{D}\mathcal{I}(\hat{z}(a)) + \ell_* \in \partial\mathcal{R}(0)$. Hence,

$$\begin{aligned}
&\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\hat{\mathcal{E}}(a, \hat{z}(a)), \partial\mathcal{R}(0)) \\
&\qquad \le \big\|-\mathrm{D}\hat{\mathcal{E}}(a, \hat{z}(a)) + (\mathrm{D}\mathcal{I}(\hat{z}(a)) - \ell_*)\big\|_{\mathbb{V}^*} \le c_{\mathbb{V}}\big(\|\ell\|_{L^\infty(0,T;\mathcal{V}^*)} + \mathrm{Var}_{\mathcal{V}^*}(\ell, [0, T])\big).
\end{aligned}$$

Combining the estimates derived for the cases (a) and (b) with the estimate derived for $(0, S)\backslash G$ we ultimately arrive at (5.4). Now, (5.5) is an immediate consequence of (5.1) and the estimate (5.4).

Finally, thanks to Proposition C.5, $\hat{z} \in C_{\mathrm{weak}}([0, S]; \mathcal{Z})$, and hence, $\mathrm{D}\mathcal{I}(\hat{z}(\cdot)) \in C_{\mathrm{weak}}([0, S]; \mathcal{Z}^*)$ (by assumption (2.12)). Together with the uniform bound of $\mathrm{D}\mathcal{I}(\hat{z}(\cdot))$ in $\mathcal{V}^*$ the last assertion of claim (3) follows. □

Remark 5.2. Let $(S, \hat{t}, \hat{z}, \hat{\ell})$ be a solution associated with $(\mathcal{I}, \mathcal{R}, \ell, z_0)$ in the sense of Definition 4.6. Let $\hat{\ell}_\pm$ be the left resp. the right continuous version of $\hat{\ell}$. Then $(S, \hat{t}, \hat{\ell}_\pm, \hat{z})$ is a solution associated with $(\mathcal{I}, \mathcal{R}, \ell, z_0)$ in the sense of Definition 4.6, as well.

This can be seen as follows: $\hat{\ell}$ and its left or right continuous version differ in at most countably many points. Thus, (4.14)–(4.15) are valid after replacing $\hat{\ell}$ with $\hat{\ell}_\pm$. Let $G_\pm := \{\, s \in [0, S]\,;\, \mathrm{m}(\hat{\ell}_\pm(s), \hat{z}(s)) > 0\,\}$. Clearly, $G \subseteq G_\pm$ and the sets differ on a set of measure zero, only. Since $\hat{z} \in C([0, S]; \mathcal{V})$ (cf. Proposition C.5), for every $s \in [0, S]$ we have $\int_0^s \langle \hat{\ell}(r), \mathrm{d}\hat{z}(r)\rangle_{\mathcal{V}^*,\mathcal{V}} = \int_0^s \langle \hat{\ell}_\pm(r), \mathrm{d}\hat{z}(r)\rangle_{\mathcal{V}^*,\mathcal{V}}$. This is due to the identity $\int_a^b \langle \chi_{s_*}(r), \mathrm{d}g(r)\rangle_{\mathcal{V}^*,\mathcal{V}} = g(s_*+) - g(s_*-)$ that is valid for every $s_* \in [a, b]$ and every regulated function $g \in G([a, b]; \mathcal{V})$, [Tvr89, Proposition 2.1].

Here, $\chi_{s_*}(s) = 0$ if $s \neq s_*$ and $\chi_{s_*}(s_*) = 1$. Hence, the energy dissipation identity (4.17) remains unaffected by a switch from $\hat{\ell}$ to $\hat{\ell}_\pm$.

As a consequence of the weak continuity of $\mathrm{D}\mathcal{I}(\hat{z}(\cdot))$ in $\mathcal{V}^*$ (see Lemma 5.1) with the same arguments as in the proof of Theorem 4.5 it follows that $G_\pm$ is open. Thus, $\mathrm{D}\mathcal{I}(\hat{z}(\cdot)) \in L^\infty_{\mathrm{loc}}(G_\pm;\mathcal{V}^*)$. Moreover, condition (4.16) holds with $G_\pm$ instead of G. It remains to show that $\hat{z} \in W^{1,1}_{\mathrm{loc}}(G_\pm;\mathcal{V})$. Let $K \Subset G_\pm$ be compact. Then, again by lower semicontinuity, $\inf_{s\in K} \mathrm{m}(\hat{\ell}_\pm(s),\hat{z}(s)) =: c > 0$ which in turn implies (using the normalization property (4.16)) that $\|\hat{z}'(s)\|_{\mathcal{V}} \leq c$ a.e. on K. Since $z \in W^{1,1}_{\mathrm{loc}}(G;\mathcal{V})$ this implies $\hat{z} \in W^{1,\infty}(K \cap G;\mathcal{V})$ and thus ultimately $\hat{z} \in W^{1,1}_{\mathrm{loc}}(G_\pm;\mathcal{V})$.

