On De Giorgi's lemma for variational interpolants in metric and Banach spaces*

Alexander Mielke† and Riccarda Rossi‡

31 August 2024. Revised 11 March 2026

Abstract

Variational interpolants are an indispensable tool for the construction of gradient-flow solutions via the Minimizing Movement Scheme. The De Giorgi lemma provides the associated discrete energy-dissipation inequality. It was originally developed for metric gradient systems. Drawing from this theory we study the case of generalized gradient systems in Banach spaces, where a refined theory allows us to extend the validity of the discrete energy-dissipation inequality and to establish it as an equality. For the latter we have to impose the condition of radial differentiability of the dissipation potential. Several examples are discussed to show how sharp the results are.

Keywords: Generalized gradient systems, minimizing movement scheme, variational interpolants, discrete energy-dissipation inequality, radial differentiability.

Dedicated to Giuseppe Savaré on the occasion of his sixtieth birthday, with gratitude and friendship

1 Introduction

The Minimizing Movement Scheme (MMS) was introduced by De Giorgi in [De93] for constructing solutions for gradient flows in abstract spaces. Since then, the MMS has developed into a versatile tool for analyzing gradient systems in Hilbert spaces, Banach space, and metric spaces. In this paper, we address the specific tool called "variational interpolant", also called "De Giorgi interpolant" that was first introduced in [Amb95, Lem. 2.5] and further developed in [AGS05]. A generalization to the Banach spaces was done in [MRS13, Lem. 6.1]. Variational interpolants generalize the idea of piecewise affine interpolants in linear spaces, or geodesic interpolants in geodesic spaces, in such a way that they turn out to be applicable in more general situations, namely in general metric spaces. However, even in the cases of geodesic spaces, including Banach and Hilbert

*Research of AM partially supported by DFG via the Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689) subproject "DistFell". Research of RR partially supported by GNAMPA (Indam) and MIUR via the MIUR-PRIN Grant 2020F3NCPX "Mathematics for industry 4.0 (Math4I4)".

†WIAS Berlin, Anton-Wilhelm-Amo-Str. 39, 10117 Berlin, Germany, alexander.mielke@wias-berlin.de

‡DIMI, Università degli studi di Brescia, Italy, riccarda.rossi@unibs.it

spaces, they are useful if the energy functional is not geodesically semi-convex. In general, variational interpolants are no longer continuous in time and hence, the desired discrete energy-dissipation estimate is more difficult to obtain. It is exactly this estimate, which is established in the so-called "*De Giorgi lemma*". The purpose of this paper is twofold: (i) we generalize the validity of the lemma in the Banach setup and (ii) we discuss the question why and when the discrete energy-dissipation estimate is an *equality*.

To be more precise, we now introduce our approach in more detail by comparing the theory in metric spaces $(M, \mathcal{D})$ and in Banach spaces $(X; \|\cdot\|)$ in parallel. It has to be emphasized that the Banach setup we are going to address is not contained in the metric one. Indeed, while in the latter we shall essentially confine the discussion to the case in which the dissipation mechanism is encoded by the interplay of the metric $\mathcal{D}$ with a convex *scalar* function $\psi : [0,\infty[\to [0,\infty[$, in the former we will address *general* dissipation potentials $\mathcal{R} : X \to [0,\infty]$, for instance $\mathcal{R}(v) = \int_\Omega (\frac{1}{2}v^2+\frac{1}{4}v^4)\,\mathrm{d}x$ for $X = \mathrm{L}^4(\Omega)$.

Following [RMS08] (see also [Mie23]) we consider a generalized metric gradient system $(M, \mathcal{E}, \mathcal{D}, \psi)$, subsequently abbreviated by gMGS; see Definition 2.1 for the precise definition. For a given initial value $u^\circ \in M$ with $\mathcal{E}(u^\circ) < \infty$, it is the aim to construct a curve $u : [0,\infty[\to M$ of maximal slope emanating from u°, i.e. u must satisfy for all $t > 0$

$$\mathcal{E}(u(t)) + \int_0^t \Big(\psi\big(|u'|(s)\big) + \psi^*\big(|\partial\mathcal{E}|(u(s))\big)\Big)\,\mathrm{d}s = \mathcal{E}(u(0)) \quad \text{and} \quad u(0) = u^\circ, \tag{1.1a}$$

where $|u'| \geq 0$ denotes the metric speed of u and $|\partial\mathcal{E}|(u) \geq 0$ denotes the metric slope, see [AGS05], and ψ^* is the convex conjugate of ψ. The case of general dissipation functions $\psi : [0,\infty[\to [0,\infty[$ (lower semicontinuous, convex, $\psi(0) = 0$), and superlinear) was introduced in [RMS08], the choice $\psi(r) = \frac{1}{2}r^2$ gives the classical notion of curve of maximal slope of [Amb95], while $\phi(r) = \frac{1}{p}r^p$ leads to p-curves of maximal slopes as in [AGS05].

For a generalized Banach-space gradient system $(X, \mathcal{E}, \mathcal{R})$, subsequently abbreviated by gBGS and precisely defined in Section 3.1, the aim is to find *energy-dissipation balance (EDB) solutions* $u : [0,\infty[\to X$, which are defined via the following identity

$$\begin{aligned} \mathcal{E}(u(t)) + \int_0^t \Big(\mathcal{R}\big(u'(s)\big) + \mathcal{R}^*(-\xi(s))\Big)\,\mathrm{d}s = \mathcal{E}(u(0)) \text{ for all } t > 0, \\ \text{and } \xi(s) \in \mathrm{D}\mathcal{E}(u(s)) \text{ for a.a. } s \geq 0, \end{aligned} \tag{1.1b}$$

where u' is the distributional derivative of $u \in \mathrm{AC}([0,T]; X)$, and the convex conjugate $\mathcal{R}^* : X^* \to [0,\infty)$ is evaluated along a selection $\xi :]0,\infty[\to X^*$ in the multivalued Fréchet subdifferential $\mathrm{D}\mathcal{E}(u) \subset X^*$ of $\mathcal{E}$, see Section 3.1.

With an initial value $u^\circ \in M$ and a time step $\tau > 0$ the metric and the Banach-space MMS are defined via $u^0_\tau = u^\circ$ and

$$u^k_\tau \quad \text{minimizes} \quad M \ni u \mapsto \tau\psi\big(\tfrac{1}{\tau}\mathcal{D}(u, u^{k-1}_\tau)\big) + \mathcal{E}(u) \qquad \text{for all } k \in \mathbb{N}; \tag{1.2a}$$

$$u^k_\tau \quad \text{minimizes} \quad X \ni u \mapsto \tau\mathcal{R}\big(\tfrac{1}{\tau}(u-u^{k-1}_\tau)\big) + \mathcal{E}(u) \qquad \text{for all } k \in \mathbb{N}. \tag{1.2b}$$

Variational interpolants $\widetilde{u}^\tau$ are defined for all $t \in [0,\infty[$, satisfy $\widetilde{u}^\tau(k\tau) = u^k_\tau$, and are determined by a variational condition: for all $k \in \mathbb{N}_0$ and $\sigma \in]0,\tau[$, we ask for

$$\widetilde{u}^\tau(k\tau+\sigma) \quad \text{minimizes} \quad M \ni u \mapsto \sigma\psi\big(\tfrac{1}{\sigma}\mathcal{D}(u, u^{k-1}_\tau)\big) + \mathcal{E}(u); \tag{1.3a}$$

$$\widetilde{u}^\tau(k\tau+\sigma) \quad \text{minimizes} \quad X \ni u \mapsto \sigma\mathcal{R}\big(\tfrac{1}{\sigma}(u-u^{k-1}_\tau)\big) + \mathcal{E}(u). \tag{1.3b}$$

In general, one cannot hope to choose the variational interpolant $t \mapsto \widetilde{u}^\tau(t)$ as a continuous function. However, by classical selection theorems for measurable multivalued mappings, it is possible to choose a measurable selection, see Section 3.1.

The De Giorgi lemma, which was first published in [Amb95, Lem. 2.5], provides a discrete counterpart to the energy-dissipation balances in (1.1a) and (1.1b), namely for all $\sigma \in \left]0,\tau\right]$ we have the so-called *De Giorgi estimates*

$$\mathcal{E}\big(\widetilde{u}^\tau(k\tau{+}\sigma)\big) + \sigma\,\psi\big(\tfrac{1}{\sigma}\mathcal{D}(\widetilde{u}^\tau(k\tau{+}\sigma),u_\tau^k)\big) + \int_0^\sigma \psi^*\big(|\partial\mathcal{E}|(\widetilde{u}^\tau(k\tau{+}\rho)\big)\,\mathrm{d}\rho \le \mathcal{E}(u_\tau^k); \tag{1.4a}$$

$$\mathcal{E}\big(\widetilde{u}^\tau(k\tau{+}\sigma)\big) + \sigma\,\mathcal{R}\big(\tfrac{1}{\sigma}(\widetilde{u}^\tau(k\tau{+}\sigma){-}u_\tau^k)\big) + \int_0^\sigma \mathcal{R}^*\big({-}\xi(k\tau{+}\rho)\big)\,\mathrm{d}\rho \le \mathcal{E}(u_\tau^k) \tag{1.4b}$$

$$\text{for some } \xi(t) \in \mathrm{D}\mathcal{E}(\widetilde{u}^\tau(t)) \text{ for a.a. } t>0.$$

For establishing (1.1) via a suitable limit passage, it would be enough to have (1.4) for $\sigma = \tau$, and then adding the results over all subintervals, but we will see that it is very instructive to keep $\sigma \in \left]0,\tau\right]$ on the left-hand side as an independent variable.

For general differentiable dissipation functions ψ, the De Giorgi estimate (1.4a) was first established in [RMS08, Lem. 4.5], while $\psi(r) = \frac{1}{p}r^p$ is treated in [AGS05]. The Banach-space case (1.4b) appears first in [MRS13, Lem. 6.1], but the result therein relies on the *condition of radial differentiability* of $\mathcal{R}$, namely

$$\forall\, v \in X: \quad \text{the function } \left]0,\infty\right[\ni \lambda \mapsto \mathcal{R}(\lambda v) \text{ is differentiable.} \tag{1.5}$$

This condition is equivalent to the fact that for all $\xi_1, \xi_2 \in \partial\mathcal{R}(v)$ (with $\partial\mathcal{R}: X \rightrightarrows X^*$ the convex subdifferential of $\mathcal{R}$), there holds $\mathcal{R}^*(\xi_1) = \mathcal{R}^*(\xi_2)$, cf. Proposition 3.11.

So far, our general overview and introduction shows a complete analogy between the metric case and the Banach-space setting. Even the condition of radial differentiability of $\mathcal{R}$ corresponds to the condition of differentiability of ψ. However, the methods for establishing the so-called *De Giorgi estimates* (1.4a) and (1.4b) involve quite different techniques. In particular, for gBGS we can exploit the linear structure of X and thus obtain an Euler-Lagrange equation for the minimizers $\widetilde{u}_\sigma := \widetilde{u}^\tau(k\tau{+}\sigma)$ (keep k fixed, w.l.o.g. $k=0$), namely

$$0 \in \partial\mathcal{R}\big(\tfrac{1}{\sigma}(\widetilde{u}_\sigma{-}u_\tau^k)\big) + \widetilde{\xi}_\sigma \quad\text{and}\quad \widetilde{\xi}_\sigma \in \mathrm{D}\mathcal{E}(\widetilde{u}_\sigma). \tag{1.6}$$

Indeed, to see a first nontrivial fact, we may assume that $\sigma \mapsto \widetilde{u}_\sigma$ and $\sigma \mapsto \mathcal{E}(\widetilde{u}_\sigma)$ are absolutely continuous and such that the chain rule relation $\frac{\mathrm{d}}{\mathrm{d}\sigma}\mathcal{E}(\widetilde{u}_\sigma) = \langle \widetilde{\xi}_\sigma, \frac{\mathrm{d}}{\mathrm{d}\sigma}\widetilde{u}_\sigma\rangle$ holds. Then, the Euler-Lagrange equation (1.6) gives the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\mathcal{R}\big(\tfrac{1}{\sigma}(\widetilde{u}_\sigma{-}u_\tau^k)\big) = \big\langle {-}\widetilde{\xi}_\sigma, \frac{1}{\sigma}\frac{\mathrm{d}}{\mathrm{d}\sigma}\widetilde{u}_\sigma - \frac{1}{\sigma^2}(\widetilde{u}_\sigma{-}u_\tau^k)\big\rangle.$$

Thus, differentiating the right-hand side of (1.4b) with respect to σ gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\sigma}\,\mathrm{RHS}(1.4\mathrm{b}) &= \big\langle \widetilde{\xi}_\sigma, \frac{\mathrm{d}}{\mathrm{d}\sigma}\widetilde{u}_\sigma\big\rangle + \mathcal{R}\big(\tfrac{1}{\sigma}(\widetilde{u}_\sigma{-}u_\tau^k)\big) \\ &\quad + \big\langle {-}\widetilde{\xi}_\sigma, \frac{\mathrm{d}}{\mathrm{d}\sigma}\widetilde{u}_\sigma - \tfrac{1}{\sigma}(\widetilde{u}_\sigma{-}u_\tau^k)\big\rangle + \mathcal{R}^*\big({-}\widetilde{\xi}_\sigma\big) \\ &= \mathcal{R}\big(\tfrac{1}{\sigma}(\widetilde{u}_\sigma{-}u_\tau^k)\big) + \mathcal{R}^*\big({-}\widetilde{\xi}_\sigma\big) - \big\langle {-}\widetilde{\xi}_\sigma, \tfrac{1}{\sigma}(\widetilde{u}_\sigma{-}u_\tau^k)\big\rangle \overset{(1.6)}{=} 0,\end{aligned}$$

by the Fenchel equivalence $\mu \in \partial\mathcal{R}(v) \;\Leftrightarrow\; \mathcal{R}(v) + \mathcal{R}^*(\mu) = \langle \mu, v\rangle$.

This observation (which we shall revisit in Section 4.5), motivates our first main result, see Theorem 4.2, that the De Giorgi estimate (1.4b) is indeed an equality, then called *De Giorgi identity*. This result is established for all measurable variational interpolants, under the sole additional assumption of radial differentiability of $\mathcal{R}$, cf. (1.5). However, for this version of the De Giorgi identity, it is essential to be more specific with the choice of $\widetilde{\xi}_\sigma \in \mathrm{D}\mathcal{E}(\widetilde{u}_\sigma)$: one has to restrict to those $\xi \in \mathrm{D}\mathcal{E}(\widetilde{u}_\sigma)$ that minimize $\mathcal{R}^*(-\xi)$ subject to the constraint of satisfying the Euler-Lagrange equation (1.6), see (3.17) for the precise definition. We refer to Example 3.9 for a very simple case, where this restriction is essential for the validity of the De Giorgi estimate as an identity.

For the case of a general $\mathcal{R}$, dropping radial differentiability, we are able to establish the De Giorgi estimate (1.4b) if X is a reflexive Banach space. In fact, along with (1.4b) we shall also obtain a refined estimate, involving a force selection $\widetilde{\xi}_\sigma \in \mathrm{D}\mathcal{E}(\widetilde{u}_\sigma)$ that also satisfies the Euler-Lagrange equation (1.6). In fact, we shall refer to (1.4b) as the *simple De Giorgi estimate*, which will be enhanced to the *improved De Giorgi estimate* keeping track of (1.6). Both, the *simple* and the *improved* De Giorgi estimates will be proved in our second main result, Theorem 4.12. For this, we use a Yosida-Moreau regularization $\mathcal{R}_\eta$ of $\mathcal{R}$ with an equivalent norm $\|\cdot\|$ such that $u \mapsto \|u\|^2$ is differentiable. Then, $\mathcal{R}_\eta$ is differentiable, in particular also radially differentiable, and for the corresponding gBGS $(X, \mathcal{E}, \mathcal{R}_\eta)$ the De Giorgi identity holds thanks to Theorem 4.2. It can be shown that in the limit passage $\eta \to 0^+$ the De Giorgi estimate survives.

While Section 3 introduces the definitions and conditions for the case of gradient systems in Banach space that will then be the focus of Section 4, we start in Section 2 with the metric case. The missing Euler-Lagrange equation is replaced by a purely metric identity, not involving the slope $|\partial\mathcal{E}|$ but rather the functions d^+ or d^- defined via

$$\mathsf{d}_\rho^-(u^\circ) = \inf\big\{\, \mathcal{D}(u^\circ, u) \,\big|\, u \in \mathrm{Argmin}\big(\mathcal{D}(u^\circ,\cdot)^2/(2\rho) + \mathcal{E}(\cdot)\big) \,\big\},$$
$$\mathsf{d}_\rho^+(u^\circ) = \sup\big\{\, \mathcal{D}(u^\circ, u) \,\big|\, u \in \mathrm{Argmin}\big(\mathcal{D}(u^\circ,\cdot)^2/(2\rho) + \mathcal{E}(\cdot)\big) \,\big\}.$$

For gMGS with differentiable ψ, the *metric energy identity* takes the form

$$\mathcal{E}(\widetilde{u}_\sigma) + \sigma\, \psi\big(\tfrac{1}{\sigma}\mathcal{D}(u^\circ, \widetilde{u}_\sigma)\big) + \int_{\rho=0}^{\sigma} \psi^*\Big(\psi'\big(\tfrac{1}{\rho}\mathsf{d}_\rho^\pm(u^\circ)\big)\Big)\,\mathrm{d}\rho = \mathcal{E}(u^\circ). \tag{1.7}$$

For $\psi(r) = \frac{1}{p}r^p$ the identity was established in [AGS05, Thm. 3.1.4], whereas the general case is contained in [Mie23, Sec. 4.2].

