Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 0946 – 8633

Balanced Viscosity (BV) solutions to infinite-dimensional rate-independent systems

Alexander Mielke [1,2], Riccarda Rossi [3], Giuseppe Savaré[4]

submitted: September 23, 2013

[1] Weierstraß-Institut
Mohrenstraße 39
10117 Berlin
Germany
E-Mail: alexander.mielke@wias-berlin.de

[2] Institut für Mathematik
Humboldt-Universität zu Berlin
Rudower Chaussee 25
12489 Berlin-Adlershof
Germany

[3] DICATAM – Sezione di Matematica
Università di Brescia
via Valotti 9
I–25133 Brescia
Italy
E-Mail: riccarda.rossi@ing.unibs.it

[4] Dipartimento di Matematica "F. Casorati"
Università di Pavia
Via Ferrata 1
27100 Pavia
Italy
E-Mail: giuseppe.savare@unipv.it

No. 1845

Berlin 2013

2010 *Mathematics Subject Classification.* 49Q20, 58E99.

Key words and phrases. Doubly nonlinear equations, generalized gradient flows, rate-independent systems, vanishing-viscosity limit, variational Gamma convergence, energy-dissipation balance, arclength parameterized solutions.

A.M. has been partially supported by the ERC grant no.267802 AnaMultiScale. R.R. and G.S. have been partially supported by a MIUR-PRIN'10-11 grant for the project "Calculus of Variations".

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract. Balanced Viscosity solutions to rate-independent systems arise as limits of regularized rate-independent flows by adding a superlinear vanishing-viscosity dissipation.

We address the main issue of proving the existence of such limits for infinite-dimensional systems and of characterizing them by a couple of variational properties that combine a local stability condition and a balanced energy-dissipation identity.

A careful description of the jump behavior of the solutions, of their differentiability properties, and of their equivalent representation by time rescaling is also presented.

Our techniques rely on a suitable chain-rule inequality for functions of bounded variation in Banach spaces, on refined lower semicontinuity-compactness arguments, and on new BV-estimates that are of independent interest.

Contents

1. Introduction

This paper concerns the asymptotic behavior of the solutions $u_\varepsilon : [0,T] \to V$, $\varepsilon \downarrow 0$, of singularly perturbed doubly nonlinear evolution equations of the type

$$\partial\Psi_\varepsilon(\dot u_\varepsilon(t)) + \partial\mathcal{E}_t(u_\varepsilon(t)) \ni 0 \quad \text{in } V^*, \quad t \in (0,T). \tag{1.1}$$

Here $(V, \|\cdot\|)$ is a Banach space satisfying the Radon-Nikodým property (e.g. a reflexive space, see [DiU77]), $\partial\mathcal{E}$ is the Fréchet subdifferential of a time-dependent energy functional $\mathcal{E} : [0,T] \times V \to (-\infty, +\infty]$, and $\Psi_\varepsilon : V \to [0,+\infty)$ is a family of convex and superlinear dissipation potentials; the main coercivity and structural assumptions on $\Psi_\varepsilon, \mathcal{E}$ will be discussed in Section 2.1.

The main feature we want to address here is the degeneration of the superlinear character of Ψ_ε as $\varepsilon \downarrow 0$, approximating a degree-1 positively homogeneous convex potential $\Psi : V \to [0,+\infty)$,

$$\Psi(\lambda v) = \lambda\Psi(v) \quad \text{for every } v \in V,\ \lambda \geq 0; \qquad \Psi(v) > 0 \quad \text{if } v \neq 0. \tag{1.2}$$

An important example motivating our investigation is the vanishing quadratic approximation

$$\Psi_\varepsilon(v) = \Psi(v) + \frac{\varepsilon}{2}\|v\|^2, \quad \text{associated with the viscous potential} \quad \Phi(v) := \frac{1}{2}\|v\|^2. \tag{1.3}$$

The superlinear case. Equations of the type (1.1) arise in several contexts, ranging from thermomechanics to the modeling of rate-independent evolution. In the realm of these applications, (1.1) may be interpreted as *generalized balance relation*, balancing viscous and potential forces.

The analysis of (1.1) when the energy $\mathcal{E}$ has the typical form

$$\mathcal{E}_t(u) = \mathcal{E}(u) - \langle \ell(t), u\rangle, \quad \text{with } \ell : [0,T] \to V^* \text{ smooth and } \mathcal{E} : V \to (-\infty, +\infty] \textit{ convex}$$

goes back to the seminal papers [CoV90, Col92]. Therein, the existence of absolutely continuous solutions to the Cauchy problem for (1.1) was proved by means of maximal monotone operator techniques. Existence and approximation results for a broad class of *nonconvex* energies, also featuring a singular dependence on time, have been recently obtained in [MRS13], relying on various contributions from the theory of curves of Maximal Slope [DGMT80, MST89] and from the variational approach to gradient flows [De 93, RoS06, AGS08, RMS08].

Positively* 1*-homogeneous dissipations: energetic solutions. Since Ψ is positively homogeneous of degree 1, when $\varepsilon = 0$ the formal limit of (1.1)

$$\partial\Psi(\dot{u}(t)) + \partial\mathcal{E}_t(u(t)) \ni 0 \quad \text{in } V^*, \quad t \in (0,T), \tag{1.4}$$

describes a rate-independent evolution. In this case, even for convex energies $\mathcal{E}_t(\cdot)$ one cannot expect the existence of absolutely continuous solutions to (1.1): in general, they may be only BV with respect to time and in fact have jumps, so that even the precise meaning of the differential inclusion (1.4) is a delicate question.

This has called for *weak-variational* characterization of the solutions of (1.4), leading to the concept of *energetic solution* to the rate-independent system $(V, \mathcal{E}, \Psi)$: it dates back to [MiT99] and was further developed in [MiT04, DFT05], see also [Mie05, Mie11] and the references therein.

In this setting, $u : [0,T] \to V$ is an energetic solution to equation (1.4), if it complies with the *global stability* (S) and with the *energy balance* (E) conditions

$$\forall\, v \in V : \quad \mathcal{E}_t(u(t)) \le \mathcal{E}_t(v) + \Psi(v - u(t)) \quad \text{for all } t \in [0,T], \tag{S}$$

$$\mathrm{Var}_\Psi(u; [0,t]) + \mathcal{E}_t(u(t)) = \mathcal{E}_0(u(0)) + \int_0^t \partial_t\mathcal{E}_s(u(s))\,\mathrm{d}s \quad \text{for all } t \in [0,T], \tag{E}$$

where $\mathrm{Var}_\Psi(u; [a,b])$ is the total variation induced by $\Psi(\cdot)$ on the interval $[a,b] \subset [0,T]$, viz.

$$\mathrm{Var}_\Psi(u; [a,b]) := \sup\Big\{ \sum_{m=1}^{M} \Psi\big(u(t_m) - u(t_{m-1})\big) : a = t_0 < t_1 < \cdots < t_{M-1} < t_M = b \Big\}. \tag{1.5}$$

The energetic formulation (S)–(E) has several strong points: it is derivative-free, it bypasses all the technical differentiability issues on V (related to the validity of the Radon-Nikodým property), on u (related to its behavior on the Cantor-Jump set), and on the energy $\mathcal{E}$ (related to its Frechét subdifferential). Furthermore, it provides nice existence-stability results under simple coercivity and time-regularity assumptions.

Nonetheless, in the case of *nonconvex* energies it is now well known [MRS09, MRS12a, Mie11, MiZ12, RoS13] that the global stability condition (S) involves a variational characterization of the jump behavior of the system, that is affected by the whole energetic landscape of $\mathcal{E}$.

Positively* 1*-homogeneous dissipations: the vanishing-viscosity approach. The by now well-established *vanishing-viscosity* approach aims to find good local conditions describing rate-independent evolution (and in particular the behavior of the solutions at jumps). It also leads to a clarification of the connections with the metric-variational theory of gradient flows.

While referring to [MRS12a] for a more detailed survey, here we recall the works where the vanishing-viscosity analysis is carried out via the *reparameterization technique* introduced in [EfM06]. They range from applicative contexts in material modeling (such as crack propagation

[KMZ08, KZM10], Cam-Clay and non-associative plasticity [DMDS11, DDS12, BFM12], and damage [KRZ13]), to the analysis of parabolic PDEs with rate-independent dissipation terms [MiZ12].

Abstract rate-independent systems in a finite-dimensional setting have been studied by [MRS09, MRS12a]. In particular, the vanishing-viscosity limit of gradient systems of the type (1.1) has been studied in [MRS12a] when V is a *finite*-dimensional space and $\mathcal{E} \in \mathrm{C}^1([0,T] \times V)$. Here we aim to generalize the results from [MRS12a] to the present *nonsmooth, infinite-dimensional* setting.

A simple prototype of the situation we have in mind (see also Section 5) is

$$V = L^2(\Omega), \quad \Psi_\varepsilon(v) = \int_\Omega |v| + \frac{\varepsilon}{2}|v|^2 \,\mathrm{d}x, \quad \mathcal{E}_t(u) = \int_\Omega \left(\frac{1}{2}|\nabla u|^2 + W(u) - \ell(t)u\right) \mathrm{d}x \tag{1.6}$$

where Ω is a bounded open subset of $\mathbb{R}^d$, $\ell \in \mathrm{C}^1([0,T];L^2(\Omega))$ and $W \in \mathrm{C}^1(\mathbb{R})$ is, e.g., a double-well type nonlinearity. The abstract subdifferential inclusion (1.1) leads to the nonlinear parabolic equation

$$\varepsilon\,\partial_t u + \mathrm{Sign}\big(\partial_t u\big) - \Delta u + W'(u) = \ell \quad \text{in } \Omega \times (0,T), \tag{1.7}$$

for which the vanishing-viscosity limit $\varepsilon \downarrow 0$ was in fact analyzed in [MiZ12], based on the *reparameterization* technique and on the concept of *parameterized solution*, from [EfM06].

In this work we will propose a direct characterization of the limit evolution, in the same spirit of conditions (S)–(E), and we will show how it is related to a parameterized formulation. A particular emphasis will be on the crucial property encoded in the *balanced energy–dissipation identities*, both in the original and in the rescaled time variables. The notion of *Balanced Viscosity* (BV) *solution* to a rate-independent system tries to capture this essential feature.

Balanced Viscosity (BV) ***solutions.*** Let us briefly describe what we mean by a *balanced viscosity* (BV) *solution to the rate-independent system* (RIS) $(V, \mathcal{E}, \Psi, \Phi)$, where now also the viscosity correction induced by Φ characterizes the evolution. To simplify the exposition in this introduction, we suppose that Ψ is V-coercive, i.e. $\Psi(v) \geq c\|v\|$ for all $v \in V$ and for a constant $c > 0$.

A crucial role is played by the dual convex set

$$K^* := \{\xi \in V^* : \langle \xi, v\rangle \leq \Psi(v) \text{ for every } v \in V\} \tag{1.8}$$

whose support function is Ψ. Following [MRS12a], we say that a curve $u \in \mathrm{BV}([0,T];V)$ is a BV solution to the RIS $(V, \mathcal{E}, \Psi, \Phi)$, if it fulfills the following *local stability condition*

$$K^* + \partial\mathcal{E}_t(u(t)) \ni 0 \quad \text{for all}\, t \in [0,T] \setminus \mathrm{J}_u, \tag{$\mathrm{S}_{\mathrm{loc}}$}$$

where J_u is the jump set of u, and the *Energy-Dissipation Balance*

$$\mathit{Var}_{\mathfrak{f}}(u;[0,t]) + \mathcal{E}_t(u(t)) = \mathcal{E}_0(u(0)) + \int_0^t \partial_t \mathcal{E}_s(u(s))\,\mathrm{d}s \quad \text{ for all } t \in [0,T]. \tag{$\mathrm{E}_{\mathfrak{f}}$}$$

Like (E), ($\mathrm{E}_{\mathfrak{f}}$) as well balances at every evolution time $t \in [0,T]$ the energy dissipated by the system and the current energy, with the initial energy and the work of the external forces. However, in ($\mathrm{E}_{\mathfrak{f}}$) dissipation is measured by the total variation functional $\mathit{Var}_{\mathfrak{f}}$. While referring to the forthcoming Definition 3.6 for a precise formula, we may mention here that the main difference of $\mathit{Var}_{\mathfrak{f}}$ with respect to Var_Ψ concerns the contribution of the jumps. In fact, in the definition of $\mathit{Var}_{\mathfrak{f}}$ the cost $\Psi(u(t_+) - u(t_-))$ of the transition from the left limit $u(t_-)$ to the right limit $u(t_+)$ at a time $t \in \mathrm{J}_u$ is replaced by the *Finsler dissipation cost*

$$\Delta_{\mathfrak{f}_t}(u_0, u_1) := \inf\left\{ \int_0^1 \mathfrak{f}_t(\vartheta;\dot\vartheta)\,\mathrm{d}r : \vartheta \in \mathrm{AC}([0,1];V),\ \vartheta(0) = u(t_-),\ \vartheta(1) = u(t_+)\right\}, \tag{1.9}$$

where

$$\mathfrak{f}_t(\vartheta;\dot\vartheta) = \Psi(\dot\vartheta) + \mathfrak{e}_t(\vartheta)\|\dot\vartheta\|, \quad \mathfrak{e}_t(\vartheta) := \inf\left\{\|\xi - z\|_* : \xi \in -\partial\mathcal{E}_t(\vartheta),\ z \in K^*\right\}. \tag{1.10}$$

Formula (1.10) clearly shows that the Finsler dissipation cost (1.9) (and thus the total variation $Var_{\mathfrak{f}}$) encompasses both *rate-independent* effects through $\Psi(\cdot)$, and *viscous* effects through $\|\cdot\|$. The latter are active whenever $\mathfrak{e}_t(\vartheta) > 0$, precisely when the local stability condition ($\mathrm{S}_{\mathrm{loc}}$) is violated, since $K^* + \partial\mathcal{E}_t(u) \ni 0$ if and only if $\mathfrak{e}_t(u) = 0$. Ultimately, by virtue of ($\mathrm{E}_{\mathfrak{f}}$), *viscous* dissipation enters in the description of the energetic behavior of the system at jumps.

The link between the particular structure of (1.10) and the vanishing-viscosity approximation (1.1) can be better understood by recalling the strucure of the energy-dissipation balance satisfied by the solutions to the viscous evolution:

$$\mathcal{E}_t(u_\varepsilon(t))+\int_0^t \Big(\Psi_\varepsilon(\dot u_\varepsilon)+\Psi_\varepsilon^*(\xi_\varepsilon)\Big)\,\mathrm{d}r = \mathcal{E}_0(u_\varepsilon(0))+\int_0^t \partial_t\mathcal{E}_r(u_\varepsilon(r)), \quad \xi_\varepsilon(r) \in -\partial\mathcal{E}_r(u_\varepsilon(r)). \tag{1.11}$$

It turns out that $\mathfrak{f}_t$ admits the variational representation

$$\mathfrak{f}_t(\vartheta,\dot\vartheta) = \inf\Big\{\Psi_\varepsilon(\dot\vartheta) + \Psi_\varepsilon^*(\xi) : \xi \in -\partial\mathcal{E}_t(\vartheta),\ \varepsilon > 0\Big\}. \tag{1.12}$$

This feature is in some sense reflected by the so-called *optimal jump transitions* connecting $u(t_-)$ and $u(t_+)$: they are curves $\vartheta \in \mathrm{AC}([0,1];V)$ which attain the infimum in formula (1.9) and keep track of the asymptotic profile of the converging solutions u_ε around a jump point. By means of a careful rescaling technique, we will show that optimal transitions fulfill the doubly nonlinear equation

$$\partial\Psi(\dot\vartheta(r)) + \partial\Phi(\varepsilon(r)\dot\vartheta(r)) + \partial\mathcal{E}_t(\vartheta(r)) \ni 0 \quad \text{for a.a. } r \in (0,1) \tag{1.13}$$

for some map $r \mapsto \varepsilon(r) \in [0,+\infty)$.

Lack of differentiability and non-coercive rate-independent dissipations. Up to now, for the sake of simplicity, we have overlooked one crucial issue in the analysis of the rate-independent equation (1.4), namely the lack of differentiability of the limiting solution u when Ψ is not coercive with respect to the norm $\|\cdot\|$ on V (as in the example (1.6)). Even the introduction of a weaker norm cannot avoid this technical issue, since in many interesting examples norms of L^1-type do not comply with the Radon-Nikodým property.

This fact leads to significant technical difficulties, in that Ψ-absolutely continuous curves need not be pointwise differentiable with respect to time. Hence, for example formulae (1.9)–(1.10) need to be carefully modified by introducing the convenient notion of the metric Ψ-derivative, and differential inclusions like (1.13) have to be suitably interpreted.

On the other hand, we will show that under slightly stronger assumptions on the energy functional $\mathcal{E}$, limiting solutions still belong to $\mathrm{BV}([0,T];V)$ even in the case of a degenerate rate-independent dissipation Ψ. For this class of V-parameterizable solutions we can recover a more precise differential characterization, and several expressions take a simpler form.

Main results and plan of the paper. In this paper we provide existence and approximation results for Balanced Viscosity solutions to the RIS $(V,\mathcal{E},\Psi,\Phi)$ under quite general conditions on the dissipation potentials Ψ, Φ and on the energy functional $\mathcal{E}$, enlisted in Section 2.1. Let us mention in advance that, our standing assumptions on $\mathcal{E}$ guarantee the lower semicontinuity, coercivity, uniform subdifferentiability of the functional $u \mapsto \mathcal{E}_t(u)$, and (sufficient) smoothness of the time-dependent function $t \mapsto \mathcal{E}_t(u)$. In §2.2 we provide some preliminary results on absolutely continuous and BV curves, while the main existence and structural properties of viscous gradient systems are recalled in § 2.3–§ 2.4.

In Section 3 we present our main results concerning Balanced Viscosity solutions. The Finsler cost (1.9) and its related total variation are discussed in § 3.1. In **Theorem 3.9** we state the relative compactness of viscous solutions $(u_\varepsilon)_\varepsilon$ to (1.1) with respect to pointwise convergence, and we show that any limit point as $\varepsilon \downarrow 0$ is a BV solution. A similar result (**Theorem 3.10**) addresses the passage to the limit in the time-incremental minimization scheme [De 93] for the viscous problem: given a time step $\tau > 0$, the uniform partition $t_n := n\tau$, $n = 0, \cdots, N_\tau$, of the time interval $[0,T]$ so that $\tau(N_\tau-1)<T\le\tau N_\tau$, and an initial datum $\mathrm{U}^0_{\tau,\varepsilon}$, the scheme produces discrete sequences $(\mathrm{U}^n_{\tau,\varepsilon})$, $n\in\mathbb{N}$, by solving the minimization problem

$$\mathrm{U}^n_{\tau,\varepsilon}\in\operatorname*{Argmin}_{\mathrm{U}\in V}\left\{\tau\Psi_\varepsilon\Big(\frac{\mathrm{U}-\mathrm{U}^{n-1}_{\tau,\varepsilon}}{\tau}\Big)+\mathcal{E}_{t_n}(\mathrm{U})\right\}\quad\text{for } n=1,\cdots,N_\tau. \tag{IP$_{\varepsilon,\tau}$}$$

As τ, $\varepsilon\downarrow0$ with $\tau/\varepsilon\downarrow0$ we will prove that the piecewise affine interpolants (see (7.25)) $(\mathrm{U}_{\tau,\varepsilon})_{\tau,\varepsilon}$ of the discrete values $\mathrm{U}^n_{\tau,\varepsilon}$ converge (up to subsequences) to a BV solution of the RIS $(V,\mathcal{E},\Psi,\Phi)$. Under slightly stronger assumptions on the energy functional $\mathcal{E}$, **Theorems 3.21 and Corollary 3.23** show that the limits obtained by this variational scheme belong to $\mathrm{BV}([0,T];V)$ and are V-parameterizable, a distinguished class of solutions studied in § 3.4. Other important properties of BV solutions are discussed in § 3.2 and 3.3: the latter is focused in particular on the notion of **optimal jump transition**, a useful tool to describe the asymptotic profile of the solution u_ε around a jump limit point.

We discuss parameterized solutions in Section 4: **Theorem 4.3** provides the main existence and convergence result, the tight connections with BV solutions are clarified in **Theorem 4.7**, and the case of V-parameterized solutions is investigated in Section 4.2.

Section 5 is devoted to a series of examples, where we discuss the validity of the abstract conditions on the energy enucleated in § 2.1, and in particular of the chain-rule inequality. Furthermore, Example 5.2 shows that there exist BV solutions which are not V-parameterizable. Most of the proofs and of the technical tools are collected in the last three sections. **Section 6** is devoted to the main theme of the chain-rule inequalities in the parameterized (§ 6.1) and BV setting (§ 6.2).

Section 7 contains the main stability, compactness, and lower semicontinuity results that lie at the core of our proofs. In § 7.1 and § 7.2 we alternate the parameterized and the non-parameterized point of view to describe the limit of various integral functionals. The crucial lower semicontinuity result in the BV setting is **Proposition 7.3**, where we adapt ideas introduced in [MRS12b]. The proofs of the main Theorems are eventually collected in §7.3. The crucial BV estimate for the discrete Minimizing Movements leading to **V-parameterizable solutions** are collected in § 7.4.

2. Notation, assumptions and preliminary results

2.1. **The energy-dissipation framework.** Throughout the present paper we will suppose that

$$(V,\|\cdot\|)\text{ is a separable Banach space satisfying the Radon-Nikodým property.} \tag{2.1}$$

This means that absolutely continuous curves with values in V are $\mathscr{L}^1$-a.e. differentiable, see Section 2.2. This condition is certainly satisfied if V is reflexive or if it is the dual of a separable Banach space, see [DiU77]. With $\|\cdot\|_*$ we will denote the dual norm in V^*, while $\langle\cdot,\cdot\rangle$ stands for the duality pairing between V^* and V.

Rate-independent and viscous dissipation. On V are defined two

$$\text{continuous convex dissipation potentials }\Psi,\Phi:V\to[0,+\infty),\text{ strictly positive in }V\setminus\{0\}. \tag{D.0}$$

The *"rate-independent"* potential Ψ is positively 1-homogeneous (a "gauge" functional, [Roc70])

$$\Psi(\lambda v)=\lambda\Psi(v)\quad\text{for all }\ \lambda\ge0\text{ and }\ v\in V. \tag{D.1}$$

Notice that if $\Psi(-v)=\Psi(v)$ for every $v\in V$, then Ψ is a norm in V; we will say that Ψ is coercive if $\Psi(v)\ge c\|v\|$ for every $v\in V$ and some constant $c>0$. However, in general we will not assume any coercivity on Ψ, so that the sublevel sets $\{v\in V:\Psi(v)\le r\}$ are not bounded.

Coercivity will be recovered by the addition of a *"viscous"* dissipation potential Φ of the form

$$\begin{gathered}\Phi(v)=F(\|v\|)\quad\text{for }F\in\mathrm{C}^1([0,+\infty))\text{ convex, with}\\ F(r)>0\text{ for }r>0,\ F(0)=F'(0)=0,\ \lim_{r\uparrow+\infty}F'(r)=+\infty.\end{gathered} \tag{D.2}$$

We then consider a vanishing-viscosity family $\Psi_\varepsilon : V \to [0,+\infty)$, $\varepsilon > 0$, of dissipation potentials approximating Ψ:

$$\Psi_\varepsilon(v) := \Psi(v) + \varepsilon^{-1}\Phi(\varepsilon v) =: \varepsilon^{-1}\Psi_1(\varepsilon v), \quad \Psi_0(v) := \Psi(v) = \lim_{\varepsilon\downarrow 0} \Psi_\varepsilon(v) = \inf_{\varepsilon>0} \Psi_\varepsilon(v). \tag{2.2}$$

Observe that the whole theory is restricted to the case $\Psi_\varepsilon(v) < +\infty$. Indeed, allowing for $\Psi_\varepsilon(v) = +\infty$ as in unidirectional processes such as damage, hardening, or fracture (cf., e.g., [DFT05, MiR06, MaM09, KRZ13, BFM12]) would give rise to additional complications, which we prefer not to address in this paper. Still, a typical situation that is relevant in elastoplasticity is given by the choices $V = L^p(\Omega;\mathbb{R}^m)$ for $p \in (1,\infty)$, $\|v\| = (\int_\Omega |v(x)|^p \,\mathrm{d}x)^{1/p}$, $\Psi(v) = \int_\Omega \sigma_Y |v(x)| \,\mathrm{d}x$, and $F(r) = \nu r^p$. In particular, Ψ_ε has the simple form $\Psi_\varepsilon(v) = \int_\Omega \sigma_Y |v(x)| + \varepsilon^{p-1}\nu |v(x)|^p \,\mathrm{d}x$.

Subdifferential of the rate-independent dissipation and the dual convex stability set. Ψ is the support function of the w^*-closed and bounded convex subset of V^*

$$K^* := \Big\{\xi \in V^* : \langle \xi, w\rangle \le \Psi(w) \text{ for every } w \in V\Big\} \subset V^*, \quad \Psi(v) = \sup_{\xi\in K^*} \langle \xi, v\rangle, \tag{2.3}$$

which will play a prominent role in the following. K^* is related to Ψ by two different important relations: first of all, it is the proper domain of the conjugate function of Ψ^*:

$$\Psi^*(\xi) := \sup_{v\in V} \big(\langle \xi, v\rangle - \Psi(v)\big) = \mathrm{I}_{K^*}(\xi) = \begin{cases} 0 & \text{if } \xi \in K^*, \\ +\infty & \text{otherwise.} \end{cases} \tag{2.4}$$

Second, K^* can be characterized in terms of the subdifferential $\partial\Psi : V \rightrightarrows V^*$ of Ψ, defined as

$$\xi \in \partial\Psi(v) \quad \Leftrightarrow \quad \langle \xi, w - v\rangle \le \Psi(w) - \Psi(v) \quad \forall w \in V, \tag{2.5}$$

so that

$$K^* = \partial\Psi(0); \qquad \xi \in \partial\Psi(v) \quad \Leftrightarrow \quad \xi \in K^* \text{ and } \langle \xi, v\rangle = \Psi(v). \tag{2.6}$$

The energy functional and its subdifferential. We shall consider a time-dependent

$$\text{lower semicontinuous energy functional } \mathcal{E} : [0,T] \times D \to \mathbb{R},\ D \subset V. \tag{E.0}$$

To simplify some formulae, we will set $\mathcal{E}_t(u) = +\infty$ if $u \notin D$ and we will assume the following properties:

Coercivity: the map

$$u \mapsto \mathcal{G}(u) := \Psi(u) + \sup_{t\in[0,T]} \mathcal{E}_t(u) \quad \text{has compact sublevels in } V, \\ \text{i.e. for every } E > 0 \text{ the set } D_E := \{u \in D : \mathcal{G}(u) \le E\} \text{ is compact.} \tag{E.1}$$

Power-control: for all $u \in D$ the function $t \mapsto \mathcal{E}_t(u)$ is differentiable on $[0,T]$ with derivative $\mathcal{P}_t(u) := \partial_t \mathcal{E}_t(u)$ satisfying for a constant $C_P \ge 0$

$$|\mathcal{P}_t(u)| \le C_P\big(\Psi(u) + \mathcal{E}_t(u)\big), \qquad \limsup_{w\to u, w\in D_E} \mathcal{P}_t(w) \le \mathcal{P}_t(u) \tag{E.2}$$

for every $(t,u) \in (0,T) \times D$, $E > 0$.

Ψ-uniform subdifferentiability: for every $E > 0$ there exists an upper semicontinuous map $\omega^E : [0,T] \times D_E \times D_E \to \mathbb{R}$, with $\omega^E_\cdot(u,u) \equiv 0$ for every $u \in D_E$, such that

$$\mathcal{E}_t(v) \ge \mathcal{E}_t(u) + \langle \xi, v - u\rangle - \omega^E_t(u,v)\Psi_\wedge(v-u) \qquad \forall\, t \in [0,T],\ u, v \in D_E, \quad \xi \in \partial\mathcal{E}_t(u), \tag{E.3}$$

where

$$\Psi_\wedge(w) := \min\Big(\Psi(w), \Psi(-w)\Big). \tag{2.7}$$

Recall that the Fréchet subdifferential of $\mathcal{E}_t$ is the possibly multivalued map $\partial\mathcal{E}_t : V \rightrightarrows V^*$ defined at $u \in D$ by

$$\xi \in \partial\mathcal{E}_t(u) \quad \Longleftrightarrow \quad \xi \in V^*, \quad \mathcal{E}_t(v) - \mathcal{E}_t(u) - \langle \xi, v-u \rangle \geq o(\|v-u\|) \text{ as } v \to u \text{ in } V, \tag{2.8}$$

Thus (E.3) prescribes a uniform and specific form for the remainder infinitesimal term on the right-hand side of (2.8). For later use, we observe that (E.2) and the Gronwall Lemma yield

$$0 \leq \Psi(u) + \mathcal{E}_s(u) \leq \mathcal{G}(u) \leq \exp(C_P T)\big(\Psi(u) + \mathcal{E}_t(u)\big) \qquad \text{for all } s, t \in [0,T],\ u \in D. \tag{2.9}$$

Since $\mathcal{E}$ is lower semicontinuous, (2.9) joint with (E.1) yields that the maps

$$u \mapsto \Psi(u) + \mathcal{E}_t(u) \quad \text{have compact sublevels in } V \quad \text{for every } t \in [0,T]. \tag{2.10}$$

Remark 2.1. Most of the results of the present paper could be extended to the cases when Ψ depends on the state of the system (as in [MRS13]), or it is replaced by a distance on D (as in [RMS08, MRS09]) and when the viscous correction Φ is a general convex superlinear functional (as in [MRS12a]). We have chosen the current simpler structure to focus on the main features and techniques of the vanishing-viscosity method in the infinite-dimensional setting.

