Viscous corrections of the Time Incremental Minimization Scheme and Visco-Energetic Solutions to Rate-Independent Evolution Problems

Luca Minotti, Giuseppe Savaré *

June 13, 2016

Abstract

We propose the new notion of Visco-Energetic solutions to rate-independent systems $(X, \mathcal{E}, \mathsf{d})$ driven by a time dependent energy $\mathcal{E}$ and a dissipation quasi-distance d in a general metric-topological space X.

As for the classic Energetic approach, solutions can be obtained by solving a modified time Incremental Minimization Scheme, where at each step the dissipation (quasi-)distance d is incremented by a viscous correction δ (e.g. proportional to the square of the distance d), which penalizes far distance jumps by inducing a localized version of the stability condition.

We prove a general convergence result and a typical characterization by Stability and Energy Balance in a setting comparable to the standard energetic one, thus capable to cover a wide range of applications. The new refined Energy Balance condition compensates the localized stability and provides a careful description of the jump behavior: at every jump the solution follows an optimal transition, which resembles in a suitable variational sense the discrete scheme that has been implemented for the whole construction.

Contents

*Università di Pavia. email: luca.minotti01@universitadipavia.it, giuseppe.savare@unipv.it. G.S. has been partially supported by PRIN10/11 grant from MIUR for the project *Caculus of Variations* and by IMATI-CNR

1 Introduction

Since the pioneering papers [36, 34], energetic solutions (also called irreversible quasi-static evolutions in the fracture models studied in [13, 10, 11]) to rate-independent evolutionary systems driven by time-dependent functionals have played a crucial role and provided a unifying framework for many different applied models, such as shape memory alloys [36, 4], crack propagation [11, 10] elastoplasticity [25, 14, 8, 9, 23], damage in brittle materials [32, 5, 26] delamination [19], ferroelectricity [38], and superconductivity [46]. We refer to the recent monograph [33] for a complete discussion and overview of the theory and its applications.

In its simplest *metric* formulation, a Rate-Independent System (R.I.S.) $(X, \mathcal{E}, \mathsf{d})$ can be described by a metric space (X, d) and a time-dependent energy functional $\mathcal{E} : [0, T] \times X \to \mathbb{R}$. Energetic solutions can be obtained as a limit of piecewise constant interpolant of discrete solutions U^n_τ obtained by recursively solving the time Incremental Minimization scheme

$$\min_{U \in X} \mathcal{E}(t^n_\tau, U) + \mathsf{d}(U^{n-1}_\tau, U). \tag{IM$_\mathsf{d}$}$$

The main aim of the present paper is to study general *viscous corrections* of (IM$_\mathsf{d}$)

$$\min_{U \in X} \mathcal{E}(t^n_\tau, U) + \mathsf{d}(U^{n-1}_\tau, U) + \delta(U^{n-1}_\tau, U), \tag{IM$_{\mathsf{d},\delta}$}$$

obtained by perturbing the distance d by a "viscous" penalization term $\delta : X \times X \to [0, \infty)$, which should induce a better localization of the minimizers. A typical choice is the quadratic correction $\delta(u, v) := \frac{\mu}{2}\mathsf{d}^2(u, v)$, for some $\mu > 0$.

We will show that solutions generated by the scheme ($\mathrm{IM}_{\mathsf{d},\delta}$) exhibit a sort of intermediate behaviour between Energetic and Balanced Viscosity solutions [28], since they retain the great structural robustness of the former and allow for a more localized response typical of the latter. Before explaining these novel features, let us briefly recall a few basic facts concerning Energetic and Balanced Viscosity solutions.

Energetic solutions. Energetic solutions to the R.I.S. $(X, \mathcal{E}, \mathsf{d})$ are curves $u : [0, T] \to X$ with bounded variation that are characterized by two variational conditions, called *stability* (S_{d}) and *Energy Balance* (E_{d}):

$$\mathcal{E}(t, u(t)) \le \mathcal{E}(t, v) + \mathsf{d}(u(t), v) \quad \text{for every } v \in X,\ t \in [0, T], \tag{S_{d}}$$

$$\mathcal{E}(t, u(t)) + \mathrm{Var}_{\mathsf{d}}(u, [0, t]) = \mathcal{E}(0, u_0) + \int_0^t \mathcal{P}(r, u(r))\,\mathrm{d}r \quad \text{for every } t \in [0, T]. \tag{E_{d}}$$

In (E_{d}) $\mathrm{Var}_{\mathsf{d}}(u, [0, t])$ denotes the usual pointwise total variation of u on the interval $[0, t]$ (see (2.6)) and $\mathcal{P}(t, u) = \partial_t \mathcal{E}(t, u)$ is the partial derivative of the energy $\mathcal{E}$ with respect to (w.r.t.) time, which we assume to be continuous and satisfying the uniform bound

$$|\mathcal{P}(t, x)| \le C_0\big(\mathcal{E}(t, x) + C_1\big) \quad \text{for every } x \in X \tag{1.1}$$

for some constants $C_0, C_1 \ge 0$.

As we mentioned, one of the strongest features of the energetic approach is the possibility to construct energetic solutions by solving the time *Incremental Minimization scheme* (IM_{d}) (also called *Minimizing Movement method* in the De Giorgi approach to metric gradient flows, see [2]). If $\mathcal{E}$ has compact sublevels then for every ordered partition $\tau = \{t_\tau^0 = 0, t_\tau^1, \cdots, t_\tau^{N-1}, t_\tau^N = T\}$ of the interval $[0, T]$ with variable time step $\tau^n := t_\tau^n - t_\tau^{n-1}$ and for every initial choice $U_\tau^0 = u(0)$ we can construct by induction an approximate sequence $(U_\tau^n)_{n=0}^N$ solving (IM_{d}).

If $\overline{U}_\tau$ denotes the left-continuous piecewise constant interpolant of $(U_\tau^n)_n$ which takes the value U_τ^n on the interval $(t^{n-1}, t_\tau^n]$, then the family of discrete solutions $\overline{U}_\tau$ has limit curves with respect to pointwise convergence as the maximum of the step sizes $|\tau| = \max \tau^n$ vanishes, and every limit curve u is an energetic solution.

A second important fact concerns the mutual interaction between the Stability and the Energy Balance conditions (S_{d})-(E_{d}): it is possible to prove that for every curve u satisfying (S_{d}), (E_{d}) is in fact equivalent to the Energy-Dissipation inequality

$$\mathcal{E}(t, u(t)) + \mathrm{Var}_{\mathsf{d}}(u, [0, t]) \le \mathcal{E}(0, u_0) + \int_0^t \mathcal{P}(r, u(r))\,\mathrm{d}r \quad \text{for every } t \in [0, T]. \tag{1.2}$$

When

$$X = \mathbb{R}^d, \quad \mathsf{d}(x, y) := \alpha|y - x|,\ \alpha > 0, \quad \mathcal{E}(t, \cdot) \text{ is sufficiently smooth,} \tag{1.3}$$

and $\mathrm{D}_x^2\mathcal{E}(t, x) \ge \lambda I$, $\lambda > 0$, so that $\mathcal{E}(t, \cdot)$ is uniformly convex, then it is possible to prove that energetic solutions are continuous and can be equivalently characterized by the doubly nonlinear evolution inclusion

$$\alpha\,\partial\psi(\dot{u}(t)) + \mathrm{D}\mathcal{E}(t, u(t)) \ni 0, \quad \psi(v) := |v|. \tag{1.4}$$

Even simple 1-dimensional nonconvex examples, e.g. when the energy has the form

$$\mathcal{E}(t,x) := W(x) - \ell(t)x \text{ for a double well potential such as } W(x) = (x^2-1)^2,\ x \in \mathbb{R}, \quad (1.5)$$

show that energetic solutions have jumps, preventing the violation of the global stability condition. In fact, combining stability and energy balance, it is possible to check that at every jump point $t \in \mathrm{J}_u$, the left and right limits $u(t-)$, $u(t+)$ of a solution u satisfy the energetic jump conditions

$$\mathsf{d}(u(t-),u(t)) = \mathcal{E}(t,u(t-)) - \mathcal{E}(t,u(t)), \quad \mathsf{d}(u(t),u(t+)) = \mathcal{E}(t,u(t)) - \mathcal{E}(t,u(t+)), \quad (1.6)$$

which are strongly influenced by the global energy landscape of $\mathcal{E}$. This reflects the global constraint imposed by the stability condition, whose violation induces the jump (see e.g. [16, Ex. 6.3], [27, Ex. 1]).

For instance, in the case of example (1.3)-(1.5) with $\ell \in \mathrm{C}^1([0,T])$ strictly increasing with $u_0 < -1$ and $\ell(0) = \alpha + W'(u_0)$, it is possible to prove [45] that an energetic solution u is an increasing selection of the equation

$$\alpha + \partial W^{**}(u(t)) \ni \ell(t) \quad (1.7)$$

where W^{**} is the convex envelope $W^{**}(x) = \big((x^2-1)_+\big)^2$, independently of the parameter $\alpha > 0$.

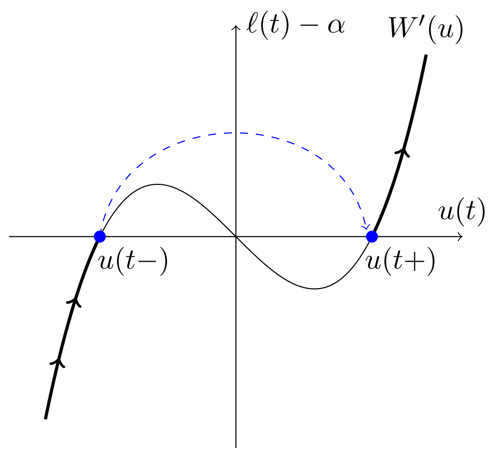

Figure 1: Energetic solution for a double-well energy W with an increasing load ℓ, see (1.7).

Balanced Viscosity solutions. In order to obtain a formulation where local effects are more relevant (see [24, Sec. 6], [11, 43, 12, 21, 39]) various kinds of corrections have been considered. A natural one introduces a viscous correction to the incremental minimization scheme (IM_{d}), penalizing the square of the distance from the previous step

$$\min_{U \in X} \mathcal{E}(t_\tau^n, U) + \mathsf{d}(U_\tau^{n-1}, U) + \frac{\varepsilon^n}{2\tau^n}\mathsf{d}^2(U_\tau^{n-1}, U), \quad (\mathrm{IM}_{\mathsf{d},\varepsilon})$$

for a parameter $\varepsilon^n = \varepsilon^n(\tau) \downarrow 0$ with $\frac{\varepsilon^n(\tau)}{|\tau|} \uparrow +\infty$. In the previous Euclidean framework (1.3), (IM_{d}) corresponds to the discretization of the generalized gradient flow

$$\alpha\partial\psi(\dot{u}(t)) + \varepsilon\dot{u}(t) + \mathrm{D}\mathcal{E}(t, u(t)) \ni 0. \tag{1.8}$$

Such kinds of approximations have been studied in a series of contributions [44, 27, 28, 30], also dealing with more general corrections in metric and linear settings. Under suitable smoothness and lower semicontinuity assumptions involving the metric slope of $\mathcal{E}$ it is possible to prove that all the limit curves satisfy a local stability assumption and a modified Energy Balance, involving an augmented total variation that encodes a more refined description of the jump behaviour of u: roughly speaking, a jump between $u(t-)$ and $u(t+)$ occurs only when these values can be connected by a rescaled solution ϑ of (1.8), where the energy is frozen at the jump time t (see the next section 2.4):

$$\alpha\partial\psi(\dot{\vartheta}(s)) + \dot{\vartheta}(s) + \mathrm{D}\mathcal{E}(t, \vartheta(s)) \ni 0. \tag{1.9}$$

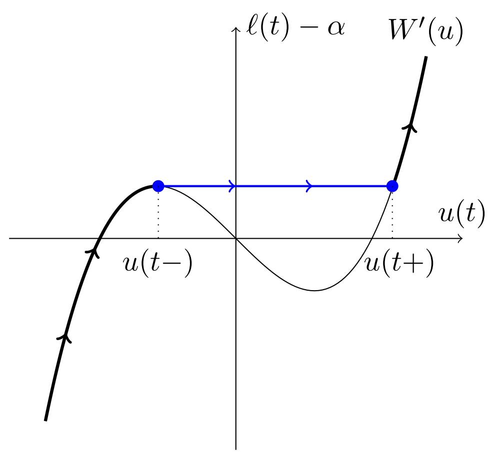

Figure 2: BV solution for a double-well energy W with an increasing load ℓ. The blue line denotes the path described by the optimal transition ϑ solving (1.9).

One of the main technical difficulties of the theory of Balanced Viscosity solutions is related to the properties of the slope of $\mathcal{E}$, which can be difficult to check when highly nonsmooth-nonconvex energies are involved.

More degenerate situations when $\alpha = 0$ can also be considered, both from the continuous (see [1]) and the discrete point of view (see [3], who considers a different dependence with respect to time, given by a time-dependent linear constraint): the main difficulty here relies on the loss of time-compactness, since simple estimates of the total variation of the approximating curves are missing.

Viscous corrections of the Incremental Minimization Scheme The present paper introduces and studies an intermediate situation between Energetic and Balanced Viscosity solutions, when one keeps constant the ratio $\mu := \varepsilon^n/\tau^n$ in ($\mathrm{IM}_{\mathsf{d},\varepsilon}$). In this way the metric dissipation d is corrected by an extra viscous penalization term $\delta(u, v) := \frac{\mu}{2}\mathsf{d}^2(u, v)$ which induces a localization of the minimizer, tuned by the parameter $\mu > 0$. At each step n we thus propose to solve a modified Incremental Minimization scheme of the form

$$\begin{aligned} &\text{select } U_\tau^n \in \mathrm{M}(t_\tau^n, U_\tau^{n-1}), \qquad \text{where for every } t \in [0,T] \text{ and } x \in X \\ &\mathrm{M}(t,x) := \operatorname*{argmin}_{y \in X} \Big\{ \mathcal{E}(t,y) + \mathsf{d}(x,y) + \delta(x,y) \Big\}, \qquad \delta(x,y) := \frac{\mu}{2}\mathsf{d}^2(x,y). \end{aligned} \tag{IM$_{\mathsf{d},\delta}$}$$

Even if much more general viscous corrections δ could be considered (see Section 3.1), in this Introduction we will choose the simpler quadratic one for ease of exposition. Notice that in the finite dimensional case (1.3) when the Hessian of the energy is bounded from below, i.e. $\mathrm{D}_x^2\mathcal{E}(t,x) \geq -\lambda I$ for every t, x and some $\lambda \geq 0$, the choice $\mu\alpha^2 \geq \lambda$ yields a *convex* incremental problem (IM$_{\mathsf{d},\delta}$), which could greatly help in the effective computation of the solution. Differently from [3], we do not need to construct U_τ^n by freezing the time variable at t_τ^n and iterating the minimization scheme to converge to a critical point: after each incremental minimization step the energy is immediately updated to the new value at the time t_τ^{n+1}.

Since $\delta \geq 0$, it is not difficult to check that the family of discrete solutions $\overline{U}_\tau$ has uniformly bounded d-total variation and takes value in a compact set of X, so that it always admits limit curves $u \in \mathrm{BV}_{\mathsf{d}}([0,T];X)$. The difficult task here concerns the characterization of such limit curves. One of the main problems underlying the simple scheme (IM$_{\mathsf{d},\delta}$) is the loss of the triangle inequality for the total dissipation

$$\mathsf{D}(u,v) := \mathsf{d}(u,v) + \delta(u,v) = \mathsf{d}(u,v) + \frac{\mu}{2}\mathsf{d}^2(u,v). \tag{1.10}$$

In the case of Energetic solutions, the triangle inequality of d lies at the core of two crucial properties:

a) every solution U^n of the minimization step (IM$_{\mathsf{d}}$) satisfies the stability condition (S$_{\mathsf{d}}$) at $t = t_\tau^n$;

b) the computation of the total variation of a piecewise constant map $\overline{U}_\tau$ associated with some partition τ involves only consecutive points, i.e.

$$\mathrm{Var}_{\mathsf{d}}(\overline{U}_\tau, [0,T]) = \sum_{n=1}^{N} \mathsf{d}(U_\tau^{n-1}, U_\tau^n)$$

and the total variation functional $u \mapsto \mathrm{Var}_{\mathsf{d}}(u,[0,T])$ is lower semicontinuous w.r.t. pointwise convergence, so that at least an Energy inequality corresponding to (E$_{\mathsf{d}}$) can be easily deduced from the corresponding version at the discrete level.

Such properties fail in the case of the augmented dissipation D of (1.10). In particular, even in the finite-dimensional setting (1.3) with $\alpha = 1$, it is easy to check that e.g. Lipschitz curves $u : [0,T] \to \mathbb{R}^d$ can be approximated by piecewise constant interpolants $\overline{U}_\tau$ on uniform partitions $\tau = \{nT/N\}_{n=0}^N$ with $U_\tau^n = u(nT/N)$ and $|\tau| = T/N$, so that

$$\lim_{|\tau|\downarrow 0} \sum_{n=1}^{N} |U_\tau^{n-1} - U_\tau^n| + \frac{\mu}{2}|U_\tau^{n-1} - U_\tau^n|^2 = \mathrm{Var}_{\mathsf{d}}(u,[0,T]) = \int_0^T |\dot{u}(t)|\,\mathrm{d}t. \tag{1.11}$$

Visco-Energetic solutions. Nevertheless, by using more refined arguments and guided by the results obtained in the Balanced Viscosity approach, we are able to obtain a precise variational characterization of the limit curves (called *Visco-Energetic solutions*), still stated in terms of suitably adapted stability and energy balance conditions.

Concerning stability, we obtain a natural generalization of $(\mathrm{S_d})$

$$\mathcal{E}(t,u(t)) \le \mathcal{E}(t,v) + \mathsf{D}(u(t),v) \quad \text{for every } v \in X,\ t \in [0,T] \setminus \mathrm{J}_u, \tag{$\mathrm{S_D}$}$$

which is naturally associated with the D-*stable set*

$$\mathscr{S}_{\mathsf{D}} := \Big\{(t,x) : \mathcal{E}(t,x) \le \mathcal{E}(t,y) + \mathsf{D}(x,y) \quad \text{for every } y \in X\Big\}. \tag{1.12}$$

Notice that in the finite dimensional case (1.3) when μ is sufficiently big so that $\mathrm{D}^2_x\mathcal{E}(t,x) \ge -\mu\alpha^2 I$, $(\mathrm{S_d})$ is in fact a local condition, which can be restated as

$$|\mathrm{D}_x\mathcal{E}(t,u(t))| \le \alpha \quad \text{for every } t \in [0,T] \setminus \mathrm{J}_u. \tag{1.13}$$

The right replacement of the Energy Balance condition is harder to formulate. Since $(\mathrm{S_D})$ is weaker than $(\mathrm{S_d})$, it is clear that the Energy Dissipation inequality (1.2) (which still trivially holds for limits of $(\mathrm{IM}_{\mathsf{d},\delta})$) will not be enough to recover the energy balance: in particular, important pieces of information are lost along the jumps. A heuristic idea, which one can figure out by the direct analysis of simple cases such as (1.3)-(1.5), is that jump transitions between $u(t-)$ and $u(t+)$ should be described by discrete trajectories $\vartheta : Z \to X$ defined in a subset $Z \subset \mathbb{Z}$ such that each value $\vartheta(n) \in \mathrm{M}(t,\vartheta(n-1))$ is a minimizer of the "frozen" incremental problem at time t with datum $\vartheta(n-1)$. In the simplest cases $Z = \mathbb{Z}$, the left and right jump values $u(t\pm)$ are the limit of $\vartheta(n)$ as $n \to \pm\infty$, but more complicated situations can occur, when Z is a proper subset of $\mathbb{Z}$ or one has to deal with concatenation of (even countable) discrete transitions and sliding parts parametrized by a continuous variable, where the stability condition $(\mathrm{S_D})$ holds.

In order to capture all of these possibilities, we will introduce a quite general notion of transition parametrized by a continuous map $\vartheta : E \to X$ defined in an arbitrary compact subset of $\mathbb{R}$ such that $\vartheta(\min E) = u(t-)$ and $\vartheta(\max E) = u(t+)$. The cost of such kind of transition results from the contribution of three parts: the first one is the usual total variation $\mathrm{Var}_{\mathsf{d}}(\vartheta, E)$ (see the next (2.6)). The second contribution arises at each "gap" in E, i.e. a bounded connected component $I = (I^-, I^+)$ of $\mathbb{R} \setminus E$: denoting by $\mathfrak{H}(E)$ the collection of all these intervals, we will set

$$\mathrm{GapVar}_\delta(\vartheta, E) := \sum_{I \in \mathfrak{H}(E)} \delta(\vartheta(I^-), \vartheta(I^+)). \tag{1.14}$$

The last contribution detects if ϑ violates the stability condition at $s \in E$: it is defined as the sum

$$\sum_{\substack{s \in E \\ s < \max E}} \mathcal{R}(t, \vartheta(s)) \tag{1.15}$$

where $\mathcal{R}$ is the residual stability function

$$\mathcal{R}(t,x) := \max_{y \in X} \mathcal{E}(t,x) - \mathcal{E}(t,y) - \mathsf{D}(x,y) = \mathcal{E}(t,x) - \min_{y \in X} \Big(\mathcal{E}(t,y) + \mathsf{D}(x,y)\Big). \tag{1.16}$$

Since it is easy to check that $\mathcal{R}(t,\theta)=0$ if and only if $(t,\theta)\in\mathscr{S}_{\mathsf{D}}$, $\mathcal{R}(t,\cdot)$ provides a measure of the violation of the stability constraint.

The total cost of a transition $\vartheta : E\to X$ at a jump time t is therefore

$$\mathrm{Trc}(t,\vartheta,E) := \mathrm{Var}_{\mathsf{d}}(\vartheta,E)+\mathrm{GapVar}_{\delta}(\vartheta,E)+\sum_{\substack{s\in E\\ s<\max E}}\mathcal{R}(t,\vartheta(s)), \tag{1.17}$$

and the corresponding cost c for a jump from $u(t-)$ to $u(t+)$ passing through the value $u(t)$ is given by

$$\begin{aligned}\mathsf{c}(t,u(t-),u(t),u(t+)) := \inf\Big\{ &\mathrm{Trc}(t,\vartheta,E) : \vartheta\in \mathrm{C}(E,X),\ \vartheta(E)\ni u(t),\\ &\vartheta(\min E)=u(t-),\ \vartheta(\max E)=u(t+)\Big\},\end{aligned} \tag{1.18}$$

where the infimum is attained whenever there is at least one admissible transition with finite cost. Notice that the cost c is always bigger than the corresponding value computed by the dissipation distance d, i.e. the quantity

$$\Delta_{\mathsf{c}}(t,u(t-),u(t),u(t+)) := \mathsf{c}(t,u(t-),u(t),u(t+)) - \mathsf{d}(u(t-),u(t)) - \mathsf{d}(u(t),u(t+)) \tag{1.19}$$

is nonnegative. c always controls the energy dissipation along the jump, i.e.

$$\mathrm{Trc}(t,\vartheta,E)\ge \mathsf{c}(t,u(t-)),u(t),u(t+))\ge \mathcal{E}(t,u(t-))-\mathcal{E}(t,u(t+)). \tag{1.20}$$

With these notions at our disposal, we can eventually write the Energy Balance condition for Visco-Energetic solutions

$$\mathcal{E}(t,u(t))+\mathit{Var}_{\mathsf{d},\mathsf{c}}(u,[0,t])=\mathcal{E}(0,u_0)+\int_0^t \mathcal{P}(r,u(r))\,\mathrm{d}r\quad\text{for every } t\in[0,T], \tag{$\mathrm{E}_{\mathsf{d},\mathsf{c}}$}$$

where the augmented total variation $\mathit{Var}_{\mathsf{d},\mathsf{c}}(u,[a,b])$ differs from the usual one $\mathrm{Var}_{\mathsf{d}}(u,[a,b])$ by an extra contribution at the jump points $t\in \mathrm{J}_u$:

$$\mathit{Var}_{\mathsf{d},\mathsf{c}}(u,[a,b]) := \mathrm{Var}_{\mathsf{d}}(u,[a,b])+\sum_{t\in \mathrm{J}_u\cap[a,b]}\Delta_{\mathsf{c}}(t,u(t-),u(t),u(t+)). \tag{1.21}$$

As in the case of energetic solutions, once the stability condition (S_{D}) is satisfied, it is sufficient to check the Energy-Dissipation inequality associated with ($\mathrm{E}_{\mathsf{d},\mathsf{c}}$), since the extra term appearing in the definition of $\mathit{Var}_{\mathsf{d},\mathsf{c}}(u,[0,t])$ (1.21) provides the right correction to compensate the weaker stability property. At each jump point we thus obtain the Visco-Energetic jump conditions corresponding to (1.6)

$$\mathsf{c}(t,u(t-),u(t)) = \mathcal{E}(t,u(t-))-\mathcal{E}(t,u(t)),\quad \mathsf{c}(t,u(t),u(t+)) = \mathcal{E}(t,u(t))-\mathcal{E}(t,u(t+)). \tag{1.22}$$

One of the beautiful aspects of this formulation is that for each $t\in \mathrm{J}_u$ there always exists an optimal transition $\vartheta : E\to X$ connecting $u(t-)$ to $u(t+)$ and passing through $u(t)$ such that

$$\mathcal{E}(t,u(t-))-\mathcal{E}(t,u(t+))=\mathrm{Trc}(t,\vartheta,E). \tag{1.23}$$

In the case when $(t,\vartheta(s))\notin\mathscr{S}_{\mathsf{D}}$ for some $s\in E$ we can prove that s is isolated and denoting by $s_- := \max E\cap(-\infty,s)$ we recover the property

$$\vartheta(s)\in \mathrm{M}(t,\vartheta(s_-)), \tag{1.24}$$

which provides an important description of optimal transitions (see Figure 3 for a simple example).

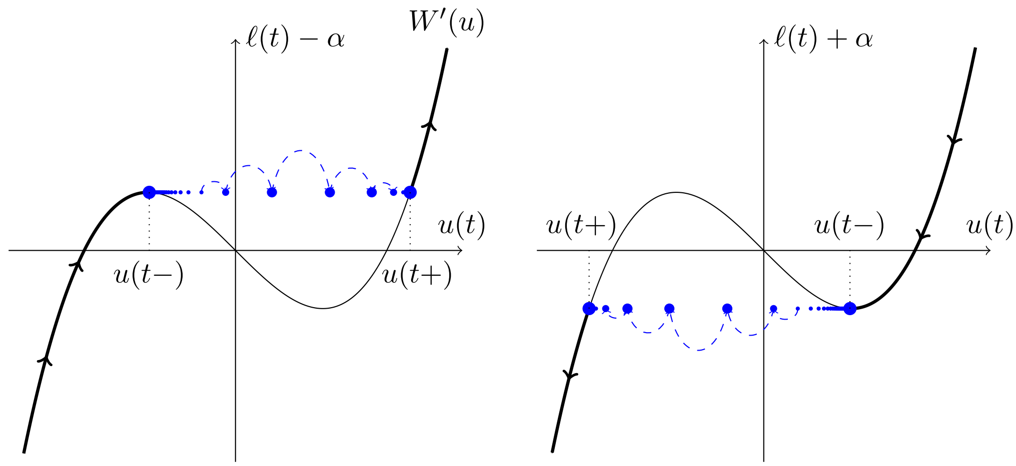

Figure 3: Visco-Energetic solutions for a double-well energy W with an increasing (on the left) and a decreasing (on the right) load ℓ and the choices $\mu\alpha^2 \geq -\min W''$. Jumps occur when u reaches a local maximum (or minimum) of W as in the Balanced Viscosity case. The optimal transitions ϑ are infinite sequences of jumps.

Further generalizations and plan of the paper. In the paper we try to develop the ideas above at the highest level of generality, hoping that the present visco-energetic theory can reach the same power of the energetic one. In particular

- we separate the roles of the dissipation distance and of the topology, by considering a general metric-topological setting, where the compactness assumptions are stated in terms of a weaker topology σ (see Section 2.1).

- we consider general lower semicontinuous asymmetric quasi-distances d, possibly taking the value $+\infty$ (2.1). As in the energetic framework, in this case a further closedness condition involving the stable set will play a crucial role: in the visco-energetic setting, we will need the closure of the Q-quasi stable sets, $Q \geq 0$, of the points $x \in X$ satisfying

$$\mathcal{E}(t,x) \leq \mathcal{E}(t,y) + \mathsf{D}(x,y) + Q \quad \text{for every } y \in X. \tag{1.25}$$

  Notice that (1.25) reduces to the definition of the stable set when $Q = 0$.

- we try to relax the assumptions concerning the time-differentiability of $\mathcal{E}$, thus allowing for super-differentiable energies: this is particularly useful to cover the important case of product spaces $X := Y \times Z$ (see Section 4.5) where d controls only the Z-component and one has to deal with reduced/marginal energies

$$\widetilde{\mathcal{E}}(t,z) := \min_{y\in Y} \mathcal{E}(t,y,z). \tag{1.26}$$

- we consider quite general viscous corrections δ, not necessarily obtained as a function of the quasi-distance d (see Section 3.1).

In the preliminary section 2 we briefly recall the canonical metric-topological setting, how to deal with BV and regulated functions, the properties of the energy $\mathcal{E}$ and its power $\mathcal{P}$, the basic framework of Energetic and Balanced Viscosity solutions.

We collect our main results in section 3: we start by discussing admissible viscous corrections δ and we introduce in full detail the associated viscous jump cost c, relying on generalized transitions, and the residual stability function $\mathcal{R}$ (Section 3.2). ***Section 3.3 contains the precise definition of VE solutions, their basic characterizations, and the main existence theorem 3.9.*** In the case when the dissipation distance d is not continuous w.r.t. σ, the properties of $\mathcal{R}$ will play a crucial role, so we investigate them in Section 3.4 and we will apply these results to elucidate the structure of optimal jump transitions in Section 3.5.

Examples and applications are collected in Section 4, starting from the simplest convex or 1-dimensional cases, and moving towards more complicated situations, where δ may depend on an accessory distance d_* (Section 4.3), d is degenerate but still separates the stable set (Section 4.4), X is a product space and we will have to introduced a reduced marginal energy as in (1.26) (Section 4.5).