Clearly, the metric De Giorgi estimate (1.4a) follows easily from (1.7) by inserting the slope inequality

$$|\partial\mathcal{E}|(\widetilde{u}_\sigma) \leq \psi'\big(\tfrac{1}{\sigma}\mathcal{D}(\widetilde{u}_\sigma, u_\tau^k)\big)\, \big|\partial(-\mathcal{D}(u^\circ,\cdot)\big|(\widetilde{u}_\sigma) \leq \psi'\big(\tfrac{1}{\sigma}\mathcal{D}(\widetilde{u}_\sigma, u_\tau^k)\big), \tag{1.8}$$

see Proposition 2.9. The latter can be seen as a metric counterpart of the Euler-Lagrange equation. If one of the two inequalities in (1.8) is strict, then the De Giorgi identity is lost. The last inequality is strict, if $\mathcal{D}$ is not a length distance, which means that the metric De Giorgi estimate can only hold in geodesic spaces. However, even there the first inequality may be strict. In Theorem 2.15 we show that full continuity and a uniform slope estimate are sufficient to establish the De Giorgi identity in geodesic metric spaces.

2 The metric case

Following [Mie23], we specify in the following definition the notion of metric gradient system we will be working with hereafter.

Definition 2.1 *We call a quadruple* $(M, \mathcal{E}, \mathcal{D}, \psi)$ *a* generalized metric gradient system *(most often abbreviated to* gMGS*), if*

1. $(M, \mathcal{D})$ *is a complete metric space;*
2. $\mathcal{E} : M \to (-\infty, \infty]$ *is a proper (i.e., with non-empty domain* $\mathrm{dom}(\mathcal{E})$*) lsc functional;*
3. $\psi : \mathbb{R} \to [0, \infty)$ *is proper, convex, with* $\psi(0) = 0$ *and* $\lim_{r\to\infty} \frac{\psi(r)}{r} = \infty$.

For later use, we introduce the energy sublevels

$$S_E := \{u \in M \ : \ \mathcal{E}(u) \le E\}, \qquad E > 0. \tag{2.1}$$

Remark 2.2 *Most often, a* gMGS *is in fact individuated by a* quintuple $(M, \mathscr{T}, \mathcal{E}, \mathcal{D}, \psi)$, *where, mimicking the setup considered in [AGS05], in addition to the topology induced by the metric* $\mathcal{D}$*, a second (Hausdorff) topology* $\mathscr{T}$ *is considered on* M*. Typically,* $\mathscr{T}$ *is related to 'coercivity' properties of the energy functional, as it turns out to be the topology w.r.t. which the sublevel sets* S_E*, or the sublevels of a perturbation of* $\mathcal{E}$*, are compact. Although weaker than the topology induced by* $\mathcal{D}$*,* $\mathscr{T}$ *is related to it by the following compatibility condition (here* $\overset{\mathscr{T}}{\rightharpoonup}$ *denotes convergence with respect to* $\mathscr{T}$*):*

$$(u_n, v_n) \overset{\mathscr{T}}{\rightharpoonup} (u, v) \implies \lim_{n\to\infty} \mathcal{D}(u_n, v_n) \ge \mathcal{D}(u, v).$$

Nonetheless, we have opted for omitting the role of $\mathscr{T}$ *in the discussion of the metric case in order to avoid overburdening it, on the one hand, and to highlight the purely metric flavor of the arguments, on the other hand. Instead, in the Banach setup it will be convenient to encompass the weak topology in the picture.*

In the setup of a gMGS $(M, \mathcal{E}, \mathcal{D}, \psi)$, the classical notion of curve of maximal slope is extended by the following definition (cf. [Mie23, Def. 4.8]), which brings into play the convex conjugate $\psi^* : [0, \infty[\to [0, \infty[$, $\psi^*(s) = \sup_{r\in[0,\infty[}(sr - \psi(r))$, of ψ. To simplify the arguments, we fix an arbitrary $T > 0$ and confine the discussion to evolutions on the compact time interval $[0, T]$.

Definition 2.3 *Given a generalized metric gradient system* $(M, \mathcal{E}, \mathcal{D}, \psi)$*, we say that* $u : [0, T] \to M$ *is a* curve of maximal slope *if* $u \in \mathrm{AC}([0, T]; M)$ *and it satisfies for every* $0 \le s \le t \le T$

$$\mathcal{E}(u(t)) + \int_s^t \Big(\psi\big(|u'|(r)\big) + \psi^*\big(|\partial\mathcal{E}|(u(r))\big)\Big)\,\mathrm{d}r = \mathcal{E}(u(s))\,. \tag{2.2}$$

Remark 2.4 *In [RMS08, Prob. 2.6] an alternative definition for the above concept was given, imposing for the curve* $u \in \mathrm{AC}([0, T]; M)$ *the* pointwise estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(u(t)) \le -\psi\big(|u'|(t)\big) - \psi^*\big(|\partial\mathcal{E}|(u(t))\big) \qquad \text{for a.a. } t \in (0, T). \tag{2.3}$$

In fact, if the slope $|\partial\mathcal{E}|$ *is a* strong upper gradient *according to the terminology of [AGS05] (namely, if a suitable* chain-rule *inequality holds along* u*), then* (2.3) *is in fact equivalent to the energy-dissipation balance* (2.2)*.*

The Minimizing Movement Scheme for constructing curves of maximal slope fulfilling the initial condition $u(0) = u_0$, for an assigned initial datum $u_0 \in M$, then reads as follows: given a time step $\tau > 0$, inducing a (uniform, without loss of generality) partition $\mathscr{P}_\tau = \{t_\tau^k\}_{k=1}^{K_\tau}$ of the interval $[0,T]$, starting from $u_\tau^0 := u_0$ find $(u_\tau^k)_{k=1}^{K_\tau}$ such that

$$u_\tau^k \quad \text{minimizes} \quad M \ni u \mapsto \left(\tau\psi\big(\frac{1}{\tau}\mathcal{D}(u_\tau^{k-1}, u)\big) + \mathcal{E}(u) \right) \quad \text{for } k \in \{1,\ldots,K_\tau\}\,. \tag{MMS}$$

That is why, from now on we will study the properties of the single-step minimum problem

$$\operatorname*{Min}_{u\in M} \Phi_\sigma(u^\circ; u) \quad \text{with} \quad \Phi_\sigma(u^\circ; u) := \sigma\psi\big(\frac{1}{\sigma}\mathcal{D}(u^\circ, u)\big) + \mathcal{E}(u), \quad \sigma > 0, \tag{2.4}$$

for a fixed $u^\circ \in M$. We will also use the following notation

$$\phi(u^\circ;\sigma) = \inf\left\{ \Phi_\sigma(u^\circ;u) \,\middle|\, u \in M \right\} \text{ and } J_\sigma(u^\circ) = \operatorname{Argmin}\left\{ \Phi_\sigma(u^\circ;u) \,\middle|\, u \in M \right\}. \tag{2.5}$$

for the associated value functional and the set of minimizers (which we will assume non-empty, cf. (2.6) below). It is also significant to introduce the following quantities

$$\mathsf{d}_\sigma^-(u^\circ) := \inf\left\{ \mathcal{D}(u^\circ,u) \,\middle|\, u \in J_\sigma(u^\circ) \right\} \quad \text{and} \quad \mathsf{d}_\sigma^+(u^\circ) := \sup\left\{ \mathcal{D}(u^\circ,u) \,\middle|\, u \in J_\sigma(u^\circ) \right\}.$$

Throughout this section, we will work under the following assumptions.

Hypothesis 2.5 (Conditions for generalized metric gradient systems) *We assume that*

- $\psi \in \mathrm{C}^1(\mathbb{R})$ *is strictly convex;*
- $\mathcal{E}$ *is bounded from below by* E_0*, namely* $\inf_{u\in M}\mathcal{E}(u) > E_0 > 0$*;*
- *there exists* $\sigma_* > 0$ *such that*

$$J_\sigma(u^\circ) \neq \emptyset \text{ for all } \sigma \in (0,\sigma_*) \text{ and all } u^\circ \in \mathrm{dom}(\mathcal{E}). \tag{2.6}$$

Remark 2.6 *Whenever the generalized metric gradient system is individuated by a quintuple* $(M,\mathscr{T},\mathcal{E},\mathcal{D},\psi)$ *such that the topology* $\mathscr{T}$ *is compatible with* $\mathcal{D}$, (2.6) *follows from a coercivity property of the following type: there exists* $\sigma_* > 0$ *such that for all* $\sigma \in (0,\sigma_*)$ *and all* $u^\circ \in M$*, for all sequences* $(u_n)_n \subset M$ *we have the implication*

$$\sup_n \Phi_\sigma(u^\circ;u_n) < +\infty \implies (u_n)_n \text{ admits a } \mathscr{T}\text{-converging subsequence.} \tag{2.7}$$

Then, the direct method *yields the existence of minimizers for* (2.4)*.*

Since this paper is focused on the one-step minimum problem (2.4) and its 'starting' point u° is fixed, throughout most of the paper (up to a few exceptions) we will omit to indicate it in the notation for the functions and sets related to (2.4). Thus, we will simply write

$$\Phi_\sigma(u), \quad \phi(\sigma), \quad J_\sigma, \quad \mathsf{d}_\sigma^\pm\,.$$

We can in fact enhance (2.6) by observing that

$$\text{there exists a } \textit{measurable} \text{ selection } (0,\sigma_*) \ni \sigma \mapsto \widetilde{u}_\sigma \in J_\sigma \tag{2.8}$$

To show this, let us consider the multivalued mapping

$$\boldsymbol{J} : (0,\infty) \rightrightarrows X, \qquad \boldsymbol{J}(\sigma) := J_\sigma. \tag{2.9}$$

It is easy to check that $\boldsymbol{J}$ is upper semicontinuous from $\mathbb{R}$ to $(M,\mathcal{D})$ in that it fulfills for every $(\sigma_n)_n$, $\sigma \in (0,\infty)$

$$\sigma_n \to \sigma \implies \operatorname*{Ls}_{n\to\infty} \boldsymbol{J}(\sigma_n) \subset \boldsymbol{J}(\sigma)\,, \tag{2.10a}$$

where the *Kuratowski upper limit* (cf., e.g., [AmT04, Def. 4.4.13]) of the sequence of closed sets $(\boldsymbol{J}(\sigma_n))_n$ is defined by

$$u \in \operatorname*{Ls}_{n\to\infty} \boldsymbol{J}(\sigma_n) \iff \exists\,(\sigma_{n_k})_k,\ (u_k)_k \text{ with } \begin{cases} u_k \in \boldsymbol{J}(\sigma_{n_k}) \text{ for all } k\in\mathbb{N}\,, \\ \lim_{k\to\infty}\mathcal{D}(u,u_k) = 0\,. \end{cases} \tag{2.10b}$$

Since $\boldsymbol{J}$ is upper semicontinuous, its graph is a Borel subset of $(0,\infty)\times X$, and the von Neumann-Aumann selection theorem [CaV77, Thm. 3.22] applies, yielding (2.8).

The variational interpolant. Thanks to condition (2.6) the MMS does admit solutions $(u_\tau^k)_{k=1}^{N_\tau}$. We are then in a position to precisely introduce the notion of interpolant of the values $(u_\tau^k)_{k=1}^{N_\tau}$ we will focus on hereafter.

Definition 2.7 (De Giorgi variational interpolant) *We denote by* $\widetilde{u}_\tau : [0,T] \to M$ *any measurable function obtained by setting*

$$\begin{aligned} &\widetilde{u}_\tau(0) := u_\tau^0, \\ &\widetilde{u}_\tau(r) \in J_r(u_\tau^{k-1}) = \operatorname{Argmin}\big\{\, \Phi_r(u_\tau^{k-1};u) \,\big|\, u\in M \,\big\} \quad \textit{if } t = t_\tau^{k-1}{+}r. \end{aligned} \tag{2.11}$$

The cornerstone of the proof of the convergence, as $\tau \downarrow 0$, of (a subsequence of) the sequence $(\widetilde{u}_\tau)_\tau$ to a curve of maximal slope is the discrete estimate obtained by applying the De Giorgi estimate (1.4a) to the interpolant $\widetilde{u}_\tau$. The next section revolves around the validity of (1.4a) as an equality.

2.1 Metric energy identity and the De Giorgi estimate: statements and examples

The following identity was established in [AGS05, Sec. 3.1, eqn. (3.1.27)] for the case $\psi(\delta) = \delta^p/p$ and in [Mie23, Thm. 4.17] for general differentiable scalar dissipation potentials ψ.

Proposition 2.8 (Metric energy identity) *Under Hypothesis 2.5, any measurable selection* $(0,\sigma_*) \ni \sigma \to \widetilde{u}_\sigma \in J_\sigma(u^\circ)$ *fulfills*

$$\mathcal{E}(\widetilde{u}_\sigma) + \sigma\,\psi\big(\frac{1}{\sigma}\mathcal{D}(u^\circ,\widetilde{u}_\sigma)\big) + \int_0^\sigma \psi^*\big(\psi'\big(\frac{1}{\rho}\mathsf{d}_\rho^{\pm}(u^\circ)\big)\big)\,\mathrm{d}\rho = \mathcal{E}(u^\circ). \tag{2.12}$$

There is a straightforward way to relate the metric energy identity to De Giorgi estimate, and that is throughout the following result in which the slope at $\widetilde{u}_\sigma$ is estimated in terms of the slope of the distance function $\mathcal{D}(u^\circ,\cdot)$. In fact, (2.13) below extends to the setup of a gMGS $(M,\mathcal{E},\mathcal{D},\psi)$, the slope estimate proved in [AGS05, Lem. 3.1.3] in the quadratic case $\psi(r) = \frac{1}{2}r^2$. For completeness, we also record that a version of (2.13) was proved in [RMS08, Lemma 4.4] for non-differentiable dissipation potentials ψ. Note that this estimate is the metric counterpart of the Euler-Lagrange equation in the Banach-space setting, and hence it is less precise for a general gMGS.

Proposition 2.9 (Slope estimate) *Under Hypothesis 2.5,*

$$\big|\partial\mathcal{E}\big|(\widetilde{u}_\sigma) \le \psi'\big(\frac{1}{\sigma}\mathcal{D}(u^\circ,\widetilde{u}_\sigma)\big)\;\big|\partial(-\mathcal{D})(u^\circ,\cdot)\big|(\widetilde{u}_\sigma) \le \psi'\big(\frac{1}{\sigma}\mathcal{D}(u^\circ,\widetilde{u}_\sigma)\big). \tag{2.13}$$

Proof. We observe that for arbitrary $v \in M$ there holds

$$\begin{aligned}\mathcal{E}(\widetilde{u}_\sigma) - \mathcal{E}(v) &= \Phi_\sigma(\widetilde{u}_\sigma) - \Phi_\sigma(v) - \sigma\psi\big(\frac{1}{\sigma}\mathcal{D}(u^\circ,\widetilde{u}_\sigma)\big) + \psi\big(\frac{1}{\sigma}\mathcal{D}(u^\circ,v)\big)\\ &\overset{(1)}{\le} \sigma\Big(\psi\big(\frac{1}{\sigma}\mathcal{D}(u^\circ,v)\big) - \psi\big(\frac{1}{\sigma}\mathcal{D}(u^\circ,\widetilde{u}_\sigma)\big)\Big)\\ &\overset{(2)}{\le} \psi'\big(\frac{1}{\sigma}\mathcal{D}(u^\circ,\widetilde{u}_\sigma)\big)\,\Big(-\mathcal{D}(u^\circ,\widetilde{u}_\sigma) + \mathcal{D}(u^\circ,v)\Big).\end{aligned}$$

where (1) is due to $\widetilde{u}_\sigma \in J_\sigma$, whereas (2) follows by the convexity of ψ. Taking the positive part on both sides (using $\psi' \ge 0$), dividing by $\mathcal{D}(\widetilde{u}_\sigma, v)$, and taking the limsup for $v \to \widetilde{u}_\sigma$ gives the first estimate in (2.13).

By the triangle inequality for the distance $\mathcal{D}$, for any fixed $u^\circ \in M$ the functions $\mathcal{D}(u^\circ,\cdot)$ and $-\mathcal{D}(u^\circ,\cdot)$ have slope less or equal 1. Therefore, using $|-\partial\mathcal{D}(u^\circ,\cdot)| \le 1$, the second estimate in (2.13) follows. ∎

We can combine Propositions 2.8 and 2.9 and obtain the De Giorgi lemma for gMGS $(M,\mathcal{E},\mathcal{D},\psi)$, cf. also [RMS08, Lemma 4.5].

