THE BREZIS–EKELAND–NAYROLES PRINCIPLE IN METRIC SPACES: TIME-CONTINUOUS SETTING

PIERRE–CYRIL AUBIN–FRANKOWSKI, GIACOMO ENRICO SODINI, AND ULISSE STEFANELLI

Abstract. Based on a variational characterization of the local slope of a functional, we extend the celebrated Brezis–Ekeland–Nayroles null-minimization principle to curves of maximal slope in metric spaces. In the time-continuous setting, we check that the ensuing global-in-time functionals admit minimizers.

1. Introduction

The Brezis–Ekeland–Nayroles principle [10, 11, 36, 37] is a variational approach to dissipative evolution systems, characterizing solutions as null-minimizers of suitable nonnegative functionals on trajectories. To introduce the principle, consider for simplicity a doubly nonlinear flow

$$\partial\hat\psi(u') + \partial\phi(u) = 0 \quad \text{a.e. in } [0,T] \tag{1.1}$$

where $u : [0,T] \to V$ is a trajectory in a reflexive Banach space V, the prime denotes differentiation, and ϕ, $\hat\psi : V \to [0,\infty)$ are convex functionals with single-valued subdifferentials $\partial\phi$, $\partial\hat\psi : V \to V^*$ (dual). By introducing a second trajectory $a : [0,T] \to V^*$, equation (1.1) can be equivalently written as the system $-a = \partial\hat\psi(u')$ and $a = \partial\phi(u)$ almost everywhere in $[0,T]$.

One can then characterize $\partial\phi : V \to V^*$ of ϕ at a point u with $\phi(u) < \infty$ as $a \in \partial\phi(u)$ iff $-\langle a,u\rangle + \phi(u) + \phi^*(a) \le 0$ where $\phi^*(a) = \sup_{v\in V} \langle a, v\rangle - \phi(v)$ is the classical *Fenchel conjugate* and $\langle\cdot,\cdot\rangle$ is the duality pairing between V^* and V. In the Banach case one hence has

$$a \in \partial\phi(u) \;\Leftrightarrow\; -\langle a, u-v\rangle + \phi(u) - \phi(v) \le 0 \quad \forall v \in V. \tag{1.2}$$

Using this characterization, the Brezis–Ekeland-Nayroles principle for doubly nonlinear flows [50, Thm. 1.2] asserts that solutions u to (1.1) correspond to null-minimizers of

$$\mathcal{H}_{\rm Banach}(a,u) = \left(\phi(u(T)) - \phi(u(0)) + \int_0^T \hat\psi(u')\,{\rm d}r + \int_0^T \hat\psi^*(a)\,{\rm d}r\right)^+ \\ + \int_0^T \sup_{v\in V}\big(-\langle a, u-v\rangle + \phi(u) - \phi(v)\big)\,{\rm d}r,$$

2010 *Mathematics Subject Classification.* 35K55.
Key words and phrases. Curves of maximal slope · Minimizing movements · Brezis–Ekeland–Nayroles variational principle.

denoting $x^+ = x \vee 0$. The second term in $\mathcal{H}_{\rm Banach}$ is 0 iff $a = \partial\phi(u)$ and, in this case, the first term is 0 iff $-a = \partial\hat\psi(u')$.

We are interested in extending the Brezis–Ekeland–Nayroles principle to evolution equations in metric spaces. Originating from the pioneering contributions [28, 39], flows in metric spaces have specifically received a strong development, motivated by their wide range of possible applications. The fundamental reference in the field is the monograph by Ambrosio, Gigli, & Savaré [2]. In a metric space, however, no notion of derivative or duality is available and one has to resort to *scalar* fields: the *metric derivative* $|u'|$ of u and $u \mapsto |\partial\phi|(u)$ the *local slope* of ϕ.

A prominent notion of dissipative evolution in the complete metric space (U, d) is that of *curve of maximal slope*. This is a trajectory $u : [0, T] \to U$ solving

$$\phi(u(t)) - \phi(u(0)) + \int_0^t \psi\big(|u'|(r)\big)\,{\rm d}r + \int_0^t \psi^*\big(|\partial\phi|(u(r))\big)\,{\rm d}r = 0 \quad \forall t \in [0, T] \tag{1.3}$$

where the functional $\phi : U \to [0, \infty]$ and the strictly convex, smooth function $\psi : [0, \infty) \to [0, \infty)$ with Fenchel conjugate ψ^* are given.

In the classical *geodesically-convex* case (a setting which we will nonetheless soon abandon) our extension of the classical Brezis–Ekeland–Nayroles principle is based on the global variational characterization of the local slope (which in this case coincides with the global slope) given by

$$\alpha = |\partial\phi|(u) \;\Leftrightarrow\; \alpha = \arg\min\big\{\overline{\alpha} \geq 0 : -\overline{\alpha}d(u, v) + \phi(u) - \phi(v) \leq 0 \;\forall v \in U\big\}, \tag{1.4}$$

see Proposition 3.3 below. By letting

$$E(\alpha, u) = \max_{v\in U}\big\{ - \alpha d(u, v) + \phi(u) - \phi(v)\big\}$$

(see Lemma 3.1 below), we define the analogue functional $\mathcal{H}$ to $\mathcal{H}_{\rm Banach}$ over entire trajectories $u : [0, T] \to U$ and $\alpha : [0, T] \to [0, \infty)$ as

$$\mathcal{H}(\alpha, u) = \left(\phi(u(T)) - \phi(u(0)) + \int_0^T \psi\big(|u'|^p(r)\big)\,{\rm d}r + \int_0^T \psi^*\big(\alpha(r)\big)\,{\rm d}r\right)^+ + \int_0^T E(\alpha(r), u(r))\,{\rm d}r.$$

The *metric Brezis–Ekeland–Nayroles principle* ensues by checking in Theorem 4.1 that

$$\boxed{u \text{ is a curve of maximal slope } \Leftrightarrow \mathcal{H}(\alpha, u) = \min \mathcal{H} = 0} \tag{1.5}$$

for some α. In this case, $\alpha = |\partial\phi|(u)$ almost everywhere in $[0, T]$. This provides a variational characterization of curves of maximal slope, opening the way to applying the methods of the calculus of variations to the evolution problem. In particular, the vectorial variable $a(t) \in V^*$ in $\mathcal{H}_{\rm Banach}$ is replaced by the scalar one $\alpha(t) \geq 0$ in $\mathcal{H}$ and the duality $-\langle a, u - v\rangle$ is replaced by $-\alpha d(u, v)$. As an effect of this generalization, the analysis of $\mathcal{H}$ is new even in the Banach space case.

In this paper, we analyze the above approach in the time-continuous setting. After checking characterization (1.5) (Theorem 4.1), we prove that $\mathcal{H}$ is lower semicontinuous (Proposition 5.2) and that it admits minimizers (Proposition 5.3). In this paper, we do not explicitly prove that these minimizers are in fact null-minimizers, i.e., $\min_{K(u^0)} \mathcal{H} = 0$. This is however true, as curves of maximal slope do exist, see also [3].

Besides $\mathcal{H}$, other functionals may be used to obtain alternative metric Brezis–Ekeland–Nayroles principles. In this paper, we also investigate the functional $\mathcal{G}$ given by

$$\mathcal{G}(\alpha, u) = \begin{cases} \phi(u(T)) - \phi(u(0)) + \int_0^T \psi\big(|u'|(r)\big)\mathrm{d}r + \int_0^T \psi^*\big(\alpha(r)\big)\,\mathrm{d}r \\ \qquad \text{if } \displaystyle\int_0^T E(\alpha(r), u(r))\,\mathrm{d}r = 0, \\ \infty \quad \text{otherwise,} \end{cases}$$

and prove that the characterization (1.5) holds for $\mathcal{G}$ as well (Theorem 4.1), $\mathcal{G}$ is lower semicontinuous (Proposition 5.2), and that it admits minimizers (Proposition 5.3). Moreover, we also consider a penalized version of $\mathcal{G}$. For all $\delta > 0$, we let the approximating functional $\mathcal{G}_\delta$ be given by

$$\mathcal{G}_\delta(\alpha, u) = \phi(u(T)) - \phi(u(0)) + \int_0^T \psi\big(|u'|(r)\big)\mathrm{d}r + \int_0^T \psi^*\big(\alpha(r)\big)\,\mathrm{d}r \\ + \frac{1}{\delta}\int_0^T E(\alpha(r), u(r))\,\mathrm{d}r,$$

we show that minimizers of $\mathcal{G}_\delta$ converge as $\delta \to 0$ to a curve of maximal slope up to subsequences (Proposition 5.6).

Related works. Originally introduced in the linear parabolic setting, the Brezis–Ekeland–Nayroles approach has been extended to cover gradient flows [23] and their quadratic perturbations [20], long-time dynamics [29], compactness [60], and time discretizations [52]. Extensions to general monotone and semimonotone flows [56, 58], doubly nonlinear flows [50, 55, 59], rate-independent flows [57], Hamiltonian flows [12, 22, 21, 24], and second-order in time problems [30, 31] are also available. A far-reaching recast of the Brezis–Ekeland–Nayroles principle in terms of (anti-)selfdual Lagrangians can be found in [19]. Applications have been developed to stochastic PDEs [6, 9], optimal control [4, 42], nonlinear diffusion [32, 41], elastoplasticity [13, 51], conservation laws [40], identification [5], and deep learning [14], among many others.

Beside the variants of the Brezis–Ekeland–Nayroles principle mentioned above, global-in-time variational approaches to dissipative evolution problems have a long tradition, dating back at least to Biot's work on irreversible thermodynamics [7] and Gurtin's principle for viscoelasticity and elastodynamics [25, 26], see also [27]. In [54], solutions to doubly nonlinear problems are obtained as minimal elements of a certain partial-order relation on the trajectories. Related to the specific form of the first term in $\mathcal{H}$, one should also mention the *Energy-Dissipation Principle*, which has attracted attention as an effective tool to identify limits and check stability, see [17, 18, 33, 34, 47, 48] among many others. An alternative global variational approach is the so-called *Weighted Inertia-Dissipation-Energy* principle, where solutions are obtained as limits of subsequences of minimizers of parametrized functionals over trajectories [53]. In the setting of gradient flows, the WIDE principle has already been extended to the case of metric spaces in [44, 45].

In the second paper in this series [3], we focus on *minimizing-movements*, a reference method to construct curves of maximal slope through time discretization [1]. Therein we introduce new schemes that arise as *variational integrators* of time-continuous functionals $\mathcal{H}$, $\mathcal{G}$, and $\mathcal{G}_\delta$. These schemes are shown to admit solutions, whose time-interpolants converge up to subsequences to curves of maximal slope. General conditional convergence results are also deduced, applying in particular to the case of incremental quasiminimizers. In fact, these schemes may serve as a-posteriori convergence tests, independently of the method used to produce the discrete solutions.