The last sections contain all the proofs and the relevant properties of the transition cost and the Viscous Incremental Minimization scheme. Section 5 is devoted to the properties of the cost Trc of a transition ϑ and to the existence of optimal transitions.

Section 6 contains the crucial lower energy estimates along jumps and along arbitrary BV curves satisfying the stability condition (S_D), thus proving that

$$\mathcal{E}(t,u(t)) + \mathrm{Var}_{\mathsf{d},\mathsf{c}}(u,[s,t]) \geq \mathcal{E}(0,u_s) + \int_s^t \mathcal{P}(r,u(r))\,\mathrm{d}r \quad \text{for every } s,t \in [0,T],\ s \leq t, \tag{1.27}$$

whenever (S_D) holds in $[0,T]$.

The last Section 7 contains all the main steps of the proof, which follows a canonical strategy: discrete estimates for the Viscous Incremental Minimization scheme ($\mathrm{IM}_{\mathsf{d},\delta}$), compactness, energy-dissipation inequality

$$\mathcal{E}(t,u(t)) + \mathrm{Var}_{\mathsf{d},\mathsf{c}}(u,[0,t]) \leq \mathcal{E}(0,u_0) + \int_0^t \mathcal{P}(r,u(r))\,\mathrm{d}r \quad \text{for every } t \in [0,T], \tag{1.28}$$

obtained by the lower semicontinuity results of Section 5, and conclusion by reinforcing energy convergence at each time t thanks to (1.27).

Acknowledgment

We would like to thank Riccarda Rossi for insightful discussions on the whole subject and for valuable comments on the manuscript.

Part of this paper has been written while the second author was visiting the Erwin Schrödinger Institute for Mathematics and Physics (Vienna), whose support is gratefully acknowledged.

List of notation

(X, σ)	The reference topological Hausdorff space, Section 2.1
d	the asymmetric (quasi-)distance on X, (2.1)
d separates U	(2.2)
$\mathrm{Var}_{\mathsf{d}}(u, E)$	pointwise total variation w.r.t. d of $u : E \to X$, (2.6)
(σ, d)-regulated functions	Definition 2.3
J_u	Jump set of a (σ, d)-regulated function u, (2.11)
$\mathrm{BV}_{\sigma,\mathsf{d}}([a, b]; X)$	Space of (σ, d)-regulated function with bounded variation, 2.3
$\Delta_{\mathsf{e}}(t, u^-, u^+)$	incremental cost function associated with e, (2.13)
$\mathrm{Jmp}_{\mathsf{e}}(u, [a, b])$	incremental jump variation induced by Δ_{e}, Definition 2.5
$Var_{\mathsf{d},\mathsf{e}}(u, [a, b])$	Augmented total variation, Definition 2.5
$\mathcal{E}(t, u)$	the energy functional, Section 2.2
$\mathcal{P}(t, u)$	the power functional, $\partial_t \mathcal{E}(t, u)$, Section 2.2
$\mathcal{F}(t, u), \mathcal{F}_0(u)$	perturbed energy through the distance d, (2.18)
$\mathsf{D}(u, v), \delta(u, v)$	modified viscous dissipations, (3.2)
$(X, \mathcal{E}, \mathsf{d}, \delta)$	the basic Visco-Energetic Rate-Independent System
$\mathscr{S}_{\mathsf{D}}, \mathscr{S}_{\mathsf{D}}(t)$	the D-stable set and its sections, Definition 3.2
$\mathcal{R}(t, u)$	the residual stability function, Definition 3.4
$\mathrm{M}(t, u)$	the set of minimizers of the Incremental Minimization scheme, (3.32)
τ	a finite partition $\{t^0_\tau, t^1_\tau, \cdots, t^N_\tau\}$ of the time interval $[0, T]$, see page 18
U^n_τ	discrete solutions to the time incremental minimization scheme ($\mathrm{IM}_{\mathsf{d},\delta}$)
$\overline{U}_\tau$	left continuous piecewise constant interpolant of the values U^n_τ
E^-, E^+	infimum and supremum of a set $E \subset \mathbb{R}$, Section 3.2
$\mathfrak{H}(E)$	collection of the bounded connected components of $\mathbb{R} \setminus E$, Section 3.2
$\mathfrak{P}_f(E)$	collection of all the finite subset of E
$\mathrm{C}_{\sigma,\mathsf{d}}(E; X)$	σ- and d- continuous functions $f : E \to X$, (3.20)
$\mathrm{Trc}(t, \vartheta, E)$	the transition cost, Definition 3.5
$\mathrm{GapVar}_\delta(\vartheta, E)$	one of the component of the transition cost, Definition 3.5
$\mathsf{c}(t, u^-, u^+)$	the Visco-Energetic jump dissipation cost, Definition 3.6

Assumptions:

$\langle$A$\rangle$	Energy and power, Page 16
$\langle$B$\rangle$	Admissible viscous corrections, Page 21
$\langle$C$\rangle$	Closure and separation of the (quasi)-stable set, Page 25

2 Notation, assumptions and preliminary results

In this section we recall some notation and properties related to asymmetric (quasi-)distances in topological spaces, regulated BV functions and *Energetic* and *Balanced Viscosity* (BV) *solutions* of a general rate-independent system.

2.1 The metric-topological setting.

Let (X,σ) be a Hausdorff topological space satisfying the first axiom of countability; we will fix a reference point $x_o \in X$ and a time interval $[0,T] \subset \mathbb{R}$, $T > 0$.

Asymmetric dissipation distances. The first basic object characterizing a Rate-Independent System (R.I.S.) is

$$\begin{gathered}\text{a l.s.c. } \textit{asymmetric (quasi-)distance } \mathsf{d} : X \times X \to [0,\infty], \text{ satisfying} \\ \mathsf{d}(x,x) = 0, \quad \mathsf{d}(x_o,x) < \infty, \quad \mathsf{d}(x,z) \le \mathsf{d}(x,y) + \mathsf{d}(y,z) \quad \text{for every } x,y,z \in X.\end{gathered} \tag{2.1}$$

We say that a subset $U \subset X$ is d-*bounded* if $\sup_{u\in U} \mathsf{d}(x_o,u) < \infty$. We say that d separates the points of $U \subset X$ if

$$u,v \in U, \quad \mathsf{d}(u,v) = 0 \qquad \Rightarrow \qquad u = v. \tag{2.2}$$

We will often deal with subsets of the product space $Y := \mathbb{R} \times X$, which will be endowed with the product topology $\sigma_{\mathbb{R}}$, the asymmetric distance

$$\mathsf{d}_{\mathbb{R}}((s,x),(t,y)) := |t-s| + \mathsf{d}(x,y) \quad \text{and the distinguished point } y_o := (0,x_o). \tag{2.3}$$

Notice that $\mathscr{U} \subset \mathbb{R} \times X$ is separated by $\mathsf{d}_{\mathbb{R}}$ if d separates the points of all its sections $\mathscr{U}(t) := \{u \in X : (t,u) \in \mathscr{U}\}$, $t \in \mathbb{R}$.

The relation between σ and d will be clarified by the following Lemma: we will typically choose W as a sequentially compact subset of X or $\mathbb{R} \times X$.

Lemma 2.1 *Let (Z,σ_Z) and (W,σ_W) be Hausdorff topological spaces satisfying the first axiom of countability; we suppose that W is sequentially compact and it is endowed with a l.s.c. asymmetric quasi-distance d_W as in* (2.1) *and we fix an accumulation point of $z_0 \in Z$, with neighborhood basis $\mathcal{N}(z_0)$.*

i) If $v : Z \to W$ satisfies

$$\lim_{z\to z_0} \mathsf{d}_W(v(z),v(z_0)) \wedge \mathsf{d}_W(v(z_0),v(z)) = 0, \qquad \bigcap_{N\in\mathcal{N}(z_0)} \overline{v(N)} \quad \textit{is separated by } \mathsf{d}_W, \tag{2.4}$$

then $\lim_{z\to z_0} v(z) = v(z_0)$.

ii) If $v : Z \to W$ satisfies

$$\lim_{z,z'\to z_0} \mathsf{d}_W(v(z),v(z')) \wedge \mathsf{d}_W(v(z'),v(z)) = 0, \qquad \bigcap_{N\in\mathcal{N}(z_0)} \overline{v(N)} \quad \textit{is separated by } \mathsf{d}_W, \tag{2.5}$$

then there exists the limit $\bar v := \lim_{z\to z_0} v(z)$ and $\lim_{z\to z_0} \mathsf{d}_W(v(z),\bar v) \wedge \mathsf{d}_W(\bar v,v(z)) = 0$.

Proof. We will prove the claim ii), since the proof of point i) is completely analogous. We set $\mathsf{d}_{W,\wedge}(w,w') := \mathsf{d}_W(w,w') \wedge \mathsf{d}_W(w',w)$. Since W is sequentially compact and σ_Z, σ_W satisfy the first countability axiom, in order to prove the existence of the limit it is sufficient to show that whenever sequences $z_n', z_n'' \to z_0$ with $z_n', z_n'' \in Z$ and $v(z_n') \to v'$, $v(z_n'') \to v''$ then $v' = v''$.

By the first of (2.5) for every $\varepsilon > 0$ we find $\bar{n} \in \mathbb{N}$ such that $\mathsf{d}_{W,\wedge}(v(z_n'), v(z_m'')) \le \varepsilon$ for $n, m \ge \bar{n}$; since $\mathsf{d}_{W,\wedge}$ is σ_W lower semicontinuous, we obtain

$$\mathsf{d}_{W,\wedge}(v(z_n'), v) \le \liminf_{m \to \infty} \mathsf{d}_{W,\wedge}(v(z_n'), v(z_m'')) \le \varepsilon \quad \text{for every } n \ge \bar{n}.$$

Passing to the limit as $n \to \infty$ we obtain $\mathsf{d}_{W,\wedge}(v', v'') \le \varepsilon$; since $\varepsilon > 0$ is arbitrary we get $\mathsf{d}_{W,\wedge}(v', v'') = 0$. Since v', v'' belong to $\bigcap_{N \in \mathcal{N}(z_0)} \overline{v(N)}$ which is separated by d_W we conclude that $v' = v''$. $\square$

Remark 2.2 (The metric setting) A typical situation occurs when d is a distance on X, i.e. it is symmetric, finite, and separates the points of X, so that (X, d) is a standard metric space, and σ is the induced topology. In this case, which we will simply call *the metric setting*, part of the previous discussion and of the next developments can be stated in a much simpler form. A less restrictive notion in the case of an asymmetric distance (thus not inducing a topology) is the left-continuity property, that we will introduce in formula (2.23) below.

Pointwise total variation and (σ, d)-regulated functions. Let $E \subset \mathbb{R}$ be an arbitrary subset; we will denote by $\mathfrak{P}_f(E)$ the collection of all the finite subsets of E and we will set $E^- := \inf E$, $E^+ := \sup E$. The *pointwise* total variation $\mathrm{Var}_{\mathsf{d}}(u, E)$ of a function $u : E \to X$ is defined in the usual way by

$$\mathrm{Var}_{\mathsf{d}}(u, E) := \sup \left\{ \sum_{j=1}^{M} \mathsf{d}\left(u(t_{j-1}), u(t_j)\right) : t_0 < t_1 < \ldots < t_M,\ \{t_i\}_{i=0}^M \in \mathfrak{P}_f(E) \right\}; \qquad (2.6)$$

we set $\mathrm{Var}_{\mathsf{d}}(u, \emptyset) := 0$.

If $\mathrm{Var}_{\mathsf{d}}(u, E) < \infty$ then u belongs to the space $\mathrm{BV}_{\mathsf{d}}(E, X)$ of function with *bounded variation* and we can define the function

$$V_u(t) := \mathrm{Var}_{\mathsf{d}}(u, E \cap [E^-, t]), \quad t \in [E^-, E^+], \qquad (2.7)$$

which is monotone non decreasing and satisfies

$$\mathsf{d}(u(t_0), u(t_1)) \le \mathrm{Var}_{\mathsf{d}}(u, [t_0, t_1]) = V_u(t_1) - V_u(t_0) \quad \text{for every } t_0, t_1 \in E,\ t_0 \le t_1. \qquad (2.8)$$

When d is a distance and (X, d) is a complete metric space, it is well known that every function u with bounded variation is *regulated*, i.e. it admits left and right d-limits at every time t and the jump set coincides with the jump set of V_u. In our weaker framework, regulated functions should also take into account the σ topology (which could also be non-metrizable). We propose the following definition.

Definition 2.3 ((σ, d)-regulated functions) *We say that $u : [a, b] \to X$ is (σ, d)-regulated if for every $t \in [a, b]$ there exist the left and right limits $u(t\pm)$ (here we adopt the convention*

$u(a-) := u(a)$ and $u(b+) := u(b)$) w.r.t. the σ topology, satisfying

$$
\begin{aligned}
u(t-) &= \lim_{s\uparrow t} u(s), \quad \lim_{s\uparrow t} \mathsf{d}(u(s), u(t-)) = 0, \\
u(t+) &= \lim_{s\downarrow t} u(s), \quad \lim_{s\downarrow t} \mathsf{d}(u(t+), u(s)) = 0,
\end{aligned}
\tag{2.9}
$$

and

$$
\mathsf{d}(u(t-), u(t)) = 0 \;\Rightarrow\; u(t-) = u(t); \qquad \mathsf{d}(u(t), u(t+)) = 0 \;\Rightarrow\; u(t) = u(t+). \tag{2.10}
$$

The pointwise jump set $\mathrm{J}_u^{\pm}$ of a (σ, d)-regulated function u is defined by

$$
\mathrm{J}_u^- := \{t \in [a,b] : u(t-) \neq u(t)\}, \quad \mathrm{J}_u^+ := \{t \in [a,b] : u(t) \neq u(t+)\}, \quad \mathrm{J}_u := \mathrm{J}_u^- \cup \mathrm{J}_u^+. \tag{2.11}
$$

We will denote by $\mathrm{BV}_{\sigma,\mathsf{d}}([a,b]; X)$ the space of (σ, d)-regulated functions with finite d-total variation. In this case $\mathrm{J}_u^{\pm}$ coincide with the corresponding jump sets $\mathrm{J}_{V_u}^{\pm}$ of the real monotone function V_u. In particular, $\mathrm{J}_u = \mathrm{J}_u^- \cup \mathrm{J}_u^+$ is at most countable.

Notice that for a monotone function $V : [a,b] \to \mathbb{R}$ the jump set J_V coincides with $\{t \in [a,b] : V(t-) \neq V(t+)\}$.

The following simple lemma, that lies behind [22, Assumption (A4), Theorems 3.2, 3.3] and [30, Section 2.2], provides a sufficient condition for a function $u \in \mathrm{BV}_{\mathsf{d}}(D; X)$, D being a dense subset of $[a,b]$, to admit a unique (σ, d) regulated extension to $[a,b]$; when $D = [a,b]$ it still provides interesting σ-continuity properties of u.

Lemma 2.4 *Let D be a dense subset of $[a,b]$, let u be a curve in $\mathrm{BV}_{\mathsf{d}}(D; X)$ with $\mathrm{J}_{V_u} \subset D$, and let $\mathscr{U} \subset [a,b] \times X$ a sequentially compact set separated by $\mathsf{d}_{\mathbb{R}}$ such that $u(t) \in \mathscr{U}(t)$ for every $t \in D \setminus \mathrm{J}_{V_u}$.*

If $\mathscr{U}$ is sequentially compact and $\mathsf{d}_{\mathbb{R}}$ separates its points then u admits a unique extension $\tilde{u}$ to a function in $\mathrm{BV}_{\sigma,\mathsf{d}}([a,b]; X)$, with $\mathrm{Var}_{\mathsf{d}}(u, D) = \mathrm{Var}_{\mathsf{d}}(\tilde{u}, [a,b])$. In particular, when $D = [a,b]$ we get $u = \tilde{u} \in \mathrm{BV}_{\sigma,\mathsf{d}}([a,b]; X)$.

Proof. We fix $t \in (a,b]$ and apply Lemma 2.1 ii), with $Z := D \cap [a,t)$, $z_0 := t$, observing that for $r, s \in Z$ with $r \leq s$

$$
\mathsf{d}(u(r), u(s)) \leq V_u(s) - V_u(r), \quad \lim_{s,r\uparrow t} V_u(s) - V_u(r) = 0. \tag{2.12}
$$

Once the existence of the limit has been established, the previous estimate and the lower semicontinuity of d show that $\lim_{s\uparrow t} \mathsf{d}(u(s), u) = 0$. The argument for the existence of the right limit is completely analogous. If $t \in [a,b] \setminus D$, we can extend u by setting $\tilde{u}(t) := \lim_{s\to t,\, s\in D} u(s)$, since $t \notin \mathrm{J}_{V_u}$. It is easy to check that $\tilde{u} \in \mathrm{BV}_{\mathsf{d}}([a,b]; X)$, $V_{\tilde{u}} = V_u$, $\tilde{u}(t\pm) = u(t\pm)$, and $\mathrm{J}_{\tilde{u}} \subset D$.

If $t \notin \mathrm{J}_{\tilde{u}}^-$ then the above argument shows $\lim_{s\uparrow t} \mathsf{d}(\tilde{u}(s), \tilde{u}(t)) = 0$ so that $\lim_{s\uparrow t} V_{\tilde{u}}(s) = V_{\tilde{u}}(t)$ and t is a left continuity point for V_u. On the other hand, if $t \in \mathrm{J}_{\tilde{u}}^-$, $\tilde{u}(t-), \tilde{u}(t) = u(t) \in \mathscr{U}(t)$, the separation property yields $\mathsf{d}(\tilde{u}(t-), \tilde{u}(t)) > 0$ and

$$
\mathsf{d}(\tilde{u}(t-), \tilde{u}(t)) \leq \liminf_{s\uparrow t} \mathsf{d}(\tilde{u}(s), \tilde{u}(t)) \leq \liminf_{s\uparrow t} V_{\tilde{u}}(t) - V_{\tilde{u}}(s) = V_{\tilde{u}}(t) - V_{\tilde{u}}(t-)
$$

so that $t \in \mathrm{J}_{V_u}^-$.

Finally, in order to show that $\mathrm{Var}_{\mathsf{d}}(u, D) = \mathrm{Var}_{\mathsf{d}}(\tilde{u}, [a, b])$, we consider a finite subset $P = \{t_j\}_{j=0}^N$ of $[a, b]$ with $t_j \le t_{j+1}$ and we fix a small $\varepsilon > 0$. Since $\mathrm{J}_{\tilde{u}} = \mathrm{J}_u \subset D$, every $t_j \in P$ can be approximated by points $t_j^\pm \in D$ such that $t_{j-1} < t_j^- \le t_j \le t_j^+ < t_{j+1}$ and $\mathsf{d}(\tilde{u}(t_j^-), \tilde{u}(t_j)) + \mathsf{d}(\tilde{u}(t_j), \tilde{u}(t_j^+)) \le \varepsilon/(N+1)$ (we just set $t_j^\pm := t_j$ whenever $t_j \in P \cap D$). Thus we have a new partition (with possible repetitions) $t_0^- \le t_0^+ < t_1^- \le t_1^+, \cdots, < t_N^- \le t_N^+$ in D and

$$\begin{aligned}\sum_{j=1}^N \mathsf{d}(\tilde{u}(t_{j-1}), \tilde{u}(t_j)) &\le \sum_{j=0}^N \mathsf{d}(\tilde{u}(t_j^-), \tilde{u}(t_j)) + \mathsf{d}(\tilde{u}(t_j), \tilde{u}(t_j^+)) + \sum_{j=1}^N \mathsf{d}(\tilde{u}(t_{j-1}^+), \tilde{u}(t_j^-)) \\ &\le \varepsilon + \sum_{j=0}^N \mathsf{d}(\tilde{u}(t_j^-), \tilde{u}(t_j^+)) + \sum_{j=1}^N \mathsf{d}(\tilde{u}(t_{j-1}^+), \tilde{u}(t_j^-)) \le \varepsilon + \mathrm{Var}_{\mathsf{d}}(u, D).\end{aligned}$$

Since $\varepsilon > 0$ and P are arbitrary we conclude. □

Augmented total variation associated with a transition cost. In some cases, such as for Balanced Viscosity or Visco-Energetic solutions, we will need a modified notion of total variation, increased by a further contribution along the jumps of the function.

Such a contribution can be described by a function $\mathsf{e} : [0, T] \times X \times X \to [0, +\infty]$ satisfying

$$\Delta_{\mathsf{e}}(t, u_-, u_+) = \mathsf{e}(t, u_-, u_+) - \mathsf{d}(u_-, u_+) \ge 0 \quad \text{for every } t \in [0, T],\ u_\pm \in X. \tag{2.13}$$

We will also use the notation $\Delta_{\mathsf{e}}(t, u_-, u, u_+)$

$$\Delta_{\mathsf{e}}(t, u_-, u, u_+) := \Delta_{\mathsf{e}}(t, u_-, u) + \Delta_{\mathsf{e}}(t, u, u_+). \tag{2.14}$$

Definition 2.5 (Augmented total variation) *Let* $\mathsf{e}, \Delta_{\mathsf{e}}$ *be as in* (2.13). *For every* (σ, d)-*regulated curve* $u \in \mathrm{BV}_{\mathsf{d}}([0, T]; X)$ *and every subinterval* $[a, b] \subset [0, T]$, *the* incremental jump variation *of* u *on* $[a, b]$ *induced by* Δ_{e} *is*

$$\begin{aligned}\mathrm{Jmp}_{\Delta_{\mathsf{e}}}(u, [a, b]) := \Delta_{\mathsf{e}}(a, u(a), u(a+)) + \Delta_{\mathsf{e}}(b, u(b-), u(b)) \\ + \sum_{t \in \mathrm{J_u} \cap (a,b)} \Delta_{\mathsf{e}}(t, u(t-), u(t), u(t+)),\end{aligned} \tag{2.15}$$

and the corresponding augmented total variation *is*

$$\mathrm{Var}_{\mathsf{d},\mathsf{e}}(u, [a, b]) := \mathrm{Var}_{\mathsf{d}}(u, [a, b]) + \mathrm{Jmp}_{\Delta_{\mathsf{e}}}(u, [a, b]). \tag{2.16}$$

Notice that $Var_{\mathsf{d},\mathsf{e}}(u, [a, b]) \ge \mathrm{Var}_{\mathsf{d}}(u, [a, b])$ and they coincide when $\mathrm{J_u} = \emptyset$ or when $\mathsf{e} = \mathsf{d}$. As for the d-total variation, $Var_{\mathsf{d},\mathsf{e}}$ satisfies the additive property

$$Var_{\mathsf{d},\mathsf{e}}(u, [a, b]) + Var_{\mathsf{d},\mathsf{e}}(u, [b, c]) = Var_{\mathsf{d},\mathsf{e}}(u, [a, c]) \quad \text{whenever } a \le b \le c. \tag{2.17}$$

2.2 The Energy functional

In this section we briefly recall one of the possible settings for energetic solutions to a rate-independent system (R.I.S.) $(X, \mathcal{E}, \mathsf{d})$, following the approach of [22]. Besides the asymmetric dissipation distance d we have introduced in the previous section, variationally driven rate-independent evolutions are characterized by a time-dependent *energy functional* $\mathcal{E} : [0,T] \times X \to \mathbb{R}$. A few basic properties will also involve the perturbed functionals

$$\mathcal{F}(t,x) := \mathcal{E}(t,x) + \mathsf{d}(x_o, x), \quad \mathcal{F}_0(x) := \mathcal{F}(0,x), \quad x \in X \tag{2.18}$$

and the collection of their sublevels $\big\{(t,x) \in [0,T] \times X : \mathcal{F}(t,x) \le C\big\}$ in $[0,T] \times X$. We will always make the following standard assumptions [22, 44, 33], where we also allow some flexibility in the choice of the power $\mathcal{P}$ (see [16, 18, 29, 31].

Assumption $\langle$A$\rangle$ *The R.I.S.* $(X, \mathcal{E}, \mathsf{d})$ *satisfy*

$\langle$**A.1**$\rangle$ **Lower semicontinuity and compactness.** $\mathcal{E}$ *is* σ*-l.s.c. on all the sublevels of* $\mathcal{F}$*, which are* $\sigma_{\mathbb{R}}$*-sequentially compact in* $[0,T] \times X$.

$\langle$**A.2**$\rangle$ **Power-control.** *There exists a map* $\mathcal{P} : [0,T] \times X \to \mathbb{R}$ *(the "time superdifferential" of the energy) upper semicontinuous on the sublevels of* $\mathcal{F}$ *satisfying*

$$\liminf_{s\uparrow t} \frac{\mathcal{E}(t,x) - \mathcal{E}(s,x)}{t-s} \ge \mathcal{P}(t,x) \ge \limsup_{s\downarrow t} \frac{\mathcal{E}(s,x) - \mathcal{E}(t,x)}{s-t} \tag{2.19}$$

$$|\mathcal{P}(t,x)| \le C_P \mathcal{F}(t,x), \tag{2.20}$$

for a constant $C_P > 0$ *and for every* $(t,x) \in [0,T] \times X$.

Notice that whenever $t \mapsto \mathcal{E}(t,x)$ is differentiable at some $t_0 \in [0,T]$ (2.19) yields $\mathcal{P}(t_0,x) = \partial_t \mathcal{E}(t_0,x)$. On the other hand, (2.19) shows that $t \to \mathcal{E}(t,x)$ is upper semicontinuous (and thus continuous by $\langle$A.1$\rangle$ and bounded) on $[0,T]$. It follows that $t \mapsto \mathcal{F}(t,x)$ is bounded so that there exists a suitable constant C providing $|\mathcal{P}(t,x)| \le C$ for every $t \in [0,T]$.

(2.19) then shows that $t \mapsto \mathcal{E}(t,x)$ is Lipschitz continuous differentiable a.e.; by estimating $\mathcal{E}(t_1,x) = \mathcal{E}(t_0,x) + \int_{t_0}^{t_1} \mathcal{P}(s,x)\,\mathrm{d}s$ with (2.20) and applying Gronwall's lemma, we obtain

$$\mathcal{F}(t_1,x) \le \mathcal{F}(t_0,x) \exp(C_P|t_1 - t_0|) \quad \text{for every } t_0, t_1 \in [0,T]. \tag{2.21}$$

This estimate will be the basis for the a priori estimate of the stored and the dissipated energies. In particular it implies that for every $t \in [0,T]$ and for every $y \in X$ the map

$$x \mapsto \mathcal{E}(t,x) + \mathsf{d}(y,x) \quad \text{has bounded sequentially compact sublevels in } (X,\sigma). \tag{2.22}$$

Moreover, it would not be difficult to check that $\langle$A.2$\rangle$ yields

$$\begin{aligned}
\mathcal{P}(t,x) &= \frac{\partial^-}{\partial t}\mathcal{E}(t,x) = \lim_{s\uparrow t} \frac{\mathcal{E}(t,x) - \mathcal{E}(s,x)}{t-s} && \text{for every } t \in [0,T],\ x \in X,\\
\mathcal{P}(t,x) &= \lim_{s\to t} \frac{\mathcal{E}(t,x) - \mathcal{E}(s,x)}{t-s} && \text{for a.e. } t \in [0,T], \text{ for every } x \in X.
\end{aligned}$$

Remark 2.6 (Left continuity of d and upper semicontinuity of $\mathcal{P}$) The upper semicontinuity condition of $\mathcal{P}$ stated in $\langle$A.2$\rangle$ could be relaxed if we know more properties on d (and on the viscous correction δ, see $\langle$B$\rangle$). An example is provided by this condition, that can occasionally replace $\langle$A.2$\rangle$:

$\langle$**A.2'**$\rangle$ d *is left-continuous on the sublevels of* $\mathcal{F}_0$*, i.e.*

$$\mathcal{F}_0(x_n) \le C, \quad x_n \stackrel{\sigma}{\to} x \quad \Rightarrow \quad \mathsf{d}(x_n, v) \to \mathsf{d}(x, v) \tag{2.23}$$

and the map $\mathcal{P} : [0, T] \times X \to \mathbb{R}$ *satisfies* (2.19), (2.20) *and the conditional upper-semicontinuity*

$$(t_n, x_n) \stackrel{\sigma}{\to} (t, x), \quad \mathcal{E}(t_n, x_n) \to \mathcal{E}(t, x) \quad \Rightarrow \quad \limsup_{n\uparrow\infty} \mathcal{P}(t_n, x_n) \le \mathcal{P}(t, x). \tag{2.24}$$

Remark 2.7 (Extended-valued energies and distances) Our setting is equivalent to considering an energy functional $\widetilde{\mathcal{E}} : [0, T] \times \widetilde{X} \to \mathbb{R} \cup \{+\infty\}$ possibly assuming the value $+\infty$, since any reasonable formulation of $\langle$A.2$\rangle$ yields that the proper domain $\mathrm{D}(\widetilde{\mathcal{E}}(t, \cdot)) := \{u \in \widetilde{X} : \widetilde{\mathcal{E}}(t, u) < +\infty\}$ should be independent of time, thanks to (2.21). Also the assumption $\mathsf{d}(x_o, u) < \infty$ is not restrictive. In fact, it is sufficient to choose x_o as the initial datum $\bar{u}$ of the evolution problem and consider the restriction of $\tilde{\mathcal{E}}$ and $\tilde{\mathsf{d}}$ to the set $X := \{v \in \mathrm{D}(\widetilde{\mathcal{E}}(0, \cdot)) : \widetilde{\mathsf{d}}(x_o, v) < \infty\}$.

2.3 Energetic solutions to rate-independent problems

Hereafter we recall the notion of *energetic solution* (see [37], [35]) to the Rate-Independent System (R.I.S.) $(X, \mathcal{E}, \mathsf{d})$. Let us first introduce the notion of the d-stable set $\mathscr{S} \subset [0, T] \times X$ associated with $\mathcal{E}$

$$\mathscr{S}_{\mathsf{d}} := \Big\{(t, u) \in [0, T] \times X : \mathcal{E}(t, u) \le \mathcal{E}(t, v) + \mathsf{d}(u, v) \quad \text{for every } v \in X\Big\} \tag{2.25}$$

with its time-dependent sections $\mathscr{S}_{\mathsf{d}}(t) := \{u \in X : (t, u) \in \mathscr{S}_{\mathsf{d}}\}$.