Proposition 5.3. *Assume* (4.1). *The set $\mathcal{L}(\ell, z_0)$ is compact in the following sense: For every sequence $(S_n, \hat{t}_n, \hat{z}_n, \hat{\ell}_n)_{n\in\mathbb{N}} \subset \mathcal{L}(\ell, z_0)$ there exists a (not relabeled) subsequence and a tuple $(S, \hat{t}, \hat{z}, \hat{\ell}) \in \mathcal{L}(\ell, z_0)$ such that*

$$S_n \to S, \quad \hat{t}_n \overset{*}{\rightharpoonup} \hat{t} \text{ weakly* in } W^{1,\infty}(0,S),\ \hat{t}(S) = T, \tag{5.8}$$

$$\hat{z}_n \rightharpoonup \hat{z} \text{ weakly* in } L^\infty(0,S;\mathcal{Z}) \text{ and uniformly in } C([0,S];\mathcal{V}), \tag{5.9}$$

$$\hat{\ell}_n \overset{*}{\rightharpoonup} \hat{\ell}, \quad \mathrm{D}\mathcal{I}(\hat{z}_n) \overset{*}{\rightharpoonup} \mathrm{D}\mathcal{I}(\hat{z}) \text{ weakly* in } L^\infty(0,S;\mathcal{V}^*), \tag{5.10}$$

and for every $s \in [0, S]$

$$\hat{t}_n(s) \to \hat{t}(s), \quad \hat{z}_n(s) \rightharpoonup \hat{z}(s) \text{ weakly in } \mathcal{Z}\ , \tag{5.11}$$

$$\mathrm{D}\mathcal{I}(\hat{z}_n(s)) \rightharpoonup \mathrm{D}\mathcal{I}(\hat{z}(s)) \text{ weakly in } \mathcal{V}^*, \quad \hat{\ell}_n(s) \rightharpoonup \hat{\ell}(s) \text{ weakly in } \mathcal{V}^*. \tag{5.12}$$

Proof. Let $(S_n, \hat{t}_n, \hat{z}_n, \hat{\ell}_n)_{n\in\mathbb{N}} \subset \mathcal{L}(\ell, z_0)$ and let $(G_n)_n \subset [0, S]$ be the corresponding open sets according to Definition 4.6. Thanks to Lemma 5.1 the bounds (5.2)–(5.5) hold uniformly with respect to n and G_n. Hence, up to a subsequence, $S_n \to S$ for some $S > 0$. Again, if $S > S_n$ we extend all functions by their constant value at S_n. Having in mind the normalization condition (4.16), with Lemma C.5, part (b), there exists $\hat{z} \in AC^\infty([0, S];\mathcal{X}) \cap L^\infty((0, S);\mathcal{Z})$, $\hat{t} \in W^{1,\infty}(0, S)$ and $\hat{\ell} \in BV([0, S];\mathcal{V}^*)$ such that (up to extracting a further subsequence) the convergences in (5.8)–(5.12) hold. Thereby, the convergences of the sequence $\hat{\ell}_n$ follows again from the Banach space valued version of Helly's selection principle [BP86], while the convergences of $\mathrm{D}\mathcal{I}$ follow by the same arguments as in the proof of Proposition 4.2. Moreover, again by the same arguments as in Proposition 4.2 the continuity of $s \mapsto \mathcal{I}(\hat{z}(s))$ ensues. Observe further that the function $s \mapsto \mathrm{D}\mathcal{I}(\hat{z}(s))$ belongs to $C_{\mathrm{weak}}([0, S];\mathcal{V}^*)$.