Theorem 2.10 (De Giorgi lemma for gMGS**)** *Under Hypothesis 2.5, any measurable selection* $(0,\sigma_*) \ni \sigma \mapsto \widetilde{u}_\sigma \in J_\sigma$ *fulfills*

$$\mathcal{E}(\widetilde{u}_\sigma) + \sigma\,\psi\big(\frac{1}{\sigma}\mathcal{D}(u^\circ,\widetilde{u}_\sigma)\big) + \int_0^\sigma \psi^*\big(\big|\partial\mathcal{E}\big|(\widetilde{u}_\rho)\big)\,\mathrm{d}\rho \le \mathcal{E}(u^\circ). \tag{2.14}$$

With the aim of improving (2.14) to an equality, taking into account the slope estimate, it is thus natural to check for cases in which $|\partial\mathcal{D}(u^\circ,\cdot)| = 1$. This is true for a *geodesic distance.*

Lemma 2.11 *Suppose in addition that the metric space* $(M,\mathcal{D})$ *is a* geodesic space*. Then, for every* $u^\circ \in M$ *we have* $|\partial\mathcal{D}(u^\circ,\cdot)|(u) = 1$ *for every* $u \neq u^\circ$*.*

Proof. Clearly, it suffices to show that $|\partial\mathcal{D}(u^\circ,\cdot)|(u) \ge 1$. For this, let us consider the constant-speed geodesic γ connecting u° to u, such that $\mathcal{D}(\gamma_t,\gamma_s) = (t-s)\mathcal{D}(u^\circ,u)$ for all $0 \le s \le t \le 1$. Then,

$$\begin{aligned}|\partial\mathcal{D}(u^\circ,\cdot)|(u) &= \limsup_{v\to u} \frac{(\mathcal{D}(u^\circ,u) - \mathcal{D}(u^\circ,v))^+}{\mathcal{D}(u,v)}\\ &\ge \limsup_{t\to 0^+} \frac{(\mathcal{D}(u^\circ,u) - \mathcal{D}(u^\circ,\gamma_t))^+}{\mathcal{D}(u,\gamma_t)} = \lim_{t\to 0^+} \frac{\mathcal{D}(u^\circ,u) - t\mathcal{D}(u^\circ,u)}{(1-t)\mathcal{D}(u^\circ,u)} = 1\,.\end{aligned} \tag{2.15}$$

∎

Remark 2.12 *We highlight that, for the validity of the above result it is sufficient for* $(M,\mathcal{D})$ *to be just a* length space *[BBI01, Chap. 2], namely with the property that the distance* $\mathcal{D}(u^\circ,u)$ *between any two points* u°, $u \in M$ *can be realized as the infimum of the lengths of the curves joining them. Then, for the lower estimate in* (2.15) *the role of the constant-speed geodesic* γ *connecting* u° *to* u *can be played by a curve with length arbitrarily close to* $\mathcal{D}(u^\circ,u)$*.*

In contrast, for the non-geodesic distance

$$\mathcal{D}(u,w) := \min\{|u-w|_2, R\} \quad \text{on } \mathbb{R}^n \tag{2.16}$$

with $R > 0$ a given constant (indeed, note that $(\mathbb{R}^n, \mathcal{D})$ is not a length space, either), we have $|\partial(\pm\mathcal{D})(u^\circ,\cdot)|(u) = 0$ whenever $|u-u^\circ| > R$.

We next provide two examples showing that, without a geodesic metric and without a continuous slope, we cannot expect (2.14) to hold as an equality.

Example 2.13 (Failure of equality in (2.14) **if** $(M,\mathcal{D})$ **not geodesic)** *We consider the* quadratic *metric gradient system* $(M,\mathcal{E},\mathcal{D})$ *with*

$$M = \mathbb{R}, \quad \mathcal{E}(u) = \frac{1}{2}u^2, \quad \mathcal{D}(u,w) = \min\{|w-u|, 1\}, \quad \psi(\delta) = \delta^2/2.$$

Starting with $u^\circ > 1$ *and setting* $\sigma_* = \big((u^\circ)^2-1\big)^{-1/2}$ *we obtain*

$$J_\sigma = \mathop{\mathrm{Argmin}}\Big\{ \frac{1}{\sigma}\mathcal{D}(u^\circ,u)^2 + \mathcal{E}(u) \,\Big|\, u \in \mathbb{R} \Big\} = \begin{cases} \frac{u^\circ}{1+\sigma} & \text{for } \sigma < \sigma_*, \\ \big\{\frac{u^\circ}{1+\sigma}, 0\big\} & \text{for } \sigma = \sigma_*, \\ 0 & \text{for } \sigma > \sigma_*. \end{cases}$$

We can calculate all terms in (2.14) *and find equality as long as* $\sigma \le \sigma_*$*; but strict inequality holds for* $\sigma > \sigma_*$.

Note that $|\partial\mathcal{E}|(\widetilde{u}_\sigma) = |\partial\mathcal{E}|(0) = 0$ *for* $\sigma > \sigma_*$ *but* $\psi'\big(\frac{1}{\sigma}\mathcal{D}(u^\circ,\widetilde{u}_\sigma)\big) = \frac{1}{\sigma}\mathcal{D}(u^\circ,0) = 1/\sigma \gneqq 0$*. Hence, the first estimate in Proposition 2.9 holds as an equality, but the second estimate is strict because of* $\big|\partial(-\mathcal{D})(u^\circ,\cdot)\big|(0) = 0 < 1$.

Our second counterexample involves a *discontinuous* slope functional.

Example 2.14 (Failure of equality in (2.14) **if the slope of** $\mathcal{E}$ **is not continuous)** *We consider the (again, quadratic) metric gradient system*

$$M = \mathbb{R}, \quad \mathcal{E}(u) = \max\{u, 0\}, \quad \mathcal{D}(u,w) = |u-w|, \quad \psi(\delta) = \delta^2/2.$$

Starting at $u^\circ = 1$ *we find the unique variational interpolant* $\widetilde{u}_\sigma = \max\{1-\sigma, 0\}$*. The curve* $\sigma \mapsto \widetilde{u}_\sigma$ *is absolutely continuous but the slope along the curve is discontinuous, namely*

$$|\partial\mathcal{E}|(\widetilde{u}_\sigma) = 1 \;\; \text{for } \sigma \in [0,1[\quad \text{and} \quad |\partial\mathcal{E}|(\widetilde{u}_\sigma) = 0 \;\; \text{for } \sigma \ge 1.$$

This time, the first estimate in Proposition 2.9 is strict for $\sigma \ge 1$*, and hence* (2.14) *is also a strict inequality for* $\sigma > 1$.

2.2 Equality in the De Giorgi estimate

The discussion in Section 2.1 has highlighted the link between two properties (one related to the geometry of the space, the other to the driving energy and its slope), and equality in the De Giorgi Lemma. With Theorem 2.15 we now prove that the joint validity of such properties is a sufficient, albeit rather strong, condition for a gMGS to guarantee that all measurable variational interpolants satisfy estimate (2.14) with equality. The idea is that, if $|\partial\mathcal{E}|$ is upper semicontinuous, points where the slope is 'too big' are avoided by minimizers.

Theorem 2.15 *Consider a* gMGS $(M, \mathcal{E}, \mathcal{D}, \psi)$ *satisfying Hypothesis 2.5 such that, additionally,*

1. $(M, \mathcal{D})$ *is a geodesic space, and*
2. $\mathcal{E}$ *is continuous on* M*,*
3. $|\partial\mathcal{E}|$ *is upper semicontinuous on* M*.*

Then, for any measurable selection $(0, \sigma_*) \ni \sigma \mapsto \widetilde{u}_\sigma \in J_\sigma$ *the relation* (2.14) *holds with equality.*

Proof. It suffices to show that the slope estimate from Proposition 2.9 holds with equality and combine this with Lemma 2.11. Since the upper estimate $|\partial\mathcal{E}|(\widetilde{u}_\sigma) \leq \psi'\big(\frac{1}{\tau}\mathcal{D}(u^\circ, \widetilde{u}_\sigma)\big)$ is clear, it remains to show $|\partial\mathcal{E}|(\widetilde{u}_\sigma) \geq \psi'\big(\frac{1}{\sigma}\mathcal{D}(u^\circ, \widetilde{u}_\sigma)\big)$.

We consider the geodesic curve $[0, 1] \ni \theta \mapsto \gamma_\theta$ with $\gamma_0 = u^\circ$ and $\gamma_1 = \widetilde{u}_\sigma$. Since $\gamma_1 = \widetilde{u}_\sigma$ is a global minimizer for $\Phi_\sigma(\cdot)$, we can compare with $u = \gamma_\theta$ and obtain

$$\begin{aligned}\mathcal{E}(\gamma_\theta) - \mathcal{E}(\gamma_1) &= \mathcal{E}(\gamma_\theta) - \mathcal{E}(\widetilde{u}_\sigma) \geq \sigma\,\psi\big(\frac{1}{\sigma}\mathcal{D}(u^\circ, \widetilde{u}_\sigma)\big) - \sigma\,\psi\big(\frac{1}{\sigma}\mathcal{D}(u^\circ, \gamma_\theta)\big)\\ &\geq \psi'\big(\frac{1}{\sigma}\mathcal{D}(u^\circ, \gamma_\theta)\big)\,\big(\mathcal{D}(\gamma_0, \gamma_1) - \mathcal{D}(\gamma_0, \gamma_\theta)\big) = \psi'\big(\frac{1}{\sigma}\mathcal{D}(u^\circ, \gamma_\theta)\big)\,\mathcal{D}(\gamma_\theta, \gamma_1),\end{aligned}$$

where we used the convexity of ψ and the fact that $(\gamma_\theta)_\theta$ is a geodesic. Thus, we have

$$S_\theta := \frac{\mathcal{E}(\gamma_\theta) - \mathcal{E}(\gamma_1)}{\mathcal{D}(\gamma_\theta, \gamma_1)} \geq \psi'\big(\frac{1}{\sigma}\mathcal{D}(u^\circ, \gamma_\theta)\big) \quad \text{for all } \theta \in [0, 1).$$

Now fix θ and define the function $h_\theta(s) = \mathcal{E}(\gamma_s) - S_\theta\mathcal{D}(\gamma_s, \gamma_1)$ for $s \in [\theta, 1]$. Since $\mathcal{E}$ is continuous, h_θ is continuous, too, and satisfies $h_\theta(\theta) = h_\theta(1) = \mathcal{E}(\gamma_1)$. Hence, there exists a maximizer $t_\theta \in (\theta, 1)$ of h_θ. For this t_θ and $s \in (t_\theta, 1]$ we have, by construction,

$$\mathcal{E}(\gamma_{t_\theta}) - \mathcal{E}(\gamma_s) \geq S_\theta(\mathcal{D}(\gamma_{t_\theta}, \gamma_1)) - \mathcal{D}(\gamma_s, \gamma_1)\big) = S_\theta\mathcal{D}(\gamma_{t_\theta}, \gamma_s).$$

Dividing by $\mathcal{D}(\gamma_{t_\theta}, \gamma_s)$ and taking the limit $s \searrow t_\theta$ we find

$$|\partial\mathcal{E}|(\gamma_{t_\theta}) \geq S_\theta \geq \psi'\big(\frac{1}{\sigma}\mathcal{D}(u^\circ, \gamma_\theta)\big).$$

In the limit $\theta \nearrow 1$ we have $\gamma_{t_\theta} \to \gamma_1 = \widetilde{u}_\sigma$, and the upper semicontinuity of $|\partial\mathcal{E}|$ yields

$$|\partial\mathcal{E}|(\widetilde{u}_\sigma) = |\partial\mathcal{E}|(\gamma_1) \geq \limsup_{\theta\nearrow 1}|\partial\mathcal{E}|(\gamma_{t_\theta}) \geq \lim_{\theta\nearrow 1}\psi'\big(\frac{1}{\sigma}\mathcal{D}(u^\circ, \gamma_\theta)\big) = \psi'\big(\frac{1}{\sigma}\mathcal{D}(u^\circ, \widetilde{u}_\sigma)\big),$$

which is the desired opposite inequality to (2.12), and the assertion follows. ■

3 Banach case: examples and preliminary results

3.1 Setup

We consider a generalized Banach-space gradient system (gBGS, for short) $(X, \mathcal{E}, \mathcal{R})$ where the state space

$$X \text{ is a reflexive Banach space.} \tag{X}$$

We will thus study the minimum problem

$$\operatorname*{Min}_{u\in X} \Phi_\sigma(u) \qquad \text{with} \;\; \Phi_\sigma(u) := \sigma\mathcal{R}\Big(\frac{1}{\sigma}(u-u^\circ)\Big) + \mathcal{E}(u), \quad \sigma>0, \tag{3.1}$$

for a fixed $u^\circ \in \mathrm{dom}(\mathcal{E}) \subset X$.

We now collect our working assumptions on $\mathcal{E}$ and $\mathcal{R}$, which partly mirror those collected in Definition 2.3 and Hypothesis 2.5. At the same time, they clearly reflect the underlying Banach setup, involving the weak topology on X, in addition to the norm topology, in conditions (3.4) below (cf. also Remark 2.6). As already mentioned in the Introduction, in the Banach setting we will allow for *nonsmooth* energies, and thus work with the *Fréchet subdifferential* $\partial\mathcal{E}$ of $\mathcal{E}$ in place of its Gâteau derivative $\mathrm{D}\mathcal{E}$. We recall that the multivalued operator $\partial\mathcal{E} : X \rightrightarrows X^*$ is defined at $u \in \mathrm{dom}(\mathcal{E})$ by

$$\xi \in \partial\mathcal{E}(u) \quad \text{if and only if} \quad \mathcal{E}(w) - \mathcal{E}(u) \geq \langle \xi, w-u\rangle + \mathrm{o}(\|w-u\|_X) \text{ as } w \to u\,. \tag{3.2}$$

Then, in (3.4b) we ask for closedness of the graph of $\partial\mathcal{E}$, w.r.t. the weak topology of $X\times X^*$, along sequences with bounded energy.

Hypothesis 3.1 (Conditions for gBGS**)** *We assume that*

- *The dissipation potential* $\mathcal{R} : X \to [0,\infty]$ *has a proper domain* $\mathrm{dom}(\mathcal{R})$ open *in* X*, it is lower semicontinuous, convex, and fulfills the following conditions:*

$$\mathcal{R}(0) = 0, \qquad \lim_{\|v\|\to\infty} \frac{\mathcal{R}(v)}{\|v\|} = \infty. \tag{3.3}$$

- *The energy functional* $\mathcal{E} : X \to (-\infty,\infty]$ *is proper, bounded from below, and weakly-sequentially lower semicontinuous, i.e. for all* $(u_n)_n \subset X$

$$u_n \rightharpoonup u \text{ in } X \implies \liminf_{n\to\infty} \mathcal{E}(u_n) \geq \mathcal{E}(u)\,, \tag{3.4a}$$

and $\partial\mathcal{E} : X \rightrightarrows X^*$ *is closed on energy sublevels (cf. notation* (2.1)*) , i.e.*

$$\forall\, E>0: \left\{ \begin{array}{cc} (u_n,\xi_n) \rightharpoonup (u,\xi) & \text{in } X\times X^*, \\ u_n \in S_E,\ \xi_n \in \partial\mathcal{E}(u_n) & \text{for all } n\in\mathbb{N} \end{array} \right\} \implies \xi \in \partial\mathcal{E}(u)\,. \tag{3.4b}$$

We pin down some elementary, but important, consequences of our assumptions on $\mathcal{R}$, for its convex conjugate

$$\mathcal{R}^* : X^* \to [0,\infty], \qquad \mathcal{R}^*(\xi) = \sup_{v\in X}(\langle\xi,v\rangle - \mathcal{R}(v)),$$

which is, by construction, lower semicontinuous and convex on X^*.

Lemma 3.2 *The functional* $\mathcal{R}^* : X^* \to [0,\infty]$ *has domain* $\mathrm{dom}(\mathcal{R}^*) = X^*$*, it fulfills* $\mathcal{R}^*(0)=0$*, and the* coercivity *estimate*

$$\exists\, r, K_r > 0 \;\forall\, \xi \in X^* : \quad \mathcal{R}^*(\xi) \geq r\|\xi\|_* - K_r\,. \tag{3.5}$$

Furthermore, $\mathcal{R}$ *is locally Lipschitz on* $\mathrm{dom}(\mathcal{R})$*.*

Proof. Since $\mathcal{R}$ has superlinear growth at infinity, $\mathcal{R}^*$ has full domain (cf., e.g., [Bre73, Prop. 2.11, 2.14]). It can be immediately checked that $\mathcal{R}^*(0) = 0$. To show (3.5), we observe that since 0 belongs to the open set $\mathrm{dom}(\mathcal{R})$, $\mathcal{R}$ is continuous in 0 [EkT99, Chap. I, Cor. 2.5] and, thus, there exists $r > 0$ such that $\overline{B_r(0)} \subset \mathrm{dom}(\mathcal{R})$ and $\mathcal{R}$ is bounded from above on $\overline{B_r(0)}$. Therefore,

$$\mathcal{R}^*(\xi) \geq \langle \xi, \ell r w\rangle - \mathcal{R}(\ell r w) \qquad \text{for all } w \in \overline{B_1(0)},\ \ell \in [0,1]\,.$$

Taking into account that $\mathcal{R}(\ell r w) \leq \ell \mathcal{R}(rw) \leq \mathcal{R}(rw)$, we thus conclude

$$r\|\xi\|_* = r \sup_{w\in\overline{B_1(0)}} |\langle \xi, w\rangle| \leq \mathcal{R}^*(\xi) + K \qquad \text{with } K_r = \sup_{v\in\overline{B_r(0)}} \mathcal{R}(v) < \infty\,.$$

The last assertion follows from [EkT99, Chap. I, Cor. 2.4], which states that $\mathcal{R}$ is locally Lipschitz on (the open) $\mathrm{dom}(\mathcal{R})$ if and only if there exists a non-empty convex set over which $\mathcal{R}$ is bounded from above which, in our case, is the aforementioned ball $B_r(0)$. ■

Remark 3.3 *For most part of the subsequent analysis, the open-domain condition for $\mathcal{R}$ will be sufficient. One the one hand it implies the coercivity property* (3.5) *for $\mathcal{R}^*$, and on the other hand it implies local Lipschitz continuity of $\mathcal{R}$ in all points of the domain. Only for specific results (see Lemma 4.7 and Proposition 4.14 ahead), we will have to require that $\mathcal{R}^*$ has superlinear growth at infinity, which in particular implies that $\mathcal{R}$ has full domain.*

It remains an open problem, though, to deal with the case in which $\mathrm{dom}(\mathcal{R})$ *is a general proper subset of X, not necessarily open, which occurs, for instance, when the gradient system $(X, \mathcal{E}, \mathcal{R})$ models the unidirectional evolution of inelastic processes in solid mechanics like damage.*

Remark 3.4 *It is often significant to consider dissipation potentials that also depend on the state variable, i.e. $\mathcal{R} = \mathcal{R}(u, v)$, but this generalization would be irrelevant for the study of the properties of the single-step minimum problem* (3.1) *for a fixed $u^\circ \in X$. Indeed, in the state-dependent case the dissipation term would be simply replaced by $\mathcal{R}\left(u^\circ, \frac{1}{\sigma}(u - u^\circ)\right)$.*

We emphasize that the closedness condition (3.4b) assumes only *weak* convergence in X on sequences $(u_n)_n$. However, the additional assumptions $(u_n)_n \subset S_E$ and the existence of a *bounded* sequence $(\xi_n)_n$ such that $\xi_n \in \partial\mathcal{E}(u_n)$ for all $n \in \mathbb{N}$, often grants extra compactness properties to the sequence $(u_n)_n$.