Definition 2.8 (Energetic solutions) *A curve* $u \in \mathrm{BV}_{\mathsf{d}}([0, T]; X)$ *is an energetic solution of the R.I.S.* $(X, \mathcal{E}, \mathsf{d})$ *if for all* $t \in [0, T]$ *it satisfies the global stability condition*

$$u(t) \in \mathscr{S}_{\mathsf{d}}(t), \quad \textit{i.e.} \quad \mathcal{E}(t, u(t)) \le \mathcal{E}(t, v) + \mathsf{d}(u(t), v) \qquad \textit{for every } v \in X, \tag{S$_\mathsf{d}$}$$

and the energetic balance

$$\mathcal{E}(t, u(t)) + \mathrm{Var}_{\mathsf{d}}(u, [0, t]) = \mathcal{E}(0, u(0)) + \int_0^t \mathcal{P}(s, u(s))\mathrm{d}s. \tag{E$_\mathsf{d}$}$$

The Existence of energetic solutions in such a general framework is one of the main results of MAINIK-MIELKE: it requires the *closedness* of the stable set $\mathscr{S}_{\mathsf{d}}$ and a non-degeneracy of d in each sections $\mathscr{S}_{\mathsf{d}}(t)$ [22, Thm. 4.5], two conditions which are always satisfied in the simpler metric setting of Remark 2.2.

Theorem 2.9 (Existence of Energetic solutions) *Let us assume that*

$$\mathscr{S}_{\mathsf{d}} \textit{ is } \sigma\textit{-closed in } [0, T] \times X \textit{ and separated by } \mathsf{d}. \tag{2.26}$$

Then for every $\bar{u} \in \mathscr{S}_{\mathsf{d}}(0)$ *there exists at least one energetic solution to the R.I.S.* $(X, \mathcal{E}, \mathsf{d})$ *satisfying* $u(0) = \bar{u}$.

One of the main features of the definition above concerns the jump behaviour of a solution: assuming (2.26), every energetic solution u is (σ, d)-regulated and satisfies the following *jump conditions* at every jump point $t \in \mathrm{J_u}$:

$$\begin{aligned} \mathcal{E}(t, u(t-)) - \mathcal{E}(t, u(t)) &= \mathsf{d}(u(t-), u(t)), \\ \mathcal{E}(t, u(t)) - \mathcal{E}(t, u(t+)) &= \mathsf{d}(u(t), u(t+)), \\ \mathcal{E}(t, u(t-)) - \mathcal{E}(t, u(t+)) &= \mathsf{d}(u(t-), u(t+)). \end{aligned} \tag{2.27}$$

The time incremental minimization scheme. The most powerful method to construct energetic solutions to the R.I.S. $(X, \mathcal{E}, \mathsf{d})$ and to prove their existence is provided by the time incremental minimization scheme.

We consider ordered finite partitions $\tau \subset [0, T]$ whose points will be denoted by t^n_τ for integers n between 0 and $N = N(\tau)$, $0 = t^0_\tau < t^1_\tau < \ldots < t^{N-1}_\tau < t^N_\tau = T$, and we set $|\tau| := \max_n t^n_\tau - t^{n-1}_\tau$. In order to find good approximations U^n_τ of $u(t^n_\tau)$ we choose an initial value $U^0_\tau \approx u_0$ and solve the time incremental minimization scheme

$$U^n_\tau \in \operatorname*{argmin}_{U \in X} \Big\{ \mathsf{d}(U^{n-1}_\tau, U) + \mathcal{E}(t^n_\tau, U) \Big\}. \tag{IM$_\mathsf{d}$}$$

Setting

$$\overline{U}_\tau(t) := U^n_\tau \quad \text{if } t \in (t^{n-1}_\tau, t^n_\tau], \tag{2.28}$$

it is possible to find a sequence of partitions τ_k with $|\tau_k| \downarrow 0$ such that

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

and u is an energetic solution starting from u_0.

2.4 Viscosity approximation and Balanced Viscosity (BV) solutions

A different approach to solve *rate-independent* problems is to use a viscous approximation of the dissipation distance. For the sake of simplicity, we present this approach in the metric setting of Remark 2.2 with uniform partitions (i.e. $t^n_\tau = n|\tau|$), starting from the viscous regularization of the incremental minimization scheme (IM$_\mathsf{d}$) by a quadratic perturbation generated by the same distance d, i.e. a term

$$\delta_\tau(u, v) := \frac{1}{2}\mu(|\tau|)\mathsf{d}^2(u, v), \quad u, v \in X, \quad \mu : (0, \infty) \to (0, \infty), \quad \lim_{r \downarrow 0} \mu(r) = +\infty. \tag{2.29}$$

The viscous incremental problem is therefore to find $U^1_\tau, \ldots, U^N_\tau$ such that

$$U^n_\tau \in \operatorname*{argmin}_{U \in X} \Big\{ \mathsf{d}(U^{n-1}_\tau, U) + \delta_\tau(U^{n-1}_\tau, U) + \mathcal{E}(t^n, U) \Big\}. \tag{IM$_{\mathsf{d},\delta_\tau}$}$$

Setting as in (2.28) $\overline{U}_\tau(t) := U^n_\tau$ if $t \in (t^{n-1}_\tau, t^n_\tau]$, we can study the limit of the discrete solutions when $|\tau| \downarrow 0$.

The scaling of the factor $\mu(\tau)$ in (2.29) can be justified by observing that when $X = \mathbb{R}^d$, $\mathsf{d}(u, v) := |u - v|$ and $\mathcal{E}$ is a C^1 function, (IM$_{\mathsf{d},\delta_\tau}$) naturally arises as the implicit discretization of the differential inclusion

$$\operatorname{sign}(u'(t)) + \varepsilon u'(t) + \mathrm{D}_u \mathcal{E}(t, u(t)) \ni 0, \quad \operatorname{sign}(v) := \begin{cases} v/|v| & \text{if } v \neq 0, \\ \{w \in \mathbb{R}^d : |w| \leq 1\} & \text{if } v = 0, \end{cases} \tag{2.30}$$

with the choice $\mu(|\tau|) := \varepsilon/|\tau|$. Therefore, if $\lim_{|\tau|\downarrow 0}\mu(|\tau|) = +\infty$ one can heuristically expect that the limits of discrete solutions $\overline{U}_\tau$ to the incremental minimization problems $\mathrm{IM}_{\mathsf{d},\delta_\tau}$ coincide with the limit trajectories of (2.30) as $\varepsilon \downarrow 0$.

Under quite general assumptions on $\mathcal{E}$ it is possible to prove that limit solutions u of suitable subsequences of $\overline{U}_\tau$ satisfy a local stability condition, which replaces the global one (S_d), and a modified energy balance defined in terms of an augmented total variation $\mathit{Var}_{\mathsf{d},\mathsf{v}}(u,[0,T]) \geq \mathrm{Var}_\mathsf{d}(u,[0,T])$ as in Definition 2.5. Both involve the *metric slope* of $\mathcal{E}$, defined as

$$|\mathrm{D}\mathcal{E}|(t,u) := \limsup_{v\to u} \frac{(\mathcal{E}(t,u)-\mathcal{E}(t,v))_+}{\mathsf{d}(u,v)}. \tag{2.31}$$

$\mathit{Var}_{\mathsf{d},\mathsf{v}}$ is associated with the *minimal transition cost* between u_0 and $u_1 \in X$ at the time t:

$$\begin{aligned}\mathsf{v}(t,u_0,u_1) := \inf\Big\{ \int_{r_0}^{r_1} |\dot\theta|(r)\,(|\mathrm{D}\mathcal{E}|(t,\theta(r)) \vee 1)\,\mathrm{d}r : \\ \theta \in \mathrm{AC}([r_0,r_1];X,\mathsf{d}),\ \theta(r_0)=u_0,\ \theta(r_1)=u_1\Big\},\end{aligned} \tag{2.32}$$

where $|\dot\theta|$ denotes the *metric derivative* of θ, see [2, Sect. 1] and [44, Sect. 2.1] in the asymmetric case; clearly $\mathsf{v}(t,u_0,u_1) \geq \mathsf{d}(u_0,u_1)$. Based on (2.16), we can now specify the concept of *Balanced Viscosity* (BV) *solution* to the rate-independent system $(X,\mathcal{E},\mathsf{d})$.

Definition 2.10 (Balanced Viscosity (BV) solutions) *A curve* $u \in \mathrm{BV}_\mathsf{d}([0,T];X)$ *is a* BV *solution of the rate-independent system* $(X,\mathcal{E},\mathsf{d})$ *with the viscous dissipation* (2.29) *if it satisfies the local stability*

$$|\mathrm{D}\mathcal{E}|(t,u(t)) \leq 1 \quad \textit{for every} \quad t \in [0,T]\setminus \mathrm{J_u}, \tag{$\mathrm{S}_{\mathsf{d},\mathrm{loc}}$}$$

and the energy balance

$$\mathcal{E}(t,u(t)) + \mathit{Var}_{\mathsf{d},\mathsf{v}}(u,[0,t]) = \mathcal{E}(0,u(0)) + \int_0^t \mathcal{P}(s,u(s))\mathrm{d}s \quad \textit{for all } t\in[0,T]. \tag{$\mathrm{E}_{\mathsf{d},\mathsf{v}}$}$$

The viscous total variation induced by v in the energy balance ($\mathrm{E}_{\mathsf{d},\mathsf{v}}$) (instead of the canonical one induced by the distance d) compensates for the lack of information in the local stability condition ($\mathrm{S}_{\mathsf{d},\mathrm{loc}}$). In particular, it is possible to prove that a curve $u \in \mathrm{BV}_\mathsf{d}([0,T];X)$ is a BV solution of the rate-independent system $(X,\mathcal{E},\mathsf{d})$ if and only if it satisfies the local stability condition ($\mathrm{S}_{\mathsf{d},\mathrm{loc}}$), the localized energy dissipation inequality

$$\mathcal{E}(t,u(t)) + \mathrm{Var}_\mathsf{d}(u,[s,t]) \leq \mathcal{E}(s,u(s)) + \int_s^t \mathcal{P}(r,u(r))\,\mathrm{d}r \quad \text{for every } 0\leq s\leq t\leq T$$

and the following jump conditions at each point $t \in \mathrm{J_u}$ (see [30, Theorem 3.13]):

$$\begin{aligned}\mathcal{E}(t,u(t-)) - \mathcal{E}(t,u(t)) &= \mathsf{v}(t,u(t-),u(t)),\\ \mathcal{E}(t,u(t)) - \mathcal{E}(t,u(t+)) &= \mathsf{v}(t,u(t),u(t+)),\\ \mathcal{E}(t,u(t-)) - \mathcal{E}(t,u(t+)) &= \mathsf{v}(t,u(t-),u(t+)).\end{aligned} \tag{2.33}$$

3 Visco-Energetic (VE) solutions

As we have seen in section 2, the choice $\mu(\tau) \equiv 0$ in the incremental minimization problem ($\mathrm{IM}_{\mathsf{d},\delta_\tau}$) corresponds to ($\mathrm{IM}_{\mathsf{d}}$) and leads to the notion of *Energetic solutions*, while the case when $\mu(\tau) \uparrow +\infty$ as $\tau \downarrow 0$ corresponds to *Balanced Viscosity solutions.* In the present paper, we want to study the asymptotic behaviour of the incremental minimization scheme in the case when $\mu(\tau) \equiv \mu$ is a constant, and to find an appropriate variational characterization for the corresponding limit trajectories.

3.1 Viscous correction of the incremental minimization scheme

In order to cover a wide spectrum of possible applications with the greatest flexibility, we are considering here general viscous corrections modeled by a lower semicontinuous map

$$\delta : X \times X \to [0, +\infty] \quad \text{with } \delta(x, x) = 0 \text{ for every } x \in X, \tag{3.1}$$

and the corresponding modified dissipation

$$\mathsf{D}(x, y) := \mathsf{d}(x, y) + \delta(x, y) \qquad \text{for all } x, y \in X. \tag{3.2}$$

As in the previous section, d and $\mathcal{E}$ will be a dissipation distance and a time-dependent energy functional satisfying Assumptions $\langle \mathrm{A} \rangle$ in the metric-topological setting introduced in Section 2.1. Our starting point is the following modified variational scheme.

Definition 3.1 (The viscous incremental minimization scheme.) *Starting from $U^0_\tau \in X$, find recursively $U^1_\tau, \dots, U^N_\tau$ such that U^n_τ minimizes*

$$U \mapsto \mathsf{D}(U^{n-1}_\tau, U) + \mathcal{E}(t_n, U) = \mathsf{d}(U^{n-1}_\tau, U) + \delta(U^{n-1}_\tau, U) + \mathcal{E}(t^n, U). \tag{$\mathrm{IM}_{\mathsf{d},\delta}$}$$

Since d and δ are lower semicontinuous, the existence of a minimizer for the problem ($\mathrm{IM}_{\mathsf{d},\delta}$) follows from condition $\langle \mathrm{A.1} \rangle$. Of course, not every continuous function δ will provide an admissible viscous correction; the trivial example $\delta = \mathsf{d}$ (which doubles the dissipation distance) shows that we should impose some sufficiently strong vanishing condition of $\delta(x, y)$ in the neighborhoods of points where $\mathsf{d}(x, y) = 0$ in $X \times X$.

A quite general admissibility criterion is related to the notion of *global* D*-stability*, which can be easily imagined from the corresponding property (S_{d}) of the energetic case. We will also introduce the weaker notion of quasi-stability, which will turn out to be very useful later on.

Definition 3.2 (Quasi D-stability and D-stable set) *Let $Q \geq 0$; we say that $(t, x) \in [0, T] \times X$ is a (D, Q)-quasi-stable point if it satisfies*

$$\mathcal{E}(t, x) \leq \mathcal{E}(t, y) + \mathsf{D}(x, y) + Q \qquad \textit{for every } y \in X. \tag{3.3}$$

In the case $Q = 0$, i.e. when

$$\mathcal{E}(t, x) \leq \mathcal{E}(t, y) + \mathsf{D}(x, y) \qquad \textit{for every } y \in X, \tag{3.4}$$

we say that (t, x) is D-stable. We call $\mathscr{S}_{\mathsf{D}}$ the stable set (i.e. the collection of all the D-stable points) and $\mathscr{S}_{\mathsf{D}}(t)$ its section at time t.

As in the case of energetic solutions, we expect that the D-stability condition will play a crucial role. A first important point concerns the admissibility criterion for the viscous correction δ: in addition to basic compatibility properties between δ and d, we will essentially require that D-stable points satisfy a sort of local d-stability. To better understand the next condition, let us first notice that (3.4) can be equivalently formulated as

$$\sup_{y\neq x} \frac{\mathcal{E}(t,x)-\mathcal{E}(t,y)}{\mathsf{D}(x,y)} \le 1. \tag{3.5}$$

Assumption $\langle$B$\rangle$ (Admissible viscous corrections) *An admissible viscous correction $\delta : X\times X \to [0,+\infty]$ for the R.I.S. $(X,\mathcal{E},\mathsf{d})$ satisfies the following conditions:*

$\langle$B.1$\rangle$ **d-compatibility.** *For every $x,y,z\in X$*

$$\mathsf{d}(x,y)=0 \quad\Rightarrow\quad \delta(z,y)\le\delta(z,x) \quad \textit{and} \quad \delta(x,z)\le\delta(y,z). \tag{3.6}$$

$\langle$B.2$\rangle$ **Left d-continuity.** *For every sequence x_n and every $x\in X$ we have*

$$\sup_n \mathcal{F}_0(x_n)<\infty,\quad x_n\stackrel{\sigma}{\to} x,\quad \mathsf{d}(x_n,x)\to 0 \quad\Rightarrow\quad \lim_{n\to\infty}\delta(x_n,x)=0. \tag{3.7}$$

$\langle$B.3$\rangle$ **D-stability yields local d-stability.** *For every $(t,x)\in\mathscr{S}_{\mathsf{D}}$, $M>1$ there exists $\eta>0$ and a neighborhood U of x in X such that*

$$\mathcal{E}(s,y)\le\mathcal{E}(s,x)+M\mathsf{d}(v,x) \quad \textit{for every } (s,y)\in\mathscr{S}_{\mathsf{D}},\ s\in(t-\eta,t],\ y\in U,\ \mathsf{d}(y,x)\le\eta. \tag{3.8}$$

Equivalently,

$$\limsup_{\substack{(s,y)\to(t,x),\ \mathsf{d}(y,x)\to 0\\ (s,y)\in\mathscr{S}_{\mathsf{D}},\ s\le t}} \frac{\mathcal{E}(s,y)-\mathcal{E}(s,x)}{\mathsf{d}(y,x)}\le 1. \tag{3.9}$$

$\langle$B.1$\rangle$ is a minimal compatibility condition between d and δ. Notice that $\langle$B.1$\rangle$ is trivially satisfied by any monotonically increasing function of d or if d separates the points of X, e.g. in the simpler metric setting of Remark 2.2.

Let us now see an important example of admissible viscous corrections in which assumption $\langle$B$\rangle$ is satisfied. A further example will be discussed in Section 4.3.

Example 3.3 If $\delta : X\times X\to[0,+\infty)$ satisfies

$$\lim_{\substack{y\to x\\ \mathsf{d}(y,x)\to 0}} \frac{\delta(y,x)}{\mathsf{d}(y,x)}=0 \quad \text{for every } x\in\mathscr{S}_{\mathsf{D}}(t),\quad t\in[0,T], \tag{3.10}$$

then it is immediate to check that D satisfies $\langle$B.3$\rangle$. In particular, any function of the form

$$\delta(x,y)=h(\mathsf{d}(x,y)) \quad \text{for a nondecreasing } h\in \mathrm{C}([0,\infty)) \text{ with } \quad \lim_{r\downarrow 0} h(r)/r=0 \tag{3.11}$$

provides an admissible correction satisfying $\langle$B$\rangle$. A typical choice is the quadratic correction

$$\delta(x,y):=\frac{\mu}{2}\mathsf{d}^2(x,y),\quad \mu>0. \tag{3.12}$$

3.2 Transition costs and augmented total variation.

Let us first notice that the quasi-stability condition (3.3) (and therefore the stability condition (3.4)) can be equivalently characterized through a sort of *residual function* that we introduce in the definition below.

Definition 3.4 *For every $t \in [0,T]$ and $x \in X$ the residual stability function is defined by*

$$\mathcal{R}(t,x) : = \sup_{y\in X}\{\mathcal{E}(t,x) - \mathcal{E}(t,y) - \mathsf{D}(x,y)\} \tag{3.13}$$

$$= \mathcal{E}(t,x) - \inf_{y\in X}\{\mathcal{E}(t,y) + \mathsf{D}(x,y)\}; \tag{3.14}$$

$\mathcal{R}(t,x)$ provides the minimal constant $Q \geq 0$ such that (t,x) is (D,Q)-quasi-stable:

$$\mathcal{R}(t,x) := \min\Big\{Q \geq 0 : \mathcal{E}(t,x) \leq \mathcal{E}(t,y) + \mathsf{D}(x,y) + Q \quad \text{for every } y \in X\Big\}. \tag{3.15}$$

By choosing $y := x$ in (3.13) we can immediately check that the residual function $\mathcal{R}$ is non-negative, i.e.

$$\mathcal{R}(t,x) \geq 0 \quad \text{for all } t \in [0,T],\ x \in X.$$

$\mathcal{R}$ provides a measure of the failure of the stability condition (3.4), since for every $x \in X$, $t \in [0,T]$ we get

$$\mathcal{E}(t,x) \leq \mathcal{E}(t,y) + \mathsf{D}(x,y) + \mathcal{R}(t,x) \tag{3.16}$$

and

$$\mathcal{R}(t,x) = 0 \quad \Longleftrightarrow \quad x \in \mathscr{S}_{\mathsf{D}}(t). \tag{3.17}$$

Notice that when D is σ-continuous then

$$\mathcal{R} \text{ is } \sigma_{\mathbb{R}}\text{-lower semicontinuous} \tag{3.18}$$

or, equivalently,

$$\text{for every } Q \geq 0 \text{ the } (\mathsf{D},Q)\text{-quasi-stable set is } \sigma\text{-closed.} \tag{3.19}$$

We will see that this property will play a crucial role in our general setting and corresponds to the closedness property of the stable set (2.26) in the energetic framework.

As in the case of Balanced Viscosity solutions, we expect that the jumps of a limit trajectory of the viscous incremental minimization scheme 3.1 can be characterized by a class of curves minimizing a suitable transition cost.

The main novelty here is represented by the fact that such curves are parametrized by continuous maps $\vartheta : E \to X$, defined on a compact subset E of $\mathbb{R}$, which in general may have a more complicated structure than an interval. We will also require that ϑ satisfies a natural continuity condition with respect to d

$$\forall\, \varepsilon > 0\ \exists\, \eta > 0 : \quad \mathsf{d}(\vartheta(s_0), \vartheta(s_1)) \leq \varepsilon \quad \text{for every } s_0, s_1 \in E,\ s_0 \leq s_1 \leq s_0 + \eta. \tag{3.20}$$

The class of curves satisfying (3.20) will be denoted by $\mathrm{C}_{\mathsf{d}}(E,X)$ and we will set $\mathrm{C}_{\sigma,\mathsf{d}}(E,X) := \mathrm{C}(E,X) \cap \mathrm{C}_{\mathsf{d}}(E,X)$. In order to get a precise description of this (pseudo-) total variation, we have to introduce a dissipation cost.

Hereafter for every subset $E \subset \mathbb{R}$ we will call $E^- := \inf E$, $E^+ := \sup E$; whenever E is compact, we will denote by $\mathfrak{H}(E)$ the (at most) countable collection of the connected components of the open set $[E^-, E^+] \setminus E$: each element of $\mathfrak{H}(E)$ (the "holes" of E) is therefore an open interval of $\mathbb{R}$. We also denote by $\mathfrak{P}_f(E)$ the collection of all the finite subsets of E.

Definition 3.5 (Transition cost) *Let $E \subset \mathbb{R}$ compact and $\vartheta \in \mathrm{C}_{\sigma,\mathsf{d}}(E;X)$. For every $t \in [0,T]$ we define the* transition cost function $\mathrm{Trc}(t,\vartheta,E)$ *by*

$$\mathrm{Trc}(t,\vartheta,E) := \mathrm{Var}_{\mathsf{d}}(\vartheta,E) + \mathrm{GapVar}_{\delta}(\vartheta,E) + \sum_{s\in E\setminus\{E^+\}} \mathcal{R}(t,\vartheta(s)) \tag{3.21}$$

where the first term is defined as the usual total variation (2.6)*, the second one is*

$$\mathrm{GapVar}_{\delta}(\vartheta,E) := \sum_{I\in\mathfrak{H}(E)} \delta(\vartheta(I^-),\vartheta(I^+)),$$

and the third term is

$$\sum_{s\in E\setminus\{E^+\}} \mathcal{R}(t,\vartheta(s)) := \sup\left\{\sum_{s\in P}\mathcal{R}(t,\vartheta(s)) : P \in \mathfrak{P}_f(E\setminus\{E^+\})\right\},$$

where the sum is defined as 0 *if $E \setminus \{E^+\} = \emptyset$.*

We adopt the convention $\mathrm{Trc}(t,\vartheta,\emptyset) := 0$. It is not difficult to check that the transition cost $\mathrm{Trc}(t,\vartheta,E)$ is additive with respect to E:

$$\mathrm{Trc}(t,\vartheta,E\cap[a,c]) = \mathrm{Trc}(t,\vartheta,E\cap[a,b]) + \mathrm{Trc}(t,\vartheta,E\cap[b,c]) \quad \text{for every } a<b<c. \tag{3.22}$$

It will be proved (see Proposition 6.3) that for every $t\in[0,T]$ and for every $\vartheta \in \mathrm{C}(E;X)$

$$\mathcal{E}(t,\vartheta(E^+)) + \mathrm{Trc}(t,\vartheta,E) \geq \mathcal{E}(t,\vartheta(E^-)). \tag{3.23}$$

The dissipation cost $\mathsf{c}(t,u_0,u_1)$ induced by the function Trc is defined by minimizing $\mathrm{Trc}(t,\vartheta,E)$ among all the transitions ϑ connecting u_0 to u_1:

Definition 3.6 (Jump dissipation cost and augmented total variation) *Let $t \in [0,T]$ be fixed and let us consider $u_0,u_1 \in X$. We set*

$$\mathsf{c}(t,u_0,u_1) := \inf\left\{\mathrm{Trc}(t,\vartheta,E) : E \Subset \mathbb{R},\ \vartheta \in \mathrm{C}_{\sigma,\mathsf{d}}(E;X),\ \vartheta(E^-)=u_0,\ \vartheta(E^+)=u_1\right\}, \tag{3.24}$$

with the incremental dissipation cost $\Delta_{\mathsf{c}}(t,u_0,u_1) := \mathsf{c}(t,u_0,u_1) - \mathsf{d}(u_0,u_1)$. The corresponding augmented total variation $\mathrm{Var}_{\mathsf{d},\mathsf{c}}$ is then defined according to Definition 2.5.

Since $\mathcal{R}$ and GapVar_δ are positive, as in the case of BV solutions, it is immediate to check that

$$\mathsf{c}(t,u_0,u_1) \geq \mathsf{d}(u_0,u_1) \quad \text{for every } u_0,u_1 \in X.$$

Moreover, from (3.23), it easily follows that

$$\mathsf{c}(t,u_0,u_1) \geq \mathcal{E}(t,u_0) - \mathcal{E}(t,u_1). \tag{3.25}$$

As in the case of $\mathit{Var}_{\mathsf{d},\mathsf{e}}$ for Balanced Viscosity solutions, $\mathit{Var}_{\mathsf{d},\mathsf{c}}$ is not a *standard* total variation functional: for instance, it is not induced by any distance on X. Nevertheless, $\mathit{Var}_{\mathsf{d},\mathsf{c}}$ enjoys the nice additivity property (2.17).

3.3 Visco-Energetic solutions

We can now give our precise definition of *Visco-Energetic solution* of the rate-independent system $(X, \mathcal{E}, \mathsf{d}, \delta)$. We will always assume that the energy functional satisfy the standard assumptions $\langle \mathrm{A} \rangle$ and δ is an admissible viscous correction (i.e. $\langle \mathrm{B} \rangle$ hold).

Definition 3.7 (Visco-Energetic (VE) solutions) *We say that a (σ, d)-regulated curve $u : [0,T] \to X$ is a* Visco-Energetic (VE) solution *of the rate-independent system $(X, \mathcal{E}, \mathsf{d}, \delta)$ if it satisfies the stability condition*

$$u(t) \in \mathscr{S}_{\mathsf{D}}(t) \quad \textit{for every } t \in [0,T] \setminus \mathrm{J_u}, \tag{$\mathrm{S_D}$}$$

and the energetic balance

$$\mathcal{E}(t, u(t)) + \mathit{Var}_{\mathsf{d},\mathsf{c}}(u, [0,t]) = \mathcal{E}(0, u(0)) + \int_0^t \mathcal{P}(s, u(s))\mathrm{d}s \tag{$\mathrm{E_{d,c}}$}$$

for every $t \in [0,T]$, where c is the jump dissipation cost (3.24).

As in the case of energetic and BV solutions, it is not difficult to see that the energy balance ($\mathrm{E_{d,c}}$) holds on any subinterval of $[t_0, t_1]$ of $[0,T]$:

$$\mathcal{E}(t_1, u(t_1)) + \mathit{Var}_{\mathsf{d},\mathsf{c}}(u, [t_0,t_1]) = \mathcal{E}(t_0, u(t_0)) + \int_{t_0}^{t_1} \mathcal{P}(s, u(s))\mathrm{d}s.$$

Indeed, this follows from the additivity property (2.17) for the augmented total variation $\mathit{Var}_{\mathsf{d},\mathsf{c}}$. Moreover, if a curve $u \in \mathrm{BV}_{\sigma,\mathsf{d}}([0,T]; X)$ satisfies the stability condition ($\mathrm{S_D}$), then a *chain-rule* inequality holds:

$$\mathcal{E}(t_1, u(t_1)) + \mathit{Var}_{\mathsf{d},\mathsf{c}}(u, [t_0,t_1]) \geq \mathcal{E}(t_0, u(t_0)) + \int_{t_0}^{t_1} \mathcal{P}(s, u(s))\mathsf{d}s. \tag{3.26}$$

As a direct consequence, we have a characterization of VE solutions in terms of a single, global in time, energy-dissipation inequality or of a d-energy-dissipation inequality combined with a precise description of the jump behaviour. The proof can be easily adapted from [28, Prop. 4.2, Thm. 4.3].

Proposition 3.8 (Sufficient criteria for VE solutions) *Let $u \in \mathrm{BV}_{\sigma,\mathsf{d}}([0,T]; X)$ be a curve satisfying the stability condition* ($\mathrm{S_D}$). *Then u is a* VE *solution of the rate-independent system $(X, \mathcal{E}, \mathsf{d}, \delta)$ if and only if it satisfies one of the following equivalent characterizations:*

i) u satisfies the (d, c)-energy-dissipation inequality

$$\mathcal{E}(T, u(T)) + \mathit{Var}_{\mathsf{d},\mathsf{c}}(u, [0,T]) \leq \mathcal{E}(0, u(0)) + \int_0^T \mathcal{P}(s, u(s))\mathrm{d}s. \tag{3.27}$$

ii) u satisfies the d-energy-dissipation inequality

$$\mathcal{E}(t, u(t)) + \mathrm{Var}_{\mathsf{d}}(u, [s,t]) \leq \mathcal{E}(s, u(s)) + \int_s^t \mathcal{P}(r, u(r))\mathrm{d}r \quad \textit{for all } s \leq t \in [0,T] \tag{3.28}$$

and the following jump conditions at each point $t \in \mathrm{J_u}$

$$\begin{aligned} \mathcal{E}(t, u(t-)) - \mathcal{E}(t, u(t)) &= \mathsf{c}(t, u(t-), u(t)), \\ \mathcal{E}(t, u(t)) - \mathcal{E}(t, u(t+)) &= \mathsf{c}(t, u(t), u(t+)), \\ \mathcal{E}(t, u(t-)) - \mathcal{E}(t, u(t+)) &= \mathsf{c}(t, u(t-), u(t+)). \end{aligned} \tag{3.29}$$

Existence of VE solutions. As in the energetic and BV cases, existence of Visco-Energetic solutions can be obtained by proving the convergence of discrete solutions to the incremental minimization scheme 3.1. Besides the canonical assumptions $\langle \mathrm{A} \rangle$ and $\langle \mathrm{B} \rangle$ we will further suppose that the following properties hold.