The characterization of the limit function $\hat{\ell}$ follows by similar arguments as in the proof of Proposition 4.2. Indeed, since for the functions ℓ, ℓ_-, ℓ_+ in each $t \in [0, T]$ the (strong) left and right limits exist and belong to $\{\ell_-(t), \ell_+(t)\}$ and since $\hat{\ell}_n(s) \in \{\ell(\hat{t}_n(s)), \ell_-(\hat{t}_n(s)), \ell_+(\hat{t}_n(s))\}$, the limit $\hat{\ell}(s)$ belongs to $\{\ell(\hat{t}(s)), \ell_-(\hat{t}(s)), \ell_+(\hat{t}(s))\}$ and we even have strong convergence $\hat{\ell}_n(s) \to \hat{\ell}(s)$ in $\mathcal{V}^*$. If $t_* \in [0, T]$ is a point of continuity of ℓ, then from the above, for all $s \in \hat{t}^{-1}(t_*)$ we have $\hat{\ell}(s) = \ell(t_*) = \ell_-(t_*) = \ell_+(t_*)$. Assume now that t_* is not a point of continuity of ℓ with $\ell_-(t_*) \neq \ell_+(t_*)$. Let $s \in [a, b] := \hat{t}^{-1}(t_*)$ with $\hat{\ell}(s) = \ell_+(t_*)$. A proof by contradiction shows that there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have $\hat{t}_n(s) \geq \hat{t}(s) = t_*$. Moreover, by monotonicity of the functions $\hat{t}_n$, for all $n \geq n_0$ and all $r \in [s, b]$ we have $\hat{t}_n(r) \geq t_*$. Hence, $\hat{\ell}_n(r) \to \ell_+(t_*)$, as well. Let $s_+ := \inf\{\, s \in [a, b]\,;\, \hat{\ell}(s) = \ell_+(t_*)\,\}$. In a similar way we define $s_- := \sup\{\, s \in [a, b]\,;\, \hat{\ell}(s) = \ell_-(t_*)\,\}$ and obtain $\hat{\ell}_n(r) \to \ell_-(t_*)$ for all $r \in [a, s_-)$. Thus we have shown that $\hat{\ell}(s) = \ell_-(s)$ if $s \in [a, s_-)$ and $\hat{\ell}(s) = \ell_+(s)$ if $s \in (s_+, b]$. Assume finally that $s_- < s_+$. Then $\hat{\ell}(s) = \ell(t_*)$ for all $s \in (s_-, s_+)$ and for each pair $s_1 < s_2 \in (s_-, s_+)$ there exists $n_0 \in \mathbb{N}$ such that $\hat{\ell}_n(s_1) = \hat{\ell}_n(s_2) = \ell(t_*)$ for all $n \geq n_0$ (proof by contradiction). This implies in particular that $\hat{t}_n(s_1) = \hat{t}_n(s_2) = t_*$ for all $n \geq n_0$ and that $s_1 = s_{*,n}$ and $s_2 = s_{*,n}$ for all n with $s_{*,n}$ from (4.11). But this is a contradiction. Hence, $s_- = s_+$ in this case. For the case $\ell_-(t_*) = \ell_+(t_*) \neq \ell(t_*)$ the arguments can be easily adapted. To summarize, we finally have shown that $\hat{\ell}$ is of the structure (4.11).

It remains to prove that $(S,\hat{t},\hat{z},\hat{\ell})\in\mathcal{L}(\ell,z_0)$. Here, we follow mainly the proof of Theorem 4.5. Due to Proposition C.2 the complementarity relation (4.15) is satisfied by the limit tuple.

Energy dissipation estimate (4.17), $\leq$: Starting from (4.17) written for every n, by lower semicontinuity, the Helly convergence Theorem [MM05, Theorem 3.2], Lemma C.3, and Proposition B.1, we obtain $\liminf_n \mathcal{I}(\hat{z}_n(s))\geq\mathcal{I}(\hat{z}(s))$, $\liminf_n\int_0^s\mathcal{R}[\hat{z}_n'](r)\,\mathrm{d}r\geq\int_0^s\mathcal{R}[\hat{z}'](r)\,\mathrm{d}r$ and $\int_0^s\langle\hat{\ell}_n(r),\mathrm{d}\hat{z}_n(r)\rangle\to\int_0^s\langle\hat{\ell}(r),\mathrm{d}\hat{z}(r)\rangle$.

Let $G:=\{\,s\in[0,S]\,;\,\mathrm{m}(\hat{\ell}(s),\hat{z}(s))>0\,\}$. Like in the proof of Theorem 4.5 it follows that G is open with $0\notin G$. Let $K\subset G$ be compact. With the very same arguments as in the proof of Theorem 4.5 there exists $n_0\in\mathbb{N}$ such that for all $n\geq n_0$ we have $K\subset G_n$ and $\sup_{n\geq n_0}\|\hat{z}_n'\|_{L^\infty(K;\mathcal{V})}<\infty$. Hence, each subsequence of $(\hat{z}_n)_n$ contains a subsubsequence that converges weakly$*$ in $W^{1,\infty}(K;\mathcal{V})$ to $\hat{z}$, whence $\hat{z}\in W^{1,\infty}_{\mathrm{loc}}(G;\mathcal{V})$ and in fact the whole sequence converges. By Proposition C.1 we therefore have the analogue to (4.20). In summary, we have proved (4.17) with $\leq$ instead of equality. By similar arguments we obtain (4.16) with $\geq$ instead of equality. The very same arguments as in the proof of Theorem 4.5 yield the opposite estimate in (4.17) as well as the normalization condition (4.16). Hence, in summary the limit tuple $(S,\hat{t},\hat{z},\hat{\ell})$ belongs to the solution set $\mathcal{L}(\ell,z_0)$. □

