ESAIM: COCV 18 (2012) 36–80
DOI: 10.1051/cocv/2010054

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

BV SOLUTIONS AND VISCOSITY APPROXIMATIONS OF RATE-INDEPENDENT SYSTEMS*

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

Abstract. In the nonconvex case, solutions of rate-independent systems may develop jumps as a function of time. To model such jumps, we adopt the philosophy that rate-independence should be considered as limit of systems with smaller and smaller viscosity. For the finite-dimensional case we study the vanishing-viscosity limit of doubly nonlinear equations given in terms of a differentiable energy functional and a dissipation potential that is a viscous regularization of a given rate-independent dissipation potential. The resulting definition of "BV solutions" involves, in a nontrivial way, both the rate-independent and the viscous dissipation potential, which play crucial roles in the description of the associated jump trajectories. We shall prove general convergence results for the time-continuous and for the time-discretized viscous approximations and establish various properties of the limiting BV solutions. In particular, we shall provide a careful description of the jumps and compare the new notion of solutions with the related concepts of energetic and local solutions to rate-independent systems.

Mathematics Subject Classification. 49Q20, 58E99.

Received October 14, 2009. Revised August 11, 2010.
Published December 23, 2010.

1. Introduction

Rate-independent evolutions occur in several contexts. We refer the reader to [32] and the forthcoming monograph [36] for a survey of rate-independent modeling and analysis in a wide variety of applications, which may pertain to very different and far-apart branches of mechanics and physics. Rate-independent systems present very distinctive common features, because of their hysteretic character [26,52]. Driven by external loadings on a time scale much slower than their internal scale, such systems respond to changes in the external actions invariantly for time-rescalings. Thus, they show (almost) no intrinsic time-scale. This kind of behavior is encoded in the simplest, but still significant, example of rate-independent evolution, namely the *doubly nonlinear*

Keywords and phrases. Doubly nonlinear, differential inclusions, generalized gradient flows, viscous regularization, vanishing-viscosity limit, vanishing-viscosity contact potential, parameterized solutions.

* *A.M. has been partially supported by DFG, Research Unit FOR 797 "MicroPlast", Mi 459/5. R.R. and G.S. have been partially supported by a MIUR-PRIN'06 grant for the project "Variational methods in optimal mass transportation and in geometric measure theory".*

[1] Weierstraß-Institut, Mohrenstraße 39, 10117 Berlin, Germany. mielke@wias-berlin.de

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

[3] Dipartimento di Matematica, Università di Brescia, via Valotti 9, 25133 Brescia, Italy. riccarda.rossi@ing.unibs.it

[4] Dipartimento di Matematica "F. Casorati", Università di Pavia, 27100 Pavia, Italy. giuseppe.savare@unipv.it

Article published by EDP Sciences

© EDP Sciences, SMAI 2010

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

differential inclusion

$$\partial\Psi_0(u'(t)) + \mathrm{D}\mathcal{E}_t(u(t)) \ni 0 \quad \text{in } X^* \quad \text{for a.a. } t \in (0,T). \tag{DN$_0$}$$

For the sake of simplicity, we will consider here the case where X is a finite dimensional linear space. The energy $\mathcal{E} : [0,T] \times X \to \mathbb{R}$ is a functional, $\mathrm{D}\mathcal{E}_t$ denotes its differential with respect to the variable $u \in X$, and $\Psi_0 : X \to [0,+\infty)$ is a convex, nondegenerate, dissipation potential, hereafter supposed *positively homogeneous of degree* 1. Thus, (DN$_0$) is invariant for time rescalings, making the system rate-independent.

By rate-independence the range of $\partial\Psi_0$ equals $K^* := \partial\Psi_0(0)$, which is a proper subset of X. Hence, if $\mathcal{E}_t(\cdot)$ is not strictly convex, one cannot expect the existence of an absolutely continuous solution of (DN$_0$). Over the past decade, this fact has motivated the development of suitable notions of weak solutions to (DN$_0$). In the mainstream of [18,40,42], the present paper aims to contribute to this issue. Relying on the *vanishing-viscosity approach*, we shall propose the notion of BV solution to (DN$_0$) and thoroughly analyze it.

To better motivate the use of vanishing viscosity and highlight the features of the concept of BV solution, in the next paragraphs we shall briefly recall the other main weak solvability notions for (DN$_0$). For the sake of simplicity, we shall focus on the particular case

$$\Psi_0(v) = |v|, \qquad \text{for some norm } |\cdot| \text{ on } X. \tag{1.1}$$

Energetic and local solutions. The first approach to a rigorous weak formulation of (DN$_0$) goes back to [37] and the subsequent [38,41], which advanced the notion of *energetic solution* to the rate-independent system (DN$_0$). In the simplified case (1.1), this solution concept consists of the following relations, holding for all $t \in [0,T]$:

$$\forall\, z \in X : \qquad \mathcal{E}_t(u(t)) \le \mathcal{E}_t(z) + |z - u(t)|, \tag{S}$$

$$\mathcal{E}_t(u(t)) + \mathrm{Var}(u;[0,t]) = \mathcal{E}_0(u(0)) + \int_0^t \partial_t\mathcal{E}_s(u(s))\,\mathrm{d}s. \tag{E}$$

The energy identity (E) balances at every time $t \in [0,T]$ the dissipated energy $\mathrm{Var}(u;[0,t])$ (the latter symbol denotes the total variation of the solution $u \in \mathrm{BV}([0,T];X)$ on the interval $[0,t]$), with the stored energy $\mathcal{E}_t(u(t))$, the initial energy, and the work of the external forces. On the other hand, (S) is a stability condition, for it asserts that the change from the current state $u(t)$ to another state z brings about a gain of potential energy smaller than the dissipated energy. Since the competitors for $u(t)$ range in the whole space X, (S) is in fact a *global stability* condition.

The energetic formulation (S)–(E) only involves the (assumedly smooth) power of the external forces $\partial_t\mathcal{E}$, and is otherwise *derivative-free*. Thus, it is well suited to jumping solutions. Furthermore, as shown in [29,32], it is amenable to analysis in very general ambient spaces, even with no underlying linear structure. Because of its flexibility, this concept has been exploited in a variety of applicative contexts, like, for instance, shape memory alloys [5,34,41], crack propagation [10,12,13], elastoplasticity [14,15,21,30], damage in brittle materials [6,20,28,35,50], delamination [25], and ferroelectricity [39].

On the other hand, in the case of nonconvex energies condition (S) turns out to be a strong requirement, for it may lead the system to change instantaneously in a very drastic way, jumping into very far-apart energetic configurations (see, for instance, [31], Ex. 6.1, [23], Ex. 6.3, and [42], Ex. 1). On the discrete level, global stability is reflected in the global minimization scheme giving raise to approximate solutions by time-discretization. Indeed, for a fixed time-step $\tau > 0$, inducing a partition $\{0 = t_0 < t_1 < \ldots < t_{N-1} < t_N = T\}$ of the interval $[0,T]$, one constructs *discrete solutions* $(\mathrm{U}_\tau^n)_{n=1}^N$ of (S)–(E) by setting $\mathrm{U}_\tau^0 := u_0$ and then solving recursively the variational incremental scheme

$$\mathrm{U}_\tau^n \in \operatorname*{Argmin}_{\mathrm{U}\in X} \Big\{ |\mathrm{U} - \mathrm{U}_\tau^{n-1}| + \mathcal{E}_{t_n}(\mathrm{U}) \Big\} \qquad \text{for } n = 1,\ldots,N. \tag{IP$_\tau$}$$

However, a scheme based on *local minimization* would be preferable, both in view of numerical analysis and from a modeling perspective, see the discussions in [31], Section 6, and, in the realm of crack propagation, [11,28,44].

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

An alternative, purely energetic approach is developed in [28], where minimization is restricted to ε-accessible sets, which are defined *via* continuous ε-slides along which the energy may only increase by ε.

As pointed out in [11], local minimization may be enforced by perturbing the variational scheme (IP_τ) with a term, modulated by a *viscosity* parameter ε, which penalizes the squared distance from the previous step $\mathrm{U}^{n-1}_{\tau,\varepsilon}$

$$\mathrm{U}^n_{\tau,\varepsilon} \in \operatorname*{Argmin}_{\mathrm{U}\in X}\left\{|\mathrm{U}-\mathrm{U}^{n-1}_{\tau,\varepsilon}| + \varepsilon\frac{\|\mathrm{U}-\mathrm{U}^{n-1}_{\tau,\varepsilon}\|^2}{\tau} + \mathcal{E}_{t_n}(\mathrm{U})\right\} \quad \text{for } n=1,\dots,N, \tag{$\mathrm{IP}_{\tau,\varepsilon}$}$$

and depends on a second norm $\|\cdot\|$, typically Hilbertian, on the space X. In a infinite-dimensional setting, one may think of $X = L^2(\Omega)$, with Ω a domain in $\mathbb{R}^d$, $d \geq 1$, and $|\cdot|$, $\|\cdot\|$ the L^1 and L^2 norms, respectively. On the time-continuous level, (IP_τ) corresponds to the *viscous doubly nonlinear* equation

$$\begin{gathered}\partial\Psi_\varepsilon(u_\varepsilon'(t)) + \mathrm{D}\mathcal{E}_t(u_\varepsilon(t)) \ni 0 \quad \text{in } X^* \quad \text{for a.a. } t\in(0,T),\\ \text{with} \quad \Psi_\varepsilon(v) = |v| + \frac{\varepsilon}{2}\|v\|^2\end{gathered} \tag{DN_ε}$$

(see [8,9] for the existence of solutions $u_\varepsilon \in \mathrm{AC}([0,T];X)$). Then, the idea would be to consider the solutions to (DN_0) arising in the passage to the limit, in the discrete scheme ($\mathrm{IP}_{\tau,\varepsilon}$), as ε and τ tend to 0 *simultaneously*, keeping $\tau \ll \varepsilon$. One can guess that, at least formally, this procedure should be equivalent to considering the limit of the solutions to (DN_ε) as $\varepsilon \downarrow 0$.

Vanishing viscosity has by now become an established selection criterion for mechanically feasible weak solvability notions of rate-independent evolutions. We refer the reader to [27] for rate-independent problems with convex energies and discontinuous inputs, and, in more specific applied contexts, to [16] for elastoplasticity with softening, to [19] for general material models with nonconvex elastic energies, to [17] for a non-associative plasticity model for cam clay, and [23,24,51] for crack propagation. Since the energy functionals involved in such applications are usually nonsmooth and nonconvex, the passage to the limit mostly relies on lower semicontinuity arguments. Let us illustrate the latter in the prototypical case (DN_ε). The key observation is that (DN_ε) is equivalent (see the discussion in Sect. 2.4) to the *ε-energy identity*

$$\mathcal{E}_t(u_\varepsilon(t)) + \int_0^t \Big(|u_\varepsilon'(s)|\,\mathrm{d}s + \frac{\varepsilon}{2}\|u_\varepsilon'(s)\|^2 + \frac{1}{2\varepsilon}\mathrm{dist}_*\big(-\mathrm{D}\mathcal{E}_s(u_\varepsilon(s)),K^*\big)^2\Big)\mathrm{d}s = \mathcal{E}_0(u(0)) + \int_0^t \partial_t\mathcal{E}_s(u_\varepsilon(s))\mathrm{d}s \tag{1.2}$$

for all $t\in[0,T]$, where the term

$$\mathrm{dist}_*\big(-\mathrm{D}\mathcal{E}_t(u(t)),K^*\big) := \min_{z\in K^*}\|-\mathrm{D}\mathcal{E}_t(u(t))-z\|_* \quad \text{with} \quad K^* = \big\{z\in X^* : |z|_*\leq 1\big\}, \tag{1.3}$$

measures the distance, with respect to the dual norm $\|\cdot\|_*$, of $-\mathrm{D}\mathcal{E}_t(u(t))$ from the set K^*. The term defined in (1.3) is penalized in (1.2) by the coefficient $1/(2\varepsilon)$. Thus, passing to the limit in (1.2) as $\varepsilon\downarrow 0$, one finds

$$\mathrm{dist}_*(-\mathrm{D}\mathcal{E}_t(u(t)),K^*) = 0 \qquad \text{for a.a. } t\in(0,T).$$

Hence,

$$-\mathrm{D}\mathcal{E}_t(u(t)) \in K^*, \quad i.e. \quad |-\mathrm{D}\mathcal{E}_t(u(t))|_* \leq 1 \qquad \text{for a.a. } t\in(0,T), \tag{1.4}$$

which is a *local version* of the global stability (S). Furthermore, *via* lower semicontinuity (1.2) yields the *energy inequality*

$$\mathcal{E}_t(u(t)) + \mathrm{Var}(u;[0,t]) \leq \mathcal{E}_0(u(0)) + \int_0^t \partial_t\mathcal{E}_s(u(s))\,\mathrm{d}s \quad \text{for all } t\in[0,T]. \tag{1.5}$$

Conditions (1.4)–(1.5) give raise to the notion of *local solution* of the rate-independent system (DN_0).

While the local stability (1.4) is physically more realistic than (S), its combination with the energy inequality (1.5) turns out to provide an unsatisfactory description of the solution at jumps (see the discussion after

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Prop. 2.7 and in [42], Sect. 5.2). In order to capture the jump dynamics, the energetic behavior of the system in a jump regime has to be revealed. From this perspective, it seems to be crucial to recover from (1.2), as $\varepsilon \downarrow 0$, an *energy identity*, rather than an energy inequality. Thus, the passage to the limit has to somehow keep track of the limit of the term

$$\int_0^t \left(\frac{\varepsilon}{2} \|u_\varepsilon'(s)\|^2 + \frac{1}{2\varepsilon} \mathrm{dist}_*\big(-\mathrm{D}\mathcal{E}_s(u_\varepsilon(s)), K^*\big)^2 \right) \mathrm{d}s,$$

which in fact encodes the contribution of the viscous dissipation $\frac{\varepsilon}{2}\|u_\varepsilon'\|^2$, completely missing in (1.5).

BV solutions. Moving from these considerations, it is natural to introduce the *vanishing-viscosity contact potential* (which is related to the bipotential discussed in [7], see Sect. 3.1) induced by Ψ_ε, *i.e.* the quantity

$$\begin{aligned} \mathfrak{p}(v,w) := \inf_{\varepsilon>0} \Big(\Psi_\varepsilon(v) + \Psi_\varepsilon^*(w) \Big) &= \inf_{\varepsilon>0} \left(|v| + \frac{\varepsilon}{2}\|v\|^2 + \frac{1}{2\varepsilon}\mathrm{dist}_*^2(w, K^*) \right) \\ &= |v| + \|v\| \, \mathrm{dist}_*(w, K^*) \quad \text{for } v \in X,\ w \in X^*. \end{aligned} \tag{1.6}$$

Then, the ε-energy identity (1.2) yields the inequality

$$\mathcal{E}_t(u_\varepsilon(t)) + \int_0^t \mathfrak{p}\left(u_\varepsilon'(s), -\mathrm{D}\mathcal{E}_s(u_\varepsilon(s))\right) \mathrm{d}s \le \mathcal{E}_0(u(0)) + \int_0^t \partial_t \mathcal{E}_s(u_\varepsilon(s)) \mathrm{d}s, \tag{1.7}$$

see Section 3.1. Passing to the limit in (1.7), in Theorem 4.9 we shall prove that, up to a subsequence, the solutions (u_ε) of the viscous equation (DN_ε) converge, as $\varepsilon \downarrow 0$, to a curve $u \in \mathrm{BV}([0,T];X)$ satisfying the local stability (1.4) and the following improved *energy inequality*

$$\mathcal{E}_t(u(t)) + \mathrm{Var}_{\mathfrak{p},\mathcal{E}}(u;[0,t]) \le \mathcal{E}_0(u(0)) + \int_0^t \partial_t \mathcal{E}_s(u(s)) \, \mathrm{d}s. \tag{1.8}$$

Without going into details (see Def. 3.4 later on), we may point out that (1.8) features a notion of (pseudo)-total variation (denoted by $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$) induced by the vanishing-viscosity contact potential $\mathfrak{p}$ (1.6) and the energy $\mathcal{E}$. The main novelty is that a BV-curve obeying the local stability condition (1.4) always satisfies the opposite inequality in (1.8), which yields the energy balance

$$\mathcal{E}_t(u(t)) + \mathrm{Var}_{\mathfrak{p},\mathcal{E}}(u;[0,t]) = \mathcal{E}_0(u(0)) + \int_0^t \partial_t \mathcal{E}_s(u(s)) \, \mathrm{d}s. \tag{$\mathrm{E}_{\mathfrak{p},\mathcal{E}}$}$$

In fact, $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$ provides a finer description of the dissipation $\Delta_{\mathfrak{p},\mathcal{E}}$ of u along a jump between two values u_- and u_+ at time t. It involves not only the quantity $|u_+ - u_-|$ related to the dissipation potential (1.1), but also the viscous contribution induced by the vanishing-viscosity contact potential $\mathfrak{p}$ through the formula

$$\begin{aligned} \Delta_{\mathfrak{p},\mathcal{E}}(t;u_-,u_+) := \inf \Bigg\{ &\int_{r_0}^{r_1} \mathfrak{p}(\dot\vartheta(r), -\mathrm{D}\mathcal{E}_t(\vartheta(r))) \, \mathrm{d}r : \\ &\vartheta \in \mathrm{AC}([r_0,r_1];X),\ \vartheta(r_0) = u_-,\ \vartheta(r_1) = u_+ \Bigg\}. \end{aligned} \tag{1.9}$$

By a rescaling technique, it is possible to show that, in a jump point, the system may switch to a *viscous behavior*, which is in fact reminiscent of the viscous approximation (DN_ε). In particular, when the jump point is of viscous type, the infimum in (1.9) is attained and the states u_- and u_+ are connected by some transition curve $\vartheta : [r_0, r_1] \to X$, fulfilling the *viscous* doubly nonlinear equation

$$\partial\Psi_0(\vartheta'(r)) + \vartheta'(r) + \mathrm{D}\mathcal{E}_t(\vartheta(r)) \ni 0 \quad \text{in } X^* \quad \text{for a.a. } r \in (r_0, r_1)$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

in the case the norm $\|\cdot\|$ is Euclidean and we use its differential to identify X with X^*. The combination of (1.4) and (1.8) yields the notion of *BV solution* to the rate-independent system $(X, \mathcal{E}, \mathfrak{p})$. This concept was first introduced in [42], in the case the ambient space X is a finite-dimensional *manifold* $\mathcal{X}$, and both the rate-independent and the viscous approximating dissipations depend on one single Finsler distance on $\mathcal{X}$. In this paper, while keeping to a Banach framework, we shall considerably broaden the class of rate-independent and viscous dissipation functionals, *cf.* Remark 2.3. Moreover, the notion of BV solution shall be presented here in a more compact form than in [42], amenable to a finer analysis and, hopefully, to further generalizations.

Let us now briefly comment on our main results. First of all, we are going to show in Theorems 4.3 and 4.7 that the concept of BV solution completely encompasses the solution behavior in both a purely rate-independent, non-jumping regime, and in jump regimes, where the competition between dry-friction and viscous effects is highlighted. Indeed, from (1.4) and (1.8) it is possible to deduce suitable energy balances at jumps (*cf.* conditions ($\mathrm{J_{BV}}$) in Thm. 4.3).

Then, in Theorem 4.9 we shall prove that, along a subsequence, the viscous approximations arising from (DN_ε) converge as $\varepsilon \downarrow 0$ to a BV solution. Next, our second main result, Theorem 4.10, states that, up to a subsequence, also the discrete solutions $\mathrm{U}_{\tau,\varepsilon}$ constructed *via* the ε-discretization scheme ($\mathrm{IP}_{\tau,\varepsilon}$) converge to a BV solution $u \in \mathrm{BV}([0,T];X)$ of (DN_0) as $\varepsilon \downarrow 0$ and $\tau \downarrow 0$ simultaneously, provided that the respective convergence rates are such that

$$\lim_{\varepsilon,\,\tau\downarrow 0} \frac{\varepsilon}{\tau} = +\infty.$$

For the proof of these results we present, in Section 5, the notion of *parameterized rate-independent solutions*, which was introduced in [18] and further developed in [40,42]. Using an energetic-like arclength parameterization, the passage to the vanishing viscosity limit can be handled with straightforward lower semicontinuity results, relying on an *augmented vanishing-viscosity contact potential* $\mathfrak{P}$, see Proposition 6.2. However, to achieve the main goal of proving direct convergence of the viscous solutions u_ε (without reparameterization), a much more refined lower semicontinuity result has to be established in Section 6.3.

The main idea of parameterized rate-independent solutions is to suitably reparameterize the approximate viscous curves (u_ε) in order to capture, in the vanishing-viscosity limit, the viscous transition paths at jump points. This leads to performing an asymptotic analysis as $\varepsilon \downarrow 0$ of the graphs of the functions u_ε, in the extended phase space $[0,T]\times X$. For every $\varepsilon > 0$ the graph of u_ε can be parameterized by a couple of functions $(\mathsf{t}_\varepsilon, \mathsf{u}_\varepsilon)$, where t_ε is the (strictly increasing) rescaling function and $\mathsf{u}_\varepsilon := u_\varepsilon \circ \mathsf{t}_\varepsilon$ is the rescaled solution. In Theorem 5.5 we assert that, up to a subsequence, the functions $(\mathsf{t}_\varepsilon, \mathsf{u}_\varepsilon)$ converge as $\varepsilon \downarrow 0$ to a parameterized rate-independent solution *i.e.*, a curve $(\mathsf{t},\mathsf{u}) : [0,\mathsf{S}] \to [0,T]\times X$ fulfilling

$$\mathsf{t} : [0,\mathsf{S}] \to [0,T] \text{ is nondecreasing, } \mathsf{t}'(s) + |\mathsf{u}'(s)| > 0 \quad \text{for a.a. } s \in (0,\mathsf{S}), \tag{1.10a}$$

$$\left.\begin{array}{rcl} \mathsf{t}'(s) > 0 & \Longrightarrow & |-\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))|_* \le 1 \\ |\mathsf{u}'(s)| > 0 & \Longrightarrow & |-\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))|_* \ge 1 \end{array}\right\} \quad \text{for a.a. } s \in (0,\mathsf{S}), \tag{1.10b}$$

and the *energy identity*

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\mathsf{t}(s),\mathsf{u}(s)) - \partial_t\mathcal{E}(\mathsf{t}(s),\mathsf{u}(s))\,\mathsf{t}'(s) = -|\mathsf{u}'(s)| - \|\mathsf{u}'(s)\|\mathrm{dist}_*(-\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)), K^*) \quad \text{for a.a. } s \in (0,\mathsf{S}). \tag{1.10c}$$

As already pointed out in [18,42], like the notion of BV solution, relations (1.10) comprise both the purely rate-independent evolution as well as the viscous transient regime at jumps. The jump regime corresponds to the case $-\mathrm{D}\mathcal{E}_\mathsf{t}(\mathsf{u}) \notin K^*$, where the system does not obey the local stability constraint (1.4) any longer and switches to viscous behavior, see also Remark 5.6 below.

As a matter of fact, Theorem 5.7 shows that parameterized rate-independent solutions may be viewed as the "continuous counterpart" to BV solutions. With a suitable transformation, it is possible to associate with every

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

parameterized rate-independent solution a BV solution, and conversely. One advantage of the parameterized notion is that it avoids the technicalities related to BV functions. For instance, it is more easily amenable to a stability analysis (*cf.* [42], Rem. 6). Furthermore, in [40] a refined vanishing-viscosity analysis has been developed based on a slightly different reparameterization technique. There, an infinite-dimensional (L^1, L^2)-framework is used, where (DN_ε) is replaced by a class of nonlinear evolutionary PDEs.

Generalizations and future developments. Above we have focused on dissipation functionals of the type (1.1) and $\Psi_\varepsilon(v) = |v| + \frac{\varepsilon}{2}\|v\|^2$ as in (DN_ε) for ease of presentation only, in order to highlight the main variational argument leading to the notion of BV solution. Indeed, the analysis developed in this paper is targeted to a general

$$\text{positively 1-homogeneous, convex dissipation } \Psi_0 : X \to [0, +\infty),$$

(*cf.* (2.1)), and considers a fairly wide class of approximate viscous dissipation functionals Ψ_ε, defined by conditions (Ψ.1)–(Ψ.3) in Section 2.3. Furthermore, at the price of just technical complications, our results could be extended to the case of a *Finsler-like* family of dissipation functionals $\Psi_\varepsilon(u, \cdot)$, depending on the state variable $u \in X$, and satisfying uniform bounds and Mosco-continuity with respect to u, see [42], Section 2 and [47], Sections 6 and 8.

The extension to *infinite-dimensional* ambient spaces and *nonsmooth* energies is crucial for application of the concept of BV solution to PDE systems modeling rate-independent evolutions in continuum mechanics. A first step in this direction is to generalize the known existence results for doubly nonlinear equations, driven by a viscous dissipation, to nonconvex and nonsmooth energy functionals in infinite dimensions. As shown in [46,47], in the nonsmooth and nonconvex case one can replace the energy differential $\mathrm{D}\mathcal{E}_t$ with a suitable notion of *subdifferential* $\partial\mathcal{E}_t$. Accordingly, instead of continuity of $\mathrm{D}\mathcal{E}_t$, one asks for closedness of the multivalued subdifferential $\partial\mathcal{E}_t$ in the sense of graphs. These ideas shall be further advanced in the forthcoming work [43]. Therein, exploiting techniques from nonsmooth analysis, we shall also tackle energies that do not depend smoothly on time (this is relevant for rate-independent applications, see *e.g.* [24,27]).

On the other hand, the requirement that the ambient space X is finite-dimensional could be replaced by suitable compactness (of the sublevels of the energy) and reflexivity assumptions. The latter topological requirement in fact ensures that X has the so-called Radon-Nikodým property, *i.e.* that absolutely continuous curves with values in X are almost everywhere differentiable. The vanishing-viscosity analysis in spaces which do not enjoy this property requires a subtler approach involving metric arguments (see *e.g.* [47], Sect. 7), or *ad hoc* stronger estimates [40]. See also [17,33] for some preliminary approaches to BV solutions for PDE problems.

Plan of the paper. Section 2 is devoted to an extended presentation of energetic and local solutions to rate-independent systems. In particular, after fixing the setup of the paper in Section 2.1, in Section 2.2 we recall the definition of energetic solution, show its differential characterization and the related variational time-incremental scheme. We develop the vanishing-viscosity approach in Sections 2.3 and 2.4, thus arriving at the notion of local solution (see Sect. 2.5), which also admits a differential characterization.

In Section 3 we introduce the concept of vanishing-viscosity contact potential and thoroughly analyze its properties, as well as the induced (pseudo-) total variation. With these ingredients, in Section 4 we present the notion of BV solution. We show that BV solutions also admit a differential characterization, and that they provide a careful description of the energetic behavior of the system, see Section 4.2. Then, in Section 4.3 we state our main results on the convergence of the viscous solutions u_ε towards BV solutions u.

While Section 5 develops the fundamental tool of parameterized rate-independent solutions, Section 6 contains some technical results which lie at the core of our theory. Appendices A and B provide further details on the vanishing-viscosity contact potential and the optimal transitions, respectively.

2. Energetic *versus* local solutions, and their viscous regularizations

In this section, we will briefly recall the notion of *energetic solutions* and show that viscous regularizations of rate-independent systems give raise to *local solutions*. The latter concept is rather weak, since it still contains

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

all energetic solutions, and it is the aim of this work to identify more properties of the limits obtained in the vanishing-viscosity limit.

2.1. **Rate-independent setting: dissipation and energy functionals**

We let

$$(X, \|\cdot\|_X) \quad \text{be a finite-dimensional normed vector space,}$$

endowed with a gauge function Ψ_0, namely a

$$\text{non-degenerate, positively 1-homogeneous, convex dissipation } \Psi_0 : X \to [0, +\infty), \tag{2.1}$$

i.e. Ψ_0 satisfies $\Psi_0(v) > 0$ if $v \neq 0$, and

$$\Psi_0(v_1 + v_2) \leq \Psi_0(v_1) + \Psi_0(v_2), \quad \Psi_0(\lambda v) = \lambda \Psi_0(v) \quad \text{for every } \lambda \geq 0,\ v, v_1, v_2 \in X.$$

In particular, there exists a constant $\eta > 0$ such that

$$\eta^{-1}\|v\|_X \leq \Psi_0(v) \leq \eta \|v\|_X \quad \text{for every } v \in X.$$

Since Ψ_0 is 1-homogeneous, its subdifferential $\partial\Psi_0 : X \rightrightarrows X^*$ can be characterized by

$$\partial\Psi_0(v) := \Big\{ w \in X : \langle w, z\rangle \leq \Psi_0(z) \text{ for every } z \in X, \quad \langle w, v\rangle = \Psi_0(v) \Big\} \subset X^*. \tag{2.2}$$

Moreover, $\partial\Psi_0$ takes its values in the convex set $K^* \subset X^*$, given by

$$K^* = \partial\Psi_0(0) := \big\{ w \in X^* : \langle w, z\rangle \leq \Psi_0(z) \quad \forall\, z \in X \big\} \supset \partial\Psi_0(v) \quad \text{for every } v \in X. \tag{2.3}$$

The following properties are well-known (see *e.g.* [45]):

K1. K^* is the proper domain of the Legendre transform Ψ_0^* of Ψ_0, since

$$\Psi_0^*(w) = \mathrm{I}_{K^*}(w) = \begin{cases} 0 & \text{if } w \in K^*, \\ +\infty & \text{otherwise.} \end{cases} \tag{2.4}$$

K2. Ψ_0 is the support function of K^*, since

$$\Psi_0(v) = \sup_{w \in K^*} \langle w, v\rangle \quad \text{for every } v \in X, \tag{2.5}$$

and K^* is the polar set of the unit ball $K := \big\{ v \in X : \Psi_0(v) \leq 1 \big\}$ associated with Ψ_0.