The paper is organized as follows. We specify our assumptions in Section 2. The variational characterization of the local slope (1.4) is then discussed in Section 3. We present the metric Brezis–Ekeland–Nayroles principle in Section 4 and prove the existence of minimizers in Section 5. Finally, in Section 6 we present a further generalization of the metric Brezis–Ekeland–Nayroles principle under even weaker assumptions on the driving functional ϕ.

2. Setting

In this section, we collect assumptions (2.6)-(2.12) and fix the notation. To start with, we assume that

$$(U, d) \text{ is a complete, separable, path-connected metric space, and } T \in (0, \infty). \tag{2.6}$$

The driving functional ϕ is requested to be

$$\phi : U \to (-\infty, \infty] \text{ proper and with compact sublevels.} \tag{2.7}$$

A consequence of the compactness of the sublevels of ϕ is that ϕ is lower semicontinuous and bounded from below. In particular, (2.7) implies that the *effective domain* $D(\phi) \coloneqq \{u \in U : \phi(u) < \infty\}$ is nonempty and that ϕ admits a global minimizer. With no loss of generality, we henceforth assume that

$$\min \phi = 0. \tag{2.8}$$

For given $p \in (1, \infty)$, following the notation in [2], the set $AC^p([0, T]; U)$ indicates *p-absolutely continuous* curves, namely, trajectories $u : [0, T] \to U$ such that there exists $\eta \in L^p(0, T)$ with

$$d(u(s), u(t)) \le \int_s^t \eta(r)\,\mathrm{d}r \quad \forall 0 \le s \le t < T.$$

If $u \in AC^p([0, T]; U)$, the limit

$$|u'|(t) \coloneqq \lim_{s \to t} \frac{d(u(s), u(t))}{|t - s|}$$

exists for almost all $t \in (0, T)$, see [2, Thm. 1.1.2, p. 24], and is called the *metric derivative* of u at t. In this case, the map $t \mapsto |u'|(t)$ is in $L^p(0, T)$ and coincides with the minimal function $\eta \in L^p(0, T)$ satisfying the above bound.

The *local slope* $|\partial\phi|(u)$ and the *global slope* $\mathfrak{l}_\phi(u)$ [2, 15, 16] of ϕ at $u \in U$ are defined via

$$|\partial\phi|(u) \coloneqq \limsup_{v\to u} \frac{(\phi(u)-\phi(v))^+}{d(u,v)}, \quad \mathfrak{l}_\phi(u) \coloneqq \sup_{v\neq u} \frac{(\phi(u)-\phi(v))^+}{d(u,v)},$$

respectively, and we indicate the corresponding domains as $D(|\partial\phi|) \coloneqq \{u \in D(\phi) : |\partial\phi|(u) < \infty\}$ and $D(\mathfrak{l}_\phi) \coloneqq \{u \in D(\phi) : \mathfrak{l}_\phi(u) < \infty\}$. Note that the definition of $|\partial\phi|(u)$ implicitly requires that u is not isolated which actually cannot occur as U is path-connected by (2.7).

Being defined via a supremum, the global slope $\mathfrak{l}_\phi$ is lower semicontinuous. On the other hand, the local slope $|\partial\phi|$ is defined via a lim sup and is not, in general, lower semicontinuous. In order to approximate $|\partial\phi|$, we introduce a penalized formulation of $\mathfrak{l}_\phi$ through a $\mu \geq 0$. For some given $q > 1$ and for all $\mu \geq 0$, we define the *μ-slope* $|\partial\phi|_\mu(u)$ of ϕ at u as

$$|\partial\phi|_\mu(u) \coloneqq \sup_{v\neq u} \left(\frac{(\phi(u)-\phi(v))^+}{d(u,v)} - \mu d^{q-1}(u,v) \right). \tag{2.9}$$

Note that $|\partial\phi|_\mu$ is lower semicontinuous [46, Prop. 2.7] and that $|\partial\phi|_0 = \mathfrak{l}_\phi$.

We say that a function $g : U \to [0,\infty]$ is a *strong upper gradient* for ϕ [2, Def. 1.2.1, p. 27] if for all $u \in AC^p([0,T];U)$, the map $t \mapsto g(u(t))$ is Borel and

$$|\phi(u(t)) - \phi(u(s))| \leq \int_s^t g(u(r))\,|u'|(r)\,\mathrm{d}r \quad \forall 0 \leq s \leq t \leq T.$$

If $t \mapsto g(u(t))|u'|(t) \in L^1(0,T)$ the latter entails that $\phi \circ u \in W^{1,1}(0,T)$ and $(\phi\circ u)' \leq g(u)|u'|$ almost everywhere in $[0,T]$. Note that $|\partial\phi|_\mu$ is a strong upper gradient for all $\mu \in [0,\infty)$ based on [2, proof of Corollary 2.4.10]. In particular, $\mathfrak{l}_\phi$ is a strong upper gradient.

In what follows, we will assume that

$$\exists \mu_0 \in [0,\infty) \text{ such that } |\partial\phi| = |\partial\phi|_{\mu_0}. \tag{2.10}$$

The existence of such μ_0 entails in particular that $|\partial\phi|$ is lower semicontinuous and a strong upper gradient for ϕ. Assumption (2.10) in particular holds if ϕ is (λ,q)-*generalized-geodesically convex* for some $\lambda \in \mathbb{R}$ [46] (see also [38, 49] for alternative but comparable definitions). In this case, one can choose $\mu_0 = \lambda^-/q$, where $\lambda = -\lambda \vee 0$. In Section 6, we drop this assumption (2.10).

The initial value u^0 is asked to fulfill

$$u^0 \in D(\phi). \tag{2.11}$$

The dissipative character of the evolution is quantified via the function $\psi : \mathbb{R} \to [0,\infty)$ satisfying

$$\begin{aligned}
&\psi \in C^1(\mathbb{R}) \text{ convex with} \\
&\psi(0) = \min\psi = 0 \text{ such that there exist } p \in (1,\infty) \text{ and } c_\psi \in (0,\infty) \text{ with} \\
&\psi(t) + \psi^*(t^*) \geq c_\psi|t|^p + c_\psi|t^*|^{p'} - \frac{1}{c_\psi} \quad \forall t,\, t_* \in \mathbb{R}
\end{aligned} \tag{2.12}$$

Here, $1/p + 1/p' = 1$ and ψ^* is the classical Fenchel conjugate $\psi^*(t^*) = \sup_t(t^*t - \psi(t))$ for $t^* \in \mathbb{R}$. The reference example is $\psi(t) = |t|^p/p$. By duality one has that

$$\psi(t) + \psi^*(t^*) \le C_\psi(1 + |t|^p + |t^*|^{p'}) \quad \forall t,\, t_* \in \mathbb{R} \tag{2.13}$$

for some $C_\psi \in (0,\infty)$ depending on c_ψ and p. Indeed, by (2.12) evaluated first at $t^* = 0$ and then at $t = 0$, we have

$$\psi(t) \ge c_\psi|t|^p - \frac{1}{c_\psi}, \quad \psi^*(t^*) \ge c_\psi|t^*|^{p'} - \frac{1}{c_\psi} \quad \forall t, t^* \in \mathbb{R},$$

so that

$$\psi(t) = \psi^{**}(t) \le \sup_{t^*}\{tt^* - c_\psi|t^*|^{p'}\} + \frac{1}{c_\psi} = \frac{1}{p(c_\psi p')^{p-1}}|t|^p + \frac{1}{c_\psi} \quad \forall t \in \mathbb{R},$$

$$\psi^*(t^*) \le \sup_{t}\{tt^* - c_\psi|t|^{p}\} + \frac{1}{c_\psi} = \frac{1}{p'(c_\psi p)^{p'-1}}|t|^{p'} + \frac{1}{c_\psi} \quad \forall t^* \in \mathbb{R},$$

which give (2.13) with $C_\psi := \max\{2c_\psi^{-1}, (p'(c_\psi p)^{p'-1})^{-1}, (p(c_\psi p')^{p-1})^{-1}\}$. Note that the convexity of ψ and the fact that $0 = \psi(0) = \min\psi$ imply that ψ is nondecreasing on $[0,\infty)$. On the other hand, the differentiability of ψ ensures that ψ^* is strictly convex, which, together with $0 = \psi^*(0) = \min\psi^*$, gives that ψ^* is strictly increasing on $[0,\infty)$.

Assumptions (2.6)–(2.7) and (2.10)–(2.12) are tacitly assumed throughout the paper, with the exception of Section 6 where only the assumption (2.10) is dropped.

In this paper, we focus on the following classical notion [35] in studying the evolution of gradient flows.

Definition 2.1 (Curve of maximal slope)**.** *The trajectory $u \in AC^p([0,T];U)$ is said to be a* curve of maximal slope *if, $u(0) = u^0$, $\phi \circ u \in W^{1,1}(0,T)$, and*

$$\phi(u(t)) + \int_0^t \psi(|u'|(r))\,\mathrm{d}r + \int_0^t \psi^*(|\partial\phi|(u(r)))\,\mathrm{d}r = \phi(u(0)) \quad \forall t \in [0,T]. \tag{2.14}$$

The latter is a slight generalization of the concept of (p,q)*-curves of maximal slope* from [2, Def. 1.3.2, p. 32], which indeed correspond to the choice $\psi(t) = |t|^p/p$.

To conclude this preliminary section, let us collect here some compactness tools, which are used in the following.

Lemma 2.2 (Compactness tools)**.** *Let $(u_j)_j \subset AC^p([0,T];U)$ have $|u_j'| \to \eta$ weakly in $L^p(0,T)$.*

(a) *In case that*

$$\sup_j \sup_{t\in[0,T]} \phi(u_j(t)) < \infty,$$

there exist $u \in AC^p([0,T];U)$ with $|u'| \le \eta$ almost everywhere in $[0,T]$ and a not relabeled sequence such that $u_j(t) \to u(t)$ for all $t \in [0,T]$.

(b) *In case that, for some $t_0 \in [0,T]$, $(u_j(t_0))_j$ is precompact and*

$$\sup_j \int_0^T \phi(u_j(t))\,\mathrm{d}t < \infty,$$

there exist $u \in AC^p([0,T];U)$ with $|u'| \le \eta$ almost everywhere in $[0,T]$ and a not relabeled sequence such that $u_j(t) \to u(t)$ for almost all $t \in [0,T]$ and $u_j(t_0) \to u(t_0)$.