Appendix A. Properties of $\mathcal{R}$

We collect here the properties of the dissipation $\mathcal{R}:\mathcal{X}\to[0,\infty)$ and related quantities which are used throughout the paper. Since $\mathcal{R}$ is positively one-homogeneous functional, it holds

$$\eta\in\partial\mathcal{R}(v)\quad\Leftrightarrow\quad\begin{cases}\langle\eta,v\rangle=\mathcal{R}(v)\\ \langle\eta,w\rangle\leq\mathcal{R}(w)\quad\text{for all } w\in\mathcal{Z}\,.\end{cases}$$

It follows from (2.8) that $\partial\mathcal{R}(0)\subset\mathcal{V}^*$ and bounded in $\mathcal{V}^*$-norm (see for instance [Kne18]).

For $\varepsilon>0$, let $\mathcal{R}_\varepsilon:\mathcal{V}\to[0,+\infty)$, $\mathcal{R}_\varepsilon(v):=\mathcal{R}(v)+\frac{\varepsilon}{2}\langle\mathbb{V}v,v\rangle$ be the viscous regularized dissipation potential. Its Fenchel-Moreau conjugate with respect to $\mathcal{V}-\mathcal{V}^*$ duality, $\mathcal{R}_\varepsilon^*:\mathcal{V}^*\to[0,+\infty)$, is defined by $\mathcal{R}_\varepsilon^*(\eta)=\sup\{\langle\eta,v\rangle_{\mathcal{V}^*,\mathcal{V}}-\mathcal{R}_\varepsilon(v):\ v\in\mathcal{V}\}$ and can be explicitly described by

$$\mathcal{R}_\varepsilon^*(\eta)=\frac{1}{2\varepsilon}\big(\operatorname{dist}_{\mathbb{V}}(\eta,\partial\mathcal{R}(0))\big)^2\,.$$

By $\operatorname{dist}_{\mathbb{V}}(\cdot,\partial\mathcal{R}(0))$ we denote the distance of an element of $\mathcal{V}^*$ to $\partial\mathcal{R}(0)(\subset\mathcal{V}^*)$ measured in the norm induced by the operator $\mathbb{V}$: for $\eta\in\mathcal{V}^*$,

$$\operatorname{dist}_{\mathbb{V}}(\eta,\partial\mathcal{R}(0)):=\inf\{\|\eta-\xi\|_{\mathbb{V}^*}:\ \xi\in\partial\mathcal{R}(0)\}\,,\tag{A.1}$$

where $\|\sigma\|_{\mathbb{V}^*}^2=\langle\sigma,\mathbb{V}^{-1}\sigma\rangle$.

Appendix B. Kurzweil integrals and convergence

In this section we use the terminology from [KL09]. Let $\mathcal{W}$ be a Banach space and let $G([a,b];\mathcal{W})$ denote the space of regulated functions $f:[a,b]\to\mathcal{W}$, i.e. the space of those functions for which there exist both one-sided limits $f(t+),f(t-)\in\mathcal{W}$ in every $t\in[a,b]$, see [Die69, KL09]. For functions $f:[a,b]\to\mathcal{W}^*$ and $g:[a,b]\to\mathcal{W}$ we denote with $\int_a^b\langle f(t),\mathrm{d}g(t)\rangle$ ($\langle\cdot,\cdot\rangle$ the dual pairing of $\mathcal{W}$) the Kurzweil integral of f with respect to g. According to [KL09, Theorem 1.9] (see also [Tvr89, Section 2]), the Kurzweil integral of f with respect to g exists provided that $f\in G([a,b];\mathcal{W}^*)$ and $g\in BV([a,b];\mathcal{W})$ or vice versa, i.e. $f\in BV([a,b];\mathcal{W}^*)$ and $g\in G([a,b];\mathcal{W})$. In both cases the following estimate is valid

$$\left|\int_a^b\langle f(t),\mathrm{d}g(t)\rangle\right|\leq\min\Big\{\|f\|_{L^\infty(a,b;\mathcal{W}^*)}\operatorname{Var}_{\mathcal{W}}(g,[a,b]),\\ (\|f(a)\|_{\mathcal{W}^*}+\|f(b)\|_{\mathcal{W}^*}+\operatorname{Var}_{\mathcal{W}^*}(f,[a,b]))\,\|g\|_{L^\infty(a,b;\mathcal{W})}\Big\}\,.\tag{B.1}$$