K3. K^* is the unit ball of the support function Ψ_{0*} of K:

$$K^* = \big\{ w \in X^* : \Psi_{0*}(w) \leq 1 \big\}, \quad \text{with} \quad \Psi_{0*}(w) = \sup_{v \in K} \langle w, v\rangle = \sup_{v \neq 0} \frac{\langle w, v\rangle}{\Psi_0(v)}. \tag{2.6}$$

K4. In the even case (*i.e.*, when $\Psi_0(v) = \Psi_0(-v)$ for all $v \in X$) Ψ_0 is an equivalent norm for X, Ψ_{0*} is its dual norm, and K and K^* are their respective unit balls.

Further, we consider a *smooth* energy functional

$$\mathcal{E} \in \mathrm{C}^1([0,T] \times X),$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

which we suppose bounded from below (and thus, up to the addition of a constant, nonnegative) and with energy-bounded time derivative

$$\exists\, C > 0\ \forall\, (t,u) \in [0,T] \times X\ : \qquad |\partial_t \mathcal{E}_t(u)| \le C \mathcal{E}_t(u). \tag{2.7}$$

The rate-independent system associated with the energy functional $\mathcal{E}$ and the dissipation potential Ψ_0 can formally be described by the *rate-independent doubly nonlinear* differential inclusion

$$\partial \Psi_0(u'(t)) + \mathrm{D}\mathcal{E}_t(u(t)) \ni 0 \quad \text{in } X^* \quad \text{for a.a. } t \in (0,T). \tag{DN$_0$}$$

As already mentioned in the introduction, for nonconvex energies the solutions to (DN$_0$) may exhibit discontinuities in time. The first weak solvability notion for (DN$_0$) is the concept of *energetic solution* to the rate-independent system $(X, \mathcal{E}, \Psi_0)$, see [37,38,41] and the survey [32]. In what follows, we shall use triplets of the form $(X, \mathcal{E}, \Xi)$ to indicate an evolutionary system on the state space X that is driven by the energy functional $\mathcal{E}$ and the dissipation functional Ξ. According to the choice of Ξ, we shall address different solution notions.

2.2. Energetic solutions and variational incremental scheme

Definition 2.1 (energetic solution)**.** A curve $u \in \mathrm{BV}([0,T];X)$ is an energetic solution of the *rate-independent system* $(X, \mathcal{E}, \Psi_0)$ if for all $t \in [0,T]$ the *global stability* (S) and the *energy balance* (E) hold:

$$\forall\, z \in X : \qquad \mathcal{E}_t(u(t)) \le \mathcal{E}_t(z) + \Psi_0(z - u(t)), \tag{S}$$

$$\mathcal{E}_t(u(t)) + \mathrm{Var}_{\Psi_0}(u;[0,t]) = \mathcal{E}_0(u(0)) + \int_0^t \partial_t \mathcal{E}_s(u(s))\,\mathrm{d}s. \tag{E}$$

BV functions. Hereafter, we shall consider functions of bounded variation *pointwise defined at every time* $t \in [0,T]$, such that the *pointwise* total variation with respect to Ψ_0 (any equivalent norm of X can be chosen) $\mathrm{Var}_{\Psi_0}(u;[0,T])$ is finite, where

$$\mathrm{Var}_{\Psi_0}(u;[a,b]) := \sup\left\{ \sum_{m=1}^{M} \Psi_0\big(u(t_m) - u(t_{m-1})\big) : a = t_0 < t_1 < \cdots < t_{M-1} < t_M = b \right\}.$$

Notice that a function $u \in \mathrm{BV}([0,T];X)$ admits left and right limits at every $t \in [0,T]$:

$$u(t_-) := \lim_{s \uparrow t} u(s), \quad u(t_+) := \lim_{s \downarrow t} u(s) \ \text{ with the convention } u(0_-) := u(0),\ u(T_+) := u(T). \tag{2.8}$$

Its *pointwise* jump set J_u is the at most countable set defined by

$$\mathrm{J}_u := \big\{ t \in [0,T] : u(t_-) \ne u(t) \text{ or } u(t) \ne u(t_+) \big\} \supset \text{ess-J}_u := \big\{ t \in [0,T] : u(t_-) \ne u(t_+) \big\}. \tag{2.9}$$

We denote by u' the distributional derivative of u, and recall that u' is a Radon vector measure with finite total variation $|u'|$. It is well known [3] that u' can be decomposed into the sum of the three mutually singular measures

$$u' = u'_{\mathscr{L}} + u'_{\mathrm{Ca}} + u'_{\mathrm{J}}, \quad u'_{\mathscr{L}} = \dot{u}\,\mathscr{L}^1, \quad u'_{\mathrm{co}} := u'_{\mathscr{L}} + u'_{\mathrm{Ca}}. \tag{2.10}$$

Here, $u'_{\mathscr{L}}$ is the absolutely continuous part with respect to the Lebesgue measure $\mathscr{L}^1$, whose Lebesgue density $\dot{u}$ is the usual pointwise (and $\mathscr{L}^1$-a.e. defined) derivative, u'_{J} is a discrete measure concentrated on ess-$\mathrm{J}_u \subset \mathrm{J}_u$, and u'_{Ca} is the so-called Cantor part, still satisfying $u'_{\mathrm{Ca}}(\{t\}) = 0$ for every $t \in [0,T]$. Therefore $u'_{\mathrm{co}} = u'_{\mathscr{L}} + u'_{\mathrm{Ca}}$ is the diffuse part of the measure, which does not charge J_u. In the following, it will be useful to use a nonnegative

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

and diffuse reference measure μ on $(0,T)$, such that $\mathscr{L}^1$ and $u'_{\rm Ca}$ are absolutely continuous w.r.t. μ: just to fix our ideas, we set

$$\mu := \mathscr{L}^1 + |u'_{\rm Ca}|. \tag{2.11}$$

With a slight abuse of notation, for every $(a,b) \subset (0,T)$ we denote by $\int_a^b \mathrm{d}\Psi_0(u'_{\rm co})$ the integral

$$\int_a^b \mathrm{d}\Psi_0(u'_{\rm co}) := \int_a^b \Psi_0\left(\frac{\mathrm{d}u'_{\rm co}}{\mathrm{d}\mu}\right)\mathrm{d}\mu = \int_a^b \Psi_0(\dot{u})\,\mathrm{d}\mathscr{L}^1 + \int_a^b \Psi_0\left(\frac{\mathrm{d}u'_{\rm Ca}}{\mathrm{d}|u'_{\rm Ca}|}\right)\mathrm{d}|u'_{\rm Ca}|. \tag{2.12}$$

Since Ψ_0 is 1-homogeneous, the above integral is independent of μ, provided $u'_{\rm co}$ is absolutely continuous w.r.t. μ.

Towards a differential characterization of energetic solutions. Let us first of all point out that (S) is stronger than the local stability condition

$$-\mathrm{D}\mathcal{E}_t(u(t)) \in K^* \quad \text{for every } t \in [0,T]\setminus \mathrm{J}_u, \tag{$\mathrm{S}_{\rm loc}$}$$

which can be formally deduced from (DN_0) and (2.3). Indeed, the global stability (S) yields for every $z = u(t)+hv \in X$ and $h>0$

$$\langle -\mathrm{D}\mathcal{E}_t(u(t)), hv\rangle + o(|h|) \le \mathcal{E}_t(u(t)) - \mathcal{E}_t(u(t)+hv) \le h\Psi_0(v)$$

and therefore, dividing by h and passing to the limit as $h\downarrow 0$, one gets

$$\langle -\mathrm{D}\mathcal{E}_t(u(t)), v\rangle \le \Psi_0(v) \quad \text{for every } v\in X,$$

so that ($\mathrm{S}_{\rm loc}$) holds by (2.6). We obtain more insight into (E) by representing the Ψ_0-variation $\mathrm{Var}_{\Psi_0}(u;[a,b])$ in terms of the distributional derivative u' of u. In fact, recalling (2.10)–(2.12), we have

$$\mathrm{Var}_{\Psi_0}(u;[a,b]) := \int_a^b \mathrm{d}\Psi_0(u'_{\rm co}) + \mathrm{Jmp}_{\Psi_0}(u;[a,b]),$$

where the jump contribution $\mathrm{Jmp}_{\Psi_0}(u;[a,b])$ can be described by

$$\begin{aligned}\mathrm{Jmp}_{\Psi_0}(u;[a,b]) &:= \Delta_{\Psi_0}(u(a),u(a_+)) + \Delta_{\Psi_0}(u(b_-),u(b))\\ &\quad + \sum_{t\in \mathrm{J}_u\cap(a,b)} \big(\Delta_{\Psi_0}(u(t_-),u(t)) + \Delta_{\Psi_0}(u(t),u(t_+))\big),\\ \text{with } \Delta_{\Psi_0}(v_0,v_1) &:= \Psi_0(v_1-v_0).\end{aligned} \tag{2.13}$$

As usual in rate-independent evolutionary problems, u is pointwise *everywhere* defined and the jump term $\mathrm{Jmp}_{\Psi_0}(u;[\cdot,\cdot])$ takes into account the value of u at every time $t\in J_u$. Therefore, if u is not continuous at t, this part may yield a strictly bigger contribution than the total mass of the distributional jump measure $u'_{\rm J}$ (which gives rise to the so-called *essential variation*).

The following result provides an equivalent characterization of energetic solutions: besides the global stability condition (S), it involves a BV formulation of the differential inclusion (DN_0) (*cf.* the *subdifferential formulation* of [38]) and a *jump condition* at any jump point of u.

Proposition 2.2. *A curve $u\in \mathrm{BV}([0,T];X)$ satisfying the global stability condition (S) is an energetic solution of the rate-independent system $(X,\mathcal{E},\Psi_0)$ if and only if it satisfies the differential inclusion*

$$\partial\Psi_0\left(\frac{\mathrm{d}u'_{\rm co}}{d\mu}(t)\right) + \mathrm{D}\mathcal{E}_t(u(t)) \ni 0 \quad \textit{for } \mu\textit{-a.e. } t\in[0,T], \ \textit{where } \mu := \mathscr{L}^1 + |u'_{\rm Ca}|, \tag{$\mathrm{DN}_{0,\mathrm{BV}}$}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

and the jump conditions

$$\mathcal{E}_t(u(t)) - \mathcal{E}_t(u(t_-)) = -\Delta_{\Psi_0}(u(t_-), u(t)), \mathcal{E}_t(u(t_+)) - \mathcal{E}_t(u(t)) = -\Delta_{\Psi_0}(u(t), u(t_+)),$$
$$\mathcal{E}_t(u(t_+)) - \mathcal{E}_t(u(t_-)) = -\Delta_{\Psi_0}(u(t_-), u(t_+)) \tag{J_{ener}}$$

for every $t \in J_u$ *(recall convention* (2.8) *in the case* $t = 0, T$*).*

We shall simply sketch the proof, referring to the arguments for the forthcoming Proposition 2.7 for all details.

Proof. By the additivity property of the total variation $\mathrm{Var}_{\Psi_0}(u; [\cdot, \cdot])$, (E) yields for every $0 \le t_0 < t_1 \le T$

$$\mathrm{Var}_{\Psi_0}(u; [t_0, t_1]) + \mathcal{E}_{t_1}(u(t_1)) = \mathcal{E}_{t_0}(u(t_0)) + \int_{t_0}^{t_1} \partial_t \mathcal{E}_t(u(t))\, \mathrm{d}t. \tag{E'}$$

Arguing as in the proof of Proposition 2.7 later on, one can see that the global stability (S) and (E') yield the differential inclusion ($DN_{0,BV}$) and conditions (J_{ener}).

Conversely, repeating the arguments of Proposition 2.7 one can verify that ($DN_{0,BV}$) and (J_{ener}) imply (E). □

Incremental minimization scheme. Existence of energetic solutions can be proved by solving a minimization scheme, which is also interesting as construction of an effective approximation of the solutions.

For a given time-step $\tau > 0$ we consider a uniform partition (for simplicity) $0 = t_0 < t_1 < \cdots < t_{N-1} < T \le t_N$, $t_n := n\tau$, of the time interval $[0, T]$, and an initial value $\mathrm{U}_\tau^0 \approx u_0$. In order to find good approximations of $\mathrm{U}_\tau^n \approx u(t_n)$, we solve the incremental minimization scheme

$$\text{find } \mathrm{U}_\tau^1, \dots, \mathrm{U}_\tau^N \quad \text{such that} \quad \mathrm{U}_\tau^n \in \operatorname*{Argmin}_{\mathrm{U} \in X} \Big\{ \Psi_0(\mathrm{U} - \mathrm{U}_\tau^{n-1}) + \mathcal{E}_{t_n}(\mathrm{U}) \Big\}. \tag{IP_τ}$$

Setting

$$\overline{\mathrm{U}}_\tau(t) := \mathrm{U}_\tau^n \quad \text{if } t \in (t_{n-1}, t_n], \tag{2.14}$$

it is possible to find a suitable vanishing sequence of step sizes $\tau_k \downarrow 0$ (see, *e.g.*, [32,38] for all calculations), such that

$$\exists \lim_{k \to +\infty} \overline{\mathrm{U}}_{\tau_k}(t) =: u(t) \quad \text{for every } t \in [0, T],$$

and u is an energetic solution of $(X, \mathcal{E}, \Psi_0)$.

2.3. **Viscous approximations of rate-independent systems**

In the present paper we want to study a different approach to approximate and solve (DN_0): the main idea is to replace the linearly growing dissipation potential Ψ_0 with a suitable *convex and superlinear* "viscous" regularization $\Psi_\varepsilon : X \to [0, +\infty)$ of Ψ_0, depending on a "small" parameter $\varepsilon > 0$ and "converging" to Ψ_0 in a suitable sense as $\varepsilon \downarrow 0$. Solving the doubly nonlinear differential inclusion (we use the notation $\dot{u}$ for the time derivative of u, when u is absolutely continuous)

$$\partial \Psi_\varepsilon(\dot{u}_\varepsilon(t)) + \mathrm{D}\mathcal{E}_t(u_\varepsilon(t)) \ni 0 \quad \text{in } X^* \quad \text{for a.a. } t \in (0, T), \tag{DN_ε}$$

one can consider the sequence (u_ε) as a good approximation of the solution u of (DN_0) as $\varepsilon \downarrow 0$.

There is also a natural discrete counterpart to (DN_ε), which regularizes the incremental minimization problem (IP_τ). We simply substitute Ψ_0 by Ψ_ε in (IP_τ), recalling that now the time-step τ should explicitly appear, since Ψ_ε is no longer 1-homogeneous. The viscous incremental problem is therefore

$$\text{find } \mathrm{U}_{\tau,\varepsilon}^1, \dots, \mathrm{U}_{\tau,\varepsilon}^N \quad \text{such that} \quad \mathrm{U}_{\tau,\varepsilon}^n \in \operatorname*{Argmin}_{\mathrm{U} \in X} \left\{ \tau \Psi_\varepsilon \left(\frac{\mathrm{U} - \mathrm{U}_{\tau,\varepsilon}^{n-1}}{\tau} \right) + \mathcal{E}_{t_n}(\mathrm{U}) \right\}. \tag{$IP_{\tau,\varepsilon}$}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

As in (2.14), we set

$$\overline{\mathrm{U}}_{\tau,\varepsilon}(t) := \mathrm{U}^n_{\tau,\varepsilon} \quad \text{for } t \in (t_{n-1}, t_n]$$

and study the limit of the discrete solutions $\overline{\mathrm{U}}_{\tau,\varepsilon}$ for $(\tau,\varepsilon) \to 0$, imposing an additional restriction on the behavior of the quotient ε/τ, see Theorem 4.10 below.

The choice of the viscosity approximation Ψ_ε. In this work we restrict to the particular case that the potential Ψ_ε is obtained by rescaling from a given Ψ as follows:

$$\Psi : X \to [0,+\infty) \text{ is convex}, \quad \Psi(0) = 0, \qquad \lim_{\|v\|_X \uparrow +\infty} \frac{\Psi(v)}{\|v\|_X} = +\infty, \tag{Ψ.1}$$

$$\Psi_\varepsilon(v) := \varepsilon^{-1}\Psi(\varepsilon v) \quad \text{for every } v \in X,\ \varepsilon > 0, \tag{Ψ.2}$$

$$\Psi_0(v) = \lim_{\varepsilon \downarrow 0} \Psi_\varepsilon(v) = \lim_{\varepsilon \downarrow 0} \varepsilon^{-1}\Psi(\varepsilon v) \quad \text{for every } v \in X. \tag{Ψ.3}$$

The last relation shows that Ψ_0 can easily be recovered from Ψ_ε. In fact, by convexity of Ψ and the fact that $\Psi(0) = 0$, the map $\varepsilon \mapsto \varepsilon^{-1}\Psi(\varepsilon v)$ is nondecreasing for all $v \in X$. Hence,

$$\Psi_0(v) \le \Psi_\varepsilon(v) \quad \text{for all } v \in X \text{ and all } \varepsilon > 0. \tag{2.15}$$

Furthermore, the superlinearity in (Ψ.1) implies that

$$v \mapsto \partial\Psi_\varepsilon(v) := \partial\Psi(\varepsilon v) \quad \text{is a surjective map.}$$

Denoting by $\Psi^*, \Psi^*_\varepsilon$ the conjugate functions of Ψ and Ψ_ε, we also have

$$0 = \Psi^*(0) \le \Psi^*(\xi) < +\infty, \qquad \Psi^*_\varepsilon(\xi) = \varepsilon^{-1}\Psi^*(\xi) \quad \text{for every } \xi \in X^*,\ \varepsilon > 0, \tag{2.16}$$

and, by (Ψ.3), for every $\xi \in X^*$

$$\Psi^*(\xi) = 0 \quad \Leftrightarrow \quad \Psi^*_0(\xi) = 0 \quad \Leftrightarrow \quad \xi \in K^*. \tag{2.17}$$

Due to (2.15), there holds

$$\Psi^*_\varepsilon(\xi) \le \Psi^*_0(\xi) \quad \text{for all } v \in X,\ \varepsilon > 0. \tag{2.18}$$

The following examples show that the choice (Ψ.2) still provides a great flexibility and covers several interesting cases.

Example 2.3.

(1) Ψ_0-viscosity: The simplest example, still absolutely nontrivial [42], is to consider

$$\begin{gathered}\Psi(v) := \Psi_0(v) + \frac{1}{2}\big(\Psi_0(v)\big)^2, \quad \Psi_\varepsilon(v) := \Psi_0(v) + \frac{\varepsilon}{2}\big(\Psi_0(v)\big)^2, \\ \partial\Psi_\varepsilon(v) = \Big(1 + \varepsilon\Psi_0(v)\Big)\partial\Psi_0(v).\end{gathered} \tag{2.19}$$

A similar regularization can be obtained by choosing a real convex and superlinear function $F_V : [0,+\infty) \to [0,+\infty)$, with $F_V(0) = F_V'(0) = 0$, and setting

$$\Psi(v) := \Psi_0(v) + F_V(\Psi_0(v)) = F(\Psi_0(v)) \quad \text{with} \quad F(r) := r + F_V(r). \tag{2.20}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

(2) Quadratic or p-viscosity induced by a norm $\|\cdot\|$: The most interesting case involves an arbitrary norm $\|\cdot\|$ on X and considers for $p>1$

$$\Psi(v)=\Psi_0(v)+\frac{1}{p}\|v\|^p,\quad \Psi_\varepsilon(v)=\Psi_0(v)+\frac{\varepsilon^{p-1}}{p}\|v\|^p,\quad \partial\Psi_\varepsilon(v)=\partial\Psi_0(v)+\varepsilon^{p-1}J_p(v), \tag{2.21}$$

where J_p is the p-duality map associated with $\|\cdot\|$. In particular, if $\|\cdot\|$ is a Hilbertian norm and $p=2$, then J_2 is the Riesz isomorphism and we can choose $J_2(v)=v$ by identifying X with X^*. Hence, (DN_ε) reads

$$\partial\Psi_\varepsilon(\dot u_\varepsilon(t))+\varepsilon\dot u_\varepsilon(t)+\mathrm{D}\mathcal{E}_t(u_\varepsilon(t))\ni 0\quad \text{in } X^*\quad \text{for a.a. } t\in(0,T).$$

For the general p-viscosity, the incremental problem ($\mathrm{IP}_{\tau,\varepsilon}$) looks for $\mathrm{U}^n_{\tau,\varepsilon}$ which recursively minimizes

$$\mathrm{U}\mapsto\Psi_0(\mathrm{U}-\mathrm{U}^{n-1}_{\tau,\varepsilon})+\frac{\varepsilon}{p\tau}\|\mathrm{U}-\mathrm{U}^{n-1}_{\varepsilon,\tau}\|^p+\mathcal{E}_{t_n}(\mathrm{U}).$$

(3) Additive viscosity: More generally, we can choose a convex "viscous" potential $\Psi_V: X\to[0,+\infty)$ satisfying

$$\lim_{\varepsilon\downarrow0}\varepsilon^{-1}\Psi_V(\varepsilon v)=0,\quad \lim_{\lambda\uparrow+\infty}\lambda^{-1}\Psi_V(\lambda v)=+\infty\quad \text{for all } v\in X, \tag{2.22}$$

and set

$$\Psi(v):=\Psi_0(v)+\Psi_V(v),\quad \Psi_\varepsilon(v):=\Psi_0(v)+\varepsilon^{-1}\Psi_V(\varepsilon v),\quad \partial\Psi_\varepsilon(v)=\partial\Psi_0+\partial\Psi_V(\varepsilon v). \tag{2.23}$$

Case (3) of the above example, which contains the two earlier cases, is the typical situation that motivates our investigation as it has a natural physical interpretation. However, other vanishing-viscosity limits might be mathematically relevant, like $\Psi_\varepsilon(v)=\frac{1}{1+\varepsilon}\|v\|^{1+\varepsilon}_{1+\varepsilon}$, which is not of the form ($\Psi.2$).

2.4. **Viscous energy identity**

Since Ψ has a superlinear growth, the results of [8,9] ensure that for every $\varepsilon>0$ and initial datum $u_0\in X$ there exists at least one solution $u_\varepsilon\in\mathrm{AC}([0,T];X)$ to equation (DN_ε), fulfilling the Cauchy condition $u_\varepsilon(0)=u_0$.

In order to capture its asymptotic behavior as $\varepsilon\downarrow0$, we split equation (DN_ε) in a simple system of two conditions, involving an auxiliary variable $w_\varepsilon:[0,T]\to X^*$ and a scalar function $p_\varepsilon:[0,T]\to\mathbb{R}$

$$\partial\Psi_\varepsilon(\dot u_\varepsilon(t))\ni w_\varepsilon\qquad \text{for a.a. } t\in(0,T), \tag{2.24a}$$

$$\mathrm{D}\mathcal{E}_t(u_\varepsilon(t))=-w_\varepsilon(t),\qquad \partial_t\mathcal{E}_t(u_\varepsilon(t))=-p_\varepsilon(t)\qquad \text{for all } t\in[0,T]. \tag{2.24b}$$

The classical characterization of the subdifferential of Ψ_ε yields that the first condition (2.24a) is equivalent to

$$\Psi_\varepsilon(\dot u_\varepsilon(t))+\Psi^*_\varepsilon(w_\varepsilon(t))=\langle w_\varepsilon(t),\dot u_\varepsilon(t)\rangle\qquad \text{for a.a. } t\in(0,T). \tag{2.25}$$

On the other hand, the chain rule for the C^1 functional $\mathcal{E}$ shows that along any absolutely continuous curve u_ε we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u_\varepsilon(t))=\langle\mathrm{D}\mathcal{E}_t(u_\varepsilon(t)),\dot u_\varepsilon(t)\rangle+\partial_t\mathcal{E}_t(u_\varepsilon(t))=-\langle w_\varepsilon(t),\dot u_\varepsilon(t)\rangle-p_\varepsilon(t)\quad \text{for a.a. } t\in(0,T). \tag{2.26}$$

Thus, equation (2.24a) is equivalent to the *energy identity*

$$\int_{t_0}^{t_1}\Big(\Psi_\varepsilon\left(\dot u_\varepsilon(r)\right)+\Psi^*_\varepsilon\left(w_\varepsilon(r)\right)+p_\varepsilon(r)\Big)\,\mathrm{d}r+\mathcal{E}_{t_1}(u_\varepsilon(t_1))=\mathcal{E}_{t_0}(u_\varepsilon(t_0)), \tag{2.27}$$

for every $0\le t_0\le t_1\le T$.

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Remark 2.4 (the role of Ψ_ε^*)**.** In the general case of additive viscosity (2.23) (where $\Psi(v) = \Psi_0(v) + \Psi_V(v)$), the inf-sup convolution formula [22], Theorem 2.3.2, yields

$$\Psi_\varepsilon^*(\xi) = \inf_{\substack{\xi_1+\xi_2=\xi \\ \xi_1,\xi_2\in X^*}} \left\{ \mathrm{I}_{K^*}(\xi_1) + \frac{1}{\varepsilon}\Psi_V^*(\xi_2) \right\} = \frac{1}{\varepsilon} \min_{z\in K^*} \Psi_V^*(\xi - z).$$

In particular, when $\Psi_V(\xi) := \frac{1}{2}\|v\|^2$ for some norm $\|\cdot\|$ of X, one finds

$$\Psi_\varepsilon^*(\xi) = \frac{1}{2\varepsilon} \min_{z\in K^*} \|\xi - z\|_*^2,$$

where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$. Thus, for all $\xi \in X^*$ the functional $2\varepsilon\Psi_\varepsilon^*(\xi)$ is the squared distance of ξ from K^* with respect to $\|\cdot\|_*$. This shows that, in the viscous regularized equation (DN_ε), the (local) stability condition $w(t) = -\mathrm{D}\mathcal{E}_t(u(t)) \in K^*$ has been replaced by the contribution of the penalizing term

$$\frac{1}{2\varepsilon}\int_0^T \min_{z\in K^*} \|-\mathrm{D}\mathcal{E}_t(u_\varepsilon(t)) - z\|_*^2 \,\mathrm{d}t$$

in the energy identity (2.27).

2.5. Pointwise limit of viscous approximations and local solutions

Using (2.7), it is not difficult to show that the viscous solutions u_ε of (DN_ε) satisfy the *a priori* bound

$$\int_0^T \Big(\Psi_\varepsilon(\dot u_\varepsilon(t)) + \Psi_\varepsilon^*(w_\varepsilon(t))\Big)\mathrm{d}t \le C, \quad \text{with } \ w_\varepsilon(t) = -\mathrm{D}\mathcal{E}_t(u_\varepsilon(t)) \ \text{ for all } t\in[0,T]. \tag{2.28}$$

Therefore, Helly's compactness theorem yields that, up to the extraction of a suitable subsequence, the sequence (u_ε) pointwise converges to a BV curve u, with $w_\varepsilon(t) \to w(t) = -\mathrm{D}\mathcal{E}_t(u(t))$ as $\varepsilon \downarrow 0$. By (Ψ.3), (2.15), and the explicit representation formula in (2.16), $\Psi_\varepsilon^*(\cdot)$ is converging monotonically increasing to $\Psi_0^* = \mathrm{I}_{K^*}$ as $\varepsilon \downarrow 0$, so that

$$\liminf_{\varepsilon\downarrow 0} \Psi_\varepsilon^*(w_\varepsilon(t)) \ge \Psi_0^*(w(t)) = \mathrm{I}_{K^*}(w(t)) = \begin{cases} 0 & \text{if } w(t)\in K^*, \\ +\infty & \text{otherwise.} \end{cases} \tag{2.29}$$

We infer that the limit curve u satisfies the (local) stability condition ($\mathrm{S}_{\mathrm{loc}}$). On the other hand, passing to the limit in (2.27) one gets the energy inequality

$$\mathcal{E}_{t_1}(u(t_1)) + \mathrm{Var}_{\Psi_0}(u;[t_0,t_1]) \le \mathcal{E}_{t_0}(u(t_0)) + \int_{t_0}^{t_1} \partial_t\mathcal{E}_t(u(t))\,\mathrm{d}t \quad \text{for } 0\le t_0 < t_1 \le T. \tag{$\mathrm{E}'_{\mathrm{ineq}}$}$$

The above discussion motivates the concept of *local solution*, first introduced in [51], Definition 3.1 (see also [42], Sect. 5.2, and the comparison in [33]).

Definition 2.5 (local solutions)**.** A curve $u \in \mathrm{BV}([0,T];X)$ is called a *local solution* of the rate-independent system $(X,\mathcal{E},\Psi_0)$ if it satisfies the *local stability* condition

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

and the *energy dissipation inequality* ($\mathrm{E}'_{\mathrm{ineq}}$).

As explained before Definition 2.5, the following result is an immediate consequence of the vanishing-viscosity limit.

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Lemma 2.6. *Pointwise limits $u : [0,T] \times X$ of subsequences of solutions u_ε of (DN_ε) are local solutions in the sense of Definition 2.5.*

Local solutions admit the following differential characterization.