Nonetheless, it would also be possible to bypass (3.4b) by working with the so-called *limiting subdifferential* of $\mathcal{E}$, defined at a given $u \in \mathrm{dom}(\mathcal{E})$ as the set of all ξ that are weak limits of sequences $(\xi_n)_n$, with $\xi_n \in \partial\mathcal{E}(u_n)$ for every $n \in \mathbb{N}$, and $u_n \to u$ with $\sup_n \mathcal{E}(u_n) < \infty$. Gradient flows and generalized gradient systems featuring this subdifferential notion were analyzed in [RoSa06] and [MRS13], respectively.

For the minimum problem (3.1), we will stick to notation (2.5), namely

- $\phi = \phi(\sigma)$ for the value functional associated with the above minimum problem; in fact, we shall also refer to $\phi(u^\circ; \cdot)$ as *marginal function.* We remark for later use that, in analogy to the metric case in [AGS05], it was proved in [MRS13, Lem. 6.1] that
$$\lim_{\sigma\downarrow 0} \phi(\sigma) = \mathcal{E}(u^\circ)\,, \tag{3.6}$$
where the superlinearity of $\mathcal{R}$ in (3.3) is essential.

- and J_σ for the set of minimizers.

Additionally, in the Banach setup under Hypothesis 3.1 every $\widetilde{u}_\sigma \in J_\sigma$ satisfies the Euler-Lagrange equation for (3.1), namely

$$0 \in \partial\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) + \widetilde{\xi}_\sigma \quad \text{with } \widetilde{\xi}_\sigma \in \partial\mathcal{E}(\widetilde{u}_\sigma). \tag{3.7}$$

This follows as in [MRS13, Prop. 4.2] (which, in turn, relies on results from [Mor06]), whose proof also uses that $\mathcal{R}$ is locally Lipschitz around the point $\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)$; note that this is granted by Lemma 3.2. In fact, [MRS13, Prop. 4.2] would apply also if we worked with the notion of *limiting subdifferential* previously mentioned.

Furthermore, under Hypothesis 3.1 the following holds:

1. We have $J_\sigma \neq \emptyset$ for all $u^\circ \in \mathrm{dom}(\mathcal{E})$ and all $\sigma > 0$.

   For this, it suffices to observe that any infimizing sequence for $\Phi_\sigma(\cdot)$ is bounded in X (thanks to the facts that $\mathcal{E}$ is bounded from below and $\mathcal{R}$ has superlinear growth), and to resort to the weak lower semicontinuity of $\mathcal{E}$ and $\mathcal{R}$.

2. The multivalued mapping $\boldsymbol{J}\colon (0,\infty) \rightrightarrows X;\ \sigma \mapsto \boldsymbol{J}(\sigma) := J_\sigma$ (cf. (2.9)) is upper semicontinuous from $\mathbb{R}$ to X, in the sense that inclusion (2.10a) holds (even for the Kuratowski upper limit $\mathrm{Ls}^{\mathrm{weak}}_{n\to\infty} \boldsymbol{J}(\sigma_n)$ defined in terms of the *weak* topology on X). Therefore, by [CaV77, Thm. 3.22] we may conclude the existence of a (strongly) *measurable* selection $(0,\infty) \ni \sigma \to \widetilde{u}_\sigma \in J_\sigma$.

We conclude this section with the example of a gBGS $(X, \mathcal{E}, \mathcal{R})$ fulfilling the conditions from Hypothesis 3.1. To keep the exposition simple, we confine the discussion to a Gâteaux differentiable energy $\mathcal{E}$, where $\partial\mathcal{E}(u)$ is always a singleton. This is not a significant restriction when we revisit Example 3.5 later on, because our focus will rather be on the properties of the dissipation potential $\mathcal{R}$. Nonetheless, it would not be difficult to adjust the conditions in such a way as to allow for a nonsmooth, but λ-convex, potential W in (3.9a) below.

Example 3.5 We consider a gBGS $(X, \mathcal{E}, \mathcal{R})$ such that

- $X = \mathrm{L}^p(\Omega)$, $p > 1$, with a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$;
- the dissipation potential $\mathcal{R} : \mathrm{L}^p(\Omega) \to [0,\infty)$ is defined by $\mathcal{R}(v) = \int_\Omega \mathrm{R}(v(x))\,\mathrm{d}x$ with

$$\begin{gathered}\mathrm{R} : \mathbb{R} \to [0,\infty) \text{ convex, s.t. } \mathrm{R}(0) = 0 \text{ and} \\ \exists\, \kappa,\ K > 0\ \ \forall\, x, y \in \mathbb{R} :\ \begin{cases} \mathrm{R}(x) \geq \kappa|x|^p - K, \\ \mathrm{R}^*(y) \geq \kappa|y|^{p'} - K, \end{cases}\end{gathered} \tag{3.8}$$

  where p' is the dual exponent to p.

- the energy functional $\mathcal{E} : \mathrm{L}^p(\Omega) \to (-\infty, \infty]$ features a (possibly) nonconvex, but lower order, perturbation of the Dirichlet integral, i.e. it is defined via

$$\mathcal{E}(u) := \begin{cases} \int_\Omega \frac{1}{2}|\nabla u|^2 + W(u)\,\mathrm{d}x & \text{for } u \in \mathrm{H}^1_0(\Omega) \text{ and } W(u) \in \mathrm{L}^1(\Omega)\,, \\ \infty & \text{otherwise}\,. \end{cases} \tag{3.9a}$$

  Along the footsteps of [RMS08, Sec. 7], we require for the potential energy density $W : \mathbb{R} \to \mathbb{R}$ that $W \in \mathrm{C}^2(\mathbb{R})$ and

$$\exists\, C_W > 0\ \exists\, s_p \in (1, \tfrac{p_d}{p'})\ \forall\, r \in \mathbb{R} :\ \begin{cases} W''(r) \geq -C_W\,, \\ W(r) \geq -C_W\,, \\ |W'(r)| \leq C_W(1 + |r|^{s_p})\,. \end{cases} \tag{3.9b}$$

As shown in [RMS08, Sec. 7] (cf. also [MRS23, Sec. 4.2]), $\mathcal{E}$ complies with the lower semicontinuity and closedness conditions (3.4), where $\partial\mathcal{E}(u) = \{ -\Delta u + W'(u) \} \subset \mathrm{H}^{-1}(\Omega) = \mathrm{H}^1_0(\Omega)^*$.

3.2 The De Giorgi estimate and identity in the Banach setup

In order to state the Banach-space versions of the metric energy identity (2.12) and of the De Giorgi estimate (2.14), we are naturally led to introduce the following object, which corresponds to the "slope part of the dissipation" in (2.14).

Definition 3.6 ($\mathcal{R}$-slope of the energy) *Let $u \in \mathrm{dom}(\partial\mathcal{E})$. The quantity*

$$\mathcal{S}_{\mathcal{R}}(u) := \inf \left\{ \mathcal{R}^*(-\xi) \,\middle|\, \xi \in \partial\mathcal{E}(u) \right\} \tag{3.10}$$

is called $\mathcal{R}$-slope *of the energy functional $\mathcal{E}$ at u.*

Indeed, as a consequence of the coercivity property (3.5) of $\mathcal{R}^*$, guaranteeing that infimizing sequences for the above minimum problem are bounded in X^*, and of the closedness property (3.4b), the infimum in (3.10) is attained, i.e. we have that

$$\mathcal{S}_{\mathcal{R}}(u) = \min \left\{ \mathcal{R}^*(-\xi) \,\middle|\, \xi \in \partial\mathcal{E}(u) \right\}. \tag{3.11a}$$

To see this, consider $\xi_n \in \partial\mathcal{E}(u)$ with $\mathcal{R}^*(-\xi_n) \to \mathcal{S}_{\mathcal{R}}(u)$, then the coercivity of $\mathcal{R}^*$ allows to extract a subsequence (not relabeled) such that $\xi_n \rightharpoonup \xi$ and the sequence $(u, \xi_n) \rightharpoonup (u, \xi)$ satisfies the assumptions of (3.4b) with energy level $E = \mathcal{E}(u)$. Hence, we have $\xi \in \partial\mathcal{E}(u)$ and $\mathcal{S}_{\mathcal{R}}(u) = \mathcal{R}^*(-\xi)$ follows. Likewise, it is immediate to check that, for every $(u_n)_n$, $u \in X$ in some energy sublevel (recall that the closedness property (3.4b) holds under an energy bound), we have

$$(u_n)_n,\, u \in S_E, \quad u_n \rightharpoonup u \quad \Longrightarrow \quad \liminf_{n\to\infty} \mathcal{S}_{\mathcal{R}}(u_n) \geq \mathcal{S}_{\mathcal{R}}(u). \tag{3.11b}$$

Obviously, one may expect that the $\mathcal{R}$-slope $\mathcal{S}_{\mathcal{R}}(\widetilde{u}_\sigma)$, evaluated along a (measurable) selection $\sigma \mapsto \widetilde{u}_\sigma \in J_\sigma$, will play the role of the term $\psi^*\big(\big|\partial\mathcal{E}\big|(\widetilde{u}_\sigma)\big)$ in the Banach-space version of (2.14). However, a more careful comparison with the metric identity (2.12), featuring the quantity

$$\tfrac{1}{\sigma}\mathsf{d}_\sigma^- = \inf \left\{ \|\tfrac{1}{\sigma}(u - u^\circ)\| \,\middle|\, u \in J_\sigma \right\},$$

suggests that the notion of slope has to be adjusted. In fact, it is expedient to bring into the picture the additional structure available in the Banach setup, namely the fact that every $u \in J_\sigma$ fulfills the Euler-Lagrange equation (3.7).

We thus set forth a "*conditioned* slope part of the dissipation", defined by the minimization of the dual dissipation potential $\mathcal{R}^*$ over selections $\xi \in \partial\mathcal{E}(u)$ that *additionally* satisfy the Euler-Lagrange equation. Accordingly, we introduce a multivalued operator which encodes the validity of (3.7). We shall refer to these two objects as *conditioned $\mathcal{R}$-slope* and *conditioned subdifferential*, respectively. Notice that the conditioned $\mathcal{R}$-slope is naturally defined on the graph of the multivalued operator $\boldsymbol{J}$, namely on the set

$$\mathrm{Graph}(\boldsymbol{J}) = \{(\sigma, u) \in (0,\infty) \times X \;:\; u \in \boldsymbol{J}(\sigma) = J_\sigma\}.$$

Definition 3.7 (*Conditioned* subdifferential / slope of energy) *The multivalued mapping* $\partial_{\mathcal{R}}\mathcal{E} : \mathrm{Graph}(\boldsymbol{J}) \rightrightarrows X^*$ *defined by*

$$\partial_{\mathcal{R}}\mathcal{E}(\sigma; u) := \Big\{ \xi \in X^* \,\Big|\, \xi \in \partial\mathcal{E}(u) \ \textit{and} \ -\xi \in \partial\mathcal{R}\big(\frac{1}{\sigma}(u - u^\circ)\big) \Big\} \tag{3.12}$$

is called conditioned subdifferential of the energy $\mathcal{E}$. *The quantity*

$$\mathcal{C}_{\mathcal{R}}(\sigma; u) := \inf\big\{ \mathcal{R}^*(-\xi) \,\big|\, \xi \in \partial_{\mathcal{R}}\mathcal{E}(\sigma; u) \big\} \quad \textit{for all } (\sigma, u) \in \mathrm{Graph}(\boldsymbol{J}) \tag{3.13}$$

is called conditioned $\mathcal{R}$-slope *of* $\mathcal{E}$ *at* $u \in \mathrm{dom}(\partial\mathcal{E})$.

Although with slight abuse, the notation for $\partial_{\mathcal{R}}\mathcal{E}$ highlights the geometry induced by the dissipation potential through the Euler-Lagrange equation (3.7).

Clearly, since every $u \in J_\sigma$ fulfills the Euler-Lagrange equation (3.7), we have $\partial_{\mathcal{R}}\mathcal{E}(\sigma; u) \neq \emptyset$ and thus $\mathcal{C}_{\mathcal{R}}(\sigma; u) < \infty$ for all $(\sigma, u) \in \mathrm{Graph}(\boldsymbol{J})$. Obviously, in general there holds

$$\mathcal{C}_{\mathcal{R}}(\sigma; u) \geq \mathcal{S}_{\mathcal{R}}(u) \qquad \text{for all } (\sigma, u) \in \mathrm{Graph}(\boldsymbol{J}), \tag{3.14}$$

which seems to suggest $\mathcal{C}_{\mathcal{R}}$ as the 'right' object for the attainment of an equality in the De Giorgi estimate, cf. also Example 3.9 below.

We are now in the position to precisely introduce the estimate/identity whose validity we are are going to address hereafter.

Definition 3.8 *We say that a (strongly) measurable selection* $(0, \infty) \ni \sigma \mapsto \widetilde{u}_\sigma \in J_\sigma$ *fulfills, on some interval* $(0, \sigma_*)$,

the ***simple De Giorgi estimate*** *if for all* $\sigma \in (0, \sigma_*)$ *we have*

$$\mathcal{E}(\widetilde{u}_\sigma) + \sigma\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) + \int_{\rho=0}^{\sigma} \mathcal{S}_{\mathcal{R}}(\widetilde{u}_\rho)\,\mathrm{d}\rho \leq \mathcal{E}(u^\circ); \tag{3.15}$$

the ***improved De Giorgi estimate*** *if for all* $\sigma \in (0, \sigma_*)$ *we have*

$$\mathcal{E}(\widetilde{u}_\sigma) + \sigma\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) + \int_{\rho=0}^{\sigma} \mathcal{C}_{\mathcal{R}}(\rho; \widetilde{u}_\rho)\,\mathrm{d}\rho \leq \mathcal{E}(u^\circ); \tag{3.16}$$

the ***De Giorgi identity*** *if* (3.16) *holds as an equality, i.e. for all* $\sigma \in (0, \sigma_*)$*:*

$$\mathcal{E}(\widetilde{u}_\sigma) + \sigma\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) + \int_{\rho=0}^{\sigma} \mathcal{C}_{\mathcal{R}}(\rho; \widetilde{u}_\rho)\,\mathrm{d}\rho = \mathcal{E}(u^\circ)\,. \tag{3.17}$$

Later on, in Section 4.3 we will provide a sufficient condition for the attainment of the infimum in the definition (3.13) of $\mathcal{C}_{\mathcal{R}}$, which is also related to the validity of the analog of the lower semicontinuity properties (3.11b). This will lead to the existence of measurable selections

$$(0, \infty) \ni \sigma \mapsto \widetilde{\xi}_\sigma \in \partial_{\mathcal{R}}\mathcal{E}(\sigma; \widetilde{u}_\sigma), \quad \text{with } \sigma \mapsto \widetilde{u}_\sigma \in \boldsymbol{J}(\sigma) \text{ a measurable selection,}$$

such that estimate (3.16) rephrases as

$$\mathcal{E}(\widetilde{u}_\sigma) + \sigma\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) + \int_0^{\sigma} \mathcal{R}^*(-\widetilde{\xi}_\rho)\,\mathrm{d}\rho \leq \mathcal{E}(u^\circ)\,,$$

and analogously for (3.17).

We conclude this section by using Example 2.14, revisited in the Banach setup, to convey the idea that the finer information encoded in the conditioned $\mathcal{R}$-slope of the energy may play a key role for the attainment of the De Giorgi identity.