The results (a) and (b) above correspond to the Ascoli–Arzelà theorem and a metric version of the Aubin–Lions lemma, and are proved in [2, Prop. 3.3.1] and [43, Thm. 2]. The only difference between the result (b) above and the one presented in [43, Thm. 2] is the convergence at the additional point t_0, which we briefly comment here. Since $(u_j(t_0))_j$ is assumed to be precompact, we can first extract a not relabeled subsequence such that $u_j(t_0) \to \overline{u} \in U$. On the other hand, the result by [43, Thm. 2] gives the existence of $u \in AC^p([0,T];U)$ with $|u'| \le \eta$ almost everywhere in $[0,T]$ and of a not relabeled sequence such that $u_j(t) \to u(t)$ for almost all $t \in [0,T]$. In particular, we have

$$\begin{aligned} d(u_j(t_0), u(t_0)) &\le d(u_j(t_0), u_j(t)) + d(u_j(t), u(t)) + d(u(t), u(t_0)), \\ &\le \int_{t_0 \wedge t}^{t_0 \vee t} |u_j'|(r)\mathrm{d}r + d(u_j(t), u(t)) + d(u(t), u(t_0)) \end{aligned}$$

for every $t \in [0,T]$ such that $u_j(t) \to u(t)$. Passing to the limit as $j \to \infty$, we get

$$d(\overline{u}, u(t_0)) \le \int_{t_0 \wedge t}^{t_0 \vee t} \eta(r)\mathrm{d}r + d(u(t), u(t_0))$$

and passing to the limit as $t \to t_0$, we obtain $u(t_0) = \overline{u}$.

3. Variational characterization of $|\partial\phi|(u)$

This section pertains to some introductory discussion concerning the variational characterization of the slope $|\partial\phi|$. Let us start by defining the functional $\overline{E} : [0,\infty) \times U \times D(\phi) \to (-\infty, \infty]$ as

$$\overline{E}(\alpha, u, v) := -\alpha d(u,v) + \phi(u) - \phi(v) - \mu_0 d^q(u,v).$$

We first observe the following.

Lemma 3.1 (Reduced functional)**.** *For all $(\alpha, u) \in [0,\infty) \times U$ there exists the maximum*

$$E(\alpha, u) := \max_{v \in D(\phi)} \overline{E}(\alpha, u, v) \tag{3.15}$$

which defines the reduced functional $E : [0,\infty) \times U \to [0,\infty]$.

Proof. If $u \notin D(\phi)$, then $\overline{E}(\alpha,u,v) = \infty$ for every $(\alpha, v) \in [0,\infty) \times D(\phi)$, so that the maximum in (3.15) is ∞ and is achieved at any $v \in D(\phi)$. Assume now that $u \in D(\phi)$. As ϕ is lower semicontinuous, for all $(\alpha, u) \in [0,\infty) \times D(\phi)$ the map $v \in U \mapsto \overline{E}(\alpha,u,v) = -\alpha d(u,v) + \phi(u) - \phi(v) - \mu_0 d^q(u,v) \in [-\infty, \infty)$ is upper semicontinuous. Moreover, it is finite iff $v \in D(\phi)$. Letting $u_* \in \arg\min \phi$ and using $\phi(u_*) = 0$ by (2.8), we can reduce the maximization in (3.15) to those $v \in D(\phi)$ such that

$$-\alpha d(u,v) - \phi(v) - \mu_0 d^q(u,v) \ge -\alpha d(u,u_*) - \mu_0 d^q(u,u_*).$$

Hence, one can maximize over the compact set $\{v \in D(\phi) : \phi(v) \le \alpha d(u,u_*) + \mu_0 d^q(u,u_*)\}$. The existence of a maximizer follows by the Direct Method. □

Note that we are not claiming uniqueness for problem (3.15). This nonetheless follows if $v \mapsto \alpha d(u,v) + \phi(v) + \mu_0 d^q(u,v)$ is *strictly* geodesically convex for all (α, u). This holds independently of $\alpha \geq 0$ if ϕ is strictly geodesically convex and both $v \mapsto d(u,v)$ and $v \mapsto \mu_0 d^q(u,v)$ are convex along any geodesic.

Note that $E(\alpha,u) \geq \overline{E}(\alpha,u,u) = 0$ for all $\alpha \geq 0$ and $u \in D(\phi)$. Moreover, E is convex with respect to α, being the supremum of linear functions in α. We have $E(\alpha,u) = \infty$ for all $\alpha \geq 0$ iff $u \notin D(\phi)$. We collect some properties of E in the following.

Proposition 3.2 (Properties of E)**.** *The reduced functional $E : [0,\infty) \times U \to [0,\infty]$ is lower semicontinuous and there exists a Borel map $\widehat{v} : [0,\infty) \times U \to U$ such that $E(\alpha,u) = \overline{E}(\alpha,u,\widehat{v}(\alpha,u))$ for all $(\alpha,u) \in [0,\infty) \times U$.*

Proof. The lower semicontinuity is immediate as the maps $(\alpha,u) \in [0,\infty) \times U \mapsto \overline{E}(\alpha,u,v)$ are lower semicontinuous for all $v \in D(\phi)$ and $E(\alpha,u)$ is defined as a maximum on v. This also implies that E is a Borel map.

Let $\Gamma \subset [0,\infty) \times U \times U$ be defined as

$$\Gamma \coloneqq \{(\alpha,u,v) \in [0,\infty) \times U \times D(\phi) : \overline{E}(\alpha,u,v) = E(\alpha,u)\}.$$

Since $\overline{E}$ and E are Borel maps, then Γ is a Borel set. Its sections $\Gamma_{\alpha,u}$, given, for every $(\alpha,u) \in [0,\infty) \times U$, by

$$\Gamma_{\alpha,u} \coloneqq \{v \in U : (\alpha,u,v) \in \Gamma\} = \{v \in D(\phi) : \overline{E}(\alpha,u,v) = E(\alpha,u)\},$$

are not empty by Lemma 3.1 and compact by the upper semicontinuity of $D(\phi) \ni v \mapsto \overline{E}(\alpha,u,v)$ and by the compactness of the sublevels of ϕ. By [8, Theorem 6.9.6] we obtain the existence of a Borel map $\widehat{v} : [0,\infty) \times U \to U$ such that its graph is contained in Γ. □

Under assumption (2.10), for all $u \in U$ one has that

$$|\partial\phi|(u) \leq \alpha \;\Leftrightarrow\; \overline{E}(\alpha,u,v) = -\alpha d(u,v) + \phi(u) - \phi(v) - \mu_0 d^q(u,v) \leq 0 \quad \forall v \in U. \tag{3.16}$$

As $E(\alpha,u) \geq 0$ for all $u \in D(\phi)$, the latter yields

$$|\partial\phi|(u) \leq \alpha \;\Leftrightarrow\; E(\alpha,u) = 0. \tag{3.17}$$

This implies the following variational characterization, which corresponds to a metric version of the *Fenchel identity* in Banach spaces.

Proposition 3.3 (Characterization)**.** *Given $u \in D(\phi)$, we have that*

$$u \in D(|\partial\phi|) \;\Leftrightarrow\; A_u = \{\alpha \geq 0 : E(\alpha,u) = 0\} \neq \emptyset.$$

In this case, $|\partial\phi|(u) = \min A_u$.

Proof. If $u \in D(|\partial\phi|)$ one has that $E(|\partial\phi|(u),u) = 0$ from (3.17), hence $|\partial\phi|(u) \in A_u \neq \emptyset$. Conversely, if $\alpha \in A_u$ then $E(\alpha,u) = 0$ so that $|\partial\phi|(u) \leq \alpha < \infty$ and $u \in D(|\partial\phi|)$. In this case, again (3.17) guarantees that $|\partial\phi|(u)$ is minimal in A_u. □

Proposition 3.3 ensures that for all $u \in D(\phi)$ one can variationally characterize $|\partial\phi|(u)$ as follows

$$|\partial\phi|(u) = \operatorname*{arg\,min}_{E(\alpha,u)=0} \psi^*(\alpha), \tag{3.18}$$

where we recall that ψ^* from (2.12) is strictly increasing on $[0,\infty)$. As E is nonnegative, this provides the possibility of penalizing the equation $\alpha = |\partial\phi|(u)$ by considering an unconstrained minimization problem.

Proposition 3.4 (Penalization)**.** *For all $u \in D(\phi)$ and $\delta > 0$ one can find*

$$\alpha_\delta \in \operatorname*{arg\,min}_{\alpha\geq 0} \left\{ \psi^*(\alpha) + \frac{1}{\delta} E(\alpha, u) \right\}. \tag{3.19}$$

If $u \in D(|\partial\phi|)$, we have that $\alpha_\delta \to |\partial\phi|(u)$ as $\delta \to 0$.

Proof. Let us first check that the minimization problem (3.19) can be solved. As E is nonnegative and finite on $D(\phi)$ and ψ^* is coercive, one has that $\alpha \in [0,\infty) \mapsto \psi^*(\alpha) + E(\alpha,u)/\delta$ is coercive, as well. The existence of α_δ for all $\delta > 0$ follows then by the lower semicontinuity of ψ^* and by Proposition 3.2.

Assume now that $u \in D(|\partial\phi|)$. By comparing with $|\partial\phi|(u)$ and using the fact that $E(|\partial\phi|(u), u) = 0$ we obtain that

$$\psi^*(\alpha_\delta) + \frac{1}{\delta} E(\alpha_\delta, u) \leq \psi^*(|\partial\phi|(u)) + \frac{1}{\delta} E(|\partial\phi|(u), u) = \psi^*(|\partial\phi|(u)). \tag{3.20}$$

On the one hand, as E is nonnegative this proves that α_δ is bounded independently of δ, so that $\alpha_\delta \to \alpha$ (up to some not relabeled subsequence) as $\delta \to 0$. On the other hand, as ψ^* is nonnegative, inequality (3.20) implies that

$$0 \leq E(\alpha, u) \leq \liminf_{\delta\to 0} E(\alpha_\delta, u) \leq \liminf_{\delta\to 0} \delta\, \psi^*(|\partial\phi|(u)) = 0$$

proving that $E(\alpha, u) = 0$, or $|\partial\phi|(u) \leq \alpha$ by (3.17). Hence, we get that $\psi^*(|\partial\phi|(u)) \leq \psi^*(\alpha)$ since ψ^* is (strictly) increasing on $[0,\infty)$. Again by (3.20), we get

$$\psi^*(\alpha) \leq \liminf_{\delta\to 0} \psi^*(\alpha_\delta) \leq \psi^*(|\partial\phi|(u))$$

which implies $\alpha = |\partial\phi|(u)$, as ψ^* is strictly increasing on $[0,\infty)$. The above argument ensures that any subsequence of $(\alpha_\delta)_\delta$ has a converging subsequence to the same limit. This proves that the whole sequence $(\alpha_\delta)_\delta$ converges to $|\partial\phi|(u)$ and no extraction is actually necessary. $\square$

4. Metric Brezis–Ekeland–Nayroles principle

In this section, we introduce the metric version of the Brezis–Ekeland–Nayroles principle (Theorem 4.1) and discuss some of its properties. Let us endow $[0,T]$ with the Lebesgue measure and indicate by $M([0,T];U)$ and $M([0,T];[0,\infty)\times U)$ the sets of Borel measurable maps with values in U and $[0,\infty)\times U$, respectively. Owing to Proposition 3.2, for all $(\alpha,u) \in M([0,T];[0,\infty)\times U)$ we have that $t \in [0,T] \mapsto E(\alpha(t),u(t))$ is measurable and $v := \widehat{v}(\alpha,u) \in M([0,T];U)$ is such that $E(\alpha(t),u(t)) = \overline{E}(\alpha(t),u(t),v(t))$ for all $t \in [0,T]$. Recalling that E is nonnegative, we define $\mathcal{E} : M([0,T];[0,\infty)\times U) \to [0,\infty]$ as

$$\mathcal{E}(\alpha,u) = \int_0^T E(\alpha(r),u(r))\,\mathrm{d}r.$$

Analogously to the static setting, we have that $\mathcal{E}$ is nonnegative and $\mathcal{E}(\alpha, u) < \infty$ implies that $u \in D(\phi)$ almost everywhere in $[0, T]$.