Proposition B.1. *For $n \in \mathbb{N}$ let $z, z_n \in C([a,b];\mathcal{W})$, $\ell, \ell_n \in BV([a,b];\mathcal{W}^*)$ and assume that $(z_n)_n$ converges uniformly to z. Assume further that*

$$\sup_{n\in\mathbb{N}} \Big(\|\ell_n\|_{L^\infty((a,b),\mathcal{W}^*)} + \mathrm{Var}_{\mathcal{W}^*}(\ell_n,[a,b]) \Big) =: C < \infty$$

and that $\ell_n(t) \rightharpoonup \ell(t)$ weakly$*$ in $\mathcal{W}^*$ for every $t \in [a,b]$. Then $\int_a^b \langle \ell_n(t), \mathrm{d}z_n(t)\rangle \to \int_a^b \langle \ell(t), \mathrm{d}z(t)\rangle$.

Proof. Let $(\ell_n)_n, \ell$ be given according to Proposition B.1. Observe that by lower semicontinuity we obtain $\|\ell\|_{L^\infty((a,b),\mathcal{W}^*)} + \mathrm{Var}_{\mathcal{W}^*}(\ell,[a,b]) \le C$. Assume first that $z \in C^1([a,b];\mathcal{W})$. By [KL09, Prop. 1.10], we have $\int_a^b \langle \ell_n(t), \mathrm{d}z(t)\rangle = (L)\int_a^b \langle \ell_n(t), \dot z(t)\rangle\,\mathrm{d}t$, where the right hand side denotes the Lebesgue integral. Due to the assumptions, the integrand converges pointwise for every t and is uniformly bounded with respect to t and n. Hence, by Lebesgue's Theorem we have $(L)\int_a^b \langle \ell_n, \dot z\rangle\,\mathrm{d}t \to (L)\int_a^b\langle \ell, \dot z\rangle\,\mathrm{d}t = \int_a^b \langle \ell(t), \mathrm{d}z(t)\rangle$.

Since $C^1([a,b];\mathcal{W})$ is dense in $C([a,b];\mathcal{W})$ with respect to the sup norm, this convergence carries over to the case $z \in C([a,b];\mathcal{W})$ in the usual way. Indeed, let $z \in C([a,b];\mathcal{W})$ and choose $\varepsilon > 0$ arbitrarily. Let $\widetilde z \in C^1([a,b];\mathcal{W})$ with $\|z - \widetilde z\|_{L^\infty((a,b);\mathcal{W})} \le \varepsilon/3$. Let $n_\varepsilon \in \mathbb{N}$ such that we have $\left|\int_a^b \langle \ell_n, \mathrm{d}\widetilde z\rangle - \int_a^b\langle \ell, \mathrm{d}\widetilde z\rangle\right| \le \varepsilon C/3$ for all $n \ge n_\varepsilon$. By (B.1), for all $n \ge n_\varepsilon$ it follows

$$\left|\int_a^b \langle \ell_n, \mathrm{d}(\widetilde z - z)\rangle\right| \;\le\; \left(\|\ell_n(a)\|_{\mathcal{W}^*} + \|\ell_n(b)\|_{\mathcal{W}^*} + \mathrm{Var}_{\mathcal{W}^*}(\ell_n,[a,b])\right)\|\widetilde z - z\|_{L^\infty(a,b;\mathcal{W})} \;\le\; \tfrac{C\varepsilon}{3},$$

and similar for ℓ instead of ℓ_n. Thus,

$$\left|\int_a^b \langle \ell_n - \ell, \mathrm{d}z\rangle\right| \le \left|\int_a^b \langle \ell_n, \mathrm{d}(z-\widetilde z)\rangle\right| + \left|\int_a^b \langle \ell_n - \ell, \mathrm{d}\widetilde z\rangle\right| + \left|\int_a^b \langle \ell, \mathrm{d}(z - \widetilde z)\rangle\right| \le C\varepsilon\,,$$

which proves the convergence with $z \in C([a,b],\mathcal{W})$ fixed. Finally, the very same argument provides the general statement of Proposition B.1. ☐