Assumption $\langle \mathbf{C} \rangle$ (Closure and separation properties for the stable set)

$\langle \mathbf{C.1} \rangle$ *For every $Q \geq 0$ the (D, Q)-quasistable sets* (3.3) *have σ-closed intersections with the sublevels of $\mathcal{F}$.*

$\langle \mathbf{C.2} \rangle$ *The D-stable set is separated by d.*

In the viscous setting these assumptions correspond to (2.26) in the energetic one. $\langle \mathrm{C.1} \rangle$ is always satisfied in the case when $\langle \mathrm{A.2'} \rangle$ holds, in particular in the simpler metric case considered in Remark 2.2. Notice that $\langle \mathrm{C.1} \rangle$ is equivalent to assume that the residual stability function of Definition 3.4

$$\mathcal{R} \text{ is } \sigma\text{-l.s.c. on the sublevels of } \mathcal{F}. \tag{C.1'}$$

Our main result is stated in the following theorem.

Theorem 3.9 (Convergence of the $(\mathrm{IM}_{\mathsf{d},\delta})$ scheme and existence of VE solutions)

Let us suppose that Assumptions $\langle \mathrm{A} \rangle$, $\langle \mathrm{B} \rangle$, $\langle \mathrm{C} \rangle$ *hold. Let $u_0 \in X$ be fixed and let $\overline{U}_\tau$ be the family of piecewise left-continuous constant interpolants of discrete solutions U_τ^n of* $(\mathrm{IM}_{\mathsf{d},\delta})$, *with*

$$\mathsf{d}(x_o, U_\tau^0) \leq C, \quad U_\tau^0 \xrightarrow{\sigma} u_0, \qquad \mathcal{E}(0, U_\tau^0) \to \mathcal{E}(0, u_0) \quad \textit{as } |\tau| \downarrow 0. \tag{3.30}$$

Then for all sequence of partitions $k \mapsto \tau(k)$ with $|\tau(k)| \downarrow 0$ there exist a (not relabeled) subsequence and a limit curve $u \in \mathrm{BV}_{\sigma,\mathsf{d}}([0,T]; X)$ such that

$$\overline{U}_{\tau(k)}(t) \xrightarrow{\sigma} u(t), \quad \mathcal{E}(t, \overline{U}_{\tau(k)}(t)) \to \mathcal{E}(t, u(t)) \quad \textit{as } k \to \infty \quad \textit{for every } t \in [0,T],$$

and u is a Visco-Energetic solution of the rate-independent system $(X, \mathcal{E}, \mathsf{d}, \delta)$ starting from u_0.

The proof of Theorem 3.9 will follow a standard structure, that will be exploited in Section 7, strongly relying on the basic and preliminary results of Sections 5, 6.

- We will first derive discrete stability estimates in Section 7.1 for the solution of the incremental minimization problem $(\mathrm{IM}_{\mathsf{d},\delta})$: here only Assumption $\langle \mathrm{A} \rangle$ will play an important role.
- We will prove a preliminary convergence result by refined compactness arguments (Section 7.2), where we combine the lower semicontinuity of the residual stability function $\mathcal{R}$ $\langle \mathrm{C.1} \rangle$ with the d-separation of the D-stable set $\mathscr{S}_{\mathsf{D}}$ $\langle \mathrm{C.2} \rangle$; in this way, we will prove that every limit curve satisfies the stability condition $(\mathrm{S}_{\mathsf{D}})$.
- We will then obtain the energy-dissipation inequality (3.27) by proving the lower semicontinuity of the augmented total variation $Var_{\mathsf{d},\mathsf{c}}$ (section 7.3): this property strongly relies on the lower semicontinuity of the jump dissipation cost c, which will be thoroughly studied in Section 5: here the minimal compatibility properties $\langle \mathrm{B.1} \rangle$-$\langle \mathrm{B.2} \rangle$ of the viscous correction δ will enter in the game.

- The whole argument will be concluded by showing that along arbitrary stable curves the decay rate of the energy can always be controlled by the power integral and the augmented variation $Var_{\mathsf{d},\mathsf{c}}$ (3.26): this topic will be discussed in Section 6 and strongly depends on $\langle$B.3$\rangle$.

3.4 The residual stability function $\mathcal{R}$

Let us briefly discuss a few properties of the residual stability function $\mathcal{R}$. We first introduce the Moreau-Yosida regularization of $\mathcal{E}$ and its associated minimal set.

Definition 3.10 (Moreau-Yosida regularization and minimal set) *Let us suppose that $\mathcal{E}$ satisfies $\langle$A.1$\rangle$. The D-Moreau-Yosida regularization $\mathcal{Y} : [0,T] \times X \to \mathbb{R}$ of $\mathcal{E}$ is defined by*

$$\mathcal{Y}(t,x) := \min_{y \in X} \mathcal{E}(t,y) + \mathsf{D}(x,y). \tag{3.31}$$

For every $t \in [0,T]$ and $x \in X$ the minimal set is

$$\mathrm{M}(t,x) := \operatorname*{argmin}_X \mathcal{E}(t,\cdot) + \mathsf{D}(x,\cdot) = \Big\{ y \in X : \mathcal{E}(t,y) + \mathsf{D}(x,y) = \mathcal{Y}(t,x) \Big\}. \tag{3.32}$$

Notice that by $\langle$A.1$\rangle$ $\mathrm{M}(t,x) \neq \emptyset$ for every t,x. It is clear that

$$\mathcal{R}(t,x) = \mathcal{E}(t,x) - \mathcal{Y}(t,x). \tag{3.33}$$

In the next Lemma we collect a list of useful properties, connecting $\mathcal{R}$, $\mathcal{Y}$ and M.

Lemma 3.11 *Let us suppose that* Assumption $\langle$A$\rangle$ *holds. Then*

i)
$$\mathcal{E}(t,y) + \mathsf{D}(x,y) + \mathcal{R}(t,x) \geq \mathcal{E}(t,x) \quad \text{for every } t \in [0,T], \quad x,y \in X \tag{3.34}$$
and equality holds in (3.34) *if and only if $y \in \mathrm{M}(t,x)$.*

ii) The map $(t,x) \mapsto \mathcal{Y}(t,x)$ is $\sigma_{\mathbb{R}}$-lower semicontinuous on every sublevel of $\mathcal{F}_0$.

iii) If $(t_n,x_n) \xrightarrow{\sigma_{\mathbb{R}}} (t,x)$ and $\mathsf{D}(x_n,x) \to 0$ as $n \to \infty$, we have
$$\limsup_{n\to\infty} \mathcal{Y}(t_n,x_n) \leq \mathcal{Y}(t,x), \quad \liminf_{n\to\infty} \mathcal{R}(t_n,x_n) \geq \mathcal{R}(t,x). \tag{3.35}$$
In particular, if $y \mapsto \mathsf{D}(y,x)$ is σ-continuous (on the sublevels of $\mathcal{F}_0$) then $\mathcal{R}$ is lower semicontinuous (on the sublevels of $\mathcal{F}_0$).

iv) If $(t_n,x_n) \xrightarrow{\sigma_{\mathbb{R}}} (t,x)$ and $\mathcal{E}(t_n,x_n) \to \mathcal{E}(t,x)$ then
$$\limsup_{n\to\infty} \mathcal{R}(t_n,x_n) \leq \mathcal{R}(t,x). \tag{3.36}$$
If moreover $\liminf_{n\to\infty} \mathcal{R}(t_n,x_n) \geq \mathcal{R}(t,x)$ then any limit point y of a sequence $y_n \in \mathrm{M}(t_n,x_n)$ belongs to $\mathrm{M}(t,x)$.

v) If $(t_n,x_n) \xrightarrow{\sigma_{\mathbb{R}}} (t,x)$ and $\mathsf{D}(x_n,x) \to 0$ as $n \to \infty$ with $\mathsf{d}(x_o,x_n) \leq C$, we have
$$\lim_{n\to\infty} \mathcal{R}(t_n,x_n) = 0 \quad \Rightarrow \quad \lim_{n\to\infty} \mathcal{E}(t_n,x_n) = \mathcal{E}(t,x). \tag{3.37}$$

vi) $\mathcal{R}$ *is lower semicontinuous on the sublevels of* $\mathcal{F}_0$ *if and only if for every* d*-bounded sequence* (t_n, x_n) *converging to* (t, x) *in* $[0,T]\times X$ *with* $\lim_{n\to\infty} \mathcal{E}(t_n, x_n) = \bar{\mathcal{E}} = \mathcal{E}(t,x)+\eta$, $\eta \geq 0$ *there exists* $y \in \mathrm{M}(t,x)$ *and a sequence* y_n *such that*

$$\liminf_{n\to\infty} \Big(\mathcal{E}(t, y_n) + \mathsf{D}(x_n, y_n)\Big) \leq \mathcal{E}(t,y) + \mathsf{D}(x,y) + \eta. \tag{3.38}$$

Remark 3.12 (Mutual recovery sequences) The characterization vi) of the lower semicontinuity of $\mathcal{R}$ (a crucial property in view of (C.1')) is strongly related to the *mutual recovery sequence* condition which typically characterizes the closure of the d-stable set in the energetic case (see e.g. [33, Lemma 2.1.14]). Notice however that the sequence (t_n, x_n) in vi) is not assumed to be stable.

Proof of Lemma 3.11. i) is an immediate consequence of the definition.

ii) Let us consider a sequence $(t_n, x_n)_n$ with $\mathcal{F}_0(x_n) \leq C$, $\mathcal{Y}(t, x_n) \leq Y$ for every $n \in \mathbb{N}$ and $(t_n, x_n) \xrightarrow{\sigma_\mathbb{R}} (t, x)$ as $n \to \infty$. Let $y_n \in \mathrm{M}(t_n, x_n)$ so that $\mathcal{Y}(t_n, x_n) = \mathcal{E}(t_n, y_n)+\mathsf{D}(x_n, y_n) \leq Y$; it is easy to check (see also Theorem 7.1) that $\mathcal{F}_0(y_n) \leq C'$ so that it is not restrictive to assume by $\langle$A.1$\rangle$ (up to extracting a not relabeled subsequence) that $y_n \xrightarrow{\sigma} y$. The lower semicontinuity of $\mathcal{E}$, d and δ yield

$$\mathcal{Y}(t,x) \leq \mathcal{E}(t,y) + \mathsf{D}(x,y) \leq \liminf_{n\uparrow\infty} \mathcal{E}(t_n, y_n) + \mathsf{D}(x_n, y_n) \leq Y.$$

iii) Let $y \in \mathrm{M}(t,x)$ so that $\mathcal{Y}(t,x) = \mathcal{E}(t,y) + \mathsf{D}(x,y)$; by definition

$$\mathcal{Y}(t_n, x_n) \leq \mathcal{E}(t_n, y) + \mathsf{D}(x_n, y) \leq \Big|\int_{t_n}^t \mathcal{P}(r, y)\,\mathrm{d}r\Big| + \mathcal{E}(t,y) + \mathsf{D}(x_n, y);$$

passing to the limit as $n \uparrow \infty$ and recalling $\langle$B.2$\rangle$ we get the first property of (3.35). The second one follows by (3.33) and the lower semicontinuity of $\mathcal{E}$.

iv) (3.36) is a consequence of i), (3.33), and the convergence of the energy. The last statement follows immediately by i).

v) (3.34) yields

$$\mathcal{E}(t_n, x_n) \leq \mathcal{E}(t_n, x) + \mathsf{D}(x_n, x) + \mathcal{R}(t_n, x_n)$$

so that $\limsup_{n\to\infty} \mathcal{E}(t_n, x_n) \leq \mathcal{E}(t, x)$. Since x_n belongs to a sublevel of $\mathcal{F}_0$ the σ-lower semicontinuity of $\mathcal{E}$ yields (3.37).

vi) If property (3.38) holds for every sequence $(x_n)_n$, up to extracting a further subsequence it is not restrictive to suppose that $\mathcal{R}(t_n, x_n)$ is converging, so that

$$\begin{aligned}
\lim_{n\to\infty} \mathcal{R}(t_n, x_n) &\geq \limsup_{n\to\infty} \mathcal{E}(t_n, x_n) - \mathcal{E}(t_n, y_n) - \mathsf{D}(x_n, y_n)\\
&\geq \mathcal{E}(t,x) + \eta - \liminf_{n\to\infty} \Big(\mathcal{E}(t_n, y_n) + \mathsf{D}(x_n, y_n)\Big)\\
&\overset{(3.38)}{\geq} \mathcal{E}(t,x) - \mathcal{E}(t,y) - \mathsf{D}(x,y) = \mathcal{R}(t,x).
\end{aligned}$$

Conversely, let us suppose that $\mathcal{R}$ is lower semicontinuous and let (t_n, x_n) be a sequence satisfying $\lim_{n\to\infty} \mathcal{E}(t_n, x_n) = \bar{\mathcal{E}} = \mathcal{E}(t,x) + \eta$, $\eta \geq 0$. We pick up any $y_n \in \mathrm{M}(t_n, x_n)$ obtaining

$$\begin{aligned}
\limsup_{n\uparrow\infty} \mathcal{E}(t, y_n) + \mathsf{D}(x_n, y_n) &= \limsup_{n\uparrow\infty} \mathcal{E}(t_n, y_n) + \mathsf{D}(x_n, y_n) \\
&= \limsup_{n\uparrow\infty} \mathcal{E}(t_n, x_n) - \Big(\mathcal{E}(t_n, x_n) - \mathcal{E}(t_n, y_n) - \mathsf{D}(x_n, y_n)\Big) \\
&= \mathcal{E}(t,x) + \eta - \liminf_{n\to\infty} \Big(\mathcal{E}(t_n, x_n) - \mathcal{E}(t_n, y_n) - \mathsf{D}(x_n, y_n)\Big) \\
&= \mathcal{E}(t,x) + \eta - \liminf_{n\to\infty} \mathcal{R}(t_n, x_n) = \mathcal{E}(t,x) + \eta - \mathcal{R}(t,x) \\
&\leq \mathcal{E}(t,x) + \eta - \Big(\mathcal{E}(t,x) - \mathcal{E}(t,y) - \mathsf{D}(x,y)\Big) = \mathcal{E}(t,y) + \mathsf{D}(x,y) + \eta.
\end{aligned}$$

□

3.5 Optimal jump transitions

Thanks to the jump conditions given by (3.29), we can give a finer description of the behaviour of Visco-Energetic solutions along jumps. The crucial notion is provided by the following definition.

Definition 3.13 (Optimal transitions) *Let $t \in [0,T]$ and u_-, $u_+ \in X$. We say that a curve $\vartheta \in \mathrm{C}_{\sigma,\mathsf{d}}(E; X)$, E being a compact subset of $\mathbb{R}$, is an optimal transition between u_- and u_+ if*

$$u_- = \vartheta(E^-), \quad u_+ = \vartheta(E^+), \quad \mathsf{c}(t, u_-, u_+) = \mathrm{Trc}(t, \vartheta, E). \tag{3.39}$$

ϑ is tight *if for every $I \in \mathfrak{H}(E)$ $\vartheta(I^-) \neq \vartheta(I^+)$. ϑ is a*

$$\textit{pure jump transition, if } E \setminus \{E^-, E^+\} \textit{ is discrete,} \tag{3.40}$$

$$\textit{sliding transition, if } \mathcal{R}(t, \theta(r)) = 0 \quad \textit{for every } r \in E, \tag{3.41}$$

$$\textit{viscous transition, if } \mathcal{R}(t, \theta(r)) > 0 \quad \textit{for every } r \in E \setminus \{E^{\pm}\}. \tag{3.42}$$

We will say that a compact set E is almost discrete if $E \setminus \{E^-, E^+\}$ is discrete.

It is easy to check that almost discrete compact sets E can be parametrized by sequences $n \mapsto e_n$ defined in a compact interval Z of $\mathbb{Z} \cup \{\pm\infty\}$ and continuous at $n = \pm\infty$ whenever those points belong to Z.

The main interest of optimal transitions derives from the next result, whose proof follows immediately from Corollary 5.5 later on.

Theorem 3.14 *Under the same assumptions $\langle$A$\rangle$, $\langle$B$\rangle$, $\langle$C$\rangle$ of Theorem 3.9, if $u \in \mathrm{BV}_{\sigma,\mathsf{d}}([0,T]; X)$ is a Visco-Energetic solution to the rate-independent system $(X, \mathcal{E}, \mathsf{d}, \delta)$, then for every $t \in \mathrm{J_u}$ there exists a tight optimal transition $\vartheta \in \mathrm{C}_{\sigma,\mathsf{d}}(E, X)$ between $u(t-)$ and $u(t+)$ such that*

$$u(t) \in \vartheta(E) \quad \textit{and} \quad \mathcal{E}(t, u(t-)) - \mathcal{E}(t, u(t+)) = \mathrm{Trc}(t, \vartheta, E). \tag{3.43}$$

Remark 3.15 If $\langle$C.1$\rangle$ holds and $\vartheta \in \mathrm{C}_{\sigma,\mathsf{d}}(E, X)$ is a transition with finite cost $\mathrm{Trc}(t, \vartheta, E) < \infty$, then the set

$$E_{\mathcal{R}} := \Big\{ r \in E \setminus \{E^+\} : \mathcal{R}(t, \theta(r)) > 0 \Big\} \text{ is discrete, i.e. all its points are isolated.} \tag{3.44}$$

Indeed, since $\mathcal{R}$ is lower semicontinuous by $\langle$C.1$\rangle$, there exists $\eta > 0$ such that $\mathcal{R}(t, \vartheta(r)) \geq \frac{1}{2}\mathcal{R}(t, \vartheta(r_0)) > 0$ for every $r_0 \in E_{\mathcal{R}}$ and $|r - r_0| < \eta$. On the other hand, the finiteness of the transition cost yields $\sum_{s \in E \setminus \{E^+\}} \mathcal{R}(t, \vartheta(s)) < \infty$ so that $E_{\mathcal{R}} \cap \{r \in \mathbb{R} : |r - r_0| \leq \eta\}$ is finite.

We have another interesting characterization of optimal Visco-Energetic transitions. Whenever a set $E \subset \mathbb{R}$ is given, we will use the notation

$$r_E^- := \sup \big(E \cap (-\infty, r)\big) \cup \{E^-\}, \quad r_E^+ := \inf \big(E \cap (r, +\infty)\big) \cup \{E^+\}, \quad r \in \mathbb{R}. \tag{3.45}$$

Theorem 3.16 *A curve $\vartheta \in \mathrm{C}_{\sigma,\mathsf{d}}(E, X)$ with $\vartheta(E) \ni u(t)$ is an optimal transition between $u(t-)$ and $u(t+)$ satisfying* (3.43) *if and only if it satisfies*

$$\mathrm{Var}_{\mathsf{d}}(\vartheta, E \cap [r_0, r_1]) \leq \mathcal{E}(t, \vartheta(r_0)) - \mathcal{E}(t, \vartheta(r_1)) \quad \textit{for every } r_0, r_1 \in E,\ r_0 \leq r_1, \tag{3.46}$$

and

$$\vartheta(r) \in \mathrm{M}(t, \vartheta(r_E^-)) \quad \textit{for every } r \in E \setminus \{E^-\}. \tag{3.47}$$

Proof. By the additivity property of Trc (3.22) and the energy inequality (3.23) it is easy to check that (3.43) yields

$$\mathrm{Trc}(t, \vartheta, E \cap [a, b]) = \mathcal{E}(t, \vartheta(a)) - \mathcal{E}(t, \vartheta(b)) \quad \text{for every } a, b \in E,\ a < b. \tag{3.48}$$

Since $\mathrm{Var}_{\mathsf{d}}(\vartheta, E \cap [r_0, r_1]) \leq \mathrm{Trc}(t, \vartheta, E \cap [r_0, r_1])$ we get (3.46); particularizing (3.48) to the case of $a = r_E^-$, $b = r$, we also get

$$\mathsf{D}(\vartheta(r_E^-), \vartheta(r)) + \mathcal{R}(t, \vartheta(r_E^-)) = \mathcal{E}(t, \vartheta(r_E^-)) - \mathcal{E}(t, \vartheta(r)) \tag{3.49}$$

showing that $\vartheta(r) \in \mathrm{M}(t, \vartheta(r_E^-))$ by Lemma 3.11 i). Notice that when $r = r_E^-$ (3.48) simply yields $\mathcal{R}(t, \vartheta(r)) = 0$, i.e. $\vartheta(r) \in \mathscr{S}_{\mathsf{D}}(t)$.

In order to prove the converse implication, we fix $\varepsilon > 0$ and we consider a finite subset $H \subset \mathfrak{H}(E)$ such that

$$\sum_{I \in H} \delta(\vartheta(I^-), \vartheta(I^+)) \geq \mathrm{GapVar}_{\delta}(t, \vartheta, E) - \varepsilon, \tag{3.50}$$

and let $H_\pm := \{I^\pm : I \in H\}$. Since $E_{\mathcal{R}} \subset H_-$ by Remark 3.15, we can choose H sufficiently big so that

$$\sum_{I \in H} \mathcal{R}(t, \vartheta(I^-)) \geq \sum_{s \in E \setminus \{E^+\}} \mathcal{R}(t, \vartheta(s)) - \varepsilon. \tag{3.51}$$

Let us consider now an arbitrary finite part $F = \{E^- = s_0 < s_1 < \cdots < s_N = E^+\} \subset E$ containing $H_- \cup H_+$ such that

$$\sum_{n=1}^{N} \mathsf{d}(\vartheta(s_{n-1}), \vartheta(s_n)) \geq \mathrm{Var}_{\mathsf{d}}(\vartheta, E) - \varepsilon \tag{3.52}$$

and let $k \mapsto n(k)$ be an increasing sequence such that $H_- = \{s_{n(k)} : 1 \le k \le K\}$. Notice that for every $I \in H$ if $I^- = s_{n(k)}$ then $I^+ = s_{n(k)+1}$, since $I^+ \subset F$. Setting $n(0) = 0$, $n(K+1) = N$, we have

$$
\begin{aligned}
\mathcal{E}(t,u(t-)) - \mathcal{E}(t,u(t+)) &= \sum_{k=0}^{K} \mathcal{E}(t,\vartheta(s_{n(k)})) - \mathcal{E}(t,\vartheta(s_{n(k+1)})) = \mathcal{E}(t,\vartheta(s_0)) - \mathcal{E}(t,\vartheta(s_{n(1)})) \\
&\quad + \sum_{k=1}^{N} \Big(\mathcal{E}(t,\vartheta(s_{n(k)})) - \mathcal{E}(t,\vartheta(s_{n(k)+1}))\Big) + \sum_{k=1}^{N} \Big(\mathcal{E}(t,\vartheta(s_{n(k)+1})) - \mathcal{E}(t,\vartheta(s_{n(k+1)}))\Big) \\
&\overset{(3.47)}{\ge} \mathrm{Var}_{\mathsf{d}}(\vartheta, E \cap [s_0, s_{n(1)}]) + \sum_{k=1}^{N} \mathcal{R}(t,\vartheta(s_{n(k)})) + \sum_{k=1}^{N} \delta(\vartheta(s_{n(k)}),\vartheta(s_{n(k)+1})) \\
&\quad + \sum_{k=1}^{N} \mathsf{d}(\vartheta(s_{n(k)}),\vartheta(s_{n(k)+1})) + \sum_{k=1}^{N} \mathrm{Var}_{\mathsf{d}}(\vartheta, E \cap [s_{n(k)+1}, s_{n(k+1)}]) \\
&\ge \sum_{n=1}^{N} \mathsf{d}(\vartheta(s_{n-1}),\vartheta(s_n)) + \sum_{k=1}^{N} \mathcal{R}(t,\vartheta(s_{n(k)})) + \sum_{k=1}^{N} \delta(\vartheta(s_{n(k)}),\vartheta(s_{n(k)+1})) \ge \mathrm{Trc}(t,\vartheta,E) - 3\varepsilon,
\end{aligned}
$$

where the last inequality results from (3.50), (3.51), and (3.52). Since $\varepsilon > 0$ is arbitrary, by recalling (3.23) we get (3.43). □

Corollary 3.17 (Representation of optimal transitions of viscous type) *We can always represent an optimal viscous transition between $u(t-)$ and $u(t+)$ as a finite or countable sequence $n \mapsto \vartheta(n)$ defined in a compact interval Z of $\mathbb{Z} \cup \{\pm\infty\}$ satisfying*

$$
\vartheta(n) \in \mathrm{M}(t,\vartheta(n-1)) \quad \textit{for every } n \in Z \setminus \{Z^-\}, \quad \vartheta(Z^\pm) = u(t\pm), \tag{3.53}
$$

and the continuity conditions (whenever $\pm\infty \in Z$)

$$
\begin{gathered}
\lim_{n\to\pm\infty} \vartheta(n) = u(t\pm), \quad \lim_{n\to-\infty} \mathsf{d}(u(t-),\vartheta(n)) = 0, \quad \lim_{n\to+\infty} \mathsf{d}(\vartheta(n),u(t+)) = 0, \\
\lim_{n\downarrow-\infty} \mathcal{E}(t,\vartheta(n)) = \mathcal{E}(t,u(t-)), \quad \lim_{n\uparrow+\infty} \mathcal{E}(t,\vartheta(n)) = \mathcal{E}(t,u(t+)).
\end{gathered} \tag{3.54}
$$

An optimal transition ϑ can be decomposed in a canonical way into (at most countable) collections of sliding and pure jump transitions.

Proposition 3.18 *Let $\vartheta \in \mathrm{C}_{\sigma,\mathsf{d}}(E,X)$ be an optimal transition between u_- and u_+. Then there exist disjoint closed intervals $(S_j)_{j\in\sigma}$ and almost discrete compact sets $\{V_k\}_{k\in\nu}$, with $\sigma,\nu \subset \mathbb{N}$, such that*

$$
E = (\cup_{j\in\sigma} S_j) \cup \overline{(\cup_{k\in\nu} V_k)} \tag{3.55}
$$

and

$$
\vartheta_{|S_j} \textit{ is of sliding type,} \quad \vartheta_{|V_k} \textit{ is of pure jump type.} \tag{3.56}
$$

Proof. We set for every $r \in [E^-, E^+]$

$$
E_0 := \bigcup_{I\in\mathfrak{H}(E)} \{I^-, I^+\}, \quad E_1 := E \setminus E_0, \quad E_0(r) := E \cap [r^-_{E_1}, r^+_{E_1}].
$$

Notice that E_0 contains all the isolated points of E (in particular it contains $E_{\mathfrak{R}}$). If $r \in E_0$, $E_0(r)$ is the closure of the "maximal component" of E_0 containing r, in the sense that all the other points of E_0 are separated from r by some accumulation point in E_1. The restriction of ϑ to $E_0(r)$ is of pure jump type and $E_0(r)$ is almost discrete.

We first decompose E_0 in the disjoint countable union of $\cup_{k\in\nu} V_k \cap E_0$ where V_k is of the form $E_0(r)$ for some $r \in E_0$. We then set

$$V = \overline{E_0}, \quad S := E \setminus (V \cup \{E^\pm\}),$$

observing that

$$S = (E^-, E^+) \setminus \overline{\bigcup_{I\in\mathfrak{H}(E)} I}.$$

We can now decompose the set S, open in $\mathbb{R}$, as the disjoint union of its connected components (a_j, b_j), $j \in \sigma$, and we set $S_j := [a_j, b_j]$ obtaining (3.55). Since $E_{\mathfrak{R}} \subset E_0$ we also get (3.56). $\square$

As we have seen in Remark 3.15, if an optimal transition $\vartheta : E \to X$ is of viscous type, then the set $E \setminus \{E^-, E^+\}$ is discrete. In general it may happen that E is homeomorphic to a finite set of $\mathbb{Z}$ or to infinite intervals of the form $\{-\infty\} \cup -\mathbb{N}$, $\mathbb{N} \cup \{+\infty\}$ or even to $\mathbb{Z} \cup \{\pm\infty\}$. We can be more precise in the case when the functional

$$u \mapsto \mathcal{E}(t, u) + \mathsf{D}(u_0, u) \quad \text{admits a unique minimizer in } X \text{ for every } u_0 \in X. \tag{3.57}$$

This happens, e.g. , if X is a linear space and we choose a sufficiently strong viscous correction δ so that the map $u \mapsto \mathcal{E}(t, u) + \mathsf{D}(u_0, u)$ is strictly convex.

Proposition 3.19 *Let $\vartheta : E \to X$ be a tight optimal transition between u_- and u_+.*

i) If the energy and the dissipation satisfy (3.57) *then every $r \in E \setminus (E_{\mathfrak{R}} \cup \{E^+\})$ (in particular $r = E^-$ when u_- is stable) is a right accumulation point of E, i.e. there exists a sequence $r_k \in E \cap (r, \infty)$ such that $r_k \downarrow r$.*

ii) If X is a vector space, d is the distance induced by a norm on X, $\delta(u, v) := \frac{\mu}{2}\mathsf{d}^2(u, v)$ as in (3.12) *and $\mathcal{E}$ is Gateaux differentiable in X then every $r \in E \setminus (E_{\mathfrak{R}} \cup \{E^-\})$ (in particular $r = E^+$ when u_+ is stable) is a left accumulation point of E, i.e. there exists a sequence $r_k \in E \cap (-\infty, r)$ such that $r_k \uparrow r$.*

Proof. Let us consider i) and let us suppose by contradiction that there exists $s \in E$ such that $(r, s) \in \mathfrak{H}(E)$. Since $\vartheta(r) \in \mathscr{S}_{\mathsf{D}}(t)$ we have $\vartheta(r) \in \mathrm{M}(t, \vartheta(r))$; on the other hand, (3.47) yields $\vartheta(s) \in \mathrm{M}(t, \vartheta(r))$ so that (3.57) yields $\vartheta(r) = \vartheta(s)$ which contradicts the tightness of ϑ.