By recalling that $u^0 \in D(\phi)$ from (2.11), we indicate by

$$K(u^0) := \{(\alpha, u) \in L^{p'}(0,T) \times AC^p([0,T]; U) : u(0) = u^0\} \tag{4.21}$$

the set of *admissible trajectories*. Moreover, we define the functionals $\mathcal{F} : L^{p'}(0,T) \times AC^p([0,T]; U) \to (-\infty, \infty]$ and $\mathcal{G}$, $\mathcal{G}_\delta, \mathcal{H} : L^{p'}(0,T) \times AC^p([0,T]; U) \to (-\infty, \infty]$ for $\delta > 0$ as

$$\begin{aligned}
&\mathcal{F}(\alpha, u) = \phi(u(T)) - \phi(u(0)) + \int_0^T \psi(|u'|(r))\,\mathrm{d}r + \int_0^T \psi^*(\alpha(r))\,\mathrm{d}r,\\
&\mathcal{G}(\alpha, u) = \begin{cases} \mathcal{F}(\alpha, u) & \text{if } \mathcal{E}(\alpha, u) = 0,\\ \infty & \text{otherwise,} \end{cases}\\
&\mathcal{G}_\delta(\alpha, u) = \mathcal{F}(\alpha, u) + \frac{1}{\delta}\mathcal{E}(\alpha, u),\\
&\mathcal{H}(\alpha, u) = (\mathcal{F}(\alpha, u))^+ + \mathcal{E}(\alpha, u).
\end{aligned}$$

We are now ready to present our metric version of the Brezis–Ekeland–Nayroles principle [10, 11, 36, 37], delivering a characterization of curves of maximal slope. As a preparation, let us remark that both $\mathcal{H}$ and $\mathcal{G}$ are nonnegative on $K(u^0)$. Indeed, $\mathcal{H} \geq 0$ trivially follows as $\mathcal{E} \geq 0$. Assume by contradiction that $\mathcal{G}(\alpha, u) < 0$ for some $(\alpha, u) \in K(u^0)$. As $\mathcal{E}(\alpha, u) = 0$, we have that $|\partial\phi|(u) \leq \alpha$ almost everywhere in $[0, T]$ by (3.17), and $\mathcal{F}(\alpha, u) = \mathcal{G}(\alpha, u) < 0$. As $\psi^*(|\partial\phi|(u)) \leq \psi^*(\alpha)$ almost everywhere in $[0, T]$, we have $\mathcal{F}(|\partial\phi|(u), u) \leq \mathcal{F}(\alpha, u) < 0$. On the other hand, as $|\partial\phi|$ is a strong upper gradient, one finds $\phi \circ u \in W^{1,1}(0, T)$ and the chain rule yields

$$\begin{aligned}
\mathcal{F}(|\partial\phi|(u), u) &= \phi(u(T)) - \phi(u(0)) + \int_0^T \psi(|u'|(r))\,\mathrm{d}r + \int_0^T \psi^*(|\partial\phi|(u(r)))\,\mathrm{d}r\\
&\geq \int_0^T \big(- |\partial\phi|(u(r))|u'|(r) + \psi(|u'|(r)) + \psi^*(|\partial\phi|(u(r))\big)\mathrm{d}r \geq 0
\end{aligned}$$

from Fenchel's inequality, leading to a contradiction.

Theorem 4.1 (Null-minimization)**.** *The following are equivalent:*

(i) *u is a curve of maximal slope,*
(ii) *$\mathcal{G}(\alpha, u) = \min_{K(u^0)} \mathcal{G} = 0$ for some $\alpha \in L^{p'}(0, T)$,*
(iii) *$\mathcal{H}(\alpha, u) = \min_{K(u^0)} \mathcal{H} = 0$ for some $\alpha \in L^{p'}(0, T)$.*

Proof. Let u be a curve of maximal slope. Then $u(0) = u^0$, $|\partial\phi|(u) \in L^{p'}(0, T)$, and (2.14) for $t = T$ gives $\mathcal{F}(|\partial\phi|(u), u) = 0$. Since in particular $u \in D(|\partial\phi|)$ almost everywhere in $[0, T]$, we have that $\mathcal{E}(|\partial\phi|(u), u) = 0$ and $\mathcal{G}(\alpha, u) = \mathcal{H}(\alpha, u) = 0$ follows by choosing $\alpha = |\partial\phi|(u)$. Hence (i) implies both (ii) and (iii).

We now check (ii) $\Rightarrow$ (i). Let $(\alpha, u) \in K(u^0)$ with $\mathcal{G}(\alpha, u) = 0$. This implies that $\mathcal{E}(\alpha, u) = 0$, so that $\mathcal{F}(\alpha, u) = \mathcal{G}(\alpha, u) = 0$. As $|\partial\phi|(u) \leq \alpha$ almost everywhere in $[0, T]$ by (3.17), we have proved that $u \in D(|\partial\phi|)$ almost everywhere in $[0, T]$

and $|\partial\phi|(u) \in L^{p'}(0,T)$. As $|\partial\phi|$ is a strong upper gradient for ϕ, this implies that $\phi \circ u \in W^{1,1}(0,T)$. In particular, for all $t \in [0,T]$ we have that

$$\begin{aligned}\phi(u^0) - \phi(u(t)) &\le |\phi(u(t)) - \phi(u^0)| \le \int_0^t |\partial\phi|(u(r))\,|u'|(r)\,\mathrm{d}r \\ &\le \int_0^t \psi(|u'|(r))\,\mathrm{d}r + \int_0^t \psi^*(|\partial\phi|(u(r)))\,\mathrm{d}r\end{aligned}$$

giving the inequality

$$\phi(u(t)) - \phi(u^0) + \int_0^t \psi(|u'|(r))\,\mathrm{d}r + \int_0^t \psi^*(|\partial\phi|(u(r)))\,\mathrm{d}r \ge 0. \tag{4.22}$$

On the other hand, from $|\partial\phi|(u) \le \alpha$ almost everywhere in $[0,T]$ we deduce that $\psi^*(|\partial\phi|(u)) \le \psi^*(\alpha)$ almost everywhere in $[0,T]$, hence $\mathcal{F}(|\partial\phi|(u),u) \le \mathcal{F}(\alpha,u) = 0$. By using the fact that

$$\begin{aligned}&-\phi(u(T)) + \phi(u(t)) \le |\phi(u(T)) - \phi(u(t))| \le \int_t^T |\partial\phi|(u(r))\,|u'|(r)\,\mathrm{d}r \\ &\quad\le \int_t^T \psi(|u'|(r))\,\mathrm{d}r + \int_t^T \psi^*(|\partial\phi|(u(r)))\,\mathrm{d}r,\end{aligned} \tag{4.23}$$

we can check that

$$\begin{aligned}&\phi(u(t)) - \phi(u^0) + \int_0^t \psi(|u'|(r))\,\mathrm{d}r + \int_0^t \psi^*(|\partial\phi|(u(r)))\,\mathrm{d}r \\ &\quad= \mathcal{F}(|\partial\phi|(u),u) - \phi(u(T)) + \phi(u(t)) \\ &\quad- \int_t^T \psi(|u'|(r))\,\mathrm{d}r - \int_t^T \psi^*(|\partial\phi|(u(r)))\,\mathrm{d}r \le 0.\end{aligned}$$

where in the last inequality we have used $\mathcal{F}(|\partial\phi|(u),u) \le \mathcal{F}(\alpha,u) = 0$ and (4.23). This proves the converse inequality to (4.22). As $t \in [0,T]$ is arbitrary, (2.14) follows and u is a curve of maximal slope.

Let us now prove that (iii) ⇒ (i). As $\mathcal{H}(\alpha,u) = 0$ we again have that $\mathcal{E}(\alpha,u) = 0$, as well as $\mathcal{F}(\alpha,u) \le 0$. As $|\partial\phi|(u) \le \alpha$ almost everywhere in $[0,T]$, this implies that $\mathcal{F}(|\partial\phi|(u),u) \le \mathcal{F}(\alpha,u) \le 0$. We can hence obtain (4.22)–(4.23), proving again that u is a curve of maximal slope. □

Remark 4.2 ($\mathcal{G}_\delta$ may be negative)**.** *Before moving on, let us remark that we cannot expect Theorem 4.1 to hold for $\delta > 0$ as $\mathcal{G}_\delta$ may take negative values. Indeed, if $\delta > 0$ it might be convenient to choose some $\alpha < |\partial\phi|(u)$ making $\mathcal{F}(\alpha,u)$ negative, despite having a positive $\mathcal{E}(\alpha,u)$. Let us provide an elementary example in this direction. Set $U = [0,\infty)$, $\phi(u) = \lambda u^2/2$ for some $\lambda > 0$, and $\psi(t) = t^2/2$. In this case, we readily compute that $|\partial\phi|(u) = \lambda u$. For all $\alpha \ge 0$ and $u \ge 0$ we have that $E(\alpha,u) = ((\lambda u - \alpha)^+)^2/(2\lambda)$. One then considers*

$$[0,\infty] \ni \alpha \mapsto g_u(\alpha) := \frac{1}{2}\alpha^2 + \frac{1}{\delta}E(\alpha,u) = \frac{1}{2}\alpha^2 + \frac{1}{2\delta\lambda}((\lambda u - \alpha)^+)^2$$

which is minimized at $\alpha_u := \lambda u/(1+\lambda\delta)$ *with value* $g_u(\alpha_u) = \lambda^2 u^2/(2(1+\lambda\delta))$. *Hence, for all* $u \in AC^2([0,T];U) = H^1(0,T;U)$ *with* $u > 0$ *we compute*

$$\begin{aligned}
\mathcal{G}_\delta(\alpha_u, u) &= \mathcal{F}(\alpha_u, u) + \frac{1}{\delta}\mathcal{E}(\alpha_u, u) \\
&= \frac{\lambda}{2}u^2(T) - \frac{\lambda}{2}u^2(0) + \frac{1}{2}\int_0^T |u'(r)|^2\,\mathrm{d}r + \frac{1}{2(1+\lambda\delta)}\int_0^T |\lambda u(r)|^2\,\mathrm{d}r \\
&< \frac{\lambda}{2}u^2(T) - \frac{\lambda}{2}u^2(0) + \frac{1}{2}\int_0^T |u'(r)|^2\,\mathrm{d}r + \frac{1}{2}\int_0^T |\lambda u(r)|^2\,\mathrm{d}r \\
&= \mathcal{F}(\lambda u, u) \le \mathcal{G}_\delta(\lambda u, u).
\end{aligned}$$

By choosing $u_0 = 1$*, setting* $t \in [0,T] \mapsto \overline{u}(t) := \mathrm{e}^{-\lambda t}$*, and observing that* $\overline{u}' + \lambda\overline{u} = 0$ *on* $[0,T]$ *and* $(\lambda\overline{u}, \overline{u}) \in K(u^0)$*, one has that* $\mathcal{G}_\delta(\lambda\overline{u}, \overline{u}) = 0$*. This implies that* $\mathcal{G}_\delta(\alpha_{\overline{u}}, \overline{u}) < \mathcal{G}_\delta(\lambda\overline{u}, \overline{u}) = 0$.