Proposition 2.7 (differential characterization of local solutions)**.** *A curve $u \in \mathrm{BV}([0,T];X)$ is a* local solution *of the rate-independent system $(X, \mathcal{E}, \Psi_0)$ if and only if it satisfies the BV differential inclusion*

$$\partial\Psi_0\left(\frac{\mathrm{d}u'_{\mathrm{co}}}{d\mu}(t)\right) + \mathrm{D}\mathcal{E}_t(u(t)) \ni 0 \quad \text{for } \mu\text{-a.e. } t \in [0,T], \quad \mu := \mathscr{L}^1 + |u'_{\mathrm{Ca}}|, \tag{$\mathrm{DN}_{0,\mathrm{BV}}$}$$

and the jump inequalities

$$\begin{aligned} &\mathcal{E}_t(u(t)) - \mathcal{E}_t(u(t_-)) \le -\Delta_{\Psi_0}(u(t_-),u(t)), \mathcal{E}_t(u(t_+)) - \mathcal{E}_t(u(t)) \le -\Delta_{\Psi_0}(u(t),u(t_+)),\\ &\mathcal{E}_t(u(t_+)) - \mathcal{E}_t(u(t_-)) \le -\Delta_{\Psi_0}(u(t_-),u(t_+)), \end{aligned} \tag{$\mathrm{J}_{\mathrm{local}}$}$$

at each jump time $t \in \mathrm{J}_u$.

Proof. At every point $t \in (0,T)$ where $\mathrm{d}u'_{\mathrm{co}}(t)/\mathrm{d}\mu = 0$, the differential inclusion ($\mathrm{DN}_{0,\mathrm{BV}}$) reduces to the local stability condition ($\mathrm{S}_{\mathrm{loc}}$). In the general case, ($\mathrm{DN}_{0,\mathrm{BV}}$) follows by differentiation of ($\mathrm{E}'_{\mathrm{ineq}}$). Indeed, the latter procedure provides the following inequality between the distributional derivative $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t))$ of the map $t \mapsto \mathcal{E}_t(u(t))$ and the Ψ_0-total variation measure $\Psi_0(u'_{\mathrm{co}}) := \Psi_0\big(\mathrm{d}u'_{\mathrm{co}}/\mathrm{d}\mu\big)\mu$ for $\mu := u'_{\mathrm{Ca}} + \mathscr{L}^1$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) + \Psi_0(u'_{\mathrm{co}}) - \partial_t\mathcal{E}_t(u(t))\mathscr{L}^1 \le 0. \tag{2.30}$$

Applying the chain-rule formula for the composition of the C^1 functional $\mathcal{E}$ and the BV curve u (see [2] and [3], Thm. 3.96) and taking into account the fact that u'_{co} and u'_{J} are mutually singular, we obtain from (2.30) that

$$\left\langle -\mathrm{D}\mathcal{E}_t(u(t)), \frac{\mathrm{d}u'_{\mathrm{co}}}{\mathrm{d}\mu}\right\rangle \mu \ge \Psi_0(u'_{\mathrm{co}}) = \Psi_0\left(\frac{\mathrm{d}u'_{\mathrm{co}}}{\mathrm{d}\mu}\right)\mu. \tag{2.31}$$

Combining (2.31) with the local stability condition ($\mathrm{S}_{\mathrm{loc}}$), in view of the characterization (2.2) of $\partial\Psi_0$ and of (2.3) we finally conclude ($\mathrm{DN}_{0,\mathrm{BV}}$). Localizing ($\mathrm{E}'_{\mathrm{ineq}}$) around a jump point t we get the inequalities ($\mathrm{J}_{\mathrm{local}}$).

Conversely, let us suppose that a BV curve u satisfies ($\mathrm{DN}_{0,\mathrm{BV}}$) and ($\mathrm{J}_{\mathrm{local}}$). The local stability condition is an immediate consequence of ($\mathrm{DN}_{0,\mathrm{BV}}$), which yields $-\mathrm{D}\mathcal{E}_t(u(t)) \in K^*$ for $\mathscr{L}^1$-a.e. $t \in [0,T]$ and therefore, by continuity, at every point of $[0,T] \setminus \mathrm{J}_u$.

In order to get ($\mathrm{E}'_{\mathrm{ineq}}$), we again apply the chain rule for the composition $\mathcal{E}$ and u, obtaining

$$\mathcal{E}_{t_1}(u(t_1)) + \int_{t_0}^{t_1}\left\langle -\mathrm{D}\mathcal{E}_t(u(t)), \frac{\mathrm{d}u'_{\mathrm{co}}}{\mathrm{d}\mu}\right\rangle \mathrm{d}\mu(t) - \mathrm{Jmp}(\mathcal{E};[t_0,t_1]) = \mathcal{E}_{t_0}(u(t_0)) + \int_{t_0}^{t_1}\partial_t\mathcal{E}_t(u(t))\,\mathrm{d}t, \tag{2.32}$$

where

$$\begin{aligned} &\mathrm{Jmp}(\mathcal{E};[t_0,t_1]) = E_+(t_0) + E_-(t_1) + \sum_{t\in \mathrm{J}_u\cap(t_0,t_1)}\Big(E_-(t) + E_+(t)\Big)\\ &E_-(t) := \mathcal{E}_t(u(t)) - \mathcal{E}_t(u(t_-)), \quad \text{and } E_+(t) := \mathcal{E}_t(u(t_+)) - \mathcal{E}_t(u(t)). \end{aligned}$$

By ($\mathrm{DN}_{0,\mathrm{BV}}$) we have

$$\int_{t_0}^{t_1}\left\langle -\mathrm{D}\mathcal{E}_t(u(t)), \frac{\mathrm{d}u'_{\mathrm{co}}}{\mathrm{d}\mu}\right\rangle \mathrm{d}\mu(t) = \int_{t_0}^{t_1}\Psi_0\left(\frac{\mathrm{d}u'_{\mathrm{co}}}{\mathrm{d}\mu}(t)\right)\mathrm{d}\mu(t) = \int_{t_0}^{t_1}\mathrm{d}\Psi_0(u'_{\mathrm{co}}), \tag{2.33}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

whereas (J_{local}) yields for every $t \in J_u$

$$E_-(t) \le -\Delta_{\Psi_0}(u(t_-), u(t)), \quad E_+(t) \le -\Delta_{\Psi_0}(u(t), u(t_+)). \tag{2.34}$$

Hence, $-\mathrm{Jmp}(\mathcal{E}; [t_0, t_1]) \ge \mathrm{Jmp}_{\Psi_0}(u; [t_0, t_1])$ and therefore (E'_{ineq}) follows from (2.32). □

Unlike the case of energetic solutions (*cf.* Prop. 2.2), a precise description of the behavior of local solutions at jumps in missing here. In fact, the jump inequalities (J_{local}) are not sufficient to get an energy balance and do not completely capture the jump dynamics. As a consequence, the set of local solutions is much larger than one expects. In particular, all energetic solutions are local solutions, in addition to the fundamentally different solutions obtained as vanishing-viscosity limits, see the discussions in [42], Section 5.2, or [33], Section 2.6.

In order to get more precise insight into the jump properties and to understand the correct energy balance along them, we have to introduce a finer description of the dissipation. It is related to an extra contribution to the jump part of $\mathrm{Var}_{\Psi_0}(u; [\cdot, \cdot])$, which can be better described by using the *vanishing-viscosity contact potential* induced by the coupling Ψ, Ψ^*. We describe this notion in the next section.

3. Vanishing-viscosity contact potentials and Finsler dissipation distances

3.1. Heuristics for the concept of vanishing-viscosity contact potential

Suppose for the moment being that, in a given time interval $[r_0, r_1]$, the energy $\mathcal{E}_t(\cdot) = \mathcal{E}(\cdot)$ does not change w.r.t. time. If $\vartheta \in \mathrm{AC}([r_0, r_1]; X)$ is a solution of (DN_ε) connecting $u_0 = \vartheta(r_0)$ to $u_1 = \vartheta(r_1)$, then the energy release between the initial and the final state is, by the energy identity (2.27),

$$\mathcal{E}(u_0) - \mathcal{E}(u_1) = \int_{r_0}^{r_1} \Big(\Psi_\varepsilon(v) + \Psi_\varepsilon^*(w) \Big)\, \mathrm{d}t, \text{ with } v(t) = \dot\vartheta(t) \quad \text{and} \quad w(t) = -\mathrm{D}\mathcal{E}(\vartheta(t)) \quad \text{for a.a. } t \in (0, T). \tag{3.1}$$

If one looks for a lower bound of the right-hand side in the above energy identity, *which is independent of* $\varepsilon > 0$, it is natural to introduce the functional $\mathfrak{p} : X \times X^* \to [0, +\infty)$ defined by

$$\mathfrak{p}(v, w) := \inf_{\varepsilon > 0} \left(\Psi_\varepsilon(v) + \Psi_\varepsilon^*(w) \right) = \inf_{\varepsilon > 0} \left(\varepsilon^{-1} \Psi(\varepsilon v) + \varepsilon^{-1} \Psi^*(w) \right) \quad \text{for } v \in X,\ w \in X^*.$$

We obtain

$$\mathcal{E}(u_0) - \mathcal{E}(u_1) \ge \int_{r_0}^{r_1} \mathfrak{p}(v, w)\, \mathrm{d}t \quad \text{with} \quad v(t) = \dot\vartheta(t) \ \text{ and } \ w(t) = -\mathrm{D}\mathcal{E}(\vartheta(t)). \tag{3.2}$$

Since $\mathfrak{p}(\cdot, \cdot)$ is positively 1-homogeneous with respect to its first variable, the right-hand side expression in (3.2) is in fact independent of (monotone) time rescalings. Moreover, by construction the vanishing-viscosity contact potential $\mathfrak{p}(\cdot, \cdot)$ has the basic properties

$$\mathfrak{p}(v, w) \ge \langle w, v \rangle, \qquad \mathfrak{p}(v, w) \ge \Psi_0(v) \quad \text{for every } v \in X,\ w \in X^*. \tag{3.3}$$

Therefore, if $\tilde\vartheta \in \mathrm{AC}([r_0, r_1]; X)$ is another arbitrary curve connecting u_0 to u_1, the chain rule (2.26) for $\mathcal{E}$ yields

$$\mathcal{E}(u_0) - \mathcal{E}(u_1) = \int_{r_0}^{r_1} \langle \tilde w(t), \tilde v(t) \rangle\, \mathrm{d}t \overset{(3.3)}{\le} \int_{r_0}^{r_1} \mathfrak{p}(\tilde v(t), \tilde w(t))\, \mathrm{d}t, \tag{3.4}$$

where $\tilde v$ denotes the time derivative of $\tilde\vartheta$ and $\tilde w = -\mathrm{D}\mathcal{E}(\tilde\vartheta)$. It follows that, in a time regime in which the energy functional $\mathcal{E}$ does not change with respect to time, for every $\varepsilon > 0$ any viscous solution of (DN_ε) (and, therefore, any suitable limit of viscous solutions) should attain the minimum dissipation, measured in terms of the vanishing-viscosity contact potential $\mathfrak{p}$. Moreover, this dissipation always provides an upper bound for the energy release, reached exactly along viscous curves and their limits.

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Remark 3.1. In some of the cases discussed in Example 2.3, the vanishing-viscosity contact potential $\mathfrak{p}$ admits a more explicit representation.

(1) We first consider the Ψ_0-*viscosity* case (2.20), where $\Psi(v) := F(\Psi_0(v))$, $F : [0,+\infty) \to [0,+\infty)$ being a real convex superlinear function with $F(0) = 0, F'(0) = 1$. We use the 1-homogeneous support function Ψ_{0*} from (2.6) and obtain $\Psi^*(w) = F^*(\Psi_{0*}(w))$. This implies, for all $(v,w) \in X \times X^*$,

$$\mathfrak{p}(v,w) = \Psi_0(v)\,\max(1,\Psi_{0*}(w)) = \begin{cases} \Psi_0(v) & \text{if } w \in K^*, \\ \Psi_0(v)\,\Psi_{0*}(w) & \text{if } w \notin K^*. \end{cases} \tag{3.5}$$

(2) In the additive viscosity case (2.23) one has, for all $(v,w) \in X \times X^*$,

$$\mathfrak{p}(v,w) = \Psi_0(v) + \mathfrak{p}_V(v,w), \text{ where } \mathfrak{p}_V(v,w) = \inf_{\varepsilon>0}\left(\varepsilon^{-1}\Psi_V(\varepsilon v) + \varepsilon^{-1}\inf_{z\in K^*}\Psi_V^*(w-z)\right). \tag{3.6}$$

In particular, if $\Psi_V(v) = F_V(\|v\|)$ for some norm $\|\cdot\|$ of X and a convex and superlinear function $F_V : [0,+\infty) \to [0,+\infty)$ with $F_V(0) = F_V'(0) = 0$, we obtain, for all $(v,w) \in X \times X^*$,

$$\mathfrak{p}(v,w) = \Psi_0(v) + \mathfrak{p}_V(v,w) \quad \text{with} \quad \mathfrak{p}_V(v,w) = \|v\| \min_{z\in K^*}\|w-z\|_*. \tag{3.7}$$

Notice that in (3.5) and (3.7) the form of the vanishing-viscosity contact potential $\mathfrak{p}$ *does not depend* on the choice of F and F_V, respectively, but only on the chosen viscosity norm.

By the 1-homogeneity of $\mathfrak{p}(\cdot,w)$ and these variational properties, it is then natural to introduce the following Finsler dissipation distance, which replaces the previous rate-independent distance $(v_0,v_1) \mapsto \Delta_{\Psi_0}(v_0,v_1) = \Psi_0(v_1-v_0)$.

Definition 3.2 (Finsler dissipation distance)**.** For a fixed $t \in [0,T]$, the (possibly asymmetric) Finsler distance induced by $\mathfrak{p}$ and (the differential of) $\mathcal{E}$ at the time t is for every $u_0,u_1 \in X$ given by

$$\Delta_{\mathfrak{p},\mathcal{E}}(t;u_0,u_1) := \inf\left\{\int_{r_0}^{r_1}\mathfrak{p}(\dot\vartheta(r),-\mathrm{D}\mathcal{E}_t(\vartheta(r)))\,\mathrm{d}r : \vartheta \in \mathrm{AC}([r_0,r_1];X),\ \vartheta(r_0)=u_0,\ \vartheta(r_1)=u_1\right\}. \tag{3.8}$$

The Finsler distance $\Delta_{\mathfrak{p},\mathcal{E}}(t;\cdot,\cdot)$ is obviously related to an infinitesimal Finsler metric $\mathcal{F}_t : X \times X \to \mathbb{R}$, with $\mathcal{F}_t(u,\dot u) = \mathfrak{p}(\dot u, -\mathrm{D}\mathcal{E}_t(u))$.

Remark 3.3. Since $\mathfrak{p}(v,w) \geq \Psi_0(v)$ by (3.3), a simple time rescaling argument shows that the infimum in (3.8) is always attained by a *Lipschitz curve* $\vartheta \in \mathrm{Lip}([r_0,r_1];X)$ with constant $\mathfrak{p}$-speed, in particular such that $\mathfrak{p}(\dot\vartheta(r),-\mathrm{D}\mathcal{E}_t(\vartheta(r))) \equiv 1$ for a.a. $r \in (r_0,r_1)$.

By the heuristic discussion developed throughout (3.1)–(3.4), the distance $\Delta_{\mathfrak{p},\mathcal{E}}$ is the natural candidate for improving the potential Ψ_0, and the related distance $\Delta_{\Psi_0}(v_0,v_1) = \Psi_0(v_1-v_0)$, in the jump contributions (2.13) and in the jump conditions ($\mathrm{J}_{\mathrm{ener}}$). The second relation of (3.3) implies

$$\Delta_{\mathfrak{p},\mathcal{E}}(t;u_0,u_1) \geq \Delta_{\Psi_0}(u_0,u_1) \quad \text{for every } u_0,u_1 \in X. \tag{3.9}$$

The notion of jump variation arising from such replacements is precisely stated as follows.

Definition 3.4 (the total variation induced by $\Delta_{\mathfrak{p},\mathcal{E}}$)**.** Let $u \in \mathrm{BV}([0,T];X)$ a given curve, let u'_{co} be the diffuse part of its distributional derivative u', and let J_u be its pointwise jump set (2.9). For every subinterval $[a,b] \subset [0,T]$ the *jump variation* of u induced by $(\mathfrak{p},\mathcal{E})$ on $[a,b]$ is

$$\begin{aligned}\mathrm{Jmp}_{\mathfrak{p},\mathcal{E}}(u;[a,b]) := {}& \Delta_{\mathfrak{p},\mathcal{E}}(a;u(a),u(a_+)) + \Delta_{\mathfrak{p},\mathcal{E}}(b;u(b_-),u(b)) \\ & + \sum_{t\in \mathrm{J}_u\cap(a,b)}\big(\Delta_{\mathfrak{p},\mathcal{E}}(t;u(t_-),u(t)) + \Delta_{\mathfrak{p},\mathcal{E}}(t;u(t),u(t_+))\big),\end{aligned} \tag{3.10}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

and the (pseudo-)total variation induced by $(\mathfrak{p}, \mathcal{E})$ is

$$\mathrm{Var}_{\mathfrak{p},\mathcal{E}}(u;[a,b]) := \int_a^b \mathrm{d}\Psi_0(u'_{\mathrm{co}}) + \mathrm{Jmp}_{\mathfrak{p},\mathcal{E}}(u;[a,b]). \tag{3.11}$$

Remark 3.5 (the (pseudo-)total variation $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$)**.** Let us mention that $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$ enjoys some of the properties of the usual total variation functionals, but it is not lower semicontinuous w.r.t. pointwise convergence. In fact, it is not difficult to see that its lower semicontinuous envelope is simply Var_{Ψ_0}. Furthermore, $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$ is not induced by any distance on X. Indeed, we have used a *slanted* font in the notation Var to stress this fact. In order to recover a more standard total variation in a metric setting, one has to work in the extended space $\mathscr{X} := [0,T] \times X$ and add the local stability constraint $-\mathrm{D}\mathcal{E}_t(u(t)) \in K^*$ on the "continuous" part of the trajectories. We shall discuss this point of view in Section 6.

In view of inequality (3.9) between the Finsler dissipation $\Delta_{\mathfrak{p},\mathcal{E}}$ and Δ_{Ψ_0}, the notion of total variation associated with $\Delta_{\mathfrak{p},\mathcal{E}}$ provides an upper bound for Var_{Ψ_0}, namely

$$\forall\, u \in \mathrm{BV}([0,T];X),\ [a,b] \subset [0,T]: \qquad \mathrm{Var}_{\mathfrak{p},\mathcal{E}}(u;[a,b]) \geq \mathrm{Var}_{\Psi_0}(u;[a,b]). \tag{3.12}$$

3.2. **Vanishing-viscosity contact potentials**

While postponing the definition of BV solutions related to $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$ to the next section, we prepare the main properties of the vanishing-viscosity contact potential $\mathfrak{p}$

$$\mathfrak{p}(v,w) := \inf_{\varepsilon>0} \left(\Psi_\varepsilon(v) + \Psi_\varepsilon^*(w)\right) = \inf_{\varepsilon>0} \left(\varepsilon^{-1}\Psi(\varepsilon v) + \varepsilon^{-1}\Psi^*(w)\right) \quad \text{for } v \in X,\ w \in X^*, \tag{3.13}$$

which partly matches the notion of bipotential introduced by [7], see Remark 3.8. In this subsection we provide only those results needed in the subsequent analysis, while in Appendix A we will give a more detailed analysis of the properties of $\mathfrak{p}$.

We will also need the so-called contact set associated with $\mathfrak{p}$.

Definition 3.6 (contact set $\Sigma_{\mathfrak{p}}$ of $\mathfrak{p}$)**.** The *contact set* $\Sigma_{\mathfrak{p}} \subset X \times X^*$ of the vanishing-viscosity contact potential $\mathfrak{p}$ is defined as

$$\Sigma_{\mathfrak{p}} := \Big\{(v,w) \in X \times X^* : \mathfrak{p}(v,w) = \langle w, v\rangle\Big\}. \tag{3.14}$$

To prove the main properties of $\mathfrak{p}$ we use the following simple representation:

$$\mathfrak{p}(v,w) = \mathfrak{q}(v,\Psi^*(w)) \quad \text{with} \quad \mathfrak{q}(v,\rho) = \inf_{\varepsilon>0} \frac{1}{\varepsilon}\Big(\Psi(\varepsilon v) + \rho\Big). \tag{3.15}$$

It is easy to see that $\mathfrak{q}(\cdot,\rho) : X \to \mathbb{R}$ is convex and 1-homogeneous. Moreover, for every $v \in X$ the mapping $\mathfrak{q}(v,\cdot) : [0,+\infty) \to [0,+\infty)$ is nondecreasing.

Theorem 3.7 (intrinsic properties of $\mathfrak{p}$)**.** *The functional $\mathfrak{p} : X \times X^* \to [0,+\infty)$ defined by (3.13) is continuous and satisfies the following properties:*

($\mathfrak{p}$1) *For every $v \in X$ and $w \in X^*$ the maps $\mathfrak{p}(v,\cdot)$ and $\mathfrak{p}(\cdot,w)$ have convex sublevels.*
($\mathfrak{p}$2) *$\mathfrak{p}(v,w) \geq \langle w, v\rangle$ for every $(v,w) \in X \times X^*$, and $\mathfrak{p}(v,w) > 0$ for $v \neq 0$.*
($\mathfrak{p}$3) *For every $w \in X^*$ the map $v \mapsto \mathfrak{p}(v,w)$ is 1-homogeneous and convex on X.*
($\mathfrak{p}$4) *The contact set can be characterized as follows:*

$$(v,w) \in \Sigma_{\mathfrak{p}} \quad \Leftrightarrow \quad w \in \partial\mathfrak{p}(\cdot,w)(v). \tag{3.16}$$

Proof. We use the representation (3.15) of $\mathfrak{p}$ *via* $\mathfrak{q}$, then the continuity of $\mathfrak{p}$ and properties ($\mathfrak{p}$2) and ($\mathfrak{p}$3) follow immediately. Moreover, ($\mathfrak{p}$1) is obtained since Ψ^* is convex.

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

It remains to establish the characterization (𝔭4): because of (𝔭3), the subdifferential of $\mathfrak{p}(\cdot,w)$ is characterized *via*

$$\xi \in \partial\mathfrak{p}(\cdot,w)(v) \quad \Longleftrightarrow \quad \big(\langle \xi, v\rangle = \mathfrak{p}(v,w) \text{ and } \langle \xi, \widetilde{v}\rangle \le \mathfrak{p}(\widetilde{v},w) \text{ for all } \widetilde{v} \in X\big),$$

cf. (2.2) and (2.3). Letting $\xi = w$ and using (𝔭2), the equivalence in (3.16) follows. □

Remark 3.8. Properties (𝔭1)–(𝔭4) suggest a strong analogy between $\mathfrak{p}$ and the notion of bipotential introduced by [7]. According to [7], a bipotential is a functional $\mathfrak{b} : X \times X^* \to (-\infty, +\infty]$ that is *convex* and lower semicontinuous in both arguments separately, satisfies (𝔭2), and whose contact set fulfills a condition similar to (3.16), namely

$$(v,w) \in \Sigma_{\mathfrak{b}} \quad \Leftrightarrow \quad w \in \partial\mathfrak{b}(\cdot,w)(v) \quad \Leftrightarrow \quad v \in \partial\mathfrak{b}(v,\cdot)(w).$$

In our situation, (3.16) is a direct consequence of the homogeneity of $\mathfrak{p}$, but the convexity condition with respect to w looks too restrictive, as shown by the following simple example. Consider the case $X = X^* = \mathbb{R}^2$, with $\Psi(v) := \|v\|_1 + \Psi_V(v)$, $\|v\|_1 := |v_1| + |v_2|$, and

$$\Psi_V(v) := \frac{1}{2}v_1^2 + \frac{1}{4}v_2^4, \quad v = (v_1, v_2) \in \mathbb{R}^2, \quad \Psi_V^*(w) = \frac{1}{2}w_1^2 + \frac{3}{4}w_2^{4/3}, \quad w = (w_1, w_2) \in \mathbb{R}^2.$$

By (3.6) we have $\mathfrak{p}(v,w) = \|v\|_1 + \mathfrak{p}_V(v,w)$, with $\mathfrak{p}_V(v,w) = \inf_{\varepsilon>0} \frac{1}{\varepsilon}\Big(\Psi_V(\varepsilon v) + \Psi^*(w)\Big)$, and find

$$\Psi^*(w) = \frac{1}{2}(|w_1| - 1)_+^2 + \frac{3}{4}(|w_2| - 1)_+^{4/3}.$$

For $v = (v_1, 0)$, $w = (0, w_2)$ we obtain $\mathfrak{p}_V((v_1,0),w) = |v_1|\sqrt{2\Psi^*(w)}$, giving the nonconvex mapping

$$w_2 \mapsto \mathfrak{p}((v_1,0),(0,w_2)) = |v_1| + \sqrt{3/2}\,|v_1|\big((|w_2| - 1)_+\big)^{2/3}.$$

Let us now consider some properties of $\mathfrak{p}$ and its contact set $\Sigma_{\mathfrak{p}}$ involving explicitly the functional Ψ. Since the vanishing-viscosity contact potential $\mathfrak{p}$ is defined through the minimum procedure (3.13), the contact set is strictly related to the set of optimal $\varepsilon > 0$ attaining the minimum in (3.13).

Definition 3.9 (contact viscosities)**.** For every $(v,w) \in X \times X^*$ we introduce the set-valued function Λ (with possibly empty values)

$$\Lambda(v,w) := \Big\{\varepsilon \ge 0 : \mathfrak{p}(v,w) = \Psi_\varepsilon(v) + \Psi_\varepsilon^*(w)\Big\} \subset [0, +\infty). \tag{3.17}$$

Notice that for every $(v,w) \in X \times X^*$ the function $\varepsilon \mapsto \varepsilon^{-1}\Psi(\varepsilon v) + \varepsilon^{-1}\Psi^*(w)$ is convex on $(0, +\infty)$. Since Ψ has superlinear growth at infinity, it goes to $+\infty$ as $\varepsilon \uparrow +\infty$ if $v \neq 0$, so that

$$\text{the set } \Lambda(v,w) \text{ is always a bounded closed interval if } v \neq 0 \tag{3.18a}$$

$$\text{and does not contain 0 if } v \neq 0 \text{ and } w \notin K^*. \tag{3.18b}$$

The last property follows from $\Psi^*(w) > 0$ if $w \notin K^*$, which holds due to (2.17). For $w \in K^*$ we have $\Lambda(v,w) = \{0\}$ for every $v \in X$, while $w \notin K^*$ implies $\Lambda(0,w) = \emptyset$.

The role of $\Lambda(\cdot,\cdot)$ is described by the next lemma.

Lemma 3.10. *For every $v \neq 0$ the contact set (3.14) can be characterized by*

$$(v,w) \in \Sigma_{\mathfrak{p}} \quad \Leftrightarrow \quad \begin{cases} \textit{either} \ \ w \in \partial\Psi_0(v) \subset K^* \\ \textit{or} \ \ w \notin K^* \ \textit{and for every } \varepsilon \in \Lambda(v,w) : w \in \partial\Psi(\varepsilon v). \end{cases} \tag{3.19}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Proof. Ad "⇒". If $w \in K^*$, then $\Psi^*(w) = 0$. From $(v,w) \in \Sigma_{\mathfrak{p}}$ we deduce $\mathfrak{p}(v,w) = \Psi_0(v)$ and obtain $w \in \partial\Psi_0(v)$ as desired. If $w \notin K^*$, then $\Psi^*(w) > 0$. Using $v \neq 0$ we have $(0,+\infty) \supset \Lambda(v,w) \neq \emptyset$. Thus, we find $\varepsilon \in \Lambda(v,w)$ with $\mathfrak{p}(v,w) = \frac{1}{\varepsilon}\big(\Psi(\varepsilon v) + \Psi^*(w)\big)$. Moreover, for all $\widetilde{v} \in X$ we have

$$\frac{1}{\varepsilon}\big(\Psi(\varepsilon\widetilde{v}) + \Psi^*(w)\big) \overset{\text{(i)}}{\geq} \mathfrak{p}(\widetilde{v},w) \overset{\text{(ii)}}{\geq} \mathfrak{p}(v,w) + \langle w, \widetilde{v}-v\rangle = \frac{1}{\varepsilon}\big(\Psi(\varepsilon v) + \Psi^*(w)\big) + \langle w, \widetilde{v}-v\rangle.$$

We used that $\mathfrak{p}$ is an infimum for (i) and (3.16) for (ii) in the second estimate. Multiplying by $\varepsilon > 0$, canceling $\Psi^*(w)$, and replacing $\varepsilon\widetilde{v}$ by $\widehat{v}$ gives $\Psi(\widehat{v}) \geq \Psi(\varepsilon v) + \langle w, \widehat{v}-v\rangle$, which implies $w \in \partial\Psi(\varepsilon v)$ as desired.

Ad "⇐". If $w \in \partial\Psi_0(v) \subset K^*$, then $\mathfrak{p}(v,w) = \Psi_0(v) = \langle w, v\rangle$ by the characterization of $\partial\Psi_0$. If $w \notin K^*$, then $w \in \partial\Psi(\varepsilon v)$ gives $\Psi(\varepsilon v) + \Psi^*(w) = \langle \varepsilon v, w\rangle$. Dividing by $\varepsilon > 0$ (*cf.* (3.18b)) and using the definition of $\mathfrak{p}$ yields $\mathfrak{p}(v,w) \leq \langle w, v\rangle$. Together with ($\mathfrak{p}2$) we obtain the desired equality. □

4. BV solutions and energy-driven dissipation

4.1. BV solutions

We can now give our precise definition of BV solution of the rate-independent system $(X, \mathcal{E}, \mathfrak{p})$, driven by the vanishing-viscosity contact potential $\mathfrak{p}$ defined in (3.13) and the energy $\mathcal{E}$. Below we will show that these solutions arise as limits of solutions u_ε of the viscous system $(X, \mathcal{E}, \Psi_\varepsilon)$.

From a formal point of view, the definition simply replaces the global stability condition (S) by the local one ($\mathrm{S}_{\mathrm{loc}}$), and the Ψ_0-total variation in the energy balance (E) by the Finsler total variation $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$, see (3.11).