Lemma B.2. *Let $g \in C([a,b];\mathcal{W})$ and $f \in BV([0,S];\mathcal{W}^*)$. Then*

$$\lim_{h\searrow 0}\int_a^{b-h} \langle f(s), h^{-1}(g(s+h) - g(s))\rangle\,\mathrm{d}s = \int_a^b \langle f(s), \mathrm{d}g(s)\rangle\,. \tag{B.2}$$

In the proof we use the notation $(\triangle_h g)(s) := g(s+h) - g(s)$ for $g : [a,b] \to \mathcal{W}$ and $s, h \in \mathbb{R}$. Observe that the product rule $\triangle_h(fg)(s) = (\triangle_h f)(s)g(s+h) + f(s)(\triangle_h g)(s)$ is valid.

Proof. Let $f \in BV([a,b];\mathcal{W}^*)$ and assume first that $g \in C^1([a,b];\mathcal{W})$. Then (B.2) ensues by the Lebesgue convergence theorem and [KL09, Proposition 1.10]. Let now $g \in C([a,b];\mathcal{W})$ and $\varepsilon > 0$ be arbitrary. Then there exists $g_\varepsilon \in C^1([a,b];\mathcal{W})$ such that $\|g - g_\varepsilon\|_{C([a,b];\mathcal{W})} \le \varepsilon$. Hence, for $h > 0$ we obtain using the product rule for finite differences

$$\begin{aligned}&\int_a^{b-h} \langle f(s), h^{-1}\triangle_h(g - g_\varepsilon)(s)\rangle\,\mathrm{d}s\\ &\qquad = \tfrac{1}{h}\left(\int_{b-h}^b \langle f, g - g_\varepsilon\rangle\,\mathrm{d}s - \int_a^{a+h}\langle f, g - g_\varepsilon\rangle\,\mathrm{d}s - \int_a^{b-h}\langle \triangle_h f, (g-g_\varepsilon)(s+h)\rangle\,\mathrm{d}s\right),\end{aligned}$$

which implies that

$$\begin{aligned}&\left|\int_a^{b-h} \langle f(s), h^{-1}\triangle_h(g - g_\varepsilon)(s)\rangle\,\mathrm{d}s\right|\\ &\qquad \le \left(2\,\|f\|_{L^\infty(a,b);\mathcal{W}^*)} + h^{-1}\int_a^{b-h} \|\triangle_h f\|_{\mathcal{W}^*}\,\mathrm{d}s\right)\|g - g_\varepsilon\|_{C([a,b];\mathcal{W})}\,.\end{aligned}$$

Thanks to Lemma C.4, the right hand side is bounded by $2(\|f\|_{L^\infty((a,b);\mathcal{W}^*)} + \mathrm{Var}_{\mathcal{W}^*}(f,[a,b]))\varepsilon$. Standard arguments now finish the proof of (B.2) for arbitrary $g \in C([a,b];\mathcal{W})$. $\square$

Appendix C. Miscellaneous of useful tools

We collect the statements of results useful for our analysis.

C.1. Lower semicontinuity properties.

The following Proposition is a slight variant of [MRS09, Lemma 3.1].

Proposition C.1. *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 $\delta_n, \delta \in L^1(0,S;[0,\infty))$ with $\liminf_{n\to\infty} \delta_n(s) \geq \delta(s)$ for almost all s. Then for every $\alpha \geq 1$*

$$\liminf_{n\to\infty} \int_0^S \|v_n(s)\|_{\mathcal{V}}^\alpha\, \delta_n(s)\,\mathrm{d}s \geq \int_0^S \|v(s)\|_{\mathcal{V}}^\alpha\, \delta(s)\,\mathrm{d}s. \tag{C.1}$$

The next lemma is cited from [MRS12b, Lemma 4.3].

Lemma C.2. *Let $I \subset \mathbb{R}$ be a bounded interval and $f, g, f_n, g_n : I \to [0,\infty)$, $n \in \mathbb{N}$, measurable functions satisfying $\liminf_{n\to\infty} f_n(s) \geq f(s)$ for a.a. $s \in I$ and $g_n \rightharpoonup g$ weakly in $L^1(I)$. Then*

$$\liminf_{n\to\infty} \int_I f_n(s) g_n(s)\,\mathrm{d}s \geq \int_I f(s)g(s)\,\mathrm{d}s\,.$$

C.2. Absolutely continuous functions and BV-functions.