Example 3.9 (Example 2.14 in the Banach setup) *We now treat Example 2.14 as a* gBGS *in the Hilbert or Banach space* $X = \mathbb{R}$*:*

$$M = \mathbb{R}, \quad \mathcal{E}(u) = \max\{u, 0\} =: u^+, \quad \mathcal{R}(v) = \frac{1}{2}v^2.$$

Since $\mathcal{E}$ *is convex, its Fréchet subdifferential* $\partial\mathcal{E} : \mathbb{R} \rightrightarrows \mathbb{R}$ *coincides with the subdifferential in the sense of convex analysis, and it is set-valued with* $\partial\mathcal{E}(0) = [0,1]$*. The* $\mathcal{R}$*-slope of the energy is given by*

$$\mathcal{S}_\mathcal{R}(u) = \frac{1}{2} \text{ for } u > 0 \quad \text{and} \quad \mathcal{S}_\mathcal{R}(u) = 0 \text{ for } u \leq 0.$$

As before, starting from $u^\circ = 1$ *we calculate the variational interpolant* $\widetilde{u}_\sigma = (1-\sigma)^+ = \max\{1-\sigma, 0\}$*, so that*

$$\partial\mathcal{E}(\widetilde{u}_\sigma) = \begin{cases} \{1\} & \text{if } \sigma \in (0,1), \\ [0,1] & \text{if } \sigma \geq 1. \end{cases}$$

Therefore, for $\sigma > 0$ *we have*

$$\mathcal{E}(u^\circ) - \mathcal{E}(\widetilde{u}_\sigma) - \sigma\mathcal{R}\big(\tfrac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) = 1-(1-\sigma)^+ - \tfrac{1}{2\sigma}|(1-\sigma)^+ - 1|^2 = \begin{cases} \frac{1}{2}\sigma & \text{if } \sigma \leq 1, \\ 1-\frac{1}{2\sigma} & \text{if } \sigma \geq 1. \end{cases}$$

Hence, the estimate

$$\int_0^\sigma \mathcal{R}^*\big(-\widetilde{\xi}_\rho\big)\,\mathrm{d}\rho \leq \mathcal{E}(u^\circ) - \mathcal{E}(\widetilde{u}_\sigma) - \sigma\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) \quad \text{with } \widetilde{\xi}_\rho \in \partial\mathcal{E}(\widetilde{u}_\rho) \tag{3.18}$$

holds for $\sigma \in (0,1]$*, while for* $\sigma > 1$ *it only holds for selections* $\sigma \mapsto \widetilde{\xi}_\sigma \in \partial\mathcal{E}(\widetilde{u}_\sigma)$ *with*

$$\int_0^\sigma \mathcal{R}^*\big(-\widetilde{\xi}_\rho\big)\,\mathrm{d}\rho = \int_0^1 \frac{1}{2}\,\mathrm{d}\rho + \int_1^\sigma \frac{1}{2}\big|-\widetilde{\xi}_\rho\big|^2\,\mathrm{d}\rho \leq 1 - \frac{1}{2\sigma} \quad \Leftrightarrow \quad \int_1^\sigma \big|\widetilde{\xi}_\rho\big|^2\,\mathrm{d}\rho \leq 1 - \frac{1}{\sigma}.$$

Since $\mathcal{S}_\mathcal{R}(\widetilde{u}_\rho) = 0$ *for* $\rho \geq 1$*, (3.18) is certainly satisfied if we replace the integrand* $\mathcal{R}^*\big(-\widetilde{\xi}_s\big)$ *by* $\mathcal{S}_\mathcal{R}(\widetilde{u}_s)$*; thus the simple De Giorgi estimate (3.15) is clearly satisfied.*

However, recalling that for $\sigma \geq 1$ *we have* $\partial\mathcal{E}(\widetilde{u}_\sigma) = [0,1]$*, a wrong choice of* $\widetilde{\xi}_\sigma \in \partial\mathcal{E}(\widetilde{u}_\sigma)$ *can violate (3.18).*

Finally, we consider the Euler-Lagrange equation (3.7) in the current Banach setting:

$$0 \in \widetilde{\xi}_\sigma + \partial\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) \quad \text{with } \widetilde{\xi}_\sigma \in \partial\mathcal{E}(\widetilde{u}_\sigma).$$

Since $\mathcal{R}$ *is smooth with* $\partial\mathcal{R}(v) = \{v\}$*, there is a unique solution, namely*

$$\widetilde{\xi}_\sigma = -\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ) = \begin{cases} 1 & \text{for } \sigma \in (0,1]\,, \\ 1/\sigma & \text{for } \sigma \geq 1\,. \end{cases}$$

In particular, we have

$$\begin{aligned} \text{for } \sigma \in (0,1]: \quad & \partial_\mathcal{R}\mathcal{E}(\sigma; \widetilde{u}_\sigma) = \{1\} \quad \text{and} \quad \mathcal{S}_\mathcal{R}(\widetilde{u}_\sigma) = \mathcal{C}_\mathcal{R}(\sigma; \widetilde{u}_\sigma) = \frac{1}{2}, \\ \text{for } \sigma > 1: \quad & \partial_\mathcal{R}\mathcal{E}(\sigma; \widetilde{u}_\sigma) = \big\{\frac{1}{\sigma}\big\} \quad \text{and} \quad 0 = \mathcal{S}_\mathcal{R}(\widetilde{u}_\sigma) \lneqq \mathcal{C}_\mathcal{R}(\sigma; \widetilde{u}_\sigma) = \frac{1}{2\sigma^2}\,. \end{aligned}$$

Moreover, equality in (3.18) *can be achieved only by choosing the special selection*

$$\widetilde{\xi}_\sigma \in \partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma) \qquad \textit{for a.a. } \sigma > 0\,.$$

Thus, the 'good selections' of $\partial\mathcal{E}(\widetilde{u}_\sigma)$ for the purpose of De Giorgi identity is prescribed by the Euler-Lagrange equation (3.7)*.*

In summary, we have the simple De Giorgi estimate (3.15) *for $\mathcal{S}_{\mathcal{R}}$ (with strict inequality for $\sigma > 1$), and we have the De Giorgi identity* (3.17) *for $\mathcal{C}_{\mathcal{R}}$.*

3.3 Radially differentiable potentials

This section revolves around a structural property for dissipation potentials that will be crucial for obtaining the De Giorgi identity.

Definition 3.10 *We say that a dissipation potential $\mathcal{R} : X \to [0,\infty]$ is* radially differentiable *if*

$$\forall\, v \in \mathrm{dom}(\mathcal{R}) : \quad (0,1) \ni \lambda \mapsto f(v;\lambda) := \mathcal{R}(\lambda v) \textit{ is differentiable.} \tag{3.19}$$

Observe that, if $\mathrm{dom}(\mathcal{R})$ is *open*, then (3.19) in fact implies that $\lambda \mapsto f(v;\lambda)$ is well-defined and differentiable on the interval $(0, 1+\epsilon)$ for some $\epsilon > 0$.

Clearly, both differentiability on the one hand, and positive homogeneity (i.e., $\mathcal{R}(\lambda v) = \lambda^p \mathcal{R}(v)$ for some $p \geq 1$) on the other, are sufficient conditions for (3.19). Thus, linear combinations of convex differentiable, or positively homogeneous, potentials are radially differentiable.

We are now going to show that radial differentiability is equivalent to another structural property that was used in [MRS13] to prove the De Giorgi estimate (3.16).

Proposition 3.11 *A dissipation potential $\mathcal{R}$ is radially differentiable if and only if*

$$\xi_1, \xi_2 \in \partial\mathcal{R}(v) \quad \Longrightarrow \quad \mathcal{R}^*(\xi_1) = \mathcal{R}^*(\xi_2). \tag{3.20}$$

In view of the well-known convex-analysis relation

$$\mathcal{R}(v) + \mathcal{R}^*(\xi) = \langle \xi, v\rangle \qquad \text{for all } \xi \in \partial\mathcal{R}(v),$$

condition (3.20) is, in turn, equivalent to the property that

$$\xi_1, \xi_2 \in \partial\mathcal{R}(v) \quad \Longrightarrow \quad \langle \xi_1, v\rangle = \langle \xi_2, v\rangle.$$

This ensures that the (a priori multivalued) mapping (whose domain is clearly given by $\mathrm{dom}(\partial\mathcal{R})$)

$$\mathfrak{P} : X \rightrightarrows [0,\infty), \qquad \mathfrak{P}(v) := \langle \xi, v\rangle \text{ for all } \xi \in \partial\mathcal{R}(v), \quad \text{ is single-valued} \tag{3.21}$$

which will prove useful in Section 4.2 ahead.

Let us now address the proof of the equivalence of (3.20) and (3.19): We start by observing that for every $v \in \mathrm{dom}(\mathcal{R})$ the mapping $f(v;\cdot)$ is convex. Its subdifferential in the sense of convex analysis $\partial f(v;\cdot) : (0,\infty) \rightrightarrows \mathbb{R}$ is given by

$$\partial f(v;\lambda) = [f'_-(v;\lambda), f'_+(v;\lambda)] \qquad \text{for all } \lambda > 0,\ v \in \mathrm{dom}(\mathcal{R})\,, \tag{3.22a}$$

where $f'_\pm(v;\cdot)$ are the one-sided derivatives of the mapping $f(v;\cdot)$,

$$\begin{aligned} f'_+(v;\lambda) &:= \lim_{h\to 0^+} \frac{1}{h}\big(f(v;\lambda+h) - f(v;\lambda)\big), \\ f'_-(v;\lambda) &:= \lim_{h\to 0^+} \frac{1}{h}\big(f(v;\lambda) - f(v;\lambda-h)\big)\,, \end{aligned} \tag{3.22b}$$

(the above limits exist by convexity of $f(v;\cdot)$). Then, we have the following result.

Lemma 3.12 *For every $v \in \mathrm{dom}(\mathcal{R})$ and $\lambda > 0$*

$$f'_+(v;\lambda) = \max_{\xi\in\partial\mathcal{R}(\lambda v)} \langle \xi, v\rangle = \mathcal{R}(\lambda v) - \min_{\xi\in\partial\mathcal{R}(\lambda v)} \mathcal{R}^*(\xi), \tag{3.23a}$$

$$f'_-(v;\lambda) = \min_{\xi\in\partial\mathcal{R}(\lambda v)} \langle \xi, v\rangle = \mathcal{R}(\lambda v) - \max_{\xi\in\partial\mathcal{R}(\lambda v)} \mathcal{R}^*(\xi). \tag{3.23b}$$

Proof. It can be easily checked that, for any $v \in \mathrm{dom}(\mathcal{R})$ and $\lambda > 0$, there holds

$$\ell \in \partial f(v;\lambda) \ \text{ if and only if } \ \exists\, \xi \in \partial\mathcal{R}(\lambda v) \text{ s.t. } \ell = \langle \xi, v\rangle\ .$$

Combining this with (3.22a), we immediately deduce (3.23). ■

We are now in a position to carry out the

Proof of Proposition 3.11. The mapping $f(v;\cdot)$ from (3.19) is differentiable at $\lambda = 1$ if and only if $f'_-(v;1) = f'_+(v;1)$. This is in turn equivalent, by (3.23), to the fact that $\min_{\xi\in\partial\mathcal{R}(v)} \mathcal{R}^*(\xi) = \max_{\xi\in\partial\mathcal{R}(v)} \mathcal{R}^*(\xi)$, i.e. (3.19). ■

Remark 3.13 *Let us focus on 'metric-like' dissipation potentials of the form*

$$\mathcal{R}(v) := \psi(\|v\|_X) \qquad \textit{for every } v \in X \tag{3.24}$$

with $\psi : [0,\infty) \to [0,\infty)$ convex with superlinear growth at infinity, as in Definition 2.1. For the associated mapping $f(v;\cdot)$, $v \in X$ (note that $\mathrm{dom}(\mathcal{R}) = X$, in this case), we have

$$\partial f(v;\lambda) = \|v\|_X \partial\psi(\|\lambda v\|_X) \qquad \textit{for all } \lambda > 0\,.$$

Therefore, $\mathcal{R}$ is radially differentiable in the sense of (3.19) if and only if $\partial\psi$ is single-valued. We thus retrieve the smoothness requirement on ψ from Hypothesis 2.5.

Remark 3.14 (Example 3.5 revisited.) *Let us get back to the dissipation potential $\mathcal{R} : \mathrm{L}^p(\Omega) \to [0,\infty)$, $p \geq 1$, from Example 3.5, i.e. $\mathcal{R}(v) = \int_\Omega \mathrm{R}(v(x))\,\mathrm{d}x$, with the dissipation density $\mathrm{R} : \mathbb{R} \to [0,\infty)$ satisfying conditions (3.8) (so that $\mathrm{dom}(\mathcal{R}) = \mathrm{L}^p(\Omega)$). In that setting, for a given $v \in \mathrm{dom}(\partial\mathcal{R})$ we have that*

$$\xi \in \partial\mathcal{R}(v) \quad \Longleftrightarrow \quad \xi(x) \in \partial\mathrm{R}(v(x)) \ \ \textit{for a.a. } x \in \Omega\,.$$

Therefore, it is natural to address the relations between radial differentiability of $\mathcal{R}$ and radial differentiability of R, which is clearly equivalent to differentiability of R in $\mathbb{R}\backslash\{0\}$. We now check that

$$\mathcal{R} \textit{ is radially differentiable } \Longleftrightarrow \mathrm{R} \textit{ is differentiable in } \mathbb{R}\backslash\{0\}. \tag{3.25}$$

Indeed, suppose that $\mathrm{R}:\mathbb{R}\to[0,\infty)$ *is radially differentiable. Then, in view of the characterization provided by Prop. 3.11, for all* $v\in\mathrm{dom}(\partial\mathcal{R})$ *and* $\xi_1,\ \xi_2\in\partial\mathcal{R}(v)$ *we have*

$$\begin{aligned}&\mathrm{R}^*(\xi_1(x))=\mathrm{R}^*(\xi_2(x)) \quad \textit{for a.a. } x\in\Omega \qquad \textit{and thus}\\ &\mathcal{R}^*(\xi_1)=\int_\Omega \mathrm{R}^*(\xi_1(x))\,\mathrm{d}x=\int_\Omega \mathrm{R}^*(\xi_2(x))\,\mathrm{d}x=\mathcal{R}^*(\xi_2)\,,\end{aligned}$$

hence $\mathcal{R}$ *is radially differentiable.*

The converse implication holds since, if R *is not radially differentiable, then there exist* $\hat{v},\hat{\xi}_1,\hat{\xi}_2\in\mathbb{R}$ *such that* $\hat{\xi}_i\in\partial\mathrm{R}(\hat{v})$, $i=1,2$, *and* $\mathrm{R}^*(\hat{\xi}_1)\neq\mathrm{R}^*(\hat{\xi}_2)$. *Then, defining* $v(x)\equiv\hat{v}$ *and* $\xi_i(x)\equiv\hat{\xi}_i$ *for almost all* $x\in\Omega$, *we obtain a triple* $(v,\xi_1,\xi_2)\in\mathrm{L}^p(\Omega)\times\mathrm{L}^{p'}(\Omega)\times\mathrm{L}^{p'}(\Omega)$ *for which* (3.20) *fails to hold.*

For later use, we point out that, if $\mathcal{R}$ *is radially differentiable, then the single-valued mapping* $\mathfrak{P}:\mathrm{L}^p(\Omega)\to[0,\infty)$ *from* (3.21) *is indeed given by*

$$\mathfrak{P}(v)=\int_\Omega \mathrm{P}(v(x))\,\mathrm{d}x \qquad \textit{with } \mathrm{P}(r):=\begin{cases} r\,\mathrm{R}'(r) & \textit{if } r\in\mathbb{R}\backslash\{0\},\\ 0 & \textit{if } r=0\,.\end{cases} \tag{3.26}$$

Ultimately, as a straightforward consequence of Prop. 3.11 we have the following result.

Corollary 3.15 *Suppose that* $\mathcal{R}:X\to[0,\infty)$ *is* radially differentiable. *Then, we have that*

$$\mathcal{C}_{\mathcal{R}}(\sigma;u)=\mathcal{R}^*(-\xi) \qquad \textit{for all } \xi\in\partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma) \textit{ and } u\in J_\sigma\,.$$

In particular, as soon as a (measurable) selection $\sigma\mapsto\widetilde{u}_\sigma\in J_\sigma$ fulfills the De Giorgi estimate/identity, then (3.16)/(3.17) hold with the slope part of the time integrated dissipation, namely $\int_0^\sigma\mathcal{C}_{\mathcal{R}}(\rho;\widetilde{u}_\rho)\,\mathrm{d}\rho$, given by $\int_0^\sigma\mathcal{R}^*(-\widetilde{\xi}_\rho)\,\mathrm{d}\rho$ for *any* (measurable) selection $\sigma\mapsto\widetilde{\xi}_\sigma\in\partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma)$. Furthermore, if the De Giorgi identity holds along an interval $(0,\sigma)$, the corresponding *optimal* integrand $\mathcal{C}_{\mathcal{R}}(\rho;\widetilde{u}_\rho)$ fulfills

$$\mathcal{C}_{\mathcal{R}}(\rho;\widetilde{u}_\rho)=\frac{\mathrm{d}}{\mathrm{d}s}\big(s\mathcal{R}\big(\frac{1}{s}(\widetilde{u}_\rho-u^\circ)\big)\big)|_{s=\rho} \qquad \text{for a.a. } \rho\in(0,\sigma)\,. \tag{3.27}$$

In Theorem 4.2 ahead we will indeed prove the De Giorgi identity under the condition that the dissipation potential is radially differentiable.

3.4 Tools

In this section we collect some preliminary results that will be used in the proof of Theorem 4.2.

Radial derivatives of dissipation potentials

Our first result is in the same spirit of Lemma 3.12, in that it provides some information on the left and right derivatives of the function $g(t)=t\mathcal{R}\big(\frac{1}{t}\,v\big)$, $t>0$, which also features in (3.27).

Lemma 3.16 (Radial derivative) *For fixed* $v \in \mathrm{dom}(\mathcal{R})$ *define the function* $g(t) = t\mathcal{R}\big(\frac{1}{t}\,v\big)$. *Then, g is convex, decreasing, and satisfies* $g \in \mathrm{C}^{\mathrm{Lip}}_{\mathrm{loc}}(]0,\infty[)$. *For all* $t > 0$ *the left derivative* $g'_-(t)$ *and the right derivative* $g'_+(t)$ *exist, are non-decreasing, and satisfy*

$$\begin{aligned} g'_-(t) &:= \lim_{h\to 0+} \frac{1}{h}\big(g(t) - g(t{-}h)\big) = -\max\big\{\, \mathcal{R}^*(\eta) \,\big|\, \eta \in \partial\mathcal{R}(\frac{1}{t}v) \,\big\} \\ &\leq -\min\big\{\, \mathcal{R}^*(\eta) \,\big|\, \eta \in \partial\mathcal{R}(\frac{1}{t}v) \,\big\} = \lim_{h\to 0+} \frac{1}{h}\big(g(t{+}h) - g(t)\big) =: g'_+(t) < 0. \end{aligned}$$

Moreover, $t \mapsto g'_-(t)$ *is continuous from the left and* $t \mapsto g'_+(t)$ *is continuous from the right.*

In particular, if $\mathcal{R}$ is radially differentiable, then g is continuously differentiable with $g'(t) = g'_\pm(t) = -\mathcal{R}^*(\xi)$ for all $\xi \in \partial\mathcal{R}(\frac{1}{t}\,v)$, in accordance with Lemma 3.12.