Definition 4.1 (BV solutions, variational characterization)**.** A curve $u \in \mathrm{BV}([0,T];X)$ is a BV solution of the rate-independent system $(X, \mathcal{E}, \mathfrak{p})$ if the *local stability* ($\mathrm{S}_{\mathrm{loc}}$) and the $(\mathfrak{p}, \mathcal{E})$-*energy balance* hold:

$$-\mathrm{D}\mathcal{E}_t(u(t)) \in K^* \quad \text{for a.a. } t \in [0,T] \setminus J_u \tag{$\mathrm{S}_{\mathrm{loc}}$}$$

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

We shall see in the next Section 4.3 that any pointwise limit, as $\varepsilon \downarrow 0$, of the solutions (u_ε) of the viscous equation (DN_ε) or, as $\tau, \varepsilon \downarrow 0$, of the discrete solutions $(\overline{\mathrm{U}}_{\tau,\varepsilon})$ of the viscous incremental problems ($\mathrm{IP}_{\tau,\varepsilon}$), is a BV solution for $(X, \mathcal{E}, \mathfrak{p})$. Let us first get more insight into Definition 4.1.

Properties of BV solutions. As in the case of energetic solutions, it is not difficult to see that the energy balance ($\mathrm{E}_{\mathfrak{p},\mathcal{E}}$) holds on any subinterval $[t_0,t_1] \subset [0,T]$. Moreover, if the local stability condition ($\mathrm{S}_{\mathrm{loc}}$) holds, for checking ($\mathrm{E}_{\mathfrak{p},\mathcal{E}}$) it is sufficient to prove the corresponding inequality.

Proposition 4.2. *If $u \in \mathrm{BV}([0,T];X)$ satisfies ($\mathrm{E}_{\mathfrak{p},\mathcal{E}}$), then for every subinterval $[t_0,t_1]$*

$$\mathrm{Var}_{\mathfrak{p},\mathcal{E}}(u;[t_0,t_1]) + \mathcal{E}_{t_1}(u(t_1)) = \mathcal{E}_{t_0}(u(t_0)) + \int_{t_0}^{t_1} \partial_t\mathcal{E}_s(u(s))\,\mathrm{d}s. \tag{$\mathrm{E}'_{\mathfrak{p},\mathcal{E}}$}$$

Moreover, if u satisfies ($\mathrm{S}_{\mathrm{loc}}$), then ($\mathrm{E}_{\mathfrak{p},\mathcal{E}}$) is equivalent to the energy inequality

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

Proof. ($\mathrm{E}'_{\mathfrak{p},\mathcal{E}}$) easily follows from the additivity property

$$\forall\, 0 \leq t_0 < t_1 < t_2 \leq T: \quad \mathrm{Var}_{\mathfrak{p},\mathcal{E}}(u;[t_0,t_1]) + \mathrm{Var}_{\mathfrak{p},\mathcal{E}}(u;[t_1,t_2]) = \mathrm{Var}_{\mathfrak{p},\mathcal{E}}(u;[t_0,t_2]). \tag{4.1}$$

In order to prove that ($\mathrm{E}_{\mathfrak{p},\mathcal{E};\mathrm{ineq}}$) implies ($\mathrm{E}_{\mathfrak{p},\mathcal{E}}$), we argue as in [42], Proposition 4, taking ($\mathrm{S}_{\mathrm{loc}}$) into account. □

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

By (3.12), every BV solution is also a local solution according to Definition 2.5, *i.e.* it satisfies the local stability condition and energy inequality ($\mathrm{E}'_{\mathrm{ineq}}$). In fact, one has a more accurate description of the jump conditions, as the following theorem shows (*cf.* Props. 2.2 and 2.7).

Theorem 4.3 (differential characterization of BV solutions)**.** *A curve $u \in \mathrm{BV}([0,T];X)$ is a BV solution of the rate-independent system $(X,\mathcal{E},\mathfrak{p})$ if and only if it satisfies the doubly nonlinear differential inclusion in the BV sense*

$$\partial\Psi_0\left(\frac{\mathrm{d}u'_{\mathrm{co}}}{d\mu}(t)\right) + \mathrm{D}\mathcal{E}_t(u(t)) \ni 0 \quad \text{for } \mu\text{-a.e. } t\in[0,T], \quad \mu := \mathscr{L}^1 + |u'_{\mathrm{Ca}}|, \tag{$\mathrm{DN}_{0,\mathrm{BV}}$}$$

and the following jump conditions at each point $t \in \mathrm{J}_u$ of the jump set (2.9)

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

Proof. We start from a BV solution, which hence is a local solution. Using Proposition 2.7 it satisfies ($\mathrm{DN}_{0,\mathrm{BV}}$). The jump conditions ($\mathrm{J}_{\mathrm{BV}}$) can be obtained by localizing ($\mathrm{E}'_{\mathfrak{p},\mathcal{E}}$) around any jump time $t \in \mathrm{J}_u$.

For the converse implication first note that ($\mathrm{S}_{\mathrm{loc}}$) ensues from ($\mathrm{DN}_{0,\mathrm{BV}}$) (see the proof of Prop. 2.7). For establishing ($\mathrm{E}_{\mathfrak{p},\mathcal{E}}$) we argue as in the second part of the proof of Proposition 2.7, still applying (2.32) and (2.33), but replacing inequalities (2.34) with the identities

$$E_-(t) = -\Delta_{\mathfrak{p},\mathcal{E}}(t;u(t_-),u(t)), \quad E_+(t) = -\Delta_{\mathfrak{p},\mathcal{E}}(t;u(t),u(t_+)) \quad \text{for all } t\in \mathrm{J}_u,$$

which are due to (J_{BV}). Hence, $-\mathrm{Jmp}(\mathcal{E};[0,T]) = \mathrm{Jmp}_{\mathfrak{p},\mathcal{E}}(u;[0,T])$, and ($\mathrm{E}_{\mathfrak{p},\mathcal{E}}$) follows. □

4.2. **Jumps and optimal transitions**

This section is devoted to a refined description of the behavior of BV solutions along their jumps. Let us first introduce the notion of *optimal transition*.

Definition 4.4. Consider $t\in[0,T]$, $u_-,u_+\in X$ with $-\mathrm{D}\mathcal{E}_t(u_-)$, $-\mathrm{D}\mathcal{E}_t(u_+)\in K^*$, and $-\infty \le r_0 < r_1 \le +\infty$. An absolutely continuous curve $\vartheta : [r_0,r_1]\to X$ connecting $u_- = \vartheta(r_0)$ and $u_+ = \vartheta(r_1)$ is an $(\mathfrak{p},\mathcal{E}_t)$-*optimal transition* between u_- and u_+, if

$$\dot\vartheta(r) \neq 0 \quad \text{for a.a. } r\in(r_0,r_1); \quad \Psi_{0*}(-\mathrm{D}\mathcal{E}_t(\vartheta(r))) \ge 1 \quad \forall\, r\in[r_0,r_1], \tag{O.1}$$

$$\mathcal{E}_t(u_-) - \mathcal{E}_t(u_+) = \Delta_{\mathfrak{p},\mathcal{E}}(t;u_-,u_+) = \int_{r_0}^{r_1} \mathfrak{p}(\dot\vartheta(r), -\mathrm{D}\mathcal{E}_t(\vartheta(r)))\,\mathrm{d}r. \tag{O.2}$$

We also say that an optimal transition ϑ is of

sliding type, if	$-\mathrm{D}\mathcal{E}_t(\vartheta(r)) \in K^*$ for every $r\in[r_0,r_1]$,	($\mathrm{O}_{\mathrm{sliding}}$)
viscous type, if	$-\mathrm{D}\mathcal{E}_t(\vartheta(r)) \notin K^*$ for every $r\in(r_0,r_1)$,	($\mathrm{O}_{\mathrm{viscous}}$)
energetic type, if	$\mathcal{E}_t(u_-) - \mathcal{E}_t(u_+) = \Psi_0(u_+ - u_-)$.	($\mathrm{O}_{\mathrm{ener}}$)

We denote by $\Theta(t,u_-,u_+)$ the (possibly empty) collection of such optimal transitions, with normalized domain $[0,1]$ and *constant Finsler velocity*

$$\mathfrak{p}(\dot\vartheta(r), -\mathrm{D}\mathcal{E}_t(\vartheta(r))) \equiv \mathcal{E}_t(u_-) - \mathcal{E}_t(u_+) \qquad \text{for a.a. } r\in(0,1). \tag{4.2}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Remark 4.5. The notion of optimal transition is invariant by absolutely continuous (monotone) time rescalings with absolutely continuous inverse. Moreover, any optimal transition ϑ has finite length, it admits a reparameterization with constant Finsler velocity $\mathfrak{p}(\dot\vartheta(\cdot), -\mathrm{D}\mathcal{E}_t(\vartheta(\cdot)))$, and is a minimizer of (3.8), so that it is not restrictive to assume $\vartheta \in \Theta(t, u_-, u_+)$.

The first result on optimal transitions shows that they always lie in the contact set $\Sigma_\mathfrak{p}$.

Proposition 4.6. *Any optimal transition ϑ satisfies the* contact condition

$$(\dot\vartheta(r), -\mathrm{D}\mathcal{E}_t(\vartheta(r))) \in \Sigma_\mathfrak{p} \quad \textit{for a.a. } r \in (r_0, r_1). \tag{4.3}$$

Proof. The chain rule (CR) and (O.2) yield

$$\int_{r_0}^{r_1} \langle -\mathrm{D}\mathcal{E}_t(\vartheta(r)), \dot\vartheta(r)\rangle \,\mathrm{d}r \overset{\text{(CR)}}{=} \mathcal{E}_t(u(t_-)) - \mathcal{E}_t(u(t_+)) \overset{\text{(O.2)}}{=} \int_{r_0}^{r_1} \mathfrak{p}(\dot\vartheta(r), -\mathrm{D}\mathcal{E}_t(\vartheta(r)))\,\mathrm{d}r,$$

so that $\int_{r_0}^{r_1} \big(\mathfrak{p}(\dot\vartheta, -\mathrm{D}\mathcal{E}_t(\vartheta(r))) - \langle -\mathrm{D}\mathcal{E}_t(\vartheta(r)), \dot\vartheta(r)\rangle\big)\,\mathrm{d}r = 0$. Since the integrand is always nonnegative, it follows that (4.3) holds. □

The following characterization of BV solutions is an immediate consequence of Theorem 4.3, Proposition 2.7, and Remark 3.3.

Theorem 4.7. *A local solution $u \in \mathrm{BV}([0,T];X)$ is a BV solution according to Definition 4.1 if and only if at every jump time $t \in \mathrm{J}_u$ the initial and final values $u(t_-)$ and $u(t_+)$ can be connected by an optimal transition $\vartheta^t \in \Theta(t; u(t_-), u(t_+))$, and there exists $r \in [0,1]$ such that $u(t) = \vartheta^t(r)$.*

For the following, a detailed analysis of the properties of optimal transitions is not essential. Hence we have deferred it to Appendix B. Here we just give some helpful intuition. First, we note that all optimal transitions of energetic type are sliding and are additionally geodesics with respect to the pseudo-norm Ψ_0. An optimal transition is of sliding type if and only if

$$\partial\Psi_0(\dot\vartheta(r)) + \mathrm{D}\mathcal{E}_t(\vartheta(r)) \ni 0 \quad \text{and} \quad \Psi_{0*}(-\mathrm{D}\mathcal{E}_t(\vartheta(r))) = 1 \quad \text{for a.a. } r \in (0,1),$$

which implies that sliding optimal transitions depend only on Ψ_0, and are otherwise independent of the form of the vanishing-viscosity contact potential $\mathfrak{p}$, and thus of the particular viscosity potential Ψ.

Second, we note that optimal transitions of viscous type are characterized by

$$\partial\Psi(\varepsilon(r)\dot\vartheta(r)) + \mathrm{D}\mathcal{E}_t(\vartheta(r)) \ni 0 \quad \text{and} \quad \varepsilon(r) \in \Lambda(\dot\vartheta(r), -\mathrm{D}\mathcal{E}_t(\vartheta(r))) \quad \text{for a.a. } r \in (0,1).$$

Hence, different choices of the viscous dissipation Ψ shall give raise to a different behavior in the viscous jumping regime, see also the example in [49], Section 2.2. The latter paper sets forth a different characterization of rate-independent evolution, still oriented towards local stability, but derived from a global-in-time variational principle and not a vanishing-viscosity approach.

Finally, we note that general optimal transitions may be a concatenation of optimal transitions of different types. Since viscous types occur on open intervals, there can be at most a countable number of viscous parts, while sliding might occur on complicated closed subsets.

The next result clarifies the relationships between energetic and BV solutions.

Corollary 4.8 (energy balance and comparison with energetic solutions)**.**

(1) *A BV solution u of the rate-independent system $(X, \mathcal{E}, \mathfrak{p})$ satisfies the energy balance (E) if and only if every optimal transition associated with its jump set is of energetic type ($\mathrm{O}_{\mathrm{ener}}$).*
(2) *A BV solution u is an energetic solution if and only if it satisfies the global stability condition (S). In that case, all of its optimal transitions are of energetic type.*

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

(3) *Conversely, an energetic solution u is a BV solution if and only if, for every $t \in \mathrm{J}_u$, any jump couple $(u(t_-), u(t_+))$ can be connected by a sliding optimal transition.*

Proof. **Ad (1).** Let u be a BV solution such that every optimal transition is of energetic type ($\mathrm{O}_{\mathrm{ener}}$). Now, taking into account (J_{BV}), one sees that ($\mathrm{O}_{\mathrm{ener}}$) is equivalent to the jump conditions ($\mathrm{J}_{\mathrm{ener}}$). Then, equation ($\mathrm{DN}_{0,\mathrm{BV}}$) (which holds by Thm. 4.3) and ($\mathrm{J}_{\mathrm{ener}}$) yield the energy balance (E) (*cf.* the proofs of Props. 2.2 and 2.7). The converse implication ensues by analogous arguments.

Ad (2). The necessity is obvious, and for the sufficiency we observe that, for every jump point $t \in \mathrm{J}_u$, the global stability condition (S) (written first for $u(t_-)$ with test functions $v = u(t)$ and $v = u(t_+)$, and then for $u(t)$ with $v = u(t_+)$) yields

$$\begin{array}{ccccccc}
\Psi_0(u(t) - u(t_-)) & \geq & \mathcal{E}_t(u(t_-)) - \mathcal{E}_t(u(t)) & = & \Delta_{\mathfrak{p},\mathcal{E}}(t; u(t_-), u(t)) & \geq & \Psi_0(u(t) - u(t_-)), \\
\Psi_0(u(t_+) - u(t_-)) & \geq & \mathcal{E}_t(u(t_-)) - \mathcal{E}_t(u(t_+)) & = & \Delta_{\mathfrak{p},\mathcal{E}}(t; u(t_-), u(t_+)) & \geq & \Psi_0(u(t_+) - u(t_-)), \\
\Psi_0(u(t_+) - u(t)) & \geq & \mathcal{E}_t(u(t)) - \mathcal{E}_t(u(t_+)) & = & \Delta_{\mathfrak{p},\mathcal{E}}(t; u(t), u(t_+)) & \geq & \Psi_0(u(t_+) - u(t)),
\end{array}$$

where the intermediate equalities are due to (O.2) and the subsequent inequalities to (3.9). The resulting identities ultimately show that the transition is energetic, by the very definition ($\mathrm{O}_{\mathrm{ener}}$).

Ad (3). The condition is clearly sufficient. It is also necessary by the previous point, since energetic transitions are in particular of sliding type. □

4.3. **Vanishing-viscosity limit**

We now conclude this section by stating our main convergence results. First we show that (subsequences of) the viscous solutions of $(X, \mathcal{E}, \Psi_\varepsilon)$ converge to BV solutions. The proof will follow from the theory developed in the subsequent Sections 5 and 6. Second, we provide a convergence result for the viscous incremental problem ($\mathrm{IP}_{\tau,\varepsilon}$) for $(\tau, \varepsilon) \to 0$ under the additional assumption $\varepsilon/\tau \to +\infty$. Both results rely on the fundamental lower semicontinuity result in Proposition 6.14, which in particular shows that if $u_\varepsilon \in \mathrm{AC}([0,T]; X)$ pointwise converge to u as $\varepsilon \downarrow 0$, then

$$\liminf_{\varepsilon \downarrow 0} \int_0^T \Big(\Psi_\varepsilon(\dot{u}_\varepsilon(t)) + \Psi_\varepsilon^*(-\mathrm{D}\mathcal{E}_t(u_\varepsilon(t))) \Big) \, \mathrm{d}t \geq \mathrm{Var}_{\mathfrak{p},\mathcal{E}}(u; [0,T]). \tag{4.4}$$

The latter result is derived *via* the reparameterization technique developed in Section 5. However, we emphasize that the statements of our convergence results are totally independent of these techniques.

Theorem 4.9 (convergence of viscous approximations to BV solutions)**.** *Consider a sequence $(u_\varepsilon) \subset \mathrm{AC}([0,T]; X)$ of solutions of the viscous equation (DN_ε), with $u_\varepsilon(0) \to u_0$ as $\varepsilon \downarrow 0$.*

Then, every vanishing sequence $\varepsilon_k \downarrow 0$ admits a further subsequence (still denoted by (ε_k)), and a limit function $u \in \mathrm{BV}([0,T]; X)$ such that

$$u_{\varepsilon_k}(t) \to u(t) \quad \text{for every } t \in [0,T] \text{ as } k \uparrow +\infty, \tag{4.5}$$

and u is a BV solution for $(X, \mathcal{E}, \mathfrak{p})$ in the sense of Definition 4.1.

Proof. It follows from the discussion developed in Section 2.5 that for every sequence $\varepsilon_k \downarrow 0$ there exists a not relabeled subsequence (u_{ε_k}) such that (4.5) holds, and u complies with the local stability condition ($\mathrm{S}_{\mathrm{loc}}$). In view of Proposition 4.2, it is then sufficient to check that ($\mathrm{E}_{\mathfrak{p},\mathcal{E};\mathrm{ineq}}$) holds. The latter energy inequality is a direct consequence of the ε-energy identity (2.27) and the lower semicontinuity property (4.4) proved in Proposition 6.14 below. □

Our next result concerns the convergence of the discrete solutions to the *viscous time-incremental* problem ($\mathrm{IP}_{\tau,\varepsilon}$), as both the viscosity parameter ε and the time-step τ tend to 0. For the statement we recall the definition of left-continuous interpolants $\overline{\mathrm{U}}_{\tau,\varepsilon} : [0,T] \to X$ from (2.14).

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Theorem 4.10 (convergence of discrete solutions of the viscous incremental problems)**.** *Let* $\overline{\mathrm{U}}_{\tau,\varepsilon} : [0,T] \to X$ *be the left-continuous piecewise constant interpolants of the discrete solutions of the viscous incremental problem* ($\mathrm{IP}_{\tau,\varepsilon}$)*, with* $\mathrm{U}^0_{\tau,\varepsilon} \to u_0$ *as* $(\tau,\varepsilon) \to 0$.

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

$$(\tau_k,\varepsilon_k) \to 0 \quad \textit{and} \quad \lim_{k\downarrow 0} \frac{\varepsilon_k}{\tau_k} = +\infty \tag{4.6}$$

admit a subsequence (*still denoted by* $(\tau_k,\varepsilon_k)_k$) *and a limit function* $u \in \mathrm{BV}([0,T];X)$ *such that*

$$\overline{\mathrm{U}}_{\tau_k,\varepsilon_k}(t) \to u(t) \quad \textit{for all } t\in[0,T] \textit{ as } k\uparrow+\infty,$$

and u *is a BV solution for* $(X,\mathcal{E},\mathfrak{p})$ *in the sense of Definition* 4.1.

The reader may compare this result to [11,18,23,42,48], where similar double-limit passages were performed for specific applied problems, and conditions analogous to (4.6) were imposed.

Proof. The standard energy estimate associated with the variational problem ($\mathrm{IP}_{\tau,\varepsilon}$) yields

$$\frac{\tau}{\varepsilon}\Psi\Big(\frac{\varepsilon}{\tau}(\mathrm{U}^n_{\tau,\varepsilon}-\mathrm{U}^{n-1}_{\tau,\varepsilon})\Big)+\mathcal{E}_{t_n}(\mathrm{U}^n_{\tau,\varepsilon}) \le \mathcal{E}_{t_n}(\mathrm{U}^{n-1}_{\tau,\varepsilon}) = \mathcal{E}_{t_{n-1}}(\mathrm{U}^{n-1}_{\tau,\varepsilon}) + \int_{t_{n-1}}^{t_n} \partial_t\mathcal{E}_t(\mathrm{U}^{n-1}_{\tau,\varepsilon})\,\mathrm{d}t. \tag{4.7}$$

Thanks to (2.7), we easily get from (4.7) the following uniform bounds for every $1\le n\le N$ (here C is a constant independent of n,τ,ε)

$$\mathcal{E}_{t_n}(\mathrm{U}^n_{\tau,\varepsilon}) \le C, \quad \sum_{n=1}^N \frac{\tau}{\varepsilon}\Psi\Big(\frac{\varepsilon}{\tau}(\mathrm{U}^n_{\tau,\varepsilon}-\mathrm{U}^{n-1}_{\tau,\varepsilon})\Big) \le C, \quad \sum_{n=1}^N \Psi_0(\mathrm{U}^n_{\tau,\varepsilon}-\mathrm{U}^{n-1}_{\tau,\varepsilon}) \le C,$$

the latter estimate thanks to (Ψ.2) and (2.15).

Denoting by $\underline{\mathrm{U}}_{\tau,\varepsilon}$ the *right-continuous piecewise constant interpolant* of the discrete values $(\mathrm{U}^n_{\tau,\varepsilon})$ (*i.e.* $\underline{\mathrm{U}}_{\tau,\varepsilon}(t_n)=\mathrm{U}^n_{\tau,\varepsilon}$) and by $\mathrm{U}_{\tau,\varepsilon}$ the *piecewise affine interpolant*, we have

$$\mathcal{E}_t(\overline{\mathrm{U}}_{\tau,\varepsilon}(t)) \le C, \quad \mathrm{Var}_{\Psi_0}(\mathrm{U}_{\tau,\varepsilon};[0,T]) \le C \tag{4.8a}$$

$$\|\mathrm{U}_{\tau,\varepsilon}-\overline{\mathrm{U}}_{\tau,\varepsilon}\|_{L^\infty(0,T;X)},\ \|\mathrm{U}_{\tau,\varepsilon}-\underline{\mathrm{U}}_{\tau,\varepsilon}\|_{L^\infty(0,T;X)}, \quad \sup_n \|\mathrm{U}^n_{\tau,\varepsilon}-\mathrm{U}^{n-1}_{\tau,\varepsilon}\|_X \le C\omega\big(\tau/(C\varepsilon)\big), \tag{4.8b}$$

where $\omega(r) := \sup_{x\in X}\{\|x\|_X : r\Psi(r^{-1}x)\le 1\}$ satisfies $\lim_{r\downarrow 0}\omega(r)=0$ thanks to (Ψ.1). By Helly's theorem, these bounds show that (up to extraction of a suitable subsequence still satisfying (4.6)), the sequences $(\mathrm{U}_{\tau_k,\varepsilon_k})$, $(\overline{\mathrm{U}}_{\tau_k,\varepsilon_k})$ and $(\underline{\mathrm{U}}_{\tau_k,\varepsilon_k})$ pointwise converge to the same limit u.

By differentiating the variational characterization of $\mathrm{U}^n_{\tau,\varepsilon}$ given by ($\mathrm{IP}_{\tau,\varepsilon}$) we obtain

$$\partial\Psi_\varepsilon\Big(\frac{1}{\tau}\big(\mathrm{U}^n_{\tau,\varepsilon}-\mathrm{U}^{n-1}_{\tau,\varepsilon}\big)\Big)+W^n_{\tau,\varepsilon} \ni 0, \quad W^n_{\tau,\varepsilon} := -\mathrm{D}\mathcal{E}_{t_n}(\mathrm{U}^n_{\tau,\varepsilon}),$$

which yields in each interval $(t_{n-1},t_n]$

$$\begin{aligned}\tau\Psi_\varepsilon\big(\dot{\mathrm{U}}_{\tau,\varepsilon}\big)+\tau\Psi^*_\varepsilon(\overline{\mathrm{W}}_{\tau,\varepsilon}) &= -\big\langle \mathrm{D}\mathcal{E}_{t_n}(\mathrm{U}_{\tau,\varepsilon}(t_n)),\mathrm{U}_{\tau,\varepsilon}(t_n)-\mathrm{U}_{\tau,\varepsilon}(t_{n-1})\big\rangle\\ &= \mathcal{E}_{t_{n-1}}(\mathrm{U}_{\tau,\varepsilon}(t_{n-1}))-\mathcal{E}_{t_n}(\mathrm{U}_{\tau,\varepsilon}(t_n))+\int_{t_{n-1}}^{t_n}\partial_t\mathcal{E}_t(\underline{\mathrm{U}}_{\tau,\varepsilon}(t))\,\mathrm{d}t - R(t_n;\mathrm{U}_{\tau,\varepsilon}(t_{n-1}),\mathrm{U}_{\tau,\varepsilon}(t_n))\end{aligned}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

where $\overline{\mathrm{W}}_{\tau,\varepsilon}$ denotes the left-continuous piecewise constant interpolant of $(W^n_{\tau,\varepsilon})_{n=1}^N$ and

$$R(t;x,y) := \mathcal{E}_t(y) - \mathcal{E}_t(x) - \langle \mathrm{D}\mathcal{E}_t(y), y-x\rangle.$$

Since $\mathcal{E}$ is of class C^1, for every convex and bounded set $B \subset X$ there exists a concave modulus of continuity $\sigma_B : [0,+\infty) \to [0,+\infty)$ such that $\lim_{r\downarrow 0} \sigma_B(r) = \sigma_B(0) = 0$ and

$$R(t;x,y) \le \sigma_B(\|y-x\|_X)\|y-x\|_X \quad \text{for every } t \in [0,T],\ x,y \in B.$$

We thus obtain

$$\int_0^T \Big(\Psi_\varepsilon(\dot{\mathrm{U}}_{\tau,\varepsilon}(t)) + \Psi^*_\varepsilon(\overline{\mathrm{W}}_{\tau,\varepsilon}(t))\Big)\,\mathrm{d}t + \mathcal{E}_{t_N}(\mathrm{U}_{\tau,\varepsilon}(t_N))$$
$$\le \mathcal{E}_0(u_0) + \int_0^{t_N} \partial_t\mathcal{E}_t(\underline{\mathrm{U}}_{\tau,\varepsilon}(t))\,\mathrm{d}t + \sup_{1\le n\le N} \sigma_B(\|\mathrm{U}^n_{\tau,\varepsilon} - \mathrm{U}^{n-1}_{\tau,\varepsilon}\|_X) \sum_{n=1}^N \|\mathrm{U}^n_{\tau,\varepsilon} - \mathrm{U}^{n-1}_{\tau,\varepsilon}\|_X. \quad (4.9)$$

We pass to the limit along suitable subsequences $(\tau_k,\varepsilon_k)_k$ such that $\mathrm{U}_{\tau_k,\varepsilon_k}$, $\overline{\mathrm{U}}_{\tau_k,\varepsilon_k} \to u$ pointwise. Since $\mathrm{U}_{\tau,\varepsilon}$ and $\overline{\mathrm{U}}_{\tau,\varepsilon}$ are uniformly bounded, (4.8a), (4.8b), and (4.6) yield the convergence to 0 of the third term on the right-hand side of (4.9), which thus tends to $\mathcal{E}_0(u_0) + \int_0^T \partial_t\mathcal{E}_t(u(t))\,\mathrm{d}t$. Taking into account that $\overline{\mathrm{W}}_{\tau_k,\varepsilon_k}(t) \to w(t) = -\mathrm{D}\mathcal{E}_t(u(t))$ for all $t \in [0,T]$, and applying the lower semicontinuity result (4.4) (proved in Prop. 6.14) we see that u satisfies ($\mathrm{E}_{\mathfrak{p},\mathcal{E};\mathrm{ineq}}$). Similarly,

$$\int_0^T \Psi_0{}^*(w(t))\,\mathrm{d}t \le \liminf_{k\to+\infty} \int_0^T \Psi^*_{\varepsilon_k}(\overline{\mathrm{W}}_{\tau_k,\varepsilon_k}(t))\Big)\,\mathrm{d}t \le C,$$

which implies $w(t) = -\mathrm{D}\mathcal{E}_t(u(t)) \in K^*$ for a.a. $t \in [0,T]$, *i.e.* the local stability condition. In view of Proposition 4.2, this concludes the proof. □

5. Parameterized rate-independent solutions

In this section, we restart from the discussions in Sections 2.4 and 2.5, and adopt a different point of view, which relies on the rate-independent structure of the limit problem. The main idea, which was introduced by [18] and further developed in [40,42], is to rescale time in order to gain a uniform Lipschitz bound on the (rescaled) viscous approximations. Keeping track of the asymptotic behavior of time rescalings, one can retrieve the BV limit analyzed in Section 4. In particular, we shall recover that the limiting jump paths reflect the viscous approximation.