We follow [MRS16, Section 2.2]. Let $\mathcal{X}$ be a Banach space and let $\mathcal{R} : \mathcal{X} \to \mathbb{R}$ be convex, lower semicontinuous, positively homogeneous of degree one and with (2.8). For $1 \leq p \leq \infty$, we define the set of p-absolutely continuous functions (related to $\mathcal{R}$) as

$$AC^p([a,b];\mathcal{X}) := \big\{ z : [a,b] \to \mathcal{X}\,;\, \exists m \in L^p((a,b)),\, m \geq 0,\, \forall s_1 < s_2 \in [a,b] : \\ \mathcal{R}(z(s_2) - z(s_1)) \leq \int_{s_1}^{s_2} m(r)\,\mathrm{d}r \big\}\,. \tag{C.2}$$

Observe that thanks to (2.8) this set coincides with the one defined with $\|\cdot\|_{\mathcal{X}}$ instead of $\mathcal{R}$. Let $z \in AC^p([a,b];\mathcal{X})$. It is shown in [RMS08, Prop. 2.2], [AGS05, Thm. 1.1.2] that for almost every $s \in [a,b]$ the limits

$$\mathcal{R}[z'](s) := \lim_{h\searrow 0} \mathcal{R}((z(s+h) - z(s))/h) = \lim_{h\searrow 0} \mathcal{R}((z(s) - z(s-h))/h)$$

exist and are equal, that $\mathcal{R}[z'] \in L^p((a,b))$ and that $\mathcal{R}[z']$ is the smallest function for which the integral estimate in (C.2) is valid.

Let further $\mathrm{Var}_{\mathcal{R}}(z;[a,b])$ denote the $\mathcal{R}$-variation of $z : [a,b] \to \mathcal{X}$, i.e.

$$\mathrm{Var}_{\mathcal{R}}(z;[a,b]) := \sup_{\text{partitions of } [a,b]} \sum_{i=1}^m \mathcal{R}(z(s_i) - z(s_{i-1})).$$

A proof for the next Lemma can be found in [KT18, Lemma C.1].

Lemma C.3. *For all $p \in (1,\infty]$ and $z \in AC^p([a,b];\mathcal{X})$ we have*

$$\mathrm{Var}_{\mathcal{R}}(z,[a,b]) = \int_a^b \mathcal{R}[z'](s)\,\mathrm{d}s. \tag{C.3}$$

The following Lemma is proved in [Leo17, Theorem 2.20]:

Lemma C.4. *For every $f \in BV([a,b],\mathcal{X})$ we have*

$$\sup_{0<h<(b-a)} h^{-1} \int_a^{b-h} \|f(s+h) - f(s)\|_{\mathcal{X}} \leq \mathrm{Var}_{\mathcal{X}}(f,[a,b]). \tag{C.4}$$

C.3. **A combination of Helly's Theorem and the Ascoli-Arzelà Theorem.** The general statements of the following theorem can be found in [MRS16, AGS05]. For a proof tailored to our specific situation we refer to [KT18, Proposition D.1].

Proposition C.5. *Let $\mathcal{Z}$ be a reflexive Banach space, $\mathcal{V}, \mathcal{X}$ further Banach spaces such that (2.1) is satisfied and assume that $\mathcal{R} : \mathcal{X} \to [0, \infty)$ complies with (2.8).*

(a) The set $AC^1([a, b]; \mathcal{X}) \cap L^\infty((a, b); \mathcal{Z})$ is contained in $C([a, b]; \mathcal{V})$ and there exists $C > 0$ such that for all $z \in AC^1([a, b]; \mathcal{X}) \cap L^\infty((a, b); \mathcal{Z})$ we have

$$\|z\|_{C([a,b];\mathcal{V})} \leq C(\|z\|_{L^\infty((a,b);\mathcal{Z})} + \|\mathcal{R}[z']\|_{L^1((a,b))}).$$

(b) Let $(z_n)_n \subset AC^\infty([a, b]; \mathcal{X}) \cap L^\infty((a, b); \mathcal{Z})$ 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 $z \in AC^\infty([a, b]; \mathcal{X}) \cap L^\infty((a, b); \mathcal{Z})$ and a (not relabeled) 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 $L^\infty((a, b); \mathcal{Z}) \cap C([a, b]; \mathcal{V}) \subset C_{weak}([a, b]; \mathcal{Z})$.

C.4. **Chain rule.** The following chain rule is proved in [KT18, Prop. E.1].