5. Minimization

Theorem 4.1 links curves of maximal slope with (u-components of) minimizers of $\mathcal{G}$ and $\mathcal{H}$ on $K(u^0)$. This opens the way to providing an alternative proof of the existence of curves of maximal slope by minimization. We start by checking lower semicontinuity.

Proposition 5.1 (Lower semicontinuity of $\mathcal{E}$)**.** *Let* $\alpha_j \to \alpha$ *weakly in* $L^{p'}(0,T)$, u_j, $u \in M([0,T];U)$ *with* $u_j \to u$ *almost everywhere in* $[0,T]$*, and*

$$d(u_j(s), u_j(t)) \le \int_s^t \eta_j(r)\mathrm{d}r \quad \forall 0 \le s \le t \le T, j \in \mathbb{N}$$

for some $\eta_j \in L^1(0,T)$ *converging to* η *weakly in* $L^1(0,T)$*. Then,*

$$\mathcal{E}(\alpha, u) \le \liminf_{j\to\infty} \mathcal{E}(\alpha_j, u_j).$$

Proof. Let $v = \widehat{v}(\alpha, u) \in M([0,T];U)$ from Proposition 3.2, so that $E(\alpha, u) = \overline{E}(\alpha, u, v)$ almost everywhere in $[0,T]$. Using the triangle inequality we readily check that

$$|d(u(t), v(t)) - d(u_j(t), v(t))| \le d(u(t), u_j(t)) \to 0 \quad \text{for a.e. } t \in [0,T].$$

Let $t_0 \in [0,T]$ be such that $u_j(t_0) \to u(t_0)$. For almost all $t \in [0,T]$ we have that

$$d(u(t), u(t_0)) = \lim_{j\to\infty} d(u_j(t), u_j(t_0)) \le \lim_{j\to\infty}\int_{t\wedge t_0}^{t\vee t_0} \eta_j(r)\,\mathrm{d}r = \int_{t\wedge t_0}^{t\vee t_0} \eta(r)\,\mathrm{d}r.$$

Hence, we can bound

$$\begin{aligned}
d(u(t), u_j(t)) &\le d(u(t), u(t_0)) + d(u(t_0), u_j(t_0)) + d(u_j(t_0), u_j(t)) \\
&\le \int_{t\wedge t_0}^{t\vee t_0} \eta(r)\,\mathrm{d}r + d(u(t_0), u_j(t_0)) + \int_{t\wedge t_0}^{t\vee t_0} \eta_j(r)\,\mathrm{d}r \quad \text{for a.e. } t \in [0,T].
\end{aligned}$$

We then use that η_j converges to η to further bound the right-hand side. It is thus bounded independently of j and $t \in [0,T]$, hence the Dominated Convergence Theorem implies that $d(u_j, v) \to d(u, v)$ strongly in $L^r(0,T)$ for every $r \in (1,\infty)$.

Using the lower semicontinuity of ϕ and Fatou's Lemma as ϕ is nonnegative we can hence deduce that

$$\begin{aligned}\mathcal{E}(\alpha,u) &= \int_0^T \left(-\alpha(r)d(u(r),v(r)) + \phi(u(r)) - \phi(v(r)) - \mu_0 d^q(u(r),v(r))\right) \mathrm{d}r \\ &\le \liminf_{j\to\infty} \int_0^T \left(-\alpha_j(r)d(u_j(r),v(r)) + \phi(u_j(r)) - \phi(v(r)) - \mu_0 d^q(u_j(r),v(r))\right) \mathrm{d}r \\ &\le \liminf_{j\to\infty} \int_0^T E(\alpha_j(r),u_j(r))\,\mathrm{d}r\end{aligned}$$

by taking the maximum. This proves the claim. □

Corollary 5.2 (Lower semicontinuity of $\mathcal{G}$ and $\mathcal{H}$)**.** *Let $(\alpha_j,u_j)_j \in K(u^0)$ be such that $\alpha_j \to \alpha$ weakly in $L^{p'}(0,T)$, $u_j(t) \to u(t)$ for almost all $t \in [0,T]$ with $u(0) = u^0$, $u \in AC^p([0,T];U)$,*

$$d(u_j(s),u_j(t)) \le \int_s^t \eta_j(r)\,\mathrm{d}r \quad \forall 0 \le s \le t \le T, j \in \mathbb{N}$$

for some $\eta_j \in L^1(0,T)$ converging to η weakly in $L^1(0,T)$, and $u_j(T) \to u(T)$. Then

$$\mathcal{G}(\alpha,u) \le \liminf_{j\to\infty} \mathcal{G}(\alpha_j,u_j) \quad \text{and} \quad \mathcal{H}(\alpha,u) \le \liminf_{j\to\infty} \mathcal{H}(\alpha_j,u_j).$$

Proof. Owing to Proposition 5.1 we have that

$$0 \le \mathcal{E}(\alpha,u) \le \liminf_{j\to\infty} \mathcal{E}(\alpha_j,u_j) = 0.$$

This ensures that $|\partial\phi|(u) \in L^{p'}(0,T)$ as $|\partial\phi|(u) \le \alpha$ almost everywhere in $[0,T]$ by (3.17). Hence, $(\alpha,u) \in K(u^0)$ and we can hence pass to the lim inf in $\mathcal{F}(\alpha_j,u_j)$ as $j \to \infty$ and find

$$\begin{aligned}\mathcal{F}(\alpha,u) &= \phi(u(T)) - \phi(u^0) + \int_0^T \psi(|u'|(r))\,\mathrm{d}r + \int_0^T \psi^*(\alpha(r))\,\mathrm{d}r \\ &\le \liminf_{j\to\infty} \mathcal{F}(\alpha_j,u_j)\end{aligned} \tag{5.24}$$

whence the lower semicontinuity of $\mathcal{G}$ and $\mathcal{H}$ follows. □

We now prove that $\mathcal{G}$ and $\mathcal{H}$ can actually be minimized on $K(u^0)$. The existence of minimizers obviously requires that these functionals are proper on $K(u^0)$, see condition (5.25).

Proposition 5.3 (Minimization of $\mathcal{G}$ and $\mathcal{H}$)**.** *Assume that*

$$\inf_{K(u^0)} \mathcal{G} < \infty. \tag{5.25}$$

Then, $\mathcal{G}$ admits a minimizer in $K(u^0)$. Any minimizer (α,u) of $\mathcal{G}$ in $K(u^0)$ is such that $\alpha = |\partial\phi|(u)$ almost everywhere in $[0,T]$. If condition (5.25) holds for $\mathcal{H}$ in place of $\mathcal{G}$, then $\mathcal{H}$ admits a minimizer in $K(u^0)$.

Proof. Under condition (5.25), let $(\alpha_j, u_j)_j \in K(u^0)$ be a minimizing sequence. With no loss of generality one can assume that $\mathcal{G}(\alpha_j, u_j) \leq C$ for some $C \in (0, \infty)$. This in particular entails that

$$\phi(u_j(T)) + \int_0^T \psi(|u_j'|(r))\,\mathrm{d}r + \int_0^T \psi^*(\alpha_j(r))\,\mathrm{d}r \leq C + \phi(u^0). \tag{5.26}$$

As $\mathcal{E}(\alpha_j, u_j) = 0$, one has that $u_j \in D(|\partial\phi|)$ and $|\partial\phi|(u_j) \leq \alpha_j$ almost everywhere in $[0,T]$. This gives for all $t \in [0,T]$ that

$$\begin{aligned}\phi(u_j(t)) &\leq \phi(u^0) + |\phi(u_j(t)) - \phi(u^0)| \leq \phi(u^0) + \int_0^t |\partial\phi|(u_j(r))\,|u_j'|(r)\,\mathrm{d}r \\ &\leq \phi(u^0) + \int_0^T \alpha_j(r)\,|u_j'|(r)\,\mathrm{d}r \\ &\leq \phi(u^0) + \int_0^T \psi(|u_j'|(r))\,\mathrm{d}r + \int_0^T \psi^*(\alpha_j(r))\,\mathrm{d}r \overset{(5.26)}{\leq} C + 2\phi(u^0).\end{aligned}$$

Using (5.26) and (2.12), we can apply Lemma 2.2(a) and extract not relabeled subsequences such that $u_j(t) \to u(t)$ for all $t \in [0,T]$, as well as $\alpha_j \to \alpha$ weakly in $L^{p'}(0,T)$ and $|u_j'| \to \eta$ weakly in $L^p(0,T)$ for some $u \in AC^p([0,T];U)$ with $|u'| \leq \eta$. Proposition 5.2 ensures that $(\alpha, u) \in K(u^0)$ minimizes $\mathcal{G}$.

Given any minimizer $(\alpha, u) \in K(u^0)$ of $\mathcal{G}$, from $\mathcal{E}(\alpha, u) = 0$ one has that $u \in D(|\partial\phi|)$ and $|\partial\phi|(u) \leq \alpha$ almost everywhere in $[0,T]$. This implies that $\mathcal{G}(|\partial\phi|(u), u) \leq \mathcal{G}(\alpha, u)$, the inequality being strict if $|\partial\phi|(u) \neq \alpha$ on a nonnegligible subset of $[0,T]$ due to the strict monotonicity of ψ. In particular, $|\partial\phi|(u) = \alpha$ almost everywhere in $[0,T]$.