Concerning ii) we still argue by contradiction assuming that $(s, r) \in \mathfrak{H}(E)$. We denote by $\xi \in X^*$ the unique element of the Gateaux subdifferential of $\mathcal{E}(t, \vartheta(r))$, by N the subdifferential of $\frac{1}{2}\|\cdot\|_X^2$ and by K_* the dual unitary ball of X^*. It is not difficult to check that

$$\xi \in K_*, \quad \frac{N(\vartheta(r) - \vartheta(s))}{\|\vartheta(r) - \vartheta(s)\|} + \mu\, N(\vartheta(r) - \vartheta(s)) \ni \xi, \tag{3.58}$$

so that we obtain

$$(1 + \mu\|\vartheta(r) - \vartheta(s)\|)\|N(\vartheta(r) - \vartheta(s))\|_* \le \|\vartheta(r) - \vartheta(s)\|$$

which contradicts the fact that $\|N(x)\| = \|x\|$. $\square$

Remark 3.20 When $\mathcal{E}$ is nonsmooth a jump from a non-stable point to a stable one may happen even with the assumption of strict convexity of the functional $\mathcal{E} + D$. For instance, we can consider the example

$$X = \mathbb{R}, \quad \mathcal{E}(t,u) = a|u|, \quad d(u,v) = |u-v|, \quad \delta(u,v) = \frac{1}{2}|u-v|^2.$$

If $a > 1$, it is immediate to check that $\mathscr{S}_{\mathsf{D}}(t) = \{0\}$ for every t. If we start from a point $u_- \in (0, a-1)$ then $u_+ = 0$ belongs to $\mathrm{M}(t, u_-)$ and it is also a stable point.

4 Examples

In this section we will discuss some applications of Theorem 3.9 about existence of Visco-Energetic solutions. Let us first recall that once Assumption $\langle \mathrm{A} \rangle$ holds and

$$\mathsf{D} \text{ is left continuous on the sublevels of } \mathcal{F}_0 \text{ (see (2.23)) and } \mathsf{d} \text{ separates } X \tag{4.1}$$

conditions $\langle \mathrm{B.1} \rangle$, $\langle \mathrm{B.2} \rangle$, $\langle \mathrm{C.1} \rangle$, $\langle \mathrm{C.2} \rangle$ are automatically satisfied, so that one can just focus on the verification of the canonical compactness-regularity conditions

$$\langle \mathrm{A} \rangle \text{ and on the compatibility condition } \langle \mathrm{B.3} \rangle. \tag{4.2}$$

The latter is also satisfied if $\delta(u,v)$ is a function of d as in (3.11).

4.1 The convex case

Let us first consider the case when X is a convex subset of a vector space V and d is induced by a convex, positively 1-homogeneous functional $\psi : V \to [0, +\infty)$.

Proposition 4.1 *If $\mathsf{d}(x,y) := \psi(y-x)$ for every $x, y \in X$ and the map $x \mapsto \mathcal{E}(t,x)$ is convex in X for every $t \in [0,T]$, we have*

(i) If $u_- \in \mathscr{S}_{\mathsf{d}}(t)$ and $u_+ \in X$ satisfy the energetic jump condition $\mathcal{E}(t,u_+) + \psi(u_+ - u_-) = \mathcal{E}(t, u_-)$ then $\mathsf{c}(t, u_-, u_+) = \mathsf{d}(u_-, u_+)$.

(ii) If the viscous correction δ satisfies

$$\lim_{\theta \downarrow 0} \frac{\delta(u, (1-\theta)u + \theta v)}{\theta} = 0 \quad \textit{for every } u, v \in X \tag{4.3}$$

then $\mathscr{S}_{\mathsf{D}} = \mathscr{S}_{\mathsf{d}}$.

In particular any energetic solution $u \in \mathrm{BV}_{\mathsf{d}}([0,T]; X)$ of $(X, \mathcal{E}, \mathsf{d})$ is a VE *solution of $(X, \mathcal{E}, \mathsf{d}, \delta)$ and if (4.3) holds any* VE *solution $u \in \mathrm{BV}_{\mathsf{d}}([0,T]; X)$ of $(X, \mathcal{E}, \mathsf{d}, \delta)$ is an energetic solution of $(X, \mathcal{E}, \mathsf{d})$.*

Proof. Let us first suppose that u is an energetic solution. Since the "energetic" stability condition (S_{d}) is stronger than the corresponding "Visco-Energetic" one (S_{D}), it is sufficient to check that u satisfies the Visco-Energetic balance condition ($\mathrm{E}_{\mathsf{d},\mathsf{c}}$); since u satisfies (3.28) it is sufficient to check that (3.29) holds.

Thus let $u_-, u_+ \in \mathscr{S}_{\mathsf{d}}(t)$ with $\mathcal{E}(t, u_+) + \psi(u_+ - u_-) = \mathcal{E}(t, u_-)$. We consider the convex subset of $V \times \mathbb{R}$

$$K := \big\{(v, z) \in V \times \mathbb{R} : u_- + v \in X,\ z \le \mathcal{E}(t, u_-) - \mathcal{E}(t, u_- + v)\big\}.$$

By the Mazur-Orlicz version of Hahn-Banach Theorem [47, Theorem 1.1] there exists a linear functional $L : V \times \mathbb{R} \to \mathbb{R}$ such that

$$L(v, z) \le \psi(v) - z \quad \text{for every } (v, z) \in V \times \mathbb{R}, \qquad \inf_{(v,z)\in K} L = \inf_{(v,z)\in K} \psi(v) - z. \tag{4.4}$$

Writing $L(v, z) = \ell(v) - \alpha z$ for some $\alpha \in \mathbb{R}$ and testing the first condition of (4.4) with $v = 0$ and arbitrary $z \in \mathbb{R}$ we get $L(v, z) = \ell(v) - z$, so that

$$\ell(v) \le \psi(v) \quad \text{for every } v \in V. \tag{4.5}$$

Since u_- is d-stable, for every $(v, z) \in K$ we have

$$\psi(v) - z \ge \mathcal{E}(t, u_-) - \mathcal{E}(t, u_- + v) - z \ge 0;$$

since $(0, 0) \in K$ we conclude that $\inf_K L = 0$, which yields in particular

$$\ell(w - u_-) \ge \mathcal{E}(t, u_-) - \mathcal{E}(t, w) \quad \text{for every } w \in X. \tag{4.6}$$

Choosing $w = u_+$ in (4.6) we deduce $\ell(u_+ - u_-) = \psi(u_+ - u_-)$. Setting $\vartheta(s) := (1-s)u_- + su_+$, $s \in [0, 1]$, we immediately get $\mathcal{E}(t, \vartheta(s)) = (1-s)\mathcal{E}(t, u_-) + s\mathcal{E}(t, u_+) = \mathcal{E}(t, u_-) - \psi(\vartheta(s) - u_-)$ and

$$\begin{aligned}
\mathcal{E}(t, v) &\ge \mathcal{E}(t, u_-) - \ell(v - u_-) \\
&= \mathcal{E}(t, \vartheta(s)) - \ell(v - \vartheta(s)) + \Big(\mathcal{E}(t, u_-) - \mathcal{E}(t, \vartheta(s)) - \ell(\vartheta(s) - u_-)\Big) \\
&\ge \mathcal{E}(t, \vartheta(s)) - \ell(v - \vartheta(s)) + \Big(\psi(\vartheta(s) - u_-) - \ell(\vartheta(s) - u_-)\Big) \\
&\ge \mathcal{E}(t, \vartheta(s)) - \psi(v - \vartheta(s)),
\end{aligned}$$

so that $\vartheta(s) \in \mathscr{S}_{\mathsf{d}}(t) \subset \mathscr{S}_{\mathsf{D}}(t)$. It follows that $\mathsf{c}(t, u_-, u_+) \le \mathrm{Var}_{\mathsf{d}}(\vartheta, [0, 1]) = \mathsf{d}(u_-, u_+)$.

In order to prove the converse implication, we simply have to check that $\mathscr{S}_{\mathsf{D}} \subset \mathscr{S}_{\mathsf{d}}$. If $u \in \mathscr{S}_{\mathsf{D}}(t)$, $v \in X$ and $v_\theta := (1-\theta)u + \theta v$ with $\theta \in [0, 1]$ we have

$$\mathcal{E}(t, u) \le \mathcal{E}(t, v_\theta) + \psi(v_\theta - u) + \delta(u, v_\theta) \le (1-\theta)\mathcal{E}(t, u) + \theta\Big(\mathcal{E}(t, v) + \psi(v - u) + \frac{\delta(u, v_\theta)}{\theta}\Big),$$

which yields

$$\mathcal{E}(t, u) \le \mathcal{E}(t, v) + \psi(v - u) + \frac{\delta(u, v_\theta)}{\theta}.$$

Passing to the limit as $\theta \downarrow 0$ we conclude. $\square$

4.2 The 1-dimensional case.

In the space $X := \mathbb{R}$ consider a function $W \in \mathrm{C}^2(\mathbb{R})$ bounded from below with $-\lambda := \inf_{\mathbb{R}} W'' > -\infty$, a function $\ell \in \mathrm{C}^1([0,T])$ and positive numbers $\alpha_\pm, \mu$; the standard example for W is the double-well potential $W(u) = \frac{1}{4}(1-u^2)^2$. We set

$$\mathcal{E}(t,u) := W(u) - \ell(t)u, \quad \mathsf{d}(u,v) := \sum_{\pm} \alpha_\pm (v-u)_\pm, \quad \delta(u,v) := \frac{\mu}{2}|u-v|^2. \tag{4.7}$$

Since we are in the simplified setting recalled at the beginning of Section 4, it is easy to check that all the assumptions $\langle \mathrm{A} \rangle$, $\langle \mathrm{B} \rangle$, $\langle \mathrm{C} \rangle$ hold. A careful analysis (see [40]) shows that when ℓ is strictly increasing, the initial datum u_0 satisfies a suitable stability condition and $\mu\alpha_+^2 > \lambda$ then $u \in \mathrm{BV}([0,T];\mathbb{R})$ is a VE solution of $(X, \mathcal{E}, \mathsf{d}, \delta)$ if and only if it is nondecreasing in $[0,T]$ and

$$W'(u(t)) = \ell(t) - \alpha_+, \tag{4.8}$$

so that the evolution of u can be described in terms of the upper monotone envelope of W' starting from u_0, as in the case of Balanced Viscosity solutions, see [45] and Figure 3 in the Introduction. When ℓ is strictly decreasing then u should be non-increasing and (4.8) should be replaced by $W'(u(t)) = \ell(t) + \alpha_-$.

In the case when $0 < \mu\alpha_+^2 < \lambda$ we have a sort of intermediate behaviour between the previous situation and the energetic case, corresponding to $\mu = 0$ where increasing jumps between $u(t-) < u(t+)$ obey the Maxwell rule

$$\int_{u(t-)}^{u(t+)} \Big(W'(r) - \ell(t) + \alpha_+ \Big) \,\mathrm{d}r = 0.$$

In particular, in the visco-energetic case, an increasing jump occurs at t when we have the modified Maxwell rule

$$\int_{u(t-)}^{u_+} \Big(W'(r) - \ell(t) + \alpha_+ + \mu(r - u(t-)) \Big) \,\mathrm{d}r = 0 \quad \text{for some } u_+ > u(t-). \tag{4.9}$$

In this case, however, $u(t+)$ may differ from u_+, see Figure 4.2: we refer to [40] for a detailed analysis.

4.3 The choice of δ: α-Λ geodesic convexity.

In some situations it could be interesting to choose a viscous correction δ associated with a metric different from d: we want to show a typical example where $\langle \mathrm{B.3} \rangle$ still holds and a related application to the evolution of the Allen-Cahn energy.

Let us consider for simplicity

$$\begin{gathered} \text{the metric setting of Remark 2.2 with } \delta(u,v) := \frac{1}{2}\mathsf{d}_*^2(u,v), \\ \text{where } \mathsf{d}_* \text{ is another distance on } X \text{, continuous on each sublevel of } \mathcal{F}_0, \end{gathered} \tag{4.10}$$

so that $\langle \mathrm{B.1} \rangle$-$\langle \mathrm{B.2} \rangle$ and $\langle \mathrm{C.1} \rangle$-$\langle \mathrm{C.2} \rangle$ hold.

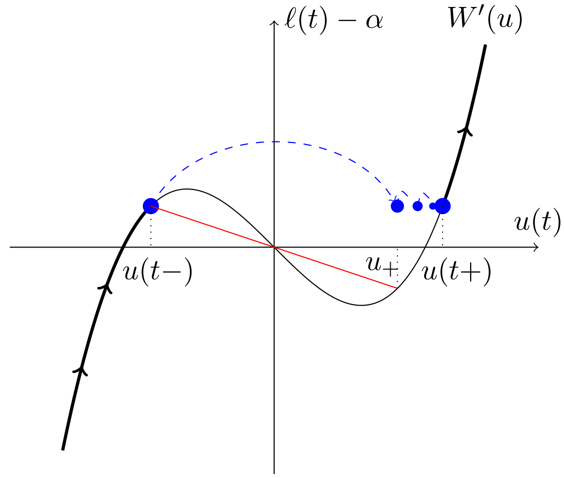

Figure 4: Visco-Energetic solutions for a double-well energy W with an increasing load ℓ and $0 < \mu\alpha^2 < -\min W''$. In this case the solution u jumps before reaching the local maximum of W and the optimal transition ϑ makes a first jump connecting $u(t-)$ with u_+ according to the modified Maxwell rule of (4.9): $u(t-)$ and u_+ corresponds to the intersection of the graph of W' with the red line, whose slope is $-\mu$. After the first jump, ϑ makes an infinite sequence of jumps accumulating to $u(t+)$.

Definition 4.2 (α-Λ convexity) *Let $\alpha > 0$, $\Lambda \geq 0$. We say that $(\mathcal{E}, \mathsf{d}, \mathsf{d}_*)$ satisfies the* weak *α-Λ convexity property on a set $S \subset X$ if for every $x, y \in S$ there exists a curve $\gamma : [0,1] \to X$ such that*

$$\mathcal{E}(t, \gamma(\theta)) \leq (1-\theta)\mathcal{E}(t,x) + \theta\mathcal{E}(t,y) - \frac{1}{2}\theta(1-\theta)\Big[\alpha\mathsf{d}_*^2(x,y) - \Lambda\mathsf{d}(x,y)\mathsf{d}_*(x,y)\Big], \tag{4.11}$$

$$\liminf_{\theta\downarrow 0} \frac{\mathsf{d}(x,\gamma(\theta))}{\theta} \leq \mathsf{d}(x,y), \quad \lim_{\theta\downarrow 0} \frac{\mathsf{d}_*(x,\gamma(\theta))}{\sqrt{\theta}} = 0. \tag{4.12}$$

We say that $(\mathcal{E}, \mathsf{d}, \mathsf{d}_*)$ satisfies the strong *α-Λ convexity property if for every $x, y \in X$ there exists a curve $\gamma : [0,1] \to X$ connecting x to y satisfying (4.11) and*

$$\mathsf{d}(\gamma(\theta), \gamma(\theta')) = |\theta - \theta'|\mathsf{d}(x,y), \quad \mathsf{d}_*(\gamma(\theta), \gamma(\theta')) = |\theta - \theta'|\mathsf{d}_*(x,y) \tag{4.13}$$

for every $\theta, \theta' \in [0,1]$.

Observe that (4.11) is a generalization of the λ-convexity along geodesics, involving two distances: see [30].

Let us show that if $(\mathcal{E}, \mathsf{d}, \mathsf{d}_*)$ satisfies the weak α-Λ convexity on $S := \cup_{t\in[0,T]}\mathscr{S}_{\mathsf{D}}(t)$ then $\langle$B.3$\rangle$ holds. In fact, if $x \in \mathscr{S}_{\mathsf{D}}(t), y \in \mathscr{S}_{\mathsf{D}}(s)$ and γ satisfies (4.11)-(4.12), then

$$\begin{aligned}
\mathcal{E}(t,x) &\overset{(3.4)}{\leq} \mathcal{E}(t,\gamma(\theta)) + \mathsf{d}(x,\gamma(\theta)) + \frac{1}{2}\mathsf{d}_*^2(x,\gamma(\theta)) \\
&\overset{(4.11)}{\leq} (1-\theta)\mathcal{E}(t,x) + \theta\mathcal{E}(t,y) - \frac{\theta(1-\theta)}{2}\mathsf{d}_*(x,y)\Big[\alpha\mathsf{d}_*(x,y) - \Lambda\mathsf{d}(x,y)\Big] + \mathsf{d}(x,\gamma(\theta)) + \frac{1}{2}\mathsf{d}_*^2(x,\gamma(\theta)).
\end{aligned}$$

Subtracting $(1-\theta)\mathcal{E}(t,x)$ and dividing by θ we obtain

$$\mathcal{E}(t,x) \leq \mathcal{E}(t,y) + \frac{\mathsf{d}(x,\gamma(\theta))}{\theta} + \frac{1}{2\theta}\mathsf{d}_*^2(x,\gamma(\theta)) - \frac{1}{2}(1-\theta)\Big[\alpha\mathsf{d}_*^2(x,y) - \Lambda\mathsf{d}(x,y)\mathsf{d}_*(x,y)\Big].$$

Passing to the limit as $\theta \downarrow 0$ and using (4.12) we get

$$\mathcal{E}(t,x) - \mathcal{E}(t,y) - \mathsf{d}(x,y) \le -\frac{\alpha}{2}\mathsf{d}_*^2(x,y) + \frac{\Lambda}{2}\mathsf{d}(x,y)\mathsf{d}_*(x,y) \le \frac{\Lambda^2}{8\alpha}\mathsf{d}^2(x,y). \tag{4.14}$$

To recover $\langle$B.3$\rangle$ is enough to divide by $\mathsf{d}(x,y)$ and to pass to the limit as $x \to y$.

As a further consequence of the above conditions we can also prove an enhanced BV estimate, which is related to a coercivity property of $\mathcal{R}$, see Lemma 7.6. The proof will be collected in the last section 7.5.

Theorem 4.3 (BV estimates w.r.t. d_*) *Let us assume that* $\langle$A$\rangle$ *holds and* $(\mathcal{E}, \mathsf{d}, \mathsf{d}_*)$ *satisfies the* strong *α-Λ convexity property. If*

$$|\mathcal{P}(t,x) - \mathcal{P}(t,y)| \le L\mathsf{d}_*(x,y) \quad \textit{if } t \in [0,T] \textit{ and } x,y \in X, \tag{4.15}$$

and (3.30) *holds, then any* VE *solution u obtained as a pointwise limit of the time incremental minimization scheme* $(\mathrm{IM}_{\mathsf{d},\delta})$ *belongs to* $\mathrm{BV}_{\mathsf{d}_*}([0,T];X)$.

Example 4.4 (VE evolution for the Allen-Cahn functional) Let us consider a bounded open Lipschitz domain $\Omega \subset \mathbb{R}^d$, a function $W \in \mathrm{C}^2(\mathbb{R})$ as in the previous Example 4.2 and let us set $X = \{u \in W^{1,2}_0(\Omega) : W(u) \in L^1(\Omega)\}$ endowed with the $L^1(\Omega)$-topology.

The distance d is the usual one induced by the L^1 norm, while δ is the squared distance induced by the L^2 norm.

$$\mathsf{d}(u,v) := \int_\Omega |u(x) - v(x)|\mathrm{d}x, \qquad \delta(u,v) := \frac{\mu}{2}\int_\Omega |u(x) - v(x)|^2\mathrm{d}x.$$

We also consider the energy functional

$$\mathcal{E}(t,u) = \begin{cases} \displaystyle\int_\Omega \left(\frac{1}{2}|\nabla u|^2 + W(u) - \ell(t)u\right)\mathrm{d}x & \text{if } u \in \mathrm{W}^{1,2}_0(\Omega); \\ +\infty & \text{otherwise,} \end{cases} \tag{4.16}$$

where $\ell \in C^1([0,T];L^2(\Omega))$. It is immediate to check that for all $u \in X$ the function $t \mapsto \mathcal{E}(t,u)$ is differentiable, with derivative

$$\mathcal{P}(t,u) = -\int_\Omega \ell'(t)u\,\mathrm{d}x$$

so that Assumptions $\langle$A$\rangle$ are satisfied since the sublevels of the energy are compact in $L^2(\Omega)$. Thus we are in the canonical metric setting and the only nontrivial assumption is $\langle$B.3$\rangle$ since δ is continuous on the sublevels of the energy. We will check that the α-Λ convexity discussed in (4.3) is satisfied. If W is λ-convex with $\lambda > 0$, we can use the estimate

$$\mathcal{E}(t,(1-\theta)u + \theta v) \le (1-\theta)\mathcal{E}(t,u) + \theta\mathcal{E}(t,v) - \frac{\theta(1-\theta)}{2}\left(\|\nabla(u-v)\|^2_{L^2(\Omega)} + \lambda\|u-v\|^2_{L^2(\Omega)}\right), \tag{4.17}$$

hence we have (4.11) with $\alpha = \lambda$ and $\Lambda = 0$. If $\lambda < 0$ we use the estimate (see [30, Example 5.1])

$$-\|\nabla(u-v)\|^2_{L^2(\Omega)} \le -(1+|\lambda|)\|u-v\|^2_{L^2(\Omega)} + M_\lambda\|u-v\|^2_{L^1(\Omega)}$$

for some $M_\lambda > 0$. Inserting this into (4.17) we obtain the generalized convexity (4.11) with $\frac{\alpha}{2} = (1+|\lambda|) + \lambda = 1 > 0$ and $\frac{\Lambda}{2} = (1+|\lambda|)M_\lambda$ and then also $\langle$B.1$\rangle$ is satisfied. We can therefore apply Theorem 3.9 and prove the existence of a Visco-Energetic solution for the rate-independent system $(X, \mathcal{E}, \mathsf{d}, \delta)$.

4.4 Product spaces and degenerate-singular distances

In many important examples the space X is a cartesian product $X = F \times Z$ (whose points can be written as $u = (\varphi, z)$, $\varphi \in F$, $z \in Z$) but d only depends on the z-component

$$\mathsf{d}(u, u') := \widetilde{\mathsf{d}}(z, z') \quad \text{if } u = (\varphi, z),\ u' = (\varphi', z'), \tag{4.18}$$

for a quasi-distance $\widetilde{\mathsf{d}}$ separating Z. In these cases it is natural to consider a viscous correction $\delta(u, u') = \widetilde{\delta}(z, z')$ which still depends only on z (but more general interesting situations can occur, see e.g. [11] or [17, 42] where an alternate minimization scheme has been studied): therefore, even if $\widetilde{\mathsf{d}}$ separates Z, the distance d does not separate X.

It may happen that for every $z \in Z$ the set

$$\Phi(t, z) := \operatorname*{argmin}_F \mathcal{E}(t, \cdot, z) \tag{4.19}$$

contains only one point. Since δ and D do not depend on φ, one can easily check that

$$(\varphi, z) \in \mathscr{S}_{\mathsf{D}}(t) \quad \Rightarrow \quad \varphi \in \Phi(t, z), \tag{4.20}$$

and $\mathsf{d}_{\mathbb{R}}$ separates $\mathscr{S}_{\mathsf{D}}$. As an example, we consider the following model discussed in [22, Sect. 6.2] (we refer to [22] and [41] for the interpretation and more details).

Example 4.5 (A delamination problem) Let O be a sufficiently regular open connected domain of $\mathbb{R}^d$, $\Gamma_{\rm dir} \subset \partial O$ with positive surface measure and let $\Gamma \subset O$ be a piecewise smooth hypersurface, such that $\Omega := O \setminus \Gamma$ is still connected. Let $\phi_{\rm dir} \in H^1(\Omega; \mathbb{R}^d)$ and $F := \big\{\varphi \in H^1(\Omega; \mathbb{R}^d) : \varphi = \phi_{\rm dir} \text{ on } \Gamma_{\rm dir}\big\}$ endowed with the weak topology σ_F of H^1 and we set $Z := L^\infty(\Gamma; [0, 1])$ endowed with the weak* topology σ_Z.

We thus define

$$\mathcal{E}(t, \varphi, z) := \int_\Omega \mathsf{W}(\mathrm{D}\varphi(x))\,\mathrm{d}x + \int_\Gamma z(x)\,\mathsf{Q}([\![\varphi]\!](x))\,\mathrm{d}\mathcal{H}^{d-1}(x) - \langle \ell(t), \varphi\rangle \tag{4.21}$$

where $\mathsf{W} : \mathbb{R}^{d\times d} \to \mathbb{R}$ is the quadratic form of linearized elasticity, $\mathsf{Q} : \mathbb{R}^d \to \mathbb{R}^d$ is a nonnegative quadratic form, $[\![\varphi]\!] \in H^{1/2}(\Gamma; \mathbb{R}^d)$ denotes the jump of the deformation of φ across Γ, and $\ell \in \mathrm{C}^1([0, T]; (H^1(\Omega))')$. We eventually introduce the dissipation $\mathsf{d}((\varphi, z), (\varphi', z')) := \widetilde{\mathsf{d}}(z, z')$ with

$$\widetilde{\mathsf{d}}(z, z') := \int_\Gamma \psi(z'(x) - z(x))\,\mathrm{d}\mathcal{H}^{d-1}(x), \quad \text{where} \quad \psi(r) := \begin{cases} r & \text{if } r \geq 0 \\ +\infty & \text{otherwise,} \end{cases} \tag{4.22}$$

and the viscous correction $\delta((\varphi, z), (\varphi', z')) = h(\widetilde{\mathsf{d}}(z, z'))$ as in (3.11).

Arguing as in [22] and taking into account Example 3.3 it is easy to check that $\langle \mathrm{A}\rangle$ and $\langle \mathrm{B}\rangle$ are satisfied; also the separation property $\langle \mathrm{C.2}\rangle$ follows by the above remarks since the set $\Phi(t, z)$ defined by (4.19) contains only one element.

The only property that remains to be checked is the closure of the (D, Q)-quasi stable set $\langle \mathrm{C.1}\rangle$. By (C.1') we can apply Lemma 3.11 vi): if $(\varphi_n, z_n) \xrightarrow{\sigma} (\varphi, z)$ in $F \times Z$ with $\mathcal{E}(t, \varphi_n, z_n) \to \mathcal{E}(t, \varphi, z) + \eta$ and (φ', z') is a minimizer of $\mathcal{E}(t, \cdot) + \mathsf{D}((\varphi, z), \cdot)$ in $\mathrm{M}(t, (\varphi, z))$, we have $\varphi' \in \Phi(t, z')$ and $z' \geq z$, so that we can apply [22, Lemma 6.1] to find another sequence $z'_n \in Z$ satisfying $z'_n \leq z_n$, $z'_n \xrightarrow{\sigma_Z} z'$ in Z and $\widetilde{\mathsf{d}}(z_n, z'_n) \to \widetilde{\mathsf{d}}(z, z')$; this also implies that $\widetilde{\delta}(z_n, z'_n) \to \widetilde{\delta}(z, z')$. Correspondingly, we set $\varphi'_n := \Phi(t, z'_n)$ (with a slight abuse of notation, we still denote by $\Phi(t,z)$ the unique element of the set). Since the maps $z \mapsto \Phi(t,z)$ and $z \mapsto \mathcal{E}(t,\Phi(t,z),z)$ are continuous (see [22]) with respect to the topology of Z we deduce that $\mathcal{E}(t,\varphi'_n,z'_n) \to \mathcal{E}(t,\varphi',z')$. We conclude that (3.38) is satisfied since

$$\begin{aligned}\liminf_{n\to\infty}\Big(\mathcal{E}(t,\varphi'_n,z'_n)+\mathsf{D}((\varphi_n,z_n),(\varphi'_n,z'_n))\Big) &= \mathcal{E}(t,\varphi',z')+\mathsf{D}((\varphi,z),(\varphi',z'))\\ &\le \mathcal{E}(t,\varphi',z')+\mathsf{D}((\varphi,z),(\varphi',z'))+\eta. \quad \square\end{aligned}$$

4.5 Marginal energies

In the same cartesian setting of the previous Section, 4.4 let us now consider the case when the set $\Phi(t,z)$ of (4.19) contains more than one element. One can try to write a reduced model in the space Z by introducing the marginal energy functionals and its generalized power

$$\widetilde{\mathcal{E}}(t,z) := \min\Big\{\mathcal{E}(t,\varphi,z) : \varphi\in F\Big\}, \quad \widetilde{\mathcal{P}}(t,u) := \max\Big\{\mathcal{P}(t,\varphi,z) : \varphi\in\Phi(t,z)\Big\}. \tag{4.23}$$

If $\mathcal{E}$ satisfies $\langle\mathrm{A}\rangle$ one can easily prove that $\Phi(t,z)$ is compact in F for every t,z and

$$(t_n,z_n)\to(t,z), \quad \widetilde{\mathcal{E}}(t,z_n)\le C \quad\Rightarrow\quad \mathrm{Ls}_{n\to\infty}\,\Phi(t,z_n)\subset\Phi(t,z), \tag{4.24}$$

where Ls denotes the Kuratowski superior limit, see Definition 5.2. The following Lemma allows to easily check conditions $\langle\mathrm{A}\rangle$.

Lemma 4.6 *If the functionals* $\mathcal{E},\mathcal{P}$ *satisfy* Assumptions $\langle\mathrm{A.1}\rangle$, $\langle\mathrm{A.2}\rangle$ (resp. $\langle\mathrm{A.2'}\rangle$) *in* (X,σ,d) *then* $(\widetilde{\mathcal{E}},\widetilde{\mathcal{P}})$ *satisfy* Assumptions $\langle\mathrm{A.1}\rangle$, $\langle\mathrm{A.2}\rangle$ (resp. $\langle\mathrm{A.2'}\rangle$) *in* $(Z,\sigma_Z,\widetilde{\mathsf{d}})$.

Proof. Notice that if (t,z) belongs to the sublevel $\big\{\widetilde{\mathcal{F}}\le C\big\}$ then (t,φ,z) belongs to $\big\{\mathcal{F}\le C\big\}$ for every $\varphi\in\Phi(t,z)$. Property $\langle\mathrm{A.1}\rangle$ is easy to verify, so we consider $\langle\mathrm{A.2}\rangle$.