2.2. **Absolutely continuous and** BV **functions.** As in Section 2.1 let $\Psi : V \to [0,\infty)$ be a gauge function with $\Psi(v) > 0$ if $v \neq 0$ and let Z a subset of V. The function

$$Z \ni u, v \mapsto \Delta_\Psi(u,v) := \Psi(v-u) \quad \text{is an asymmetric continuous distance on } Z. \tag{2.11}$$

We say that a curve $u : [0,T] \to Z$ is Ψ-absolutely continuous if there exists a nonnegative function $m \in L^1(0,T)$ such that

$$\Delta_\Psi(u(t_0), u(t_1)) \leq \int_{t_0}^{t_1} m(s)\,\mathrm{d}s \quad \text{for every } 0 \leq t_0 < t_1 \leq T. \tag{2.12}$$

We denote by $\mathrm{AC}([0,T]; Z, \Psi)$ the set of all Ψ-absolutely continuous curves with values in Z. There is a minimal function m such that (2.12) holds [AGS08, RMS08], and with a slight abuse of notation we denote it by $\Psi[u']$, since it admits the expression

$$\Psi[u'](t) = \lim_{h \to 0} \Psi\Big(\frac{u(t+h) - u(t)}{h}\Big) \quad \text{for } \mathscr{L}^1\text{-a.a. } t \in (0,T), \tag{2.13}$$

so that $\Psi[u'](t) = \Psi(\dot u(t))$ whenever u is differentiable at t. Since V has the Radon-Nikodým property, this happens at $\mathscr{L}^1$-a.a. $t \in (0,T)$ ($\mathscr{L}^1$ denoting the Lebesgue measure on $(0,T)$), when Ψ is coercive: if this is the case and $Z = V$, we will simply write $u \in \mathrm{AC}(0,T;V)$.

$\mathrm{Var}_\Psi(u;[a,b])$ is the *pointwise* total variation induced by Ψ on the interval $[a,b] \subset [0,T]$, viz.

$$\mathrm{Var}_\Psi(u;[a,b]) := \sup\Big\{ \sum_{m=1}^{M} \Psi\big(u(t_m) - u(t_{m-1})\big) : a = t_0 < t_1 < \cdots < t_{M-1} < t_M = b \Big\}. \tag{2.14}$$

If $Z \subset V$, $\mathrm{BV}([0,T]; Z, \Psi)$ will denote the set of all curves $u : [0,T] \to Z$ with finite Ψ-total variation in $[0,T]$. When $\Psi := \|\cdot\|$ we will simply write $\mathrm{BV}([0,T];V)$ and we will omit the index Ψ in the symbol of the total variation. Notice that $\mathrm{BV}([0,T];V) \subset \mathrm{BV}([0,T];V,\Psi)$ for every choice of Ψ, whereas the opposite inclusion only holds when Ψ is coercive on V.

To every $u \in \mathrm{BV}([0,T];Z,\Psi)$ we can associate the nondecreasing scalar function $\mathrm{V} : \mathbb{R} \to [0,\infty)$

$$\mathrm{V}(t) := \begin{cases} 0 & \text{if } t \leq 0, \\ \mathrm{Var}_\Psi(u;[0,t]) & \text{if } t \in (0,T), \\ \mathrm{Var}_\Psi(u;[0,T]) & \text{if } t \geq T \end{cases} \quad \text{with distributional derivative} \quad \mu = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{V}. \tag{2.15}$$

The finite Borel measure μ is supported in $[0,T]$ and it can be decomposed into the sum $\mu = \mu_{\mathrm{d}} + \mu_{\mathrm{J}}$ of a diffuse part μ_{d} (such that $\mu_{\mathrm{d}}(\{t\}) = 0$ for every $t \in \mathbb{R}$), and a jump part μ_{J} concentrated in a countable set $\mathrm{J}_u \subset [0,T]$.

When Z is compact (or when Ψ is coercive), for every $\delta > 0$ there exists a constant $M_\delta > 0$ such that (recall (2.7) for the definition of $\Psi_\wedge$)

$$\|u - v\| \le \delta + M_\delta\, \Psi_\wedge(v - u) \quad \text{for every } u, v \in Z. \tag{2.16}$$

By introducing the continuous and concave modulus of continuity

$$\Omega_Z : [0, +\infty) \to [0, +\infty), \quad \Omega_Z(r) := \inf_{\delta > 0} \delta + M_\delta\, r \quad \text{so that } \lim_{r \downarrow 0} \Omega_Z(r) = 0, \tag{2.17}$$

(2.16) rewrites as

$$\|u - v\| \le \Omega_Z(\Psi_\wedge(u - v)) \quad \text{for every } u, v \in Z. \tag{2.18}$$

If (2.16) holds, it is easy to show that a function $u \in \mathrm{BV}([0, T]; Z, \Psi)$ is continuous in $[0, T] \setminus \mathrm{J}_u$ and its left and right limits exist at every $t \in [0, T]$:

$$u(t_-) := \lim_{s \uparrow t} u(s), \;\; u(t_+) := \lim_{s \downarrow t} u(s) \;\; \text{with the convention } u(0_-) := u(0), \; u(T_+) := u(T), \tag{2.19}$$

so that J_u admits the representation

$$\mathrm{J}_u := \big\{ t \in [0, T] : u(t_-) \ne u(t) \text{ or } u(t) \ne u(t_+) \big\} \tag{2.20}$$

and

$$\mu_{\mathrm{J}}(\{t\}) = \Psi(u(t) - u(t_-)) + \Psi(u(t_+) - u(t)) \quad \text{for every } t \in \mathrm{J}_u. \tag{2.21}$$

Furthermore, μ_{d} admits the Lebesgue decomposition $\mu_{\mathrm{d}} = \mu_{\mathscr{L}} + \mu_{\mathrm{C}}$ with $\mu_{\mathscr{L}} \ll \mathscr{L}^1$ and $\mu_{\mathrm{C}} \perp \mathscr{L}^1$. The density of $\mu_{\mathscr{L}}$ with respect to $\mathscr{L}^1$ is provided by the same formula (2.13) and one has

$$u \in \mathrm{AC}([0, T]; Z, \Psi) \text{ if and only if } \mu_{\mathrm{J}} = \mu_{\mathrm{C}} \equiv 0, \text{ with } \quad \mathrm{Var}_\Psi(u; [a, b]) = \int_a^b \Psi[u'](t)\, \mathrm{d}t. \tag{2.22}$$

In this case, when Z is compact or Ψ coercive, u is a continuous curve. In general we have

$$\mathrm{Var}_\Psi(u; [a, b]) = \mu_{\mathrm{d}}([a, b]) + \mathrm{Jmp}_\Psi(u; [a, b]), \tag{2.23}$$

where the jump contribution $\mathrm{Jmp}_\Psi(u; [a, b])$ can be described by

$$\begin{aligned} \mathrm{Jmp}_\Psi(u; [a, b]) &:= \Delta_\Psi(u(a), u(a_+)) + \Delta_\Psi(u(b_-), u(b)) \\ &\quad + \sum_{t \in \mathrm{J}_u \cap (a, b)} \big(\Delta_\Psi(u(t_-), u(t)) + \Delta_\Psi(u(t), u(t_+)) \big), \\ &= \Delta_\Psi(u(a), u(a_+)) + \Delta_\Psi(u(b_-), u(b)) + \mu_{\mathrm{J}}((a, b)). \end{aligned} \tag{2.24}$$

Remark 2.2 (Scalar vs. vector measures)**.** If $u \in \mathrm{BV}(0, T; V)$ all the previous definitions have an important vector counterpart in terms of the vector measure $u'_{\mathscr{D}}$ associated with the distributional derivative of u: $u'_{\mathscr{D}}$ is a Radon vector measure on $(0, T)$ with values in V, with finite total variation $\|u'_{\mathscr{D}}\|$. The measure $u'_{\mathscr{D}}$ can be decomposed into the sum of the three mutually singular measures

$$u'_{\mathscr{D}} = u'_{\mathscr{L}} + u'_{\mathrm{C}} + u'_{\mathrm{J}}, \qquad u'_{\mathrm{d}} := u'_{\mathscr{L}} + u'_{\mathrm{C}}, \tag{2.25}$$

where $u'_{\mathscr{L}}$ is its absolutely continuous part with respect to $\mathscr{L}^1$. u'_{J} is a discrete measure concentrated on J_u, and u'_{C} is the so-called Cantor part, still satisfying $u'_{\mathrm{C}}(\{t\}) = 0$ for every $t \in [0, T]$. Therefore $u'_{\mathrm{d}} = u'_{\mathscr{L}} + u'_{\mathrm{C}}$ is the diffuse part of the measure, which does not charge J_u.

Since V has the Radon-Nikodým property, u is differentiable $\mathscr{L}^1$-a.e. in $(0, T)$ (we denote by $\dot u$ its derivative), and we can express u'_{d} in terms of its density $\boldsymbol{n}$ with respect to its total variation $\|u'_{\mathrm{d}}\|$ as

$$u'_{\mathrm{d}} = \boldsymbol{n} \|u'_{\mathrm{d}}\| \quad \text{where } \|\boldsymbol{n}\| = 1 \; \|u'_{\mathrm{d}}\|\text{-a.e.}, \quad u'_{\mathscr{L}} = \dot u\, \mathscr{L}^1, \quad \boldsymbol{n} = \frac{\dot u}{\|\dot u\|} \quad \|u'_{\mathscr{L}}\|\text{-a.e..} \tag{2.26}$$

The relation to the previously introduced measures μ_{d}, μ_{C}, and $\mu_{\mathscr{L}}$ is

$$\mu_{\mathrm{d}} = \Psi(\boldsymbol{n}) \|u'_{\mathrm{d}}\|, \quad \mu_{\mathrm{C}} = \Psi(\boldsymbol{n}) \|u'_{\mathrm{C}}\|, \quad \mu_{\mathscr{L}} = \Psi(\boldsymbol{n}) \|u'_{\mathscr{L}}\| = \Psi(\dot u) \mathscr{L}^1. \tag{2.27}$$

2.3. **Two useful properties from the theory of gradient systems.** The assumptions on the dissipation potentials Ψ and Φ and on the energy $\mathcal{E}$ stated in the previous section yield two important consequences, stated in Theorem 2.3 below, that play a crucial role in the variational approach to gradient systems and rate-independent evolutions.

Before stating them, let us recall that for every map $\Lambda : V \to (-\infty, +\infty]$ bounded from below by a continuous and affine function, $\Lambda^* : V^* \to (-\infty, +\infty]$ will denote the conjugate

$$\Lambda^*(\xi) := \sup_{v\in V} \langle \xi, v\rangle - \Lambda(v). \tag{2.28}$$

For the functional Φ in (D.2) we have

$$\Phi^*(\xi) = F^*\big(\|\xi\|_*\big), \quad \text{where} \quad F^*(s) = \sup_{r\geq 0} rs - F(r), \tag{2.29}$$

so that, by the inf-convolution duality formula (see e.g. [IoT79, Thm. 1, p. 178]) and the monotonicity of F^* we find

$$\Psi_\varepsilon^*(\xi) = \frac{1}{\varepsilon} \min_{z\in K^*} \Phi^*(\xi - z) = \frac{1}{\varepsilon} \min_{z\in K^*} F^*\big(\|\xi - z\|_*\big) = \frac{1}{\varepsilon} F^*\Big(\min_{z\in K^*} \|\xi - z\|_*\Big). \tag{2.30}$$

Theorem 2.3 ([MRS13, Prop. 2.4])**.** *Under the assumptions of Section 2.1 the following properties hold.*

Chain rule**:** *For every* $u \in \mathrm{AC}([0,T];V)$ *and* $\xi \in L^1(0,T;V^*)$ *with*

$$\begin{gathered} \sup_{t\in[0,T]} |\mathcal{E}_t(u(t))| < +\infty, \quad \xi(t) \in -\partial\mathcal{E}_t(u(t)) \text{ for a.a. } t \in (0,T), \text{ and} \\ \int_0^T \Psi_\varepsilon(\dot u(t))\,\mathrm{d}t < +\infty, \quad \int_0^T \Psi_\varepsilon^*(\xi(t))\,\mathrm{d}t < +\infty, \end{gathered} \tag{2.31}$$

the map $t \mapsto \mathcal{E}_t(u(t))$ *is absolutely continuous and*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) = -\langle \xi(t), u'(t)\rangle + \mathcal{P}_t(u(t)) \quad \text{for a.a.}\, t \in (0,T)\,. \tag{2.32}$$

Strong-Weak closedness of the graph of $(\mathcal{E}, \partial\mathcal{E})$**:** *For all sequences* $(t_n) \subset [0,T]$, $(u_n) \subset V$ *and* $(\xi_n) \subset V^*$ *we have the following condition:*

$$\begin{aligned} &\text{if } t_n \to t \text{ in } [0,T], \ \ u_n \to u \text{ in } V, \ \ \xi_n \rightharpoonup \xi \text{ in } V^*, \ \xi_n \in \partial\mathcal{E}_{t_n}(u_n), \\ &\text{and if } \mathcal{E}_{t_n}(u_n) \to \mathscr{E} \text{ in } \mathbb{R}, \qquad \text{then} \quad \xi \in \partial\mathcal{E}_t(u) \ \text{ and } \ \mathscr{E} = \mathcal{E}_t(u). \end{aligned} \tag{2.33}$$

Furthermore, (2.33) *implies that* $\partial\mathcal{E}_t(u)$ *is a weakly***-closed, convex subset (possibly empty) of* V^*.

2.4. **Variational gradient systems.** We recall an application of the general existence and approximation result of [MRS13] for the Cauchy problem associated with (1.1).

Theorem 2.4 ([MRS13])**.** *Let us assume that* (D.0)–(D.2) *and* (E.0)–(E.3) *hold. Then, for every* $u_{0,\varepsilon} \in D$ *there exists a curve* $u_\varepsilon \in \mathrm{AC}([0,T];V)$ *solving* (1.1) *and fulfilling the Cauchy condition* $u(0) = u_{0,\varepsilon}$*. More precisely, there exists a function* $\xi_\varepsilon \in L^1(0,T;V^*)$ *fulfilling*

$$\xi_\varepsilon(t) \in -\partial\mathcal{E}_t(u_\varepsilon(t)), \quad \xi_\varepsilon(t) \in \partial\Psi_\varepsilon(\dot u_\varepsilon(t)) \quad \text{for a.a. } t \in (0,T), \tag{2.34}$$

and the energy identity *for all* $0 \leq s \leq t \leq T$

$$\int_s^t \Big(\Psi_\varepsilon(\dot{u_\varepsilon}(r)) + \Psi_\varepsilon^*(\xi_\varepsilon(r))\Big)\,\mathrm{d}r + \mathcal{E}_t(u_\varepsilon(t)) = \mathcal{E}_s(u_\varepsilon(s)) + \int_s^t \mathcal{P}_r(u_\varepsilon(r))\,\mathrm{d}r. \tag{2.35}$$

Minimizing Movement solutions. Theorem 2.4 was proved in [MRS13, Thm. 4.4] by passing to the limit in the time-discretization scheme $(\mathrm{IP}_{\varepsilon,\tau})$, see the last paragraph of the introduction. Here we quote the main convergence result:

Theorem 2.5 (Minimizing Movement solutions to (1.1))**.** *Under our standard assumptions* (D.0)–(D.2) *and* (E.0)–(E.3)*, Problem* $(\mathrm{IP}_{\varepsilon,\tau})$ *has at least a solution* $(\mathrm{U}^n_{\tau,\varepsilon})_{n=0}^{N_\tau}$*. For every* $\varepsilon > 0$ *there exist a sequence* $\tau_k \downarrow 0$ *as* $k \to \infty$ *and a limit solution* $u_\varepsilon \in \mathrm{AC}([0,T];V)$ *to* (2.34) *and* (2.35) *such that the piecewise affine interpolants* U_τ *satisfy*

$$\mathrm{U}_{\tau_k,\varepsilon} \to u_\varepsilon \quad \textit{in } V, \textit{ uniformly in } [0,T]. \tag{2.36}$$

Since solutions obtained as such limits have special properties, we will call them *Minimizing Movement solutions* according to [De 93] (see also [AGS08]).

3. Balanced Viscosity (BV) solutions

Throughout this section we will keep to the notation and assumptions of Section 2.1, in particular we will suppose that Ψ, Φ fulfill (D.0)–(D.2) and that $\mathcal{E}$ complies with (E.0)–(E.3).

After a discussion of the main concepts of *contact potential* and *Finsler dissipation cost* in §3.1, we will introduce the notion of Balanced Viscosity (BV) solutions in §3.2 and we will present the main results related to this crucial concept. The distinguished subclass of V-parameterizable solutions will be considered in the last part § 3.4.

3.1. **Finsler dissipation functionals.**

As in [MRS12a], the *vanishing-viscosity contact potential* $\mathfrak{p} : V \times V^* \to [0,+\infty)$ induced by the dissipation potentials Ψ_ε is

$$\mathfrak{p}(v,\xi) := \inf_{\varepsilon>0} \Big(\Psi_\varepsilon(v) + \Psi_\varepsilon^*(\xi) \Big), \qquad v \in V,\ \xi \in V^*. \tag{3.1}$$

The representation formula (2.30) for Ψ_ε^* and the fact that

$$\inf_{\varepsilon>0} \varepsilon^{-1}\Big(F(\varepsilon r) + F^*(s) \Big) = rs \quad \text{for every } r, s \geq 0,$$

yield the useful splitting of $\mathfrak{p}$:

$$\mathfrak{p}(v,\xi) = \Psi(v) + \|v\| \min_{z \in K^*} \|\xi - z\|_*. \tag{3.2}$$

Remark 3.1 (More general viscous dissipations and contact potentials)**.** The particular form (D.2) of Φ allows for the simple representation (3.2) of $\mathfrak{p}$, which is useful to understand the role played by the two different viscosities. The general case concerning arbitrary convex superlinear functions Φ has been analyzed in [MRS12a] and almost all the crucial properties can also be adapted to the present infinite-dimensional setting. Here we just mention that every contact potential is convex and degree-1 homogeneous with respect to its first variable and it fulfills the Fenchel inequality

$$\mathfrak{p}(v,\xi) \geq \langle \xi, v\rangle, \quad \text{and} \quad \begin{cases} \mathfrak{p}(v,\xi) \geq \Psi(v) & \text{for all } (v,\xi) \in V \times V^*, \\ \mathfrak{p}(v,\xi) = \Psi(v) & \text{if and only if } \xi \in K^*. \end{cases} \tag{3.3}$$

Next, we associate with $\mathfrak{p}$ and with the Fréchet subdifferential $\partial\mathcal{E}$ the time-dependent family of Finsler dissipation functionals

$$\mathfrak{f} : [0,T] \times D \times V \to [0,+\infty], \quad \mathfrak{f}_t(u;v) := \inf \Big\{ \mathfrak{p}(v,\xi) : \xi \in -\partial\mathcal{E}_t(u) \Big\}, \tag{3.4}$$

where we adopt the standard convention $\inf \emptyset = +\infty$. Notice that when $\partial\mathcal{E}_t(u) \neq \emptyset$ the inf in formula (3.4) is attained; moreover, the functional $v \mapsto \mathfrak{f}_t(u;v)$ is lower semicontinuous, convex, and positively 1-homogeneous.

In accord with (3.2) it will also be useful to split $\mathfrak{f}_t(u;v)$ into the sum of the dissipation $\Psi(v)$ (independent of u) and of the correction term induced by the viscous norm $\|\cdot\|$ and $\partial\mathcal{E}$, viz.

$$\mathfrak{f}_t(u;v) = \Psi(v) + \mathfrak{e}_t(u)\|v\|, \quad \mathfrak{e}_t(u) := \inf \Big\{ \|\xi - z\|_* : \xi \in -\partial\mathcal{E}_t(u),\ z \in K^* \Big\}. \tag{3.5}$$

By (2.33), for every $E > 0$ the function $\mathfrak{e} : [0,T] \times D \to [0,\infty]$ satisfies the crucial properties

$$\mathfrak{e} \text{ is l.s.c. in } [0,T] \times D_E \quad \text{and} \quad \mathfrak{e}_t(u) = 0 \iff K^* + \partial \mathcal{E}_t(u) \ni 0, \tag{3.6}$$

where D_E denotes the E-sublevel of the energy, cf. (E.1).

If Ψ were coercive on V, then the Finsler cost associated to $\mathfrak{f}_t$ could be simply defined as

$$\Delta_{\mathfrak{f}_t}(u_0,u_1) := \inf\Big\{ \int_{r_0}^{r_1} \mathfrak{f}_t(\vartheta(r);\dot\vartheta(r))\,\mathrm{d}r : \vartheta \in \mathrm{AC}([r_0,r_1];V),\ \vartheta(r_i) = u_i,\ i = 0,1\Big\}, \tag{3.7}$$

and it would be possible to show that the infimum in (3.7) is attained whenever the cost is finite. Notice that, since $\mathfrak{f}_t(u;\cdot)$ is positively 1-homogeneous, the choice of the interval $[r_0,r_1]$ in (3.7) is irrelevant and one can also assume that the competing curves ϑ belong to $\mathrm{Lip}([r_0,r_1];V)$.

On the other hand, since Ψ is not coercive in general, the definition (3.7) has to be conveniently adapted to cover the case of curves ϑ that may lack differentiability at every time. The next definition focuses on this aspect (see §2.2 for BV and AC curves with respect to Ψ).

Definition 3.2 (Admissible curves)**.** *A curve $\vartheta : [r_0,r_1] \to V$ is called* admissible *if it belongs to $\mathrm{AC}([r_0,r_1];D_E,\Psi)$ for some $E > 0$, and if its restriction to the (relatively) open set*

$$G_t = G_t[\vartheta] := \big\{ r \in [r_0,r_1] : \mathfrak{e}_t(\vartheta(r)) > 0 \big\}, \tag{3.8}$$

belongs to $\mathrm{AC}_{\mathrm{loc}}(G_t[\vartheta];V)$. We call $\mathcal{T}_t(u_0,u_1)$ the class of all admissible transition curves *$\vartheta : [0,1] \to V$ such that $\vartheta(i) = u_i$, $i = 0,1$, and we set*

$$\mathfrak{f}_t[\vartheta;\vartheta'](r) := \begin{cases} \mathfrak{f}_t(\vartheta(r);\dot\vartheta(r)) = \Psi(\dot\vartheta(r)) + \mathfrak{e}_t(\vartheta(r))\|\dot\vartheta(r)\| & \text{if } r \in G_t[\vartheta], \\ \Psi[\vartheta'](r) & \text{if } r \in [0,1] \setminus G_t[\vartheta]. \end{cases} \tag{3.9}$$

Remark 3.3. Let us add a few comments on the previous definition. First of all, as we discussed in Section 2.2, we notice that the continuity of ϑ follows from the compactness of D_E in V and the fact that Ψ is continuous and nondegenerate, so that $\Psi(v) = 0 \ \Rightarrow \ v = 0$.

Once ϑ is continuous, the l.s.c. property of $\mathfrak{e}$ stated in (3.6) implies that the set $G_t[\vartheta]$ defined in (3.8) is open. Since V has the Radon-Nikodým property, ϑ is differentiable $\mathscr{L}^1$-a.e. in $G_t[\vartheta]$. It is immediate to see that for every admissibile curve ϑ

$$\int_0^1 \mathfrak{f}_t[\vartheta;\vartheta'](r)\,\mathrm{d}r = \int_0^1 \Psi[\vartheta'](r)\,\mathrm{d}r + \int_{G_t[\vartheta]} \mathfrak{e}_t(\vartheta(r))\|\dot\vartheta(r)\|\,\mathrm{d}r. \tag{3.10}$$

We are now in the position to extend the definition (3.7) of $\Delta_{\mathfrak{f}}$.

Definition 3.4 (Finsler dissipation cost)**.** *Let $t \in [0,T]$ be fixed and let us consider $u_0, u_1 \in D$. The (possibly asymmetric) Finsler cost induced by $\mathfrak{f}$ at the time t is given by*

$$\Delta_{\mathfrak{f}_t}(u_0,u_1) := \inf_{\vartheta \in \mathcal{T}_t(u_0,u_1)} \int_0^1 \mathfrak{f}_t[\vartheta,\vartheta'](r)\,\mathrm{d}r \tag{3.11}$$

$$= \inf_{\vartheta \in \mathcal{T}_t(u_0,u_1)} \int_0^1 \Psi[\vartheta'](r)\,\mathrm{d}r + \int_{G_t[\vartheta]} \mathfrak{e}_t(\vartheta(r))\|\dot\vartheta(r)\|\,\mathrm{d}r, \tag{3.12}$$

with the usual convention of setting $\Delta_{\mathfrak{f}_t}(u_0,u_1) = +\infty$ if $\mathcal{T}_t(u_0,u_1)$ is empty.

Let us notice that in general $\Delta_{\mathfrak{f}_t}(\cdot,\cdot)$ is not symmetric, unless Ψ is symmetric, and that

$$\Delta_{\mathfrak{f}_t}(u_0,u_1) \ge \Delta_\Psi(u_0,u_1) \quad \text{for every } u_0,u_1 \in D,\ t \in [0,T]. \tag{3.13}$$

This follows from the fact that in (3.12) we have

$$\int_0^1 \Psi[\vartheta'](r)\,\mathrm{d}r = \mathrm{Var}_\Psi(\vartheta;[0,1]) \ge \Psi(u_1 - u_0) = \Delta_\Psi(u_0,u_1).$$

In the next important result we collect a few crucial properties of the Finsler dissipation cost, namely the existence of optimal transition paths and the lower semicontinuity properties needed in what follows. Theorem 3.5 will be proved in Section 7.2.

Theorem 3.5. *Let* (D.0)–(D.2) *and* (E.0)–(E.3) *hold. Let* $t \in [0,T]$, $E > 0$ *and* $u_-, u_+ \in D_E$.

(F1) *If* $\Delta_{\mathfrak{f}_t}(u_-, u_+) < \infty$ *there exists a transition path* $\vartheta \in \mathcal{T}_t(u_-, u_+)$ *attaining the infimum in* (3.12). *Moreover*

$$\Delta_{\mathfrak{f}_t}(u_-, u_+) \geq \Big|\mathcal{E}_t(u_-) - \mathcal{E}_t(u_+)\Big|. \tag{3.14}$$

(F2) *If* $u_{0,n}, u_{1,n} \in D_E$, $n \in \mathbb{N}$, *then*

$$\lim_{n\to\infty} u_{0,n} = u_-, \quad \lim_{n\to\infty} u_{1,n} = u_+ \quad \Longrightarrow \quad \liminf_{n\to\infty} \Delta_{\mathfrak{f}_t}(u_{0,n}, u_{1,n}) \geq \Delta_{\mathfrak{f}_t}(u_-, u_+). \tag{3.15}$$

(F3) *If* $u_n \in \mathrm{AC}([\alpha_n, \beta_n]; V)$, $\tilde u_n : [\alpha_n, \beta_n] \to D_E$ *measurable,* $\xi_n \in L^1(\alpha_n, \beta_n; V^*)$, $\varepsilon_n > 0$, $n \in \mathbb{N}$, *are sequences satisfying*

$$\lim_{n\to\infty} \sup_{r\in[\alpha_n,\beta_n]} \|\tilde u_n(r) - u_n(r)\| = 0, \quad \xi_n(r) \in -\partial\mathcal{E}_r(\tilde u_n(r)) \quad \textit{for a.a. } r \in (\alpha_n, \beta_n), \tag{3.16}$$

$$\lim_{n\to\infty} u_n(\alpha_n) = u_-, \ \lim_{n\to\infty} u_n(\beta_n) = u_+, \quad \lim_{n\to\infty} \alpha_n = \lim_{n\to\infty} \beta_n = t, \tag{3.17}$$

and

$$\lim_{n\to\infty} \varepsilon_n = 0, \quad \Delta := \lim_{n\to\infty} \int_{\alpha_n}^{\beta_n} \Big(\Psi_{\varepsilon_n}(\dot u_n) + \Psi^*_{\varepsilon_n}(\xi_n)\Big)\,\mathrm{d}r < \infty, \tag{3.18}$$

then there exist an increasing subsequence $(n_k)_k \subset \mathbb{N}$, *increasing and surjective time rescalings* $\mathsf{t}_{n_k} \in \mathrm{AC}([0,1]; [\alpha_{n_k}, \beta_{n_k}])$, *and an admissible transition* $\vartheta \in \mathcal{T}_t(u_-, u_+)$ *such that*

$$\lim_{k\to\infty} u_{\varepsilon_{n_k}} \circ \mathsf{t}_{n_k} = \vartheta \quad \textit{strongly in } V\textit{, uniformly on } [0,1], \qquad \int_0^1 \mathfrak{f}_t[\vartheta, \vartheta'](r)\,\mathrm{d}r \leq \Delta. \tag{3.19}$$

In particular, whenever (3.16) *and* (3.17) *hold, along any sequence* $\varepsilon_n \downarrow 0$ *we have*

$$\liminf_{n\to\infty} \int_{\alpha_n}^{\beta_n} \Big(\Psi_{\varepsilon_n}(\dot u_n) + \Psi^*_{\varepsilon_n}(\xi_n)\Big)\,\mathrm{d}r \geq \Delta_{\mathfrak{f}_t}(u_-, u_+). \tag{3.20}$$

Solutions to (1.1), with $\tilde u_n = u_n$, provide a particularly important example of sequences in assertion (F3) of Thm. 3.5. Notice that by (3.14) the Finsler cost controls the amount of energy dissipation between two arbitrary points at a fixed time t. On the other hand, (3.20) shows that $\Delta_{\mathfrak{f}}$ captures the concentration of the asymptotic energy dissipation of a family of solutions to the viscous gradient flow (2.34).

We now use the Finsler cost $\Delta_{\mathfrak{f}}$ to characterize the minimal dissipated energy along *any* curve $u \in \mathrm{BV}_\Psi([0,T]; V)$, by means of a suitable notion of total variation, which involves $\Delta_{\mathfrak{f}}$ to measure the contributions due to the jumps of u (recall (2.23) and (2.24)).

Definition 3.6 (Jump and total variation induced by $\mathfrak{f}$)**.** *Let* $E > 0$ *and* $u \in \mathrm{BV}([0,T]; D_E, \Psi)$ *be a given curve with jump set* J_u. *For every subinterval* $[a,b] \subset [0,T]$ *the* jump variation *of* u *induced by* $\mathfrak{f}$ *on* $[a,b]$ *is*

$$\begin{aligned}\mathrm{Jmp}_{\mathfrak{f}}(u; [a,b]) := {} & \Delta_{\mathfrak{f}_a}(u(a), u(a_+)) + \Delta_{\mathfrak{f}_b}(u(b_-), u(b)) \\ & + \sum_{t \in \mathrm{J}_u \cap (a,b)} \big(\Delta_{\mathfrak{f}_t}(u(t_-), u(t)) + \Delta_{\mathfrak{f}_t}(u(t), u(t_+))\big).\end{aligned} \tag{3.21}$$

The $\mathfrak{f}$*-total variation induced of* u *on* $[a,b]$ *for* $a < b$ *is*

$$\begin{aligned}\mathit{Var}_{\mathfrak{f}}(u; [a,b]) &:= \mathrm{Var}_\Psi(u; [a,b]) - \mathrm{Jmp}_\Psi(u; [a,b]) + \mathrm{Jmp}_{\mathfrak{f}}(u; [a,b]) && (3.22) \\ &= \mu_{\mathrm{d}}(a,b) + \mathrm{Jmp}_{\mathfrak{f}}(u; [a,b]). && (3.23)\end{aligned}$$

Remark 3.7. As already pointed out in [MRS12a, Rmk. 3.5], $\mathit{Var}_{\mathfrak{f}}$ is not a *standard* total variation functional: for instance, it is not induced by any distance on V, and it is not lower semicontinuous with respect to pointwise convergence in V, unless a further local stability constraint is imposed.