We now turn our attention to $\mathcal{H}$. Under condition (5.25), now written for $\mathcal{H}$ instead of $\mathcal{G}$, the bound in (5.26) still holds and we also have, using the 2^{q-1}-subadditivity of $|\cdot|^q$,

$$\begin{aligned}&\phi(u_j(T)) + \int_0^T \psi(|u_j'|(r))\,\mathrm{d}r + \int_0^T \psi^*(\alpha_j(r))\,\mathrm{d}r + \int_0^T \phi(u_j(t))\,\mathrm{d}t \\ &\quad\leq C + \int_0^T \alpha_j(t) d(u_j(t), u_*)\,\mathrm{d}t + \phi(u^0) + \mu_0 \int_0^T d^q(u_j(t), u_*)\,\mathrm{d}t \\ &\quad\leq C + \frac{1}{p'}\int_0^T |\alpha_j|^{p'}(t)\,\mathrm{d}t + \frac{1}{p}\int_0^T d^p(u_j(t), u^0)\,\mathrm{d}t + \phi(u^0) \\ &\qquad + 2^{q-1}\mu_0 \int_0^T d^q(u_j(t), u^0)\,\mathrm{d}t + 2^{q-1}\mu_0 T d^q(u_j(t), u_*),\end{aligned}$$

where $u_* \in U$ is such that $\phi(u_*) = 0$. Moreover, for every $\sigma > 1$, $t \in [0,T]$ and $j \in \mathbb{N}$, we have

$$d^\sigma(u_j(t), u^0) \leq \left(\int_0^t |u_j'|(r)\,\mathrm{d}r\right)^\sigma \leq T^{\sigma/p'} \left(\int_0^T |u_j'|^p(r)\,\mathrm{d}r\right)^{\sigma/p}$$

which is uniformly bounded due to (5.26) and (2.12). We can thus apply Lemma 2.2(b) twice, once with $t_0 = 0$ and once with $t_0 = T$, and extract not relabeled subsequences such that $u_j(t) \to u(t)$ for a.e. $t \in [0,T]$, $u(0) = u^0$, $u_j(T) \to u(T)$, as well as $\alpha_j \to \alpha$ weakly in $L^{p'}(0,T)$ and $|u_j'| \to \eta$ weakly in $L^p(0,T)$ for some

$u \in AC^p([0,T];U)$ with $|u'| \leq \eta$. Proposition 5.2 ensures that $(\alpha, u) \in K(u^0)$ minimizes $\mathcal{H}$. $\square$

Remark 5.4 (Validity of condition (5.25)). *Let us now comment on condition* (5.25). *This may of course be met, as curves of maximal slope $\overline{u}$ do exist in this setting, see* [3]. *Specifically, we have that $\mathcal{G}(|\partial\phi|(\overline{u}), \overline{u}) = \mathcal{H}(|\partial\phi|(\overline{u}), \overline{u}) = 0$ by Theorem 4.1 and* (5.25) *trivially holds true for both $\mathcal{G}$ and $\mathcal{H}$.*

Not relying on the existence of curves of maximal slope, (5.25) *also holds in the regular case $u^0 \in D(|\partial\phi|)$ by simply choosing $(\overline{\alpha}, \overline{u}) = (|\partial\phi|(u^0), u^0)$ so that $\mathcal{G}(|\partial\phi|(u^0), u^0) = \mathcal{H}(|\partial\phi|(u^0), u^0) = T\psi^*(|\partial\phi|(u^0)) < \infty$. More generally, it suffices to find any curve $u \in AC^p([0,T];U)$ with $u(0) = u^0$ and $|\partial\phi|(u) \in L^{p'}(0,T)$. This can be done by resorting to the homogeneous case $\psi(t) = |t|^p/p$. Indeed, existence of curves of maximal slope for $\psi(t) = |t|^p/p$ under assumptions* (2.6)–(2.11) *is known* [2, Rem. 3.2.5, p. 69]. *Letting $\overline{u}$ be one of such curves, one has that $(|\partial\phi|(\overline{u}), \overline{u}) \in K(u^0)$ and*

$$
\begin{aligned}
\mathcal{G}(|\partial\phi|(\overline{u}), \overline{u}) &= \phi(\overline{u}(T)) - \phi(u^0) + \int_0^T \psi(|\overline{u}'|(r))\,\mathrm{d}r + \int_0^T \psi^*(|\partial\phi|(\overline{u}(r)))\,\mathrm{d}r \\
&\leq \phi(\overline{u}(T)) - \phi(u^0) + \int_0^T C_\psi\big(|\overline{u}'|^p(r) + |\partial\phi|^{p'}(\overline{u}(r)) + 1\big)\,\mathrm{d}r < \infty, \\
\mathcal{H}(|\partial\phi|(\overline{u}), \overline{u}) &= \left(\phi(\overline{u}(T)) - \phi(u^0) + \int_0^T \psi(|\overline{u}'|(r))\,\mathrm{d}r + \int_0^T \psi^*(|\partial\phi|(\overline{u}(r)))\,\mathrm{d}r\right)^+ \\
&\leq \phi(\overline{u}(T)) + \int_0^T C_\psi\big(|\overline{u}'|^p(r) + |\partial\phi|^{p'}(\overline{u}(r)) + 1\big)\,\mathrm{d}r < \infty.
\end{aligned}
$$

Remark 5.5 (Null-minimization). *Checking $\min_{K(u^0)} \mathcal{G}(\alpha, u) = \min_{K(u^0)} \mathcal{H}(\alpha, u) = 0$ without resorting to the existence of curves of maximal slope is more demanding. One is asked to find trajectories u_ε which are arbitrarily close to curves of maximal slope, in the precise quantitative sense given by $\mathcal{G}(\alpha_\varepsilon, u_\varepsilon) < \varepsilon$ or $\mathcal{H}(\alpha_\varepsilon, u_\varepsilon) < \varepsilon$, for some $\alpha_\varepsilon \in L^{p'}(0,T)$. We are not in the position of presenting a general argument to this effect. Nonetheless, we expect that this could be achieved by considering some interpolation of discrete solutions of a suitable minimizing-movement scheme for time partitions of small diameter* [3].

We conclude by showing the possibility of finding a curve of maximal slope by a penalization approach.

Proposition 5.6 (Penalization via $\mathcal{G}_\delta$). *For all $\delta > 0$, let $(\alpha_\delta, u_\delta) \in K(u^0)$ minimize $\mathcal{G}_\delta$ and assume that*

$$
\inf_{K(u^0)} \mathcal{G}_\delta \to 0 \quad \text{as} \quad \delta \to 0. \tag{5.27}
$$

Then, we have that $\alpha_\delta \to \alpha$ weakly in $L^{p'}(0,T)$, $u_\delta \to u$ almost everywhere in $[0,T]$, and $|u_\delta'| \to \eta$ weakly in $L^p(0,T)$, up to a not relabeled subequence, where u is a curve of maximal slope.

Proof. From minimality and (5.27), for all $\varepsilon > 0$ we have that $\mathcal{G}_\delta(\alpha_\delta, u_\delta) < \varepsilon$ for δ small enough, which implies that

$$\begin{aligned} 0 &\le \mathcal{F}(\alpha_\delta, u_\delta) - \phi(u^0) \\ &\le \phi(u_\delta(T)) + \int_0^T \psi(|u_\delta'|(t))\,\mathrm{d}t + \int_0^T \psi^*(\alpha_\delta(t))\,\mathrm{d}t \le \phi(u^0) + \varepsilon. \end{aligned} \tag{5.28}$$

At the same time, we have that $\mathcal{E}(\alpha_\delta, u_\delta)/\delta = \mathcal{G}_\delta(\alpha_\delta, u_\delta) - \mathcal{F}(\alpha_\delta, u_\delta) \le \varepsilon + \phi(u^0)$ ensuring that

$$\int_0^T \phi(u_\delta(t))\,\mathrm{d}t \le \int_0^T \alpha_\delta(t) d(u_\delta(t), u_*)\,\mathrm{d}t + \delta\varepsilon + \delta\phi(u^0) + \mu_0 \int_0^T d^q(u_j(t), u_*)\,\mathrm{d}t,$$

where $u_* \in U$ is such that $\phi(u_*) = 0$. By using the bounds in (5.28) and arguing as in the proof of Proposition 5.3, we have that the above right-hand side is bounded independently of δ as $\delta \to 0$. We can thus apply Lemma 2.2(b) twice, once with $t_0 = 0$ and once with $t_0 = T$, and extract not relabeled subsequences such that $u_j(t) \to u(t)$ for a.e. $t \in [0,T]$, $u(0) = u^0$, $u_j(T) \to u(T)$, as well as $\alpha_j \to \alpha$ weakly in $L^{p'}(0,T)$ and $|u_j'| \to \eta$ weakly in $L^p(0,T)$ for some $u \in AC^p([0,T];U)$ with $|u'| \le \eta$ almost everywhere in $[0,T]$. Thus $\mathcal{F}(\alpha, u) \le \liminf_{\delta\to 0} \mathcal{F}(\alpha_\delta, u_\delta)$ and $\mathcal{E}(\alpha, u) \le \liminf_{\delta\to 0} \mathcal{E}(\alpha_\delta, u_\delta)$. This implies that

$$0 \le \mathcal{E}(\alpha, u) \le \liminf_{\delta\to 0} \delta(\varepsilon + \phi(u^0)) = 0.$$

We hence conclude that $\mathcal{G}(\alpha, u) = \mathcal{F}(\alpha, u) \le \liminf_{\delta\to 0} \mathcal{G}_\delta(\alpha_\delta, u_\delta) < \varepsilon$ and the assertion follows by taking $\varepsilon \to 0$ and using Theorem 4.1. □

Before closing this section, let us note that in fact $\mathcal{G}_\delta \to \mathcal{G}$ in the Γ-convergence sense with respect to the topology specified in Corollary 5.2. Indeed, for all $(\alpha_\delta, u_\delta)_\delta$ with $\alpha_\delta \to \alpha$ weakly in $L^{p'}(0,T)$, $u_\delta(t) \to u(t)$ for almost all $t \in [0,T]$ with $u(0) = u^0$, $d(u_\delta(t), u_\delta(s)) \le \int_s^t \eta_\delta(r)\,\mathrm{d}r$ for all $0 \le s < t \le T$ with $\eta_\delta \to \eta$ weakly in $L^p(0,T)$, and $u_\delta(T) \to u(T)$, we have that $\mathcal{F}(\alpha, u) \le \liminf_{\delta\to 0} \mathcal{F}(\alpha_\delta, u_\delta)$. By assuming with no loss of generality that $\sup_j \mathcal{G}_\delta(\alpha_\delta, u_\delta) < \infty$, we readily get that $\mathcal{E}(\alpha, u) = 0$. Hence, $\mathcal{G}(\alpha, u) \le \liminf_{\delta\to 0} \mathcal{G}_\delta(\alpha_\delta, u_\delta)$. On the other hand, $\mathcal{G}_\delta$ converges to $\mathcal{G}$ pointwise, which directly entails the existence of recovery sequences.

6. An extension without assumption (2.10)

This last section is devoted to an extension of the variational characterization of Proposition 3.3, and hence of the Brezis–Ekeland–Nayroles principle of Theorem 4.1, not relying on the validity of assumption (2.10) concerning the non-asymptotic formulation of the local slope.

For the sake of notational convenience, within this section we still indicate by $\mu_0 \in [0,\infty)$ the smallest μ_0 such that $|\partial\phi| = |\partial\phi|_{\mu_0}$ if such a value exists, and we set $\mu_0 = \infty$ otherwise. Let us recall that $|\partial\phi|$ is a strong upper gradient for ϕ if $\mu_0 < \infty$. In case $\mu_0 = \infty$, this is to be additionally assumed as it is needed in the proof of the Brezis–Ekeland–Nayroles principle. Henceforth, in addition to (2.6)–(2.8) and (2.11)–(2.12), we ask that

$$|\partial\phi| \text{ is a strong upper gradient for } \phi. \tag{6.29}$$

Note again that the latter is weaker than (2.10), which is not assumed here.