The upper semicontinuity of $\widetilde{\mathcal{P}}$ follows immediately by (4.24): selecting $\varphi_n\in\Phi(t_n,z_n)$ so that $\widetilde{\mathcal{P}}(t_n,z_n)=\mathcal{P}(t_n,\varphi_n,z_n)$ and observing that (t_n,φ_n,z_n) belong to a sublevel of $\mathcal{F}$, we can suppose that φ_n converges to some $\varphi\in\Phi(t,z)$ so that the upper semicontinuity of $\mathcal{P}$ yields

$$\limsup_{n\to\infty}\widetilde{\mathcal{P}}(t_n,z_n)=\limsup_{n\to\infty}\mathcal{P}(t_n,\varphi_n,z_n)\le\mathcal{P}(t,\varphi,z)\le\widetilde{\mathcal{P}}(t,z).$$

In the case of $\langle\mathrm{A.2'}\rangle$, we observe that if $\widetilde{\mathcal{E}}(t_n,z_n)\to\widetilde{\mathcal{E}}(t,z)$ and $\Phi(t_n,z_n)\ni\varphi_n\to\varphi$ as in the above argument, we have

$$\widetilde{\mathcal{E}}(t,z)\le\mathcal{E}(t,\varphi,z)\le\liminf_{n\to\infty}\mathcal{E}(t_n,\varphi_n,z_n)=\liminf_{n\to\infty}\widetilde{\mathcal{E}}(t_n,z_n)=\widetilde{\mathcal{E}}(t,z)$$

so that $\mathcal{E}(t_n,\varphi_n,u_n)\to\mathcal{E}(t,\varphi,u)$ and we can apply the conditional upper semicontinuity of $\mathcal{P}$.

Concerning (2.20) we observe that for some $\varphi\in\Phi(t,z)$

$$|\widetilde{\mathcal{P}}(t,z)|=|\mathcal{P}(t,\varphi,z)|\le C_P\mathcal{F}(t,\varphi,z)=C_P\widetilde{\mathcal{F}}(t,z).$$

As for (2.19), since

$$\liminf_{s\uparrow t}\frac{\widetilde{\mathcal{E}}(t,z)-\widetilde{\mathcal{E}}(s,z)}{t-s}\ge\liminf_{s\uparrow t}\frac{\mathcal{E}(t,\varphi,z)-\mathcal{E}(s,\varphi,z)}{t-s}\ge\mathcal{P}(t,\varphi,z) \quad\text{for every }\varphi\in\Phi(t,z),$$

so that

$$\liminf_{s\uparrow t} \frac{\widetilde{\mathcal{E}}(t,z)-\widetilde{\mathcal{E}}(s,z)}{t-s} \ge \widetilde{\mathcal{P}}(t,z);$$

the corresponding right lim sup inequality of (2.19) follows by the same argument. $\square$

Let us consider for the sake of simplicity the case when $\widetilde{\mathsf{d}}$ is left continuous and $\widetilde{\delta} = h(\widetilde{\mathsf{d}})$.

Theorem 4.7 *Let us suppose that the energy functionals* $\mathcal{E}, \mathcal{P}$ *satisfy* Assumptions $\langle$A.1$\rangle$, $\langle$A.2'$\rangle$, $\widetilde{\mathsf{d}}$ *separates* Z *and* $\widetilde{\delta} = h(\widetilde{\mathsf{d}})$ *as in* (3.11). *Then for every* $z_0 \in Z$ *there exists a* VE *solution to the R.I.S.* $(Z, \widetilde{\mathcal{E}}, \widetilde{\mathsf{d}}, \widetilde{\delta})$. *Equivalently, there exist a map* $z \in \mathrm{BV}_{\sigma_Z,\widetilde{\mathsf{d}}}([0,T];Z)$ *and a map* $\varphi : [0,T] \to F$ *(which is measurable, if* F *is Souslin) such that* $\varphi(t) \in \Phi(t, z(t))$ *for every* $t \in [0,T]$*,*

$$\mathcal{E}(t,\varphi(t),z(t)) \le \mathcal{E}(t,\varphi',z') + \widetilde{\mathsf{d}}(z(t),z') + \widetilde{\delta}(z(t),z') \quad \textit{for every } t \in [0,T] \setminus \mathrm{J}_z, \tag{4.25}$$

$$\mathcal{E}(t,\varphi(t),z(t)) + \mathit{Var}_{\widetilde{\mathsf{d}},\widetilde{\mathsf{c}}}(z,[s,t]) = \mathcal{E}(s,\varphi(s),z(s)) + \int_s^t \mathcal{P}(r,\varphi(r),z(r))\,\mathrm{d}r. \tag{4.26}$$

The proof is immediate by applying Theorem 3.9 to the R.I.S. $(Z, \widetilde{\mathcal{E}}, \widetilde{\mathsf{d}}, \widetilde{\delta})$ and recalling the remarks stated at the beginning of Section 4. We then select

$$\varphi(t) \in \Phi(t, z(t)) \quad \text{such that } \mathcal{P}(t,\varphi(t),z(t)) = \widetilde{\mathcal{P}}(t,z(t)); \tag{4.27}$$

by the Von Neumann-Aumann selection Theorem [6, Section III.6] φ can also be supposed to be measurable, if F is a Souslin space (in particular, if σ_F is metrizable, since the sets $\Phi(t,z(t))$ are contained in a compact set).

An interesting application of the above result concerns a material model driven by a nonconvex elastic energy, discussed in [14, Sect. 4] in the framework of energetic evolutions.

Example 4.8 (A material model with a nonconvex elastic energy) We consider a Lipschitz and bounded open set $\Omega \subset \mathbb{R}^d$, a compact set $K \subset \mathbb{R}^m$, two exponents $\alpha, p > 1$ and two maps

$$\varphi_{\mathrm{dir}} \in \mathrm{C}^1([0,T]; W^{1,p}(\Omega)), \quad \ell \in \mathrm{C}^1([0,T](W^{1,p}(\Omega))'). \tag{4.28}$$

The spaces F and Z are defined by

$$F := W^{1,p}_0(\Omega), \quad Z := \Big\{ z \in W^{1,\alpha}(\Omega;\mathbb{R}^m) : z(x) \in K \Big\} \tag{4.29}$$

endowed with their weak topologies and the energy functional is

$$\mathcal{E}(t,\varphi,z) := \int_\Omega W(\mathrm{D}\varphi(x) + \mathrm{D}\varphi_{\mathrm{dir}}(t,x), z(x))\,\mathrm{d}x + \lambda \int_\Omega |\mathrm{D}z(x)|^\alpha\,\mathrm{d}x - \langle \ell, \varphi + \varphi_{\mathrm{dir}} \rangle, \tag{4.30}$$

where $W \in \mathrm{C}(\mathbb{R}^{d\times d} \times K; \mathbb{R}_+)$ is C^1 and quasiconvex with respect to its first variable and satisfies

$$c|D|^p - C \le W(D,z) \le C(1+|D|^p) \quad \text{for every } D \in \mathbb{R}^{d\times d},\ z \in K \tag{4.31}$$

for some constants $0 < c < C < \infty$. $\widetilde{\mathsf{d}}$ is an asymmetric distance on Z satisfying

$$C^{-1}\|z - z'\|_{L^1} \le \widetilde{\mathsf{d}}(z,z') \le C\|z-z'\|_{L^1} \quad \text{for every } z, z' \in Z. \tag{4.32}$$

Notice that in this case

$$\mathcal{P}(t,\varphi,z)=\int_\Omega \mathrm{D}W(\varphi+\varphi_{\mathrm{dir}}(t))\cdot \mathrm{D}\partial_t\varphi_{\mathrm{dir}}(t)\,\mathrm{d}x-\langle\partial_t\ell(t),\varphi+\varphi_{\mathrm{dir}}(t)\rangle-\langle\ell(t),\partial_t\varphi_{\mathrm{dir}}(t)\rangle, \quad (4.33)$$

satisfies the assumptions stated in ⟨A.2'⟩ thanks to an argument of [10], see [14, Prop. 4.4].

By choosing a viscous correction as in (3.11) we can therefore apply Theorem 4.7 and prove the existence of a VE solution. We refer to [41] for more details.

5 Main structural properties of the viscous dissipation cost

In this section we will prove some relevant properties of the viscous transition and dissipation costs $\mathrm{Trc}(t,\vartheta,E)$ and $\mathsf{c}(t,u_0,u_1)$ introduced in Definition 3.5 and 3.6. They lie at the core of the structure of Visco-Energetic solutions and of our existence proof.

5.1 Additivity and Invariance by rescaling

A first simple fact concerns the possibility of performing suitable rescaling of the domain E of a transition $\vartheta : E\to X$ without affecting the cost. This is related to the following additivity property of $\mathrm{Var}_{\mathsf{d}}$ and GapVar_δ: for every $a,b,c\in E$ with $a<b<c$ we have

$$\begin{aligned}\mathrm{Var}_{\mathsf{d}}(\vartheta,E\cap[a,c])&=\mathrm{Var}_{\mathsf{d}}(\vartheta,E\cap[a,b])+\mathrm{Var}_{\mathsf{d}}(\vartheta,E\cap[b,c]),\\ \mathrm{GapVar}_\delta(\vartheta,E\cap[a,c])&=\mathrm{GapVar}_\delta(\vartheta,E\cap[a,b])+\mathrm{GapVar}_\delta(\vartheta,E\cap[b,c]).\end{aligned}\quad (5.1)$$

Lemma 5.1 *Let $E\subset\mathbb{R}$ compact and $\vartheta\in \mathrm{C}_{\sigma,\mathsf{d}}(E,X)$ with*

$$\mathrm{Var}_{\mathsf{d}}(\vartheta,E)+\mathrm{GapVar}_\delta(\vartheta,E)=C. \quad (5.2)$$

There exists a compact set $\tilde E$ with $\tilde E^-=0,\tilde E^+=C+1$ and a bijective Lipschitz map $\mathsf{s}:\tilde E\to E$ such that the new transition $\tilde\vartheta:=\vartheta\circ\mathsf{s}$ satisfies

$$\mathrm{Var}_{\mathsf{d}}(\tilde\vartheta,\tilde E\cap[r_0,r_1])+\mathrm{GapVar}_\delta(\tilde\vartheta,\tilde E\cap[r_0,r_1])\le|r_0-r_1|\quad \textit{for every } r_0,r_1\in\tilde E,\ r_0<r_1, \quad (5.3)$$

$$\mathrm{Var}_{\mathsf{d}}(\vartheta,E)=\mathrm{Var}_{\mathsf{d}}(\tilde\vartheta,\tilde E),\qquad \mathrm{GapVar}_\delta(\vartheta,E)=\mathrm{GapVar}_\delta(\tilde\vartheta,\tilde E). \quad (5.4)$$

Moreover, for every $\mathsf{t}:E\to[0,T]$ setting $\tilde{\mathsf{t}}:=\mathsf{t}\circ\mathsf{s}$ we have

$$\sum_{s\in E\setminus\{E^+\}}\mathcal{R}(\mathsf{t}(s),\vartheta(s))=\sum_{r\in\tilde E\setminus\{\tilde E^+\}}\mathcal{R}(\tilde{\mathsf{t}}(r),\tilde\vartheta(r)). \quad (5.5)$$

Proof. We define $\mathsf{r}:E\to[0,C+1]$ by

$$\mathsf{r}(s):=\frac{s-E^-}{E^+-E^-}+\mathrm{Var}_{\mathsf{d}}(\vartheta,E\cap[E^-,s])+\mathrm{GapVar}_\delta(\vartheta,E\cap[E^-,s]); \quad (5.6)$$

it is not too difficult to check that r is continuous and strictly increasing so that $\tilde E:=\mathsf{r}(E)$ is compact. Moreover

$$|s_1-s_0|\le|E^+-E^-||\mathsf{r}(s_1)-\mathsf{r}(s_0)|\quad \text{for every } s_0,s_1\in[E^-,E^+], \quad (5.7)$$

so that r admits a Lipschitz continuous inverse s defined in $[0,C+1]$, which satisfies (5.3) by construction, thanks to (5.1), and (5.4)-(5.5). □

5.2 Lower semicontinuity of the transition and dissipation cost

Since the viscous transition cost c involve curves $\vartheta : E \to X$ defined in general compact parametrization domains $E \subset \mathbb{R}$, it will be crucial to study its lower semicontinuity along sequence of transition curves ϑ_k defined in *varying domains* E_k.

Let us first recall the notion of convergence in the sense of Kuratowski in a Hausdorff topological space (Y, ρ) satisfying the first axiom of countability.

Definition 5.2 (Kuratowski convergence) *Let $(A_k)_k$ be a sequence of subsets of Y. The Kuratowski limit inferior (resp. limit superior) of A_k, as $k \to \infty$ are defined by:*

$$\operatorname*{Li}_{k\to\infty} A_k := \left\{ a \in Y : \exists\, a_k \in A_k \textit{ such that } a_k \xrightarrow{\rho} a \right\}, \tag{5.8}$$

$$\operatorname*{Ls}_{k\to\infty} A_k := \left\{ a \in Y : \exists\, n \mapsto k_n \textit{ increasing, and } \ a_{k_n} \in A_{k_n} : a_{k_n} \xrightarrow{\rho} a \right\}. \tag{5.9}$$

We say that $A_k \xrightarrow{K} A$ in the Kuratowski sense if $A = \operatorname*{Li}_{k\to\infty} A_k = \operatorname*{Ls}_{k\to\infty} A_k$.

Recall that Kuratowski convergence coincides with Γ-convergence of the indicator functions $i_k := i_{A_k}$ associated with the sets A_k [7, Chapter 4], where in general

$$i_A(x) := \begin{cases} 0 & \text{if } x \in A, \\ +\infty & \text{if } x \notin A. \end{cases} \tag{5.10}$$

Whenever Y is a metric space and A_k, A are compact sets, then Kuratowski convergence coincides with the convergence induced by the Hausdorff distance.

Let us now consider a sequence $\vartheta_k \in \mathrm{C}(E_k, X)$, where E_k is compact subset of $\mathbb{R}$. In order to study the asymptotic behaviour of ϑ_k to some limit curve $\vartheta \in \mathrm{C}(E, X)$ we can simply consider the Kuratowski convergence of the graphs $\mathrm{graph}(\vartheta_k)$ to $\mathrm{graph}(\vartheta)$ in $\mathbb{R} \times X$ (see e.g. [20]). Notice that

$$\mathrm{graph}(\vartheta) \subset \mathrm{Li}_{k\uparrow\infty}\, \mathrm{graph}(\vartheta_k) \quad \Leftrightarrow \quad \forall\, s \in E\ \exists\, s_k \in E_k : \quad s_k \to s,\ \vartheta_k(s_k) \to \vartheta(s). \tag{5.11}$$

This weak condition is sufficient to prove the lower semicontinuity of the function $\mathrm{Trc}(t, \vartheta, E)$ as stated in the next Theorem, which also covers a slightly more general situation that will turn out to be useful in what follows.

Theorem 5.3 (Lower semicontinuity of Trc**)** *Let $\vartheta \in \mathrm{C}(E, X)$, $t \in \mathbb{R}$, and let $\vartheta_k \in \mathrm{C}(E_k, X)$, $\mathsf{t}_k : E_k \to \mathbb{R}$, $k \in \mathbb{N}$, be sequences of functions satisfying* (5.11). *We have the following lower semicontinuity properties.*
a)

$$\mathrm{Var}_{\mathsf{d}}(\vartheta, E) \le \liminf_{k\uparrow\infty} \mathrm{Var}_{\mathsf{d}}(\vartheta_k, E_k). \tag{5.12}$$

b) If $\langle \mathrm{C}.1 \rangle$ *holds and*

$$\lim_{k\to\infty} \sup_{s\in E_k} |\mathsf{t}_k(s) - t| = 0, \quad \vartheta_k(E_k) \subset F, \textit{ where } F \textit{ is a sublevel of } \mathcal{F}_0, \tag{5.13}$$

then for every $t \in [0, T]$

$$\sum_{s \in E \setminus \{E^+\}} \mathcal{R}(t, \vartheta(s)) \le \liminf_{k\uparrow\infty} \sum_{s \in E_k \setminus \{E_k +\}} \mathcal{R}(\mathsf{t}_k(s), \vartheta_k(s)) \tag{5.14}$$

c) If there exists a modulus of continuity $\omega : [0,\infty) \to [0,\infty)$ with $\omega(0) = 0$ such that

$$\mathsf{d}(\vartheta_k(x), \vartheta_k(y)) \le \omega(y - x) \quad \textit{for every } k \in \mathbb{N} \textit{ and } x, y \in E_k,\ x \le y, \tag{5.15}$$

and δ satisfies $\langle$B.1$\rangle$, then

$$\mathrm{GapVar}_\delta(\vartheta, E) \le \liminf_{k\uparrow\infty} \mathrm{GapVar}_\delta(\vartheta_k, E_k). \tag{5.16}$$

d) If $\langle$B.1$\rangle$, $\langle$C.1$\rangle$, (5.13) and (5.15) hold, then

$$\mathrm{Trc}(t, \vartheta, E) \le \liminf_{k\uparrow\infty} \mathrm{Trc}(t, \vartheta_k, E_k). \tag{5.17}$$

Proof. Clearly d) is a consequence of the first three properties a), b), and c). Let us prove each of them.

Lower semicontinuity of the total variation. Let $E^- = s^0 < s^1 < \cdots < s^N = E^+$ be a finite subset of E. By (5.11), for every s^j there exists a sequence $s^j_k \in E_k$ such that $s^j_k \to s^j$ and $\vartheta_k(s^j_k) \to \vartheta(s^j)$. If k is big enough, we can also assume $s^{j-1}_k \le s^j_k$ for every j and then

$$\sum_{j=1}^N \mathsf{d}(\vartheta_k(s^{j-1}_k), \vartheta_k(s^j_k)) \le \mathrm{Var}_{\mathsf{d}}(\vartheta_k, E_k).$$

In particular, taking the liminf and recalling that d is lower semicontinuous we obtain

$$\sum_{j=1}^N \mathsf{d}(\vartheta(s^{j-1}), \vartheta(s^j)) \le \liminf_k \mathrm{Var}_{\mathsf{d}}(\vartheta_k). \tag{5.18}$$

Since (5.18) holds for every choice of $\{s^1, \cdots, s^N\} \subset E$, by taking the supremum among all the finite subsets of E we obtain (5.12).

Semicontinuity of the residual sum. Since the viscous residual functional $\mathcal{R}(t, \cdot)$ is σ-lower semicontinuos on $[0,T] \times F$ and positive, we can argue as in the previous step:

$$\sum_{j=0}^{N-1} \mathcal{R}(t, \vartheta(s^j)) \le \liminf_{k\uparrow\infty} \sum_{j=0}^{N-1} \mathcal{R}(\mathsf{t}_k(s^j_k), \vartheta_k(s^j_k)) \le \liminf_{k\uparrow\infty} \sum_{s \in E_k \setminus \{E_k +\}} \mathcal{R}(\mathsf{t}_k(s), \vartheta_k(s)).$$

(5.14) then follows by taking the supremum of the left hand side.

Lower semicontinuity of GapVar_δ**.** We first prove the following property:

$$\text{for every } I \in \mathfrak{H}(E)\ \exists\ I_k \in \mathfrak{H}(E_k): \quad \lim_{k\uparrow\infty} I_k^- = I^-, \quad \lim_{k\uparrow\infty} I_k^+ = I^+. \tag{5.19}$$

Indeed, consider an increasing family of compact intervals $C^h \uparrow I$, $h \in \mathbb{N}$ and two sequences $s^\pm_k \in E_k$ such that $s^\pm_k \to I^\pm$ and $\vartheta_k(s^\pm_k) \xrightarrow{\sigma} \vartheta(I^\pm)$ (they exist by (5.11), since $I^\pm \in E \subset \mathrm{Li}_{k\uparrow\infty} E_k$). We will have $E_k \cap C^h = \emptyset$ for k sufficiently big, since otherwise C^h should intersect $\mathrm{Ls}_{k\to\infty} E_k = E$. Denoting by I^h_k the connected component of $\mathbb{R} \setminus E_k$ intersecting C^h, since $E_k^- \le s_k^- \le (I^h_k)^- \le \min C^h$ and $E_k^+ \ge s_k^+ \ge (I^h_k)^+ \ge \max C^h$, we clearly have

$$I^- = \lim_{k\uparrow\infty} s_k^- \le \liminf_{k\to\infty} (I^h_k)^- \le (C^h)^-, \quad I^+ = \lim_{k\uparrow\infty} s_k^+ \ge \limsup_{k\to\infty} (I^h_k)^+ \ge (C^h)^+. \tag{5.20}$$

Since $\lim_{h\uparrow+\infty}(C^h)^\pm = I^\pm$, a standard diagonal argument yields (5.19).

Let us now choose a subsequence $n \mapsto k_n$ such that

$$\vartheta_{k_n}(I^\pm_{k_n}) \xrightarrow{\sigma} \theta^\pm, \quad \liminf_{k\uparrow\infty} \delta\left(\vartheta_k(I^-_k), \vartheta_k(I^+_k)\right) = \lim_{n\to\infty} \delta\left(\vartheta_{k_n}(I^-_{k_n}), \vartheta_{k_n}(I^+_{k_n})\right) \geq \delta(\theta^-, \theta^+) \tag{5.21}$$

Since $0 \leq I^-_k - s^-_k \to 0$ and $0 \leq s^+_k - I^+_k \to 0$ as $k \to \infty$, (5.15) and the lower semicontinuity of d yield

$$\mathsf{d}(\vartheta(I^-), \theta^-) \leq \liminf_{n\to\infty} \mathsf{d}(\vartheta_{k_n}(s^-_{k_n}), \vartheta_{k_n}(I^-_{k_n})) = 0,$$
$$\mathsf{d}(\theta^+, \vartheta(I^+)) \leq \liminf_{n\to\infty} \mathsf{d}(\vartheta_{k_n}(I^+_{k_n}), \vartheta_{k_n}(s^+_{k_n})) = 0.$$

$\langle$B.1$\rangle$ then yields $\delta(\vartheta(I^-), \vartheta(I^+)) \leq \delta(\theta^-, \vartheta(I^+)) \leq \delta(\theta^-, \theta^+)$ so that (5.21) yields

$$\liminf_{k\uparrow\infty} \delta\left(\vartheta_k(I^-_k), \vartheta_k(I^+_k)\right) \geq \delta(\vartheta(I^-), \vartheta(I^+)). \tag{5.22}$$

Since this holds for every connected component of $\mathbb{R} \setminus E$, if we consider a finite collection of disjoint open intervals $(I_n)_{n=1}^N \in \mathfrak{H}(E)$, we can find sequences $I_{n,k} \in \mathfrak{H}(E_k)$ as in (5.20) with $I_{n_0,k} \cap I_{n_1,k} = \emptyset$ for distinct indices $n_0, n_1 \in \{1, \cdots, N\}$. Then

$$\sum_{n=1}^N \delta(\vartheta(I^-_n), \vartheta(I^+_n)) \leq \liminf_{k\uparrow\infty} \sum_{n=1}^N \delta(\vartheta_k(I^-_{n,k})\vartheta_k(I^+_{n,k})) \leq \liminf_{k\uparrow\infty} \mathrm{GapVar}_\delta(\vartheta_k). \tag{5.23}$$

Taking the supremum of the right hand side with respect to finite collections in $\mathfrak{H}(E)$ we eventually obtain (5.16). $\square$

5.3 Existence of optimal transitions

In this section we will show that whenever $\mathsf{c}(t, u_-, u_+)$ is finite there exists an optimal transition ϑ attaining the infimum in (3.24). This results from a standard application of the Direct Method in Calculus of Variations and the following compactness property, which somehow combines Kuratowski and Arzelà-Ascoli Theorems, see [2, Prop. 3.3.1].

Theorem 5.4 (Compactness) *Let $F \subset X$ be a sequentially compact subset of X, C be a compact subset of $\mathbb{R}$, $t \in [0,T]$ and let $\mathcal{R} : [0,T] \times F \to [0, +\infty]$ be a σ-l.s.c. function such that $S(t) := \{x \in F : \mathcal{R}(t, x) = 0\}$ is separated by d.*

If $\mathsf{t}_k : E_k \to [0,T]$ and $\vartheta_k \in \mathrm{C}_{\sigma,\mathsf{d}}(E_k, F)$ are sequences of functions with $E_k \subset C$, satisfying the d-equicontinuity property (5.15) and the uniform bounds

$$\sup_k \sum_{s\in E_k\setminus\{E^\pm_k\}} \mathcal{R}(\mathsf{t}_k(s), \vartheta_k(s)) = R < \infty, \quad \lim_{k\to\infty} \sup_{s\in E_k} |\mathsf{t}_k(s) - t| = 0, \tag{5.24}$$

then there exist a subsequence $n \mapsto k_n$, a compact set $E \subset \mathbb{R}$ and a function $\vartheta \in \mathrm{C}_{\sigma,\mathsf{d}}(E, F)$ such that as $n \uparrow \infty$:

1) $E_{k_n} \xrightarrow{K} E$,

2) $\mathrm{graph}(\vartheta) \subset \mathrm{Li}_{n\uparrow\infty}\,\mathrm{graph}(\vartheta_{k_n})$,

3) whenever $s_{k_n} \in E_{k_n}$ converges to s with $\mathcal{R}(\mathsf{t}_{k_n}(s_{k_n}), \vartheta_{k_n}(s_{k_n})) \to 0$ then $\vartheta_{k_n}(s_{k_n}) \to \vartheta(s)$,

4) $\vartheta_{k_n}(E^\pm_{k_n}) \to \vartheta(E^\pm)$.

Notice that whenever d separates the points of F (e.g. when (F, d) is a metric space) then we can choose $\mathcal{R} \equiv 0$ so that (5.24) is always satisfied.

Proof. Let us now introduce the functions $r_k : \mathbb{R} \to [0, \infty]$ defined by

$$r_k(s) := \begin{cases} \mathcal{R}(\mathsf{t}_k(s), \vartheta_k(s)) & \text{if } s \in E_k, \\ +\infty & \text{if } s \in \mathbb{R} \setminus E_k. \end{cases} \tag{5.25}$$

It is not difficult to check that r_k are lower semicontinuous. Compactness of Γ-convergence [7, Theorem 8.5] provides a further subsequence (still no relabelled) and a lower semicontinuous limit function $r : \mathbb{R} \to [0, \infty]$ such that $\Gamma\text{-}\lim_{k\to\infty} r_k = r$. By using the bounds $i_{E_k} \le r_k \le i_{E_k} + R$, one can easily check that the compact set $E := \{s \in \mathbb{R} : r(s) \le R\}$ coincides with the Kuratowski limit of E_k. It is not difficult to check, arguing as in the second step of the proof of Theorem 5.3, that

$$\sum_{s\in E} r(s) \le \liminf_{k\uparrow\infty} \sum_{s\in E_k} r_k(s) \le R. \tag{5.26}$$

It follows that the (relatively) open set $B := \{s \in E : r(s) > 0\}$ is at most countable and every point of B is isolated in E.

Since E is separable, we can find a countable set A dense in $E \setminus B$ and containing $E^\pm$. For every $s \in A \cup B \setminus \{E^\pm\}$ there exists a sequence $s_k(s) \in E_k$ such that $s_k(s) \to s$ and $r_k(s_k(s)) = \mathcal{R}(\mathsf{t}_k(s), \vartheta_k(s_k(s)) \to r(s)$. When $s = E^\pm$ we just choose $s_k(s) := E^\pm_k$.

Since the maps ϑ_k take values in the sequentially compact set F, by a diagonal argument we can find a subsequence $n \mapsto k(n)$ and a function $\vartheta : A \cup B \to F$ such that

$$\vartheta_{k(n)}(s_{k(n)}(s)) \xrightarrow{\sigma} \vartheta(s), \quad \mathsf{d}(\vartheta(s), \vartheta(s')) \le \omega(s' - s) \quad \text{for every } s, s' \in A \cup B,\ s' \ge s. \tag{5.27}$$

We now extend ϑ to the closure of A: since $\vartheta(A) \subset F$ and $\overline{A} \subset S(t)$, it is sufficient to apply Lemma 2.1.

In order to prove 2) for every $s \in E$ we have to exhibit a sequence $s_{k(n)}(s) \in E_{k(n)}$ converging to s such that $\vartheta_{k(n)}(s_{k(n)}(s)) \to \vartheta(s)$. Such a property is satisfied by construction whenever $s \in A \cup B$. On the other hand, every point of $s \in E \setminus B$ is limit of sequences $s_k \in E_k$ with $\mathcal{R}(\mathsf{t}_k(s_k), \vartheta_k(s_k)) \to 0$. We denote by θ the limit of $\vartheta_{k'(n)}(s_{k'(n)}(s))$, where $k'(n)$ is a subsequence of $k(n)$. By the lower semicontinuity of $\mathcal{R}$ we get $\theta \in S(t)$. From (5.15) we deduce that for every $r \in A$, $r \le s$,

$$\begin{aligned} \mathsf{d}(\vartheta(r), \theta) &\le \liminf_{n\uparrow\infty} \mathsf{d}(\vartheta_{k'(n)}(s_{k'(n)}(r)), \vartheta_{k'(n)}(s_{k'(n)}(s))) \\ &\le \liminf_{n\uparrow\infty} \omega(s_{k'(n)}(s) - s_{k'(n)}(r)) \le \omega(s - r). \end{aligned}$$

Since A is dense in $E \setminus (A \cup B)$, the previous inequality yields $\mathsf{d}(\theta, \vartheta(s)) = 0$ so that $\theta = \vartheta(s)$.