Nevertheless, $\mathit{Var}_{\mathfrak{f}}$ enjoys the nice additivity property

$$\mathit{Var}_{\mathfrak{f}}(u; [a,b]) + \mathit{Var}_{\mathfrak{f}}(u; [b,c]) = \mathit{Var}_{\mathfrak{f}}(u; [a,c]) \quad \text{whenever} \quad 0 \leq a < b < c \leq T. \tag{3.24}$$

3.2. **Balanced Viscosity (BV) solutions.** Based on Definition 3.6, we can now specify the concept of *Balanced Viscosity* (BV) *solution* to the rate-independent system generated by $(V, \mathcal{E}, \Psi, \Phi)$: the global stability condition in the definition of *energetic solutions* is replaced by the *local* stability condition $(\mathrm{S_{loc}})$, and the energy balance features the total variation functional $\mathit{Var}_{\mathfrak{f}}$. As usual, we will always assume that Ψ, Φ fulfill (D.0)–(D.2) and that $\mathcal{E}$ complies with (E.0)–(E.3).

Definition 3.8 (BV solutions)**.** *A curve $u \in \mathrm{BV}([0,T]; D, \Psi)$ is a* BV solution of the rate-independent system $(V, \mathcal{E}, \Psi, \Phi)$ *if the* local stability $(\mathrm{S_{loc}})$ *and the* $(\mathrm{E_{\mathfrak{f}}})$-energy balance *hold:*

$$K^* + \partial\mathcal{E}_t(u(t)) \ni 0 \quad \text{for all} \quad t \in [0,T] \setminus \mathrm{J}_u, \tag{$\mathrm{S_{loc}}$}$$

$$\mathit{Var}_{\mathfrak{f}}(u; [0,t]) + \mathcal{E}_t(u(t)) = \mathcal{E}_0(u(0)) + \int_0^t \mathcal{P}_s(u(s))\,\mathrm{d}s \qquad \text{for all } t \in (0,T]. \tag{$\mathrm{E_{\mathfrak{f}}}$}$$

Every BV solution u to the RIS $(V, \mathcal{E}, \Psi, \Phi)$ satisfies the energy balance in each subinterval

$$\mathit{Var}_{\mathfrak{f}}(u; [s,t]) + \mathcal{E}_t(u(t)) = \mathcal{E}_s(u(s)) + \int_s^t \mathcal{P}_r(u(r))\,\mathrm{d}r \quad \text{for every } 0 \le s < t \le T, \tag{3.25}$$

thanks to $(\mathrm{E_{\mathfrak{f}}})$ and the additivity (3.24) of the total variation functional $\mathit{Var}_{\mathfrak{f}}$.

Before studying other properties and characterizations of balanced viscosity solutions, let us first present our main existence and convergence results.

Main existence and convergence results. Our first result states the convergence in the vanishing-viscosity limit $\varepsilon \downarrow 0$ of solutions to (1.1) to a BV solution of the rate-independent system $(V, \mathcal{E}, \Psi, \Phi)$. As a byproduct, we can prove in this way the existence of BV solutions. Let us emphasize that Definition 3.8 of BV solutions is only inspired by the vanishing-viscosity approach but otherwise completely independent of it. We postpone the proofs to Section 7.3.

The reader should be aware that, here and in what follows, we will call a sequence $(\varepsilon_k)_k$ converging to 0 simply a *vanishing sequence*.

Theorem 3.9 (Existence of BV solutions and convergence of viscous approximations)**.** *If* (D.0)–(D.2) *and* (E.0)–(E.3) *hold, then for every $u_0 \in D$ there exists a* BV *solution u of the RIS $(V, \mathcal{E}, \Psi, \Phi)$.*

Moreover for every family $(u_\varepsilon, \xi_\varepsilon)_\varepsilon \subset \mathrm{AC}([0,T]; V) \times L^1(0,T; V^*)$ of solutions of the doubly nonlinear equation (2.34) *with*

$$u_\varepsilon(0) \to u_0 \quad \text{in } V \text{ and} \quad \mathcal{E}_0(u_\varepsilon(0)) \to \mathcal{E}_0(u_0) \quad \text{as } \varepsilon \downarrow 0 \tag{3.26}$$

and for every vanishing sequence $(\varepsilon_k)_k$ there exist $E > 0$, a further (not relabeled) subsequence, and a limit function $u \in \mathrm{BV}([0,T]; D_E, \Psi)$ such that as $k \to \infty$

$$u_{\varepsilon_k}(t) \to u(t) \quad \text{in } V \text{ for all } t \in [0,T], \tag{3.27}$$

$$\lim_{k\to\infty} \mathcal{E}_t(u_{\varepsilon_k}(t)) = \mathcal{E}_t(u(t)) \quad \text{for all } t \in [0,T], \tag{3.28}$$

$$\mathrm{Var}_{\mathfrak{f}}(u; [s,t]) = \lim_{k\to\infty} \mathrm{Var}_{\mathfrak{f}}(u_{\varepsilon_k}; [s,t]) = \lim_{k\to\infty} \int_s^t \Big(\Psi_{\varepsilon_k}(\dot u_{\varepsilon_k}(r)) + \Psi^*_{\varepsilon_k}(-\xi_{\varepsilon_k}(r))\Big)\,\mathrm{d}r \tag{3.29}$$

for all $0 \le s < t \le T$. Any pointwise limit function u obtained in this way is a BV *solution to the RIS $(V, \mathcal{E}, \Psi, \Phi)$.*

Let us emphasize that, in view of the above result, *every* limit point u of solutions $(u_\varepsilon)_\varepsilon$ of (2.34) such that (3.27)–(3.29) hold is a BV solution.

The next theorem concerns the convergence of the discrete solutions of the *viscous* time-incremental problem $(\mathrm{IP}_{\varepsilon,\tau})$, as *both* the viscosity parameter ε *and* the time-step τ tend to zero. Similar results for the finite-dimensional case were obtained in [MRS12a, Thm. 4.10].

Theorem 3.10 (Discrete-viscous approximations converge to BV solutions)**.** *Assume that* (D.0)–(D.2) *and* (E.0)–(E.3) *hold. Let* $u_0 \in D$ *be fixed, and let* $(\mathrm{U}_{\tau,\varepsilon})_{\tau,\varepsilon}$ *be a family of piecewise affine interpolants of discrete solutions* $(\mathrm{U}^n_{\tau,\varepsilon})_{n,\tau,\varepsilon}$ *to* $(\mathrm{IP}_{\varepsilon,\tau})$*, with*

$$\mathrm{U}^0_{\tau,\varepsilon} \to u_0 \quad \text{in } V \text{ and } \quad \mathcal{E}_0(\mathrm{U}^0_{\tau,\varepsilon}) \to \mathcal{E}_0(u_0) \quad \text{as } \tau, \varepsilon \downarrow 0. \tag{3.30}$$

Then for all sequences $(\tau_k, \varepsilon_k)_{k\in\mathbb{N}}$ *satisfying*

$$\lim_{k\to\infty} \varepsilon_k = \lim_{k\to\infty} \frac{\tau_k}{\varepsilon_k} = 0, \tag{3.31}$$

there exists $E > 0$*, a (not relabeled) subsequence and a curve* $u \in \mathrm{BV}([0,T]; D_E, \Psi)$ *such that*

$$\overline{\mathrm{U}}_{\tau_k,\varepsilon_k}(t) \to u(t) \quad \text{in } V \text{ for all } t \in [0,T], \tag{3.32}$$

$$\mathcal{E}_t(\overline{\mathrm{U}}_{\tau_k,\varepsilon_k}(t)) \to \mathcal{E}_t(u(t)) \quad \text{for all } t \in [0,T], \tag{3.33}$$

as $k \to \infty$*, and the limit* u *is a* BV *solution to the RIS* $(V, \mathcal{E}, \Psi, \Phi)$*.*

We now aim to shed more light onto the definition and the properties of BV solutions: first of all, we derive a characterization of BV solutions in terms of a one-sided version of the energy identity $(\mathrm{E}_\mathfrak{f})$, based on the *chain-rule* inequality stated in Theorem 3.11. A second characterization is given through a "metric" subdifferential inclusion and a set of jump conditions.

Chain-rule inequalities and characterizations of* BV *solutions. The next result is the infinite-dimensional analogue of [MRS09, Prop. 4] and is especially adapted to rate-independent systems. In particular, the fact that $\mathit{Var}_\mathfrak{f}$ is not a true total variation functional is here compensated by assuming that u fulfills the local stability condition $(\mathrm{S}_{\mathrm{loc}})$.

Theorem 3.11 (A chain-rule inequality for BV curves)**.** *If* $u \in \mathrm{BV}([0,T]; D_E, \Psi)$*,* $E > 0$*, satisfies the local stability condition* $(\mathrm{S}_{\mathrm{loc}})$ *and* $\mathrm{Var}_\mathfrak{f}(u; [0,T]) < \infty$*, then the map* $t \mapsto e(t) := \mathcal{E}_t(u(t))$ *belongs to* $\mathrm{BV}([0,T])$ *and satisfies the following chain-rule inequality:*

$$\Big| e(t_1) - e(t_0) - \int_{t_0}^{t_1} \mathcal{P}_t(u(t))\mathrm{d}t \Big| \le \mathrm{Var}_\mathfrak{f}(u; [t_0, t_1]) \quad \text{for all } 0 \le t_0 \le t_1 \le T. \tag{3.34}$$

If moreover $u \in \mathrm{BV}([0,T]; V)$ *and* $\xi : [0,T] \to K^*$ *is a Borel map such that* $\xi(t) \in -\partial\mathcal{E}_t(u(t))$ *for every* $t \in [0,T] \setminus \mathrm{J}_u$ *then the diffuse part* e'_d *of the distributional derivative* $e'_\mathscr{D}$ *of* e *can be represented as (recall* (2.26)*)*

$$e'_\mathrm{d} = -\langle \xi, \boldsymbol{n}\rangle \|u'_\mathrm{d}\| + \mathcal{P}_\cdot(u)\mathscr{L}^1 = -\langle \xi, \boldsymbol{n}\rangle \|u'_\mathrm{C}\| + \Big(-\langle \xi, \dot{u}\rangle + \mathcal{P}_\cdot(u)\Big)\mathscr{L}^1, \tag{3.35}$$

where $\boldsymbol{n}$ *is as in* (2.26)*, and* u'_d*,* u'_C *are from* (2.25)*.*

Indeed, (3.34) is the counterpart to the *parameterized* chain-rule inequality which shall be stated in Theorem 4.4 ahead. Both Theorems will be proved in Section 6.

As a direct consequence of Theorem 3.11 we have a *characterization* of BV solutions in terms of a single, global in time, energy-dissipation inequality.

Corollary 3.12 (A global energy-dissipation inequality characterizing BV solutions)**.** *A curve* $u \in \mathrm{BV}([0,T]; D_E, \Psi)$ *for some* $E > 0$ *is a* BV *solution to the RIS* $(V, \mathcal{E}, \Psi, \Phi)$ *if and only if it satisfies the local stability* $(\mathrm{S}_{\mathrm{loc}})$ *and the one-sided global in time version of* $(\mathrm{E}_\mathfrak{f})$*, viz.*

$$\mathrm{Var}_\mathfrak{f}(u; [0,T]) + \mathcal{E}_T(u(T)) \le \mathcal{E}_0(u(0)) + \int_0^T \mathcal{P}_s(u(s))\,\mathrm{d}s\,. \tag{$\mathrm{E}_{\mathfrak{f},\mathrm{ineq}}$}$$

Proof. In order to deduce the energy balance $(\mathrm{E}_\mathfrak{f})$ from $(\mathrm{E}_{\mathfrak{f},\mathrm{ineq}})$, we define $a(t) := \mathcal{E}_t(u(t)) - \int_0^t \mathcal{P}_s(u(s))\,\mathrm{d}s$ and $v(t) := \mathit{Var}_\mathfrak{f}(u; [0,t])$ such that $(\mathrm{E}_{\mathfrak{f},\mathrm{ineq}})$ takes the form $a(T) + v(T) \le a(0) + v(0)$, because $v(0) = 0$. The additivity (3.24) gives $\mathit{Var}_\mathfrak{f}(u; [s,t]) = v(t) - v(s)$, so that the chain-rule estimate (3.34) rephrases as $|a(t) - a(s)| \le v(t) - v(s)$ for all $0 \le s \le t \le T$. This implies the monotonicity $a(t) + v(t) \ge a(s) + v(s)$, and we conclude $a(t) + v(t) = a(0) + v(0)$ for all t, which is $(\mathrm{E}_\mathfrak{f})$. □

The importance of using the viscous total variation induced by $\mathfrak{f}$ (instead of the simpler one associated with Ψ) is clarified by the next result, characterizing the jump conditions.

Theorem 3.13 (Local stability, (Ψ)-energy dissipation and jump conditions)**.** *A curve* $u \in \mathrm{BV}([0,T];D_E,\Psi)$ *is a* BV *solution of the RIS* $(V,\mathcal{E},\Psi,\Phi)$ *if and only if it satisfies the local stability condition* $(\mathrm{S_{loc}})$*, the* (Ψ)*-energy dissipation inequality*

$$\mathrm{Var}_\Psi(u;[s,t]) + \mathcal{E}_t(u(t)) \le \mathcal{E}_s(u(s)) + \int_s^t \mathcal{P}_r(u(r))\,\mathrm{d}r \quad \textit{for every } 0 \le s < t \le T, \tag{$\mathrm{E}_{\Psi,\mathrm{ineq}}$}$$

and the following jump conditions at each point $t \in \mathrm{J}_u$ *of the jump set* (2.20)

$$\begin{aligned} \mathcal{E}_t(u(t)) - \mathcal{E}_t(u(t_-)) &= -\Delta_{\mathfrak{f}_t}(u(t_-),u(t)),\\ \mathcal{E}_t(u(t_+)) - \mathcal{E}_t(u(t)) &= -\Delta_{\mathfrak{f}_t}(u(t),u(t_+)),\\ \mathcal{E}_t(u(t_+)) - \mathcal{E}_t(u(t_-)) &= -\Delta_{\mathfrak{f}_t}(u(t_-),u(t_+)) = -\Big(\Delta_{\mathfrak{f}_t}(u(t_-),u(t)) + \Delta_{\mathfrak{f}_t}(u(t),u(t_+))\Big). \end{aligned} \tag{$\mathrm{J_{BV}}$}$$

Proof. If u is a BV solution to $(V,\mathcal{E},\Psi,\Phi)$, then $(\mathrm{E}_{\Psi,\mathrm{ineq}})$ is a trivial consequence of the energy balance (3.25) since $\mathit{Var}_{\mathfrak{f}}(u;[s,t]) \ge \mathrm{Var}_\Psi(u;[s,t])$ for every interval $[s,t]$. The jump conditions $(\mathrm{J_{BV}})$ follow by writing (3.25) in the intervals $[t,t+\eta]$ or $[t-\eta,t]$ for small $\eta > 0$ and then passing to the limit as $\eta \downarrow 0$.

In order to prove the converse implication, let suppose that $\mathrm{J}_u = (t_n)_n \subset (0,T)$ and let us call $0 = \mathsf{t}_0 < \mathsf{t}_1 < \cdots < \mathsf{t}_N < \mathsf{t}_{N+1} = T$ an ordered subdivision of $[0,T]$ such that $\{\mathsf{t}_1,\mathsf{t}_2,\cdots,\mathsf{t}_N\}$ is a permutation of $\{t_1,t_2,\cdots,t_N\} \subset \mathrm{J}_u$.

Writing $(\mathrm{E}_{\Psi,\mathrm{ineq}})$ in each interval $[\mathsf{t}_i+\eta,\mathsf{t}_{i+1}-\eta]$ for sufficiently small $\eta > 0$ and taking the limit as $\eta \downarrow 0$, also recalling $\mathrm{Var}_\Psi(u;[a,b]) \ge \mu_{\mathrm{d}}(a,b)$ (cf. (2.23)), we get

$$\mu_{\mathrm{d}}(\mathsf{t}_i,\mathsf{t}_{i+1}) \le \mathcal{E}_{\mathsf{t}_i}(u(\mathsf{t}_{i,+})) - \mathcal{E}_{\mathsf{t}_{i+1}}(u(\mathsf{t}_{i+1,-})) + \int_{\mathsf{t}_i}^{\mathsf{t}_{i+1}} \mathcal{P}_s(u(s))\,\mathrm{d}s. \tag{3.36}$$

From $(\mathrm{J_{BV}})$ and (3.13) we obtain

$$\begin{aligned} \Delta_{\mathfrak{f}_{\mathsf{t}_i}}(u(\mathsf{t}_i),u(\mathsf{t}_{i+})) + \mu_{\mathrm{d}}(\mathsf{t}_i,\mathsf{t}_{i+1}) + \Delta_{\mathfrak{f}_{\mathsf{t}_{i+1}}}(u(\mathsf{t}_{i+1,-}),u(\mathsf{t}_{i+1}))&\\ \le \mathcal{E}_{\mathsf{t}_i}(u(\mathsf{t}_i)) - \mathcal{E}_{\mathsf{t}_{i+1}}(u(\mathsf{t}_{i+1})) + \int_{\mathsf{t}_i}^{\mathsf{t}_{i+1}} \mathcal{P}_s(u(s))\,\mathrm{d}s&, \end{aligned}$$

so that summing up all the contributions (recalling that $u(\mathsf{t}_{0,+}) = u(\mathsf{t}_0) = u(0)$ and $u(\mathsf{t}_{N,-}) = u(\mathsf{t}_N) = u(T)$) we get

$$\mu_{\mathrm{d}}(0,T) + \sum_{i=1}^N \Delta_{\mathfrak{f}_{\mathsf{t}_i}}(u(\mathsf{t}_{i,-}),u(\mathsf{t}_i)) + \Delta_{\mathfrak{f}_{\mathsf{t}_i}}(u(\mathsf{t}_i),u(\mathsf{t}_{i,+})) \le \mathcal{E}_0(u(0)) - \mathcal{E}_T(u(T)) + \int_0^T \mathcal{P}_s(u(s))\,\mathrm{d}s.$$

If J_u is finite we get $(\mathrm{E}_{\mathfrak{f},\mathrm{ineq}})$ choosing $N = \#(\mathrm{J}_u)$ and recalling (2.23) and (2.24). If J_u is infinite, we simply pass to the limit as $N \uparrow +\infty$. We leave to the reader the obvious modifications in the case $\mathrm{J}_u \cap \{0,T\} \ne \emptyset$. □

The jump conditions $(\mathrm{J_{BV}})$ should be compared with the general estimate (3.14), that at every jump point $t \in \mathrm{J}_w$ of an arbitrary curve $w \in \mathrm{BV}([0,T];D_E,\Psi)$ rephrases as

$$\Big|\mathcal{E}_t(w(t_+)) - \mathcal{E}_t(w(t))\Big| \le \Delta_{\mathfrak{f}_t}(w(t),w(t_+)), \quad \Big|\mathcal{E}_t(w(t)) - \mathcal{E}_t(w(t_-))\Big| \le \Delta_{\mathfrak{f}_t}(w(t_-),w(t)). \tag{3.37}$$

We extend now the differential characterization of BV solutions in [MRS12a, Thm. 4.3] to the present setting.

Theorem 3.14 (Differential characterization of BV solutions)**.** *Let* $u \in \mathrm{BV}([0,T];V)$ *with distributional derivative decomposed as in Remark 2.2. Then* u *is a* BV *solution of the RIS* $(V,\mathcal{E},\Psi,\Phi)$ *if and only if it satisfies the doubly nonlinear differential inclusion in the* BV *sense*

$$\partial\Psi\Big(\frac{\mathrm{d}u'_{\mathrm{d}}}{\mathrm{d}\lambda}(t)\Big) + \partial\mathcal{E}_t(u(t)) \ni 0 \quad \textit{for } \lambda\textit{-a.a. } t \in (0,T) \quad \textit{with } \lambda = \|u'_{\mathrm{C}}\| + \mathscr{L}^1, \tag{$\mathrm{DN_{BV}}$}$$

and the jump conditions ($\mathrm{J_{BV}}$)*. In particular* ($\mathrm{DN_{BV}}$) *yields the pointwise inclusion*

$$\partial\Psi\big(\dot u(t)\big) + \partial\mathcal{E}_t(u(t)) \ni 0 \quad \text{for } \mathscr{L}^1\text{-a.a. } t \in (0,T). \tag{$\mathrm{DN}_{\mathscr{L}}$}$$

Proof. We briefly recall the argument presented in [MRS12a, Prop. 2.7, Thm. 4.3]. Let us first notice that ($\mathrm{DN_{BV}}$) yields the local stability condition, since the support of λ is the full interval $[0,T]$ and K^* contains the range of $\partial\Psi$. By the distributional chain rule (3.35) we get

$$e_{\mathrm{d}}' = -\Psi(\boldsymbol{n})\|u_{\mathrm{d}}'\| + \mathcal{P}_{\cdot}(u)\mathscr{L}^1 \overset{(2.27)}{=} -\mu_{\mathrm{d}} + \mathcal{P}_{\cdot}(u)\mathscr{L}^1.$$

Combining this information with the jump conditions ($\mathrm{J_{BV}}$) and recalling formula (3.23) for $Var_{\mathfrak{f}}$ we get ($\mathrm{E}_{\mathfrak{f}}$).

Conversely, if u is a solution then ($\mathrm{E}_{\Psi,\mathrm{ineq}}$) yields

$$e_{\mathrm{d}}' + \Psi(\boldsymbol{n})\|u_{\mathrm{d}}'\| - \mathcal{P}_{\cdot}(u)\mathscr{L}^1 \le 0 \quad \text{in } \mathscr{D}'(0,T).$$

Recalling (3.35) we thus obtain for $-\xi \in \partial\mathcal{E}_t(u(t)) \cap K^*$

$$\Big(\langle -\xi, \boldsymbol{n}\rangle + \Psi(\boldsymbol{n})\Big)\|u_{\mathrm{d}}'\| \le 0 \quad \text{in } \mathscr{D}'(0,T),$$

which yields the inclusion ($\mathrm{DN_{BV}}$) $\|u_{\mathrm{d}}'\|$-a.e. in $(0,T)$, and in particular $\mathscr{L}^1$-a.e. in the set $\|\dot u\| > 0$. For $\mathscr{L}^1$-a.a. points of the set $\|\dot u\| = 0$ the local stability condition still provides ($\mathrm{DN_{BV}}$). □

3.3. **Optimal jump transitions.** Thanks to the jump conditions given by ($\mathrm{J_{BV}}$), we can give a finer description of the behavior of BV solutions along jumps. The crucial notion is provided by the following definition.

Definition 3.15 (Optimal transitions)**.** *Let $t \in [0,T]$ and u_-, $u_+ \in D$ with*

$$K^* + \partial\mathcal{E}_t(u_-) \ni 0, \qquad K^* + \partial\mathcal{E}_t(u_+) \ni 0. \tag{3.38}$$

We say that an admissible curve $\vartheta \in \mathcal{T}_t(u_-, u_+)$ is an $\mathfrak{f}_t$-optimal transition *between u_- and u_+ if*

$$\mathcal{E}_t(u_-) - \mathcal{E}_t(u_+) = \Delta_{\mathfrak{f}_t}(u_-, u_+) = \mathfrak{f}_t[\vartheta, \vartheta'](r) > 0 \quad \text{for a.a. } r \in (0,1), \tag{3.39}$$

and we denote by $\mathcal{O}_t(u_-, u_+)$ the (possibly empty) collection of such optimal transitions.

We say that ϑ is of

$$\text{sliding type, if} \quad \mathfrak{e}_t(\vartheta(r)) = 0 \quad \text{for every } r \in [r_0, r_1], \tag{3.40}$$

$$\text{viscous type, if} \quad \mathfrak{e}_t(\vartheta(r)) > 0 \quad \text{for every } r \in (r_0, r_1). \tag{3.41}$$

The main interest of optimal transitions derives from the next result, whose proof follows immediately from Theorem 3.5 by a simple rescaling argument.

Proposition 3.16. *If $u \in \mathrm{BV}([0,T]; V, \Psi)$ is a* BV *solution to the rate-independent system $(V, \mathcal{E}, \Psi, \Phi)$, then for every $t \in \mathrm{J}_u$ there exists an $\mathfrak{f}_t$-optimal transition $\vartheta^t \in \mathcal{O}_t(u(t_-), u(t_+))$ such that $u(t) = \vartheta^t(r)$ for some $r \in [0,1]$.*

We now provide a characterization of *sliding* and *viscous* optimal transitions in terms of doubly nonlinear differential inclusions.

Proposition 3.17 (The structure of optimal transitions)**.** *Let $t \in [0,T]$ and u_-, $u_+ \in D$ fulfilling* (3.38) *be given and let $\vartheta \in \mathcal{T}_t(u_-, u_+)$ be an admissible transition curve with constant normalized velocity $\mathfrak{f}_t[\vartheta, \vartheta'](r) \equiv c > 0$ for a.a. $r \in (0,1)$. Then*

(1) ϑ is an optimal transition of sliding type if and only if it satisfies

$$\exists \xi(r) \in -\partial\mathcal{E}_t(\vartheta(r))) \cap K^* \quad \text{for every } r \in [0,1], \tag{3.42}$$

$$\frac{\mathrm{d}}{\mathrm{d}r}\mathcal{E}_t(\vartheta(r)) + \Psi[\vartheta'] = 0 \quad \text{for a.a. } r \in (0,1). \tag{3.43}$$

In particular, when ϑ is differentiable $\mathscr{L}^1$-a.e. in $(0,1)$, (3.42) *and* (3.43) *are equivalent to*

$$\partial\Psi(\dot\vartheta(r)) + \partial\mathcal{E}_t(\vartheta(r)) \ni 0 \quad \text{for a.a. } r \in (0,1). \tag{3.44}$$

(2) ϑ *is an optimal transition of viscous type if and only if it is differentiable* $\mathscr{L}^1$*-a.e. in* $(0,1)$ *and there exists maps* $\xi \in L^1(0,1;V^*)$*, and* $\varepsilon : (0,1) \to (0,+\infty)$ *such that*

$$\xi(r) \in \Big(\partial\Psi(\dot\vartheta(r)) + \partial\Phi(\varepsilon(r)\dot\vartheta(r))\Big) \cap \Big(- \partial\mathcal{E}_t(\vartheta(r))\Big) \quad \text{for a.a. } r \in (0,1); \tag{3.45}$$

in particular,

$$\begin{aligned} &\varepsilon(r) = \Lambda_t(\vartheta(r);\dot\vartheta(r)) \quad \text{for a.a. } r \in (0,1), \\ &\text{where } \Lambda_t(\vartheta;v) := (F^*)'(\mathfrak{e}_t(\vartheta))/F(\|v\|) \quad \vartheta \in D,\ v \in V \setminus \{0\}. \end{aligned} \tag{3.46}$$

Equivalently, there exists an absolutely continuous, surjective time rescaling $\mathsf{r} : (s_0,s_1) \to (0,1)$*, with* $-\infty \le s_0 < s_1 \le +\infty$ *and* $\dot{\mathsf{r}}(s) > 0$ *for* $\mathscr{L}^1$*-a.a.* $s \in (s_0,s_1)$*, such that the rescaled transition* $\theta(s) := \vartheta(\mathsf{r}(s))$ *satisfies the viscous differential inclusion*

$$\partial\Psi(\dot\theta(s)) + \partial\Phi(\dot\theta(s)) + \partial\mathcal{E}_t(\theta(s)) \ni 0 \quad \text{for a.a. } s \in (s_0,s_1). \tag{3.47}$$

(3) If ϑ *is an optimal transition, then it can be decomposed in a canonical way into an (at most) countable collection of optimal* sliding and viscous *transitions. Namely, there exist (uniquely determined) disjoint open intervals* $(S_j)_{j\in\sigma}$ *and* $(V_k)_{k\in\upsilon}$ *of* $(0,1)$*, with* $\sigma,\upsilon \subset \mathbb{N}$*, such that* $(0,1) \subset \big(\cup_{j\in\sigma} S_j\big) \cup \overline{\big(\cup_{k\in\upsilon} V_k\big)}$ *and*

$$\vartheta|_{S_j} \quad \text{is of sliding type,} \quad \vartheta|_{V_k} \quad \text{is of viscous type.}$$

Proof. (1) It is easy to check that if an admissible transition ϑ satisfies (3.42)–(3.43) then ϑ is an optimal transition of sliding type. Indeed, by the chain rule of Theorem 2.3 $r \mapsto \mathcal{E}_t(\vartheta(r))$ is absolutely continuous, and integrating (3.43) we get (3.39). The converse implication is even easier by combining the chain rule along ϑ, the fact that $\mathfrak{f}_t[\vartheta,\vartheta'] = \Psi[\vartheta']$, and (3.39).

(2) Similarly, if $\vartheta,\varepsilon,\xi$ satisfy (3.45), the chain rule yields

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}r}\mathcal{E}_t(\vartheta(r)) &= -\langle \xi(r),\dot\vartheta(r)\rangle = -\Psi_{\varepsilon(r)}(\varepsilon(r)\dot\vartheta(r)) - \Psi^*_{\varepsilon(r)}(\vartheta(r)) \\ &\le -\Psi(\dot\vartheta(r)) - \frac{1}{\varepsilon(r)}F(\varepsilon(r)\|\dot\vartheta(r)\|) - \frac{1}{\varepsilon(r)}F^*(\mathfrak{e}_t(\vartheta(r))) \\ &\le -\Psi(\dot\vartheta(r)) - \mathfrak{e}_t(\vartheta(r))\|\dot\vartheta(r)\| = -\mathfrak{f}_t(\vartheta(r),\dot\vartheta(r)) = -c < 0. \end{aligned}$$

Integrating in time we get one inequality of (3.39); the converse one is always true. Then, all the above inequalities are in fact equalities: in particular $\mathfrak{e}_t(\vartheta(r)) > 0$ in $(0,1)$, since $F(r) > 0$ if $r > 0$ by (D.0). We then conclude that ϑ is an optimal transition of viscous type.