The first step to extend the results of the previous sections is to obtain an alternative variational representation of the local slope $|\partial\phi|$ covering the case $\mu_0 = \infty$, as well. Let us start by proving the following properties of the μ-slope $|\partial\phi|_\mu$ defined in (2.9).

Proposition 6.1 (Properties of the μ-slope)**.** *For all $u \in U$ one has that*

$$\text{the map } \mu \in [0,\infty) \mapsto |\partial\phi|_\mu(u) \text{ is nonincreasing,} \tag{6.30}$$

$$|\partial\phi|(u) \leq |\partial\phi|_\mu(u) \leq |\partial\phi|_0(u) = \mathfrak{l}_\phi(u), \tag{6.31}$$

$$\lim_{\mu\to\infty} |\partial\phi|_\mu(u) = |\partial\phi|(u), \tag{6.32}$$

$$\lim_{\mu\to 0} |\partial\phi|_\mu(u) = |\partial\phi|_0(u) = \mathfrak{l}_\phi(u). \tag{6.33}$$

In case $\mu_0 < \infty$ we have that $|\partial\phi| = |\partial\phi|_\mu$ for all $\mu \geq \mu_0$.

Proof. The monotonicity (6.30) of the map $\mu \mapsto |\partial\phi|_\mu(u)$ readily follows from the definition of the μ-slope.

For all $v \in U$ with $v \neq u$ one has that

$$\frac{(\phi(u)-\phi(v))^+}{d(u,v)} - \mu d^{q-1}(u,v) \leq |\partial\phi|_\mu(u) \leq |\partial\phi|_0(u) = \mathfrak{l}_\phi(u).$$

Taking the lim sup as $v \to u$ proves (6.31).

Fix $u \in D(|\partial\phi|)$. Since $\mu : (0,\infty) \mapsto |\partial\phi|_\mu(u) \in [0,\infty]$ is nonincreasing, the limits in (6.32)–(6.33) exist. For all $\varepsilon > 0$ we can find $r > 0$ such that

$$\frac{(\phi(u)-\phi(v))^+}{d(u,v)} \leq |\partial\phi|(u) + \varepsilon \quad \forall v \in B_r := \{w \in U : 0 < d(u,w) < r\}.$$

Note that r is independent of μ. We hence have that

$$\frac{(\phi(u)-\phi(v))^+}{d(u,v)} - \mu d^{q-1}(u,v) \leq |\partial\phi|(u) + \varepsilon \quad \forall v \in B_r.$$

On the other hand

$$\begin{aligned}\frac{(\phi(u)-\phi(v))^+}{d(u,v)} - \mu d^{q-1}(u,v) &\leq \frac{(\phi(u)-\phi(v))^+}{r} - \mu r^{q-1} \\ &\leq \frac{\phi(u)}{r} - \mu r^{q-1} \quad \forall v \in U \setminus B_r.\end{aligned}$$

As $\lim_{\mu\to\infty}(-\mu r^{q-1}) = -\infty$, for μ large enough, the supremum of the above left-hand side is taken in B_r. This in particular proves that $\lim_{\mu\to\infty} |\partial\phi|_\mu(u) \leq |\partial\phi|(u) + \varepsilon$ and (6.32) follows as ε is arbitrary.

Let $u \in U$ and $(v_k)_k \subset U$ with $v_k \neq u$ be such that

$$\mathfrak{l}_\phi(u) = \lim_{k\to\infty} \frac{(\phi(u)-\phi(v_k))^+}{d(u,v_k)}.$$

Thus, for every $\mu \in (0,\infty)$, we have

$$\begin{aligned}\frac{(\phi(u)-\phi(v_k))^+}{d(u,v_k)} &= \frac{(\phi(u)-\phi(v_k))^+}{d(u,v_k)} - \mu d^{q-1}(u,v_k) + \mu d^{q-1}(u,v_k) \\ &\leq |\partial\phi|_\mu(u) + \mu d^{q-1}(u,v_k).\end{aligned}$$

Passing first to the limit as $\mu \to 0$ and then as $k \to \infty$, we get $\mathfrak{l}_\phi(u) \le \lim_{\mu\to 0} |\partial\phi|_\mu(u)$.

In case $\mu_0 < \infty$ one has that $|\partial\phi|_{\mu_0}(u) = |\partial\phi|(u)$. As $\mu \mapsto |\partial\phi|_\mu(u)$ is nonincreasing by (6.30), from (6.31) we get $|\partial\phi|(u) \le |\partial\phi|_\mu(u) \le |\partial\phi|_{\mu_0}(u) = |\partial\phi|(u)$ for all $\mu \ge \mu_0$. □

We now introduce a modification $\tilde{E}$ of the functional E, for which the analogue of (3.17) holds. Arguing as in Lemma 3.1, we define the functionals $\overline{E}_\mu : [0,\infty) \times U \times D(\phi) \to (-\infty,\infty]$ and $E_\mu : [0,\infty) \times U \to [0,\infty]$ as

$$\begin{aligned} \overline{E}_\mu(\alpha,u,v) &= -\alpha d(u,v) + \phi(u) - \phi(v) - \mu d^q(u,v), \\ E_\mu(\alpha,u) &= \max_{v\in D(\phi)} \overline{E}_\mu(\alpha,u,v). \end{aligned}$$

Note that, if $u \in D(\phi)$ both $\alpha \in [0,\infty) \mapsto E_\mu(\alpha,u)$ and $\mu \in [0,\infty) \mapsto E_\mu(\alpha,u)$ are nonincreasing, convex and finite, hence continuous on $(0,\infty)$. Both functions are instead identically equal to ∞ if $u \notin D(\phi)$. Proposition 3.3, now applied to E_μ, ensures that

$$|\partial\phi|_\mu(u) \le \alpha \;\Leftrightarrow\; E_\mu(\alpha,u) = 0. \tag{6.34}$$

Set $h : [0,\infty) \to [0,\infty]$ to be lower semicontinuous and nondecreasing, with $h(0) = 0$ and $h(0+) > 0$, a possible choice being the indicator function of 0, namely $h(0) = 0$ and $h(r) = \infty$ for all $r > 0$. Using such h we define $\tilde{E} : [0,\infty) \times U \to [0,\infty]$ as

$$\tilde{E}(\alpha,u) = \sup_{\varepsilon\in\mathbb{Q}_+} \inf_{\mu_0 \ge \mu \in \mathbb{Q}_+} h(E_\mu(\alpha+\varepsilon,u)) \tag{6.35}$$

where $\mathbb{Q}_+ := \mathbb{Q} \cap (0,\infty)$. We have the following variational characterization of the local slope.

Proposition 6.2 (Characterization 2)**.** *For every $(\alpha,u) \in [0,\infty) \times U$, we have that*

$$|\partial\phi|(u) \le \alpha \;\Leftrightarrow\; \tilde{E}(\alpha,u) = 0. \tag{6.36}$$

Proof. The case $\mu_0 < \infty$ can be checked by easily adapting the proof of Proposition 3.3.

Set $\mu_0 = \infty$. Assume $|\partial\phi|(u) \le \alpha$ and fix any $\varepsilon \in \mathbb{Q}_+$. As $|\partial\phi|_\mu(u) \to |\partial\phi|(u)$ from (6.32), there exists $\mu \in \mathbb{Q}_+$ such that $|\partial\phi|_\mu(u) \le \alpha + \varepsilon$. Then, (6.34) implies that $E_\mu(\alpha+\varepsilon,u) = 0$, so that $\inf_{\mu\in\mathbb{Q}_+} h(E_\mu(\alpha+\varepsilon,u)) = 0$ and $\tilde{E}(\alpha,u) = 0$ follows as ε is arbitrary.

Assume now that $\tilde{E}(\alpha,u) = 0$. Then, for all $\varepsilon \in \mathbb{Q}_+$ there exists $\mu \in \mathbb{Q}_+$ such that $E_\mu(\alpha+\varepsilon,u) = 0$. This implies that $|\partial\phi|_\mu(u) \le \alpha + \varepsilon$ by (6.34). As $|\partial\phi|(u) \le |\partial\phi|_\mu(u)$ by (6.31), we deduce that $|\partial\phi|(u) \le \alpha$ by taking $\varepsilon \to 0$. □

Let us now introduce the time-continuous functionals, based on the functional $\tilde{E}$. Note that for all $\mu, \varepsilon > 0$ the map $(\alpha,u) \mapsto E_\mu(\alpha+\varepsilon,u)$ is Borel, which implies that $\tilde{E}$ is Borel, as well. This allows to define the functionals $\tilde{\mathcal{E}}$, $\tilde{\mathcal{G}}$, $\tilde{\mathcal{H}}$:

$L^p(0,T) \times AC^p([0,T];U) \to [0,\infty]$ as

$$\tilde{\mathcal{E}}(\alpha,u) = \int_0^T \tilde{E}(\alpha(t),u(t))\,\mathrm{d}t,$$
$$\tilde{\mathcal{G}}(\alpha,u) = \begin{cases} \mathcal{F}(\alpha,u) & \text{if } \tilde{\mathcal{E}}(\alpha,u)=0,\\ \infty & \text{otherwise,} \end{cases}$$
$$\tilde{\mathcal{H}}(\alpha,u) = (\mathcal{F}(\alpha,u))^+ + \tilde{\mathcal{E}}(\alpha,u).$$

Using these functionals, we can again state a Brezis–Ekeland–Nayroles principle as in Theorem 4.1.

Proposition 6.3 (Null-minimization 2)**.** *The following are equivalent:*

(i) *u is a curve of maximal slope,*
(ii) $\tilde{\mathcal{G}}(\alpha,u) = \min_{K(u^0)} \tilde{\mathcal{G}} = 0$ *for some* $\alpha \in L^{p'}(0,T)$*,*
(iii) $\tilde{\mathcal{H}}(\alpha,u) = \min_{K(u^0)} \tilde{\mathcal{H}} = 0$ *for some* $\alpha \in L^{p'}(0,T)$*.*

We omit the proof of Proposition 6.3, as it coincides with that of Theorem 4.1, up to using the characterization (6.36) instead of (3.17).

Before closing this section, let us point out that, in case $\mu_0 < \infty$, the above functionals play the same role as the ones of the previous sections. More precisely, if $\mu_0 < \infty$, one has that the functionals $\tilde{E}$ and E share the same sublevel 0, which in particular implies that $\tilde{\mathcal{G}} = \mathcal{G}$. This is an immediate consequence of Proposition 6.2 and of the characterization (3.17).

Proposition 6.4 (Case $\mu_0 < \infty$)**.** *If $\mu_0 < \infty$ one has that*

$$\tilde{E}(\alpha,u) = 0 \;\Leftrightarrow\; E(\alpha,u) = 0.$$

Acknowledgement. This research is partially supported by the Austrian Science Fund (FWF) project 10.55776/PAT1408325. For open access purposes, the authors have applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

Competing Interests. The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

Tool and computational resource disclosure. The authors used a LLM as a language support tool. Moreover, the proof of the limit (6.32) has been inspired by a LLM comment.