The proof of 3) follows by a completely analogous argument. 4) is a consequence of the fact that $E^\pm \in A$ and $s_k(E^\pm) = E^\pm_k \to E^\pm$ as $k \to \infty$. $\square$

Corollary 5.5 (Existence of optimal transitions) *Let us assume that* $\langle$A.1$\rangle$, $\langle$B.1$\rangle$, $\langle$C$\rangle$ *hold and let* $t \in [0, T], u^\pm \in X$ *with* $\mathsf{c}(t, u^-, u^+) = C < \infty$. *Then there exists an optimal transition* $\vartheta \in \mathrm{C}_{\sigma,\mathsf{d}}(E, X)$ *connecting* u^- *and* u^+, *namely*

$$\vartheta(E^-) = u^-, \quad \vartheta(E^+) = u^+, \quad \mathsf{c}(t, u^-, u^+) = \mathrm{Trc}(t, \vartheta, E). \tag{5.28}$$

Proof. Let $\vartheta_k \in \mathrm{C}_{\sigma,\mathsf{d}}(E_k, X)$ be an optimizing sequence of transitions with $\vartheta_k(E^\pm) = u^\pm$ and $\mathrm{Trc}(t, \vartheta_k, E_k) \to \mathsf{c}(t, u^-, u^+)$ as $k \uparrow \infty$. By Lemma 5.1 it is not restrictive to assume that $E_k^- = 0$, $E_k^+ \le C$ for a sufficiently big constant C and that (5.3) holds uniformly. In particular $\mathsf{d}(u^-, \vartheta_k(r)) \le C$ for every $k \in \mathbb{N}$ and $r \in E_k$ so that $\vartheta_k(E_k)$ is uniformly bounded. Moreover, the next Theorem 6.3 shows that $\mathcal{E}(t, \vartheta_k(r)) \le C$ so that $\vartheta_k(E_k)$ is contained in a sublevel of $\mathcal{F}_0$. Applying Theorem 5.4 we can extract a subsequence converging to a limit transition $\vartheta \in \mathrm{C}_{\sigma,\mathsf{d}}(E, X)$ with $\vartheta(E^\pm) = u^\pm$. By Theorem 5.3 we have $\mathrm{Trc}(t, \vartheta, E) \le \liminf_{k\to\infty} \mathrm{Trc}(t, \vartheta_k, E_k) = \mathsf{c}(t, u^-, u^+)$, so that ϑ is optimal. $\square$

By a similar argument, we obtain

Corollary 5.6 (Lower semicontinuity of the cost c) *Let us assume that $\langle$A$\rangle$, $\langle$B.1$\rangle$, $\langle$C$\rangle$ hold, let F be a sublevel of $\mathcal{F}_0$ and let $(u_k^\pm)_k \subset F$ be sequences of points converging to $u^\pm$. Let $\vartheta_k \in \mathrm{C}_{\sigma,\mathsf{d}}(E_k, X)$ and $\mathsf{t}_k : E_k \to [0, T]$ satisfy $\vartheta_k(E_k^\pm) = u^\pm$ and $\lim_{k\to\infty} \sup_{s\in E_k} |\mathsf{t}_k(s) - t| = 0$. Then*

$$\liminf_{k\to\infty} \Big(\mathrm{Var}_{\mathsf{d}}(\vartheta_k, E_k) + \mathrm{GapVar}_\delta(\vartheta_k, E_k) + \sum_{s\in E_k} \mathcal{R}(\mathsf{t}_k(s), \vartheta_k(s)) \Big) \ge \mathsf{c}(t, u^-, u^+). \tag{5.29}$$

In particular, if $t_k \to t$

$$\liminf_{k\to\infty} \mathsf{c}(t_k, u_k^-, u_k^+) \ge \mathsf{c}(t, u^-, u^+). \tag{5.30}$$

6 Energy inequalities

We can now prove the energetic inequality stated in (3.23). Our proof is based on the following elementary Lemma, see [15] for similar arguments.

Lemma 6.1 *Let $E \subset \mathbb{R}$ be a compact set with $E^- < E^+$, let $L(E)$ be the set of limit points of E. We consider a function $f : E \to \mathbb{R}$ upper semicontinuous and continuous on the left and a function $g \in \mathrm{C}(E)$ strictly increasing, satisfying the following two conditions:*
i) for every $I \in \mathfrak{H}(E)$

$$\frac{f(I^+) - f(I^-)}{g(I^+) - g(I^-)} \le 1; \tag{6.1}$$

ii) for every $t \in L(E)$ which is an accumulation point of $L(E) \cap (-\infty, t)$ we have

$$\liminf_{s\uparrow t,\ s\in L(E)} \frac{f(t) - f(s)}{g(t) - g(s)} \le 1. \tag{6.2}$$

Then the map $s \mapsto f(s) - g(s)$ is non increasing in E; in particular

$$f(E^+) - f(E^-) \le g(E^+) - g(E^-). \tag{6.3}$$

Proof. By replacing E with $E \cap [E^-, s]$ it is easy to see that our thesis is in fact equivalent to (6.3). In order to prove it, it is not restrictive to assume that $f(E^-) = g(E^-) = 0$ and E contains at least three points (otherwise (6.3) follows by (6.1)).

We argue by contradiction, supposing that

$$\gamma := \frac{f(E^+)}{g(E^+)} > 1$$

and we consider the map $h(s) := f(s) - \gamma g(s)$, $s \in E$. Since $h(E^-) = h(E^+) = 0$, h takes its maximum at some point $\bar s \in E \cap (E^-, E^+)$. Since

$$f(s) - \gamma g(s) \le f(\bar s) - \gamma g(\bar s) \quad \text{for every } s \in E,$$

we obtain

$$\frac{f(\bar s) - f(s)}{g(\bar s) - g(s)} \ge \gamma > 1 \quad \text{for every } s \in E \cap [E^-, \bar s). \tag{6.4}$$

(6.1) shows that $\bar s$ cannot be the right extremum I^+ for some $I \in \mathfrak{H}(E)$ and (6.2) shows that $\bar s$ is isolated in $L(E) \cap [E^-, \bar s]$. Therefore, there exists $\varepsilon > 0$ such that $(\bar s - \varepsilon, \bar s)$ contains an increasing sequence $(s_n)_n$ of isolated points of E, converging to $\bar s$. Using (6.1) and the fact that $(s_n, s_{n+1}) \in \mathfrak{H}(E)$ we get

$$f(s_{n+1}) - f(s_n) \le g(s_{n+1}) - g(s_n). \tag{6.5}$$

Summing up from $n = 1$ to $N - 1$ we obtain

$$f(s_N) - f(s_1) \le g(s_N) - g(s_1), \tag{6.6}$$

and passing to the limit as $N \uparrow \infty$ by using the left continuity of f and the continuity of g we eventually get

$$f(\bar s) - f(s_1) \le g(\bar s) - g(s_1) \tag{6.7}$$

which is in contradiction with (6.4). $\square$

As a corollary we obtain a "dual" result for functions defined on intervals.

Lemma 6.2 *Let $g : [a,b] \to \mathbb{R}$ be strictly increasing, $f : [a,b] \to \mathbb{R}$ be a left-continuous function whose restriction to $[a,b] \setminus \mathrm{J}_g$ is upper semicontinuous. If*

$$\limsup_{r \downarrow t} f(r) - f(t) \le g(t+) - g(t-) \quad \textit{for every } t \in \mathrm{J}_g, \tag{6.8}$$

and

$$\liminf_{s \uparrow t} \frac{f(t) - f(s)}{g(t-) - g(s-)} \le 1 \quad \textit{for every } t \in [a,b], \tag{6.9}$$

then the map $t \mapsto f(t) - g(t)$ is non increasing.

Proof. Let us define E as the closure of $Z := g([a,b])$. We denote by $\mathsf{s} : E \mapsto [a,b]$ the continuous map whose restriction to Z coincides with g^{-1}; $\tilde f = f \circ \mathsf{s}$ is left continuous and upper semicontinuous in Z; notice moreover that every $I \in \mathfrak{H}(E)$ is of the form $(g(t-), g(t))$ or $(g(t), g(t+))$ for some $t \in \mathrm{J}_g$. Thus if $z \in E \setminus Z$ there exists a unique $t \in \mathrm{J}_g$ such that $z \in \{g(t-), g(t+)\}$. If $z = g(t-)$ we simply set $\tilde f(z) := f(t)$; if $z = g(t+)$ we set $\tilde f(z) := \limsup_{r \downarrow t} f(r)$.

Defining $\tilde g(r) := r$, $r \in E$, it is then easy to check that we can apply Lemma 6.1 to the couple of functions $\tilde f, \tilde g$ obtaining that $r \mapsto h(r) = \tilde f(r) - r$ is nonincreasing in E. Thus composing with g we get $t \mapsto f(t) - g(t) = h \circ g$ is non increasing. $\square$

Theorem 6.3 *Suppose that* Assumption $\langle$B$\rangle$ *hold. For every $t \in [0,T]$ and $u^\pm \in X$ we have*

$$\mathcal{E}(t, u^+) + \mathsf{c}(t, u^-, u^+) \ge \mathcal{E}(t, u^-). \tag{6.10}$$

Proof. If $\mathsf{c}(t,u^-,u^+)=+\infty$ the inequality is trivial. Otherwise, let E be a compact subset of $\mathbb{R}$, $\vartheta\in \mathrm{C}_{\sigma,\mathsf{d}}(E,X)$ a continuous map such that $\vartheta(E^\pm)=u^\pm$ and $\mathrm{Trc}(t,\vartheta)<+\infty$. We want to apply the previous Lemma 6.1 with the choices

$$f(s):=-\mathcal{E}(t,\vartheta(s)),\quad g(s):=\mathrm{Trc}(t,\vartheta;E\cap[E^-,s]).$$

Notice that g is continuous since $\vartheta\in \mathrm{C}_{\sigma,\mathsf{d}}(E,X)$ and f is upper semicontinuous thanks to the lower semicontinuity of $\mathcal{E}$ and the continuity of ϑ; f is also left continuous: whenever $s_n\uparrow s$ is an increasing sequence in E, we have $\mathsf{d}(\vartheta(s_n),\vartheta(s))\to 0$ and the property $\sum_n \mathcal{R}(t,\vartheta(s_n))<\infty$ shows that $\lim_{n\to\infty}\mathcal{R}(t,\vartheta(s_n))=0$, so that we obtain $\mathcal{E}(t,\vartheta(s_n))\to\mathcal{E}(t,\vartheta(s))$ thanks to Lemma 3.11 v) and $\langle$B.2$\rangle$. It remains to check conditions (6.1) and (6.2)

(6.1) follows from the definition of $\mathcal{R}$, since

$$\mathcal{E}(t,\vartheta(I^+))+\mathcal{R}(t,\vartheta(I^-))+\mathsf{D}(\vartheta(I^-),\vartheta(I^+))\ge\mathcal{E}(t,\vartheta(I^-)),$$

and the fact that

$$\mathrm{Trc}(t,\vartheta,E\cap[E^-,I^+])-\mathrm{Trc}(t,\vartheta,E\cap[E^-,I^-])=\mathcal{R}(t,\vartheta(I^-))+\mathsf{D}(\vartheta(I^-),\vartheta(I^+)).$$

(6.2) is a direct consequence of (3.9) and of the inequality

$$\mathsf{d}(\vartheta(r),\vartheta(s))\le\mathrm{Trc}(t,\vartheta,E\cap[E^-,s])-\mathrm{Trc}(t,\vartheta,E\cap[E^-,r])=g(s)-g(r);$$

recall that the set $E_{\mathcal{R}}=\{r:\mathcal{R}(t,\vartheta(r))>0\}$ is discrete, so that for every point $s\in L(E)$ we have $\vartheta(s)\in\mathscr{S}_{\mathsf{D}}(t)$ and (3.9) can be applied. To conclude the proof it is sufficient to take the infimum over admissible curves ϑ. □

A direct consequence of Proposition 6.3 is a description of the behaviour of a bounded variation curve on its jump points.

Corollary 6.4 *Let $u\in\mathrm{BV}_{\sigma,\mathsf{d}}([0,T];X)$. Then for every $t\in\mathrm{J_u}$ the following inequalities hold:*

$$\begin{aligned}\mathcal{E}(t,u(t+))+\mathsf{c}(t,u(t),u(t+))&\ge\mathcal{E}(t,u(t)),\\ \mathcal{E}(t,u(t))+\mathsf{c}(t,u(t-),u(t))&\ge\mathcal{E}(t,u(t-)),\\ \mathcal{E}(t,u(t+))+\mathsf{c}(t,u(t-),u(t+))&\ge\mathcal{E}(t,u(t-)).\end{aligned}\tag{6.11}$$

The chain rule inequality (3.26) is a consequence of (6.11). Indeed, we can recover the inequality also at the continuity points of u with a similar trick, applying Lemma 6.2.

Theorem 6.5 *Let us suppose that $\langle$B$\rangle$ and $\langle$A.1$\rangle$, (2.19), (2.20), (2.24) hold (these properties are verified if $\langle$A.2$\rangle$ or $\langle$A.2'$\rangle$ hold). Let $u\in\mathrm{BV}_{\sigma,\mathsf{d}}([0,T];X)$ satisfy $(\mathrm{S_D})$. Then for every $0\le t_0\le t_1\le T$ the following inequality holds:*

$$\mathcal{E}(t_1,u(t_1))+\mathrm{Var}_{\mathsf{d},\mathsf{c}}(u,[t_0,t_1])\ge\mathcal{E}(t_0,u(t_0))+\int_{t_0}^{t_1}\mathcal{P}(s,u(s))\mathrm{d}s.\tag{6.12}$$

Proof. We will apply Lemma 6.2 in the interval $[0,T]$ with the choices

$$f(t):=\int_0^t\mathcal{P}(s,u(s))\mathrm{d}s-\mathcal{E}(t,u(t-)),\quad g(t):=\mathrm{Var}_{\mathsf{d},\mathsf{c}}(u,[0,t])+\varepsilon t,$$

for a small parameter $\varepsilon > 0$.

Since u is continuous in $[0,T] \setminus \mathrm{J}_g$ the upper semicontinuity of f outside J_g is guaranteed by the lower semicontinuity of $\mathcal{E}$. Its left continuity is a consequence of the stability property $(\mathrm{S_D})$ of u, of Lemma 3.11 v) and of $\langle$B.2$\rangle$.

Condition (6.8) is satisfied thanks to Corollary 6.4 and the fact that at every $t \in \mathrm{J}_u$

$$\limsup_{r \downarrow t} f(r) \le \int_0^t \mathcal{P}(s,u(s))\,\mathrm{d}s - \mathcal{E}(t,u(t+)), \quad g(t+) - g(t-) = \mathsf{c}(t,u(t-),u(t)) + \mathsf{c}(t,u(t),u(t+)).$$

In order to check (6.9) let us fix a couple of times $s,t \in [0,T]$ with $s < t$ and observe that

$$\begin{aligned} f(t) - f(s) &= \mathcal{E}(s,u(s-)) - \mathcal{E}(t,u(t-)) + \int_s^t \mathcal{P}(r,u(r))\,\mathrm{d}r = \\ &= \mathcal{E}(s,u(s-)) - \mathcal{E}(s,u(t-)) + \mathcal{E}(s,u(t-)) - \mathcal{E}(t,u(t-)) + \int_s^t \mathcal{P}(r,u(r))\,\mathrm{d}r, \end{aligned}$$

$$g(t-) - g(s-) \ge \mathsf{d}(u(s-),u(t-)) + \varepsilon(t-s).$$

The conditional upper semi-continuity of $\mathcal{P}$ (2.24) (recall that the energy is left-continuous) and (2.19) yield

$$\begin{aligned} &\limsup_{s \uparrow t} \frac{1}{\varepsilon(t-s)} \Big(\mathcal{E}(s,u(t-)) - \mathcal{E}(t,u(t-)) + \int_s^t \mathcal{P}(r,u(r))\,\mathrm{d}r \Big) \\ &\le \frac{1}{\varepsilon} \Big(-\liminf_{s \uparrow t} \frac{\mathcal{E}(t,u(t-)) - \mathcal{E}(s,u(t-))}{(t-s)} + \limsup_{s \uparrow t} \ \fint_s^t \mathcal{P}(r,u(r))\,\mathrm{d}r \Big) \\ &\le \frac{1}{\varepsilon} \Big(-\mathcal{P}(t,u(t-)) + \mathcal{P}(t,u(t-)) \Big) \le 0. \end{aligned}$$

On the other hand, from assumption $\langle$B.3$\rangle$ and the stability property $(\mathrm{S_D})$ we have

$$\limsup_{s \uparrow t} \frac{\mathcal{E}(s,u(s-)) - \mathcal{E}(s,u(t-))}{\mathsf{d}(u(s-),u(t-))} \le 1.$$

□

7 Convergence proof for the discrete approximations

In this section we will prove existence of a Visco-Energetic solution, stated in Theorem 3.9. We will always suppose that the energy $\mathcal{E}$ satisfies assumptions $\langle$A$\rangle$ (where we will also consider the case $\langle$A.2'$\rangle$), that the viscous correction δ is admissible according to $\langle$B$\rangle$, and that conditions $\langle$C$\rangle$ hold.

7.1 Discrete estimates

Hereafter, τ will be a given partition of $[0,T]$. We obtain some preliminary estimates for the minimizing movement scheme.

Theorem 7.1 (Discrete estimates) *Let $U_\tau^0 \in X$ be given so that*

$$\mathcal{F}_0(U_\tau^0) = \mathcal{E}(0,U_\tau^0) + \mathsf{d}(x_0,U_\tau^0) \le C_0. \tag{7.1}$$

Then every solution U_τ^n of the incremental problem (IM$_{\mathsf{d},\delta}$) *starting from U_τ^0 satisfies a discrete version of stability* (S$_{\mathsf{D}}$) *and energy balance* (E$_{\mathsf{d,c}}$)*, namely for every $n = 1, \ldots, N$ we have*

$$\mathcal{E}(t_\tau^n, U_\tau^n) \leq \mathcal{E}(t_\tau^n, V) + \mathsf{d}(U_\tau^n, V) + \delta(U_\tau^{n-1}, V), \tag{7.2}$$

$$\mathcal{E}(t_\tau^n, U_\tau^n) + \mathsf{D}(U_\tau^{n-1}, U_\tau^n) + \mathcal{R}(t_\tau^n, U_\tau^{n-1}) = \mathcal{E}(t_\tau^{n-1}, U_\tau^{n-1}) + \int_{t_\tau^{n-1}}^{t_\tau^n} \mathcal{P}(s, U_\tau^{n-1}) \mathrm{d}s. \tag{7.3}$$

Moreover, there exist constants C_1, C_2 depending only on C_0 (of (7.1)*), on C_P (of* (2.20)*), and on T, such that*

$$\mathcal{F}(t_\tau^n, U_\tau^n) \leq C_0 \mathrm{e}^{C_P t_\tau^n} \leq C_0 \mathrm{e}^{C_P T}, \quad \mathsf{d}(x_0, U_\tau^n) \leq C_1, \tag{7.4}$$

$$\sum_{j=1}^{N} \mathsf{D}(U_\tau^{j-1}, U_\tau^j) + \mathcal{R}(t_\tau^j, U_\tau^{j-1}) \leq C_2. \tag{7.5}$$

Proof. Since U_τ^n is a minimizer for (IM$_{\mathsf{d},\delta}$), the estimate

$$\mathcal{E}(t_\tau^n, U_\tau^n) + \mathsf{d}(U_\tau^{n-1}, U_\tau^n) + \delta(U_\tau^{n-1}, U_\tau^n) \leq \mathcal{E}(t_\tau^n, V) + \mathsf{d}(U_\tau^{n-1}, V) + \delta(U_\tau^{n-1}, V),$$

holds for every $V \in X$. Using the triangle inequality and $\delta(U_\tau^{n-1}, U_\tau^n) \geq 0$, we have proved the discrete stability (7.2).

From the minimality of U_τ^n and the definition of $\mathcal{R}$ we have:

$$\mathcal{R}(t_\tau^n, U_\tau^{n-1}) = \mathcal{E}(t_\tau^n, U_\tau^{n-1}) - \mathcal{E}(t_\tau^n, U_\tau^n) - \mathsf{D}(U_\tau^{n-1}, U_\tau^n)$$

and since

$$\mathcal{E}(t_\tau^n, U_\tau^{n-1}) = \mathcal{E}(t_\tau^{n-1}, U_\tau^{n-1}) + \int_{t_\tau^{n-1}}^{t_\tau^n} \mathcal{P}(s, U_\tau^{n-1}) \mathrm{d}s$$

we have also proved the discrete energy balance (7.3).

Using ⟨A.1⟩ and (2.21) in the power term and denoting by $\tau^n := t_\tau^n - t_\tau^{n-1}$ we get

$$\int_{t_\tau^{n-1}}^{t^n} \mathcal{P}(s, U_\tau^{n-1}) \mathrm{d}s \leq \Big(\mathsf{d}(x_o, U_\tau^{n-1}) + \mathcal{E}(t_\tau^{n-1}, U_\tau^{n-1}) \Big) (\mathrm{e}^{C_P \tau^n} - 1).$$

Then summing up $\mathsf{d}(x_o, U_\tau^{n-1})$ to both terms of the inequality (7.3) and using the triangle inequality (2.1) we have

$$\mathcal{E}(t_\tau^n, U_\tau^n) + \mathsf{d}(x_o, U_\tau^n) \leq \Big(\mathsf{d}(x_o, U_\tau^{n-1}) + \mathcal{E}(t_\tau^{n-1}, U_\tau^{n-1}) \Big) \mathrm{e}^{C_P \tau^n}.$$

A simple induction argument yields

$$\mathcal{E}(t_\tau^n, U_\tau^n) + \mathsf{d}(x_o, U_\tau^n) \leq \Big(\mathcal{E}(0, U_\tau^0) + \mathsf{d}(x_o, U_\tau^0) \Big) \mathrm{e}^{C_P t_\tau^n}.$$

This also yields $\mathsf{d}(x_0, U_\tau^n) \leq C_1$ where $C_1 := \sup\{\mathsf{d}(x_0, v) : \mathcal{F}_0(v) \leq C_0 \mathrm{e}^{2C_p T}\}$.

Finally, we estimate the dissipated energy via

$$\begin{aligned} &\sum_{j=1}^{N} \mathsf{D}(U_\tau^{j-1}, U_\tau^j) + \mathcal{R}(t_\tau^j, U_\tau^{j-1}) \\ &\qquad \leq \mathcal{F}_0(U_\tau^0) - \mathcal{F}(t_\tau^N, U_\tau^N) + \sum_{j=1}^{N} \mathcal{F}(t_\tau^{j-1}, U_\tau^{j-1}) \big(\mathrm{e}^{C_p \tau^j} - 1 \big) + \mathsf{d}(x_0, U_\tau^N) \\ &\qquad \leq \mathcal{F}_0(U_\tau^0) + \mathcal{F}_0(U_\tau^0) \sum_{j=1}^{N} (\mathrm{e}^{C_P t_\tau^j} - \mathrm{e}^{C_P t_\tau^{j-1}}) + C_1, \leq C_0 \mathrm{e}^{C_P T} + C_1 \end{aligned}$$

and the proof is complete with $C_2 := C_0 \mathrm{e}^{C_P T} + C_1$. $\square$

7.2 Compactness

We introduce the functions

$$\mathsf{t}_\tau(t) := t_\tau^n, \quad \tilde{\mathsf{t}}_\tau(t) := t_\tau^{n+1}, \quad \overline{U}_\tau(t) = U_\tau^n, \quad \mathsf{n}_\tau(t) := n \qquad \text{whenever } t \in (t_\tau^{n-1}, t_\tau^n], \tag{7.6}$$

so that (7.3) can be rewritten as

$$\begin{aligned} \mathcal{E}(\mathsf{t}_\tau(t), \overline{U}_\tau(t)) + \mathrm{Var}_{\mathsf{d}}(\overline{U}_\tau, [s,t]) + \sum_{j=\mathsf{n}_\tau(s)}^{\mathsf{n}_\tau(t)-1} \delta(U_\tau^j, U_\tau^{j+1}) + \mathcal{R}(t_\tau^{j+1}, U_\tau^j) \\ = \mathcal{E}(\mathsf{t}_\tau(s), \overline{U}_\tau(s)) + \int_{\mathsf{t}_\tau(s)}^{\mathsf{t}_\tau(t)} \mathcal{P}(r, \overline{U}_\tau(r))\,\mathrm{d}r. \end{aligned} \tag{7.7}$$

Notice that the variation function V_τ associated with $\overline{U}_\tau$ can be written as

$$V_\tau(t) := \mathrm{Var}_{\mathsf{d}}(\overline{U}_\tau, [0,t]) = \sum_{j=0}^{\mathsf{n}_\tau(t)-1} \mathsf{d}(U_\tau^j, U_\tau^{j+1}). \tag{7.8}$$

Similarly, we introduce the nondecreasing function $W_\tau : [0,T] \to [0,\infty)$

$$W_\tau(t) := \sum_{j=0}^{\mathsf{n}_\tau(t)-1} \mathsf{D}(U_\tau^j, U_\tau^{j+1}) + \mathcal{R}(t_\tau^{j+1}, U_\tau^j) = V_\tau(t) + \sum_{j=0}^{\mathsf{n}_\tau(t)-1} \delta(U_\tau^j, U_\tau^{j+1}) + \mathcal{R}(t_\tau^{j+1}, U_\tau^j). \tag{7.9}$$

Notice that $W_\tau - V_\tau$ is still a nonnegative and nondecreasing function.

Theorem 7.2 (Compactness) *Let $u_0 \in X$ be fixed and let $(\overline{U}_\tau)$ be a family of piecewise constant left-continuous interpolants of the discrete solutions U_τ^n of* $(\mathrm{IM}_{\mathsf{d},\delta})$ *starting from $U_\tau^0 \in X$, with*

$$\mathcal{F}_0(U_\tau^0) \le C_0, \quad U_\tau^0 \xrightarrow{\sigma} u_0 \text{ in } X, \quad \mathcal{E}(0, U_\tau^0) \to \mathcal{E}(0, u_0) \text{ as } \tau \downarrow 0. \tag{7.10}$$

Let V_τ, W_τ be defined as in (7.8) *and* (7.9)*. Then for all sequences of partitions $k \mapsto \tau(k)$ with $\lim_{k\to\infty} |\tau(k)| = 0$ there exist*
- a (not relabeled) subsequence $k \mapsto \tau(k)$,
- a limit curve $u \in \mathrm{BV}_{\sigma,\mathsf{d}}([0,T]; X)$,
- nondecreasing functions $V, W : [0,T] \to [0,+\infty)$ with $W - V$ nondecreasing,
- a real function $\mathsf{E} \in \mathrm{BV}([0,T])$,
- a set $\mathscr{C} \subset [0,T]$ with $\mathscr{L}^1([0,T] \setminus \mathscr{C}) = 0$
such that

$$V_{\tau(k)}(t) \to V(t), \quad W_{\tau(k)}(t) \to W(t), \quad \text{for every } t \in [0,T], \tag{7.11}$$

$$\mathcal{R}(\tilde{\mathsf{t}}_{\tau(k)}(t), \overline{U}_{\tau(k)}(t)) \to 0 \quad \text{for every } t \in \mathscr{C}, \tag{7.12}$$

$$\overline{U}_{\tau(k)}(t) \xrightarrow{\sigma} u(t) \quad \text{for every } t \in \mathscr{C} \cup \mathrm{J}_W \tag{7.13}$$

$$\mathsf{d}(u(s), u(t)) \le V(t) - V(s), \quad \text{for every } 0 \le s \le t \le T, \tag{7.14}$$

$$u(t) \in \mathscr{S}_{\mathsf{D}}(t) \quad \textit{for every } t \notin \mathrm{J}_V, \tag{7.15}$$

$$\mathcal{E}(\mathsf{t}_{\tau(k)}(t), \overline{U}_{\tau(k)}(t)) \to \mathsf{E}(t) \geq \mathcal{E}(t, u(t)), \quad \textit{for every } t \in [0,T], \tag{7.16}$$

where

$$\mathsf{E}(t) = \mathcal{E}(t, u(t)) \quad \textit{for every } t \in \mathscr{C} \quad \textit{if } \langle \mathrm{A.2'} \rangle \textit{ holds}, \tag{7.17}$$

$$\mathsf{E}(t) + W(t) \leq \mathsf{E}(s) + W(s) + \int_s^t \mathcal{P}(r, u(t))\,\mathrm{d}r \quad \textit{for every } 0 \leq s \leq t \leq T. \tag{7.18}$$

Moreover, for every further subsequence $k \mapsto \tau'(k)$

$$\lim_{k\to\infty} \mathcal{R}(\tilde{\mathsf{t}}_{\tau'(k)}(t), \overline{U}_{\tau'(k)}(t)) = 0 \quad \Rightarrow \quad \lim_{k\to\infty} \overline{U}_{\tau'(k)}(t) = u(t) \quad \textit{for every } t \notin \mathrm{J}_V. \tag{7.19}$$

Finally, for every $t \in \mathrm{J}_W$ *there exist sequences* $(t_k^\pm)_k$ *such that* $t_k^- \uparrow t$, $t_k^+ \downarrow t$ *as* $k \uparrow \infty$ *and*

$$V_{\tau(k)}(t_k^\pm) \to V(t\pm), \quad W_{\tau(k)}(t_k^\pm) \to W(t\pm), \quad \overline{U}_{\tau(k)}(t_k^\pm) \xrightarrow{\sigma} u(t\pm). \tag{7.20}$$

Proof. Let us first observe that the values of $\overline{U}_\tau$ belong to the bounded sequentially compact set $F := \{v \in X : \mathcal{F}_0(v) \leq C_0 \mathrm{e}^{C_P T}\}$.

$\triangleright$ (7.11): (7.5) shows that the functions W_τ (and thus a fortiori V_τ) are uniformly bounded in $[0,T]$, so that Helly's theorem provides pointwise convergence (up to a subsequence) of $W_{\tau(k)}, V_{\tau(k)}$ to some increasing functions W, V, with $W - V$ also increasing and $\mathrm{J}_V \subset \mathrm{J}_W$.

$\triangleright$ (7.12): (7.5) yields

$$\int_0^{T-|\tau|} \mathcal{R}(\tilde{\mathsf{t}}_\tau(s), \overline{U}_\tau(s))\,\mathrm{d}s \leq C_2|\tau|, \tag{7.21}$$

so that there exists a subsequence $k \mapsto \tau(k)$ and a subset $\mathscr{C} \subset [0,T]$ of full measure such that (7.12) holds.

$\triangleright$ (7.13), (7.14), (7.15), (7.19): By a standard diagonal argument, we can choose a dense countable set $\mathscr{C}_1 \subset \mathscr{C}$, a subsequence (still denoted by $\tau(k)$), and a limit function $u : \mathscr{C}_1 \cup \mathrm{J}_W \to X$ such that

$$u_{\tau(k)}(t) \xrightarrow{\sigma} u(t), \quad \mathsf{d}(u(s), u(t)) \leq V(t) - V(s) \quad \text{for every } 0 \leq s \leq t \leq T,\ s,t \in \mathscr{C}_1 \cup \mathrm{J}_W.$$

Clearly $\mathrm{J}_{V_u} \subset \mathrm{J}_V \subset \mathrm{J}_W$ and $u(t) \in \mathscr{S}_{\mathsf{D}}(t)$ for every $t \in \mathscr{C}_1$ by (7.12) and the lower semicontinuity of $\mathcal{R}$. Since $\mathscr{S}_{\mathsf{D}}$ is separated by $\mathsf{d}_{\mathbb{R}}$, applying Lemma 2.4 we can extend u to a function (still denoted by u) in $\mathrm{BV}_{\sigma,\mathsf{d}}([0,T];X)$. The closure of $\mathscr{S}_{\mathsf{D}}$ and the fact that $\mathrm{J}_{V_u} \subset \mathrm{J}_V$ yield (7.15).