The converse implication follows from the fact that

$$\mathfrak{e}_t(\vartheta)\|\dot\vartheta\| = \frac{1}{\varepsilon}F(\varepsilon\|\dot\vartheta\|) + \frac{1}{\varepsilon}F^*(\mathfrak{e}_t(\vartheta)) \quad \text{if } \varepsilon = \Lambda_t(\vartheta,\dot\vartheta).$$

Observing that $\dot\vartheta$ is locally bounded in $(0,1)$ so that $r \mapsto 1/\varepsilon(r)$ is also locally bounded, in order to get (3.47) we simply operate the absolutely continuous time rescaling

$$\mathsf{s}(r) := \int_{1/2}^r \varepsilon^{-1}(r)\,\mathrm{d}r, \quad \mathsf{r} := \mathsf{s}^{-1}, \quad \theta(s) := \vartheta(\mathsf{r}(s)), \quad \dot\theta(s) = \varepsilon(\mathsf{r}(s))\dot\vartheta(\mathsf{r}(s)).$$

(3) We can simply split the parameter interval $(0,1)$ into the open sets $V := \{r : \mathfrak{e}_t(\vartheta(r)) > 0\}$, $S := [0,1] \setminus \overline{V}$, and then we consider their connected components. $\square$

As a last result, we show that optimal transitions capture the asymptotic profile of rescaled solutions to (1.1) around a jump point.

Proposition 3.18 (Asymptotic profiles and optimal transitions)**.** *Let* $\varepsilon_k \downarrow 0$ *and let* $(u_{\varepsilon_k},\xi_{\varepsilon_k})$ *be a sequence of solutions to the viscous doubly nonlinear equation* (2.34)*, so that* u_{ε_k} *converge to a* BV *solution* u *of the RIS* $(V,\mathcal{E},\Psi,\Phi)$ *as* $k \to \infty$ *according to Theorem 3.9. For every* $t \in \mathrm{J}_u$ *let* $\alpha_k < t < \beta_k$ *be two sequences such that*

$$\alpha_k \uparrow t, \quad \beta_k \downarrow t, \quad \lim_{k\to\infty} u_{\varepsilon_k}(\alpha_k) = u(t_-), \quad \lim_{k\to\infty} u_{\varepsilon_k}(\beta_k) = u(t_+). \tag{3.48}$$

Then

$$\lim_{k\to\infty}\int_{\alpha_k}^{\beta_k}\Big(\Psi_{\varepsilon_k}(\dot u_{\varepsilon_k})+\Psi^*_{\varepsilon_k}(-\xi_{\varepsilon_k})\Big)\,\mathrm{d}r=\Delta_{\mathfrak{f}_t}(u(t_-),u(t_+)),\tag{3.49}$$

and there exist a further subsequence (not relabeled), increasing and surjective time rescalings $\mathsf{t}_k\in\mathrm{AC}([0,1];[\alpha_k,\beta_k])$, *and an optimal transition* $\vartheta\in\mathcal{O}_t(u(t_-),u(t_+))$ *such that*

$$\lim_{k\to\infty}u_{\varepsilon_k}\circ\mathsf{t}_k=\vartheta\quad\textit{strongly in } V,\ \textit{uniformly on } [0,1].\tag{3.50}$$

Proof. Estimate (3.20) from Theorem 3.5 provides the inequality

$$\liminf_{k\to\infty}\int_{\alpha_k}^{\beta_k}\Big(\Psi_{\varepsilon_k}(\dot u_{\varepsilon_k})+\Psi^*_{\varepsilon_k}(\xi_{\varepsilon_k})\Big)\,\mathrm{d}r\ge\Delta_{\mathfrak{f}_t}(u(t_-),u(t_+)).$$

On the other hand, applying (3.29) to each interval $[\alpha_h,\beta_h]$ we obviously get

$$\limsup_{k\to\infty}\int_{\alpha_k}^{\beta_k}\Big(\Psi_{\varepsilon_k}(\dot u_{\varepsilon_k})+\Psi^*_{\varepsilon_k}(\xi_{\varepsilon_k})\Big)\,\mathrm{d}r\le\mathrm{Var}_{\mathfrak{f}}(u;[\alpha_h,\beta_h])\quad\text{for every } h\in\mathbb{N}.$$

Passing to the limit as $h\uparrow\infty$ we obtain (3.49). We then apply assertion (F3) of Theorem 3.5 to find an admissible transition $\vartheta\in\mathcal{T}_t(u(t_-),u(t_+))$ and rescalings t_k such that (3.19) holds. Relation (3.49) shows that ϑ is optimal. □

3.4. **V-parameterizable solutions.** In this section we will focus on a more restrictive notion of solution, exhibiting better regularity properties: they belong to $\mathrm{BV}([0,T];V)$ and at all jump points the left and the right limits can be connected by an optimal transition with finite *V-length*. Moreover, we will require that the total V-length of the connecting paths is finite.

Definition 3.19 (V-parameterizable BV solutions)**.** *A balanced viscosity solution u of the RIS $(V,\mathcal{E},\Psi,\Phi)$ (in the sense of Definition 3.8) is called V*-parameterizable *if $u\in\mathrm{BV}([0,T];V)$ and*

$$\begin{aligned}&i)\quad \forall\, t\in\mathrm{J}_u\quad \exists\,\vartheta^t\in\mathcal{O}_t(u(t_-),u(t_+))\cap\mathrm{AC}([0,1];V),\\ &ii)\quad \sum_{t\in\mathrm{J}_u}\int_0^1\|\dot\vartheta^t(r)\|\,\mathrm{d}r<\infty.\end{aligned}\tag{3.51}$$

The notion of *V-parameterizable* BV solution slightly differs from the concept of *connectable* BV *solution* introduced in [Mie11, Def. 4.21], which only requires condition i).

As one can expect, a limit curve of solutions to (1.1) satisfying a uniform $\mathrm{BV}([0,T];V)$-bound is a V-parameterizable solution.

Theorem 3.20. *Let $(u_\varepsilon)_{\varepsilon>0}$ be a family of solutions to* (1.1) *satisfying* (3.26) *at $t=0$ and the uniform bound*

$$\exists\, C>0\ \ \forall\,\varepsilon>0:\quad \mathrm{Var}(u_\varepsilon;[0,T])\le C.\tag{3.52}$$

Then any limit curve as in Theorem 3.9 is a V-parameterizable BV *solution to the RIS $(V,\mathcal{E},\Psi,\Phi)$. Similarly, let $(\mathrm{U}^n_{\tau,\varepsilon})_{\tau,\varepsilon}$ be a family of discrete solutions to* $(\mathrm{IP}_{\varepsilon,\tau})$, *satisfying* (3.30) *and* (3.31)*. If*

$$\exists\, C>0\ \ \forall\,\tau,\,\varepsilon>0:\quad \mathrm{Var}(\mathrm{U}_{\tau,\varepsilon};[0,T])=\sum_{n=1}^{N_\tau}\|\mathrm{U}^n_{\tau,\varepsilon}-\mathrm{U}^{n-1}_{\tau,\varepsilon}\|\le C,\tag{3.53}$$

then any accumulation point of the piecewise affine interpolants $\mathrm{U}_{\tau,\varepsilon}$ as in Theorem 3.10 is a V-parameterizable solution.

Proof. The proofs of the two statements are very similar, thus we only prove the first one.

Since the total variation functional is lower semicontinuous with respect to pointwise convergence, any limit curve u obtained as in Theorem 3.9 clearly belongs to $\mathrm{BV}([0,T];V)$.

In order to check i) of (3.51) we apply Proposition 3.18 and we find a sequence of rescalings $\mathsf{t}_k:[0,1]\to[\alpha^t_k,\beta^t_k]$ (we explicitly indidate the dependence of the time intervals $[\alpha_k,\beta_k]$ on t) and an optimal transition $\vartheta^t\in\mathcal{O}_t(u(t_-),u(t_+)$ with (3.48) and (3.50). This shows that

$$\mathrm{Var}(\vartheta^t;[0,1])\le\liminf_{k\to\infty}\mathrm{Var}(u_{\varepsilon_k};[\alpha^t_k,\beta^t_k])<\infty,\tag{3.54}$$

so that $\vartheta^t \in \mathrm{BV}(0,1;V)$. Since ϑ^t is also continuous, up to a further time rescaling we can obtain an optimal transition absolutely continuous in V.

A slight refinement of the above argument also provides *ii*): we consider an arbitrary finite collection of points $t_1, t_2, \dots t_h \subset \mathrm{J}_u$ and we choose a common subsequence u_{ε_k} satisfying (3.48) in each interval. For sufficiently big k so that the intervals $[\alpha_k^{t_j}, \beta_k^{t_j}]$ are disjoint, (3.54) yields

$$\sum_{j=1}^{h} \mathrm{Var}(\vartheta^{t_j};[0,1]) \le \liminf_{k\to\infty} \sum_{j=1}^{h} \mathrm{Var}(u_{\varepsilon_k};[\alpha_k^{t_j},\beta_k^{t_j}]) \le \liminf_{k\to\infty} \mathrm{Var}(u_{\varepsilon_k};[0,T]) \overset{(3.52)}{\le} C.$$

Since the number h of jump points is arbitrary, we obtain *ii).* $\square$

The next results show that one can actually prove (3.52) and (3.53) for the particular choice

$$\Phi(v) = \frac{1}{2}\|v\|^2, \quad F(r) := \frac{1}{2}r^2, \tag{3.55}$$

under slightly more restrictive assumptions on the energy functional and on the initial data: besides the usual (D.0)–(D.1) and (E.0)–(E.2), we will also assume that for every $E > 0$ there exist constant $\alpha_E, \Lambda_E, L_E > 0$ such that the energy functional satisfies the Gårding-like subdifferentiability inequality

$$\mathcal{E}_t(v) - \mathcal{E}_t(u) \ge \langle \xi, v-u\rangle + \alpha_E\|v-u\|^2 - \Lambda_E \Psi_\wedge(v-u)\|v-u\| \quad \text{if } u, v \in D_E,\ \xi \in \partial\mathcal{E}_t(u). \tag{3.56}$$

We will also require that the power functional is uniformly Lipschitz in D_E, viz.

$$|\mathcal{P}_t(u) - \mathcal{P}_t(v)| \le L_E\|u-v\| \quad \text{if } t \in [0,T],\ u, v \in D_E. \tag{3.57}$$

Then, we have the following result.

Theorem 3.21 (A priori estimates for discrete Minimizing Movements)**.** *Assume that* (3.55)–(3.57) *hold. Then any family of solutions* $(\mathrm{U}^n_{\tau,\varepsilon})$ *of* $(\mathrm{IP}_{\varepsilon,\tau})$ *fulfilling, for some constants* $E_0, Q > 0$,

$$\Psi(\mathrm{U}^0_{\tau,\varepsilon}) + \mathcal{E}_0(\mathrm{U}^0_{\tau,\varepsilon}) \le E_0, \quad \tau \le Q\varepsilon, \quad K^* + \partial\mathcal{E}_0(\mathrm{U}^0_{\tau,\varepsilon}) \ni 0, \tag{3.58}$$

satisfies estimates (3.53)*. In particular, if* (3.30), (3.31) *and* (3.58) *hold, any curve u obtained as limit of the piecewise affine interpolants* $\mathrm{U}_{\tau,\varepsilon}$ *(cf. Theorem 3.10) is a* V*-parameterizable solution.*

The proof will be given in Section 7.4. A similar priori estimate in the form $\int_0^T \|\dot u_\varepsilon(t)\|\,\mathrm{d}t \le C$ was derived in [MiZ12] for semi- and quasilinear partial differential equations with smooth nonlinearities. There Galerkin approximation and differentiation in time is used. Like in the present case, where we have to confine ourselves to Minimizing Movement solutions (cf. Corollary 3.22 below), in [MiZ12] the a priori estimate in $\mathrm{BV}([0,T];V)$ can only be shown for a suitable subclass of solutions to (1.1), cf. [MiZ12, Def. 4.3]. This establishes an interesting parallel between our Minimizing Movement approach, and the one in [MiZ12].

Corollary 3.22 (A priori estimate for Minimizing Movement solutions)**.** *Assume that* (3.55)–(3.57) *hold. Then every family* $(u_\varepsilon)_\varepsilon \subset \mathrm{AC}([0,T];V)$ *of Minimizing Movement solutions to* (1.1)*, fulfilling*

$$u_\varepsilon(0) \to u_0 \quad \text{in } V, \quad \mathcal{E}_0(u_\varepsilon(0)) \to \mathcal{E}_0(u_0), \quad K^* + \partial\mathcal{E}_0(u_\varepsilon(0)) \ni 0, \tag{3.59}$$

satisfies estimate (3.52)*. Any limit u is a* V*-parameterizable solution to the RIS* $(V, \mathcal{E}, \Psi, \Phi)$*.*

Proof. Choose $\mathrm{U}^0_{\tau,\varepsilon} = u_\varepsilon(0)$ and apply Theorem 2.5, passing to the limit in estimate (3.53). $\square$

The following result is an immediate consequence of Corollary 3.22 or Theorem 3.21.

Corollary 3.23 (Existence of V-parameterizable BV solutions)**.** *If* (3.55)–(3.57) *hold, then for every* $u_0 \in D$ *with* $K^* + \partial\mathcal{E}_0(u_0) \ni 0$ *there exists a* V*-parameterizable* BV *solution to the RIS* $(V, \mathcal{E}, \Psi, \Phi)$ *starting form* u_0.

Notice that the subdifferentiability condition (3.56) implies (E.3) as well as

$$
\begin{aligned}
\langle \eta-\xi, v-u\rangle \geq 2\alpha_E\|v-u\|^2 - 2\Lambda_E\Psi_\wedge(v-u)\,\|v-u\| - L_E|t-s|\,\|v-u\| \\
\text{whenever } \eta\in\partial\mathcal{E}_t(v),\ \xi\in\partial\mathcal{E}_s(u),\ u,v\in D_E,\ \ s,t\in[0,T].
\end{aligned}
\tag{3.60}
$$

To check (3.60), it is sufficient to write (3.56) for u and v at time s,t respectively. Adding the two inequalities and using (3.57) we get the bound (assuming $s<t$)

$$
\mathcal{E}_t(v)-\mathcal{E}_s(v)+\mathcal{E}_s(u)-\mathcal{E}_t(u) \leq \int_s^t \big(\mathcal{P}_r(v)-\mathcal{P}_r(u)\big)\,\mathrm{d}r \leq L_E\,(t-s)\|u-v\|.
$$

Observe that in (3.56), as in (E.3), we allow for a *negative* modulus of convexity in the Ψ-term, provided that it is possible to gain an even small *positive* modulus of subdifferentiability in the stronger V-norm. This is akin to the Gårding inequality for elliptic operators.

The next result provides a useful criterium on the energy functional $\mathcal{E}$ to establish the subdifferentiability condition (3.56). It is a sort of *(generalized) λ-convexity* condition, involving two norms. Notice that both (3.61) and (3.56) are required to hold on *sublevels* of $\mathcal{E}$, only.

Lemma 3.24. *Suppose that for every $E>0$ there exist constant $\alpha_E,\Lambda_E>0$ such that the energy functional $\mathcal{E}_t : V\to(-\infty,+\infty]$ satisfies*

$$
\mathcal{E}_t((1-\theta)u+\theta v) \leq (1-\theta)\mathcal{E}_t(u)+\theta\mathcal{E}_t(v)-\theta(1-\theta)\left(\alpha_E\|u-v\|^2-\Lambda_E\Psi_\wedge(u-v)\|u-v\|\right)
\tag{3.61}
$$

for every $u,v\in D_E$ and $\theta\in[0,1]$. Then its Fréchet subdifferential $\partial\mathcal{E}_t : V\rightrightarrows V^*$ satisfies (3.56).

Proof. For ξ lying in the Fréchet subdifferential $\partial\mathcal{E}_t(u)$ there holds for every v, $u\in D_E$ and $\theta\downarrow 0$

$$
\begin{aligned}
\langle \xi, \theta(v-u)\rangle + o(\theta\|v-u\|) &\leq \mathcal{E}_t((1-\theta)u+\theta v)-\mathcal{E}_t(u) \\
&\leq \theta(\mathcal{E}_t(v)-\mathcal{E}_t(u)) - \theta(1-\theta)(\alpha_E\|v-u\|^2-\Lambda_E\Psi_\wedge(v-u)\|v-u\|).
\end{aligned}
$$

Dividing both sides of the inequality by θ, the limit $\theta\downarrow 0$ yields the desired estimate (3.56). □

4. Parameterized solutions

4.1. **Vanishing-viscosity analysis, parameterized curves and solutions.** Under the working assumptions of §2.1 (in particular, (D.0)–(D.2) and (E.0)–(E.3)), in this section we will present a different approach to the vanishing-viscosity analysis of (1.1), which goes back to [EfM06] and was further developed in [MRS09, MRS12a]. The main idea is to rescale time in (1.1) and study the limiting behavior as $\varepsilon\downarrow 0$ of the *rescaled* viscous solutions. This naturally leads to the notion of parameterized solution in Definition 4.2: it is a *space-time parameterized curve*, along which the energy $\mathcal{E}$ fulfills a "parameterized" version of the energy-dissipation identity (2.35). At the end of this section, we will also discuss the parameterized counterpart to V-parameterizable BV solutions. Let us emphasize that, while parameterized solutions were developed in [EfM06, MiZ12] in their own right, we use them mainly to obtain the desired results for BV solutions.

Vanishing-viscosity analysis. Let $(u_\varepsilon)_\varepsilon$ be a family of solutions to the "viscous" doubly nonlinear equation (1.1). It follows from the energy identity (2.35) and from the variational characterization of $\mathfrak{f}$ (3.1)–(3.4) that

$$
\int_s^t \mathfrak{f}_r(u_\varepsilon(r);\dot u_\varepsilon(r))\,\mathrm{d}r + \mathcal{E}_t(u(t)) \leq \mathcal{E}_s(u(s)) + \int_s^t \mathcal{P}_r(u(r))\,\mathrm{d}r \quad \text{for all } 0\leq s\leq t\leq T,
\tag{4.1}
$$

whence, relying on the power control (E.2), we deduce that there exists a constant $C>0$ such that

$$
\mathsf{S}_\varepsilon := T + \int_0^T \mathfrak{f}_r(u_\varepsilon(r);\dot u_\varepsilon(r))\,\mathrm{d}r \leq C \quad \text{for every } \varepsilon>0.
\tag{4.2}
$$

We rescale the functions u_ε by the *energy-dissipation arclength* $\mathsf{s}_\varepsilon : [0,T] \to [0,\mathsf{S}_\varepsilon]$ of the curve u_ε, defined by

$$\mathsf{s}_\varepsilon(t) := t + \int_0^t \mathfrak{f}_r(u_\varepsilon(r); \dot u_\varepsilon(r))\,\mathrm{d}r. \tag{4.3}$$

Hence, we introduce the rescaled functions $(\mathsf{t}_\varepsilon, \mathsf{u}_\varepsilon) : [0,\mathsf{S}_\varepsilon] \to [0,T] \times V$

$$\mathsf{t}_\varepsilon(s) := \mathsf{s}_\varepsilon^{-1}(s)\,, \quad \mathsf{u}_\varepsilon(s) := u_\varepsilon(\mathsf{t}_\varepsilon(s)). \tag{4.4}$$

We write the "rescaled energy identity" fulfilled by the triple $(\mathsf{t}_\varepsilon, \mathsf{u}_\varepsilon)$ by means of the space-time Finsler dissipation functionals $\mathfrak{F}_\varepsilon, \mathfrak{G}_\varepsilon : [0,T] \times D \times [0,+\infty) \times V \to [0,+\infty)$ defined by

$$\begin{aligned} &\mathfrak{F}_\varepsilon(\mathsf{t},\mathsf{u};\alpha,\mathsf{v}) := \Psi(\mathsf{v}) + \mathfrak{G}_\varepsilon(\mathsf{t},\mathsf{u};\alpha,\mathsf{v}) - \alpha\mathcal{P}_\mathsf{t}(\mathsf{u}) \qquad \text{with} \\ &\mathfrak{G}_\varepsilon(\mathsf{t},\mathsf{u};\alpha,\mathsf{v}) := \begin{cases} \frac{\alpha}{\varepsilon}\Phi(\frac{\varepsilon}{\alpha}\mathsf{v}) + \frac{\alpha}{\varepsilon}F^*(\mathfrak{e}_t(u)) & \text{for } \alpha > 0, \\ \infty & \text{for } \alpha = 0, \end{cases} \end{aligned} \tag{4.5}$$

where we combined (2.30) for Ψ_ε^*, yielding (3.5) for $\mathfrak{f}_t$, and the monotonicity of F^* to find

$$\inf_{\xi \in -\partial\mathcal{E}_t(u)} \Psi_\varepsilon^*(\xi) = \inf_{\substack{\xi \in -\partial\mathcal{E}_t(u) \\ z \in K^*}} \frac{1}{\varepsilon}F^*(\|\xi - z\|_*) = \frac{1}{\varepsilon}F^*(\mathfrak{e}_t(u)).$$

Then, the energy identity (2.35) yields for every $0 \le s_1 < s_2 \le \mathsf{S}_\varepsilon$

$$\int_{s_1}^{s_2} \mathfrak{F}_\varepsilon(\mathsf{t}_\varepsilon(s), \mathsf{u}_\varepsilon(s); \dot{\mathsf{t}}_\varepsilon(s), \dot{\mathsf{u}}_\varepsilon(s))\,\mathrm{d}s + \mathcal{E}_{\mathsf{t}_\varepsilon(s_2)}(\mathsf{u}_\varepsilon(s_2)) = \mathcal{E}_{\mathsf{t}_\varepsilon(s_1)}(\mathsf{u}_\varepsilon(s_1)), \tag{4.6}$$

and, on account of our choice (4.3) of the reparameterization, we have the *normalization condition*

$$\dot{\mathsf{t}}_\varepsilon(s) + \mathfrak{f}_{\mathsf{t}_\varepsilon(s)}(\mathsf{u}_\varepsilon(s); \dot{\mathsf{u}}_\varepsilon(s)) \equiv 1 \quad \text{for a.a. } s \in (0, \mathsf{S}_\varepsilon)\,. \tag{4.7}$$

From (4.6) it is possible to deduce a priori estimates on the family $(\mathsf{t}_\varepsilon, \mathsf{u}_\varepsilon)_\varepsilon$, thus proving that, up to a subsequence, the functions $(\mathsf{t}_\varepsilon, \mathsf{u}_\varepsilon)$ converge in a suitable sense to a pair $(\mathsf{t}, \mathsf{u}) : [0,S] \to [0,T] \times V$ (see Thm. 4.3 for a precise statement). In view of the forthcoming lower semicontinuity Proposition 7.1, we expect that taking the limit $\varepsilon \to 0$ in (4.6) leads to the energy estimate

$$\int_{s_1}^{s_2} \mathfrak{F}(\mathsf{t}(s), \mathsf{u}(s); \dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s))\,\mathrm{d}s + \mathcal{E}_{\mathsf{t}(s_2)}(\mathsf{u}(s_2)) \le \mathcal{E}_{\mathsf{t}(s_1)}(\mathsf{u}(s_1)) \quad \text{for all } 0 \le s_1 \le s_2 \le \mathsf{S}. \tag{4.8}$$

The functional $\mathfrak{F} : [0,T] \times D \times [0,+\infty) \times V \to [0,+\infty]$ is defined by

$$\begin{aligned} &\mathfrak{F}(\mathsf{t},\mathsf{u};\alpha,\mathsf{v}) := \Psi(\mathsf{v}) + \mathfrak{G}(\mathsf{t},\mathsf{u};\alpha,\mathsf{v}) - \alpha\mathcal{P}_\mathsf{t}(\mathsf{u}) \quad \text{with} \\ &\mathfrak{G}(\mathsf{t},\mathsf{u};\alpha,\mathsf{v}) := \mathfrak{k}_\mathsf{t}(\mathsf{u})\alpha + \mathfrak{e}_\mathsf{t}(\mathsf{u})\|\mathsf{v}\| = \begin{cases} \mathfrak{k}_\mathsf{t}(\mathsf{u}) & \text{if } \alpha > 0, \\ \mathfrak{e}_\mathsf{t}(\mathsf{u})\|\mathsf{v}\| & \text{if } \alpha = 0. \end{cases} \end{aligned} \tag{4.9}$$

Here we have adopted the convention $0 \cdot (+\infty) = 0$, and $\mathfrak{k}$ is the indicator function

$$\mathfrak{k}_\mathsf{t}(\mathsf{u}) := \inf_{\xi \in -\partial\mathcal{E}_\mathsf{t}(\mathsf{u})} \mathrm{I}_{K^*}(\xi) = \mathrm{I}_{\{0\}}(\mathfrak{e}_t(u)) = \begin{cases} 0 & \text{if } K^* + \partial\mathcal{E}_\mathsf{t}(\mathsf{u}) \ni 0, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.10}$$

Hence, it would be natural to take (4.8) as definition of parameterized solution. However, as already mentioned, limit curves have to be expected in $\mathrm{AC}([0,\mathsf{S}]; V, \Psi)$, i.e. they might lose the differentiability property with respect to time. Thus, we need to develop a more refined definition.

Admissible parameterized curves and solutions. In order to properly formulate (4.8) we need to resort to the metric Ψ-derivative introduced in the beginning of Section 2.2. Based on that definition, we first introduce a suitable class of *parameterized* curves.

Definition 4.1 (Admissible parameterized curves). *We call a pair* $(\mathsf{t}, \mathsf{u}) : [\mathsf{a}, \mathsf{b}] \to [0,T] \times V$ *an* admissible parameterized curve

(1) if t *is nondecreasing and absolutely continuous,* $\mathsf{u} \in \mathrm{AC}([\mathsf{a},\mathsf{b}]; D_E, \Psi)$ *for some* $E > 0$*,*

(2) if u *is locally* V*-absolutely continuous in the open set*

$$G := \left\{ s \in [\mathsf{a},\mathsf{b}] \ : \ \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s)) > 0 \right\} \ = \ \left\{ s \in [\mathsf{a},\mathsf{b}] \ : \ K^* + \partial \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \not\ni 0 \right\}, \tag{4.11}$$

and t *is constant in each connected component of* G *(in particular* u *is differentiable* $\mathscr{L}^1$*-a.e. in* G*),*

(3) and if we have the estimate

$$\int_{\mathsf{a}}^{\mathsf{b}} \Psi[\mathsf{u}'](s)\,\mathrm{d}s + \int_G \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\|\,\mathrm{d}s < \infty. \tag{4.12}$$

For every admissible parameterized curve and all $s \in [\mathsf{a},\mathsf{b}]$ *we set*

$$\begin{aligned} \mathfrak{G}[\mathsf{t},\mathsf{u};\dot{\mathsf{t}},\dot{\mathsf{u}}](s) &:= \mathfrak{k}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s) + \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{u}(s)\|, \\ \mathfrak{F}[\mathsf{t},\mathsf{u};\mathsf{t}',\mathsf{u}'](s) &:= \Psi[\mathsf{u}'](s) + \mathfrak{G}[\mathsf{t},\mathsf{u};\dot{\mathsf{t}},\dot{\mathsf{u}}](s) - \mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s), \end{aligned} \tag{4.13}$$

where, with a slight abuse of notation, we adopted the convention to set

$$\mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\| \equiv 0 \quad \textit{if } s \notin G. \tag{4.14}$$

By $\mathscr{A}(\mathsf{a},\mathsf{b};[0,T]\times V)$ *we denote the collection of all the (admissible) parameterized curves. Furthermore, we call* (t,u)

- nondegenerate, *if* $\dot{\mathsf{t}}(s) + \Psi[\mathsf{u}'](s) > 0$ *for a.a.* $s \in (\mathsf{a},\mathsf{b})$*;*
- surjective, *if* $\mathsf{t}(\mathsf{a}) = 0, \mathsf{t}(\mathsf{b}) = T$*;*
- m-normalized *for a positive* $\mathsf{m} \in L^\infty(0,\mathsf{S})$ *(typically* $\mathsf{m} \equiv 1$*), if* (t,u) *fulfills*

$$\dot{\mathsf{t}}(s) + \Psi[\mathsf{u}'](s) + \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{u}(s)\| = \mathsf{m}(s) \quad \textit{for a.a. } s \in (\mathsf{a},\mathsf{b})\,. \tag{4.15}$$

Two (admissible) parameterized curves $s \in [\mathsf{a},\mathsf{b}] \mapsto (\mathsf{t}(s),\mathsf{u}(s))$ *and* $\sigma \in [\mathsf{c},\mathsf{d}] \mapsto (\hat{\mathsf{t}}(\sigma),\hat{\mathsf{u}}(\sigma))$ *are equivalent if there exists an* absolutely continuous and surjective change of variable $\mathsf{s} : \sigma \in [\mathsf{c},\mathsf{d}] \mapsto \mathsf{s}(\sigma) \in [\mathsf{a},\mathsf{b}]$ *such that*

$$\hat{\mathsf{t}}(\sigma) = \mathsf{t}(\mathsf{s}(\sigma)), \quad \hat{\mathsf{u}}(\sigma) = u(\mathsf{s}(\sigma)) \quad \textit{for all } \sigma \in (\mathsf{c},\mathsf{d}), \qquad \dot{\mathsf{s}}(\sigma) > 0 \quad \textit{for a.a. } \sigma \in (\mathsf{c},\mathsf{d}).$$

The above concept is nothing but the parameterized counterpart to the notion of admissible curve from Definition 3.2: a crucial feature of parameterized curves is their $\mathscr{L}^1$-a.e. differentiability on the set G.

In the next definition of parameterized solutions we will impose (a suitable version of) (4.8) as an equality. Indeed, the upper energy estimate has been motivated throughout (4.6)–(4.8) via lower semicontinuity arguments. The lower energy estimate is a consequence of the chain rule of the forthcoming Theorem 4.4.