Fundamental lemma for marginal functions

Recall that $(0,\infty) \ni \sigma \mapsto \phi(\sigma)$ is the marginal function of $(\sigma,u) \mapsto \Phi_\sigma(u)$. Our next result shows that one-sided differentiability of Φ with respect to σ provides bounds on the one-sided derivatives of ϕ, for which the behavior of Φ with respect to u is not really important. In the following result, in fact, we do not use the special form of Φ, but only its left and right differentiability with respect to σ.

Proposition 3.17 (Derivatives of marginal functions) *We have* $\phi \in \mathrm{C}^{\mathrm{Lip}}_{\mathrm{loc}}(]0,\infty[)$ *and for all* $\sigma > 0$ *the following estimates hold*

$$\liminf_{h\to 0^+} \frac{1}{h}\big(\phi(\sigma) - \phi(\sigma{-}h)\big) \geq \sup\big\{\, \mathrm{D}^-_\sigma\Phi_\sigma(w) \,\big|\, w \in J_\sigma \,\big\} =: \delta^-_\sigma, \tag{3.28a}$$

$$\limsup_{h\to 0^+} \frac{1}{h}\big(\phi(\sigma{+}h) - \phi(\sigma)\big) \leq \inf\big\{\, \mathrm{D}^+_\sigma\Phi_\sigma(w) \,\big|\, w \in J_\sigma \,\big\} =: \delta^+_\sigma \tag{3.28b}$$

with $\mathrm{D}^\pm_\sigma\Phi_\sigma(w)$ *the one-sided derivatives of the map* $\sigma \mapsto \Phi_\sigma(w)$.

Proof. Consider $0 < r < \sigma$, then by the marginal property of ϕ we have

$$\Phi_\sigma(\widetilde{u}_r) - \Phi_r(\widetilde{u}_r) \geq \Phi_\sigma(\widetilde{u}_\sigma) - \Phi_r(\widetilde{u}_r) = \phi(\sigma) - \phi(r) \geq \Phi_\sigma(\widetilde{u}_\sigma) - \Phi_r(\widetilde{u}_\sigma).$$

Hence, the local Lipschitz property of ϕ follows from that of $\Phi(\cdot,w)$. Setting $r = t{-}h$ yields

$$\frac{1}{h}\big(\phi(\sigma) - \phi(\sigma{-}h)\big) \geq \frac{1}{h}\big(\Phi_\sigma(\widetilde{u}_\sigma) - \Phi_{\sigma-h}(\widetilde{u}_\sigma)\big).$$

Taking the liminf for $h \to 0^+$ first and then the supremum over $\widetilde{u}_\sigma \in J_\sigma$ gives (3.28a).

Similarly, we can replace (r,σ) by $(\sigma,\sigma{+}h)$ to obtain

$$\frac{1}{h}\big(\phi(\sigma{+}h) - \phi(\sigma)\big) \leq \frac{1}{h}\big(\Phi_{\sigma+h}(\widetilde{u}_\sigma) - \Phi_\sigma(\widetilde{u}_\sigma)\big).$$

Taking the limsup $h \to 0^+$ first and then the infimum over $\widetilde{u}_\sigma \in J_\sigma$ yields (3.28b). ∎

4 The Banach case: results

In the upcoming Section 4.1 we will show that the De Giorgi estimate previously proved in [MRS13, Lemma 6.1] in the radially differentiable case, in fact improves to an equality. Section 4.2 will address the regularity of the variational interpolant. Then, in Sec. 4.3 we will drop the radial differentiabiilty condition and extend the validity of the De Giorgi estimate to general dissipation potentials.

4.1 Equality in the De Giorgi estimate for radially differentiable potentials

Now we return to the variational integrand where Φ is given in the form

$$\Phi_\sigma(u) = \sigma\,\mathcal{R}\big(\frac{1}{\sigma}(u{-}u^\circ)\big) + \mathcal{E}(u).$$

In particular, the left and right derivatives of $\sigma \mapsto \Phi_\sigma(u)$ exist, see Lemma 3.16, and are independent of the energy $\mathcal{E}$. Moreover, the one-sided derivatives are ordered such that $\mathrm{D}^-_\sigma\Phi(\sigma,w) \le \mathrm{D}^+_\sigma\Phi(\sigma,w) < 0$ for every $w \in J_\sigma$.

However, because of the supremum over $\mathrm{D}^-_\sigma\Phi$ and the infimum over $\mathrm{D}^+_\sigma\Phi$, in general one cannot expect to generate one chain of inequalities from the two estimates in (3.28). Even if J_σ is single-valued, one still arrived at the wrong estimate.

The only case where the marginal estimates (3.28) are useful to prove the De Giorgi identity, is exactly when $\mathcal{R}$ is radially differentiable, cf. (3.19). Then $\mathrm{D}^-_\sigma\Phi = \mathrm{D}^+_\sigma\Phi$ implies $\delta^-_\sigma \ge \delta^+_\sigma$ and (3.28) leads to

$$\liminf_{h\to 0^+} \frac{1}{h}\big(\phi(\sigma) - \phi(\sigma{-}h)\big) \ge \delta^-_\sigma \ge \delta^+_\sigma \ge \limsup_{h\to 0^+} \frac{1}{h}\big(\phi(\sigma{+}h) - \phi(\sigma)\big).$$

From this, the De Giorgi identity follows easily.

We will also address the validity of the identity involving a measurable selection $\widetilde{\xi}_\sigma \in \partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma)$ (we will often refer to $\widetilde{\xi}$ as a *force selection*). For the existence of such a selection (see Lemma 4.3), it is useful to impose the following closedness condition for the conditioned subdifferential $\partial_{\mathcal{R}}\mathcal{E} : (0,\infty){\times}X \rightrightarrows X^*$.

Hypothesis 4.1 (Closedness of $\partial_{\mathcal{R}}\mathcal{E}$) *The conditioned subdifferential $\partial_{\mathcal{R}}\mathcal{E} : (0,\infty){\times}X \rightrightarrows X^*$ is closed on energy sublevels, i.e.*

$$\forall\, E > 0 : \left\{ \begin{array}{l} (\sigma_n, u_n, \xi_n) \rightharpoonup (\sigma, u, \xi) \;\; \text{in } (0,\infty){\times}X{\times}X^*, \\ u_n \in S_E \text{ and } \xi_n \in \partial_{\mathcal{R}}\mathcal{E}(\sigma_n; u_n) \text{ for all } n \in \mathbb{N} \end{array} \right\} \implies \xi \in \partial_{\mathcal{R}}\mathcal{E}(\sigma; u)\,. \tag{4.1}$$

Theorem 4.2 (The De Giorgi identity with radial differentiability) *Consider the* gBGS $(X, \mathcal{E}, \mathcal{R})$ *satisfying Hypothesis 3.1. Fix $u^\circ \in \mathrm{dom}(\mathcal{E})$ and assume that $\mathcal{R}$ is radially differentiable, i.e.* (3.20) *or equivalently* (3.19) *holds. Then, every measurable variational interpolant* $(0,\infty) \ni \sigma \mapsto \widetilde{u}_\sigma \in J_\sigma$ *fulfills the* De Giorgi identity

$$\mathcal{E}(\widetilde{u}_\sigma) + \sigma\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma{-}u^\circ)\big) + \int_0^\sigma \mathcal{C}_{\mathcal{R}}(\rho;\widetilde{u}_\rho)\,\mathrm{d}\rho = \mathcal{E}(u^\circ)\,. \tag{4.2a}$$

If additionally Hypothesis 4.1 holds, then there exists a measurable force selection $\sigma \mapsto \widetilde{\xi}_\sigma \in \partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma)$ such that the improved De Giorgi identity with force selection *holds:*

$$\forall\,\sigma > 0 : \quad \mathcal{E}(\widetilde{u}_\sigma) + \sigma\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma{-}u^\circ)\big) + \int_0^\sigma \mathcal{R}^*(-\widetilde{\xi}_\rho)\,\mathrm{d}\rho = \mathcal{E}(u^\circ). \tag{4.2b}$$

Proof. We fix a small $r > 0$ and consider the marginal function $[r, \sigma] \ni \rho \mapsto \phi(\rho)$. According to Proposition 3.17 ϕ is Lipschitz and hence differentiable almost everywhere in $[r, \sigma]$. This implies that ϕ'_- and ϕ'_+ exist and coincide almost everywhere.

We now fix any measurable variational interpolant $\sigma \mapsto \widetilde{u}_\sigma$. Proposition 3.17 and Lemma 3.16 imply, via the radial differentiability of $\mathcal{R}$, that for almost all $\sigma \in (0, \infty)$ there holds

$$\phi'(\sigma) = -\mathcal{R}^*(\eta) \qquad \text{for all } \eta \in \partial\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big)\,.$$

Thus,

$$\phi'(\sigma) = -\mathcal{C}_\mathcal{R}(\sigma; \widetilde{u}_\sigma) \quad \text{for a.a.}\, \sigma \in (0, \infty)\,.$$

Hence, $\phi(\sigma) = \phi(r) + \int_r^\sigma \phi'(\rho)\,\mathrm{d}\rho$ can be rewritten as

$$\phi(\sigma) = \mathcal{E}(\widetilde{u}_\sigma) + \sigma\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) = \phi(u^\circ; r) - \int_r^\sigma \mathcal{C}_\mathcal{R}(\rho; \widetilde{u}_\rho)\,\mathrm{d}\rho\,.$$

The superlinearity of $\mathcal{R}$ implies $\phi(r) \to \mathcal{E}(u^\circ)$ as $r \to 0^+$, see (3.6). Therefore, taking the limit $r \to 0^+$ gives the desired De Giorgi identity (4.2a).

The last part of the assertion immediately follows from the upcoming Lemma 4.3. ■

The following result collects two straightforward consequences of Hypothesis 4.1.

Lemma 4.3 *Let the* gBGS $(X, \mathcal{E}, \mathcal{R})$ *satisfy Hypotheses 3.1 and 4.1. Then,*

1. *The* infimum *in the definition of $\partial_\mathcal{R}\mathcal{E}$ is attained, i.e.*

$$\mathcal{C}_\mathcal{R}(\sigma; u) := \min\big\{\, \mathcal{R}^*(-\xi) \,\big|\, \xi \in \partial_\mathcal{R}\mathcal{E}(\sigma; u) \,\big\} \quad \textit{for all } u \in J_\sigma; \tag{4.3}$$

2. *For every measurable selection $(0, \infty) \ni \sigma \mapsto \widetilde{u}_\sigma \in J_\sigma$ lying in some energy sublevel S_E, there exists a measurable selection*

$$\sigma \mapsto \widetilde{\xi}_\sigma \in \mathfrak{A}_\mathcal{R}(\sigma, \widetilde{u}_\sigma) := \operatorname{argmin}\big\{\, \mathcal{R}^*(-\xi) \,\big|\, \xi \in \partial_\mathcal{R}\mathcal{E}(\sigma; \widetilde{u}_\sigma) \,\big\} \tag{4.4a}$$

such that

$$\mathcal{C}_\mathcal{R}(\sigma; \widetilde{u}_\sigma) = \mathcal{R}^*(-\widetilde{\xi}_\sigma) \quad \textit{for all } \sigma > 0\,. \tag{4.4b}$$

Proof. Property (4.3) is easily checked via the *direct method* by relying on (4.1), with a similar argument as for (3.11a). For (4.4b), observe that, as a consequence of Hypothesis 4.1, the multivalued mapping

$$\mathfrak{A}_\mathcal{R} : (0, \infty) \times X \rightrightarrows X^*; \quad \mathfrak{A}_\mathcal{R}(\sigma, u) := \operatorname{argmin}\big\{\, \mathcal{R}^*(-\xi) \,\big|\, \xi \in \partial_\mathcal{R}\mathcal{E}(\sigma; u) \,\big\} \tag{4.5}$$

is upper semicontinuous with respect to convergence in $\mathbb{R}$ for σ and weak convergence for u. Therefore, for every measurable selection $(0, \infty) \ni \sigma \mapsto \widetilde{u}_\sigma \in J_\sigma$ the multivalued mapping $\sigma \mapsto \mathfrak{A}_\mathcal{R}(\sigma, \widetilde{u}_\sigma)$ is measurable, as it is given by the composition of measurable mappings. Then, [CaV77, Thm. 3.22] grants the existence of a *measurable* selection $\sigma \mapsto \widetilde{\xi}_\sigma \in \mathfrak{A}_\mathcal{R}(\sigma, \widetilde{u}_\sigma)$. ■

We conclude this section with an example in which $\mathcal{R}$ is not radially differentiable but the De Giorgi identity still holds.

Example 4.4 We consider the gBGS $(X, \mathcal{E}, \mathcal{R})$ with

$$X = \mathbb{R}, \quad \mathcal{E}(u) = \frac{1}{2}\,u^2, \quad \mathcal{R}(v) = \begin{cases} \frac{1}{2}\,v^2 & \text{for } |v| \le 1, \\ 2v^2 - \frac{3}{2} & \text{for } |v| \ge 1. \end{cases}$$

Clearly, $\mathcal{R}$ is not radially differentiable, hence (3.20) does not hold.

For $u^\circ = 6$ we obtain the unique variational interpolant

$$\widetilde{u}_\sigma = \begin{cases} \frac{24}{4+\sigma} & \text{for } \sigma \in [0,2], \\ 6-\sigma & \text{for } \sigma \in [2,5], \\ \frac{6}{1+\sigma} & \text{for } \sigma \ge 5. \end{cases}$$

It can easily be checked that identity (3.16) holds for all $\sigma > 0$, where the choice $\widetilde{\xi}_\sigma = \mathrm{D}\mathcal{E}(\widetilde{u}_\sigma) = \widetilde{u}_\sigma$ is mandatory because the Fréchet subdifferential of $\mathcal{E}$ is single-valued. For $\sigma \in [2,5]$ we have $-\widetilde{\xi}(\sigma) = -6{+}\sigma \in \partial\mathcal{R}(-1) = [-4,-1]$, where we used $\frac{1}{\sigma}(\widetilde{u}_\sigma{-}6) = -1$.

Recalling the definitions $\delta_\sigma^\pm$ from (3.28) we obtain the strict inequalities $\delta_\sigma^- = -\mathcal{R}^*(-4) = -\frac{7}{2} \lneqq \partial_\sigma\phi(\sigma) = \sigma{-}\frac{11}{2} \lneqq -\frac{1}{2} = -\mathcal{R}^*(-1) = \delta_\sigma^+$ for $\sigma \in (2,5)$.

4.2 Regularity of the variational interpolant

In this section we take a slight detour from the main theme of the paper and provide some sufficient conditions for gaining extra time regularity of the variational interpolant $\sigma \mapsto \widetilde{u}_\sigma$. This deviation will only be useful if we employ a slightly strengthened version of radial differentiability.

Furthermore, on the one hand, we will require uniform convexity of the mapping $u \mapsto \Phi_\sigma(u)$, a sufficient condition for which is, of course, uniform convexity of the energy functional $\mathcal{E}$. On the other hand, we will need to reinforce radial differentiability. Indeed, while the latter property is equivalent to the fact that the composed, a priori multivalued, mapping $\mathcal{R}^*{\circ}\partial\mathcal{R} : X \rightrightarrows X^*$ is single-valued (cf. Proposition 3.11), we will now further require that $\mathcal{R}^*{\circ}\partial\mathcal{R}$ is Lipschitz continuous (on bounded subsets of X).