References

[1] L. Ambrosio. Minimizing movements. *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5)*, 19 (1995), 191–246.
[2] L. Ambrosio, N. Gigli, G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*, second ed., Birkhäuser Verlag, Basel, 2008.
[3] P.-C. Aubin–Frankowski, G. E. Sodini, U. Stefanelli. The Brezis–Ekeland–Nayroles principle in metric spaces: minimizing-movement schemes. In preparation (2026).
[4] V. Barbu. Optimal control approach to nonlinear diffusion equations driven by Wiener noise. *J. Optim. Theory Appl.* 153 (2012), no. 1, 1–26.

[5] V. Barbu, K. Kunisch. Identification of nonlinear parabolic equations. *Control Theory Adv. Tech.* 10 (1995), no. 4, part 5, 1959–1980.
[6] V. Barbu, M. Röckner. Variational solutions to nonlinear stochastic differential equations in Hilbert spaces. *Stoch. Partial Differ. Equ. Anal. Comput.* 6 (2018), no. 3, 500–524.
[7] M. A. Biot. Variational principles in irreversible thermodynamics with application to viscoelasticity. *Phys. Rev. (2)*, 97 (1955), 1463–1469.
[8] V. I. Bogachev, *Measure theory. Vol. I, II.* Springer, Berlin, 2007.
[9] S. Boroushaki, N. Ghoussoub. A self-dual variational approach to stochastic partial differential equations. *J. Funct. Anal.* 276 (2019), no. 4, 1201–1243.
[10] H. Brezis, I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas dépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282 (1976), no. 20, Ai, A1197–A1198.
[11] H. Brezis, I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282 (1976), no. 17, Aii, A971–A974.
[12] M. Buliga, G. de Saxcé. A symplectic Brezis-Ekeland-Nayroles principle. *Math. Mech. Solids*, 22 (2017), no. 6, 1288–1302.
[13] X. Cao, A. Oueslati, N. Shirafkan, F. Bamer, B. Markert, G. de Saxcé. A non-incremental numerical method for dynamic elastoplastic problems by the symplectic Brezis–Ekeland–Nayroles principle. *Comput. Methods Appl. Mech. Engrg.* 384 (2021), 113908, 25 pp.
[14] L. Carini, M. Jensen, R. Nürnberg. Deep learning for gradient flows using the Brezis-Ekeland principle. *Arch. Math. (Brno)*, 59 (2023), no. 3, 249–261.
[15] J. Cheeger. Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.* 9 (1999), no. 4, 428–517.
[16] E. De Giorgi, A. Marino, M. Tosques. Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 68 (1980), no. 3, 180–187.
[17] P. Dondl, T. Frenzel, A. Mielke. A gradient system with a wiggly energy and relaxed EDP-convergence. *ESAIM Control Optim. Calc. Var.* 25 (2019), Paper No. 68, 45 pp.
[18] T. Frenzel, M. Liero. Effective diffusion in thin structures via generalized gradient systems and EDP-convergence. *Discrete Contin. Dyn. Syst. Ser. S*, 14 (2021), no. 1, 395–425.
[19] N. Ghoussoub. *Self-dual partial differential systems and their variational principles.* Springer Monographs in Mathematics. Springer, New York, 2009.
[20] N. Ghoussoub, R. J. McCann. A least action principle for steepest descent in a non-convex landscape. In *Partial differential equations and inverse problems*, vol. 362 of *Contemp. Math.*, Amer. Math. Soc., Providence, RI, 2004, 177–187.
[21] N. Ghoussoub, A. Moameni. Selfdual variational principles for periodic solutions of Hamiltonian and other dynamical systems. *Comm. Partial Differential Equations*, 32 (2007), no. 4-6, 771–795.
[22] N. Ghoussoub, A. Moameni. Hamiltonian systems of PDEs with selfdual boundary conditions. *Calc. Var. Partial Differential Equations*, 36 (2009), no. 1, 85–118.
[23] N. Ghoussoub, L. Tzou. A variational principle for gradient flows. *Math. Ann.* 330 (2004), no. 3, 519–549.
[24] N. Ghoussoub, L. Tzou. Iterations of anti-selfdual Lagrangians and applications to Hamiltonian systems and multiparameter gradient flows. *Calc. Var. Partial Differential Equations*, 26 (2006), no. 4, 511–534.
[25] M. E. Gurtin. Variational principles in the linear theory of viscoelasticity. *Arch. Rational Mech. Anal.* 13 (1963), 179–191.
[26] M. E. Gurtin. Variational principles for linear elastodynamics. *Arch. Rational Mech. Anal.* 16 (1964), 34–50.
[27] I. Hlaváček. Variational principles for parabolic equations. *Apl. Mat.* 14 (1969), 278–297.
[28] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* 29 (1998), no. 1, 1–17.
[29] B. Lemaire. An asymptotical variational principle associated with the steepest descent method for a convex function. *J. Convex Anal.* 3 (1996), no. 1, 63–70.
[30] M. Mabrouk. Un principe variationnel pour une équation non linéaire du second ordre en temps. *C. R. Acad. Sci. Paris Sér. I Math.* 332 (2001), no. 4, 381–386.

[31] M. Mabrouk. A variational principle for a nonlinear differential equation of second order. *Adv. in Appl. Math.* 31 (2003), no. 2, 388-419.
[32] G. Marinoschi. A duality approach to nonlinear diffusion equations. *Set-Valued Var. Anal.* 22 (2014), no. 4, 783–807.
[33] A. Mielke, A. Stephan. Coarse-graining via EDP-convergence for linear fast-slow reaction systems. *Math. Models Meth. Appl. Sci.* 30 (2020), no. 9, 1765–1807.
[34] A. Mielke, A. Montefusco, M. Peletier. Exploring families of energy-dissipation landscapes via tilting: three types of EDP convergence. *Contin. Mech. Thermodyn.* 33 (2021), no. 3, 611–637.
[35] M. Muratori, G. Savaré. Gradient flows and evolution variational inequalities in metric spaces. I: Structural properties. *J. Funct. Anal.* 278 (2020), no. 4, 108347, 67 pp.
[36] B. Nayroles. Deux théorèmes de minimum pour certains systèmes dissipatifs. *C. R. Acad. Sci. Paris Sér. A-B*, 282 (1976), no. 17, Aiv, A1035–A1038.
[37] B. Nayroles. Un théorème de minimum pour certains systèmes dissipatifs. Variante hilbertienne. *Travaux Sém. Anal. Convexe*, 6 (1976), Exp. no. 2, 22 pp.
[38] S. Ohta, W. Zhao. Gradient flows in asymmetric metric spaces and applications. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 26 (2025), no. 2, 919–991.
[39] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26 (2001), no. 1-2, 101–174.
[40] A. Poliakovsky. On a variational approach to the method of vanishing viscosity for conservation laws. *Adv. Math. Sci. Appl.* 18 (2008), no. 2, 429–451.
[41] A. Poliakovsky. Variational resolution for some general classes of nonlinear evolutions. Part I. *Asymptot. Anal.* 85 (2013), no. 1–2, 29–74.
[42] L. Portinale, U. Stefanelli. Penalization via global functionals of optimal-control problems for dissipative evolution. *Adv. Math. Sci. Appl.* 28 (2019), no. 2, 425–447.
[43] R. Rossi, G. Savaré. Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2 (2003), no. 2, 395–431.
[44] R. Rossi, G. Savaré, A. Segatti, U. Stefanelli. A variational principle for gradient flows in metric spaces. *C. R. Math. Acad. Sci. Paris*, 349 (2011), no. 23–24, 1225–1228.
[45] R. Rossi, G. Savaré, A. Segatti, U. Stefanelli. Weighted energy-dissipation principle for gradient flows in metric spaces. *J. Math. Pures Appl. (9)*, 127 (2019), 1–66.
[46] R. Rossi, A. Segatti, U. Stefanelli. Global attractors for gradient flows in metric spaces. *J. Math. Pures Appl.* 95 (2011), 204–244.
[47] E. Sandier, S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.* 57 (2004), no. 12, 1627–1672.
[48] S. Serfaty. Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discrete Contin. Dyn. Syst.* 31 (2011), no. 4, 1427–1451.
[49] S. Shimoyama. Transformation of p-gradient flows to p'-gradient flows in metric spaces. *Anal. Geom. Metr. Spaces*, 13 (2025), no. 1, Paper No. 20250032, 24 pp.
[50] U. Stefanelli. The Brezis–Ekeland principle for doubly nonlinear equations. *SIAM J. Control Optim.* 47 (2008), no. 3, 1615–1642.
[51] U. Stefanelli. A variational principle for hardening elastoplasticity. *SIAM J. Math. Anal.* 40 (2008), no. 2, 623–652.
[52] U. Stefanelli. The discrete Brezis–Ekeland principle. *J. Convex Anal.* 16 (2009), no. 1, 71–87.
[53] U. Stefanelli. The weighted inertia-energy-dissipation principle. *Math. Models Methods Appl. Sci.* 35 (2025), no. 2, 223–282.
[54] A. Visintin. A new approach to evolution. *C. R. Acad. Sci. Paris Sér. I Math.* 332 (2001), no. 3, 233–238.
[55] A. Visintin. Structural stability of doubly-nonlinear flows. *Boll. Unione Mat. Ital. (9)*, 4 (2011), no. 3, 363–391.
[56] A. Visintin. Variational formulation and structural stability of monotone equations. *Calc. Var. Partial Differential Equations*, 47 (2013), no. 1-2, 273–317.
[57] A. Visintin. Structural stability of rate-independent nonpotential flows. *Discrete Contin. Dyn. Syst. Ser. S*, 6 (2013), no. 1, 257–275.
[58] A. Visintin. On Fitzpatrick's theory and stability of flows. *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* 27 (2016), no. 2, 151–180.
[59] A. Visintin. Structural compactness and stability of doubly nonlinear flows. *Pure Appl. Funct. Anal.* 6 (2021), no. 2, 417–432.

[60] A. Visintin. Γ-compactness and Γ-stability of maximal monotone flows. *J. Math. Anal. Appl.* 506 (2022), no. 1, Paper No. 125602, 29 pp.

(P.-C. Aubin–Frankowski) CERMICS, CNRS, ENPC, Institut Polytechnique de Paris, 6 and 8 avenue Blaise-Pascal, Cité Descartes, Champs-sur-Marne, F-77455, Marne-la-Vallée cedex 2, France.

Email address: pierre-cyril.aubin@enpc.fr

URL: https://pcaubin.github.io/

(G. E. Sodini) Institute of Analysis and Scientific Computing, TU Wien, Wiedner Hauptstrasse 8-10, A-1040, Vienna, Austria.

Email address: giacomo.sodini@tuwien.ac.at

URL: https://giacomosodini.github.io/

(U. Stefanelli) Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria.

Email address: ulisse.stefanelli@univie.ac.at

URL: http://www.mat.univie.ac.at/∼stefanelli