Definition 4.2 (Parameterized solutions)**.** *A* parameterized solution of the RIS $(V,\mathcal{E},\Psi,\Phi)$ *is a surjective and nondegenerate curve* $(\mathsf{t},\mathsf{u}) \in \mathscr{A}(\mathsf{a},\mathsf{b};[0,T]\times V)$ *(cf. Def. 4.1) satisfying*

$$\int_{s_1}^{s_2} \mathfrak{F}[\mathsf{t},\mathsf{u};\mathsf{t}',\mathsf{u}']\,\mathrm{d}s + \mathcal{E}_{\mathsf{t}(s_2)}(\mathsf{u}(s_2)) = \mathcal{E}_{\mathsf{t}(s_1)}(\mathsf{u}(s_1)) \qquad \textit{for all } \mathsf{a} \le s_1 \le s_2 \le \mathsf{b}. \tag{4.16}$$

Since $\mathfrak{F}$ defined in (4.13) contains the term $\mathfrak{k}_{\mathsf{t}}(\mathsf{u})\dot{\mathsf{t}}$, the equation (4.16) encompasses the local stability condition ($\mathrm{S}_{\mathrm{loc}}$). It follows from (4.12) and the power-control condition (E.2) that, along a parameterized solution, the map $s \mapsto \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))$ is absolutely continuous on $[\mathsf{a},\mathsf{b}]$.

The main existence and convergence result. The main result of this section states that any limit curve of the rescaled family $(\mathsf{t}_\varepsilon,\mathsf{u}_\varepsilon)$ of solutions to (1.1) is a *parameterized solution.*

Theorem 4.3. *Assume* (D.0)–(D.2) *and* (E.0)–(E.3)*. Let* $(u_\varepsilon)_\varepsilon \subset \mathrm{AC}([0,T];V)$ *be a family of solutions to the doubly nonlinear equation* (1.1)*, such that*

$$u_\varepsilon(0) \to u_0 \quad \textit{in } V \textit{ and } \quad \mathcal{E}_0(u_\varepsilon(0)) \to \mathcal{E}_0(u_0) \quad \textit{as } \varepsilon \downarrow 0 \tag{4.17}$$

as in (3.26). *Choose non-decreasing surjective time-rescalings* $\mathsf{t}_\varepsilon : [0,\mathsf{S}] \to [0,T]$, *define* $\mathsf{u}_\varepsilon : [0,\mathsf{S}] \to V$ *by* $\mathsf{u}_\varepsilon(s) := u_\varepsilon(\mathsf{t}_\varepsilon(s))$ *for all* $s \in [0,\mathsf{S}]$ *and suppose that*

$$\exists\, \mathsf{m} \in L^\infty(0,\mathsf{S}) \ : \quad \mathsf{m}_\varepsilon := \dot{\mathsf{t}}_\varepsilon + \mathfrak{f}_{\mathsf{t}_\varepsilon}(\mathsf{u}_\varepsilon, \dot{\mathsf{u}}_\varepsilon) \rightharpoonup^* \mathsf{m} \quad \text{in } L^\infty(0,\mathsf{S}) \text{ and } \mathsf{m} > 0 \text{ a.e. in } (0,\mathsf{S})\,. \tag{4.18}$$

Then, there exist a subsequence $\varepsilon_k \downarrow 0$ *and a* parameterized solution $(\mathsf{t},\mathsf{u}) \in \mathrm{AC}([0,\mathsf{S}];[0,T]\times V)$ *to the RIS* $(V,\mathcal{E},\Psi,\Phi)$, *such that the following convergences hold as* $k \to \infty$:

$$(\mathsf{t}_{\varepsilon_k}, \mathsf{u}_{\varepsilon_k}) \longrightarrow (\mathsf{t},\mathsf{u}) \quad \text{in } \mathrm{C}^0([0,\mathsf{S}];[0,T]\times V), \tag{4.19}$$

$$\mathcal{E}_{\mathsf{t}_{\varepsilon_k}(s)}(\mathsf{u}_{\varepsilon_k}(s)) \longrightarrow \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \quad \text{uniformly in } [0,\mathsf{S}], \tag{4.20}$$

$$\int_{s_1}^{s_2} \Big(\Psi(\dot{\mathsf{u}}_{\varepsilon_k}) + \mathfrak{G}_{\varepsilon_k}(\mathsf{t}_{\varepsilon_k}, \mathsf{u}_{\varepsilon_k}; \dot{\mathsf{t}}_{\varepsilon_k}, \dot{\mathsf{u}}_{\varepsilon_k})\Big)\,\mathrm{d}s \longrightarrow \int_{s_1}^{s_2} \Big(\Psi[\mathsf{u}'] + \mathfrak{G}[\mathsf{t},\mathsf{u};\dot{\mathsf{t}},\dot{\mathsf{u}}]\Big)\,\mathrm{d}s \tag{4.21}$$

for all $0 \le s_1 \le s_2 \le \mathsf{S}$. *Moreover,* (t,u) *is* m*-normalized.*

We have already seen that the choice (4.3)–(4.4) provides the normalization condition (4.7), and thus (up to a multiplication factor converging to 1) the curves $(\mathsf{t}_\varepsilon, \mathsf{u}_\varepsilon)$ satisfy (4.18) with $\mathsf{m} \equiv 1$.

The proof of this result is postponed to the end of § 7.3.

Chain rule and further properties of parameterized solutions. We present now a *parametrized* version of the chain rule (2.32) (cf. also (3.34)), satisfied by admissible parameterized curves. In fact, (4.22) is a *metric-like* chain-rule inequality, since it involves the Ψ-metric derivative of the curve. A key ingredient of its proof is the *uniform subdifferentiability* condition (E.3).

Theorem 4.4 (Chain-rule inequality for parameterized curves)**.** *If* $(\mathsf{t},\mathsf{u}) \in \mathscr{A}(\mathsf{a},\mathsf{b};[0,T]\times V)$ *then the map* $s \mapsto \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))$ *is* absolutely continuous *on* $[\mathsf{a},\mathsf{b}]$ *and the following* chain-rule inequality *holds for a.a.* $s \in (\mathsf{a},\mathsf{b})$ *(recalling* (4.14)*)*

$$\Big|\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s)\Big| \le \Psi[\mathsf{u}'](s) + \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\|. \tag{4.22}$$

Moreover, if u *is a.e. differentiable, then for a.a.* $s \in (\mathsf{a},\mathsf{b})$ *we have*

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s) = -\langle \xi, \dot{\mathsf{u}}(s)\rangle \ge -\mathfrak{f}_{\mathsf{t}(s)}(\mathsf{u}(s);\dot{\mathsf{u}}(s)) \ \text{for all } \xi \in -\partial\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)). \tag{4.23}$$

We postpone the proof to Section 6.1. As a straightforward consequence of the chain-rule inequality (4.22), we can characterize parameterized solutions by a simpler *one-sided* inequality on the interval (a,b). The result below corresponds to Corollary 3.12 for BV solutions.

Corollary 4.5. *For every surjective and nondegenerate admissible curve in* $(\mathsf{t},\mathsf{u}) \in \mathscr{A}(\mathsf{a},\mathsf{b};[0,T]\times V)$ *the following three conditions are equivalent:*

i) (t,u) *is a parameterized solution of the RIS* $(V,\mathcal{E},\Psi,\Phi)$;

ii) $$\int_{\mathsf{a}}^{\mathsf{b}} \mathfrak{F}[\mathsf{t},\mathsf{u};\mathsf{t}',\mathsf{u}']\,\mathrm{d}s + \mathcal{E}_{\mathsf{t}(\mathsf{b})}(\mathsf{u}(\mathsf{b})) \le \mathcal{E}_{\mathsf{t}(\mathsf{a})}(\mathsf{u}(\mathsf{a})); \tag{4.24}$$

iii) $$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s) = -\Psi[\mathsf{u}'](s) - \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\| \ \text{for a.a. } s \in (\mathsf{a},\mathsf{b}). \tag{4.25}$$

When u is $\mathscr{L}^1$-a.e. differentiable, it is also possible to characterize parameterized solutions in terms of a doubly nonlinear differential inclusion involving the dissipation potentials Ψ and Φ (to be compared with the differential characterization of BV solutions in Theorem 3.14).

Proposition 4.6. *If* (t,u) *is a* $\mathscr{L}^1$*-a.e. differentiable parameterized solution of the RIS* $(V,\mathcal{E},\Psi,\Phi)$, *then there exist measurable functions* $\lambda : (\mathsf{a},\mathsf{b}) \to [0,+\infty)$ *and* $\xi : (\mathsf{a},\mathsf{b}) \to V^*$ *such that*

$$\xi(s) \in \Big(\partial\Psi(\dot{\mathsf{u}}(s)) + \partial\Phi(\lambda(s)\dot{\mathsf{u}}(s))\Big) \cap \Big(-\partial\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))\Big), \quad \lambda(s)\dot{\mathsf{t}}(s) = 0 \quad \text{for a.a. } s \in (\mathsf{a},\mathsf{b}). \tag{4.26}$$

Conversely, if an absolutely continuous, surjective, nondegenerate and $\mathscr{L}^1$*-a.e. differentiable curve* $(\mathsf{t},\mathsf{u}) : [\mathsf{a},\mathsf{b}] \to [0,T]\times D_E$ *satisfies* (4.26) *for some measurable maps* λ, ξ *and* $s \mapsto \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))$ *is absolutely continuous in* $[\mathsf{a},\mathsf{b}]$, *then* (t,u) *is a parameterized solution to the RIS* $(V,\mathcal{E},\Psi,\Phi)$.

The reformulation of the notion of parameterized solutions in terms of the subdifferential inclusion (4.26) reflects the following *mechanical interpretation*:

- the regime ($\dot{\mathsf{t}} > 0,\ \dot{\mathsf{u}} \equiv 0$) corresponds to *sticking*;
- the regime ($\dot{\mathsf{t}} > 0,\ \dot{\mathsf{u}} \neq 0$) corresponds to *rate-independent sliding* ($\lambda = 0$ implies the local stability $K^* + \partial\mathcal{E}_{\mathsf{t}}(\mathsf{u}) \ni 0$);
- when $\dot{\mathsf{t}} = 0$ (i.e. at a jump in the (slow) external time scale, encoded in the function t), the system may switch to a *viscous regime* (when $\lambda > 0$), and the solution follow a viscous transition path.

Proof. If (t, u) is a $\mathscr{L}^1$-a.e. differentiable parameterized solution, (4.25) and (4.23) show that for every selection $\xi \in -\partial\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))$ we have

$$\langle \xi, \dot{\mathsf{u}}(s)\rangle = \Psi(\dot{\mathsf{u}}(s)) + \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\| \quad \text{for a.a. } s \in (\mathsf{a}, \mathsf{b}). \tag{4.27}$$

If $\mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s)) = 0$ then choosing $\xi \in K^*$ we get (4.26) with $\lambda(s) = 0$. If $\mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s)) > 0$ then $\dot{\mathsf{t}}(s) = 0$ so that $\dot{\mathsf{u}}(s) \neq 0$ by the nondegeneracy condition; we obtain (4.26) by choosing $\lambda(s) = \Lambda_{\mathsf{t}(s)}(\mathsf{u}(s), \dot{\mathsf{u}}(s))$, see (3.46).

Conversely, assume (4.26) and that the energy map is absolutely continuous. If $\lambda(s) = 0$ then $\mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s)) = 0$ so that $\langle \xi, \dot{\mathsf{u}}(s)\rangle = \Psi(\dot{\mathsf{u}}(s))$. If $\lambda(s) > 0$ then $\dot{\mathsf{t}}(s) = 0$ so that $\dot{u}(s) \neq 0$ and

$$\langle \xi, \dot{\mathsf{u}}(s)\rangle = \Psi(\dot{\mathsf{u}}(s)) + \frac{1}{\lambda(s)}\Phi(\lambda(s)\dot{\mathsf{u}}(s)) + \frac{1}{\lambda(s)}\Phi^*(\xi) \geq \Psi(\dot{\mathsf{u}}(s)) + \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\| \geq \langle \xi, \dot{\mathsf{u}}(s)\rangle.$$

Hence, all the above estimates are equalities, and therefore $\mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s)) > 0$. Furthermore, (4.27) holds. Combining this with the fact that at almost all points the energy is differentiable with derivative $\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) = \mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s) - \langle \xi, \dot{\mathsf{u}}(s)\rangle$ in $L^1(\mathsf{a}, \mathsf{b})$, we conclude that (t, u) is admissible and (4.25) holds. □

Parameterized and* BV *solutions.

Proposition 4.7 (Equivalence between BV and parameterized solutions)**.**

(BVP1) *If* $(\mathsf{t}, \mathsf{u}) \in \mathscr{A}(\mathsf{a}, \mathsf{b}; [0, T] \times V)$ *is surjective and nondegenerate, then any curve*

$$u : [0, T] \to V \quad \textit{with} \quad u(t) \in \big\{\mathsf{u}(s) : \mathsf{t}(s) = t\big\} \tag{4.28}$$

belongs to $\mathrm{BV}([0, T]; D_E, \Psi)$ *for some* $E > 0$*, satisfies the local stability condition* ($\mathrm{S}_{\mathrm{loc}}$)*, and for every* $0 \leq t_0 < t_1 \leq T$ *with* G *defined as in* (4.11) *we have*

$$\mathrm{Var}_{\mathfrak{f}}(u; [t_0, t_1]) \leq \int_{\mathsf{s}(t_0)}^{\mathsf{s}(t_1)} \Psi[\mathsf{u}'](s)\,\mathrm{d}s + \int_{[\mathsf{s}(t_0), \mathsf{s}(t_1)] \cap G} \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\|\,\mathrm{d}s; \tag{4.29}$$

in particular $\mathrm{Var}_{\mathfrak{f}}(u; [0, T]) < \infty$.

(BVP2) *If* $(\mathsf{t}, \mathsf{u}) : [0, \mathsf{S}] \to [0, T] \times V$ *is a parameterized solution of the RIS* $(V, \mathcal{E}, \Psi, \Phi)$*, then any curve* $u : [0, T] \to V$ *satisfying* (4.28) *is a* BV *solution in the sense of Definition 3.8.*

(BVP3) *Conversely, if* $u \in \mathrm{BV}([0, T]; D_E, \Psi)$ *satisfies* ($\mathrm{S}_{\mathrm{loc}}$) *with* $\mathrm{Var}_{\mathfrak{f}}(u; [0, T]) < \infty$*, then there exists a nondegenerate, surjective* $(\mathsf{t}, \mathsf{u}) \in \mathscr{A}(0, \mathsf{S}; [0, T] \times V)$ *such that* (4.28) *holds and*

$$\mathrm{Var}_{\mathfrak{f}}(u; [0, T]) = \int_0^{\mathsf{S}} \Psi[\mathsf{u}'](s)\,\mathrm{d}s + \int_{[0, \mathsf{S}] \cap G} \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\|\,\mathrm{d}s. \tag{4.30}$$

Thus if u *is a* BV *solution of the RIS* $(V, \mathcal{E}, \Psi, \Phi)$ *then* (t, u) *is a parameterized solution.*

Proof. (BVP1): let $\mathsf{s} : [0, T] \to [\mathsf{a}, \mathsf{b}]$ be any inverse of t. Notice that $t \in \mathrm{J}_u$ if and only if $t \in \mathrm{J}_{\mathsf{s}}$ and $\mathsf{t}(s) \equiv t$ for every $s \in [\mathsf{s}(t_-), \mathsf{s}(t_+)]$. We can also define $\mathsf{s}(t)$ in $[\mathsf{s}(t_-), \mathsf{s}(t_+)]$ so that $u(t) = \mathsf{u}(\mathsf{s}(t))$ for every $t \in [0, T]$. By this choice it is immediate to see that $u \in \mathrm{BV}([0, T]; D_E, \Psi)$ with

$$\mathrm{Var}_\Psi(u; [t_0, t_1]) = \mathrm{Var}_\Psi(\mathsf{u}; [\mathsf{s}(t_0), \mathsf{s}(t_1)]) = \int_{\mathsf{s}(t_0)}^{\mathsf{s}(t_1)} \Psi[\mathsf{u}'](r)\mathrm{d}r \quad \text{for every } 0 \leq t_0 < t_1 \leq T.$$

On the other hand, the curve $\mathsf{u} : [\mathsf{s}(t_-), \mathsf{s}(t_+)] \to V$ is an admissible transition connecting $u(t_-)$ to $u(t_+)$ with

$$\Delta_{\mathfrak{f}_t}(u(t_-), u(t)) \le \int_{\mathsf{s}(t_-)}^{\mathsf{s}(t)} \mathfrak{f}_{\mathsf{s}(r)}[\mathsf{u}, \mathsf{u}'](r)\, \mathrm{d}r, \quad \Delta_{\mathfrak{f}_t}(u(t), u(t_+)) \le \int_{\mathsf{s}(t)}^{\mathsf{s}(t_+)} \mathfrak{f}_{\mathsf{s}(r)}[\mathsf{u}, \mathsf{u}'](r)\, \mathrm{d}r,$$

which yields (4.29). Since $\dot{\mathsf{t}} = 0$ in G, $\mathsf{t}(G)$ is $\mathscr{L}^1$-negligible, so that its complement (where the local stability condition ($\mathrm{S_{loc}}$) holds) is dense in $[0, T]$. Since $\mathfrak{e}$ is lower semicontinuous, every point in $[0, T] \setminus \mathrm{J}_u$ satisfies ($\mathrm{S_{loc}}$).

(BVP2) is now immediate: since ($\mathrm{S_{loc}}$) holds, it is sufficient to check ($\mathrm{E_{\mathfrak{f},ineq}}$); this follows by combining (4.29), (4.16), and the change of variable formula

$$\int_0^T \mathcal{P}_t(u(t))\, \mathrm{d}t = \int_0^{\mathsf{S}} \mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s)) \dot{\mathsf{t}}(s)\, \mathrm{d}s. \tag{4.31}$$

In order to prove (BVP3), we introduce the parameterization

$$\mathsf{s}(t) := t + \mathit{Var}_{\mathfrak{f}}(u; [0, t]), \quad \mathsf{S} := \mathsf{s}(T), \quad \mathrm{J}_u = \mathrm{J}_{\mathsf{s}} = (t_n)_{n \in \mathbb{N}}, \tag{4.32}$$

$$I_n := (\mathsf{s}(t_{n-}), \mathsf{s}(t_{n+})), \quad I := \bigcup_{n \in \mathbb{N}} I_n, \quad \mathsf{t} := \mathsf{s}^{-1} : [0, \mathsf{S}] \setminus I \to [0, T], \quad \mathsf{u} := u \circ \mathsf{t}. \tag{4.33}$$

It is immediate to check that t and u are Lipschitz maps. We extend t and u to I by setting

$$\mathsf{t}(s) \equiv t_n, \quad \mathsf{u}(s) := \vartheta_n(\mathsf{r}_n(s)) \quad \text{whenever } s \in I_n, \tag{4.34}$$

where $\mathsf{r}_n : \overline{I_n} \to [0, 1]$ is the unique affine and strictly increasing function mapping $\overline{I_n}$ onto $[0, 1]$ and $\vartheta_n \in \mathcal{T}_{t_n}(u(t_{n-}), u(t_{n+}))$ is an admissible transition satisfying $\vartheta_n(\mathsf{r}_n(\mathsf{s}(t_n))) = u(t_n)$ and (recall (F1) of Theorem 3.5)

$$\int_0^1 \mathfrak{f}_{t_n}[\vartheta_n; \vartheta_n'](r)\, \mathrm{d}r = \Delta_{\mathfrak{f}_{t_n}}(u(t_{n-}), u(t_n)) + \Delta_{\mathfrak{f}_{t_n}}(u(t_n), u(t_{n+})). \tag{4.35}$$

It follows that (4.28) holds with $u = \mathsf{u} \circ \mathsf{s}$ and

$$\begin{aligned}
&\int_0^{\mathsf{S}} \Psi[\mathsf{u}'](s)\, \mathrm{d}s + \int_G \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\, \|\dot{\mathsf{u}}(s)\|\, \mathrm{d}s = \mathrm{Var}_\Psi(\mathsf{u}; [0, S]) + \int_G \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\, \|\dot{\mathsf{u}}(s)\|\, \mathrm{d}s \\
&\quad = \mathrm{Var}_\Psi(u; [0, T]) + \sum_{n \in \mathbb{N}} \int_0^1 \mathfrak{e}_{t_n}(\vartheta_n(r)) \|\dot{\vartheta}_n(r)\|\, \mathrm{d}r \\
&\quad \le \mathrm{Var}_\Psi(u; [0, T]) - \mathrm{Jmp}_\Psi(u; [0, T]) + \mathrm{Jmp}_{\mathfrak{f}}(u; [0, T]) = \mathit{Var}_{\mathfrak{f}}(u; [0, T]),
\end{aligned}$$

so that (4.30) holds and $(\mathsf{t}, \mathsf{u}) \in \mathscr{A}(0, \mathsf{S}; [0, T] \times V)$.

If moreover u is a BV solution, then the chain rule from Theorem 4.4 and (4.31) yield inequality (4.24). $\square$

4.2. **V-parameterized solutions.** We consider now the special class of parameterizable solutions, corresponding to the notion introduced in §3.4, namely those for which u is absolutely continuous with values in V.

Definition 4.8. *A V-parameterized solution $(\mathsf{t}, \mathsf{u}) : [\mathsf{a}, \mathsf{b}] \to [0, T] \times V$ of the RIS $(V, \mathcal{E}, \Psi, \Phi)$ is a parameterized solution such that $\mathsf{u} \in \mathrm{AC}(\mathsf{a}, \mathsf{b}; V)$.*

Since V-parameterized solutions are differentiable $\mathscr{L}^1$-a.e., one does not have to distinguish the behavior of u in the set G of (4.11) from its complement. By adopting the "pointwise" definition (4.9) of $\mathfrak{F}$ and $\mathfrak{G}$ in place of (4.13), metric concepts are no longer needed, and expressions like (4.12) become simpler.

Proposition 4.9. *If $(\mathsf{t}, \mathsf{u}) \in \mathrm{AC}([0, \mathsf{S}]; [0, T] \times V)$ is a V-parameterized solution to the RIS $(V, \mathcal{E}, \Psi, \Phi)$ then every u satisfying (4.28) is a V-parameterizable* BV *solution. If u is a V-parameterizable* BV *solution there exists a V-parameterized solution (t, u) such that (4.28) hold.*

Proof. The approach is analogous to the proof of Proposition 4.7: In one direction it follows by the identity $\mathrm{Var}(u;[0,T]) = \int_0^{\mathsf{S}} \|\dot{\mathsf{u}}(s)\|\,\mathrm{d}s$. In the opposite one, we can simply replace (4.32) by

$$\mathsf{s}(t) := t + Var_{\mathfrak{f}}(u;[0,t]) + \mathrm{Var}(u;[0,t]), \tag{4.36}$$

choosing the optimal jump transitions according to (3.51). □

Thanks to Proposition 4.9, Corollary 3.23 implies the following result:

Corollary 4.10 (Existence of V-parameterized solutions). *If* (3.55)–(3.57) *hold, then for every* $u_0 \in D$ *with* $K^* + \partial\mathcal{E}_0(u_0) \ni 0$ *there exists a* V*-parameterized solution* $(\mathsf{t},\mathsf{u}) \in \mathrm{AC}([0,\mathsf{S}];[0,T]\times V)$ *of the RIS* $(V,\mathcal{E},\Psi,\Phi)$.

V-parameterized solutions can also be obtained as limit of rescaled solutions to (1.1) if they satisfy the uniform bound (3.52): one can simply adapt the argument discussed in §4.1, by replacing the definition (4.3) of the arclength s_ε with, e.g.,

$$\mathsf{s}_\varepsilon(t) := t + \int_0^t \mathfrak{f}_r(u_\varepsilon(r);\dot{u}_\varepsilon(r))\,\mathrm{d}r + \int_0^t \|\dot{u}_\varepsilon(r)\|\,\mathrm{d}r, \quad \mathsf{t}_\varepsilon := \mathsf{s}_\varepsilon^{-1}, \tag{4.37}$$

in order to gain a uniform control of the Lipschitz constant of the rescaled functions u_ε. The vanishing-viscosity limit in Theorem 4.3 then gives the following.

Theorem 4.11. *Let* $(u_\varepsilon)_{\varepsilon>0}$ *be a family of solutions to* (1.1) *satisfying* (3.26) *at* $t=0$ *and the uniform bound* (3.52) *(e.g. when the assumptions of Theorem 3.21 are satisfied) and let* $\mathsf{t}_\varepsilon : [0,\mathsf{S}] \to [0,T]$ *be nondecreasing and surjective time rescalings (e.g.* (4.37)*) such that* $\mathsf{u}_\varepsilon := u_\varepsilon \circ \mathsf{t}_\varepsilon$ *satisfy* (4.18) *and there exists* $C>0$ *such that* $\sup_{t\in(0,T)} \|\dot{\mathsf{u}}_\varepsilon(t)\| \le C$ *for all* $\varepsilon>0$. *Then any limit function* (t,u) *as in Theorem 4.3 is a* V*-parameterized solution.*

V-arclength parameterizations. Still keeping the assumptions (3.55)–(3.57) of Corollary 3.22, in particular the choice $\Phi(v) := \frac{1}{2}\|v\|^2$, we discuss now a different reparameterization technique for studying the limit of solutions to (1.1). Since estimate (3.52) is guaranteed, like in [EfM06, Mie11, MiZ12] we are entitled to use the V*-arclength* parameterization

$$\hat{\mathsf{s}}_\varepsilon(t) := t + \int_0^t \|\dot{u}_\varepsilon(r)\|\,\mathrm{d}r \tag{4.38}$$

and consider the rescaled functions $(\hat{\mathsf{t}}_\varepsilon,\hat{\mathsf{u}}_\varepsilon) : [0,\hat{\mathsf{S}}_\varepsilon] \to [0,T]\times V$, with $\hat{\mathsf{S}}_\varepsilon = \hat{\mathsf{s}}_\varepsilon(T)$, defined by $\hat{\mathsf{t}}_\varepsilon(s) := \hat{\mathsf{s}}_\varepsilon^{-1}(s)$ and $\hat{\mathsf{u}}_\varepsilon(s) := u_\varepsilon(\hat{\mathsf{t}}_\varepsilon(s))$. By construction we have $\dot{\hat{\mathsf{t}}}_\varepsilon(s) + \|\dot{\hat{\mathsf{u}}}_\varepsilon(s)\| = 1$ for a.a. $s \in (0,\hat{\mathsf{S}}_\varepsilon)$, and the pair $(\hat{\mathsf{t}}_\varepsilon,\hat{\mathsf{u}}_\varepsilon)$ is a solution of the "rescaled" doubly nonlinear equation

$$\partial\Psi(\dot{\hat{\mathsf{u}}}_\varepsilon(s)) + \frac{\varepsilon}{1-\|\dot{\hat{\mathsf{u}}}_\epsilon(s)\|}\partial\Phi(\dot{\hat{\mathsf{u}}}_\varepsilon(s)) + \partial\mathcal{E}_{\hat{\mathsf{t}}_\varepsilon(s)}(\hat{\mathsf{u}}_\varepsilon(s)) \ni 0 \quad \text{for a.a. } s \in (0,\hat{\mathsf{S}}_\varepsilon), \tag{4.39}$$

where we used the degree-1 homogeneity of $\partial\Phi$. As in [EfM06, MiZ12, Mie11], we observe that the viscous term in (4.39) is the subdifferential of the potential $\widehat{\Phi}$ that is defined via

$$\widehat{\Phi}(v) = f\left(\|v\|\right) \text{ with } f(x) = \begin{cases} -\log(1-x) - x & \text{if } 0 \le x < 1, \\ +\infty & \text{if } x \ge 1. \end{cases}$$

Thus, (4.39) rewrites as

$$\partial\Psi(\dot{\hat{\mathsf{u}}}_\varepsilon(s)) + \varepsilon\partial\hat{\Phi}(\dot{\hat{\mathsf{u}}}_\varepsilon(s)) + \partial\mathcal{E}_{\hat{\mathsf{t}}_\varepsilon(s)}(\hat{\mathsf{u}}_\varepsilon(s)) \ni 0 \quad \text{for a.a. } s \in (0,\hat{\mathsf{S}}_\varepsilon). \tag{4.40}$$

The sequence of dissipation potentials $\widehat{\Psi}_\varepsilon(v) := \Psi(v) + \varepsilon\widehat{\Phi}(v)$ Γ-converges monotonously, as $\varepsilon \downarrow 0$, to the limiting potential

$$\widehat{\Psi}(v) = \begin{cases} \Psi(v) & \text{if } \|v\| \le 1, \\ +\infty & \text{else.} \end{cases} \tag{4.41}$$

It was shown in [Mie11, Prop. 4.14] that, up to a subsequence, the parameterized solutions $(\hat{\mathsf{t}}_\varepsilon, \hat{\mathsf{u}}_\varepsilon)$ converge in $\mathrm{C}^0([0,\hat{\mathsf{S}}];[0,T]\times V)$ to a pair $(\hat{\mathsf{t}},\hat{\mathsf{u}}) \in \mathrm{C}^0_{\mathrm{lip}}([0,\hat{\mathsf{S}}];[0,T]\times V)$ such that $\hat{\mathsf{t}}(0)=0$, $\hat{\mathsf{t}}$ is non-decreasing, and

$$\dot{\hat{\mathsf{t}}}(s) + \|\dot{\hat{\mathsf{u}}}(s)\| \in [0,1] \quad \text{and} \quad \partial\widehat{\Psi}(\dot{\hat{\mathsf{u}}}(s)) + \partial\mathcal{E}_{\hat{\mathsf{t}}(s)}(\hat{\mathsf{u}}(s)) \ni 0 \quad \text{for a.a. } s\in(0,\hat{\mathsf{S}}). \tag{4.42}$$

An interesting feature of this approach is that it allows for a direct passage to the limit in the subdifferential inclusion (4.40), without passing through an energy identity like (4.6). By operating a suitable time rescaling, it is possible to show a correspondence between V-parameterized solutions in the sense of Definition 4.8 and in the sense of (4.42): the interested reader is referred to [Mie11, Cor. 4.22, Prop. 4.24].

However, let us stress that the technique from [EfM06, MiZ12] does not allow us to prove that the limit curve $(\hat{\mathsf{t}},\hat{\mathsf{u}})$ satisfies the normalization condition $\dot{\hat{\mathsf{t}}} + \|\dot{\hat{\mathsf{u}}}\| = 1$ a.e. in $(0,\hat{\mathsf{S}})$. Instead, our the variational approach of § 4.1, which is based on a chain-rule and energy-identity argument, guarantees the preservation of the normalization condition, cf. Theorem 4.3. Moreover, we also obtain the absolute continuity of the energy map $s\mapsto \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))$.