We collect these conditions in the following

Hypothesis 4.5 *We assume that*

1. *There exist $\overline{\lambda} > 0$ and $\sigma_* > 0$ such that for all $\sigma > 0$ the mapping $u \mapsto \Phi_\sigma(u)$ is $\overline{\lambda}$-convex, namely*

$$\begin{aligned} &\exists\, \overline{\lambda} > 0 \;\; \forall\, \sigma > 0\,,\; \forall\, u_0, u_1 \in X\,,\; \forall\, \theta \in [0,1]\;: \\ &\Phi_\sigma((1{-}\theta)u_0{+}\theta u_1) \le (1{-}\theta)\Phi_\sigma(u_0) + \theta\Phi_\sigma(u_1) - \frac{\overline{\lambda}}{2}\theta(1{-}\theta)\|u_0{-}u_1\|^2\,. \end{aligned} \tag{4.6}$$

2. *The mapping $\mathcal{R}^*{\circ}\partial\mathcal{R} : X \rightrightarrows X^*$ is Lipschitz continuous on bounded sets, i.e.,*

$$\begin{gathered} \forall\, M > 0\; \exists\, C_M > 0\; \forall\, (v_1, \xi_1),\, (v_2, \xi_2) \in X{\times}X^*,\; \max_{i=1,2}\|v_i\| \le M,\; \xi_i \in \partial\mathcal{R}(v_i)\;: \\ |\mathcal{R}^*(\xi_1){-}\mathcal{R}^*(\xi_2)| \le C_M\|u_1{-}u_2\|\,. \end{gathered} \tag{4.7}$$

Before stating our result, let us pin down two key consequences of Hypothesis 4.5:

- It follows from (4.6) (cf., e.g., [MRS13, Prop. 2.4]) that the Fréchet subdifferential $\partial\Phi_\sigma(\cdot) : X \rightrightarrows X^*$ can be characterized in terms of the following *global* estimate: for all $\sigma > 0$ and for every $u \in \mathrm{dom}(\partial\Phi_\sigma(\cdot))$ we have that

$$\begin{aligned} &\omega \in \partial\Phi_\sigma(u) \text{ if and only if} \\ &\Phi_\sigma(v) - \Phi_\sigma(u) \geq \langle \omega, v-u \rangle + \frac{\overline{\lambda}}{2}\|v-u\|^2 \quad \text{for all } v \in X\,. \end{aligned} \tag{4.8}$$

- By Prop. 3.11, (4.7) in particular implies that for every $M > 0$ the restriction of $\mathcal{R}$ to the ball $\overline{B}_M(0)$ is radially differentiable. Therefore, recalling Lemma 3.12 we have that for every $v \in \overline{B}_M(0)$ the function

$$\begin{aligned} &g(v;\rho) = \rho\mathcal{R}\left(\frac{1}{\rho}v\right) \text{ is differentiable at every } \rho > 0, \text{ with} \\ &\frac{\mathrm{d}}{\mathrm{d}\rho}g(v;\rho) = \mathcal{R}\left(\frac{1}{\rho}v\right) - \langle \eta, \frac{1}{\rho}v\rangle = \mathcal{R}^*(\eta) \qquad \text{for all } \eta \in \partial\mathcal{R}\left(\frac{1}{\rho}v\right)\,. \end{aligned} \tag{4.9}$$

We are now in a position to state the main result of this section.

Theorem 4.6 *Let the* gBGS $(X, \mathcal{E}, \mathcal{R})$ *satisfy Hypotheses 3.1 and 4.5; and let* $u^\circ \in \mathrm{dom}(\mathcal{E})$ *be fixed. Then, for all measurable selection* $(0,\infty) \ni \sigma \mapsto \widetilde{u}_\sigma \in J_\sigma$ *we have* $\widetilde{u} \in \mathrm{Lip}_{\mathrm{loc}}(]0,\infty[;X)$.

Proof. Preliminarily, we observe that, because of $\sup_{\sigma>0}\sigma\mathcal{R}((\widetilde{u}_\sigma - u^\circ)/\sigma) \leq C$, we have

$$\exists\, M > 0 \; \forall\, \sigma \in (0,\infty): \quad \|\widetilde{u}_\sigma\| \leq M. \tag{4.10}$$

Let us fix $\sigma_\# > 0$ and let $[\sigma_1, \sigma_2] \subset [\sigma_\#, \infty)$ be arbitrary. We apply (4.8) with $\sigma = \sigma_1$, $u = \widetilde{u}_{\sigma_1}$, $v = \widetilde{u}_{\sigma_2}$ and $\omega = 0$ as, indeed, $0 \in \partial\Phi_{\sigma_1}(\widetilde{u}_{\sigma_1})$ is the Euler-Lagrange equation for $\widetilde{u}_{\sigma_1} \in J_{\sigma_1}(u^\circ)$. Thus, we obtain

$$\begin{aligned} \frac{\overline{\lambda}}{2}\|\widetilde{u}_{\sigma_2} - \widetilde{u}_{\sigma_1}\|^2 &\leq \Phi_{\sigma_1}(\widetilde{u}_{\sigma_2}) - \Phi_{\sigma_1}(\widetilde{u}_{\sigma_1}) \\ &= \mathcal{E}(\widetilde{u}_{\sigma_2}) - \mathcal{E}(\widetilde{u}_{\sigma_1}) + \sigma_1\mathcal{R}\left(\frac{1}{\sigma_1}(\widetilde{u}_{\sigma_2} - u^\circ)\right) - \sigma_1\mathcal{R}\left(\frac{1}{\sigma_1}(\widetilde{u}_{\sigma_1} - u^\circ)\right)\,. \end{aligned}$$

Interchanging σ_1 and σ_2 and adding the inequalities gives

$$\begin{aligned} &\overline{\lambda}\|\widetilde{u}_{\sigma_2} - \widetilde{u}_{\sigma_1}\|^2 \\ &\leq \sigma_2\mathcal{R}\big(\frac{1}{\sigma_2}(\widetilde{u}_{\sigma_1} - u^\circ)\big) - \sigma_1\mathcal{R}\big(\frac{1}{\sigma_1}(\widetilde{u}_{\sigma_1} - u^\circ)\big) + \sigma_1\mathcal{R}\big(\frac{1}{\sigma_1}(\widetilde{u}_{\sigma_2} - u^\circ)\big) - \sigma_2\mathcal{R}\big(\frac{1}{\sigma_2}(\widetilde{u}_{\sigma_2} - u^\circ)\big) \\ &= \int_{\sigma_1}^{\sigma_2}\left[\frac{\mathrm{d}}{\mathrm{d}\rho}g(\widetilde{u}_{\sigma_1};\rho) - \frac{\mathrm{d}}{\mathrm{d}\rho}g(\widetilde{u}_{\sigma_2};\rho)\right]\mathrm{d}\rho \stackrel{(4.9)}{=} \int_{\sigma_1}^{\sigma_2}\left(\mathcal{R}^*(\eta_1^\rho) - \mathcal{R}^*(\eta_2^\rho)\right)\mathrm{d}\rho \end{aligned}$$

for every $\eta_i^\rho \in \partial\mathcal{R}(\frac{1}{\rho}\widetilde{u}_{\sigma_i})$, $i = 1, 2$. Now using the Lipschitz property, we find

$$\overline{\lambda}\|\widetilde{u}_{\sigma_2} - \widetilde{u}_{\sigma_1}\|^2 \leq \int_{\sigma_1}^{\sigma_2} C_M\|\tfrac{1}{\rho}\widetilde{u}_{\sigma_2} - \tfrac{1}{\rho}\widetilde{u}_{\sigma_1}\|\,\mathrm{d}\rho$$

due to (4.7). All in all, using that $\sigma_1 \geq \sigma_\#$ we conclude that

$$\overline{\lambda}\|\widetilde{u}_{\sigma_2} - \widetilde{u}_{\sigma_1}\|^2 \leq \frac{C_M}{\sigma_\#}(\sigma_2 - \sigma_1)\|\widetilde{u}_{\sigma_2} - \widetilde{u}_{\sigma_1}\|\,, \tag{4.11}$$

hence $\sigma \mapsto \widetilde{u}_\sigma$ is Lipschitz continuous on $[\sigma_\#, \infty)$. By the arbitrariness of $\sigma_\#$, we conclude $\widetilde{u} \in \mathrm{Lip}_{\mathrm{loc}}(]0,\infty[;X)$. ■

More on condition (4.7)**.**

We conclude this section by gaining further insight into property (4.7). The key observation is that it is implied by radial differentiability and superlinearity of $\mathcal{R}^*$.

Lemma 4.7 *Let $\mathcal{R} : X \to [0,\infty)$ satisfy Hypothesis 3.1, be radially differentiable,* and *suppose that $\mathcal{R}^*$ has superlinear growth at infinity, i,e.,*

$$\lim_{\|\xi\|_* \to \infty} \frac{\mathcal{R}^*(\xi)}{\|\xi\|_*} = \infty \,. \tag{4.12}$$

Then, $\mathcal{R}^* \circ \partial\mathcal{R} : X \to X^*$ fulfills (4.7)*.*

Proof. It follows from (4.12) that the subdifferential $\partial\mathcal{R} : X \rightrightarrows X^*$ is a bounded operator, cf. [Bre73, Prop. 2.14]. Then, (4.7) immediately follows, taking into account that $\mathcal{R}^*$ is itself Lipschitz on bounded sets since $\mathcal{R}$ has superlinear growth at infinity. ∎

We also remark that, if $\mathcal{R}^*$ is superlinear, then property (4.7) is equivalent to the Lipschitz continuity on bounded sets of the mapping $\mathfrak{P}$ from (3.21). This follows from the fact that $\mathfrak{P}(u) = \mathcal{R}(u) + (\mathcal{R}^* \circ \partial\mathcal{R})(u)$, and $\mathcal{R}$ is Lipschitz continuous on bounded sets. Exploiting this observation, we can readily prove that property (4.7) is stable under the sum of dissipation potentials.

Corollary 4.8 *Let $\mathcal{R}_i : X \to [0,\infty)$, $i = 1,2$, fulfill Hypothesis 3.1 be radially differentiable, and satisfy* (4.12)*. Then, $\mathcal{R}_1 + \mathcal{R}_2$ satisfies* (4.7)*.*

Proof. It suffices to remark (with slight abuse of notation) that

$$\mathfrak{P}_{\mathcal{R}_1+\mathcal{R}_2}(v) = \langle \partial(\mathcal{R}_1+\mathcal{R}_2)(v), v\rangle = \langle \partial\mathcal{R}_1(v)+\partial\mathcal{R}_2(v), v\rangle = \mathfrak{P}_{\mathcal{R}_1}(v)+\mathfrak{P}_{\mathcal{R}_2}(v) \quad \text{for all } v \in X\,,$$

where the validity of the sum rule $\partial(\mathcal{R}_1+\mathcal{R}_2)(v) = \partial\mathcal{R}_1(v)+\partial\mathcal{R}_2(v)$ is guaranteed, under the present assumptions, by, e.g., [IoT79, Thm. 1, p. 211]. ∎

Ultimately, p-homogeneous dissipation potentials provide examples for the validity of (4.7).

Corollary 4.9 *Let $\mathcal{R} : X \to [0,\infty)$ be positively homogeneous of degree p for some $p > 1$. Then,*

$$\mathfrak{P}(v) = p\mathcal{R}(v) \qquad \text{for all } v \in X \tag{4.13}$$

and $\mathcal{R}^* \circ \partial\mathcal{R}$ fulfills (4.7)*.*

Proof. First of all, we show (4.13) (which, in fact, also holds for $p = 1$). Indeed, it suffices to observe that, by definition of $\partial\mathcal{R}(v)$ and p-homogeneity of $\mathcal{R}$, we have for all $\xi \in \partial\mathcal{R}(v)$

$$(\lambda^p - 1)\mathcal{R}(v) = \mathcal{R}(\lambda v) - \mathcal{R}(v) \geq \langle \xi, \lambda v - v\rangle = (\lambda - 1)\mathfrak{P}(v)\,.$$

Then, we divide the above estimate by $(\lambda-1)$ for all $\lambda > 1$, and take the limit as $\lambda \to 1^+$, thus obtaining $p\mathcal{R}(v) \geq \mathfrak{P}(v)$. The converse inequality follows by dividing by $(\lambda-1)$ for all $0 < \lambda < 1$ and sending $\lambda \to 1^-$.

From (4.13) and the fact that $\mathcal{R}$ is Lipschitz continuous on bounded sets, we deduce that $\mathfrak{P}$ satisfies (4.7). ∎

Eventually, by Corollary 4.8 property (4.7) holds for linear combinations of homogeneous potentials of possibly different degrees, as well.

Example 4.10 (Example 3.5 re-revisited) *Getting back to Example 3.5, let us additionally assume that* R *is differentiable on* $\mathbb{R}\backslash\{0\}$*: then the dissipation potential* $\mathcal{R}(v) = \int_\Omega \mathrm{R}(v(x))\,\mathrm{d}x$ *is radially differentiable by Remark 3.14. Since we have required* R^* *to have* p'*-growth with* $p' > 1$*,* $\mathcal{R}^*$ *has superlinear growth at infinity, and then Lemma 4.7 applies, ensuring that property* (4.7) *holds.*

4.3 De Giorgi estimates for general dissipation potentials

We now return to the general case in which $\mathcal{R}$ is allowed to take the value ∞ with an open domain $\mathrm{dom}(\mathcal{R})$ in X (cf. Hypothesis 3.1). We will show that, if we drop the radial differentiability condition (3.20) on $\mathcal{R}$, the De Giorgi simple and improved estimates can be still retrieved for gBGS $(X, \mathcal{E}, \mathcal{R})$ featuring general dissipation potentials $\mathcal{R}$, see Theorem 4.12 ahead.

Our strategy for proving both results consists in replacing $\mathcal{R}$ by the Yosida approximations $(\mathcal{R}_\eta)_\eta$, where we use a suitable norm on the reflexive Banach space X such that all $\mathcal{R}_\eta$ are differentiable. Thus, for all $\eta > 0$ we find measurable variational interpolants

$$\sigma \mapsto \widetilde{u}^\eta_\sigma \in \boldsymbol{J}_\eta(\sigma) := \mathrm{Argmin}\Big\{ \sigma\mathcal{R}_\eta\big(\tfrac{1}{\sigma}(u{-}u^\circ)\big) + \mathcal{E}(u) \;\Big|\; u \in X \Big\} \tag{4.14}$$

satisfying the De Giorgi identity for the corresponding gBGS $(X, \mathcal{E}, \mathcal{R}_\eta)$. We will then study the limit $\eta \to 0^+$.

A related approach based on Yosida regularization was developed in [Bac21, Chap. 3] in order to extend the existence result for gBGS provided in [MRS13], *without* the radial differentiability condition that was adopted therein (recall the discussion prior to Prop. 3.11). In [Bac21] the limit $\eta \to 0^+$ was performed on the level of the solutions of the evolution equation.

Here, we will resort to Yosida approximation to get rid of radial differentiability for the very proof of the De Giorgi estimate. Hence, the limiting process $\eta \to 0^+$ will be more delicate than in [Bac21], because variational interpolants are not absolutely continuous in $[0, \sigma_0]$, in contrast to solutions of the regularized gBGS. Nevertheless, starting from our approximations $(\sigma \mapsto \widetilde{u}^\eta_\sigma)_{\eta>0}$ we will be able to find a limiting function $\sigma \mapsto \widetilde{u}_\sigma \in X$ that such that

$$\forall\, \sigma \in (0, \sigma_*)\ \exists\, (\eta^\sigma_k)_{k\in\mathbb{N}} : \quad \eta^\sigma_k \to 0 \ \text{ and } \ \widetilde{u}^{\eta^\sigma_k}_\sigma \rightharpoonup \widetilde{u}_\sigma \quad \text{for } k \to \infty,$$

see Proposition 4.15. This pointwise weak convergence along a σ-dependent subsequence will the crucial point to pass to the liminf, in the De Giorgi identity for $(X, \mathcal{E}, \mathcal{R}_\eta)$, to derive the De Giorgi estimates, also based on the Mosco-convergence as $\eta \to 0^+$ of the functionals

$$\Phi^\eta_\sigma(u) := \sigma\mathcal{R}_\eta\big(\tfrac{1}{\sigma}(u{-}u^\circ)\big) + \mathcal{E}(u) \tag{4.15}$$

involved in (4.14), to the functional Φ_σ for fixed $\sigma > 0$, which means

$$\begin{cases} \forall\, (u_\eta)_\eta, u \in X \ :\ u_\eta \rightharpoonup u \ \Rightarrow \ \liminf_{\eta\to 0^+} \Phi^\eta_\sigma(u_\eta) \geq \Phi_\sigma(u), \\ \forall\, u \in X\ \exists\, (\hat{u}_\eta)_\eta \ : \hat{u}_\eta \to u \text{ and } \limsup_{\eta\to 0^+} \Phi^\eta_\sigma(\hat{u}_\eta) \leq \Phi_\sigma(u). \end{cases} \tag{4.16}$$

This will follow easily from the Mosco-convergence of the potentials $(\mathcal{R}_\eta)_\eta$ stated in the upcoming Lemma 4.13.

For the proof of the simple De Giorgi estimate it will be sufficient to invoke the lower semicontinuity of the $\mathcal{R}$-slope of $\mathcal{E}$, namely $\mathcal{S}_\mathcal{R}$, in the limit passage in the De Giorgi identity holding for the regularized potentials. To obtain the improved the De Giorgi estimate, we will have to impose an additional condition on the lower semicontinuity of the conditioned $\mathcal{R}$-slope, namely $\mathcal{C}_{\mathcal{R}}$:

Hypothesis 4.11 (Qualified lower semicontinuity of $\mathcal{C}_{\mathcal{R}}$) *Fix $\sigma \in (0,\sigma_*)$ and $u^\circ \in \mathrm{dom}(\mathcal{E})$ and consider $u_\eta \in \boldsymbol{J}_\eta(\sigma)$ for $\eta \in]0,1[$. Then,*

$$u_\eta \rightharpoonup u \quad \Longrightarrow \quad \liminf_{\eta\to 0} \mathcal{C}_{\mathcal{R}_\eta}(\sigma;u_\eta) \geq \mathcal{C}_{\mathcal{R}}(\sigma;u). \tag{4.17}$$

We will discuss sufficient conditions for this hypothesis after stating the main result of this subsection. We emphasize that the following result is weaker than the De Giorgi identity in two ways. First we only have the inequalities instead of the equality, and second we only prove the existence of *at least one* measurable interpolant, whereas the De Giorgi identity holds for *all* measurable interpolants with $\widetilde{u}_\sigma \in \boldsymbol{J}(\sigma)$. So far, we are not able to show that all measurable interpolants have this property.

Let us also mention in advance that, mirroring the statement of Theorem 4.2, under the additional Hypothesis 4.1, we are in a position to prove the existence of a measurable force selection for which the improved De Giorgi estimate holds. An analogous statement could be given for the simple De Giorgi estimate; we choose to omit it to avoid overburdening the exposition.