5.1. **Vanishing-viscosity analysis** *via* **a rescaling argument**

For every $\varepsilon > 0$ we consider solutions u_ε of the viscous differential inclusion

$$\partial\Psi_\varepsilon(\dot{u}_\varepsilon(t)) + \mathrm{D}\mathcal{E}_t(u_\varepsilon(t)) \ni 0 \quad \text{in } X^* \quad \text{for a.a. } t \in (0,T), \qquad (\mathrm{DN}_\varepsilon)$$

which may be split into the system

$$\partial\Psi_\varepsilon(\dot{u}_\varepsilon(t)) \ni w_\varepsilon, \qquad \mathrm{D}\mathcal{E}_t(u_\varepsilon(t)) = -w_\varepsilon, \qquad \partial_t\mathcal{E}_t(u_\varepsilon(t)) = -p_\varepsilon.$$

For capturing the aforementioned limiting viscous jump paths, we do not follow the ideas of [18,40,42], but use a different time rescaling $\mathsf{s}_\varepsilon : [0,T] \to [0,\mathsf{S}_\varepsilon]$ based on the dissipation bound (2.28), instead. We define

$$\mathsf{s}_\varepsilon(t) := t + \int_0^t \Big(\Psi_\varepsilon(\dot{u}_\varepsilon(r)) + \Psi^*_\varepsilon(w_\varepsilon(r))\Big)\,\mathrm{d}r \quad \text{and} \quad \mathsf{T}_\varepsilon := \mathsf{s}_\varepsilon(T). \qquad (5.1)$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

We believe that this time rescaling is better adapted for treating general systems, in particular infinite-dimensional ones, where the rate-independent norm induced by Ψ_0 is not equivalent to the viscous norm. We refer to [17] for an application of our idea to elastoplasticity.

Thus, s_ε may be interpreted as some sort of "energy arclength" of the curve u_ε. Thanks to (2.28), the sequence (T_ε) is uniformly bounded with respect to the parameter ε. We consider the rescaled functions $(\mathsf{t}_\varepsilon, \mathsf{u}_\varepsilon) : [0, \mathsf{T}_\varepsilon] \to [0, T] \times X$ and $(\mathsf{p}_\varepsilon, \mathsf{w}_\varepsilon) : [0, \mathsf{T}_\varepsilon] \to \mathbb{R} \times X^*$ defined by

$$\begin{aligned} &\mathsf{t}_\varepsilon(s) := \mathsf{s}_\varepsilon^{-1}(s), && \mathsf{u}_\varepsilon(s) := u_\varepsilon(\mathsf{t}_\varepsilon(s)), \\ &\mathsf{p}_\varepsilon(s) := p_\varepsilon(\mathsf{t}_\varepsilon(s)) = -\partial_t \mathcal{E}_{\mathsf{t}_\varepsilon(s)}(\mathsf{u}_\varepsilon(s)), && \mathsf{w}_\varepsilon(s) := w_\varepsilon(\mathsf{t}_\varepsilon(s)) = -\mathrm{D}\mathcal{E}_{\mathsf{t}_\varepsilon(s)}(\mathsf{u}_\varepsilon(s)). \end{aligned} \tag{5.2}$$

We now study the limiting behavior as $\varepsilon \downarrow 0$ of the *reparameterized trajectories*

$$\begin{aligned} \big\{ \big(\mathsf{t}_\varepsilon(s), \mathsf{u}_\varepsilon(s)\big) \ : \ s \in [0, \mathsf{T}_\varepsilon] \big\} &\subset \mathscr{X} = [0, T] \times X, \\ \big\{ \big(\dot{\mathsf{t}}_\varepsilon(s), \dot{\mathsf{u}}_\varepsilon(s); \mathsf{p}_\varepsilon(s), \mathsf{w}_\varepsilon(s)\big) \ : \ s \in [0, \mathsf{T}_\varepsilon] \big\} &\subset \mathscr{B}, \end{aligned}$$

where we use the notation

$$\mathscr{B} := [0, +\infty) \times X \times \mathbb{R} \times X^*. \tag{5.3}$$

In order to rewrite the "rescaled energy identity" fulfilled by the triple $(\mathsf{t}_\varepsilon, \mathsf{u}_\varepsilon, \mathsf{w}_\varepsilon)$, we define, for $\varepsilon > 0$, the viscous space-time vanishing-viscosity contact potential $\mathfrak{P}_\varepsilon : \mathscr{B} \to [0, +\infty)$ by setting

$$\mathfrak{P}_\varepsilon(\alpha, \mathsf{v}; \mathsf{p}, \mathsf{w}) := \begin{cases} \alpha \Psi_\varepsilon(\mathsf{v}/\alpha) + \alpha \Psi_\varepsilon^*(\mathsf{w}) + \alpha \mathsf{p} & \text{for } \alpha > 0, \\ \infty & \text{for } \alpha = 0. \end{cases} \tag{5.4}$$

By the definition of Ψ_ε, we see that $\mathfrak{P}_\varepsilon(\alpha, \cdot, \mathsf{p}, \cdot)$ only depends on α/ε and $\alpha \mathsf{p}$ (as long as $\alpha > 0$), which reflects the time rescaling. In particular, the energy identity (2.27) yields

$$\int_{s_1}^{s_2} \mathfrak{P}_\varepsilon\big(\dot{\mathsf{t}}_\varepsilon(s), \dot{\mathsf{u}}_\varepsilon(s); \mathsf{p}_\varepsilon(s), \mathsf{w}_\varepsilon(s)\big) \, \mathrm{d}s + \mathcal{E}_{\mathsf{t}_\varepsilon(s_2)}(\mathsf{u}_\varepsilon(s_2)) = \mathcal{E}_{\mathsf{t}_\varepsilon(s_1)}(\mathsf{u}_\varepsilon(s_1)), \tag{5.5}$$

for all $0 \le s_1 \le s_2 \le \mathsf{S}_\varepsilon$, and (5.1) entails

$$\mathfrak{P}_\varepsilon\big(\dot{\mathsf{t}}_\varepsilon(s), \dot{\mathsf{u}}_\varepsilon(s); 1, \mathsf{w}_\varepsilon(s)\big) = 1 \quad \text{for a.a. } s \in (0, \mathsf{T}_\varepsilon). \tag{5.6}$$

***A priori* estimates and passage to the limit.** Due to estimate (2.28), there exists $\mathsf{S} > 0$ such that, along a (not relabeled) subsequence, we have $\mathsf{s}_\varepsilon(T) \to \mathsf{S}$ as $\varepsilon \downarrow 0$. Exploiting again (2.28), the Arzelà-Ascoli compactness theorem, and the fact that X is finite dimensional (see also the proof of [42], Thm. 3.3), we find two curves $\mathsf{t} \in W^{1,\infty}(0, \mathsf{S})$ and $\mathsf{u} \in W^{1,\infty}([0, \mathsf{S}]; X)$ such that along some subsequence we have

$$\mathsf{t}_\varepsilon \to \mathsf{t} \text{ in } \mathrm{C}^0([0, \mathsf{S}]), \qquad \dot{\mathsf{t}}_\varepsilon \rightharpoonup^* \dot{\mathsf{t}} \text{ in } L^\infty(0, \mathsf{S}), \tag{5.7a}$$

$$\mathsf{u}_\varepsilon \to \mathsf{u} \text{ in } \mathrm{C}^0([0, \mathsf{S}]; X), \qquad \dot{\mathsf{u}}_\varepsilon \rightharpoonup^* \dot{\mathsf{u}} \text{ in } L^\infty(0, \mathsf{S}; X), \tag{5.7b}$$

$$\mathsf{p}_\varepsilon \to \mathsf{p} \text{ in } \mathrm{C}^0([0, \mathsf{S}]), \qquad \mathsf{w}_\varepsilon \to \mathsf{w} \text{ in } \mathrm{C}^0([0, \mathsf{S}]; X^*), \tag{5.7c}$$

$$\mathcal{E}_{\mathsf{t}_\varepsilon}(\mathsf{u}_\varepsilon) \to \mathcal{E}_{\mathsf{t}}(\mathsf{u}) \text{ in } \mathrm{C}^0([0, \mathsf{S}]), \tag{5.7d}$$

where

$$\mathsf{p}(s) = -\partial_t \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \text{ and } \mathsf{w}(s) = -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \text{ for all } s \in [0, \mathsf{S}]. \tag{5.8}$$

Then, to pass to the limit in (5.5) we exploit a lower semicontinuity result (see Prop. 6.2), based on the fact that the sequence of functionals $(\mathfrak{P}_\varepsilon)$ Γ-converges to the *augmented vanishing-viscosity contact potential*

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

$\mathfrak{P} : \mathscr{B} \to [0, +\infty]$ (see Lem 6.1) defined by

$$\mathfrak{P}(\alpha, v; p, w) := \begin{cases} \Psi_0(v) + \mathrm{I}_{K^*}(w) + \alpha\, p & \text{if } \alpha > 0, \\ \mathfrak{p}(v, w) & \text{if } \alpha = 0. \end{cases} \tag{5.9}$$

By (5.7) and Proposition 6.2, we take the lim inf as $\varepsilon \downarrow 0$ of (5.5) and conclude that the pair (t, u) fulfills, for all $0 \le s_1 \le s_2 \le \mathsf{S}$, the estimate

$$\int_{s_1}^{s_2} \mathfrak{P}\big(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s); \mathsf{p}(s), \mathsf{w}(s)\big)\, \mathrm{d}s + \mathcal{E}_{\mathsf{t}(s_2)}(\mathsf{u}(s_2)) \le \mathcal{E}_{\mathsf{t}(s_1)}(\mathsf{u}(s_1)). \tag{5.10}$$

The notion of parameterized solutions looks much simpler than that of BV solutions, since after reparameterization one always deals with Lipschitz curves avoiding the more subtle BV setting. Hence, it is sufficient to use the comparably simple lower semicontinuity result in Proposition 6.2. To obtain a similar convergence result for BV solutions we have to use the more intricate Proposition 6.14. On the other hand, the notion of BV solutions is well suited to study the asymptotic behavior of the viscosity approximations in the original time interval. In this case, even adopting the point of view of parameterized solutions, one should keep track of the ε-dependent time rescaling to operate a final reconstruction of the limit solution. In this respect, passing to the limit in the energy identity is particularly important, in order to obtain a solution that is independent of the particular form of the rescaling. This somehow delicate procedure is hidden in the notion of BV solution and in some sense justifies its intrinsic difficulty.

5.2. **Vanishing-viscosity contact potentials and rate-independent evolution**

The augmented space-time vanishing-viscosity contact potential $\mathfrak{P}$ is closely related to the contact potential $\mathfrak{p}$ introduced by (3.13). The following result fixes some properties of $\mathfrak{P}$. Its proof, which we choose to omit, can be easily derived from the corresponding properties ($\mathfrak{p}1$)–($\mathfrak{p}4$) for the vanishing-viscosity contact potential $\mathfrak{p}$ obtained in Theorem 3.7.

Lemma 5.1 (properties of the augmented vanishing-viscosity contact potential $\mathfrak{P}$)**.** *The functional $\mathfrak{P} : \mathscr{B} \to [0, +\infty]$ defined in (5.9) is lower semicontinuous on $\mathscr{B}$ and has the following properties:*

($\mathfrak{P}1$) *For every $(\alpha, v, p, w) \in \mathscr{B}$ the maps $\mathfrak{P}(\cdot, \cdot; p, w)$ and $\mathfrak{P}(\alpha, v; \cdot, \cdot)$ have convex sublevels.*

($\mathfrak{P}2$) *For all $(\alpha, v, p, w) \in \mathscr{B}$ we have*

$$\mathfrak{P}(\alpha, v; p, w) \ge \langle w, v\rangle + \alpha p, \quad \mathfrak{P}(0, v; p, w) = \mathfrak{p}(v, w) \ge \Psi_0(v), \tag{5.11a}$$

$$\mathfrak{P}(0, v; p, w) = \Psi_0(v) \quad \Leftrightarrow \quad w \in K^* \quad \text{or} \quad v = 0. \tag{5.11b}$$

($\mathfrak{P}3$) *For all $(p, w) \in \mathbb{R} \times X^*$ the map $(\alpha, v) \mapsto \mathfrak{P}(\alpha, v; p, w)$ is 1-homogeneous and convex.*

($\mathfrak{P}4$) *The* augmented contact set $\Sigma_{\mathfrak{P}}$ *of $\mathfrak{P}$ given by*

$$\Sigma_{\mathfrak{P}} := \Big\{(\alpha, v; p, w) \in \mathscr{B} : \mathfrak{P}(\alpha, v; p, w) = \langle w, v\rangle + \alpha p\Big\} \tag{5.12}$$

can be characterized by

$$(\alpha, v; p, w) \in \Sigma_{\mathfrak{P}} \quad \Leftrightarrow \quad w \in \partial\, \mathfrak{P}(\alpha, \cdot\,; p, w)(v). \tag{5.13}$$

In particular, (5.13) does not impose any condition on p. We also have

$$\text{if } \alpha > 0, \text{ then} \qquad (\alpha, v; p, w) \in \Sigma_{\mathfrak{P}} \quad \Leftrightarrow \quad w \in \partial \Psi_0(v), \tag{5.14a}$$

$$\text{if } \alpha = 0, \text{ then} \qquad (\alpha, v; p, w) \in \Sigma_{\mathfrak{P}} \quad \Leftrightarrow \quad (v, w) \in \Sigma_{\mathfrak{p}}. \tag{5.14b}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Recalling Lemma 3.10 and Λ from (3.17), for $v \neq 0$ we have $(\alpha, v; p, w) \in \Sigma_{\mathfrak{P}}$ if and only if

$$w \in \partial\Psi_0(v) \subset K^* \quad \text{or} \quad \Big(w \notin K^*,\ \alpha = 0,\ \ w \in \partial\Psi(\varepsilon v)\Big) \quad \text{for every } \varepsilon \in \Lambda(v, w). \tag{5.15}$$

In particular, in the additive viscosity case (2.23), for $v \neq 0$ we simply have

$$(\alpha, v; p, w) \in \Sigma_{\mathfrak{P}} \quad \Longleftrightarrow \quad \exists\, \lambda \geq 0 :\ w \in \partial\Psi_0(v) + \partial\Psi_V(\lambda v) \ \text{ and } \ \alpha\lambda = 0. \tag{5.16}$$

Conclusion of the vanishing-viscosity analysis. We are now going to show that (5.10) is in fact an equality and that the parameterization condition (5.6) is retained in the limit $\varepsilon \to 0$. For this, we rely on the chain rule (2.26), which yields for a.a. $s \in (0, \mathsf{S})$

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \ &= \ \partial_t\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))\,\dot{\mathsf{t}}(s) - \langle -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)), \dot{\mathsf{u}}(s)\rangle \\ &\overset{(5.8)}{=} -\mathsf{p}(s)\dot{\mathsf{t}}(s) - \langle \mathsf{w}(s), \dot{\mathsf{u}}(s)\rangle \overset{(5.11a)}{\geq} -\mathfrak{P}(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s), \mathsf{p}(s), \mathsf{w}(s)). \end{aligned} \tag{5.17}$$

Combining (5.17) and (5.10), we conclude that the latter holds with an equality sign and, with an elementary argument, that such equality also holds in the differential form, *i.e.* for a.a. $s \in (0, \mathsf{S})$

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) = -\mathsf{p}(s)\dot{\mathsf{t}}(s) - \langle \mathsf{w}(s), \dot{\mathsf{u}}(s)\rangle = -\mathfrak{P}(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s), \mathsf{p}(s), \mathsf{w}(s)), \tag{5.18}$$

where (5.8) is used. By the definition of $\Sigma_{\mathfrak{P}}$ we have

$$\Big(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s); -\partial_t\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)), -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))\Big) \in \Sigma_{\mathfrak{P}} \quad \text{for a.a. } s \in (0, \mathsf{S}). \tag{5.19}$$

Finally, we want to show that the limit (t, u) is still parameterized in the energy-arclength sense (5.6), namely

$$\mathfrak{P}\big(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s); 1, \mathsf{w}(s)\big) = 1 \quad \text{for a.a. } s \in (0, \mathsf{S}). \tag{5.20}$$

To do so, we let $\pi_\varepsilon(s) = \mathfrak{P}_\varepsilon(\dot{\mathsf{t}}_\varepsilon(s), \dot{\mathsf{u}}_\varepsilon(s); \mathsf{p}_\varepsilon(s), \mathsf{w}_\varepsilon(s))$ and $\pi(s) = \mathfrak{P}(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s); \mathsf{p}(s), \mathsf{w}(s))$. Our convergence established above gives $\pi_\varepsilon \rightharpoonup^* \pi$ in $L^\infty(0, \mathsf{S})$. Using $\mathfrak{P}_\varepsilon(\alpha, v; 1, w) = \mathfrak{P}_\varepsilon(\alpha, v; p, w) + \alpha(1 - p)$ we find

$$1 \overset{(5.6)}{=} \mathfrak{P}_\varepsilon(\dot{\mathsf{t}}_\varepsilon, \dot{\mathsf{u}}_\varepsilon; 1, \mathsf{w}_\varepsilon) = \dot{\mathsf{t}}_\varepsilon(1 - \mathsf{p}_\varepsilon) + \pi_\varepsilon \rightharpoonup^* \dot{\mathsf{t}}(1 - \mathsf{p}) + \pi = \mathfrak{P}(\dot{\mathsf{t}}, \dot{\mathsf{u}}; 1, \mathsf{w}),$$

where "$\rightharpoonup^*$" follows from (5.7). Thus, (5.20) is established.

5.3. **Parameterized solutions of rate-independent systems**

Motivated by the discussion of the previous section, we now give the notion of *parameterized rate-independent solutions*, driven by a general augmented vanishing-viscosity contact potential $\mathfrak{P}$, satisfying conditions ($\mathfrak{P}1$)–($\mathfrak{P}3$) of Lemma 5.1. To denote these solutions we will use the triple $(X, \mathcal{E}, \mathfrak{P})$. In contrast to [18,40], we do not assume the usage of a particular arclength parameterization. We rather keep the parameterization as free as possible and thus display the rate-independence of this solution concept more clearly.

Definition 5.2 (parameterized rate-independent solutions)**.** Let $\mathfrak{P} : \mathscr{B} \to (-\infty, +\infty]$ be the augmented vanishing-viscosity contact potential (5.9). A Lipschitz continuous curve $(\mathsf{t}, \mathsf{u}) : [a, b] \to [0, T] \times X$ is called a *parameterized rate-independent solution* for $(X, \mathcal{E}, \mathfrak{P})$, if t is nondecreasing and

$$\int_{s_1}^{s_2} \mathfrak{P}(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s); \mathsf{p}(s), \mathsf{w}(s))\,\mathrm{d}s + \mathcal{E}_{\mathsf{t}(s_2)}(\mathsf{u}(s_2)) \leq \mathcal{E}_{\mathsf{t}(s_1)}(\mathsf{u}(s_1)) \quad \text{for all } a \leq s_1 \leq s_2 \leq b, \tag{5.21}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

where $\mathsf{p}(s) = -\partial_t \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))$ and $\mathsf{w}(s) = -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))$. Furthermore, we call (t,u):

- *nondegenerate*, if $\dot{\mathsf{t}}(s) + \Psi_0(\dot{\mathsf{u}}(s)) > 0$ for a.a. $s \in (a,b)$;
- *surjective*, if $\mathsf{t}(a) = 0, \mathsf{t}(b) = T$;
- *normalized*, if (t,u) satisfies (5.20).

Definition 5.2 generalizes to the present setting the notion that was first introduced in [42]. The rate-independence of this definition manifests itself in its invariance with respect to (nondecreasing, Lipschitz) time rescalings. Namely, if $(\mathsf{t},\mathsf{u}) : [a,b] \to [0,T] \times X$ is a parameterized solution and $\mathsf{s} : [\alpha,\beta] \to [a,b]$ is a Lipschitz nondecreasing map, then $(\mathsf{t}\circ\mathsf{s}, \mathsf{u}\circ\mathsf{s})$ is a parameterized solution in $[\alpha,\beta]$. Surjectivity is always preserved; nondegeneracy is maintained only if $\dot{\mathsf{s}} > 0$ a.e. in (α,β).

The next result provides several equivalent characterizations of parameterized solutions.

Proposition 5.3. *A Lipschitz continuous curve $(\mathsf{t},\mathsf{u}) : [a,b] \to [0,T] \times X$, with t nondecreasing, is a parameterized solution of $(X,\mathcal{E},\mathfrak{P})$ if and only if one of the following (equivalent) conditions is satisfied (where as usual $\mathsf{p} = -\partial_t \mathcal{E}_{\mathsf{t}}(\mathsf{u}), \mathsf{w} = -\mathrm{D}\mathcal{E}_{\mathsf{t}}(\mathsf{u})$):*

(1) *The energy inequality (5.21) holds just for $s_1 = a$ and $s_2 = b$,* i.e.

$$\int_a^b \mathfrak{P}(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s); \mathsf{p}(s), \mathsf{w}(s))\,\mathrm{d}s + \mathcal{E}_{\mathsf{t}(b)}(\mathsf{u}(b)) \le \mathcal{E}_{\mathsf{t}(a)}(\mathsf{u}(a)). \tag{5.22}$$

(2) *The energy inequality (5.21) holds in the differential form*

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) + \mathfrak{P}(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s); \mathsf{p}(s), \mathsf{w}(s)) \le 0 \quad \textit{for a.a. } s \in (a,b). \tag{5.23}$$

(3) *The energy identity holds, either in the differential form*

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) + \mathfrak{P}(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s); \mathsf{p}(s), \mathsf{w}(s)) = 0 \quad \textit{for a.a. } s \in (a,b), \tag{5.24}$$

or in the integrated form

$$\int_{s_1}^{s_2} \mathfrak{P}(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s); \mathsf{p}(s), \mathsf{w}(s))\,\mathrm{d}s + \mathcal{E}_{\mathsf{t}(s_2)}(\mathsf{u}(s_2)) = \mathcal{E}_{\mathsf{t}(s_1)}(\mathsf{u}(s_1)) \quad \textit{for } a \le s_1 \le s_2 \le b. \tag{5.25}$$

(4) *There holds*

$$\big(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s); -\partial_t \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)), -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))\big) \in \Sigma_{\mathfrak{P}} \quad \textit{for a.a. } s \in (a,b).$$

(5) *The pair (t,u) satisfies the differential inclusion*

$$\partial\mathfrak{P}\big(\dot{\mathsf{t}}(s), \cdot\,; -\partial_t \mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)), -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))\big)(\dot{\mathsf{u}}(s)) + \mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \ni 0 \quad \textit{a.e. in } (a,b). \tag{5.26}$$

In particular, (1)–(5) *imply that for a.a. $s \in (a,b)$ we have the implications*

$$\begin{aligned} \dot{\mathsf{t}}(s) > 0 \quad &\Rightarrow \quad -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \in K^*, \\ -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \in K^* \quad &\Rightarrow \quad -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \in \partial\Psi_0(\dot{\mathsf{u}}(s)). \end{aligned} \tag{5.27}$$

Moreover, if (t,u) is nondegenerate, for every Borel map $\lambda : \mathsf{J} \to [0,+\infty)$ defined in the open set $\mathsf{J} := \big\{s \in (a,b) : -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \notin K^*\big\}$ and satisfying $\lambda(s) \in \Lambda(\dot{\mathsf{u}}(s), -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)))$ for a.a. $s \in \mathsf{J}$ we have

$$-\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \in \partial\Psi(\lambda(s)\dot{\mathsf{u}}(s)), \quad \dot{\mathsf{t}}(s) = 0 \quad \textit{for a.a. } s \in \mathsf{J}. \tag{5.28}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

The proof follows from the chain rule (2.26) (arguing as for (5.17), (5.18), (5.19)), and from the characterization of the contact set $\Sigma_{\mathfrak{P}}$ of Lemma 5.1 (see also [42], Prop. 2).

Corollary 5.4 (differential characterization in the case of additive viscosity)**.** *Let $\mathfrak{P} : \mathscr{B} \to (-\infty, +\infty]$ be an augmented vanishing-viscosity contact potential satisfying conditions ($\mathfrak{P}1$)–($\mathfrak{P}3$) of Lemma 5.1, and suppose also that the augmented contact set $\Sigma_{\mathfrak{P}}$ satisfies the characterization (5.16) in the additive viscosity case (2.23) $\Psi = \Psi_0 + \Psi_V$.*

Then, a nondegenerate Lipschitz continuous curve $(\mathsf{t}, \mathsf{u}) : [a, b] \to [0, T] \times X$ is a parameterized rate-independent solution of $(X, \mathcal{E}, \mathfrak{P})$, if and only if there exists a Borel function $\lambda : (a, b) \to [0, +\infty)$ such that

$$\partial\Psi_0(\dot{\mathsf{u}}(s)) + \partial\Psi_V(\lambda(s)\dot{\mathsf{u}}(s)) + \mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)) \ni 0 \quad \text{and} \quad \lambda(s)\dot{\mathsf{t}}(s) = 0 \quad \text{for a.a. } s \in (a, b). \tag{5.29}$$

With the vanishing-viscosity analysis developed in Sections 5.1 and 5.2 we arrive at the following convergence result. Recall that it is based on the forthcoming lower semicontinuity result in Proposition 6.2, see the discussion before (5.10).

Theorem 5.5 (convergence to parameterized solutions)**.** *Let (u_n) be viscous solutions of (DN_ε) corresponding to a vanishing sequence (ε_n), let $\mathsf{t}_n : [0, \mathsf{S}] \to [0, T]$ be uniformly Lipschitz and surjective time rescalings and let $\mathsf{u}_n : [0, \mathsf{S}] \to X$ be defined as $\mathsf{u}_n(s) := u_n(\mathsf{t}_n(s))$ for all $s \in [0, \mathsf{S}]$. Suppose that*

$$\exists\, \delta > 0 \quad \forall\, n \in \mathbb{N} : \quad \mathsf{m}_n(s) := \mathfrak{P}_{\varepsilon_n}(\dot{\mathsf{t}}_n(s), \dot{\mathsf{u}}_n(s); 1, -\mathrm{D}\mathcal{E}_{\mathsf{t}_n(s)}(\mathsf{u}_n(s))) \in [\delta, \delta^{-1}]$$

for a.a. $s \in (0, \mathsf{S})$. If $(\mathsf{t}_n, \mathsf{u}_n)$ converges pointwise to (t, u) and $\mathsf{m}_n \rightharpoonup^* \mathsf{m}$ as $n \to \infty$, then (t, u) is a nondegenerate, surjective parameterized rate-independent solution according to Definition 5.2, and

$$\mathfrak{P}(\dot{\mathsf{t}}(s), \dot{\mathsf{u}}(s); 1, -\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))) = \mathsf{m}(s) \qquad \text{for a.a. } s \in (0, \mathsf{S}). \tag{5.30}$$

The convergence of the parameterization stated in (5.30) (and proved in the last part of Sect. 5.2) is useful for showing that the limit (t, u) remains nondegenerate or normalized, if this holds uniformly along the sequence $(\mathsf{t}_n, \mathsf{u}_n)$.

The following remark (to be compared with the classification ($\mathrm{O}_{\mathrm{sliding}}$), ($\mathrm{O}_{\mathrm{viscous}}$), and ($\mathrm{O}_{\mathrm{ener}}$) for optimal transitions in Definition 4.4 and the detailed analysis in Appendix B), highlights the different mechanical regimes encompassed in the notion of parameterized rate-independent solution.

Remark 5.6 (mechanical interpretation)**.** The evolution described by (5.26) in Proposition 5.3 bears the following mechanical interpretation (*cf.* [18,42]):

- The regime ($\dot{\mathsf{t}} > 0$, $\dot{\mathsf{u}} = 0$) corresponds to *sticking* (with $-\mathrm{D}\mathcal{E}_t(u) \in \partial\Psi_0(0) = K^*$).
- The regime ($\dot{\mathsf{t}} > 0$, $\dot{\mathsf{u}} \neq 0$) means *rate-independent sliding* (with $-\mathrm{D}\mathcal{E}_t(u) \in \partial\Psi_0(\dot{\mathsf{u}}) \subset K^*$).
- If $-\mathrm{D}\mathcal{E}_{\mathsf{t}}(\mathsf{u})$ cannot obey the constraint K^*, then the system switches to a *viscous regime*. The time is frozen (*i.e.*, $\dot{\mathsf{t}} = 0$), and the solution follows a viscous path. In the additive viscosity case (2.23) it is governed by the rescaled viscous equation (5.29) with $\lambda > 0$. This viscous motion can be seen as a jump in the (slow) external time scale.

We conclude this section with the main equivalence result between parameterized and BV solutions of rate-independent systems (compare with the analogous [42], Prop. 6). We postpone its proof at the very end of Section 6.3.

Theorem 5.7 (equivalence between BV and parameterized solutions)**.** *Let $(\mathsf{t}, \mathsf{u}) : [0, \mathsf{S}] \to [0, T] \times X$ be a (nondegenerate, surjective) parameterized solution of the rate-independent system $(X, \mathcal{E}, \mathfrak{P})$. For every $t \in [0, T]$ set*

$$\widehat{\mathsf{s}}(t) := \big\{ s \in [0, \mathsf{S}] : \mathsf{t}(s) = t \big\}. \tag{5.31}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Then, any curve $u : [0,T] \to X$ *such that*

$$u(t) \in \big\{ \mathsf{u}(s) : s \in \widehat{\mathsf{s}}(t) \big\} \tag{5.32}$$

is a BV solution of the rate-independent system $(X, \mathcal{E}, \mathfrak{p})$. *Conversely, if* $u : [0,T] \to X$ *is a BV solution, then there exists a nondegenerate, surjective parameterized solution* (t, u) *such that* (5.32) *holds for a time-rescaling function* $\widehat{\mathsf{s}}$ *defined as in* (5.31).

6. Auxiliary results

The focus of Section 6.1 is on Proposition 6.2, which yields the lower semicontinuity result for the integral functional associated with $\mathfrak{P}_\varepsilon$. This result was exploited in the vanishing-viscosity analysis for parameterized solutions in Section 5, see (5.10) and Theorem 5.5. In Section 6.2 we develop some auxiliary results concerning the total variation induced by time-dependent (and possibly asymmetric) Finsler norms. These results form the core of the proof of the main lower semicontinuity result of Section 6.3, namely Proposition 6.14, which is crucial for the convergence results for BV solutions in Theorems 4.9 and 4.10.