5. Examples

Throughout this section, we focus on the rate-independent system $(V,\mathcal{E},\Psi,\Phi)$ given by

$$V = L^2(\Omega), \quad \Psi(v) = \int_\Omega |v(x)|\,\mathrm{d}x, \quad \Phi(v) = \frac12\|v\|^2 = \frac12\int_\Omega |v(x)|^2\,\mathrm{d}x$$

with $\Omega\subset\mathbb{R}^d$, $d\geq 1$, a bounded Lipschitz domain, and on the following class of energy functionals $\mathcal{E}:[0,T]\times L^2(\Omega)\to(-\infty,+\infty]$

$$\mathcal{E}_t(u) = \begin{cases} \int_\Omega \big(\beta(|\nabla u|) + W(u) - \ell(t)u\big)\,\mathrm{d}x & \text{if } u\in \mathrm{W}^{1,1}(\Omega),\ \beta(|\nabla u|),\ W(u)\in L^1(\Omega),\\ +\infty & \text{otherwise.}\end{cases} \tag{5.1}$$

Hereafter, we suppose that

$$\beta:[0,+\infty)\to[0,+\infty) \text{ is convex;} \tag{5.2}$$

$$W:\mathbb{R}\to(-\infty,+\infty] \text{ is bounded from below;} \tag{5.3}$$

$$\ell\in \mathrm{C}^1([0,T];L^2(\Omega)). \tag{5.4}$$

In all of the examples we present, $\mathcal{E}$ will satisfy (E.0) and for each of them we will discuss the coercivity condition (E.1). Exploiting (5.4), it is immediate to check that for all $u\in D$ the function $t\mapsto\mathcal{E}_t(u)$ is differentiable, with derivative $\mathcal{P}_t(u) = -\int_\Omega \ell'(t)u\mathrm{d}x$ which fulfills both (E.2) and the Lipschitz estimate (3.57). In what follows, the focus will be on the uniform subdifferentiability (E.3) and on the (stronger) generalized convexity (3.61) (which yields the subdifferentiability condition (3.56) and in particular (E.3)).

We start with Example 5.1, where we provide sufficient conditions on the nonlinearities β and W guaranteeing the validity of (3.61).

Example 5.1. *We take*

$$\beta(|\nabla u|) = \frac12|\nabla u|^2 \quad \textit{and} \quad W\in\mathrm{C}^1(\mathbb{R}),\ \lambda\textit{-convex for some } \lambda\in\mathbb{R}; \tag{5.5}$$

for instance, one may think of the double-well potential $W(u) = (1-u^2)^2/4$. Clearly, $\mathcal{E}$ from (5.1) *fulfills* (E.1). *In order to check* (3.61), *we fix u, $v\in D$ and estimate, for $\theta\in[0,1]$,*

$$\mathcal{E}_t((1-\theta)u+\theta v) \leq (1-\theta)\mathcal{E}_t(u) + \theta\mathcal{E}_t(v) - \frac{(1-\theta)\theta}{2}\Big(\|\nabla(u-v)\|^2_{L^2(\Omega)} + \lambda\|u-v\|^2_{L^2(\Omega)}\Big), \tag{5.6}$$

where we used 1-convexity of β and λ-convexity of W. Hence, for $\lambda>0$ we have (3.61) *with $\alpha_E=\lambda$ and $\Lambda_E=0$. If $\lambda<0$, we use the Gagliardo-Nirenberg inequality*

$$\|w\|_{L^2(\Omega)} \leq C_{\mathrm{GN}}\big(\|w\|^{2/(d+2)}_{L^1(\Omega)}\|\nabla w\|^{d/(d+2)}_{L^2(\Omega)} + \|w\|_{L^1(\Omega)}\big) \leq \big(\tfrac{1}{1+|\lambda|}\|\nabla w\|^2_{L^2(\Omega)} + M_\lambda\|w\|^2_{L^1(\Omega)}\big)^{1/2},$$

for some $M_\lambda > 0$, which is equivalent to $-\|\nabla w\|^2_{L^2(\Omega)} \le -(1+|\lambda|)\|w\|^2_{L^2(\Omega)} + (1+|\lambda|)M_\lambda \|w\|^2_{L^1(\Omega)}$. Inserting this for $w = u-v$ into (5.6) we obtain estimate (3.61) with $\alpha_E = (1+|\lambda|) + \lambda = 1 > 0$ and $\Lambda_E = (1+|\lambda|)M_\lambda$. In particular, we have no dependence on the energy sublevel E.

In fact, it can be checked that for suitably convex functions β with the growth $\beta(|\nabla u|) \ge c_1|\nabla u|^p - c_2$ for some $c_1, c_2 > 0$ the related functional $\mathcal{E}$ in (5.1) still complies with (3.61), if $p > p_d$ for a suitable $p_d > 1$ depending on the dimension d.

Our next example treats the case in which β has only linear growth. Even taking a *convex* function W, the generalized convexity condition (3.61) is no longer guaranteed. Nonetheless, since the functional $u \mapsto \mathcal{E}_t(u)$ is convex, its Fréchet subdifferential reduces to the subdifferential in the sense of convex analysis, and (E.3) clearly holds. In this setting, we show that there exist BV solutions to the rate-independent system $(V, \mathcal{E}, \Psi, \Phi)$, which are *not V-parameterizable.*

Example 5.2. *We consider the one-dimensional domain $\Omega = (0, l)$ for some $l > 1$ and take*

$$\beta\left(\left|\tfrac{\mathrm{d}}{\mathrm{d}x}u\right|\right) = \delta\left|\tfrac{\mathrm{d}}{\mathrm{d}x}u\right| \quad \text{with } \delta > 0, \quad W(u) = \mathrm{I}_{[0,1]}(u) = \begin{cases} 0 & \text{if } u \in [0,1], \\ +\infty & \text{otherwise,} \end{cases} \tag{5.7}$$

and the external loading $\ell : [0, T] \times (0, l) \to \mathbb{R}$ with $\ell(t, x) = t+2-x$, where $0 < T \le l-1$. Observe that, thanks to the compactifying character of the total-variation contribution $\delta \int_0^l |\frac{\mathrm{d}}{\mathrm{d}x}u| \,\mathrm{d}x$, the energy $\mathcal{E}$ fulfills (E.1). We now show that the function

$$u(t, x) = \chi_{[0,a(t)]}(x) = \begin{cases} 1 & \text{for } x \in [0, a(t)], \\ 0 & \text{otherwise} \end{cases}$$

for some continuous and nondecreasing function $a : [0, T] \to [0, l]$, which will be specified later, is a BV solution to the RIS $(V, \mathcal{E}, \Psi, \Phi)$.

Concerning the energy balance $(\mathrm{E_f})$, we observe that, since $u \in \mathrm{C}^0([0, T]; L^2(0, l))$ there holds

$$\mathrm{Var}_{\mathfrak{f}}(u; [0, t]) = \mathrm{Var}_{\|\cdot\|_{L^1(0,l)}}(u; [0, t]) = a(t) - a(0) \quad \text{for all } t \in [0, 1],$$

where we also used that a is nondecreasing. Easy calculations give $\mathcal{E}_t(u(t)) = \delta - (t+2)a(t) + \frac{a^2(t)}{2}$ and $\mathcal{P}_t(u(t)) = -a(t)$, therefore $(\mathrm{E_f})$ yields the flow rule for the moving interface a:

$$\dot{a}(t)(a(t)-1-t) = 0 \quad \Rightarrow \quad a(t) = 1 + t \quad \text{for all } t \in [0, T].$$

Since $\mathcal{E}_t(\cdot)$ is convex, u fulfills the local stability $(\mathrm{S_{loc}})$ if and only if it complies with the global stability condition (S), which in the present setting reads

$$\begin{aligned} \delta - \tfrac{1}{2}(t+1)(3t+5) &= \mathcal{E}_t(u(t)) \le \mathcal{E}_t(v) + \|v-u(t)\|_{L^1(0,l)} \\ &= \int_0^l \left(\delta|\tfrac{\mathrm{d}}{\mathrm{d}x}v| + |v-\chi_{[0,t+1]}| - (t+2-x)v\right) \mathrm{d}x \\ &= \delta \int_0^l |\tfrac{\mathrm{d}}{\mathrm{d}x}v| \,\mathrm{d}x + t + 1 - \int_0^{t+1}(t+3-x)v \,\mathrm{d}x + \int_{t+1}^l (x-1-t)v \,\mathrm{d}x \end{aligned} \tag{5.8}$$

for all $v \in L^1(0, l)$ and $t \in [0, 1]$. With some calculations one can show that for all $\delta \in [0, 2]$ and $l \ge 4$ the function $u(t, x) = \chi_{[0,t+1]}(x)$ fulfills (5.8), hence it is a BV *solution. Indeed, u is a* BV *solution also in the case $\delta = 0$, in which $\mathcal{E}$ does not comply with (E.1) and our existence results Thms. 3.9 and 3.10 do not apply. Although $u \in \mathrm{C}^{\mathrm{lip}}([0, 1]; L^1(0, l))$, we have that $u \notin \mathrm{BV}([0, 1]; L^2(0, l))$, therefore it is not a V-parameterizable* BV *solution.*

We now revisit [Mie11, Ex. 4.4, 4.27], which means in our notation that $\beta \equiv 0$ and that W is of double-well type. Relying on the calculations from [Mie11], we show that as $\varepsilon \to 0$ the viscous solutions converge to a curve u, which is not a BV solution to the rate-independent system $(V, \mathcal{E}, \Psi, \Phi)$. Observe that in this case neither (E.1), nor the (parameterized) chain-rule inequality (4.22), are fulfilled.

Example 5.3. *We take $\Omega = (0, 1)$, $\beta \equiv 0$, $\ell(t, x) = t + x$, and*

$$W(u) = \begin{cases} \frac{1}{2}(u+4)^2 & \text{if } u \le -2, \\ 4 - \frac{1}{2}u^2 & \text{if } |u| < 2, \\ \frac{1}{2}(u-4)^2 & \text{if } u \ge 2, \end{cases} \tag{5.9}$$

In [Mie11, Ex. 4.4] *the unique solution to the viscous problem*

$$\mathrm{Sign}(\dot{u}_\varepsilon(t,x)) + \varepsilon\dot{u}_\varepsilon(t,x) + W'(u_\varepsilon(t,x)) \ni \ell(t,x) \quad \text{and} \quad u_\varepsilon(0,x) = -4 \tag{5.10}$$

was explicitly calculated: We have $u_\varepsilon(t,x) = V^\varepsilon(t{+}x)$, where $V^\varepsilon(\tau) = -4$ for $\tau \leq 1{+}\varepsilon$ and it coincides with the unique solution v of $\mathrm{Sign}(v'(\tau)) + \varepsilon v'(\tau) + W'(v(\tau)) \ni \tau$ *for $\tau \geq 1{+}\varepsilon$. It was shown that, on the time-interval $[0,6]$ the functions $(u_\varepsilon)_\varepsilon$ have a uniform Lipschitz bound with values in $L^1(0,1)$, whereas $\int_0^6 \|\dot{u}_\varepsilon\|_{L^2(0,1)}\,\mathrm{d}t$ tends to ∞ as $\varepsilon \to 0$ like $1/\sqrt{\varepsilon}$. Moreover, setting*

$$\bar{u}(t,x) = \max\{-4, t{+}x{-}5\} \text{ for } t{+}x \leq 3 \qquad \text{and } \bar{u}(t,x) = t{+}x{+}3 \text{ for } t{+}x > 3$$

we have $\bar{u} \in \mathrm{C}^0([0,6]; L^2(0,1)) \cap \mathrm{C}^{\mathrm{lip}}([0,6]; L^1(0,1))$ and $\sup_{t\in[0,6]} \|u_\varepsilon(t) - \bar{u}(t)\|_{L^2(0,1)} \to 0$ as $\varepsilon \to 0$, hence obviously $\mathcal{E}_t(u_\varepsilon(t)) \to \mathcal{E}_t(\bar{u}(t))$ for all $t \in [0,6]$.

It can be shown that $\bar{u}(t)$ complies with the local stability condition $(\mathrm{S}_{\mathrm{loc}})$ *for all $t \in [0,6]$. However, u does not comply with the energy balance* $(\mathrm{E}_{\mathfrak{f}})$*. In fact, by continuity of $\bar{u}$ we have $\mathrm{Var}_{\mathfrak{f}}(\bar{u}; [0,t]) = \mathrm{Var}_{\|\cdot\|_{L^1(0,l)}}(\bar{u}; [0,t])$ for all $t \in [0,6]$, and passing to the limit as $\varepsilon \to 0$ in the* viscous *energy balance* (2.35) *it can be calculated explicitly that for all $t \in [0,6]$*

$$\mathrm{Var}_{\|\cdot\|_{L^1}}(\bar{u}; [0,t]) + \mathcal{E}_t(\bar{u}(t)) - \mathcal{E}_t(\bar{u}(0)) - \int_0^t \mathcal{P}_s(\bar{u}(s))\,\mathrm{d}s = 8\max\{0, \min\{t{-}2, 1\}\} =: \rho(t). \tag{5.11}$$

Therefore, following [Mie11] *we observe that there is an additional limit dissipation ρ in* (5.11)*, and $\bar{u}$ is not a* BV *solution.*

In fact, the chain-rule inequality (4.22) *does not hold along the parameterized curve (cf. Definition 4.1) $(\mathsf{t},\mathsf{u}) \in \mathscr{A}(0,6; [0,6] \times L^2(0,1))$ given by $s \mapsto (\mathsf{t}(s), \mathsf{u}(s)) := (s, \bar{u}(s)) \in [0,6] \times L^2(0,1)$. On the one hand, since $\bar{u}$ satisfies* $(\mathrm{S}_{\mathrm{loc}})$ *on $[0,6]$, we have $\mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\|_{L^2(0,1)} \equiv 0$ on $[0,6]$. On the other hand,* (5.11) *yields for almost all $s \in (0,6)$*

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s) = -\dot{\rho}(\mathsf{t}(s)) - |\mathsf{u}'|_{L^1(0,1)}(s), \tag{5.12}$$

where $|\mathsf{u}'|_{L^1(0,1)}$ denotes the $L^1(0,1)$-metric derivative of u, cf. (2.13)*. Clearly, the right-hand side of* (5.12) *is strictly smaller than $|\mathsf{u}'|_{L^1(0,1)}(s)$ for $s \in (2,3)$.*

In the final example we recover the coercivity condition (E.1) by taking a nonzero β, with linear growth. Nonetheless, unlike Example 5.2 we only require W to be λ-convex: in this case, the chain-rule inequality (4.22) is still not valid.

Example 5.4. *We take $\Omega = (0,l)$ with $l > 2$, $\beta(|\frac{\mathrm{d}}{\mathrm{d}x}u|) = |\frac{\mathrm{d}}{\mathrm{d}x}u|$, the double-well potential W* (5.9)*, and $\ell(t,x) \equiv 2$ for all $(t,x) \in [0,T] \times (0,l)$, where $0 < T \leq l{-}2$. We show that the parameterized curve $s \in [0,T] \mapsto (\mathsf{t}(s), \mathsf{u}(s)) := (s, \bar{u}(s)) \in [0,T] \times L^2(0,l)$ with*

$$\bar{u}(t,x) := \begin{cases} 6 & \text{for } 0 \leq x \leq t+1, \\ -2 & \text{for } t+1 < x \leq l \end{cases} \qquad \text{for all } (t,x) \in [0,T] \times [0,l] \tag{5.13}$$

does not comply with the chain-rule inequality (4.22)*. Note that $\bar{u}$ satisfies $\bar{u} \in \mathrm{C}^0([0,T]; L^2(0,l)) \cap \mathrm{C}^{\mathrm{lip}}([0,T]; L^1(0,l))$ with $\|\bar{u}(t_1){-}\bar{u}(t_2)\|_{L^2(0,l)} = 8|t_1{-}t_2|^{1/2}$ and $\|\bar{u}(t_1){-}\bar{u}(t_2)\|_{L^1(0,l)} = 8|t_1{-}t_2|$. The latter implies $|\bar{u}'|_{L^1(0,l)} \equiv 8$.*

To see that the chain-rule inequality (4.22) *does not hold, we employ* (5.13) *to find*

$$\begin{aligned} \mathcal{E}_t(\bar{u}(t)) &= \mathcal{V}(\bar{u}(t)) + \int_0^l (W(\bar{u}(t,x)){-}2\bar{u}(t,x))\,\mathrm{d}x \\ &= 8 + \int_0^t (W(6){-}12)\,\mathrm{d}x + \int_t^l (W(-2){+}4)\,\mathrm{d}x \;=\; 8 + 6l - 16t, \end{aligned} \tag{5.14}$$

where we have used the notation $\mathcal{V}(u) := \int_0^l |\frac{\mathrm{d}}{\mathrm{d}x}u|\,\mathrm{d}x$ for the total variation functional on $(0,l)$. Next we show that $\bar{u}$ satisfies $(\mathrm{S}_{\mathrm{loc}})$*, i.e. $K^* + \partial\mathcal{E}_t(\bar{u}(t)) \ni 0$ for all $t \in [0,T]$. For this, we claim that*

$$\xi_t \in \partial\mathcal{E}_t(\bar{u}(t)) \qquad \text{with } \xi_t(x) = \begin{cases} \frac{1}{1+t} & \text{for } 0 < x < t{+}1, \\ \frac{-1}{l-1-t} & \text{for } t{+}1 < x < l. \end{cases} \tag{5.15}$$

To see this, we use $\mathcal{V}(\bar{u}(t)) = 8$ *and estimate, for general* $v \in \mathrm{BV}(0,l)$*, as follows:*

$$\mathcal{V}(v) - \mathcal{V}(\bar{u}(t)) \geq \operatorname*{ess\,sup}_{x\in(0,l)} v - \operatorname*{ess\,inf}_{x\in(0,l)} v \ - 8 \ \geq \ \tfrac{1}{1+t}\int_0^{1+t} v(x)\mathrm{d}x - \tfrac{1}{l-1-t}\int_{1+t}^l v(x)\mathrm{d}x \ - 8$$
$$= \int_0^l \xi_t(x)\big(v(x)-\bar{u}(t,x)\big)\mathrm{d}x \ = \ \langle \xi_t, v{-}\bar{u}(t)\rangle_{L^2(0,l)}.$$

Using the (-1)*-convexity of* W*, we obtain, for all* $v \in L^2(0,l)$*, the estimate*

$$\mathcal{E}_t(v) - \mathcal{E}_t(\bar{u}(t)) \geq \langle \xi_t, v{-}\bar{u}(t)\rangle - \tfrac{1}{2}\|v{-}\bar{u}(t)\|^2_{L^2(0,l)},$$

implying (5.15)*, cf. Definition* (2.8) *for Fréchet subdifferentials. Because of* $0 \leq t \leq T \leq l{-}2$ *we have* $\|\xi_t\|_{L^\infty} = \max\{\frac{1}{1+t}, \frac{1}{l-1-t}\} \leq 1$ *for all* $t \in [0,T]$*. Hence,* $\xi_t \in K^* = \{\ \xi \ : \ \|\xi\|_{L^\infty} \leq 1\ \}$*, and* $(\mathrm{S_{loc}})$ *is established.*

Now returning to the notation of the parameterized solution $(\mathsf{t}(s),\mathsf{u}(s)) = (s,\bar{u}(s))$ *for* $s \in [0,T]$*, we find* $\mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\|_{L^2(0,l)} \equiv 0$ *on* $[0,T]$*. Moreover,* $\mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s)) \equiv 0$ *as well, whereas* $|\mathsf{u}'|_{L^1(0,l)}(s) \equiv 8$*. Thus, on account of* (5.14) *we conclude that*

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s) = -16 \ \lneqq \ -8 = -|\mathsf{u}'|_{L^1(0,l)}(s) - \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\|_{L^2(0,l)}\,,$$

which is a contradiction to the chain-rule inequality (4.22).

6. Chain-rule inequalities for BV and parameterized curves

In this section we will collect the proof of the chain-rule inequalities stated in Theorems 3.11 and 4.4. We first consider the case of parameterized curves, hence, using the reparameterization technique of Proposition 4.7 we deduce Theorem 3.11.

6.1. **Chain rule for admissible parameterized curves: proof of Theorem 4.4.**

We split the proof in two claims.

Claim (1): *the map* $s \mapsto \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))$ *is absolutely continuous on* $[\mathsf{a},\mathsf{b}]$. First of all, we observe that, since $\sup_{s\in[\mathsf{a},\mathsf{b}]}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) =: E < \infty$, by (E.3) we have $\bar{\omega} := \sup_{r,s,\sigma}\omega_r^E(\mathsf{u}(s),\mathsf{u}(\sigma)) < \infty$. We decompose the open set G defined by (4.11) as the disjoint union of open intervals G_k. We fix $\mathsf{a} \leq r \leq s \leq \mathsf{b}$ and we consider the following cases:

- $r,s \in [0,T]\setminus G$. By (E.2) and estimate (2.9) there exists a constant $C>0$ (independent of r,s) such that

$$|\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(r)) - \mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(r))| \leq C\int_r^s \dot{\mathsf{t}}(\sigma)\mathrm{d}\sigma, \quad |\mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(s)) - \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))| \leq C\int_r^s \dot{\mathsf{t}}(\sigma)\mathrm{d}\sigma.$$

In view of (E.3), for $\xi(s) \in \partial\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))$ fulfilling $\xi(s) \in K^*$ we have

$$\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(r)) \leq \langle \xi(s), \mathsf{u}(s){-}\mathsf{u}(r)\rangle + \bar{\omega}\Psi_\wedge(\mathsf{u}(s){-}\mathsf{u}(r))$$
$$\leq \Psi(\mathsf{u}(s){-}\mathsf{u}(r)) + \bar{\omega}\Psi(\mathsf{u}(s){-}\mathsf{u}(r)) \leq (1+\bar{\omega})\int_r^s \Psi[\mathsf{u}'](\sigma)\mathrm{d}\sigma,$$

where the second inequality follows from (2.3) and the last one from (2.12) and the minimal representation $m = \Psi[\mathsf{u}']$. Analogously, arguing with $\xi(r) \in \partial\mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(r)) \cap K^*$, we have $\mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(r)) - \mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(s)) \leq (1+\bar{\omega})\int_r^s \Psi[\mathsf{u}'](\sigma)\mathrm{d}\sigma$. All in all, we conclude

$$|\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(r))| \leq C_1\int_r^s \Big(\dot{\mathsf{t}}(\sigma) + \Psi[\mathsf{u}'](\sigma)\Big)\mathrm{d}\sigma, \quad C_1 := 2(C+1+\bar{\omega}). \tag{6.1}$$

- r,s belong to the closure $\overline{G_k}$ of the same connected component $G_k = (a_k,b_k)$ for some k. It is not restrictive to assume that $r,s \in G_k$. Then $\mathsf{t} \equiv \bar{\mathsf{t}}$ is constant in G_k by (2) in Definition 4.1 and $\mathsf{u} \in \mathrm{AC}([r,s];V)$. We denote by $\partial^\circ\mathcal{E} : [0,T]\times D \rightrightarrows V^*$ the multivalued map defined by

$$\xi \in \partial^\circ\mathcal{E}_t(u) \text{ if and only if } \|\xi\|_* = \min\{\|\zeta\|_* \ : \ \zeta \in \partial\mathcal{E}_t(u)\},$$

with the usual convention that the latter quantity is $+\infty$ if $\partial\mathcal{E}_t(u)$ is empty. Since K^* is bounded in V^*, the definition of $\mathfrak{e}_{\bar t}(u)$ in (3.5) gives the estimate

$$\mathfrak{e}_{\bar t}(\mathsf{u}(\theta)) \geq \|\partial^\circ \mathcal{E}_{\bar t}(\mathsf{u}(\theta))\|_* - \mathsf{K}, \quad \text{where } \mathsf{K} := \sup\{\|z\| \,:\, z \in K^*\},$$

and we conclude that $\int_r^s \|\partial^\circ \mathcal{E}_{\bar t}(\mathsf{u}(\theta))\|\,\|\dot{\mathsf{u}}(\theta)\|\,\mathrm{d}\theta < \infty$. Hence the chain rule (analogous to Theorem 2.3, see the arguments of [AGS08, Theorem 1.2.5] and [MRS13, Proposition 2.4]) provides the absolute continuity the energy map in G_k and for $\mathscr{L}^1$-a.a. $\theta \in G_k$ we have

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\mathcal{E}_{\bar t}(\mathsf{u}(\theta)) = \langle \xi, \dot{\mathsf{u}}(\theta)\rangle \quad \text{for every } \xi \in \partial\mathcal{E}_{\bar t}(\mathsf{u}(\theta)), \tag{6.2}$$

$$\Big|\frac{\mathrm{d}}{\mathrm{d}\theta}\mathcal{E}_{\bar t}(\mathsf{u}(\theta))\Big| \leq \Psi(\dot{\mathsf{u}}(\theta)) + \mathfrak{e}_{\bar t}(\mathsf{u}(\theta))\|\dot{\mathsf{u}}(\theta)\|. \tag{6.3}$$

• $r \in G, s \in [0,T]$ with $r < s$ (or viceversa): we denote by σ the right boundary point of the interval $G_k \ni r$; combining (6.1) with the integrated form of (6.3) we obtain

$$\begin{aligned}
&\big|\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(r))\big| \leq \big|\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{E}_{\mathsf{t}(\sigma)}(\mathsf{u}(\sigma))\big| + \big|\mathcal{E}_{\mathsf{t}(\sigma)}(\mathsf{u}(\sigma)) - \mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(r))\big| \\
&\quad \leq C_P \int_\sigma^s \Big(\dot{\mathsf{t}}(\rho) + \Psi[\mathsf{u}'](\rho)\Big)\mathrm{d}\rho + \int_r^\sigma \Big(\Psi(\dot{\mathsf{u}}(\rho)) + \mathfrak{e}_{\mathsf{t}(\rho)}(\mathsf{u}(\rho))\|\dot{\mathsf{u}}(\rho)\|\Big)\,\mathrm{d}\rho \\
&\quad = \int_r^s h(\rho)\,\mathrm{d}\rho \quad \text{with } h \in L^1(0,T).
\end{aligned}$$

Claim (2): *the chain-rule inequality* (4.22) *holds.* It follows from Claim (1) there exists a set of full measure $\mathcal{T} \subset (a,b)$ such that for all $s \in \mathcal{T}$ the function t is differentiable at s, the first of (E.2) holds at s, the Ψ-metric derivative $\Psi[\mathsf{u}'](s)$ exists, and, if $s \in G$, the map u is V-differentiable at s. Hence, we evaluate the derivative of the map $\mathcal{E}_{\mathsf{t}(\cdot)}(\mathsf{u}(\cdot))$ at $s \in \mathcal{T}$: if $s \in \cup_k \overline{G_k}$ we immediately get the thesis by (6.3) (notice that $\mathscr{L}^1\big((\cup_k\overline{G_k})\setminus G\big) = 0$). If $s \in [0,T]\setminus \cup_k\overline{G_k}$ then $r = s - h \in [0,T]\setminus G$ for infinitely main values of $h > 0$, accumulating at 0. Since $\mathfrak{e}_{\mathsf{t}(r)}(\mathsf{u}(r)) = 0$ we can choose $\xi(r) \in -\partial\mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(r)) \cap K^*$ and thanks to (E.3) we have

$$\begin{aligned}
&\frac{\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(r))}{h} = \frac{\big(\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(s))\big) + \big(\mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(s)) - \mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(r))\big)}{h} \\
&\geq \langle \xi(r), \frac{1}{h}(\mathsf{u}(s) - \mathsf{u}(r))\rangle - \frac{1}{h}\omega_r(\mathsf{u}(s),\mathsf{u}(r))\Psi_\wedge(\mathsf{u}(s) - \mathsf{u}(r)) + \frac{\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{E}_{\mathsf{t}(r)}(\mathsf{u}(s))}{h} \\
&\geq -\frac{1 + \omega_r(\mathsf{u}(s),\mathsf{u}(r))}{h}\Psi(\mathsf{u}(s) - \mathsf{u}(r)) + \frac{1}{h}\int_r^s \mathcal{P}_{\mathsf{t}(\theta)}(\mathsf{u}(s))\dot{\mathsf{t}}(\theta)\,\mathrm{d}\theta
\end{aligned} \tag{6.4}$$

In the limit $r \uparrow s$, with $r \in [0,T]\setminus G$, we get the lower bound $\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s) \geq -\Psi[\mathsf{u}'](s)$. The corresponding upper bound can be obtained by choosing $r = s + h$, $h > 0$, in (6.4), and passing to the limit as $r \downarrow s$.

Whenever u is differentiable $\mathscr{L}^1$-a.e., the chain rule (4.23) follows from (6.2) and (6.4) by a similar argument. Hence, Theorem 4.4 is proved. □

By applying Theorem 4.4 to the parameterized curve $[0,1] \ni r \mapsto (t, \vartheta(r))$ associated with any admissible transition $\vartheta \in \mathcal{T}_t(u_0,u_1)$ we immediately have the desired jump estimates.

Corollary 6.1. *The jump estimates* (3.14) *and* (3.37) *hold true.*

6.2. **Chain rule for** BV **curves: proof of Theorem 3.11.** It is clearly not restrictive to assume $t_0 = 0$, $t_1 = T$. If $u \in \mathrm{BV}([0,T]; D_E, \Psi)$ satisfies the local stability condition and $\mathrm{Var}_{\mathfrak{f}}(u;[0,T]) < \infty$ as in the statement of the Theorem, we apply assertion (BVP3) of Proposition 4.7: the chain-rule inequality (3.34) follows then by the parameterized chain rule (4.22), combined with (4.30) and (4.31).