We can also prove that $\overline{U}_{\tau(k)}(t) \to u(t)$ for every $t \in \mathscr{C} \setminus \mathrm{J}_V$: in fact, (7.12) and the lower semicontinuity of $\mathcal{R}$ show that any limit point of the sequence $\{\overline{U}_{\tau(k)}(t)\}_{k\in\mathbb{N}}$ is contained in $\mathscr{S}_{\mathsf{D}}(t)$, which is separated by d. If v is an arbitrary limit point, passing to the limit in the inequalities $\mathsf{d}(\overline{U}_{\tau(k)}(s), \overline{U}_{\tau(k)}(t)) \leq V_k(t) - V_k(s)$ we get $\mathsf{d}(u(s), v) \leq V(t) - V(s)$ for every $s \in \mathscr{C}_1$; passing to the limit as $s \uparrow t$, $s \in \mathscr{C}_1$, we conclude that $\mathsf{d}(u(t), v) = 0$ which yields $v = u(t)$ since $\mathscr{S}_{\mathsf{D}}(t)$ is separated by d. The same argument yields (7.19).

$\triangleright$ (7.16): We notice that (7.7) yields for the constant $C := C_P C_0 \exp(C_P T)$

$$\mathcal{E}(\mathsf{t}_\tau(t), \overline{U}_\tau(t)) + W_\tau(t) + C\,\mathsf{t}_\tau(t) \leq \mathcal{E}(\mathsf{t}_\tau(s), \overline{U}_\tau(s)) + W_\tau(s) + C\,\mathsf{t}_\tau(s) \tag{7.22}$$

whenever $0 \le s \le t \le T$. Since $W_{\tau(k)}(t) \to W(t)$ and $\mathsf{t}_{\tau(k)}(t) \to t$ as $k \to \infty$, a further application of Helly's Theorem (and a further extraction of a subsequence) yields $\mathcal{E}(\mathsf{t}_{\tau(k)}(t), \overline{U}_{\tau(k)}(t)) \to \mathsf{E}(t)$ for every $t \in [0,T]$ and

$$\mathsf{E}(t) \ge \mathcal{E}(t, u(t)) \quad \text{for every } t \in \mathscr{C} \cup \mathrm{J}_W, \tag{7.23}$$

thanks to the lower semicontinuity of $\mathcal{E}$. Since the uniform bound

$$|\mathsf{E}(t) - \mathsf{E}(s)| \le W(t) - W(s) + C(t-s) \tag{7.24}$$

obtained by passing to the limit in (7.7) shows that E is continuous outside J_W, we conclude that $\mathsf{E}(t) \ge \mathcal{E}(t, u(t))$ holds everywhere in $[0,T]$.

$\triangleright$ (7.17): let us first notice that

$$\mathsf{E}(t) \le \limsup_{k\to\infty} \Big(\mathcal{E}(\tilde{\mathsf{t}}_{\tau(k)}(t), \overline{U}_{\tau(k)}(t)) + C|\tau(k)| \Big) = \limsup_{k\to\infty} \mathcal{E}(\tilde{\mathsf{t}}_{\tau(k)}(t), \overline{U}_{\tau(k)}(t)).$$

If $\langle$A.2'$\rangle$ holds, by (7.13), (7.12), and Lemma 3.11 v) we get

$$\lim_{k\to\infty} \mathcal{E}(\tilde{\mathsf{t}}_{\tau(k)}(t), \overline{U}_{\tau(k)}(t)) = \mathcal{E}(t, u(t)). \tag{7.25}$$

$\triangleright$ (7.18): since $\overline{U}_{\tau(k)}(t) \to u(t)$ for almost every $t \in [0,T]$, the upper semicontinuity of $\mathcal{P}$ and the fact that $\overline{U}_{\tau(k)}(t)$ is contained in a sublevel of $\mathcal{F}_0$ yield

$$\limsup_{k\to\infty} \int_{\mathsf{t}_\tau(s)}^{\mathsf{t}_\tau(t)} \mathcal{P}(r, \overline{U}_{\tau(k)}(r))\,\mathrm{d}r \le \int_s^t \mathcal{P}(r, u(r))\,\mathrm{d}r \quad \text{for every } 0 \le s < t \le T. \tag{7.26}$$

The same conclusion holds if we assume $\langle$A.2'$\rangle$ instead of $\langle$A.2$\rangle$, thanks to (7.17). (7.18) then follows by (7.7).

$\triangleright$ (7.20): this is a general property of double limits; let us check the case of t_k-. We first select a fundamental sequence of open neighborhoods N_n of $u(t-)$, a decreasing vanishing sequence ε_n and an increasing sequence $(t_n)_n$ in $[0,t) \cap \mathscr{C}_1$ so that $t_n \uparrow t$ as $n \uparrow \infty$, $u(t_n) \in N_n$ and $W(t) - W(t_n) < \varepsilon_n$, $\overline{U}_{\tau(k)}(t_n) \to u(t_n)$, $W_{\tau(k)}(t_n) \to W(t_n)$ as $k \to \infty$. We may find a strictly increasing sequence $n \mapsto \kappa(n)$ such that

$$|W_{\tau(k)}(t_n) - W(t)| < \varepsilon_n, \quad \overline{U}_{\tau(k)}(t_n) \in N_n \quad \text{for every } k \ge \kappa(n).$$

For every $k \ge \kappa(1)$ we can then define $n(k) := \min\big\{m \in \mathbb{N} : k \ge \kappa(m)\big\}$; it is not difficult to check that $n(k) \uparrow \infty$ and the sequence $k \mapsto t_k^- := t_{n(k)}$ satisfies (7.20). $\square$

7.3 Limit energy-dissipation inequality

We can now prove the energy inequality on jumps.

Lemma 7.3 *Let u_0, $\overline{U}_{\tau(k)}$, u, W, E be as in the previous Theorem 7.2. Then for every $t \in \mathrm{J}_W$ we have*

$$W(t) - W(t-) \ge \mathsf{c}(t, u(t-), u(t)), \quad W(t+) - W(t) \ge \mathsf{c}(t, u(t), u(t+)). \tag{7.27}$$

Proof. We will prove the first inequality of (7.27); the proof of the second inequality is completely analogous.

Let us fix $t \in \mathrm{J}_W$ and let us choose a sequence $t_k^- \uparrow t$ as in (7.20) so that

$$\overline{U}_{\tau(k)}(t_k^-) \xrightarrow{\sigma} u(t-), \quad \overline{U}_{\tau(k)}(t) \xrightarrow{\sigma} u(t).$$

For every $k \in \mathbb{N}$ we consider the compact set $E_k := \{n \in \mathbb{N} : \mathsf{t}_{\tau(k)}(t_k^-) \le n \le \mathsf{t}_{\tau(k)}(t)\}$ and the discrete transition $\vartheta_k : E_k \to X$ defined by $\vartheta_k(n) := U^n_{\tau(k)}$, $n \in E_k$. By construction $\vartheta_k(E_k^-) = \overline{U}_{\tau(k)}(t_k^-)$ and $\vartheta_k(E_k^+) = \overline{U}_{\tau(k)}(t)$. Moreover

$$\mathrm{Var}_{\mathsf{d}}(\vartheta_k, E_k) = \sum_{n \in E_k \setminus E_k^-} \mathsf{d}(U^{n-1}_{\tau(k)}, U^n_{\tau(k)}), \quad \mathrm{GapVar}_\delta(\vartheta_k, E_k) = \sum_{n \in E_k \setminus E_k^-} \delta(U^{n-1}_{\tau(k)}, U^n_{\tau(k)}) \tag{7.28}$$

and

$$\sum_{n \in E_k \setminus E_k^+} \mathcal{R}(t^{n+1}_{\tau(k)}, \vartheta_k(n)) = \sum_{n \in E_k \setminus E_k^+} \mathcal{R}(t^{n+1}_{\tau(k)}, U^n_{\tau(k)}) \tag{7.29}$$

so that

$$\mathrm{Var}_{\mathsf{d}}(\vartheta_k, E_k) + \mathrm{GapVar}_\delta(\vartheta_k, E_k) + \sum_{n \in E_k \setminus E_k^+} \mathcal{R}(t^{n+1}_{\tau(k)}, \vartheta_k(n)) = W_{\tau(k)}(t) - W_{\tau(k)}(t_k^-). \tag{7.30}$$

Passing to the limit and recalling Corollary 5.6 we conclude. □

Corollary 7.4 *Let u_0, $\overline{U}_{\tau(k)}$, u, V, W, E be as in the previous Theorem 7.2. Then*

$$V(t) - V(s) \ge \mathrm{Var}_{\mathsf{d}}(u, [s,t]), \quad W(t) - W(s) \ge \mathrm{Var}_{\mathsf{d},\mathsf{c}}(u, [s,t]). \tag{7.31}$$

In particular

$$\mathcal{E}(T, u(T)) + \mathit{Var}_{\mathsf{d},\mathsf{c}}(u; [0,T]) \le \mathcal{E}(0, u_0) + \int_0^T \mathcal{P}(r, u(r))\,\mathrm{d}r. \tag{7.32}$$

Proof. The first inequality of (7.31) immediately follows from (7.14).

We now consider an arbitrary ordered finite subset $\{t_1, t_2, \cdots, t_N\}$ of J_W. Using the additivity of the total variation and the fact that $V(t) - V(s) \le W(t) - W(s)$ we have

$$\begin{aligned}
&\mathrm{Var}_{\mathsf{d}}(u,[0,T]) + \sum_{j=1}^{N} \Delta_{\mathsf{c}}(t,u(t_j-),u(t_j)) + \Delta_{\mathsf{c}}(t,u(t_j),u(t_j+)) \\
&\quad\le V(t_1-) - V(0) + V(T) - V(t_N+) + \sum_{j=1}^{N-1} V(t_{j+1}-) - V(t_j+) \\
&\qquad + \sum_{j=1}^{N} \mathsf{c}(t,u(t_j-),u(t_j)) + \mathsf{c}(t,u(t_j),u(t_j+)) \\
&\quad\le V(t_1-) - V(0) + V(T) - V(t_N+) + \sum_{j=1}^{N-1} V(t_{j+1}-) - V(t_j+) \\
&\qquad + \sum_{j=1}^{N} W(t_j+) - W(t_j-) \\
&\quad\le W(t_1-) - W(0) + W(T) - W(t_N+) + \sum_{j=1}^{N-1} W(t_{j+1}-) - W(t_j+) \\
&\qquad + \sum_{j=1}^{N} W(t_j+) - W(t_j-)) = W(T) - W(0).
\end{aligned}$$

Taking the supremum with respect to all the finite subsets of J_W we conclude. □

7.4 Convergence: proof of Theorem 3.9

We can now conclude the proof of our main Theorem 3.9.

Let $\overline{U}_\tau$ be a family of piecewise constant left-continuous interpolants of the values U^n_τ of the incremental minimization scheme ($\mathrm{IM}_{\mathsf{d},\delta}$) with U^0_τ satisfying (7.10) and let $k \mapsto \tau(k)$ be any sequence of partitions with $|\tau(k)| \to 0$ as $k \to \infty$.

By Theorem 7.2 we can extract a subsequence (not relabeled) such that $\overline{U}_{\tau(k)}$ pointwise converges to a function $u \in \mathrm{BV}_{\sigma,\mathsf{d}}([0,T];X)$ in a set $\mathscr{C}$ containing J_u and with $\mathscr{L}^1\big([0,T]\setminus\mathscr{C}\big) = 0$. u satisfies the stability condition (S_{D}) by (7.14)-(7.15) and the energy inequality (3.27) by Corollary 7.4: applying Proposition 3.8 we conclude that u is a VE solution to $(X,\mathcal{E},\mathsf{d},\delta)$.

Combining (7.16), (7.18) with $s = 0$ and (7.31) we deduce that $\lim_{k\to\infty} \mathcal{E}(\mathsf{t}_{\tau(k)}(t),\overline{U}_{\tau(k)}(t)) = \mathsf{E}(t) = \mathcal{E}(t,u(t))$ for every $t \in [0,T]$. In particular $t \mapsto \mathcal{E}(t,u(t))$ is continuous in $[0,T] \setminus \mathrm{J}_V$.

Let us now prove that $\overline{U}_{\tau(k)}(t) \xrightarrow{\sigma} u(t)$ for every $t \in [0,T]$; the thesis is already true in $\mathscr{C}$, so we pick a point $t \notin \mathscr{C}$ (in particular $t \notin \mathrm{J}_V$) and we want to show that any limit u' of a converging subsequence $\overline{U}_{\tau'(k)}(t)$ coincides with $u(t)$. For every $r, s \in \mathscr{C}$ with $r < t < s$ the lower semicontinuity of d yields

$$\mathsf{d}(u(r),u') \le V(t) - V(r), \quad \mathsf{d}(u',u(s)) \le V(s) - V(t)$$

so that passing to the limit as $r \uparrow t$ and $s \downarrow t$ we get $\mathsf{d}(u(t),u') = \mathsf{d}(u',u) = 0$; by the triangle inequality and ⟨B.1⟩ we have

$$\mathsf{d}(u(t),v) \le \mathsf{d}(u',v), \quad \delta(u(t),v) \le \delta(u',v).$$

Since $\mathcal{E}(t,u(t)) = \mathsf{E}(t) \geq \mathcal{E}(t,u')$ and $u(t) \in \mathscr{S}_{\mathsf{D}}(t)$ we get for every $v \in X$

$$\mathcal{E}(t,u') \leq \mathcal{E}(t,u(t)) \leq \mathcal{E}(t,v) + \mathsf{d}(u(t),v) + \delta(u(t),v) \leq \mathcal{E}(t,v) + \mathsf{d}(u',v) + \delta(u',v)$$

so that $u' \in \mathscr{S}_{\mathsf{D}}(t)$. Since d separates $\mathscr{S}_{\mathsf{D}}(t)$ we conclude that $u' = u(t)$.

7.5 A uniform BV estimate for discrete Minimizing Movements

The aim of this section is to prove Theorem 4.3, namely a uniform bound for all discrete Minimizing Movements, under the stronger $\alpha - \Lambda$ convexity assumption of Section 4.3 and a Lipschitz property of the power term.

Let us recall that we are considering the metric setting of Remark 2.2, with

$$\delta(x,y) = \frac{1}{2}\mathsf{d}_*^2(x,y) \qquad \text{for every } x,y \in X,$$

where d_* is another continuous distance on X, the energy satisfies assumptions $\langle \mathrm{A} \rangle$, the generalized convexity (4.11) and the power term is Lipschitz, according to (4.15).

To prove Theorem 4.3 we combine two basic facts: the first one is the discrete Gronwall-like lemma of [30, Lemma 7.5].

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

$$(1+\gamma)^2 a_n^2 \leq a_{n-1}^2 + b_n a_n \quad \forall n \geq 1.$$

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

$$\sum_{n=1}^{k} a_n \leq \frac{1}{\gamma}\left(a_0 + \sum_{n=1}^{k} b_n\right).$$

The second ingredient is provided by the following estimates of the residual stability functional. We set $\mathsf{p}(x,y) := \mathsf{d}(x,y)\mathsf{d}_*(x,y)$.

Lemma 7.6 *Let us assume that $(\mathcal{E},\mathsf{d},\mathsf{d}_*)$ satisfies the* strong α-Λ *convexity property of Definition 4.2. For every $t \in [0,T]$, $x \in X$ and $y \in \mathrm{M}(t,x)$ we have*

$$2\mathcal{R}(t,x) \geq (\alpha+1)\mathsf{d}_*^2(x,y) - \Lambda\mathsf{p}(x,y) \tag{7.33}$$

If moreover $x \in \mathrm{M}(s,v)$ for some $(s,v) \in [0,T] \times X$ and (4.15) hold, then

$$2\mathcal{R}(t,x) \leq -(\alpha+1)\mathsf{d}_*^2(x,y) + \Lambda\mathsf{p}(x,y) + 2\mathsf{d}_*(v,x)\mathsf{d}_*(x,y) + 2L|t-s|\mathsf{d}_*(x,y) \tag{7.34}$$

so that

$$(2\alpha+1)\mathsf{d}_*^2(x,y) \leq 2\Lambda\mathsf{p}(x,y) + \mathsf{d}_*^2(v,x) + 2L|t-s|\mathsf{d}_*(x,y). \tag{7.35}$$

Proof. If $y \in \mathrm{M}(t,x)$ and γ is a curve connecting y to x as (4.11), we get

$$\begin{aligned}\mathcal{E}(t,y) + \mathsf{d}(x,y) + \frac{1}{2}\mathsf{d}_*^2(x,y) &\leq \mathcal{E}(t,\gamma(\theta)) + \mathsf{d}(x,\gamma(\theta)) + \frac{1}{2}\mathsf{d}_*^2(x,\gamma(\theta)) \\ &\leq (1-\theta)\mathcal{E}(t,y) + \theta\mathcal{E}(t,x) - \frac{\alpha}{2}\theta(1-\theta)\mathsf{d}_*^2(x,y) \\ &\quad + \frac{\Lambda}{2}\theta(1-\theta)\mathsf{p}(x,y) + \mathsf{d}(x,\gamma(\theta)) + \frac{1}{2}\mathsf{d}_*^2(x,\gamma(\theta)).\end{aligned}$$

We obtain by (4.13)

$$\theta(1-\theta)\frac{\alpha}{2}\mathsf{d}_*^2(x,y)+\frac{1}{2}(2\theta-\theta^2)\mathsf{d}_*^2(x,y)\le\theta\Big(\mathcal{E}(t,x)-\mathcal{E}(t,y)-\mathsf{d}(x,y)+\frac{\Lambda}{2}(1-\theta)\mathsf{p}(x,y)\Big)$$

Dividing by θ and passing to the limit as $\theta\downarrow 0$

$$\frac{\alpha+1}{2}\mathsf{d}_*^2(x,y)\le\mathcal{E}(t,x)-\mathcal{E}(t,y)-\mathsf{d}(x,y)-\frac{1}{2}\mathsf{d}_*^2(x,y)+\frac{\Lambda}{2}\mathsf{p}(x,y)=\mathcal{R}(t,x)+\frac{\Lambda}{2}\mathsf{p}(x,y)$$

which yields (7.33).

The proof of (7.34) is similar, but now we start from the minimality of x and consider the curve γ connecting x to y obtaining

$$\begin{aligned}\mathcal{E}(s,x)+\mathsf{d}(v,x)+\frac{1}{2}\mathsf{d}_*^2(v,x)&\le\mathcal{E}(s,\gamma(\theta))+\mathsf{d}(v,\gamma(\theta))+\frac{1}{2}\mathsf{d}_*^2(v,\gamma(\theta))\\&\le(1-\theta)\mathcal{E}(s,x)+\theta\mathcal{E}(s,y)-\frac{\alpha}{2}\theta(1-\theta)\mathsf{d}_*^2(x,y)\\&\qquad+\frac{\Lambda}{2}\theta(1-\theta)\mathsf{p}(x,y)+\mathsf{d}(v,x)+\mathsf{d}(x,\gamma(\theta))+\frac{1}{2}\mathsf{d}_*^2(v,\gamma(\theta)).\end{aligned}$$

so that

$$\begin{aligned}0\le\theta&\Big(\mathcal{E}(s,y)-\mathcal{E}(s,x)-\frac{\alpha}{2}(1-\theta)\mathsf{d}_*^2(x,y)\Big)\\&+\theta\Big(\frac{\Lambda}{2}(1-\theta)\mathsf{p}(x,y)+\mathsf{d}(x,y)+\frac{1}{2}\mathsf{d}_*(x,y)\big(\mathsf{d}_*(v,x)+\mathsf{d}_*(v,\gamma(\theta))\big)\Big).\end{aligned}$$

Dividing by θ and passing to the limit as $\theta\downarrow 0$ we get

$$0\le\mathcal{E}(s,y)-\mathcal{E}(s,x)-\frac{\alpha}{2}\mathsf{d}_*^2(x,y)+\frac{\Lambda}{2}\mathsf{p}(x,y)+\mathsf{d}(x,y)+\mathsf{d}_*(x,y)\mathsf{d}_*(v,x).$$

Adding $\mathcal{R}(t,x)=\mathcal{E}(t,x)-\mathcal{E}(t,y)-\frac{1}{2}\mathsf{d}_*^2(x,y)-\mathsf{d}(x,y)$ we get

$$\begin{aligned}\mathcal{R}(t,x)\le&-\frac{\alpha+1}{2}\mathsf{d}_*^2(x,y)+\frac{\Lambda}{2}\mathsf{p}(x,y)+\mathsf{d}_*(x,y)\mathsf{d}_*(v,x)\\&+\Big(\mathcal{E}(t,x)-\mathcal{E}(s,x)\Big)-\Big(\mathcal{E}(t,y)-\mathcal{E}(s,y)\Big),\end{aligned}$$

and estimating the last term by (4.15)

$$\Big(\mathcal{E}(t,x)-\mathcal{E}(s,x)\Big)-\Big(\mathcal{E}(t,y)-\mathcal{E}(s,y)\Big)=\int_s^t\Big(\mathcal{P}(r,x)-\mathcal{P}(r,y)\Big)\,\mathrm{d}r\le L|t-s|\mathsf{d}_*(x,y)$$

we obtain (7.34). (7.35) follows by combining (7.33) with (7.34) and using the elementary inequality $2\mathsf{d}_*(v,x)\mathsf{d}_*(x,y)\le\mathsf{d}_*^2(v,x)+\mathsf{d}_*^2(x,y)$. □

Proof of Theorem 4.3. If U_τ^n is a solution of the time incremental minimization scheme, we clearly have $U_\tau^{n+1}\in\mathrm{M}(t_\tau^{n+1},U_\tau^n)$ so that we can apply the previous Lemma 7.6 with $v:=U_\tau^{n-1}$, $x:=U_\tau^n$, $y:=U_\tau^{n+1}$ and $s=t_\tau^n,t=t_\tau^{n+1}$ obtaining

$$(2\alpha+1)\mathsf{d}_*^2(U_\tau^n,U_\tau^{n+1})\le\mathsf{d}_*^2(U_\tau^{n-1},U_\tau^n)+\Big(2\Lambda\mathsf{d}(U_\tau^n,U_\tau^{n+1})+2L\tau^{n+1}\Big)\mathsf{d}_*(U_\tau^n,U_\tau^{n+1}).\quad(7.36)$$

We can apply the discrete Gronwall lemma 7.5 with:

$$a_n = \mathsf{d}_*(U_\tau^n, U_\tau^{n+1}), \quad b_n := 2\Lambda \mathsf{d}(U_\tau^n, U_\tau^{n+1}) + 2L\tau^{n+1} \quad \gamma := 2\alpha$$

obtaining

$$\sum_{n=1}^{N_\tau - 1} \mathsf{d}_*(U_\tau^n, U_\tau^{n+1}) \le \frac{1}{\alpha}\Big(\frac{1}{2}\mathsf{d}_*^2(U_\tau^0, U_\tau^1) + 1 + LT + \Lambda \sum_{n=1}^{N_\tau - 1} \mathsf{d}(U_\tau^n, U_\tau^{n+1})\Big) \le \frac{(\Lambda+1)C_2 + 1 + LT}{\alpha}$$

where C_2 is the constant of (7.5): this estimate shows that the total variation $\mathrm{Var}_{\mathsf{d}_*}(\overline{U}_\tau, [0,T])$ is uniformly bounded w.r.t. τ, so that any pointwise limit of $\overline{U}_\tau$ belongs to $\mathrm{BV}_{\mathsf{d}_*}([0,T]; X)$. □

References

[1] V. Agostiniani, R. Rossi, and G. Savaré, *Rate-independent limits of gradient flows: a variational approach*, 2016.

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

[3] M. Artina, F. Cagnetti, M. Fornasier, and F. Solombrino, *Linearly constrained evolutions of critical points and an application to cohesive fractures*, ArXiv e-prints 1508.02965, (2015).

[4] F. Auricchio, A. Mielke, and U. Stefanelli, *A rate-independent model for the isothermal quasi-static evolution of shape-memory materials*, M^3AS Math. Models Meth. Appl. Sci., 18 (2008), pp. 125–164.

[5] G. Bouchitté, A. Mielke, and T. Roubíček, *A complete-damage problem at small strains*, Z. Angew. Math. Phys., 60 (2009), pp. 205–236.

[6] C. Castaing and M. Valadier, *Convex analysis and measurable multifunctions*, Springer-Verlag, Berlin, 1977. Lecture Notes in Mathematics, Vol. 580.

[7] G. Dal Maso, *An Introduction to Γ-Convergence*, vol. 8 of Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser, Boston, 1993.

[8] G. Dal Maso, A. DeSimone, and M. G. Mora, *Quasistatic evolution problems for linearly elastic-perfectly plastic materials*, Arch. Ration. Mech. Anal., 180 (2006), pp. 237–291.

[9] G. Dal Maso, A. DeSimone, M. G. Mora, and M. Morini, *Globally stable quasistatic evolution in plasticity with softening*, Netw. Heterog. Media, 3 (2008), pp. 567–614.

[10] G. Dal Maso, G. A. Francfort, and R. Toader, *Quasistatic crack growth in nonlinear elasticity*, Arch. Ration. Mech. Anal., 176 (2005), pp. 165–225.

[11] G. Dal Maso and R. Toader, *A model for the quasi-static growth of brittle fractures based on local minimization*, Math. Models Methods Appl. Sci., 12 (2002), pp. 1773–1799.

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

[13] G. Francfort and J. Marigo, *Revisiting brittle fracture as an energy minimization problem*, J. Mech. Phys. Solids, 46 (1998), pp. 1319–1342.

[14] 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), pp. 55–91.

[15] I. S. Gál, *On the fundamental theorems of the calculus*, Trans. Amer. Math. Soc., 86 (1957), pp. 309–320.

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

[17] D. Knees and M. Negri, *Convergence of alternate minimization schemes for phase field fracture and damage*, 2015.

[18] D. Knees, C. Zanini, and A. Mielke, *Crack growth in polyconvex materials*, Phys. D, 239 (2010), pp. 1470–1484.

[19] M. Kočvara, A. Mielke, and T. Roubíček, *A rate–independent approach to the delamination problem*, Math. Mech. Solids, 11 (2006), pp. 423–447.

[20] K. Kuratowski, *Sur l'espace des fonctions partielles*, Ann. Mat. Pura Appl. (4), 40 (1955), pp. 61–67.

[21] C. Larsen, *ϵ-stable quasi-static brittle fracture evolution*, Comm. Pure Appl. Math, (2009). To appear.

[22] A. Mainik and A. Mielke, *Existence results for energetic models for rate-independent systems*, Calc. Var. Partial Differential Equations, 22 (2005), pp. 73–99.

[23] ——, *Global existence for rate-independent gradient plasticity at finite strain*, J. Nonlinear Science, 19 (2009), pp. 221–248.

[24] A. Mielke, *Energetic formulation of multiplicative elasto–plasticity using dissipation distances*, Cont. Mech. Thermodynamics, 15 (2003), pp. 351–382.

[25] ——, *Existence of minimizers in incremental elasto–plasticity with finite strains*, SIAM J. Math. Analysis, 36 (2004), pp. 384–404.

[26] ——, *Complete-damage evolution based on energies and stresses*, Special Issue of Discrete Cont. Dyn. Syst. Ser. S, (2009). Submitted. WIAS preprint 1414.

[27] A. Mielke, R. Rossi, and G. Savaré, *Modeling solutions with jumps for rate-independent systems on metric spaces*, Discrete and Continuous Dynamical Systems A, 25 (2009).

[28] ——, *BV solutions and viscosity approximations of rate-independent systems*, ESAIM Control Optim. Calc. Var., 18 (2012), pp. 36–80.

[29] ——, *Variational convergence of gradient flows and rate-independent evolutions in metric spaces*, Milan J. Math., 80 (2012), pp. 381–410.

[30] ——, *Balanced viscosity (BV) solutions to infinite-dimensional rate-independent systems*, JEMS, to appear. ArXiv 1309.6291, (2013).

[31] ——, *Nonsmooth analysis of doubly nonlinear evolution equations*, Calc. Var. Partial Differential Equations, 46 (2013), pp. 253–310.

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

[33] ——, *Rate-independent systems*, vol. 193 of Applied Mathematical Sciences, Springer, New York, 2015. Theory and application.

[34] A. Mielke and F. Theil, *On rate-independent hysteresis models*, NoDEA Nonlinear Differential Equations Appl., 11 (2004), pp. 151–189.

[35] ——, *On rate-independent hysteresis models*, NoDEA Nonlinear Differential Equations Appl., 11 (2004), pp. 151–189.

[36] 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), pp. 137–177.

[37] ——, *A variational formulation of rate-independent phase transformations using an extremum principle*, Arch. Ration. Mech. Anal., 162 (2002), pp. 137–177.

[38] A. Mielke and A. Timofte, *An energetic material model for time-dependent ferroelectric behavior: existence and uniqueness*, Math. Meth. Appl. Sciences, 29 (2006), pp. 1393–1410.

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

[40] L. Minotti, *Visco-Energetic solutions to 1-dimensional rate-independent problems*, in preparation., (2016).

[41] ——, *Visco-Energetic solutions to rate-independent evolution problems*, PhD thesis, Pavia, 2016. in preparation.

[42] M. Negri, *An L^2 gradient flow and its quasi-static limit in phase-field fracture by alternate minimization*, 2016.

[43] M. Negri and C. Ortner, *Quasi-static crack propagation by Griffith's criterion*, Math. Models Methods Appl. Sci., 18 (2008), pp. 1895–1925.

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

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

[46] F. Schmid and A. Mielke, *Vortex pinning in super-conductivity as a rate-independent process*, Europ. J. Appl. Math., 16 (2005), pp. 799–808.

[47] S. Simons, *Minimax and monotonicity*, vol. 1693 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1998.