6.1. Lower semicontinuity for vanishing-viscosity contact potentials

Let us start with a lemma showing that the sequence $(\mathfrak{P}_\varepsilon)$, which is defined in (5.4), Γ-converges for $\varepsilon \downarrow 0$ to the augmented space-time contact potential $\mathfrak{P}$, defined in (5.9) (*cf.* [42], Lem. 3.1).

Lemma 6.1 (Γ-convergence of $\mathfrak{P}_\varepsilon$)**.**

Γ-liminf estimate: *For every choice of sequences* $\varepsilon_n \downarrow 0$ *and* $(\alpha_n, v_n, p_n, w_n) \to (\alpha, v, p, w)$ *in* $\mathscr{B}$, *we have*

$$\liminf_{n\to\infty} \mathfrak{P}_{\varepsilon_n}(\alpha_n, v_n; p_n, w_n) \geq \mathfrak{P}(\alpha, v; p, w). \tag{6.1}$$

Γ-limsup estimate: *For every* $(\alpha, v; p, w) \in \mathscr{B}$ *there exists* $(\alpha_\varepsilon, v_\varepsilon, p_\varepsilon, w_\varepsilon)_{\varepsilon>0}$ *such that*

$$\limsup_{\varepsilon\downarrow 0} \mathfrak{P}_\varepsilon(\alpha_\varepsilon, v_\varepsilon; p_\varepsilon, w_\varepsilon) \leq \mathfrak{P}(\alpha, v; p, w). \tag{6.2}$$

Proof. The Γ-liminf estimate is easy. If $\alpha > 0$ then, also recalling (2.29), one verifies that

$$\begin{aligned}\liminf_{n\to\infty} \mathfrak{P}_{\varepsilon_n}(\alpha_n, v_n; p_n, w_n) &\overset{(2.15)}{\geq} \liminf_{n\to\infty} \big(\Psi_0(v_n) + \alpha_n \varepsilon_n^{-1} \Psi^*(w_n) + \alpha_n p_n\big) \\ &\overset{(2.17)}{\geq} \Psi_0(v) + \mathrm{I}_K^*(w) + \alpha p = \mathfrak{P}(\alpha, v; p, w).\end{aligned}$$

If $\alpha = 0$, we use the obvious lower bound $\mathfrak{P}_{\varepsilon_n}(\alpha_n, v_n; p_n, w_n) \geq \mathfrak{p}(v_n, w_n) + \alpha_n p_n$ and the continuity of $\mathfrak{p}$, see Theorem 3.7.

To show the limsup estimate (6.2), we first consider $w \in K^*$ giving $\Psi^*(w) = 0$. With $\alpha_\varepsilon := \alpha + \sqrt{\varepsilon}$, $v_\varepsilon := v$, $p_\varepsilon := p$, and $w_\varepsilon := w$, we obtain

$$\mathfrak{P}_\varepsilon(\alpha_\varepsilon, v_\varepsilon; p_\varepsilon, w_\varepsilon) \leq \frac{\alpha + \sqrt{\varepsilon}}{\varepsilon} \Psi\left(\frac{\varepsilon}{\alpha + \sqrt{\varepsilon}} v \right) + (\alpha + \sqrt{\varepsilon}) p \overset{\varepsilon\downarrow 0}{\to} \Psi_0(v) + \alpha p = \mathfrak{P}(\alpha, v; p, w),$$

where the first passage is due to (2.18). This argument works both for $\alpha > 0$ and $\alpha = 0$.

Next we consider $w \notin K^*$. The case $\alpha > 0$ is trivial, since $\mathfrak{P}(\alpha, v; p, w) = +\infty$. For $\alpha = 0$ and $v \neq 0$ we choose $\lambda \in \Lambda(v, w)$ (see (3.17)), let $\alpha_\varepsilon := \varepsilon/\lambda$, $v_\varepsilon := v$, $p_\varepsilon := p$, and $w_\varepsilon := w$, and find

$$\mathfrak{P}_\varepsilon(\alpha_\varepsilon, v_\varepsilon; p_\varepsilon, w_\varepsilon) = \frac{1}{\lambda}\Psi(\lambda v) + \frac{1}{\lambda}\Psi^*(w) + \frac{\varepsilon}{\lambda} p = \mathfrak{p}(v, w) + \frac{\varepsilon p}{\lambda} \overset{\varepsilon\downarrow 0}{\to} \mathfrak{p}(v, w) = \mathfrak{P}(0, v; p, w).$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

In the case $\alpha = 0, v = 0$ we choose $\alpha_\varepsilon := \varepsilon^2$, $v_\varepsilon = 0$, $p_\varepsilon = p$, $w_\varepsilon = w$, obtaining

$$\mathfrak{P}_\varepsilon(\alpha_\varepsilon, v_\varepsilon; p_\varepsilon, w_\varepsilon) = \varepsilon \Psi^*(w) + \varepsilon^2 p \overset{\varepsilon\downarrow 0}{\to} 0 = \mathfrak{P}(0,0;p,w).$$

This proves the result. □

An important consequence of the previous lemma is provided by the following:

Proposition 6.2 (lower semicontinuity of the rescaled dissipation integral)**.** *For a fixed interval (s_0, s_1) choose sequences $\varepsilon_n > 0$, $\alpha_n \in L^\infty(s_0, s_1)$, $p_n \in L^1(s_0, s_1)$, $v_n \in L^1(0, T; X)$, and $w_n \in L^1(0, T; X^*)$ such that $\varepsilon_n \to 0$ and*

$$\begin{aligned} &\alpha_n \rightharpoonup^* \alpha \quad \text{in } L^\infty(s_0, s_1), \qquad p_n \to p \quad \text{in } L^1(0, T),\\ &v_n \rightharpoonup v \quad \text{in } L^1(0, T; X), \qquad w_n \to w \quad \text{in } L^1(s_0, s_1). \end{aligned}$$

Then, we have the liminf estimate

$$\liminf_{n\to\infty} \int_{s_0}^{s_1} \mathfrak{P}_{\varepsilon_n}(\alpha_n(s), v_n(s); p_n(s), w_n(s))\,\mathrm{d}s \geq \int_{s_0}^{s_1} \mathfrak{P}(\alpha(s), v(s); p(s), w(s))\,\mathrm{d}s. \tag{6.3}$$

Proof. It is sufficient to prove this result in the case $p_n \equiv p = 0$, because $\alpha_n p_n$ weakly converges in $L^1(s_0, s_1)$. By Lemma 6.1, the integrand $\tilde{\mathfrak{P}} : [0, +\infty) \times [0, +\infty) \times X \times X^* \to [0, +\infty]$ defined *via*

$$\tilde{\mathfrak{P}}(\varepsilon, \alpha, v, w) := \begin{cases} \mathfrak{P}_\varepsilon(\alpha, v; 0, w) & \text{for } \varepsilon > 0,\\ \mathfrak{P}(\alpha, v; 0, w) & \text{for } \varepsilon = 0, \end{cases}$$

is lower semicontinuous and convex in the pair (α, v). Then, inequality (6.3) follows from Ioffe's theorem (see *e.g.* [3], Thm. 5.8). □

6.2. **Asymmetric dissipations, pseudo-total variation, and extended space-time curves**

Here we discuss Finsler distances on the extended state space $\mathscr{X} := [0, T] \times X$, which are appropriate to tackle parameterized solutions. This gives a general framework to handle the previously defined Finsler distance $\Delta_{\mathfrak{p},\mathcal{E}}(t; u_0, u_1)$.

Notation. The elements of $\mathscr{X} = [0, T] \times X$ will be denoted by bold letters, namely $\boldsymbol{x} = (t, u)$. We shall denote by $\mathscr{V}$ the tangent cone $[0, +\infty) \times X$ to $\mathscr{X}$ and by $\boldsymbol{v} = (\alpha, v)$ the elements in $\mathscr{V}$.

We shall consider lower semicontinuous dissipation functionals $\mathcal{R} : \mathscr{X} \times \mathscr{V} \to [0, +\infty]$ satisfying the following properties:

$$\forall\, \boldsymbol{x} \in \mathscr{X} :\ \mathcal{R}(\boldsymbol{x}; \cdot) \text{ is convex and positively 1-homogeneous;} \tag{6.4a}$$

$$\exists\, C > 0\ \forall\, \boldsymbol{x} \in \mathscr{X},\ \boldsymbol{v} = (\alpha, v) \in \mathscr{V} :\quad \mathcal{R}(\boldsymbol{x}; \boldsymbol{v}) \geq C\|v\|_X \tag{6.4b}$$

$$\mathcal{R} \text{ is lower semicontinuous on } \mathscr{X} \times \mathscr{V}. \tag{6.4c}$$

It will be essential here that we allow $\mathcal{R}$ to take the value $+\infty$. In particular, we want to use dissipation potentials $\mathcal{R}$ constructed *via* the augmented vanishing-viscosity contact potential $\mathfrak{P}$, which assumes the value $+\infty$ in an intrinsic manner.

Note that (6.4b) allows that $\mathcal{R}(\boldsymbol{x}; (\alpha, v))$ is independent of the time-component α of $\boldsymbol{v} = (\alpha, v)$. In order to keep track of α we use $\beta \geq 0$ and set

$$\mathcal{R}_\beta(\boldsymbol{x}; \boldsymbol{v}) = \alpha\beta + \mathcal{R}(\boldsymbol{x}; \boldsymbol{v}) \quad \text{for all } \boldsymbol{x} \in \mathscr{X},\ \boldsymbol{v} = (\alpha, v) \in \mathscr{V}.$$

Hereafter, for $\beta = 0$ we shall simply use the notation $\mathcal{R}$ instead of $\mathcal{R}_0$.

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

For any dissipation $\mathcal{R}$ complying with properties (6.4), the corresponding functional $\mathcal{R}_\beta(\boldsymbol{x},\cdot)$ is 1-homogeneous and satisfies the subadditivity property

$$\mathcal{R}_\beta(\boldsymbol{x};\boldsymbol{v}_1+\boldsymbol{v}_2)\le\mathcal{R}_\beta(\boldsymbol{x};\boldsymbol{v}_1)+\mathcal{R}_\beta(\boldsymbol{x};\boldsymbol{v}_2)\quad\text{for all }\boldsymbol{x}\in\mathscr{X}\text{ and }\boldsymbol{v}_1,\,\boldsymbol{v}_2\in\mathscr{V}.$$

We interpret $\mathcal{R}$ as a time-dependent infinitesimal (pseudo-) Finsler metric, and introduce an associated distance in the canonical way by minimizing the length over all admissible curves.

Definition 6.3 (pseudo-Finsler distance induced by $\mathcal{R}$)**.** Given a dissipation function $\mathcal{R}:\mathscr{X}\times\mathscr{V}\to[0,+\infty)$ complying with (6.4), for every $\boldsymbol{x}_i=(t_i,u_i)\in\mathscr{X}$, $i=0,1$, with $0\le t_0\le t_1\le T$, we set

$$\begin{aligned}\Delta_{\mathcal{R}_\beta}(\boldsymbol{x}_0,\boldsymbol{x}_1):=\inf\Bigg\{&\int_{r_0}^{r_1}\mathcal{R}_\beta(\mathsf{x}(r);\dot{\mathsf{x}}(r))\,\mathrm{d}r,\\&\mathsf{x}=(\mathsf{t},\mathsf{u})\in\mathrm{Lip}(r_0,r_1;\mathscr{X}),\ \boldsymbol{x}(r_i)=\boldsymbol{x}_i,\ i=0,1,\ \dot{\mathsf{t}}\ge0\Bigg\}.\end{aligned}\tag{6.5}$$

If $t_0>t_1$ we set $\Delta_{\mathcal{R}_\beta}(\boldsymbol{x}_0,\boldsymbol{x}_1):=+\infty$.

Example 6.4 (dissipations induced by Ψ_0 and $\mathfrak{P}$)**.**

(1) Our first trivial example of a dissipation fulfilling properties (6.4) is given by

$$\mathcal{R}_\beta(\boldsymbol{x},\boldsymbol{v}):=\alpha\beta+\Psi_0(v)\quad\text{for }\boldsymbol{x}\in\mathscr{X},\ \boldsymbol{v}=(\alpha,v)\in\mathscr{V}.\tag{6.6}$$

In this case the associated distance is simply given by

$$\Delta_{\mathcal{R}_\beta}((t_0,u_0),(t_1,u_1))=\beta(t_1-t_0)+\Psi_0(u_1-u_0)\quad\text{for }u_0,u_1\in X,\ 0\le t_0<t_1\le T.$$

(2) Our **main application** will be provided by the dissipation induced by the augmented vanishing-viscosity contact potential $\mathfrak{P}$ and the differential $\mathrm{D}\mathcal{E}$ of the energy, namely

$$\mathcal{P}(\boldsymbol{x};\boldsymbol{v}):=\mathfrak{P}(\alpha,v;0,-\mathrm{D}\mathcal{E}_t(u))\quad\text{and}\quad\mathcal{P}_\beta(\boldsymbol{x};\boldsymbol{v}):=\alpha\beta+\mathcal{P}(\boldsymbol{x};\boldsymbol{v})=\mathfrak{P}(\alpha,v;\beta,-\mathrm{D}\mathcal{E}_t(u))\tag{6.7}$$

for $\boldsymbol{x}=(t,u)\in\mathscr{X}$ and $\boldsymbol{v}=(\alpha,v)\in\mathscr{V}$. By Lemma 5.1, this $\mathcal{P}$ complies with conditions (6.4). The Finsler distance $\Delta_{\mathfrak{p},\mathcal{E}}$ (*cf.* (3.8)) induced by $(\mathfrak{p},\mathcal{E})$ is obtained with $\mathcal{R}=\mathcal{P}$ and $\beta=0$:

$$\Delta_{\mathcal{P}}((t,u_0),(t,u_1))=\Delta_{\mathfrak{p},\mathcal{E}}(t;u_0,u_1)\quad\text{for every }u_0,u_1\in X,\tag{6.8}$$

where we used $t_0=t_1=t$ and $\mathfrak{P}(0,v;0,w)=\mathfrak{p}(v,w)$.

General properties of $\Delta_{\mathcal{R}_\beta}(\cdot,\cdot)$. It is not difficult to check that the infimum in (6.5) is attained and, by the usual rescaling argument (*cf.* Rem. 4.5), one can always choose an optimal Lipschitz curve $\mathsf{x}=(\mathsf{t},\mathsf{u})$ defined in $[0,1]$ such that

$$\mathcal{R}_1(\mathsf{x}(\cdot);\dot{\mathsf{x}}(\cdot))\ \text{ is essentially constant and equal to }\Delta_{\mathcal{R}_1}(\boldsymbol{x}_0,\boldsymbol{x}_1)=(t_1-t_0)+\Delta_{\mathcal{R}}(\boldsymbol{x}_0,\boldsymbol{x}_1).\tag{6.9}$$

Property (6.4b) yields, for every $u_0,u_1\in X$ and $0\le t_0\le t_1\le T$, the estimate

$$\beta(t_1-t_0)+C\|u_1-u_0\|_X\le\Delta_{\mathcal{R}_\beta}((t_0,u_0),(t_1,u_1)).\tag{6.10}$$

Notice that $\Delta_{\mathcal{R}_\beta}(\cdot,\cdot)$ is not symmetric but still satisfies the triangle inequality: for $\boldsymbol{x}_i=(t_i,u_i)\in\mathscr{X}$ with $t_0\le t_1\le t_2$, there holds

$$\Delta_{\mathcal{R}_\beta}(\boldsymbol{x}_0,\boldsymbol{x}_2)\le\Delta_{\mathcal{R}_\beta}(\boldsymbol{x}_0,\boldsymbol{x}_1)+\Delta_{\mathcal{R}_\beta}(\boldsymbol{x}_1,\boldsymbol{x}_2).$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Another useful property, which is a direct consequence of (6.4c), is the lower semicontinuity with respect to convergence in $\mathscr{X}$: if $\boldsymbol{x}_{i,n} = (t_{i,n}, u_{i,n}) \to \boldsymbol{x}_i = (t_i, u_i)$ in $\mathscr{X}$ as $n \uparrow +\infty$, $i = 0, 1$, then

$$\liminf_{n\uparrow+\infty} \Delta_{\mathcal{R}_\beta}(\boldsymbol{x}_{0,n}, \boldsymbol{x}_{1,n}) \geq \Delta_{\mathcal{R}_\beta}(\boldsymbol{x}_0, \boldsymbol{x}_1). \tag{6.11}$$

Indeed, assuming that the lim inf in (6.11) is finite and that, up to the extraction of a suitable subsequence, it is a limit, it is sufficient to choose an optimal sequence $\mathbf{x}_n = (\mathsf{t}_n, \mathsf{u}_n)$ of Lipschitz curves as in (6.9), which therefore satisfies a uniform Lipschitz bound and, up to the extraction of a further subsequence, converges to some Lipschitz curve $\mathbf{x} = (\mathsf{t}, \mathsf{u})$. Then, (6.11) can be proved in the same way as (6.3).

In the case of the distance $\Delta_{\mathcal{P}}$ induced by the vanishing-viscosity contact potential $\mathfrak{P}$, we have a refined lower semicontinuity result:

Lemma 6.5. *Let $u_n, w_n : [t_0, t_1] \to X$ be Borel maps and $0 < \varepsilon_n \to 0$. Suppose that u_n is absolutely continuous for every $n \in \mathbb{N}$ and that the following convergences hold as $n \to \infty$:*

$$u_n(t) \to u(t) \text{ and } w_n(t) \to w(t) \text{ for all } t \in [t_0, t_1]; \qquad \sup_{t\in[t_0,t_1]} \|w_n(t) + \mathrm{D}\mathcal{E}_t(u_n(t))\|_{X^*} \to 0.$$

Then,

$$\liminf_{n\uparrow+\infty} \int_{t_0}^{t_1} \Big(\Psi_{\varepsilon_n}(\dot{u}_n(t)) + \Psi^*_{\varepsilon_n}(w_n(t))\Big)\,\mathrm{d}t \geq \Delta_{\mathcal{P}}((t_0, u(t_0)), (t_1, u(t_1))). \tag{6.12}$$

Proof. Up to extracting a further subsequence, it is not restrictive to assume that the lim inf in (6.12) is a limit. As in (5.1) and (5.2), we set

$$\mathsf{s}_n(t) := t - t_0 + \int_{t_0}^{t} \Big(\Psi_{\varepsilon_n}(\dot{u}_n(r)) + \Psi^*_{\varepsilon_n}(w_n(r))\Big)\,\mathrm{d}r, \quad \mathsf{S}_n := \mathsf{s}_n(t_1),$$
$$\mathsf{t}_n(s) := \mathsf{s}_n^{-1}(s), \quad \mathsf{u}_n(s) := u_n(\mathsf{t}_n(s)), \quad \mathsf{w}_n(s) := w_n(\mathsf{t}_n(s)) \quad \text{for all } s \in [0, \mathsf{S}_n]$$

so that

$$\int_{t_0}^{t_1} \Big(\Psi_{\varepsilon_n}(\dot{u}_n(t)) + \Psi^*_{\varepsilon_n}(w_n(t))\Big)\,\mathrm{d}t = \int_0^{\mathsf{S}_n} \mathfrak{P}_{\varepsilon_n}(\dot{\mathsf{t}}_n(s), \dot{\mathsf{u}}_n(s); 0, \mathsf{w}_n(s))\,\mathrm{d}s. \tag{6.13}$$

Since the sequences (t_n) and (u_n) are uniformly Lipschitz, applying the Ascoli-Arzelà Theorem we can extract a (not relabeled) subsequence such that $\mathsf{S}_n \to \mathsf{S}$, and find functions $\mathsf{t} : [0, \mathsf{S}] \to [t_0, t_1]$, $\mathsf{u} : [0, \mathsf{S}] \to X$, and $\mathsf{w} : [0, \mathsf{S}] \to X^*$, such that

$$\mathsf{t}_n \to \mathsf{t}, \quad \mathsf{u}_n \to \mathsf{u}, \quad \mathsf{w}_n \to \mathsf{w} = -\mathrm{D}\mathcal{E}_{\mathsf{t}}(\mathsf{u}) \quad \text{uniformly in } [0, \mathsf{S}].$$

By construction, we have $\mathsf{t}(0) = t_0$, $\mathsf{u}(0) = u(t_0)$, $\mathsf{t}(\mathsf{S}) = t_1$, and $\mathsf{u}(\mathsf{S}) = u(t_1)$. Applying the lower semicontinuity result of Proposition 6.2, we have

$$\liminf_{n\uparrow+\infty} \int_0^{\mathsf{S}_n} \mathfrak{P}_{\varepsilon_n}(\dot{\mathsf{t}}_n(s), \dot{\mathsf{u}}_n(s); 0, \mathsf{w}_n(s))\,\mathrm{d}s \geq \Delta_{\mathcal{P}}((\mathsf{t}(0), \mathsf{u}(0)), (\mathsf{t}(\mathsf{S}), \mathsf{u}(\mathsf{S}))). \tag{6.14}$$

Combining (6.13) and (6.14), we conclude (6.12). □

The total variation associated with $\Delta_{\mathcal{R}_\beta}$. We recall Definition 3.4 for the total variation $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$ associated with the Finsler distance $\Delta_{\mathfrak{p},\mathcal{E}}$. That definition was an *ad hoc* construction, which was not obtained as a variation for a true distance. We now define $\mathrm{Var}_{\mathcal{R}_\beta}$ as the total variation associated with the distance $\Delta_{\mathcal{R}_\beta}$ and derive, using the reparameterization technique of Section 5, natural lower semicontinuity results and integral expressions. In Section 6.3 we are then able to characterize $\mathrm{Var}_{\mathcal{P}}$ it in terms of $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$, *cf.* Theorem 6.12.

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Definition 6.6 (total variation for the pseudo-Finsler distance $\Delta_{\mathcal{R}_\beta}$)**.** For every curve $\mathsf{x} = (\mathsf{t}, \mathsf{u}) : [0, \mathsf{S}] \to \mathscr{X}$ such that t is *nondecreasing* and every interval $[\mathsf{a}, \mathsf{b}] \subset [0, \mathsf{S}]$ we set

$$\mathrm{Var}_{\mathcal{R}_\beta}(\mathsf{x}; [\mathsf{a}, \mathsf{b}]) := \sup \Bigg\{ \sum_{m=1}^{M} \Delta_{\mathcal{R}_\beta}(\mathsf{x}(s_m), \mathsf{x}(s_{m-1})) : \mathsf{a} = s_0 < s_1 < \cdots < s_{M-1} < s_M = \mathsf{b} \Bigg\}. \tag{6.15}$$

For a *non-parameterized* curve $u : [0, T] \to X$ and $[a, b] \subset [0, T]$, we simply set

$$\mathrm{Var}_{\mathcal{R}_\beta}(u; [a, b]) := \mathrm{Var}_{\mathcal{R}_\beta}(\boldsymbol{u}; [a, b]), \quad \text{with } \boldsymbol{u}(t) := (t, u(t)) \in \mathscr{X}, \quad t \in [0, T]. \tag{6.16}$$

In view of (6.4b), it is immediate to check that a curve u with $\mathrm{Var}_{\mathcal{R}}(u; [0, T]) < +\infty$ belongs to $\mathrm{BV}([0, T]; X)$.

In contrast to the (pseudo)-total variation $\mathrm{Var}_{\mathsf{p},\mathcal{E}}$ defined in (3.11), the new total variation $\mathrm{Var}_{\mathcal{R}_\beta}$ is lower semicontinuous with respect to pointwise convergence (*cf.* Rem. 3.5).

Proposition 6.7 (lower semicontinuity of $\mathrm{Var}_{\mathcal{R}_\beta}(\cdot; [,a])\mathsf{b}$)**.** *If $\mathsf{x}_n = (\mathsf{t}_n, \mathsf{u}_n) : [0, \mathsf{S}] \to \mathscr{X}$ is a sequence of curves pointwise converging to $\mathsf{x} = (\mathsf{t}, \mathsf{u})$ as $n \uparrow \infty$, we have*

$$\liminf_{n\uparrow\infty} \mathrm{Var}_{\mathcal{R}_\beta}(\mathsf{x}_n; [\mathsf{a}, \mathsf{b}]) \geq \mathrm{Var}_{\mathcal{R}_\beta}(\mathsf{x}; [\mathsf{a}, \mathsf{b}]). \tag{6.17}$$

Proof. For an arbitrary subdivision $\mathsf{a} = s_0 < s_1 < \cdots < s_{M-1} < s_M = \mathsf{b}$, inequality (6.11) yields

$$\sum_{m=1}^{M} \Delta_{\mathcal{R}_\beta}(\mathsf{x}(s_m), \mathsf{x}(s_{m-1})) \leq \liminf_{n\uparrow+\infty} \sum_{m=1}^{M} \Delta_{\mathcal{R}_\beta}(\mathsf{x}_n(s_m), \mathsf{x}_n(s_{m-1})) \leq \liminf_{n\uparrow\infty} \mathrm{Var}_{\mathcal{R}_\beta}(\mathsf{x}_n; [\mathsf{a}, \mathsf{b}]).$$

Taking the supremum with respect to all subdivisions of $[\mathsf{a}, \mathsf{b}]$ we obtain (6.17). □

Lipschitz curves. The next result shows that, for Lipschitz curves, the total variation can be calculated by integrating the corresponding dissipation potential.

Proposition 6.8 (the total variation for Lipschitz curves)**.** *Given β, $L > 0$, a bounded curve $\mathsf{x} := (\mathsf{t}, \mathsf{u}) : [0, \mathsf{S}] \to \mathscr{X}$ satisfies the $\Delta_{\mathcal{R}_\beta}$-Lipschitz condition with Lipschitz constant L,* i.e.

$$\Delta_{\mathcal{R}_\beta}(\mathsf{x}(s_1), \mathsf{x}(s_2)) \leq L(s_2 - s_1) \quad \textit{for every } 0 \leq s_1 \leq s_2 \leq \mathsf{S}, \tag{6.18}$$

if and only if it is Lipschitz continuous (with respect to the usual distance in $\mathscr{X}$), t is nondecreasing, and

$$\mathcal{R}_\beta(\mathsf{x}(s); \dot{\mathsf{x}}(s)) \leq L \quad \textit{for a.a. } s \in (0, \mathsf{S}). \tag{6.19}$$

In this case, for every $\gamma \geq 0$

$$\mathrm{Var}_{\mathcal{R}_\gamma}(\mathsf{x}; [\mathsf{a}, \mathsf{b}]) = \gamma(\mathsf{t}(\mathsf{b}) - \mathsf{t}(\mathsf{a})) + \int_{\mathsf{a}}^{\mathsf{b}} \mathcal{R}(\mathsf{x}(s); \dot{\mathsf{x}}(s))\, \mathrm{d}s. \tag{6.20}$$

Proof. The sufficiency of condition (6.19) is clear. Let us now consider a curve x satisfying (6.18). By the coercivity (6.10), x is a Lipschitz curve in the usual sense and [47], Proposition 2.2, yields

$$\Delta_{\mathcal{R}_\beta}(\mathsf{x}(s_0), \mathsf{x}(s_1)) \leq \int_{s_0}^{s_1} \mathsf{m}(s)\, \mathrm{d}s, \quad \text{where} \quad \mathsf{m}(s) := \lim_{h\downarrow 0} \frac{\Delta_{\mathcal{R}_\beta}(\mathsf{x}(s), \mathsf{x}(s+h))}{h} \tag{6.21}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

is the so-called *metric derivative* of x (see [1,4]). The minimality of m and the definition of $\Delta_{\mathcal{R}_\beta}$ ensure that

$$\mathsf{m}(s) \le \mathcal{R}_\beta(\mathsf{x}(s);\dot{\mathsf{x}}(s)) \quad \text{for a.a. } s \in (0,\mathsf{S}). \tag{6.22}$$

On the other hand, since $\mathcal{R}_\beta$ is lower semicontinuous and 1-homogeneous in $\boldsymbol{v}$, for any $s \in [0,\mathsf{S}]$, such that $\mathsf{m}(s)$ and $\dot{\mathsf{x}}(s)$ exist, and for every $0<\sigma<1$ we find a constant $\delta>0$ such that

$$\mathcal{R}_\beta(\boldsymbol{x};\boldsymbol{v}) \ge \sigma\mathcal{R}_\beta(\mathsf{x}(s);\boldsymbol{v}) \quad \text{for every } \boldsymbol{v}\in\mathscr{V} \quad \text{if } |r-s|\le\delta,\ \|\boldsymbol{x}-\mathsf{x}(s)\|_{\mathscr{X}}\le\delta. \tag{6.23}$$

For $h>0$ take an optimal path $\mathsf{y}:[0,h]\to\mathscr{X}$ for the minimum in (6.5) with $\boldsymbol{x}_0=\mathsf{x}(s)$ and $\boldsymbol{x}_1=\mathsf{x}(s+h)$, *i.e.* $\Delta_{\mathcal{R}_\beta}(\mathsf{x}(s),\mathsf{x}(s+h)) = \int_0^h \mathcal{R}_\beta(\mathsf{y}(r);\dot{\mathsf{y}}(r))\,\mathrm{d}r$. Then, (6.10) implies $\|\mathsf{y}(r)-\mathsf{x}(s)\|_{\mathscr{X}}<\delta$ for $r\in[0,h]$, if h is sufficiently small. Employing (6.23) and Jensen's inequality yields

$$\int_0^h \mathcal{R}_\beta(\mathsf{y}(r);\dot{\mathsf{y}}(r))\,\mathrm{d}r \ge \sigma\int_0^h \mathcal{R}_\beta(\mathsf{x}(s);\dot{\mathsf{y}}(r))\,\mathrm{d}r \ge \sigma\mathcal{R}_\beta\left(\mathsf{x}(s);\int_0^h\dot{\mathsf{y}}(r)\,\mathrm{d}r\right) = \sigma\mathcal{R}_\beta(\mathsf{x}(s);\mathsf{x}(s+h)-\mathsf{x}(s))\,,$$

so that

$$\Delta_{\mathcal{R}_\beta}(\mathsf{x}(s),\mathsf{x}(s+h)) \ge \sigma\mathcal{R}_\beta(\mathsf{x}(s);\mathsf{x}(s+h)-\mathsf{x}(s))\,.$$

Dividing by h and passing to the limit, first $h\downarrow 0$ and then $\sigma\uparrow 1$, we obtain the opposite inequality of (6.22). Combining (6.22) (which holds as an equality) with (6.18) and (6.21), we infer (6.19). Since x satisfies the Lipschitz condition (6.18), the total variation map $s\mapsto \mathrm{Var}_{\mathcal{R}_\gamma}(\mathsf{x};[0,s])$ is Lipschitz and it is easy to check that its derivative coincides a.e. in $(0,\mathsf{S})$ with the metric derivative m defined by (6.21). Then (6.20) follows by the identity $\mathsf{m}(s)=\mathcal{R}_\gamma(\mathsf{x}(s);\dot{\mathsf{x}}(s))$. □

Proposition 6.9 (reparameterization)**.** *Let the curve $u:[0,T]\to X$ have total variation $V:=\mathrm{Var}_{\mathcal{R}}(u;[0,T])<+\infty$ (*cf.* (6.16)) and set*

$$\mathsf{s}(t) := t + \mathrm{Var}_{\mathcal{R}}(u;[0,t]) = \mathrm{Var}_{\mathcal{R}_1}(u;[0,t]) \quad \textit{for every } t\in[0,T]. \tag{6.24}$$

Then, there exists a Lipschitz parameterization $\mathsf{x}=(\mathsf{t},\mathsf{u}):[0,\mathsf{S}]\to\mathscr{X}$, with $\mathsf{S}=V+T$, such that

$$\mathcal{R}_1(\mathsf{x}(s);\dot{\mathsf{x}}(s)) = 1 \quad \textit{for a.a. } s\in(0,\mathsf{S}), \tag{6.25}$$

$$\mathsf{t}(\mathsf{s}(t)) = t, \quad \mathsf{u}(\mathsf{s}(t)) = u(t) \quad \textit{for every } t\in[0,T]. \tag{6.26}$$

In particular, for $0\le a<b\le T$ we have

$$b-a+\mathrm{Var}_{\mathcal{R}}(u;[a,b]) = \mathsf{s}(b)-\mathsf{s}(a) = \int_{\mathsf{s}(a)}^{\mathsf{s}(b)} \mathcal{R}_1(\mathsf{x}(s);\dot{\mathsf{x}}(s))\,\mathrm{d}s. \tag{6.27}$$

Proof. The proof is classical, at least when the dissipation $\mathcal{R}$ is continuous and even in its second argument: we briefly sketch the main ideas and refer to [42], Lemma 4.1.