Let us now check (3.35) in the case $u \in \mathrm{BV}([0,T];V)$. We will use the simpler change of variable formula

$$\mathsf{s}(t) := t + \mathrm{Var}(u;[0,t]), \quad \mathsf{S} := \mathsf{s}(T), \tag{6.5}$$

keeping the same notation as in (4.33) for t, u, I_n, and I. We will use two basic facts: the first property concerns the diffuse part s'_d of the distributional derivative of s and has been proved in [MRS12a, Prop. 6.11] (the proof does not rely on the finite-dimensional setting therein considered), namely

$$u'_\mathrm{d} = \boldsymbol{n}\|u'_\mathrm{d}\| = (\dot{\mathsf{u}}\circ\mathsf{s})\,\mathsf{s}'_\mathrm{d}, \quad \mathscr{L}^1_{(0,T)} = (\dot{\mathsf{t}}\circ\mathsf{s})\,\mathsf{s}'_\mathrm{d}. \tag{6.6}$$

The second fact is a general property of the distributional derivative of an increasing map, viz.

$$\mathsf{t}_\sharp\big(\mathscr{L}^1_{[0,\mathsf{S}]}\big) = \mathsf{s}'_\mathrm{d}. \tag{6.7}$$

We set

$$\mathsf{e}(s) := \begin{cases} e(\mathsf{t}(s)) = \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) & \text{if } s \in (0,\mathsf{S})\setminus I, \\ \text{affine interpolation of } e(t_{n-}), e(t_{n+}) & \text{if } s \in I_n \text{ for some } n \in \mathbb{N}, \end{cases}$$

and we extend in a similar way s in each interval I_n. Now u defined by (4.33) is absolutely continuous and arguing as in § 6.1 we can easily prove that e is absolutely continuous with derivative

$$\dot{\mathsf{e}}(s) = -\langle\xi(\mathsf{t}(s)),\dot{\mathsf{u}}(s)\rangle + \mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s) \quad \text{for } \mathscr{L}^1\text{-a.a. } s \in (0,\mathsf{S}). \tag{6.8}$$

On the other hand $\mathsf{e}(s) = e(\mathsf{t}(s))$ whenever $s \in [0,\mathsf{S}]\setminus\bigcup\overline{I_n}$. Since $\mathsf{t}(s) \equiv t_n$ and $\dot{\mathsf{t}}(s) \equiv 0$ in I_n we obtain $e(\mathsf{t}(s))\dot{\mathsf{t}}(s) = \mathsf{e}(s)\dot{\mathsf{t}}(s)$ for a.a. $s \in (0,\mathsf{S})$. Hence, for every $\zeta \in \mathrm{C}^1([0,T])$ with compact support in $(0,T)$ we obtain

$$\begin{aligned}
\int_{[0,T]}\zeta(t)\,\mathrm{d}e'_\mathrm{d}(t) &= -\int_0^T\dot\zeta(t)e(t)\,\mathrm{d}t - \sum_{t\in\mathrm{J}(u)}\zeta(t)(e(t_+)-e(t_-)) \\
&= -\int_0^{\mathsf{S}}\dot\zeta(\mathsf{t}(s))e(\mathsf{t}(s))\dot{\mathsf{t}}(s)\,\mathrm{d}s - \sum_n\zeta(t_n)(e(t_{n+})-e(t_{n-})) \\
&= -\int_0^{\mathsf{S}}\dot\zeta(\mathsf{t}(s))\mathsf{e}(s)\dot{\mathsf{t}}(s)\,\mathrm{d}s - \sum_n\zeta(t_n)(\mathsf{e}(\mathsf{s}(t_{n+}))-e(\mathsf{s}(t_{n-}))) \\
&= \int_0^{\mathsf{S}}\zeta(\mathsf{t}(s))\dot{\mathsf{e}}(s)\,\mathrm{d}s - \sum_n\int_{I_n}\zeta(\mathsf{t}(s))\dot{\mathsf{e}}(s)\,\mathrm{d}s = \int_{[0,\mathsf{S}]\setminus I}\zeta(\mathsf{t}(s))\dot{\mathsf{e}}(s)\,\mathrm{d}s \\
&\overset{(6.8)}{=} -\int_{[0,\mathsf{S}]\setminus I}\zeta(\mathsf{t}(s))\langle\xi(\mathsf{t}(s)),\dot{\mathsf{u}}(s)\rangle\mathrm{d}s + \int_{[0,\mathsf{S}]\setminus I}\zeta(\mathsf{t}(s))\mathcal{P}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s)\,\mathrm{d}s \\
&\overset{(6.7)}{=} \int_{[0,T]\setminus\mathrm{J}_u}\zeta(t)\Big(-\langle\xi(t),\dot{\mathsf{u}}(\mathsf{s}(t))\rangle + \mathcal{P}_t(u(t))\dot{\mathsf{t}}(\mathsf{s}(t))\Big)\,\mathrm{d}\mathsf{s}'_\mathrm{d}(t) \\
&\overset{(6.6)}{=} -\int_{[0,T]\setminus\mathrm{J}_u}\zeta(t)\langle\xi(t),\boldsymbol{n}\rangle\,\mathrm{d}\|u'_\mathrm{d}\|(t) + \int_0^T\zeta(t)\mathcal{P}_t(u(t))\,\mathrm{d}t.
\end{aligned}$$

Since the measure $\|u'_\mathrm{d}\|$ does not charge J_u, we get (3.35), and Theorem 3.11 is proved. □

7. Convergence proofs for the viscosity approximations

7.1. **Compactness and lower semicontinuity result for parameterized curves.** We first provide a lower semicontinuity result that will be used to prove Theorems 3.5, 3.9, and 3.10 in the next subsections.

Proposition 7.1. *Let $E, L > 0$ and for every $n \in \mathbb{N}$ let $\mathsf{t}_n \in \mathrm{AC}(\mathsf{a},\mathsf{b};[0,T])$ be nondecreasing. Assume that $\tilde{\mathsf{u}}_n : [\mathsf{a},\mathsf{b}] \to D_E$ are measurable, $G_n \subset [\mathsf{a},\mathsf{b}]$ are open (and possibly empty) subsets such that $\mathfrak{e}_{\mathsf{t}_n(s)}(\tilde{\mathsf{u}}_n(s)) = 0$ in $[\mathsf{a},\mathsf{b}] \setminus G_n$, $\mathsf{u}_n \in \mathrm{AC}([\mathsf{a},\mathsf{b}]; V, \Psi) \cap \mathrm{AC}_{\mathrm{loc}}(G_n; V)$, and there holds*

$$X_n := \sup_{s\in[\mathsf{a},\mathsf{b}]} \|\mathsf{u}_n(s) - \tilde{\mathsf{u}}_n(s)\| \to 0 \quad \text{as } n \to \infty \tag{7.1a}$$

$$\dot{\mathsf{t}}_n(s) + \Psi[\mathsf{u}_n'](s) + \mathfrak{e}_{\mathsf{t}_n(s)}(\tilde{\mathsf{u}}_n(s))\|\dot{\mathsf{u}}_n(s)\| \le L \quad \text{for } \mathscr{L}^1\text{-a.a. } s \in (\mathsf{a},\mathsf{b}), \tag{7.1b}$$

where we adopt the convention $\mathfrak{e}_{\mathsf{t}_n(s)}(\tilde{\mathsf{u}}_n(s))\|\dot{\mathsf{u}}_n(s)\| \equiv 0$ if $s \notin G_n$, as in (4.14).

Then there exist a subsequence (not relabeled) and a limit function $(\mathsf{t},\mathsf{u}) \in \mathscr{A}(\mathsf{a},\mathsf{b};[0,T] \times D_E)$ such that $(\mathsf{t}_n,\mathsf{u}_n) \to (\mathsf{t},\mathsf{u})$ uniformly in $[\mathsf{a},\mathsf{b}]$ with respect to the topology of $[0,T] \times V$. Moreover (t,u) satisfies the same bound (7.1b) and the following asymptotic properties hold as $n \to \infty$:

$$\liminf_{n\to\infty} \int_{\mathsf{a}}^{\mathsf{b}} \Psi[\mathsf{u}_n'](s)\,\mathrm{d}s \ge \int_{\mathsf{a}}^{\mathsf{b}} \Psi[\mathsf{u}'](s)\,\mathrm{d}s, \tag{7.2}$$

$$\liminf_{n\to\infty} \int_{\mathsf{a}}^{\mathsf{b}} \mathfrak{e}_{\mathsf{t}_n(s)}(\tilde{\mathsf{u}}_n(s))\|\dot{\mathsf{u}}_n(s)\|\,\mathrm{d}s \ge \int_{\mathsf{a}}^{\mathsf{b}} \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\|\,\mathrm{d}s, \tag{7.3}$$

$$\liminf_{n\to\infty} \int_{\mathsf{a}}^{\mathsf{b}} \Big(\mathfrak{k}_{\mathsf{t}_n(s)}(\tilde{\mathsf{u}}_n(s))\dot{\mathsf{t}}_n(s) + \mathfrak{e}_{\mathsf{t}_n(s)}(\tilde{\mathsf{u}}_n(s))\|\dot{\mathsf{u}}_n(s)\|\Big)\,\mathrm{d}s \ge \int_{\mathsf{a}}^{\mathsf{b}} \mathfrak{G}[\mathsf{t},\mathsf{u};\dot{\mathsf{t}},\dot{\mathsf{u}}](s)\,\mathrm{d}s. \tag{7.4}$$

If, moreover, $\mathsf{u}_n \in \mathrm{AC}([\mathsf{a},\mathsf{b}]; V)$, then

$$\liminf_{n\to\infty} \int_{G_n} \mathfrak{G}_{\varepsilon_n}(\mathsf{t}_n(s),\tilde{\mathsf{u}}_n(s);\dot{\mathsf{t}}_n(s),\dot{\mathsf{u}}_n(s))\,\mathrm{d}s \ge \int_{\mathsf{a}}^{\mathsf{b}} \mathfrak{G}[\mathsf{t},\mathsf{u};\dot{\mathsf{t}},\dot{\mathsf{u}}](s)\,\mathrm{d}s \tag{7.5}$$

for every vanishing sequence $(\varepsilon_n)_n \subset (0,\infty)$.

We will use later that the assumptions of Proposition 7.1 cover the case $(\mathsf{t}_n,\mathsf{u}_n) \in \mathscr{A}(\mathsf{a},\mathsf{b};[0,T] \times V)$ with $\tilde{\mathsf{u}}_n = \mathsf{u}_n$.

Proof. By (7.1b) the sequence t_n is uniformly Lipschitz, thus relatively compact with respect to uniform convergence.

Let C_Ψ be the continuity constant of Ψ and $\Omega := \Omega_{D_E}$ be the modulus of continuity from (2.17): since Ω is concave and $\Omega(0) = 0$ we have

$$\Omega(\lambda p) \le \lambda\Omega(p), \quad \Omega(p+q) \le \Omega(p) + \Omega(q) \quad \forall\ \lambda, p, q \ge 0. \tag{7.6}$$

Since every curve $\tilde{\mathsf{u}}_n$ takes values in the compact set D_E, we have in view of (2.18) that

$$\begin{aligned}\|\tilde{\mathsf{u}}_n(s) - \tilde{\mathsf{u}}_n(r)\| &\le \Omega\big(\Psi_\wedge(\tilde{\mathsf{u}}_n(s) - \tilde{\mathsf{u}}_n(r))\big) \le \Omega\big(\Psi(\mathsf{u}_n(s) - \mathsf{u}_n(r))\big) + 2C_\Psi\Omega(X_n)\\ &\le L\,\Omega(|s-r|) + 2C_\Psi\Omega(X_n)\end{aligned} \tag{7.7}$$

It follows from (7.1a) that

$$\limsup_{n\to\infty} \|\tilde{\mathsf{u}}_n(s) - \tilde{\mathsf{u}}_n(r)\| \le L\,\Omega(|s-r|).$$

Thus $\tilde{\mathsf{u}}_n$ is (asymptotically) uniformly equicontinuous and we can apply the Arzelà-Ascoli Theorem (in a slightly refined form, see e.g. [AGS08, Prop. 3.3.1]) to prove its uniform convergence to a limit u. Passing to the limit in (7.1b) we get an analogous estimate for (t,u).

Statement (7.2) is an immediate consequence of the lower semicontinuity of the Ψ-total variation and of its representation formula (2.22).

In order to prove (7.3) let us observe that the lower semicontinuity property of the map $\mathfrak{e}$ and the above uniform convergence guarantee that the limit function $s \mapsto \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))$ is lower semicontinuous. Thanks to (7.1a) we can find a set $\mathsf{M} \subset [\mathsf{a},\mathsf{b}]$ with $\mathscr{L}^1([\mathsf{a},\mathsf{b}] \setminus \mathsf{M}) = 0$ such that $\tilde{\mathsf{u}}_n$ converges uniformly to u in M and

$$\forall \eta > 0\ \exists \bar{n} \in \mathbb{N} : \quad \mathfrak{e}_{\mathsf{t}_n(s)}(\tilde{\mathsf{u}}_n(s)) \ge \mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s)) - \eta \quad \text{for every } n \ge \bar{n},\ s \in \mathsf{M}. \tag{7.8}$$

If G is defined as in (4.11) and $[\alpha,\beta]\subset G$, (7.8) implies that there exists a positive constant $c>0$ with $\mathfrak{e}_{\mathsf{t}_n(s)}(\tilde{\mathsf{u}}_n(s))\ge c$ for $\mathscr{L}^1$-a.a. $s\in(\alpha,\beta)$ and n sufficiently big. Estimate (7.1b) then yields that u_n are uniformly V-Lipschitz in $[\alpha,\beta]$ so that u is also Lipschitz, and therefore $\mathscr{L}^1$-a.e. differentiable. Since $[\alpha,\beta]$ is arbitrary, we conclude that u is locally absolutely continuous in G, and the following Lemma 7.2 yields the lim inf inequality (7.3).

Recalling definitions (4.9) and (4.10) for $\mathfrak{G}$ and $\mathfrak{k}$, assertion (7.4) follows if we check that

$$\liminf_{n\to\infty}\int_{\mathsf{a}}^{\mathsf{b}}\mathfrak{k}_{\mathsf{t}_n(s)}(\tilde{\mathsf{u}}_n(s))\dot{\mathsf{t}}_n(s)\,\mathrm{d}s\ge\int_{\mathsf{a}}^{\mathsf{b}}\mathfrak{k}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s)\,\mathrm{d}s,$$

which is again a consequence of Lemma 7.2 ahead.

In order to prove (7.5) let us observe that $\mathfrak{G}_\varepsilon(\mathsf{t},\tilde{\mathsf{u}};\alpha,\mathsf{v})\ge\max\big\{\frac{\alpha}{\varepsilon}F^*(\mathfrak{e}_{\mathsf{t}}(\tilde{\mathsf{u}})),\ \mathfrak{e}_{\mathsf{t}}(\tilde{\mathsf{u}})\|\mathsf{v}\|\big\}$. Splitting the integration domain into $(\mathsf{a},\mathsf{b})\setminus G$ and G a further application of Lemma 7.2 yields

$$\begin{aligned}\liminf_{n\to\infty}&\int_{\mathsf{a}}^{\mathsf{b}}\mathfrak{G}_{\varepsilon_n}(\mathsf{t}_n(s),\tilde{\mathsf{u}}_n(s);\dot{\mathsf{t}}_n(s),\dot{\mathsf{u}}_n(s))\,\mathrm{d}s\\&\ge\liminf_{n\to\infty}\int_{(\mathsf{a},\mathsf{b})\setminus G}\frac{1}{\varepsilon_n}F^*(\mathfrak{e}_{\mathsf{t}_n(s)}(\tilde{\mathsf{u}}_n(s)))\dot{\mathsf{t}}_n(s)\,\mathrm{d}s+\liminf_{n\to\infty}\int_G\mathfrak{e}_{\mathsf{t}_n(s)}(\tilde{\mathsf{u}}_n(s))\|\dot{\mathsf{u}}_n(s)\|\,\mathrm{d}s\\&\ge\int_{(\mathsf{a},\mathsf{b})\setminus G}\mathfrak{k}_{\mathsf{t}(s)}(\mathsf{u}(s))\,\dot{\mathsf{t}}(s)\,\mathrm{d}s+\int_G\mathfrak{e}_{\mathsf{t}(s)}(\mathsf{u}(s))\,\|\dot{\mathsf{u}}(s)\|\,\mathrm{d}s=\int_{\mathsf{a}}^{\mathsf{b}}\mathfrak{G}[\mathsf{t},\mathsf{u};\dot{\mathsf{t}},\dot{\mathsf{u}}](s)\,\mathrm{d}s\,.\end{aligned}$$

This concludes the proof of Proposition 7.1. □

A simple proof of the following lemma can be found, e.g., in [MRS12b, Lem. 4.3].

Lemma 7.2. *Let I be a measurable subset of $\mathbb{R}$ and let $h_n,h,m_n,m:I\to[0,+\infty]$ be measurable functions for $n\in\mathbb{N}$ that satisfy*

$$\liminf_{n\to\infty}h_n(x)\ge h(x)\quad\textit{for }\mathscr{L}^1\textit{-a.a. }x\in I,\quad m_n\rightharpoonup m\quad\textit{in }L^1(I).\tag{7.9}$$

Then

$$\liminf_{n\to\infty}\int_I h_n(x)m_n(x)\,\mathrm{d}x\ge\int_I h(x)m(x)\,\mathrm{d}x.\tag{7.10}$$

7.2. **Compactness and lower semicontinuity for non-parameterized curves.**

Proof of Theorem 3.5. To address assertion (F2) let $\vartheta_n\in\mathcal{T}_t(u_{0,n},u_{1,n})$ be a sequence of admissible transitions such that

$$\int_0^1\mathfrak{f}_t[\vartheta_n;\vartheta_n'](r)\,\mathrm{d}r\le\Delta_{\mathfrak{f}_t}(u_{0,n},u_{1,n})+\varepsilon_n\quad\text{with }\varepsilon_n\ge0\text{ and }\lim_{n\to\infty}\varepsilon_n=\varepsilon\ge0.\tag{7.11}$$

By operating the change of variable

$$\mathsf{s}_n(r):=c_n\Big(r+\int_0^r\mathfrak{f}_t[\vartheta_n;\vartheta_n'](w)\,\mathrm{d}w\Big),\quad\mathsf{r}_n:=\mathsf{s}_n^{-1}:[0,\mathsf{S}]\to[0,1],\quad\mathsf{u}_n:=\vartheta_n\circ\mathsf{r}_n:[0,\mathsf{S}]\to V,$$

where c_n is a normalization so that $\mathsf{S}:=\mathsf{s}_n(1)$ is independent of n, we see that the functions r_n are uniformly Lipschitz and the curve $s\mapsto(\mathsf{r}_n(s),\mathsf{u}_n(s))$ satisfies (7.1a)–(7.1b) with $\tilde{\mathsf{u}}_n\equiv\mathsf{u}_n$.

We can thus extract subsequences (still denoted by $\mathsf{r}_n,\mathsf{u}_n$) converging uniformly to r,u respectively. The previous Proposition 7.1 guarantees that u is an admissible transition connecting u_- to u_+ and lim inf inequalities (7.2) and (7.3) show that

$$\varepsilon+\Delta_{\mathfrak{f}_t}(u_-,u_+)\ge\liminf_{n\to\infty}\int_0^1\mathfrak{f}_t[\vartheta_n;\vartheta_n'](r)\,\mathrm{d}r\ge\int_0^{\mathsf{S}}\mathfrak{f}_t[\mathsf{u};\mathsf{u}'](r)\,\mathrm{d}r\ge\Delta_{\mathfrak{f}_t}(u_-,u_+).$$

This proves the lower semicontinuity of the Finsler cost functional. Since we may choose $0<\varepsilon_n\to\varepsilon=0$, the previous inequalities yield that u attains the infimum in (3.11), so that also assertion (F1) is proved, since the jump estimate (3.14) has been proved in Corollary 6.1.

Let us now consider the last assertion (F3): it is not restrictive to assume $u_- \neq u_+$ so that $\Delta \geq \Psi(u_+ - u_-) > 0$. For $r \in [0, \beta_n - \alpha_n]$ we set

$$\mathsf{s}_n(r) := c_n\Big(r + \int_{\alpha_n}^{\alpha_n + r} \Psi_{\varepsilon_n}(u_n(\zeta)) + \Psi^*_{\varepsilon_n}(\xi_n(\zeta))\,\mathrm{d}\zeta\Big),$$
$$\mathsf{t}_n := \mathsf{s}_n^{-1} : [0,1] \to [\alpha_n, \beta_n], \quad \mathsf{u}_n := u_n \circ \mathsf{t}_n, \quad \tilde{\mathsf{u}}_n := \tilde{u}_n \circ \mathsf{t}_n : [0,1] \to V,$$

where c_n is a normalization constant such that $\mathsf{s}_n(\beta_n - \alpha_n) = 1$. Again, it is not difficult to see that the triple $(\mathsf{t}_n, \mathsf{u}_n, \tilde{\mathsf{u}}_n)$ satisfies the assumptions of Proposition 7.1. Moreover

$$\int_{\alpha_n}^{\beta_n} \Big(\Psi_{\varepsilon_n}(u_n(r)) + \Psi^*_{\varepsilon_n}(\xi_n(r))\Big)\,\mathrm{d}r = \int_0^{\mathsf{S}} \Big(\Psi(\dot{\mathsf{u}}_n(s)) + \mathfrak{G}_{\varepsilon_n}(\mathsf{t}_n(s), \tilde{\mathsf{u}}_n(s); \dot{\mathsf{t}}_n(s), \dot{\mathsf{u}}_n(s))\Big)\,\mathrm{d}s. \quad (7.12)$$

We can thus apply Proposition 7.1 to pass to the limit obtaining an admissible limit curve $(\mathsf{t}, \mathsf{u}) \in \mathscr{A}(0,1;[0,T] \times D_E)$ such that $\mathsf{t}(s) \equiv t$, $\mathsf{u}(0) = u_-$ and $\mathsf{u}(1) = u_+$. In particular $\mathsf{u} \in \mathcal{T}_t(u_-, u_+)$ and combining (7.12) with (7.5) we get

$$\Delta = \lim_{n\to\infty} \int_0^1 \Big(\Psi(\dot{\mathsf{u}}_n(s)) + \mathfrak{G}_{\varepsilon_n}(\mathsf{t}_n(s), \mathsf{u}_n(s); \dot{\mathsf{t}}_n(s), \dot{\mathsf{u}}_n(s))\Big)\,\mathrm{d}s \geq \int_0^1 \Big(\Psi[\mathsf{u}'](s) + \mathfrak{G}[t, \mathsf{u}; 0, \dot{\mathsf{u}}](s)\Big)\,\mathrm{d}s$$
$$= \int_0^1 \Big(\Psi[\mathsf{u}'](s) + \mathfrak{e}_t(\mathsf{u}(s))\|\dot{\mathsf{u}}(s)\|\Big)\,\mathrm{d}s \geq \Delta_{\mathfrak{f}_t}(u_-, u_+).$$

This concludes the proof of Theorem 3.5. □

The next result is a counterpart to Proposition 7.1 for the lower semicontinuity, but now for the non-parameterized setting.

Proposition 7.3. *Let $E, C > 0$ and for $n \in \mathbb{N}$ let $u_n \subset \mathrm{AC}([0,T];V)$, $\tilde{u}_n : [0,T] \to D_E$, $\xi_n \to [0,T] \to V^*$ measurable, $\varepsilon_n \in (0,\infty)$ be sequences satisfying*

$$\int_0^T \Big(\Psi_{\varepsilon_n}(\dot{u}_n) + \Psi^*_{\varepsilon_n}(\xi_n)\Big)\,\mathrm{d}t \leq C, \quad \xi_n(t) \in -\partial\mathcal{E}_t(\tilde{u}_n(t)) \quad \textit{for } \mathscr{L}^1\textit{-a.a. } t \in (0,T), \quad (7.13a)$$
$$X_n := \sup_{t\in[0,T]} \|u_n(t) - \tilde{u}_n(t)\| \to 0, \quad \varepsilon_n \downarrow 0 \qquad \textit{as } n \uparrow \infty. \quad (7.13b)$$

Then there exists a subsequence (not relabeled) and a limit function $u \in \mathrm{BV}([0,T]; D_E, \Psi)$ such that convergence (3.27) *holds, u satisfies the local stability condition* ($\mathrm{S_{loc}}$)*, and*

$$\liminf_{n\to\infty} \int_r^s \Big(\Psi_{\varepsilon_n}(\dot{u}_{\varepsilon_n}(t)) + \Psi^*_{\varepsilon_n}(\xi_{\varepsilon_n}(t))\Big)\,\mathrm{d}t \geq \mathrm{Var}_{\mathfrak{f}}(u;[r,s]) \quad \textit{for every } 0 \leq r < s \leq T. \quad (7.14)$$

Proof. To obtain a pointwise convergent subsequence, we proceed as in the proof Proposition 7.1. Setting $\mathrm{V}_n(t) := \int_0^t \Psi_{\varepsilon_n}(\dot{u}_n)\,\mathrm{d}r$ and using $\tilde{u}_n(t) \in D_E$ we get a similar estimate as in (7.7):

$$\|\tilde{\mathsf{u}}_n(t) - \tilde{\mathsf{u}}_n(s)\| \leq \Omega\big(\mathrm{V}_n(t) - \mathrm{V}_n(s)\big) + 2C_\Psi \Omega(X_n) \quad \text{for every } 0 \leq s < t \leq T,\ n \in \mathbb{N}. \quad (7.15)$$

Since the functions V_n are increasing and uniformly bounded by C, by Helly's Theorem we can extract a subsequence (not relabeled) pointwise converging to the increasing function V; passing to the limit in (7.15) along such a subsequence, we obtain

$$\limsup_{n\to\infty} \|\tilde{\mathsf{u}}_n(t) - \tilde{\mathsf{u}}_n(s)\| \leq \Omega\big(\mathrm{V}(t) - \mathrm{V}(s)\big). \quad (7.16)$$

Applying the compactness result [AGS08, Prop. 3.3.1] we obtain the pointwise convergence of (a subsequence of) $\tilde{u}_n$ and thus (3.27) follows by (7.13b).

By the strong-weak closedness (2.33) of the graph of $(\mathcal{E}, \partial\mathcal{E})$ we have

$$\liminf_{n\to\infty} \Psi^*_{\varepsilon_n}(\xi_{\varepsilon_n}(t)) \geq \mathfrak{k}_t(u(t)) \quad \text{for } \mathscr{L}^1\text{-a.a. } t \in (0,T).$$

Therefore Fatou's lemma yields $\int_0^T \mathfrak{k}_t(u(t))\,\mathrm{d}t < \infty$. As $\mathfrak{k}_t(u(t)) \in \{0, +\infty\}$, we arrive at

$$\mathfrak{k}_t(u(t)) = 0 \quad \text{for } \mathscr{L}^1\text{-a.a. } t \in (0,T).$$

Since $\mathfrak{k}$ is a lower semicontinuous function we conclude that $\mathfrak{k}_t(u(t)) = 0$ for every $t \in [0,T] \setminus \mathrm{J}_u$ and also $\mathfrak{k}_t(u(t_\pm)) = 0$ whenever $t \in \mathrm{J}_u$. Thus u satisfies the local stability condition $(\mathrm{S}_{\mathrm{loc}})$.

To prove (7.14) let us introduce the nonnegative and bounded Borel measures ν_n in $[0,T]$ via

$$\nu_n := \Big(\Psi_{\varepsilon_n}(\dot u_n) + \Psi^*_{\varepsilon_n}(\xi_n)\Big)\mathscr{L}^1. \tag{7.17}$$

Possibly extracting a further subsequence, it is not restrictive to assume that $\nu_n \rightharpoonup^* \nu$ in duality with $\mathrm{C}^0([0,T])$. Since $\nu_n \geq \Psi(\dot u_n)\mathscr{L}^1$ for every interval $(\alpha,\beta) \subset [0,T]$,

$$\nu([\alpha,\beta]) \geq \limsup_{n\to\infty} \int_\alpha^\beta \Psi(\dot u_n)\,\mathrm{d}t \geq \liminf_{n\to\infty} \mathrm{Var}_\Psi(u_n;[\alpha,\beta]) \geq \mathrm{Var}_\Psi(u;[\alpha,\beta]) \geq \mu_{\mathrm{d}}([\alpha,\beta])$$

which in particular yields $\nu \geq \mu_{\mathrm{d}}$ (with μ from (2.15)).

Let us now take $t \in \mathrm{J}_u$ and two sequences $\alpha_n \uparrow t$ and $\beta_n \downarrow t$ such that

$$\lim_{n\to\infty} u_n(\alpha_n) = u(t_-), \quad \lim_{n\to\infty} u_n(\beta_n) = u(t_+).$$

Applying assertion (F3) of Theorem 3.5 and the upper semicontinuity property of weak* convergence of measures on closed sets, we get

$$\begin{aligned}\nu(\{t\}) \geq \limsup_{n\to\infty} \nu_n([\alpha_n,\beta_n]) &\geq \liminf_{n\to\infty} \int_{\alpha_n}^{\beta_n} \Big(\Psi_{\varepsilon_n}(\dot u_n) + \Psi^*_{\varepsilon_n}(\xi_n)\Big)\,\mathrm{d}t \\ &\geq \Delta_{\mathfrak{f}_t}(u(t_-),u(t_+)) = \mu_{\mathrm{J}}(\{t\}),\end{aligned} \tag{7.18}$$

and similarly

$$\limsup_{n\to\infty} \nu_n([\alpha_n,t]) \geq \Delta_{\mathfrak{f}_t}(u(t_-),u(t)), \quad \limsup_{n\to\infty} \nu_n([t,\beta_n]) \geq \Delta_{\mathfrak{f}_t}(u(t),u(t_+)). \tag{7.19}$$

It follows from (7.18) that $\nu \geq \mu$. If now $0 \leq r < s \leq T$ we can choose $r_n > r$ and $s_n < s$ such that $r_n \downarrow r$ with $u_n(r_n) \to u(r_+)$ and $s_n \uparrow s$ with $u_n(s_n) \to u(s_-)$. Eventually we have

$$\begin{aligned}\liminf_{n\to\infty} \int_r^s \Big(\Psi_{\varepsilon_n}(\dot u_n) + \Psi^*_{\varepsilon_n}(\xi_n)\Big)\,\mathrm{d}t &\geq \liminf_{n\to\infty} \nu_n([r,r_n]) + \liminf_{n\to\infty} \nu_n((r_n,s_n)) + \liminf_{n\to\infty} \nu_n([s_n,s]) \\ &\geq \Delta_{\mathfrak{f}_r}(u(r),u(r_+)) + \nu((r,s)) + \Delta_{\mathfrak{f}_s}(u(s_-),u(s)) \\ &\geq \Delta_{\mathfrak{f}_r}(u(r),u(r_+)) + \mu((r,s)) + \Delta_{\mathfrak{f}_s}(u(s_-),u(s)) \overset{(3.23)}{=} \mathrm{Var}_{\mathfrak{f}}(u;[r,s]).\end{aligned}$$

□

7.3. **Convergence of the vanishing-viscosity approximations.** Here we prove Theorem 3.9, which states that the limit u of solutions u_ε to the doubly nonlinear equations $(1.1)_\varepsilon$ are Balanced Viscosity (BV) solutions.

Proof of Theorem 3.9. Let $(u_\varepsilon)_\varepsilon \subset \mathrm{AC}([0,T];V)$ be a family of solutions to (1.1) fulfilling (3.26) at $t = 0$: in particular, $E_0 := \sup_\varepsilon \Psi(u_\varepsilon(0)) + \mathcal{E}_0(u_\varepsilon(0)) < \infty$.