Proposition C.6. *Let $z \in H^1((0, T); \mathcal{V}) \cap L^\infty((0, T); \mathcal{Z})$ and assume that $\mathrm{D}\mathcal{I}(z(\cdot)) \in L^\infty((0, T); \mathcal{V}^*)$. Then for almost all t, the mapping $t \mapsto \mathcal{I}(z(t))$ is differentiable and we have the identity*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}(z(t)) = \langle Az(t), \dot{z}(t)\rangle_{\mathcal{V}^*,\mathcal{V}} + \langle \mathrm{D}\mathcal{F}(z(t)), \dot{z}(t)\rangle_{\mathcal{V}^*,\mathcal{V}}\,.$$

Integrated version of the chain rule: Let $z \in W^{1,1}((0, T); \mathcal{V}) \cap L^\infty((0, T); \mathcal{Z})$ with $\mathrm{D}\mathcal{I}(z(\cdot)) \in L^\infty((0, T); \mathcal{V}^*)$ and assume that $t \mapsto \mathcal{I}(z(t))$ is continuous on $[0, T]$. Then for all $t_1 < t_2 \in [0, T]$

$$\mathcal{I}(z(t_2)) - \mathcal{I}(z(t_1)) = \int_{t_1}^{t_2} \langle \mathrm{D}\mathcal{I}(z(r)), \dot{z}(r)\rangle_{\mathcal{V}^*,\mathcal{V}}\,\mathrm{d}r. \tag{C.7}$$

Acknowledgments

The authors are grateful to Alex Mielke for the inspiring discussions and the permanent scientific support. This research has been partially funded by Deutsche Forschungsgemeinschaft (DFG) through the Priority Programme SPP 1962 Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization, Project P13 Simulation and Optimization of Rate-Independent Systems with Non-Convex Energies. CZ is a member of GNAMPA-INdAM. DK acknowledges the kind hospitality of the DISMA, Politecnico di Torino, and CZ acknowledges the kind hospitality of the University of Kassel.

References

[AGS05] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Basel: Birkhäuser, 2005.

[BP86] V. Barbu and Th. Precupanu. Convexity and optimization in Banach spaces. 2nd rev. and extended ed. Transl. from the Romanian., 1986.

[Die69] Jean Dieudonné. Foundations of modern analysis. Enlarged and corrected printing. New York-London: Academic Press. XV, 387 p. (1969)., 1969.

[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.

[KL02] Pavel Krejčí and Philippe Laurençot. Generalized variational inequalities. *J. Convex Anal.*, 9(1):159–183, 2002.

[KL09] Pavel Krejčí and Matthias Liero. Rate independent Kurzweil processes. *Appl. Math., Praha*, 54(2):117–145, 2009.

[Kne18] Dorothee Knees. Convergence analysis of time-discretization schemes for rate-independent systems. *ESAIM, Control Optim. Calc. Var.*, 2018. accepted.

[KR14] Pavel Krejčí and Vincenzo Recupero. Comparing BV solutions of rate independent processes. *J. Convex Anal.*, 21(1):121–146, 2014.

[KT18] Dorothee Knees and Stephanie Thomas. Optimal control of a rate-independent system constrained to parametrized balanced viscosity solutions. arxiv:1810.12572, University of Kassel, 2018.

[Leo17] Giovanni Leoni. *A first course in Sobolev spaces. 2nd edition.*, volume 181. Providence, RI: American Mathematical Society (AMS), 2nd edition edition, 2017.

[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.

[Mor77] Jean-Jacques Moreau. Evolution problem associated with a moving convex set in a Hilbert space. *J. Differential Equations*, 26(3):347–374, 1977.

[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] A. Mielke, R. Rossi, and G. 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.

[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.

[MZ14] A. Mielke and S. Zelik. On the vanishing-viscosity limit in parabolic systems with rate-independent dissipation terms. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 13(1):67–135, 2014.

[Rec11] Vincenzo Recupero. *BV* solutions of rate independent variational inequalities. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 10(2):269–315, 2011.

[Rec16] Vincenzo Recupero. Sweeping processes and rate independence. *J. Convex Anal.*, 23(3):921–946, 2016.

[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.

[RS18] Vincenzo Recupero and Filippo Santambrogio. Sweeping processes with prescribed behavior on jumps. *Ann. Mat. Pura Appl. (4)*, 197(4):1311–1332, 2018.

[Tvr89] Milan Tvrdý. Regulated functions and the Perron-Stieltjes integral. *Čas. Pěstování Mat.*, 114(2):187–209, 1989.

Dorothee Knees, Institute of Mathematics, University of Kassel, Heinrich-Plett Str. 40, 34132 Kassel, Germany. Phone: +49 0561 8044355

E-mail address: dknees@mathematik.uni-kassel.de

Department of Mathematical Sciences, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

E-mail address: chiara.zanini@polito.it