The jump set J_{s} of the function s given by (6.24) coincides with the jump set J_u (*cf.* (2.9)) of u, and s is injective in $\mathrm{C}_u := (0,T)\setminus\mathrm{J}_u$. We denote by t its inverse, defined on $\mathsf{C}_u := \mathsf{s}(\mathrm{C}_u)$ and extended to $\overline{\mathsf{C}}_u$ by its (Lipschitz) continuity. We also set $\mathsf{u}(s) := u(t)$ if $s=\mathsf{s}(t)\in\mathsf{C}_u$. Suppose now that $(\mathsf{s}_-,\mathsf{s}_+)$ is a connected component of $[0,\mathsf{S}]\setminus\overline{\mathsf{C}}_u$, corresponding to some time $\bar t\in[0,T]$ with $\bar{\mathsf{s}}=\mathsf{s}(\bar t)\in[\mathsf{s}_-,\mathsf{s}_+]$ and $\mathsf{s}_\pm=\mathsf{s}(\bar t_\pm)$. We have

$$\mathsf{u}(\mathsf{s}_-) = \lim_{s\uparrow\mathsf{s}_-}\mathsf{u}(s) = u(\bar t_-), \qquad \mathsf{u}(\mathsf{s}_+) = \lim_{s\downarrow\mathsf{s}_+}\mathsf{u}(s) = u(\bar t_+),$$
$$\bar{\mathsf{s}}-\mathsf{s}_- = \Delta_{\mathcal{R}}((\bar t,u(\bar t_-)),(\bar t,u(\bar t))), \quad \mathsf{s}_+-\bar{\mathsf{s}} = \Delta_{\mathcal{R}}((\bar t,u(\bar t)),(\bar t,u(\bar t_+))).$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

By Definition 6.3 for $\Delta_{\mathcal{R}}$, we can join $(\bar t, \mathsf{u}(\mathsf{s}_-))$ to $(\bar t, \mathsf{u}(\mathsf{s}_+))$ by a $\Delta_{\mathcal{R}}$-Lipschitz curve (still denoted by (t,u)) defined in $[\mathsf{s}_-,\mathsf{s}_+]$ with constant first component $\mathsf{t}(s)=\bar t$, and satisfying (6.9) as well as $\mathsf{u}(\bar{\mathsf{s}})=u(\bar t)$.

It is then easy to check that the final curve $\mathsf{x}=(\mathsf{t},\mathsf{u})$, obtained by "filling in" this way all the (at most countable) holes in $[0,\mathsf{S}]\setminus \mathsf{C}_u$, satisfies (6.26) and the Lipschitz condition (6.18) with $L\le 1$. Applying (6.19) and (6.20) we get

$$\int_{\mathsf{s}(a)}^{\mathsf{s}(b)} \mathcal{R}_1(\mathsf{x}(s);\dot{\mathsf{x}}(s))\,\mathrm{d}s \le \mathsf{s}(b)-\mathsf{s}(a) = \mathrm{Var}_{\mathcal{R}_1}(u;[a,b]) \le \mathrm{Var}_{\mathcal{R}_1}(\mathsf{x};[\mathsf{s}(a),\mathsf{s}(b)]) = \int_{\mathsf{s}(a)}^{\mathsf{s}(b)} \mathcal{R}_1(\mathsf{x}(s);\dot{\mathsf{x}}(s))\,\mathrm{d}s,$$

where the first inequality follows from the 1-Lipschitz condition, the subsequent identity from the definition (6.24) of s, and the last one from (6.20). □

The reparameterization of Proposition 6.9 is also useful to express the distributional derivative of u. If $\mathrm{Var}_{\mathcal{R}}(u;[0,T])<+\infty$, we can introduce the distributional derivative $\mu_{\mathcal{R}_1,u}:=\mathsf{s}'$ of s, which is a finite positive measure satisfying

$$\begin{aligned}&\int_0^T \zeta(t)\,\mathrm{d}\mu_{\mathcal{R}_1,u}(t) = -\int_0^T \dot\zeta(t)\mathsf{s}(t)\,\mathrm{d}t \quad \text{for every } \zeta\in \mathrm{C}^1_0(0,T) \text{ and}\\ &\mu_{\mathcal{R}_1,u}([a,b]) = \mathsf{s}(b)-\mathsf{s}(a) \quad \text{for } 0\le a<b\le T \text{ with } a,b\notin \mathrm{J}_u.\end{aligned}$$

Notice that a singleton $\{t\}$ has strictly positive measure if and only if $t\in \mathrm{J}_u$. More precisely

$$\begin{aligned}&\mu_{\mathcal{R}_1,u}(\{t\}) = \Delta_{\mathcal{R}}(t;u(t_-),u(t)) + \Delta_{\mathcal{R}}(t;u(t),u(t_+)) \quad \text{if } t\in \mathrm{J}_u;\\ &\mu_{\mathcal{R}_1,u}(\{t\}) = 0 \quad \text{if } t\in \mathrm{C}_u=(0,T)\setminus \mathrm{J}_u,\end{aligned}$$

with obvious modifications for $t=0,T$. As a general fact we have the representation formula (recall that t is the inverse of s)

$$\mathsf{t}_\#\left(\mathscr{L}^1|_{(0,\mathsf{S})}\right) = \mu_{\mathcal{R}_1,u}, \quad i.e. \quad \int_0^T \zeta(t)\,\mathrm{d}\mu_{\mathcal{R}_1,u}(t) = \int_0^{\mathsf{S}} \zeta(\mathsf{t}(s))\,\mathrm{d}s, \tag{6.28}$$

for every bounded Borel function $\zeta:[0,T]\to\mathbb{R}$. Since t is injective in $\mathsf{C}_u:=\mathsf{t}^{-1}(\mathrm{C}_u)\subset(0,\mathsf{S})$, a Borel subset A of C_u is $\mathscr{L}^1$-negligible if and only if $\mathsf{t}(A)$ has $\mu_{\mathcal{R}_1,u}$-measure 0. Therefore, as the derivatives $\dot{\mathsf{t}},\dot{\mathsf{u}}$ are Borel functions defined up to a $\mathscr{L}^1$-negligible subset of $(0,\mathsf{S})$, the compositions $\dot{\mathsf{t}}\circ\mathsf{s}$, $\dot{\mathsf{u}}\circ\mathsf{s}$ are well-defined in C_u. The next lemma shows that they play an important role.

Proposition 6.10. *The Lebesgue measure $\mathscr{L}^1|_{(0,T)}$ and the vector measure $u'_{\mathrm{co}}=u'_{\mathscr{L}}+u'_{\mathrm{C}}$ are absolutely continuous w.r.t. $\mu_{\mathcal{R}_1,u}$, and we have*

$$\frac{\mathrm{d}\mathscr{L}^1}{\mathrm{d}\mu_{\mathcal{R}_1,u}} = \dot{\mathsf{t}}\circ\mathsf{s} \quad \textit{and} \quad \frac{\mathrm{d}u'_{\mathrm{co}}}{\mathrm{d}\mu_{\mathcal{R}_1,u}} = \dot{\mathsf{u}}\circ\mathsf{s} \qquad \mu_{\mathcal{R}_1,u}\textit{-a.e. in } \mathrm{C}_u. \tag{6.29}$$

Proof. The absolute continuity of both measures is easy, since $\mathscr{L}^1\le\mu_{\mathcal{R}_1,u}$ by (6.28) and the total variation $\|u'\|_X$ is absolutely continuous w.r.t. $\mu_{\mathcal{R}_1,u}$ thanks to (6.10). The first identity of (6.29) can be proved as in [42], Lemma 4.1. Concerning the second one, let us set, for every smooth function $\zeta\in\mathrm{C}^\infty_0(0,T)$,

$$J_u(\zeta) := \sum_{t\in \mathrm{J}_u} \zeta(t)\big(u(t_+)-u(t_-)\big),$$

and let us observe that we have

$$-\int_{\mathsf{C}_u} (\zeta\circ\mathsf{t})'(s)\,\mathsf{u}(s)\,\mathrm{d}s = \int_{\mathsf{C}_u} \zeta(\mathsf{t}(s))\,\dot{\mathsf{u}}(s)\,\mathrm{d}s + J_u(\zeta). \tag{6.30}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Indeed, denoting by $A_t = (\mathsf{a}_t, \mathsf{b}_t) = \mathsf{t}^{-1}(t)$, $t \in \mathrm{J}_u$, the connected components of $[0, \mathsf{S}] \setminus \mathsf{C}_u$, and recalling that $\mathsf{u}(\mathsf{a}_t) = u(t_-), \mathsf{u}(\mathsf{b}_t) = u(t_+)$, we have

$$\begin{aligned} -\int_{\mathsf{C}_u} (\zeta \circ \mathsf{t})'(s)\, \mathsf{u}(s)\, \mathrm{d}s &= -\int_0^{\mathsf{S}} (\zeta \circ \mathsf{t})'(s)\, \mathsf{u}(s)\, \mathrm{d}s + \sum_{t \in \mathrm{J}_u} \int_{\mathsf{a}_t}^{\mathsf{b}_t} (\zeta \circ \mathsf{t})'(s)\, \mathsf{u}(s)\, \mathrm{d}s \\ &= \int_0^{\mathsf{S}} \zeta(\mathsf{t}(s))\, \dot{\mathsf{u}}(s)\, \mathrm{d}s - \sum_{t \in \mathrm{J}_u} \int_{\mathsf{a}_t}^{\mathsf{b}_t} (\zeta \circ \mathsf{t})(s)\, \dot{\mathsf{u}}(s)\, \mathrm{d}s + J_u(\zeta) = \int_{\mathsf{C}_u} \zeta(\mathsf{t}(s))\, \dot{\mathsf{u}}(s)\, \mathrm{d}s + J_u(\zeta). \end{aligned}$$

Therefore, we find

$$\begin{aligned} \int_0^T \zeta(t)\, \mathrm{d}u'(t) &= -\int_0^T \dot\zeta(t)\, u(t)\, \mathrm{d}t = -\int_0^{\mathsf{S}} \dot\zeta(\dot{\mathsf{t}}(s))\, u(\mathsf{t}(s))\, \dot{\mathsf{t}}(s)\, \mathrm{d}s = -\int_{\mathsf{C}_u} \dot\zeta(\mathsf{t}(s))\, u(\mathsf{t}(s))\, \dot{\mathsf{t}}(s)\, \mathrm{d}s \\ &= -\int_{\mathsf{C}_u} (\zeta \circ \mathsf{t})'(s)\, \mathsf{u}(s)\, \mathrm{d}s = \int_{\mathsf{C}_u} \zeta(\mathsf{t}(s))\, \dot{\mathsf{u}}(s)\, \mathrm{d}s + J_u(\zeta) = \int_{\mathrm{C}_u} \zeta(t)\, \dot{\mathsf{u}}(\mathsf{s}(t))\, \mathrm{d}\mu_{\mathcal{R}_1,u}(t) + J_u(\zeta), \end{aligned}$$

where the fifth identity ensues from (6.30) and the last one from (6.28). Since

$$\int_0^T \zeta(t)\, \mathrm{d}u'_{\mathrm{co}}(t) = \int_0^T \zeta(t)\, \mathrm{d}u'(t) - J_u(\zeta)$$

we conclude the second of (6.29). □

The next result establishes a representation formula for $\mathrm{Var}_{\mathcal{R}}$ that is analogous to the *ad hoc* definition for $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$ in (3.11).

Corollary 6.11 (integral expression for $\mathrm{Var}_{\mathcal{R}}$)**.** *Let $u : [0,T] \to X$ fulfill $\mathrm{Var}_{\mathcal{R}}(u;[0,T]) < +\infty$, let μ be a positive finite measure such that $\mathscr{L}^1 \ll \mu$ and $u'_{\mathrm{co}} \ll \mu$, and let us set*

$$\begin{aligned} \mathrm{Jmp}_{\mathcal{R}}(u;[a,b]) := {} & \Delta_{\mathcal{R}}(a; u(a), u(a_+)) + \Delta_{\mathcal{R}}(b; u(b_-), u(b)) \\ & + \sum_{t \in \mathrm{J}_u \cap (a,b)} \Big(\Delta_{\mathcal{R}}(t; u(t_-), u(t)) + \Delta_{\mathcal{R}}(t; u(t), u(t_+)) \Big). \end{aligned} \tag{6.31}$$

Then,

$$\mathrm{Var}_{\mathcal{R}}(u;[a,b]) = \int_a^b \mathcal{R}\left(\big(t, u(t)\big); \left(\frac{\mathrm{d}\mathscr{L}^1}{\mathrm{d}\mu}(t), \frac{\mathrm{d}u'_{\mathrm{co}}}{\mathrm{d}\mu}(t) \right) \right) \mathrm{d}\mu(t) + \mathrm{Jmp}_{\mathcal{R}}(u;[a,b]). \tag{6.32}$$

Proof. Since the expression on the right-hand side is independent of the measure μ, it is not restrictive to choose $\mu = \mu_{\mathcal{R}_1,u}$. By (6.27) we have

$$\begin{aligned} b - a + \mathrm{Var}_{\mathcal{R}}(u;[a,b]) &= \int_{(\mathsf{s}(a),\mathsf{s}(b)) \cap \mathsf{C}_u} \mathcal{R}_1(\mathbf{x}(s); \dot{\mathbf{x}}(s))\ \mathrm{d}s + \mathscr{L}^1((\mathsf{s}(a), \mathsf{s}(b)) \setminus \mathsf{C}_u) \\ &= \int_{(a,b) \cap \mathrm{C}_u} \mathcal{R}_1(\mathbf{x}(\mathsf{s}(t)); \dot{\mathbf{x}}(\mathsf{s}(t)))\ \mathrm{d}\mu + \mu([a,b] \cap \mathrm{J}_u) \\ &= \int_{(a,b) \cap \mathrm{C}_u} \mathcal{R}_1\Big(\big(t, u(t)\big); \big(\dot{\mathsf{t}}(\mathsf{s}(t)), \dot{\mathsf{u}}(\mathsf{s}(t))\big) \Big)\ \mathrm{d}\mu + \mathrm{Jmp}_{\mathcal{R}}(u;[a,b]), \end{aligned}$$

and we conclude by (6.29). □

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

6.3. **Total variation for BV solutions**

We now focus on the particular case (6.7) of Example 6.4, when the dissipation $\mathcal{R} = \mathcal{P}$ is associated with the vanishing-viscosity contact potential $\mathfrak{P}$. We use the symbol $\mathcal{P}$, which is defined *via* $\mathcal{P}((t,u);(\alpha,v)) = \mathfrak{P}(\alpha,v;0,-\mathrm{D}\mathcal{E}_t(u))$.

Theorem 6.12 (comparison between $\mathrm{Var}_{\mathcal{P}}$ and $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$)**.** *For every curve $u \in \mathrm{BV}([0,T];X)$ and every interval $[a,b] \subset [0,T]$ we have*

$$\mathrm{Var}_{\mathfrak{p},\mathcal{E}}(u;[a,b]) \le \mathrm{Var}_{\mathcal{P}}(u;[a,b]), \tag{6.33}$$

and equality holds in (6.33) if and only if u satisfies the local stability condition ($\mathrm{S}_{\mathrm{loc}}$) on (a,b), see Definition 2.5. Furthermore, if $\mathrm{Var}_{\mathcal{P}}(u;[a,b]) < +\infty$, then u satisfies ($\mathrm{S}_{\mathrm{loc}}$) on (a,b).

Proof. We compare the integral representation (6.32) for $\mathrm{Var}_{\mathcal{P}}$ with $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$ defined in (3.11). Obviously, the jump contributions in $\mathrm{Var}_{\mathfrak{p},\mathcal{E}}$ and $\mathrm{Var}_{\mathcal{P}}$ are the same by (6.8). We also observe that $\mathscr{L}^1 \ll \mu$, $\frac{\mathrm{d}\mathscr{L}^1}{\mathrm{d}\mu}(t) > 0$ for $\mathscr{L}^1$-a.a $t \in (0,T)$. Inequality (6.33) then follows by applying (6.32), noticing that for μ-a.a. $t \in [0,T]$ we have

$$\mathcal{P}\left((t,u(t));\left(\frac{\mathrm{d}\mathscr{L}^1}{\mathrm{d}\mu}(t),\frac{\mathrm{d}u'_{\mathrm{co}}}{\mathrm{d}\mu}(t)\right)\right) = \mathfrak{P}\left(\frac{\mathrm{d}\mathscr{L}^1}{\mathrm{d}\mu}(t),\frac{\mathrm{d}u'_{\mathrm{co}}}{\mathrm{d}\mu}(t);0,w(t)\right) \ge \Psi_0\left(\frac{\mathrm{d}u'_{\mathrm{co}}}{\mathrm{d}\mu}(t)\right), \tag{6.34}$$

where $w(t) = -\mathrm{D}\mathcal{E}_t(u(t))$. The latter inequality is due to (5.9) and (3.3). On the other hand, using $\frac{\mathrm{d}\mathscr{L}^1}{\mathrm{d}\mu}(t) > 0$ and the definition of $\mathfrak{P}$ in (5.9) we see that (6.34) is an equality if and only if $w(t) \in K^*$ for μ-a.a $t \in (0,T)$, *i.e.* if the local stability property ($\mathrm{S}_{\mathrm{loc}}$) holds. Similarly, we conclude the last part of the assertion, since $\mathfrak{P}(\alpha,v;p,w) < +\infty$ for $\alpha > 0$ implies $w \in K^*$. □

The following corollary of Theorem 6.12 provides a very short, but rather involved characterization of BV solutions *via* the Finsler metric $\mathcal{P}$.

Corollary 6.13 (characterization of BV solutions)**.** *A curve $u : [0,T] \to X$ is a BV solution if and only if it satisfies one of the following (equivalent) two conditions:*

$$\mathrm{Var}_{\mathcal{P}}(u;[t_0,t_1]) + \mathcal{E}_{t_1}(u(t_1)) = \mathcal{E}_{t_0}(u(t_0)) + \int_{t_0}^{t_1} \partial_t\mathcal{E}_t(u(t))\,\mathrm{d}t \quad \textit{for every } 0 \le t_0 \le t_1 \le T, \tag{6.35}$$

$$\mathrm{Var}_{\mathcal{P}}(u;[0,T]) + \mathcal{E}_T(u(T)) \le \mathcal{E}_0(u(0)) + \int_0^T \partial_t\mathcal{E}_t(u(t))\,\mathrm{d}t. \tag{6.36}$$

We now arrive at the main lower semicontinuity result of this section.

Proposition 6.14. *Suppose that $(u_\varepsilon)_{\varepsilon>0} \subset \mathrm{AC}([0,T];X)$ is a family of curves, which converges pointwise to u for $\varepsilon \downarrow 0$, and that the functions $w_\varepsilon : [0,T] \to X^*$ satisfy $\|w_\varepsilon(t) + \mathrm{D}\mathcal{E}_t(u_\varepsilon(t))\|_{X^*} \to 0$ uniformly in $[0,T]$. Then,*

$$\liminf_{\varepsilon\downarrow 0} \int_0^T \Big(\Psi_\varepsilon(\dot u_\varepsilon) + \Psi_\varepsilon^*(w_\varepsilon(t))\Big)\,\mathrm{d}t \ge \mathrm{Var}_{\mathcal{P}}(u;[0,T]) \ge \mathrm{Var}_{\mathfrak{p},\mathcal{E}}(u;[0,T]). \tag{6.37}$$

Proof. Choosing a finite partition $0 = t_0 < t_1 < t_2 < \cdots < t_N = T$ of the time interval $[0,T]$, Lemma 6.5 yields

$$\liminf_{\varepsilon\downarrow 0} \int_0^T \Big(\Psi_\varepsilon(\dot u_\varepsilon) + \Psi_\varepsilon^*(w_\varepsilon(t))\Big)\,\mathrm{d}t \ge \sum_{j=1}^N \Delta_{\mathcal{P}}((t_{j-1},u(t_{j-1})),(t_j,u(t_j))).$$

Taking the supremum of the right-hand side with respect to all partitions of $[0,T]$, we obtain the first estimate in (6.37). The second estimate was shown in Theorem 6.12. □

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

We conclude this section with the proof of Theorem 5.7, which states the equivalence between BV solutions and parameterized solutions.

Proof. Let (t,u) be a parameterized solution as in the statement of the theorem. It is easy to check directly from definitions (6.5) for $\Delta_{\mathcal{P}}$ and (6.15) for $\mathrm{Var}_{\mathcal{P}}$ that

$$\begin{aligned}\mathrm{Var}_{\mathcal{P}}(u;[0,T]) &\le \int_0^{\mathsf{S}} \mathfrak{P}(\dot{\mathsf{t}}(s),\dot{\mathsf{u}}(s);0,-\mathrm{D}\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s)))\,\mathrm{d}s\\ &\le \mathcal{E}_0(\mathsf{u}(0)) - \mathcal{E}_{\mathsf{t}(\mathsf{S})}(\mathsf{u}(\mathsf{S})) + \int_0^{\mathsf{S}} \partial_t\mathcal{E}_{\mathsf{t}(s)}(\mathsf{u}(s))\dot{\mathsf{t}}(s)\,\mathrm{d}s = \mathcal{E}_0(u(0)) - \mathcal{E}_T(u(T)) + \int_0^T \partial_t\mathcal{E}_t(u(t))\,\mathrm{d}t,\end{aligned}$$

where the second inequality ensues from (5.22). Thus, (6.36) holds, so that u is a BV solution by Corollary 6.13. The converse implication follows from Proposition 6.9. □

A. More properties of the vanishing-viscosity contact potential

A direct consequence of the representation $\mathfrak{p}(v,w) = \mathfrak{q}(v,\Psi^*(w))$ in (3.15) and the monotonicity of $\mathfrak{q}$ in its second argument is the following property:

($\mathfrak{p}5$) If for some $v_0 \in X$ and $\bar w, w \in X^*$ we have $\mathfrak{p}(v_0,\bar w) < \mathfrak{p}(v_0,w)$, then the inequality $\mathfrak{p}(v,\bar w) \le \mathfrak{p}(v,w)$ holds for every $v \in X$, and there exists $v_1 \in X$ such that $\mathfrak{p}(v_1,\bar w) < \langle w, v_1\rangle$.

Remark A.15 (a dual family of convex sets)**.** Property ($\mathfrak{p}5$) has a dual geometric counterpart: let us first observe that for every $w \in X^*$ the map $v \mapsto \mathfrak{p}(v,w)$ is a gauge function, and therefore it is the support function of the convex set

$$K_w^* := \Big\{ z \in X^* : \langle z, v\rangle \le \mathfrak{p}(v,w) \text{ for every } v \in X\Big\}, \quad i.e.\ \mathfrak{p}(v,w) = \sup\Big\{\langle z,v\rangle \,:\, z \in K_w^*\Big\}.$$

Assertion ($\mathfrak{p}5$) then says that for every couple $w, \bar w \in X$ we always have

$$\bar w \in K_w^* \text{ or } w \in K_{\bar w}^*, \quad \text{and} \quad \bar w \in K_w^* \quad \Leftrightarrow \quad \mathfrak{p}(\cdot,\bar w) \le \mathfrak{p}(\cdot,w). \tag{A.1}$$

In fact, suppose that $w \notin K_{\bar w}^*$: this means that an element $v_0 \in X$ exists such that $\langle w, v_0\rangle > \mathfrak{p}(v_0,\bar w)$. By ($\mathfrak{p}2$) in Theorem 3.7 we get $\mathfrak{p}(v_0,w) > \mathfrak{p}(v_0,\bar w)$, and therefore by ($\mathfrak{p}5$) we have $\mathfrak{p}(v,w) \ge \mathfrak{p}(v,\bar w) \ge \langle \bar w, v\rangle$ for every $v \in X$, so that $\bar w \in K_w^*$. The second statement of (A.1) is an immediate consequence of ($\mathfrak{p}2$) and of the second part of ($\mathfrak{p}5$).

Here are some further useful properties of $\mathfrak{p}$.

Lemma A.16. *If $\mathfrak{p} : X \times X^* \to [0,+\infty)$ satisfies ($\mathfrak{p}1$)–($\mathfrak{p}3$), then:*

($\mathfrak{p}6$) *for every $v \in X, w \in X^*$ we have*

$$\mathfrak{p}(v,0) + I_{K_0^*}(w) \ge \mathfrak{p}(v,w) \ge \mathfrak{p}(v,0). \tag{A.2}$$

($\mathfrak{p}7$) *The contact set can be characterized by*

$$(v,w) \in \Sigma_{\mathfrak{p}} \quad \Leftrightarrow \quad w \in \partial\mathfrak{p}(\cdot,w)(v) \quad \Leftrightarrow \quad v \in \partial I_{K_w^*}(w). \tag{A.3}$$

More generally, if $\bar w \in \partial\mathfrak{p}(\cdot,w)(v)$ then $(v,\bar w) \in \Sigma_{\mathfrak{p}}$, $\bar w \in K_w^*$, and $\mathfrak{p}(v,w) = \mathfrak{p}(v,\bar w)$. In particular, if $\bar w \in \partial K_w^*$ then $w \in K_{\bar w}^*$.

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Proof. The chain of inequalities in (A.2) is an immediate consequence of (A.1). Assertion (A.3) is a direct consequence of the fact that $v \mapsto \mathfrak{p}(v,w)$ is a gauge function and $I_{K_w^*}$ is its Legendre transform.

In order to check the last statement, given $v \in X, w \in X^*$ let us take $\bar w \in \partial \mathfrak{p}(\cdot,w)(v)$, so that $\bar w \in K_w^*$ and $\mathfrak{p}(v,w) = \langle \bar w, v\rangle$. Combining ($\mathfrak{p}2$) with (A.1) we get $\mathfrak{p}(v,w) = \mathfrak{p}(v,\bar w)$, so that $(v,\bar w) \in \Sigma_{\mathfrak{p}}$. □

Theorem A.17 (properties of $\mathfrak{p}, \Psi$ and $\Sigma_{\mathfrak{p}}$)**.**

(P1) *The vanishing-viscosity contact potential* $\mathfrak{p}$ *satisfies* $\mathfrak{p}(v,0) = \Psi_0(v)$, $K_0^* = K^*$, *and in particular*

$$\forall\,(v,w) \in X \times X^* : \quad \mathfrak{p}(v,w) \geq \langle w, v\rangle, \quad \Psi_0(v) + \mathrm{I}_{K^*}(w) \geq \mathfrak{p}(v,w) \geq \Psi_0(v) \geq 0, \tag{A.4}$$

$$\forall\, v \neq 0 : \quad \mathfrak{p}(v,w) = \Psi_0(v) \quad \Leftrightarrow \quad w \in K^* \quad \Leftrightarrow \quad \Psi^*(w) = 0. \tag{A.5}$$

(P2) *For every* $w \in X^*$, *the convex sets* K_w^* *are the sublevels of* Ψ^*

$$K_w^* = \Big\{ z \in X^* : \Psi^*(z) \leq \Psi^*(w) \Big\}, \tag{A.6}$$

and $\mathfrak{p}$ *admits the dual representation*

$$\mathfrak{p}(v,w) = \sup \Big\{ \langle z, v\rangle \ : \ z \in X^*,\ \Psi^*(z) \leq \Psi^*(w) \Big\}. \tag{A.7}$$

In particular, $\Psi^*(w_1) \leq \Psi^*(w_2)$ *for some* $w_1, w_2 \in X^*$ *if and only if* $\mathfrak{p}(v,w_1) \leq \mathfrak{p}(v,w_2)$ *for every* $v \in X$.