We combine the energy identity (2.35), written for $s = 0$ and for any $t \in (0,T]$, with the estimate for $\mathcal{P}_t$ in (E.2), obtaining

$$\begin{aligned}\Psi(u_\varepsilon(t)) + \mathcal{E}_t(u_\varepsilon(t)) &\leq \Psi(u_\varepsilon(0)) + \int_0^t \Big(\Psi_\varepsilon(\dot u_\varepsilon(r)) + \Psi^*_\varepsilon(\xi_\varepsilon(r))\Big)\,\mathrm{d}r + \mathcal{E}_t(u_\varepsilon(t)) \\ &= \Psi(u_\varepsilon(0)) + \mathcal{E}_0(u_\varepsilon(0)) + \int_0^t \mathcal{P}_r(u_\varepsilon(r))\,\mathrm{d}r \leq E_0 + C_P \int_0^t \Big(\Psi(u_\varepsilon(r)) + \mathcal{E}_r(u_\varepsilon(r))\Big)\,\mathrm{d}r.\end{aligned}$$

Applying a standard version of the Gronwall Lemma (cf. e.g. [Bré73, Lem. A.4]), we deduce that there exist constants $E, C > 0$ such that for every $\varepsilon > 0$ and $t \in [0,T]$ we have

$$\Psi(u_\varepsilon(t)) + \mathcal{E}_t(u_\varepsilon(t)) \leq E := E_0 \exp(C_P T) \quad \text{and} \quad \int_0^T \Big(\Psi_\varepsilon(\dot u_\varepsilon(r)) + \Psi^*_\varepsilon(\xi_\varepsilon(r))\Big)\,\mathrm{d}r \leq C.$$

By Proposition 7.3, for every vanishing sequence $(\varepsilon_k)_k$ there exists a further subsequence and $u \in \mathrm{BV}([0,T];D_E,\Psi)$ such that convergence (3.27) holds. By lower semicontinuity, we also have

$$\liminf_{k\to\infty} \mathcal{E}_t(u_{\varepsilon_k}(t)) \geq \mathcal{E}_t(u(t)) \quad \text{for all } t \in [0,T]. \tag{7.20}$$

Furthermore, by (E.2) we have $|\mathcal{P}_t(u_{\varepsilon_k}(t))| \leq C_P E$ for all $k \in \mathbb{N}$ and $t \in [0,T]$. Therefore, applying Fatou's Lemma we obtain

$$\limsup_{k\to\infty} \int_s^t \mathcal{P}_r(u_{\varepsilon_k}(r))\mathrm{d}r \leq \int_s^t \mathcal{P}_r(u(r))\mathrm{d}r \quad \text{for all } 0 \leq s \leq t \leq T. \tag{7.21}$$

We can now pass to the limit in the energy identity (2.35) as $k \to \infty$. Combining (7.14) $r = 0$ and $s = T$ with (7.20), we immediately get ($\mathrm{E}_{\mathfrak{f},\mathrm{ineq}}$). We thus deduce that u is a BV solution.

The energy identity ($\mathrm{E}_{\mathfrak{f}}$) satisfied by u on the interval $[0,T]$ and the elementary property of real sequences

$$a, b \in \mathbb{R}, \quad \begin{cases} \liminf\limits_{n\to\infty} a_n \geq a \\ \liminf\limits_{n\to\infty} b_n \geq b, \end{cases} \quad \limsup_{n\uparrow\infty}(a_n + b_n) \leq a + b \quad \Longrightarrow \quad \begin{cases} \lim\limits_{n\to\infty} a_n = a \\ \lim\limits_{n\to\infty} b_n = b, \end{cases} \tag{7.22}$$

yield that

$$\lim_{k\to\infty} \mathcal{E}_T(u_{\varepsilon_k}(T)) = \mathcal{E}_T(u(T)), \quad \lim_{k\to\infty} \int_0^T \Big(\Psi_{\varepsilon_k}(\dot{u}_{\varepsilon_k}) + \Psi^*_{\varepsilon_k}(\xi_{\varepsilon_k})\Big)\,\mathrm{d}r = \mathrm{Var}_{\mathfrak{f}}(u;[0,T]). \tag{7.23}$$

A further application of (7.14) on the intervals $[0,t]$ and $[t,T]$ combined with (7.20), the additivity of the total variation, and (7.22) provides convergences (3.28) and (3.29). Hence, Theorem 3.9 is proved. $\square$

Convergence of the discrete viscous approximations. Let us consider the time-incremental minimization problem ($\mathrm{IP}_{\varepsilon,\tau}$), giving rise to the discrete solutions $(U^n_{\tau,\varepsilon})_{n=1}^N$ which fulfill the *discrete Euler equation*

$$\partial\Psi\left(\frac{U^n_{\tau,\varepsilon} - U^{n-1}_{\tau,\varepsilon}}{\tau}\right) + \partial\Phi\left(\varepsilon\frac{U^n_{\tau,\varepsilon} - U^{n-1}_{\tau,\varepsilon}}{\tau}\right) + \partial\mathcal{E}_{t_n}(U^n_{\tau,\varepsilon}) \ni 0 \qquad \text{for all } 1,\ldots,N_\tau. \tag{7.24}$$

We denote by $\overline{\mathrm{U}}_{\tau,\varepsilon}$ the left-continuous piecewise constant interpolants, thus taking the value $\mathrm{U}^n_{\tau,\varepsilon}$ for $t \in (t_{n-1}, t_n]$, and by $\mathrm{U}_{\tau,\varepsilon}$ the piecewise affine interpolant

$$\mathrm{U}_{\tau,\varepsilon}(t) := \frac{t - t_{n-1}}{\tau}\mathrm{U}^n_{\tau,\varepsilon} + \frac{t_n - t}{\tau}\mathrm{U}^{n-1}_{\tau,\varepsilon} \quad \text{for } t \in [t_{n-1}, t_n], \quad n = 1,\ldots,N_\tau. \tag{7.25}$$

As in [MRS13], we also consider the *variational interpolant* $\widetilde{\mathrm{U}}_{\tau,\varepsilon}$ of the elements $(\mathrm{U}^n_{\tau,\varepsilon})_{n=1}^N$, first introduced by E. De Giorgi in the frame of the *Minimizing Movements* approach to gradient flows (see [DGMT80, De 93, Amb95, AGS08]). The functions $\widetilde{\mathrm{U}}_{\tau,\varepsilon} : [0,T] \to V$ are defined by $\widetilde{\mathrm{U}}_{\tau,\varepsilon}(0) = u_\varepsilon(0)$ and

$$\text{for } t = t_{n-1} + r \in (t_{n-1}, t_n], \quad \widetilde{\mathrm{U}}_{\tau,\varepsilon}(t) \in \mathop{\mathrm{Argmin}}_{U\in D}\left\{ r\Psi_\varepsilon\left(\frac{U - \mathrm{U}^{n-1}_{\tau,\varepsilon}}{r}\right) + \mathcal{E}_t(U)\right\}, \tag{7.26}$$

choosing the minimizer in (7.26) so that the map $t \mapsto \widetilde{\mathrm{U}}_\tau(t)$ is Lebesgue measurable in $(0,T)$. Notice that we may assume $\overline{\mathrm{U}}_{\tau,\varepsilon}(t_n) = \mathrm{U}_{\tau,\varepsilon}(t_n) = \widetilde{\mathrm{U}}_{\tau,\varepsilon}(t_n)$ for every $n = 1,\ldots,N_\tau$. Moreover, with the variational interpolants $\widetilde{\mathrm{U}}_{\tau,\varepsilon}$ we can associate a measurable function $\widetilde{\xi}_{\tau,\varepsilon} : (0,T) \to V^*$ fulfilling the Euler equation for the minimization problem (7.26), i.e.

$$\widetilde{\xi}_{\tau,\varepsilon}(t) \in -\partial\mathcal{E}_t(\widetilde{\mathrm{U}}_{\tau,\varepsilon}(t)) \cap \left(\partial\Psi_\varepsilon\left(\frac{\widetilde{\mathrm{U}}_{\tau,\varepsilon}(t) - \mathrm{U}^{n-1}_{\tau,\varepsilon}}{t - t_{n-1}}\right)\right) \qquad \forall\, t \in (t_{n-1}, t_n], \quad n = 1,\ldots,N_\tau, \tag{7.27}$$

cf. [MRS13] for further details. Finally, we also set $\bar{\mathrm{t}}_\tau(t) := t_k$ for $t \in (t_{k-1}, t_k]$. Observe that, for every $t \in [0,T]$ there holds $\bar{\mathrm{t}}_\tau(t) \downarrow t$ as $\tau \downarrow 0$.

We recall now a list of important properties of the discrete solutions, stated in [MRS13, Sec. 6].

Proposition 7.4. *For every $\varepsilon > 0$ and $\tau > 0$ the* discrete energy inequality

$$\int_{\bar{\mathsf{t}}_\tau(s)}^{\bar{\mathsf{t}}_\tau(t)} \Big(\Psi_\varepsilon(\dot{\mathrm{U}}_{\tau,\varepsilon}) + \Psi_\varepsilon^*(\widetilde{\xi}_{\tau,\varepsilon})\Big)\,\mathrm{d}r + \mathcal{E}_{\bar{\mathsf{t}}_\tau(t)}(\overline{\mathrm{U}}_{\tau,\varepsilon}(t)) \le \mathcal{E}_{\bar{\mathsf{t}}_\tau(s)}(\overline{\mathrm{U}}_{\tau,\varepsilon}(s)) + \int_{\bar{\mathsf{t}}_\tau(s)}^{\bar{\mathsf{t}}_\tau(t)} \mathcal{P}_r(\widetilde{\mathrm{U}}_{\tau,\varepsilon}(r))\mathrm{d}r \tag{7.28}$$

holds for every $0 \le s \le t \le T$. If moreover $\Psi(\mathrm{U}^0_{\tau,\varepsilon}) + \mathcal{E}_0(\mathrm{U}^0_{\tau,\varepsilon}) \le E_0$ for all $\tau > 0$ and $\varepsilon > 0$, then there exist constants $E, S > 0$ such that for every $\tau, \varepsilon > 0$ we have

$$\sup_{t\in[0,T]} \big(\mathcal{G}(\overline{\mathrm{U}}_{\tau,\varepsilon}(t)) + \mathcal{G}(\widetilde{\mathrm{U}}_{\tau,\varepsilon}(t))\big) \le E\,, \tag{7.29}$$

$$\mathrm{Var}_\Psi(\mathrm{U}_{\tau,\varepsilon};[0,T]) \le \int_0^T \Psi_\varepsilon(\dot{\mathrm{U}}_{\tau,\varepsilon}(s))\,\mathrm{d}s \le S, \qquad \int_0^T \Psi_\varepsilon^*(\widetilde{\xi}_{\tau,\varepsilon}(s))\,\mathrm{d}s \le S, \tag{7.30}$$

$$\sup_{t\in[0,T]} \Big(\|\mathrm{U}_{\tau,\varepsilon}(t) - \widetilde{\mathrm{U}}_{\tau,\varepsilon}(t)\| + \|\mathrm{U}_{\tau,\varepsilon}(t) - \overline{\mathrm{U}}_{\tau,\varepsilon}(t)\|\Big) \le S\,\omega\Big(\frac{\tau}{S\varepsilon}\Big)\,, \tag{7.31}$$

$$\text{where } \omega(r) := \sup\big\{v \in [0,\infty) : r\,F(r^{-1}v) \le 1\big\}$$

satisfies $\lim_{r\downarrow 0}\omega(r) = 0$, in view of the superlinearity of F.

Proof of Theorem 3.10. We argue exactly as in the proof of Theorem 3.9, observing that Proposition 7.4 enables us to apply Proposition 7.3 with the choices $u_k := \mathrm{U}_{\tau_k,\varepsilon_k}$, $\tilde{u}_k := \widetilde{\mathrm{U}}_{\tau_k,\varepsilon_k}$ along any sequences τ_k, ε_k satisfying (3.31).

Up to the extraction of a suitable subsequence, Proposition 7.3 shows that there exist $u \in \mathrm{BV}([0,T]; D_E, \Psi)$ satisfying the local stability condition $(\mathrm{S_{loc}})$ such that

$$\overline{\mathrm{U}}_{\tau_k,\varepsilon_k}(t),\ \mathrm{U}_{\tau_k,\varepsilon_k}(t),\ \widetilde{\mathrm{U}}_{\tau_k,\varepsilon_k}(t) \to u(t) \quad \text{in } V \text{ for all } t \in [0,T], \tag{7.32}$$

$$\sup_{t\in[0,T]} \Big(\|\mathrm{U}_{\tau_k,\varepsilon_k}(t) - \widetilde{\mathrm{U}}_{\tau_k,\varepsilon_k}(t)\| + \|\mathrm{U}_{\tau_k,\varepsilon_k}(t) - \overline{\mathrm{U}}_{\tau_k,\varepsilon_k}(t)\|\Big) \to 0. \tag{7.33}$$

We can also pass to the limit as $k \to \infty$ in the discrete energy inequality (7.28) with $s = 0$. Indeed, we use convergences (7.32), the lower semicontinuity of the energy $\mathcal{E}$, and the lim inf inequality (7.14) to obtain $(\mathrm{E}_{\mathfrak{f},\mathrm{ineq}})$. Thus, by Corollary 3.12 we conclude that u is a BV solution to the RIS $(V, \mathcal{E}, \Psi, \Phi)$.

The proof of the further energy convergence (3.33) follows by the very same lines as in the end of the proof of Theorem 3.9, see (7.22)–(7.23). Thus, Theorem 3.10 is proved. □

Proof of Theorem 4.3. Let $(\mathsf{t}_\varepsilon, \mathsf{u}_\varepsilon)_\varepsilon$ be a family of rescaled viscous solutions as in the statement of Theorem 4.3. Exploiting condition (4.18) as well as the energy identity (4.6) we can apply Proposition 7.1 in the interval $[0,\mathsf{S}]$ (with $\tilde{\mathsf{u}}_n \equiv \mathsf{u}_n$ and $G_n = [0,\mathsf{S}]$) and find a vanishing subsequence $(\varepsilon_n)_n$ and a parameterized curve (t, u) such that convergences (4.19) hold. The second part of (E.2), the closedness-continuity property (2.33), and Lemma 7.2 yield

$$\begin{aligned} &\liminf_{k\to\infty} \mathcal{E}_{\mathsf{t}_{\varepsilon_k}(s)}(\mathsf{u}_{\varepsilon_k}(s)) \ge \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \quad \text{for all } s \in [0,\mathsf{S}], \\ &\limsup_{k\to\infty} \int_{s_0}^{s_1} \mathcal{P}_r(\mathsf{u}_{\varepsilon_k}(r))\dot{\mathsf{t}}_{\varepsilon_k}(r)\mathrm{d}r \le \int_{s_0}^{s_1} \mathcal{P}_r(\mathsf{u}(r))\dot{\mathsf{t}}(r)\mathrm{d}r \quad \text{for all } 0 \le s_0 < s_1 \le \mathsf{S}. \end{aligned} \tag{7.34}$$

Combining (7.34) with (7.2) and (7.5), we pass to the limit as $\varepsilon_k \to 0$ in the energy identity (4.6) and conclude that (t, u) fulfill the energy estimate (4.24) with $\mathsf{a} = 0$ and $\mathsf{b} = \mathsf{S}$. Therefore thanks to Corollary 4.5 we have that (t, u) is a parameterized solution to the RIS $(V, \mathcal{E}, \Psi, \Phi)$.

The enhanced convergences (4.20) and (4.21) can be proved with similar arguments as in the end of the proof of Theorem 3.9.

In order to show that (t, u) satisfies the m-normalization condition (4.15), we observe that $\dot{\mathsf{t}}_\varepsilon \rightharpoonup^* \mathsf{t}$ and $\mathfrak{f}_{\mathsf{t}_{\varepsilon_k}}(\mathsf{u}_{\varepsilon_k}, \dot{\mathsf{u}}_{\varepsilon_k}) \rightharpoonup^* \mathfrak{f} = \mathsf{m} - \dot{\mathsf{t}}$ in $L^\infty(0,\mathsf{S})$. The liminf estimates (7.2) and (7.3) (localized on arbitrary intervals of $[0,\mathsf{S}]$) yield that $\mathfrak{f} \ge \mathfrak{h} := \Psi(\dot{\mathsf{u}}) + \mathfrak{G}[\mathsf{t}, \mathsf{u}; \dot{\mathsf{t}}, \dot{\mathsf{u}}]$ $\mathscr{L}^1$-a.e. in $(0,\mathsf{S})$. Moreover, $\mathfrak{f}_{\mathsf{t}_\varepsilon}(\mathsf{u}_\varepsilon, \dot{\mathsf{u}}_\varepsilon) \le \mathfrak{h}_\varepsilon := \Psi(\dot{\mathsf{u}}_\varepsilon) + \mathfrak{G}_\varepsilon(\mathsf{t}_\varepsilon, \mathsf{u}_\varepsilon; \dot{\mathsf{t}}_\varepsilon, \dot{\mathsf{u}}_\varepsilon)$ and convergence (4.21) implies

$$\mathfrak{h}_{\varepsilon_k} \rightharpoonup \mathfrak{h} \quad \text{in the sense of distributions of } \mathscr{D}'(0,\mathsf{S}),$$

so that $\mathfrak{f} \le \mathfrak{h}$. We conclude that $\mathfrak{f} = \mathfrak{h}$ and $\dot{\mathfrak{t}} + \mathfrak{h} = \mathfrak{m}$, and Theorem 4.3 is proved. $\square$

7.4. **Uniform** BV**-estimates for discrete Minimizing Movements.** The aim of this section is to prove Theorem 3.21, i.e. the uniform bound

$$\exists C > 0 \ \forall \tau > 0, \varepsilon > 0 : \quad \sum_{n=1}^{N_\tau} \|\mathrm{U}^n_{\tau,\varepsilon} - \mathrm{U}^{n-1}_{\tau,\varepsilon}\| \le C \tag{7.35}$$

for all discrete Minimizing Movements, whenever the stronger structural assumptions (3.55)–(3.57) hold and the discrete initial data satisfy (3.58). We start with an elementary discrete Gronwall-like lemma.

Lemma 7.5 (A discrete Gronwall lemma)**.** *Let $\gamma > 0$ and let $(a_n), (b_n) \subset [0, +\infty)$ be positive sequences, satisfying*

$$(1+\gamma)^2 \, a_n^2 \le a_{n-1}^2 + b_n a_n \quad \forall n \ge 1. \tag{7.36}$$

Then, for all $k \in \mathbb{N}$ there holds

$$\sum_{n=1}^k a_n \le \frac{1}{\gamma}\Big(a_0 + \sum_{n=1}^k b_n\Big). \tag{7.37}$$

Proof. We first show that assumption (7.36) yields

$$(1+\gamma) a_n \le a_{n-1} + b_n. \tag{7.38}$$

Indeed, (7.38) is trivially true if $(1+\gamma)a_n \le a_{n-1}$. If $(1+\gamma)a_n > a_{n-1}$ we divide both sides in (7.36) by $(1+\gamma)a_n$ and estimate the right-hand side by $\frac{a_{n-1}^2}{(1+\gamma)a_n} + \frac{b_n}{1+\gamma} < a_{n-1} + b_n$. Summing (7.38) from $n = 1$ to k and setting $S_k := \sum_{n=1}^k a_n$ we find $(1+\gamma) S_k \le a_0 + S_{k-1} + \sum_{n=1}^k b_n$, which yields (7.37) since $S_{k-1} \le S_k$. $\square$

Proof of Theorem 3.21. From estimate (7.29) it follows that $\mathrm{U}^n_{\tau,\varepsilon} \in D_E$ for all n and all ε, $\tau > 0$. Therefore (3.56)–(3.57) (and a fortiori (3.60)) hold with constants α_E, Λ_E, L_E.

Notice moreover that setting $\mathrm{U}^{-1}_{\tau,\varepsilon} := 0$, the discrete Euler equation (7.24) is satisfied also for $n = 0$. Let us set $\mathrm{V}^n_{\tau,\varepsilon} := \tau^{-1}(\mathrm{U}^n_{\tau,\varepsilon} - \mathrm{U}^{n-1}_{\tau,\varepsilon})$, $\Xi^n_{\tau,\varepsilon} \in -\partial\mathcal{E}_{t_n}(\mathrm{U}^n_{\tau,\varepsilon}) \cap \partial\Psi_\varepsilon(\mathrm{V}^n_{\tau,\varepsilon})$ according to (7.24). We subtract (7.24) at n from (7.24) at $n+1$, and take the duality with $\mathrm{V}^{n+1}_{\tau,\varepsilon}$, observing that the generalized convexity condition (3.60) yields

$$\langle \Xi^{n+1}_{\tau,\varepsilon} - \Xi^n_{\tau,\varepsilon}, \mathrm{V}^{n+1}_{\tau,\varepsilon}\rangle \le -2\alpha_E\tau\|\mathrm{V}^{n+1}_{\tau,\varepsilon}\|^2 + 2\tau\Lambda_E\Psi_\wedge(\mathrm{V}^{n+1}_{\tau,\varepsilon})\|\mathrm{V}^{n+1}_{\tau,\varepsilon}\| + 2\tau\|\mathrm{V}^{n+1}_{\tau,\varepsilon}\|. \tag{7.39}$$

On the other hand the homogeneity of Ψ and Φ yield

$$\begin{aligned} &\langle \partial\Psi(\mathrm{V}^{n+1}_{\tau,\varepsilon}), \mathrm{V}^{n+1}_{\tau,\varepsilon}\rangle = \Psi(\mathrm{V}^{n+1}_{\tau,\varepsilon}), \quad \langle \partial\Psi(\mathrm{V}^{n}_{\tau,\varepsilon}), \mathrm{V}^{n+1}_{\tau,\varepsilon}\rangle \le \Psi(\mathrm{V}^{n+1}_{\tau,\varepsilon}), \\ &\langle \partial\Phi(\varepsilon\mathrm{V}^{n+1}_{\tau,\varepsilon}), \mathrm{V}^{n+1}_{\tau,\varepsilon}\rangle = \varepsilon\|\mathrm{V}^{n+1}_{\tau,\varepsilon}\|^2, \quad \langle \partial\Phi(\varepsilon\mathrm{V}^{n}_{\tau,\varepsilon}), \mathrm{V}^{n+1}_{\tau,\varepsilon}\rangle \le \frac{\varepsilon}{2}\|\mathrm{V}^{n+1}_{\tau,\varepsilon}\|^2 + \frac{\varepsilon}{2}\|\mathrm{V}^{n}_{\tau,\varepsilon}\|^2. \end{aligned}$$

and therefore

$$\langle \Xi^{n+1}_{\tau,\varepsilon} - \Xi^n_{\tau,\varepsilon}, \mathrm{V}^{n+1}_{\tau,\varepsilon}\rangle \ge \frac{\varepsilon}{2}\|\mathrm{V}^{n+1}_{\tau,\varepsilon}\|^2 - \frac{\varepsilon}{2}\|\mathrm{V}^{n}_{\tau,\varepsilon}\|^2. \tag{7.40}$$

Combining (7.39) and (7.40) we get

$$\|\mathrm{V}^{n+1}_{\tau,\varepsilon}\|^2 + \frac{4\alpha\tau}{\varepsilon}\|\mathrm{V}^{n+1}_{\tau,\varepsilon}\|^2 \le \|\mathrm{V}^{n}_{\tau,\varepsilon}\|^2 + \frac{4\tau}{\varepsilon}\left(L_E + \Lambda_E\Psi_\wedge(\mathrm{V}^n_{\tau,\varepsilon})\right)\|\mathrm{V}^{n}_{\tau,\varepsilon}\|,$$

Observe that the above inequality can be rewritten in the form of (7.36) with the choices $a_n = \|\mathrm{V}^n_{\tau,\varepsilon}\|$, $b_n = \frac{4\tau}{\varepsilon}\big(L_E + \Psi(\mathrm{V}^n_{\tau,\varepsilon})\big)$, and $\gamma := (1 + 4\alpha\tau/\varepsilon)^{1/2} - 1$. Using $a_0 = \|\mathrm{V}^0_{\tau,\varepsilon}\| = 0$ and applying Lemma 7.5 elementary computations yield

$$\sum_{n=1}^{N_\tau - 1} \tau\|\mathrm{V}^n_{\tau,\varepsilon}\| \le \Big(4Q + \frac{2}{\alpha}\Big)\Big(TL_E + E\Big),$$

which is the desired estimate (3.53). $\square$

References

[AGS08] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[Amb95] L. Ambrosio. Minimizing movements. *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5)*, 19, 191–246, 1995.

[BFM12] J.-F. Babadjian, G. Francfort, and M. Mora. Quasistatic evolution in non-associative plasticity - the cap model. *SIAM J. Math. Anal.*, 44, 245–292, 2012.

[Bré73] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Publishing Co., Amsterdam, 1973. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).

[Col92] P. Colli. On some doubly nonlinear evolution equations in Banach spaces. *Japan J. Indust. Appl. Math.*, 9(2), 181–203, 1992.

[CoV90] P. Colli and A. Visintin. On a class of doubly nonlinear evolution equations. *Comm. Partial Differential Equations*, 15(5), 737–756, 1990.

[DDS12] G. Dal Maso, A. DeSimone, and F. Solombrino. Quasistatic evolution for cam-clay plasticity: properties of the viscosity solutions. *Calc. Var. Partial Differential Equations*, 2012. DOI: 10.1007/s00526-011-0443-6. In press. Published online August 2011.

[De 93] E. De Giorgi. New problems on minimizing movements. In C. Baiocchi and J. L. Lions, editors, *Boundary Value Problems for PDE and Applications*, pages 81–98. Masson, 1993.

[DFT05] G. Dal Maso, G. Francfort, and R. Toader. Quasistatic crack growth in nonlinear elasticity. *Arch. Rational Mech. Anal.*, 176, 165–225, 2005.

[DGMT80] E. De Giorgi, A. Marino, and M. Tosques. Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 68(3), 180–187, 1980.

[DiU77] J. Diestel and J. J. Uhl, Jr. *Vector measures.* American Mathematical Society, Providence, R.I., 1977. With a foreword by B. J. Pettis, Mathematical Surveys, No. 15.

[DMDS11] G. Dal Maso, A. DeSimone, and F. Solombrino. Quasistatic evolution for Cam-Clay plasticity: a weak formulation via viscoplastic regularization and time rescaling. *Calc. Var. Partial Differential Equations*, 40(1-2), 125–181, 2011.

[EfM06] M. Efendiev and A. Mielke. On the rate–independent limit of systems with dry friction and small viscosity. *J. Convex Analysis*, 13(1), 151–167, 2006.

[IoT79] A. D. Ioffe and V. M. Tihomirov. *Theory of extremal problems*, volume 6 of *Studies in Mathematics and its Applications.* North-Holland Publishing Co., Amsterdam, 1979. Translated from the Russian by Karol Makowski.

[KMZ08] D. Knees, A. Mielke, and C. Zanini. On the inviscid limit of a model for crack propagation. *Math. Models Methods Appl. Sci.*, 18(9), 1529–1569, 2008.

[KRZ13] D. Knees, R. Rossi, and C. Zanini. A vanishing viscosity approach to a rate-independent damage model. *Math. Models Methods Appl. Sci.*, 23(4), 565–616, 2013.

[KZM10] D. Knees, C. Zanini, and A. Mielke. Crack propagation in polyconvex materials. *Physica D*, 239, 1470–1484, 2010.

[MaM09] A. Mainik and A. Mielke. Global existence for rate-independent gradient plasticity at finite strain. *J. Nonlinear Science*, 19(3), 221–248, 2009.

[Mie05] A. Mielke. Evolution in rate-independent systems (Ch. 6). In C. Dafermos and E. Feireisl, editors, *Handbook of Differential Equations, Evolutionary Equations, vol. 2*, pages 461–559. Elsevier B.V., Amsterdam, 2005.

[Mie11] A. Mielke. Differential, energetic and metric formulations for rate-independent processes (Ch. 3). In L. Ambrosio and G. Savaré, editors, *Nonlinear PDEs and Applications.C.I.M.E. Summer School, Cetraro, Italy 2008*, pages 87–170. Springer, Heidelberg, 2011.

[MiR06] A. Mielke and T. Roubíček. Rate-independent damage processes in nonlinear elasticity. *M^3AS Math. Models Methods Appl. Sci.*, 16, 177–209, 2006.

[MiT99] A. Mielke and F. Theil. A mathematical model for rate-independent phase transformations with hysteresis. In H.-D. Alber, R. Balean, and R. Farwig, editors, *Proceedings of the Workshop on "Models of Continuum Mechanics in Analysis and Engineering"*, pages 117–129, Aachen, 1999. Shaker-Verlag.

[MiT04] A. Mielke and F. Theil. On rate–independent hysteresis models. *Nonl. Diff. Eqns. Appl. (NoDEA)*, 11, 151–189, 2004. (Accepted July 2001).

[MiZ12] A. Mielke and S. Zelik. On the vanishing viscosity limit in parabolic systems with rate-independent dissipation terms. *Ann. Sc. Norm. Sup. Pisa Cl. Sci. (5)*, 2012. To appear. WIAS preprint 1500.

[MRS09] A. Mielke, R. Rossi, and G. Savaré. Modeling solutions with jumps for rate-independent systems on metric spaces. *Discrete Contin. Dyn. Syst.*, 25, 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, 36–80, 2012.

[MRS12b] A. Mielke, R. Rossi, and G. Savaré. Variational convergence of gradient flows and rate-independent evolutions in metric spaces. *Milan J. Math.*, 80, 381–410, 2012.

[MRS13] A. Mielke, R. Rossi, and G. Savaré. Nonsmooth analysis of doubly nonlinear equations. *Calc. Var. Partial Differential Equations*, 46(1-2), 253–310, 2013.

[MST89] A. Marino, C. Saccon, and M. Tosques. Curves of maximal slope and parabolic variational inequalities on nonconvex constraints. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 16(2), 281–330, 1989.

[RMS08] R. Rossi, A. Mielke, and G. Savaré. A metric approach to a class of doubly nonlinear evolution equations and applications. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 7, 97–169, 2008.

[Roc70] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, 1970.

[RoS06] R. Rossi and G. Savaré. Gradient flows of non convex functionals in Hilbert spaces and applications. *ESAIM Control Optim. Calc. Var.*, 12(3), 564–614 (electronic), 2006.

[RoS13] R. Rossi and G. Savaré. A characterization of energetic and BV solutions to one-dimensional rate-independent systems. *Discrete Contin. Dyn. Syst. (S)*, 6, 167–191, 2013.