(P3) *The multivalued function* Λ *defined in* (3.17) *is upper semicontinuous,* i.e.

$$\text{if } (v_n, w_n) \to (v,w) \in X \times X^* \text{ and } \varepsilon_n \in \Lambda(v_n, w_n) \to \varepsilon, \quad \text{then} \quad \varepsilon \in \Lambda(v,w). \tag{A.8}$$

- *A pair* (v,w) *with* $v \neq 0$ *lies in the contact set* $\Sigma_{\mathfrak{p}}$ (3.14) *if and only if*

$$w \in \partial \Psi_0(v) \subset K^* \quad \text{or,} \quad \text{if } \Big(w \notin K^* \ \text{ and } \ \exists\, \varepsilon > 0 : w \in \partial \Psi(\varepsilon v) \Big). \tag{A.9}$$

The last inclusion holds exactly for $\varepsilon \in \Lambda(v,w)$. *Equivalently,*

$$(v,w) \in \Sigma_{\mathfrak{p}} \quad \Leftrightarrow \quad w \in \partial \Psi_\varepsilon(v) \quad \text{for every } \varepsilon \in \Lambda(v,w).$$

In particular, in the case of an additive viscosity $\Psi(v) = \Psi_0(v) + \Psi_V(v)$ *with* Ψ_V *satisfying* (2.22), *we simply have*

$$(v,w) \in \Sigma_{\mathfrak{p}} \quad \Longleftrightarrow \quad \exists\, \lambda \geq 0 : \quad w \in \partial \Psi_0(v) + \partial \Psi_V(\lambda v). \tag{A.10}$$

Proof. **Ad (P1).** Inequalities (A.4) are immediate consequences of the definition of $\mathfrak{p}$. The equality $\Psi_0(v) = \mathfrak{p}(v,w)$ is equivalent to the existence of a sequence $\varepsilon_k > 0$ such that (recall that $\Psi_\varepsilon(v) = \varepsilon^{-1}\Psi(\varepsilon v) \geq \Psi_0(v)$)

$$\lim_{k\to\infty} \varepsilon_k^{-1} \Psi(\varepsilon_k v) = \Psi_0(v), \quad \lim_{k\to\infty} \varepsilon_k^{-1} \Psi^*(w) = 0.$$

Using $v \neq 0$ the first inequality prevents ε_k from diverging to $+\infty$ (since Ψ is superlinear), and from the second limit we get $\Psi^*(w) = 0$, *i.e.*

$$\langle w, z\rangle \leq \Psi(z) \quad \forall\, z \in X.$$

Replacing z with εz, multiplying the previous inequality by ε^{-1}, and passing to the limit as $\varepsilon \downarrow 0$, in view of (Ψ.3) we conclude

$$\langle w, z\rangle \leq \Psi_0(z) \quad \forall\, z \in X, \quad \text{so that} \quad w \in K^*.$$

The converse implication in (A.5) is immediate.

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Ad (P2). Since the sublevels of Ψ^* are closed and convex, a duality argument shows that (A.6) is equivalent to (A.7). In order to prove the latter formula, let us observe that, if $\Psi^*(z) \le \Psi^*(w)$, then $\langle z, v\rangle \le \mathfrak{p}(v,w)$, because the Fenchel inequality yields

$$\langle z, v\rangle = \varepsilon^{-1}\langle z, \varepsilon v\rangle \le \varepsilon^{-1}\Psi(\varepsilon v) + \varepsilon^{-1}\Psi(z) = \Psi_\varepsilon(v) + \Psi^*_\varepsilon(z) \le \Psi_\varepsilon(v) + \Psi^*_\varepsilon(w) \quad \text{for every } \varepsilon > 0.$$

We show that there exists $z \in X^*$ such that $\Psi^*(z) \le \Psi^*(w)$ and $\mathfrak{p}(v,w) = \langle z, v\rangle$. Due to (A.5), if $w \in K^*$, then $\mathfrak{p}(v,w) = \Psi_0(v)$ and the thesis follows from (2.5). Hence, let us suppose that $w \notin K^*$ and $v \neq 0$. Then we can choose $\varepsilon_0 \in \Lambda(v,w)$, $\varepsilon_0 > 0$, such that

$$\mathfrak{p}(v,w) = \varepsilon_0^{-1}\Psi(\varepsilon_0 v) + \varepsilon_0^{-1}\Psi^*(w) \le \varepsilon^{-1}\Psi(\varepsilon v) + \varepsilon^{-1}\Psi^*(w) \quad \text{for every } \varepsilon > 0. \tag{A.11}$$

Choosing $z_\varepsilon \in \partial\Psi(\varepsilon v)$ we have

$$\Psi(\varepsilon v) - \Psi(\varepsilon_0 v) \le \langle z_\varepsilon, (\varepsilon - \varepsilon_0)v\rangle \quad \text{for every } \varepsilon > 0$$

so that, in view of inequality (A.11), there holds

$$\big(\varepsilon^{-1} - \varepsilon_0^{-1}\big)\Big(\Psi(\varepsilon_0 v) + \Psi^*(w)\Big) + \varepsilon^{-1}\langle z_\varepsilon, (\varepsilon - \varepsilon_0)v\rangle \ge 0 \quad \text{for every } \varepsilon > 0.$$

Dividing by $\varepsilon - \varepsilon_0$ and passing to the limit first as $\varepsilon \downarrow \varepsilon_0$ and then as $\varepsilon \uparrow \varepsilon_0$, we thus find $z_\pm \in \partial\Psi(\varepsilon_0 v)$ (accumulation points of the sequences $(z_\varepsilon : \varepsilon > \varepsilon_0)$ and $(z_\varepsilon : \varepsilon < \varepsilon_0)$, respectively), such that

$$\langle z_-, v\rangle \le \mathfrak{p}(v,w) = \varepsilon_0^{-1}\Big(\Psi(\varepsilon_0 v) + \Psi^*(w)\Big) \le \langle z_+, v\rangle. \tag{A.12}$$

On the other hand, the Fenchel identity of convex analysis yields

$$\varepsilon_0^{-1}\Psi^*(z) = \langle z, v\rangle - \varepsilon_0^{-1}\Psi(\varepsilon_0 v) \quad \text{for every } z \in \partial\Psi(\varepsilon_0 v), \tag{A.13}$$

so that the map $z \mapsto \Psi^*(z)$ is affine on $\partial\Psi(\varepsilon_0 v)$, and a comparison between (A.12) and (A.13) yields

$$\Psi^*(z_-) \le \Psi^*(w) \le \Psi^*(z_+).$$

Using formula (A.13) we can thus find $\theta \in [0,1]$ and $z_\theta := (1-\theta)z_- + \theta z_+ \in \partial\Psi(\varepsilon_0 v)$ such that

$$\Psi^*(z_\theta) = \Psi^*(w), \quad \langle z_\theta, v\rangle = \mathfrak{p}(v,w) = \varepsilon_0^{-1}\Big(\Psi(\varepsilon_0 v) + \Psi^*(w)\Big).$$

The last statement of (P2) follows easily. One implication is immediate. On the other hand, if $\Psi^*(w_1) > \Psi^*(w_2)$, then by the Hahn-Banach separation theorem we can find $\bar v \in X$ and $\delta > 0$ such that

$$\langle w_1, \bar v\rangle \ge \delta + \langle z, \bar v\rangle \quad \text{for every } z \in X^* \text{ such that } \Psi^*(z) \le \Psi^*(w_2),$$

and, therefore, by (A.7) we conclude $\mathfrak{p}(\bar v, w_1) \ge \langle w_1, \bar v\rangle \ge \delta + \mathfrak{p}(\bar v, w_2)$.

Ad (P3) and continuity of $\mathfrak{p}$. Notice that $\mathfrak{p}$ is upper semicontinuous, being defined as the infimum of a family of continuous functions. Take now converging sequences (v_n), (w_n), (ε_n) as in (A.8): we have that

$$\liminf_{n\to\infty}\big(\varepsilon_n^{-1}\Psi(\varepsilon_n v_n) + \varepsilon_n^{-1}\Psi^*(w_n)\big) \ge \Psi_\varepsilon(v) + \Psi^*_\varepsilon(w) = \begin{cases} \varepsilon^{-1}\Psi(\varepsilon v) + \varepsilon^{-1}\Psi^*(w) & \text{if } \varepsilon > 0, \\ \Psi_0(v) + \mathrm{I}^*_K(w) & \text{if } \varepsilon = 0. \end{cases}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

Since

$$\begin{aligned}\mathfrak{p}(v,w) \geq \liminf_{n\to\infty} \mathfrak{p}(v_n,w_n) &\geq \liminf_{n\to\infty} \big(\varepsilon_n^{-1}\Psi(\varepsilon_n v_n) + \varepsilon_n^{-1}\Psi^*(w_n)\big)\\ &\geq \Psi_\varepsilon(v) + \Psi_\varepsilon^*(w) \geq \mathfrak{p}(v,w),\end{aligned} \tag{A.14}$$

we obtain $\varepsilon \in \Lambda(v,w)$.

Inequality (A.14) shows that $\mathfrak{p}$ is also lower semicontinuous, since, if $v \neq 0$, any sequence $\varepsilon_n \in \Lambda(v_n,w_n)$ admits a converging subsequence, in view of (3.18a).

Ad (P4). Concerning the characterization (A.9) of $\Sigma_\mathfrak{p}$, it is easy to check that, if (v,w) satisfies (A.9), then by the Fenchel identity and formula (2.2) we have, for $w \in K^*$, $\mathfrak{p}(v,w) \geq \langle w,v\rangle = \Psi_0(v) = \mathfrak{p}(v,w)$, and, for $w \notin K^*$,

$$\mathfrak{p}(v,w) \geq \langle w,v\rangle = \varepsilon^{-1}\langle w,\varepsilon v\rangle = \varepsilon^{-1}\Psi(\varepsilon v) + \varepsilon^{-1}\Psi^*(w) \geq \mathfrak{p}(v,w)$$

so that $(v,w) \in \Sigma_\mathfrak{p}$ and $\varepsilon \in \Lambda(v,w)$. Conversely, if $\mathfrak{p}(v,w) = \langle w,v\rangle$ and $w \in K^*$, then by (A.4) $\Psi_0(v) = \langle w,v\rangle$ and therefore $w \in \partial\Psi_0(v)$. If $w \notin K^*$, then, choosing $\varepsilon \in \Lambda(v,w)$ (which is nonempty because of $v \neq 0$), we have

$$\Psi(\varepsilon v) + \Psi^*(w) = \varepsilon\mathfrak{p}(v,w) = \langle w,\varepsilon v\rangle, \quad \text{so that} \quad w \in \partial\Psi(\varepsilon v).$$

In the additive viscosity case of (2.23), (A.10) follows now from (A.9) by the sum rule of the subdifferentials and the 0-homogeneity of $\partial\Psi_0$. □

B. Detailed description of optimal transitions

The next result provides a careful description of $(\mathfrak{p},\mathcal{E}_t)$-optimal transitions, *cf.* Definition 4.4.

Theorem B.18. *Let $t \in [0,T]$, u_-, $u_+ \in X$, and $\vartheta : [0,1] \to X$ be an optimal transition in $\Theta(t;u_-,u_+)$. Then,*

(T1) *ϑ is a constant-speed minimal geodesic for the Finsler distance $\Delta_{\mathfrak{p},\mathcal{E}}(t;u_-,u_+)$, and for every $0 \leq \rho_0 < \rho_1 \leq 1$ it satisfies*

$$\begin{aligned}\mathcal{E}_t(\vartheta(\rho_0)) - \mathcal{E}_t(\vartheta(\rho_1)) &= \Delta_{\mathfrak{p},\mathcal{E}}(t;\vartheta(\rho_0),\vartheta(\rho_1))\\ &= (\rho_1-\rho_0)\Delta_{\mathfrak{p},\mathcal{E}}(t;u_-,u_+) = (\rho_1-\rho_0)\big(\mathcal{E}_t(u_-) - \mathcal{E}_t(u_+)\big).\end{aligned} \tag{B.1}$$

In particular, the map $\rho \mapsto \mathcal{E}_t(\vartheta(\rho))$ is affine.

(T2) *An optimal transition ϑ is of* sliding *type* ($\mathrm{O}_{\mathrm{sliding}}$) *if and only if it satisfies*

$$\partial\Psi_0(\dot\vartheta(r)) + \mathrm{D}\mathcal{E}_t(\vartheta(r)) \ni 0 \quad \text{for a.a. } r \in (0,1), \tag{B.2}$$

$$\Psi_{0*}(-\mathrm{D}\mathcal{E}_t(\vartheta(r))) = 1 \quad \text{for every } r \in [0,1]. \tag{B.3}$$

(T3) *An optimal transition ϑ is of* viscous *type* ($\mathrm{O}_{\mathrm{viscous}}$) *if and only if for every selection $(0,1) \ni r \mapsto \varepsilon(r) \in \Lambda(\dot\vartheta(r), -\mathrm{D}\mathcal{E}_t(\vartheta(r)))$ we have*

$$\partial\Psi(\varepsilon(r)\dot\vartheta(r)) + \mathrm{D}\mathcal{E}_t(\vartheta(r)) \ni 0 \quad \text{for a.a. } r \in (0,1). \tag{B.4}$$

Equivalently, there exists an absolutely continuous, surjective time rescaling $\mathsf{r} : (\rho_0,\rho_1) \to (0,1)$, with $-\infty \leq \rho_0 < \rho_1 \leq \infty$ and $\dot{\mathsf{r}}(s) > 0$ for $\mathscr{L}^1$ a.e. $s \in (\rho_0,\rho_1)$, such that the rescaled transition $\theta(s) := \vartheta(\mathsf{r}(s))$ satisfies the viscous differential inclusion

$$\partial\Psi(\dot\theta(s)) + \mathrm{D}\mathcal{E}_t(\theta(s)) \ni 0 \text{ for a.a. } s \in (\rho_0,\rho_1), \quad \text{with } \lim_{s\downarrow\rho_0}\theta(s) = u_-, \quad \lim_{s\uparrow\rho_1}\theta(s) = u_+. \tag{B.5}$$

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

(T4) *Any optimal transition ϑ can be decomposed in a canonical way into an (at most) countable collection of optimal* sliding and viscous *transitions. In other words, there exists (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 $(0,1)\subset\big(\cup_{j\in\sigma}S_j\big)\cup\big(\cup_{k\in\upsilon}\overline{V}_k\big)$ and*

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

(T5) *An optimal transition ϑ is of* energetic *type* ($\mathrm{O}_{\mathrm{ener}}$) *if and only if ϑ is of sliding type and it is a Ψ_0-minimal geodesic,* i.e.

$$\Psi_0(\vartheta(r_1)-\vartheta(r_0))=(r_1-r_0)\Psi_0(u_1-u_0) \quad \textit{for every } 0\le r_0<r_1\le 1. \tag{B.6}$$

If Ψ_0 has strictly convex sublevels, then ϑ is linear and $r\mapsto(\vartheta(r),\mathcal{E}_t(\vartheta(r)))$ is a linear segment contained in the graph of $\mathcal{E}_t$.
If Ψ_0 is Gâteaux-differentiable at $X\setminus\{0\}$ then

$$-\mathrm{D}\mathcal{E}_t(\vartheta(r))=\mathrm{D}\Psi_0(u_+-u_-) \quad \textit{for every } r\in[0,1].$$

In particular, the map $r\mapsto-\mathrm{D}\mathcal{E}_t(\vartheta(r))$ is constant.

Proof. **Ad (T1).** The geodesic property follows from the minimality of ϑ (*cf.* (O.2) in Def. 4.4). Then, there holds

$$\frac{\mathrm{d}}{\mathrm{d}r}\mathcal{E}_t(\vartheta(r))=-\mathfrak{p}(\dot\vartheta(r),-\mathrm{D}\mathcal{E}_t(\vartheta(r)))\equiv\mathcal{E}_t(u_+)-\mathcal{E}_t(u_-) \quad \text{for a.a. } r\in(0,1), \tag{B.7}$$

where the first identity ensues from the chain rule (2.26) for $\mathcal{E}$ and the contact condition (4.3), and the second one from (4.2). Clearly, (B.7) implies (B.1).

Ad (T2). If ϑ is of sliding type, then the contact condition (4.3), with (A.9), yields (B.2). (B.3) follows since $\dot\vartheta\neq0$ a.e. in $(0,1)$.

Ad (T3). Equation (B.4) as well follows from (A.9). Choosing $r_0\in(0,1)$ and a Borel selection $\varepsilon(r)\in\Lambda(\dot\vartheta(r),-\mathrm{D}\mathcal{E}_t(\vartheta(r)))$ (which is therefore locally bounded away from 0), we set

$$\mathsf{s}(r):=\int_{r_0}^{r}1/\varepsilon(\rho)\,\mathrm{d}\rho, \quad \mathsf{r}:=\mathsf{s}^{-1}, \tag{B.8}$$

so that r is defined in a suitable interval of $\mathbb{R}$ and satisfies

$$\dot{\mathsf{r}}(s)=\varepsilon(\mathsf{r}(s)), \quad \dot\theta(s)=\varepsilon(\mathsf{r}(s))\vartheta(\mathsf{r}(s)).$$

Ad (T4). We simply introduce the disjoint open sets

$$V:=\Big\{r\in(0,1):-\mathrm{D}\mathcal{E}_t(\vartheta(r))\notin K^*\Big\} \quad\text{and}\quad S:=(0,1)\setminus\overline{V},$$

and we consider their canonical decomposition in connected components.

Ad (T5). If ϑ is energetic, then by ($\mathrm{O}_{\mathrm{ener}}$) and (B.1) there holds $\Delta_{\mathfrak{p},\mathcal{E}}(t;u_-,u_+)=\Psi_0(u_+-u_-)$. Thus, taking into account (4.2) and (3.3) as well, we find $\mathfrak{p}(\dot\vartheta(r),-\mathrm{D}\mathcal{E}_t(\vartheta(r)))=\Psi_0(\dot\vartheta(r))$ for a.a. $r\in(0,1)$. Since its Ψ_0-velocity is constant and the total length is $\Psi_0(u_+-u_-)$, we deduce that ϑ is a constant speed minimal geodesic for Ψ_0. Conversely, the constraint $-\mathrm{D}\mathcal{E}_t(\vartheta(r))\in K^*$ satisfied by sliding transitions yields, in view of (A.5), that $\mathfrak{p}(\dot\vartheta,-\mathrm{D}\mathcal{E}_t(\vartheta(r)))=\Psi_0(\dot\vartheta(r))$ for a.a. $r\in(0,1)$. Therefore,

$$\Delta_{\mathfrak{p},\mathcal{E}}(t;u_-,u_+)=\int_0^1\Psi_0(\dot\vartheta(r))\,\mathrm{d}r=\Psi_0(u_+-u_-)$$

by the geodesic property (B.6).

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

It is well known that, if Ψ_0 has strictly convex sublevels, the related geodesics are linear segments. In order to prove the last statement, let us observe that for every $\xi \in \partial\Psi_0(u_+ - u_-) \subset K^*$ there holds

$$\int_0^1 \langle \xi, \dot\vartheta(r)\rangle \,\mathrm{d}r = \langle \xi, u_+ - u_-\rangle = \Psi_0(u_+ - u_-) = \int_0^1 \Psi_0(\dot\vartheta(r))\,\mathrm{d}r,$$

where the second equality follows from the characterization (2.2) of $\partial\Psi_0(u_+ - u_-)$. Hence,

$$\int_0^1 \Big(\Psi_0(\dot\vartheta(r)) - \langle \xi, \dot\vartheta(r)\rangle\Big)\,\mathrm{d}r = 0.$$

Since the above integrand is nonnegative (being $\xi \in K^*$), again by (2.2) we deduce that $\xi \in \partial\Psi_0(\dot\vartheta(r))$ for a.a. $r \in (0,1)$. On the other hand, if Ψ_0 is Gâteaux-differentiable outside 0, its subdifferential contains just one point. Ultimately, (B.2) (recall that ϑ is of sliding type) shows that $-\mathrm{D}\mathcal{E}_t(\vartheta(r)) = \xi$ for every $r \in [0,1]$. □

Acknowledgements. A.M. has been partially supported by DFG, Research Unit FOR 797 "MicroPlast", Mi 459/5-1. R.R. and G.S. have been partially supported by a MIUR-PRIN'06 grant for the project "Variational methods in optimal mass transportation and in geometric measure theory". The authors are grateful to an anonymous referee for several elucidating remarks, which helped us to improve the paper.

References

[1] L. Ambrosio, Minimizing movements. *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl.* **19** (1995) 191–246.
[2] L. Ambrosio and G. Dal Maso, A general chain rule for distributional derivatives. *Proc. Am. Math. Soc.* **108** (1990) 691–702.
[3] L. Ambrosio, N. Fusco and D. Pallara, *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs, Clarendon Press, Oxford (2000).
[4] L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich, 2nd edn., Birkhäuser Verlag, Basel (2008).
[5] F. Auricchio, A. Mielke and U. Stefanelli, A rate-independent model for the isothermal quasi-static evolution of shape-memory materials. *Math. Models Meth. Appl. Sci.* **18** (2008) 125–164.
[6] G. Bouchitté, A. Mielke and T. Roubíček, A complete-damage problem at small strains. *Z. Angew. Math. Phys.* **60** (2009) 205–236.
[7] M. Buliga, G. de Saxcé and C. Vallée, Existence and construction of bipotentials for graphs of multivalued laws. *J. Convex Anal.* **15** (2008) 87–104.
[8] P. Colli, On some doubly nonlinear evolution equations in Banach spaces. *Japan J. Indust. Appl. Math.* **9** (1992) 181–203.
[9] P. Colli and A. Visintin, On a class of doubly nonlinear evolution equations. *Commun. Partial Differ. Equ.* **15** (1990) 737–756.
[10] G. Dal Maso and R. Toader, A model for quasi-static growth of brittle fractures: existence and approximation results. *Arch. Ration. Mech. Anal.* **162** (2002) 101–135.
[11] G. Dal Maso and R. Toader, A model for the quasi-static growth of brittle fractures based on local minimization. *Math. Models Meth. Appl. Sci.* **12** (2002) 1773–1799.
[12] G. Dal Maso and C. Zanini, Quasi-static crack growth for a cohesive zone model with prescribed crack path. *Proc. R. Soc. Edinb., Sect. A, Math.* **137** (2007) 253–279.
[13] G. Dal Maso, G. Francfort and R. Toader, Quasistatic crack growth in nonlinear elasticity. *Arch. Ration. Mech. Anal.* **176** (2005) 165–225.
[14] G. Dal Maso, A. DeSimone and M.G. Mora, Quasistatic evolution problems for linearly elastic-perfectly plastic materials. *Arch. Ration. Mech. Anal.* **180** (2006) 237–291.
[15] G. Dal Maso, A. DeSimone, M.G. Mora and M. Morini, Globally stable quasistatic evolution in plasticity with softening. *Netw. Heterog. Media* **3** (2008) 567–614.
[16] G. Dal Maso, A. DeSimone, M.G. Mora and M. Morini, A vanishing viscosity approach to quasistatic evolution in plasticity with softening. *Arch. Ration. Mech. Anal.* **189** (2008) 469–544.
[17] G. Dal Maso, A. DeSimone and F. Solombrino, Quasistatic evolution for cam-clay plasiticity: a weak formulation *via* viscoplastic regularization and time rescaling. *Calc. Var. Partial Differential Equations* (to appear).
[18] M. Efendiev and A. Mielke, On the rate-independent limit of systems with dry friction and small viscosity. *J. Convex Analysis* **13** (2006) 151–167.
[19] A. Fiaschi, A vanishing viscosity approach to a quasistatic evolution problem with nonconvex energy. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* (to appear).

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press

[20] G. Francfort and A. Garroni, A variational view of partial brittle damage evolution. *Arch. Ration. Mech. Anal.* **182** (2006) 125–152.
[21] G. Francfort and A. Mielke, Existence results for a class of rate-independent material models with nonconvex elastic energies. *J. Reine Angew. Math.* **595** (2006) 55–91.
[22] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex analysis and minimization algorithms. II: Advanced theory and bundle methods*, *Grundlehren der Mathematischen Wissenschaften* [Fundamental Principles of Mathematical Sciences] **306**. Springer-Verlag, Berlin (1993).
[23] D. Knees, A. Mielke and C. Zanini, On the inviscid limit of a model for crack propagation. *Math. Models Meth. Appl. Sci.* **18** (2008) 1529–1569.
[24] D. Knees, C. Zanini and A. Mielke, Crack propagation in polyconvex materials. *Physica D* **239** (2010) 1470–1484.
[25] M. Kočvara, A. Mielke and T. Roubíček, A rate-independent approach to the delamination problem. *Math. Mech. Solids* **11** (2006) 423–447.
[26] P. Krejčí, Evolution variational inequalities and multidimensional hysteresis operators, in *Nonlinear differential equations* (Chvalatice, 1998), *Res. Notes Math.* **404**, Chapman & Hall/CRC, Boca Raton, FL (1999) 47–110.
[27] P. Krejčí and M. Liero, Rate independent Kurzweil processes. *Appl. Math.* **54** (2009) 117–145.
[28] C.J. Larsen, Epsilon-stable quasi-static brittle fracture evolution. *Comm. Pure Appl. Math.* **63** (2010) 630–654.
[29] A. Mainik and A. Mielke, Existence results for energetic models for rate-independent systems. *Calc. Var. PDEs* **22** (2005) 73–99.
[30] A. Mainik and A. Mielke, Global existence for rate-independent gradient plasticity at finite strain. *J. Nonlin. Sci.* **19** (2009) 221–248.
[31] A. Mielke, Energetic formulation of multiplicative elasto-plasticity using dissipation distances. *Contin. Mech. Thermodyn.* **15** (2003) 351–382.
[32] A. Mielke, Evolution in rate-independent systems (Chap. 6), in *Handbook of differential equations, evolutionary equations* **2**, C. Dafermos and E. Feireisl Eds., Elsevier B.V., Amsterdam (2005) 461–559.
[33] A. Mielke, *Differential, energetic and metric formulations for rate-independent processes*. Lecture Notes, Summer School Cetraro (in press).
[34] A. Mielke and T. Roubíček, A rate-independent model for inelastic behavior of shape-memory alloys. *Multiscale Model. Simul.* **1** (2003) 571–597.
[35] A. Mielke and T. Roubíček, Rate-independent damage processes in nonlinear elasticity. *M^3AS Math. Models Meth. Appl. Sci.* **16** (2006) 177–209.
[36] A. Mielke and T. Roubíček, *Rate-Independent Systems: Theory and Application*. (In preparation).
[37] A. Mielke and F. Theil, A mathematical model for rate-independent phase transformations with hysteresis, in *Proceedings of the Workshop on Models of Continuum Mechanics in Analysis and Engineering*, H.-D. Alber, R. Balean and R. Farwig Eds., Shaker-Verlag, Aachen (1999) 117–129.
[38] A. Mielke and F. Theil, On rate-independent hysteresis models. *NoDEA* **11** (2004) 151–189.
[39] A. Mielke and A. Timofte, An energetic material model for time-dependent ferroelectric behavior: existence and uniqueness. *Math. Meth. Appl. Sci.* **29** (2006) 1393–1410.
[40] 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.* (submitted).
[41] A. Mielke, F. Theil and V.I. Levitas, A variational formulation of rate-independent phase transformations using an extremum principle. *Arch. Ration. Mech. Anal.* **162** (2002) 137–177.
[42] A. Mielke, R. Rossi and G. Savaré, Modeling solutions with jumps for rate-independent systems on metric spaces. *Discrete Contin. Dyn. Syst.* **25** (2009) 585–615.
[43] A. Mielke, R. Rossi and G. Savaré, *Nonsmooth analysis of doubly nonlinear evolution equations*. (In preparation).
[44] M. Negri and C. Ortner, Quasi-static crack propagation by Griffith's criterion. *Math. Models Meth. Appl. Sci.* **18** (2008) 1895–1925.
[45] R.T. Rockafellar, *Convex Analysis*. Princeton University Press, Princeton (1970).
[46] R. Rossi and G. Savaré, Gradient flows of non convex functionals in Hilbert spaces and applications. *ESAIM: COCV* **12** (2006) 564–614.
[47] 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.* **7** (2008) 97–169.
[48] T. Roubíček, Rate independent processes in viscous solids at small strains. *Math. Methods Appl. Sci.* **32** (2009) 825–862.
[49] U. Stefanelli, A variational characterization of rate-independent evolution. *Math. Nachr.* **282** (2009) 1492–1512.
[50] M. Thomas and A. Mielke, Damage of nonlinearly elastic materials at small strains – Existence and regularity results. *Zeits. Angew. Math. Mech.* **90** (2009) 88–112.
[51] R. Toader and C. Zanini, An artificial viscosity approach to quasistatic crack growth. *Boll. Unione Mat. Ital.* **2** (2009) 1–35.
[52] A. Visintin, *Differential models of hysteresis*, *Applied Mathematical Sciences* **111**. Springer-Verlag, Berlin (1994).

https://doi.org/10.1051/cocv/2010054 Published online by Cambridge University Press