Theorem 4.12 (De Giorgi estimates for general $\mathcal{R}$) *Let the* gBGS $(X,\mathcal{E},\mathcal{R})$ *satisfy Hypothesis 3.1, and let $\sigma_*>0$. Then, for every $u^\circ \in \mathrm{dom}(\mathcal{E})$ there exists a measurable variational interpolant $\sigma \mapsto \widetilde{u}_\sigma \in J_\sigma \subset X$ fulfilling the* simple De Giorgi estimate

$$\forall\, \sigma \in (0,\sigma_*): \quad \mathcal{E}(\widetilde{u}_\sigma) + \sigma\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) + \int_0^\sigma \mathcal{S}_{\mathcal{R}}(\widetilde{u}_\sigma)\,\mathrm{d}\rho \leq \mathcal{E}(u^\circ). \tag{4.18}$$

f Hypothesis 4.11 holds, then this interpolant also satisfies the improved De Giorgi estimate

$$\forall\, \sigma \in (0,\sigma_*): \quad \mathcal{E}(\widetilde{u}_\sigma) + \sigma\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) + \int_0^\sigma \mathcal{C}_{\mathcal{R}}(\sigma;\widetilde{u}_\sigma)\,\mathrm{d}\rho \leq \mathcal{E}(u^\circ). \tag{4.19a}$$

Thus, under the additional Hypothesis 4.1 there exists a measurable force selection $\sigma \mapsto \widetilde{\xi}_\sigma \in \partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma)$ satisfying the improved De Giorgi estimate with force selection

$$\forall\, \sigma > 0: \quad \mathcal{E}(\widetilde{u}_\sigma) + \sigma\mathcal{R}\big(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\big) + \int_0^\sigma \mathcal{R}^*(-\widetilde{\xi}_\rho)\,\mathrm{d}\rho \leq \mathcal{E}(u^\circ)\,. \tag{4.19b}$$

We postpone the proof of this main result to Section 4.4.

Prior to discussing the rather abstract lower semicontinuity property (4.17) on the conditioned slope $\mathcal{C}_{\mathcal{R}}$, we need to pin down some of the properties of the Moreau-Yosida approximations $(\mathcal{R}_\eta)_{\eta>0}$. First of all, let us settle their definition: Since X is reflexive, Asplund's renorming theorem [Asp68], see also [BaP12, Thm. 1.105], ensures that there exists an equivalent norm on X (still denoted by $\|\cdot\|_X$), such that, correspondingly, both X and X^* are strictly convex (namely, the mappings $x \mapsto \|x\|_X^2$ and $\xi \mapsto \|\xi\|_{X^*}^2$ are strictly convex). With this, the Moreau-Yosida approximations $(\mathcal{R}_\eta)_{\eta>0}$ of $\mathcal{R}$ are defined via

$$\mathcal{R}_\eta\colon X \to [0,\infty) \qquad \mathcal{R}_\eta(v) := \inf_{w\in X}\left(\frac{\|w-v\|_X^2}{2\eta} + \mathcal{R}(w)\right).$$

The following result collects the properties of the functionals $(\mathcal{R}_\eta)_{\eta>0}$ we are going to use.

Lemma 4.13 *For all $\eta > 0$ the functional $\mathcal{R}_\eta$ is differentiable on X and has the convex conjugate $\mathcal{R}_\eta^*(\xi) = \mathcal{R}^*(\xi) + \frac{\eta}{2}\|\xi\|^2_{X^*}$ for $\xi \in X^*$. Moreover,*

1. *the family $\big(\mathcal{R}_\eta\big)_{\eta\in(0,1]}$ is equi-coercive, i.e.*

$$\forall\, S>0 \;\; \exists\, C_S>0 \;\; \forall\, \eta \in (0,1] \;\; \forall\, v \in X: \quad \mathcal{R}_\eta(v) \le S \implies \|v\| \le C_S\,; \tag{4.20}$$

2. *the family $(\mathcal{R}_\eta)_\eta$* Mosco*-converge to $\mathcal{R}$ as $\eta \downarrow 0$; in particular, we have*

$$v_\eta \rightharpoonup v \implies \mathcal{R}(v) \le \liminf_{\eta\to 0} \mathcal{R}_\eta(v_\eta) \quad \textit{and} \quad \xi_\eta \rightharpoonup \xi \implies \mathcal{R}^*(\xi) \le \liminf_{\eta\to 0} \mathcal{R}_\eta^*(\xi_\eta).$$

3. *For all $(v_\eta)_\eta \subset X$, $(\xi_\eta)_\eta \subset X^*$*

$$\left.\begin{array}{rl} v_n \to v & \textit{as } \eta \to 0^+, \\ \xi_\eta \rightharpoonup \xi & \textit{as } \eta \to 0^+, \\ \xi_\eta \in \partial\mathcal{R}_\eta(v_\eta) & \textit{for all } \eta > 0 \end{array}\right\} \implies \xi \in \partial\mathcal{R}(v)\,. \tag{4.21}$$

Proof. By strict convexity of X^*, the space X is *smooth* (cf., e.g., [BaP12, Thm. 1.101]), which is equivalent to the property that the duality mapping $\mathrm{J}_X = \partial\big(\frac{1}{2}\|\cdot\|^2_X\big) : X \rightrightarrows X^*$ is *single-valued* and one-to-one, see [BaP12, Rmk. 1.100]. Hence, the differentiability of the functionals $\mathcal{R}_\eta$ follows, cf. [Att84, Sec. 3.4.1]; in particular, we have

$$\mathcal{R}_\eta(v) = \mathcal{R}\big(W_\eta(v)\big) + \frac{\|W_\eta(v){-}v\|^2}{2\eta} \quad \text{and} \quad \mathrm{D}\mathcal{R}_\eta(v) = \frac{1}{\eta}\mathrm{J}_X\big(v{-}W_\eta(v)\big) \quad \text{for } v \in X, \tag{4.22}$$

with $W_\eta(v) = \big(\mathrm{Id}{+}\eta\mathrm{J}_X{\circ}\partial\mathcal{R}\big)^{-1}(v)$.

To show (4.20), we use (4.22) and that $\mathcal{R}$ has superlinear growth. This gives

$$\exists\, \overline{C} > 0 \;\; \forall\, \eta > 0 \;\; \forall\, v \in X: \quad \|W_\eta(v)\| \le \overline{C} + \mathcal{R}(W_\eta(v)) \le \overline{C} + \mathcal{R}_\eta(v)\,.$$

Moreover, using $\mathcal{R} \ge 0$ we find

$$\|W_\eta(v){-}v\| \le (2\eta\mathcal{R}_\eta(v))^{1/2} \le (2\mathcal{R}_\eta(v))^{1/2} \quad \text{for all } \eta \in (0,1]\,.$$

Thus, $\|v\| \le \|W_\eta(v)\| + \|W_\eta(v){-}v\| \le \overline{C} + \mathcal{R}_\eta(v) + (2\mathcal{R}_\eta(v))^{1/2}$, and (4.20) follows.

The Mosco convergence follows from the results in [Att84, Sec. 3.4]. Finally, in order to show (4.21), we note that $\xi_\eta \in \partial\mathcal{R}_\eta(v_\eta)$ equivalent to

$$\big\langle \xi_\eta, v_\eta \big\rangle = \mathcal{R}_\eta\big(v_\eta\big) + \mathcal{R}^*_\eta(\xi_\eta) = \mathcal{R}_{\eta_n}\big(v_\eta\big) + \mathcal{R}^*(\xi_\eta) + \tfrac{\eta}{2}\|\xi_\eta\|^2_{X^*}.$$

Using that $v_\eta \to v$ and $\xi_\eta \rightharpoonup \xi$ we can pass to the limit on the left-hand side. On the right-hand side we can pass to the liminf and obtain $\langle \xi, v\rangle \ge \mathcal{R}(v) + \mathcal{R}^*(\xi)$, which yields the desired result $\xi \in \partial\mathcal{R}(v)$. ∎

The following result shows that Hypothesis 4.11 can be naturally deduced if the conditioned subdifferential $\partial_{\mathcal{R}}\mathcal{E}$ has a suitable closedness property. A sufficient condition for such closedness involves the following property, which we define in general: We say that a functional $\mathcal{F} : X \to (-\infty, \infty]$ has the *weak-implies-strong property* (WIS, for short), if the following implication holds:

$$u_k \overset{X}{\rightharpoonup} u \text{ (weakly)} \;\; \text{and } \mathcal{F}(u_k) \to \mathcal{F}(u) < \infty \quad \implies \quad u_k \overset{X}{\to} u \text{ (strongly)}. \tag{4.23}$$

This property is certainly true if $\mathcal{F}$ has compact sublevels, however it is also true for uniformly convex functions, or for functionals $\mathcal{F}(u) = g(\|u\|)$ if g is strictly increasing and $\|\cdot\|$ is a uniformly convex norm. Thus, the conditions in Prop. 4.14 are sufficiently easy to check, and indicate the potential for generalizations.

Proposition 4.14 (Sufficient conditions for (4.17)) *Let the* gBGS $(X,\mathcal{E},\mathcal{R})$ *satisfy Hypothesis 3.1 and fix* $\sigma_*>0$ *and* $u^\circ\in\mathrm{dom}(\mathcal{E})$.

(A) If for all $\sigma\in(0,\sigma_*)$, *the operator* $\partial_{\mathcal{R}}\mathcal{E}:\mathrm{Graph}(\boldsymbol{J})\rightrightarrows X^*$ *satisfies*

$$\left.\begin{array}{r}(\eta_n,u_n,\xi_n)\rightharpoonup(0,u,\xi)\quad in\ [0,\infty)\times X\times X^*,\\ u_n\in\boldsymbol{J}_{\eta_n}(\sigma)\ for\ all\ n\in\mathbb{N},\\ and\ \xi_n\in\partial_{\mathcal{R}_{\eta_n}}\mathcal{E}(\sigma;u_n)\quad for\ all\ n\in\mathbb{N}\end{array}\right\}\implies\xi\in\partial_{\mathcal{R}}\mathcal{E}(\sigma;u)\,, \tag{4.24}$$

then the lower semi-continuity property (4.17) *is satisfied.*

(B) If $\mathcal{E}$ *or* $\mathcal{R}$ *have the* WIS *property and* $\mathrm{dom}(\mathcal{R})=X$ *with* $\mathcal{R}^*$ *having superlinear growth at infinity, then condition* (4.24) *holds.*

Proof. For (A) the proof is straightforward. We note that for a sequence $(\eta_n,u_n)\rightharpoonup(0,u)$ with $\alpha:=\liminf_{n\to\infty}\mathcal{C}_{\mathcal{R}_{\eta_n}}(\sigma;u_\eta)<\infty$ we find $\xi_n\in\partial_{\mathcal{R}_{\eta_n}}\mathcal{E}(\sigma;u_n)\subset X^*$ with $\mathcal{R}^*_{\eta_n}(\xi_n)\le\mathcal{C}_{\mathcal{R}_{\eta_n}}(\sigma;u_\eta)+\frac{1}{n}$. In turn, since $\mathcal{R}^*(\xi_n)\le\mathcal{R}^*_{\eta_n}(\xi_n)$, by the coercivity property (3.5) we conclude that $\|\xi_n\|_{X^*}\le C<\infty$.

Thus, after extracting a suitable subsequence (not relabeled) we may assume $\xi_n\rightharpoonup\xi$ and the closedness property (4.24) implies $\xi\in\partial_{\mathcal{R}}\mathcal{E}(\sigma;u)$. Now, using the weak lower semicontinuity of $\mathcal{R}^*$ we have

$$\mathcal{C}_{\mathcal{R}}(\sigma;u)\le\mathcal{R}^*(\xi)\le\liminf_{n\to\infty}\mathcal{R}^*_{\eta_n}(\xi_n)=\liminf_{n\to\infty}\big(\mathcal{C}_{\mathcal{R}_{\eta_n}}(\sigma;u_\eta)+\tfrac{1}{n}\big)=\alpha,$$

which is the desired result of part (A).

For part (B), let (η_n,u_n,ξ_n) be as in (4.24). We use that $u_n\in\boldsymbol{J}_{\eta_n}(\sigma)$ and $u_n\rightharpoonup u$ already imply

$$\Phi^{\eta_n}_\sigma(u_n)=\sigma\mathcal{R}_{\eta_n}\big(\tfrac{1}{\sigma}(u_n-u^\circ)\big)+\mathcal{E}(u_{\eta_n})\longrightarrow\Phi_\sigma(u)\,. \tag{4.25}$$

Indeed, the lim inf-estimate follows by (4.16), whereas for the lim sup we observe that $\Phi^{\eta_n}_\sigma(u_n)\le\Phi^{\eta_n}_\sigma(u)\le\Phi_\sigma(u)$, where the first inequality follows from $u_{\eta_n}\in\boldsymbol{J}_{\eta_n}(\sigma)$ and the second one from $\mathcal{R}_{\eta_n}\le\mathcal{R}$. We also have the individual lim inf-estimates

$$\mathcal{R}(\frac{1}{\sigma}(u-u^\circ))\le\liminf_{n\to\infty}\mathcal{R}_{\eta_n}(\frac{1}{\sigma}(u_n-u^\circ))\quad\text{and}\quad\mathcal{E}(u)\le\liminf_{n\to\infty}\mathcal{E}(u_n), \tag{4.26}$$

the first one by the Mosco-convergence of $\mathcal{R}_{\eta_n}$ to $\mathcal{R}$, and the second one by the weak lower semicontinuity of $\mathcal{E}$. Combining (4.25) and (4.26) we see that both liminfs are limits and that the inequalities must be equalities. Hence, by assumed WIS property for either $\mathcal{R}$, or $\mathcal{E}$, we find the strong convergence $u_n\to u$. Now, we need to take pass to the limit, as $n\to\infty$, in both relations: $\xi_n\in\partial\mathcal{E}(u_n)$ and $-\xi_n\in\partial\mathcal{R}_{\eta_n}\big(\frac{1}{\sigma}(u_n-u^\circ)\big)$. For the first relation, we resort to the closedness property (3.4b) for $\partial\mathcal{E}$ (observe that u_n lies in an energy sublevel since $\mathcal{E}(u_n)\le\Phi^{\eta_n}_\sigma(u_n)\le\mathcal{E}(u^\circ)$). For the second relation, we exploit the gained strong convergence $\frac{1}{\sigma}(u_n-u^\circ)\to\frac{1}{\sigma}(u-u^\circ)$ and property (4.21). All in all, we conclude $\xi\in\partial_{\mathcal{R}}\mathcal{E}(\sigma;u)$, and (4.24) is established. This proves part (B). ∎

4.4 Proof of Theorem 4.12

The first tool in the proof of Theorem 4.12 are the previously recapped properties of the Moreau-Yosida approximations $\mathcal{R}_\eta$. The second tool is the following standard result on the existence of measurable selections for Kuratowski upper limits which, like Lemma 4.3, also draws on [CaV77].

Proposition 4.15 (Measurable selection) *Consider a sequence of* $(u^k)_{k\in\mathbb{N}}$ *of measurable functions* $(0,\sigma_*)\ni\sigma\mapsto u^k_\sigma\in M$*, where* M *is a compact metric space.*

Then, there exists a measurable selection $\sigma\mapsto\widetilde{u}_\sigma$ *of the set of all limit points of* $(u^k_\sigma)_k$*, namely for all* $\sigma\in(0,\sigma_*)$ *there exists a subsequence* $(k^\sigma_n)_{n\in\mathbb{N}}$ *with* $k^\sigma_n\to\infty$ *and* $u^{k^\sigma_n}_\sigma\to\widetilde{u}_\sigma$ *in* M*.*

Furthermore, if $\mathcal{A}:(0,\sigma_*)\times M\to[0,\infty]$ *is a measurable function such that* $\mathcal{A}(\sigma,\cdot)$ *is lower semicontinuous for all* σ*, then the measurable selection* $\widetilde{u}$ *can be chosen such that additionally*

$$\mathcal{A}(\sigma,\widetilde{u}_\sigma)\leq A(\sigma):=\liminf_{k\to\infty}\mathcal{A}(\sigma,u^k_\sigma)\quad\text{for all }(0,\sigma_*).$$

Proof. As M is compact, we can define the multivalued mapping

$$\mathscr{U}:(0,\sigma_*)\rightrightarrows M,\qquad\mathscr{U}(\sigma):=\operatorname*{Ls}_{k\to\infty}\{u^k_\sigma\},\tag{4.27}$$

where Ls is the set of all limit points (in the sense of a Kuratowski upper limit, cf. (2.10b))

$$w\in\operatorname*{Ls}_{k\to\infty}\{u^k_\sigma\}\quad\Longleftrightarrow\quad\exists\,(k_n)_{n\in\mathbb{N}}:\ k_n\to\infty\text{ and }u^{k_n}_\sigma\to w.\tag{4.28}$$

Hence, each $\mathscr{U}(\sigma)$ is well defined, non-empty and closed. The results of [AuF90, Thm. 8.2.5] imply that $\mathscr{U}$ is a measurable multivalued mapping. Now, [CaV77, Thm. 3.22] grants the existence of a measurable selection for $\mathscr{U}$, i.e. a function $\sigma\mapsto\widetilde{u}_\sigma$ with $\widetilde{u}_\sigma\in\mathscr{U}(\sigma)$ for all $\sigma\in(0,\sigma_*)$.

For the second part of the statement involving $\mathcal{A}$ we define

$$\mathscr{U}_{\mathcal{A}}(\sigma):=\{\,u\in\mathscr{U}(\sigma)\,|\,\mathcal{A}(\sigma,u)\leq A(\sigma)\,\}.$$

By the definition of $\mathscr{U}_{\mathcal{A}}$ and the lsc of $\mathcal{A}$ we know have $\mathscr{U}_{\mathcal{A}}(\sigma)\neq\emptyset$ for all σ. Hence, $\mathscr{U}_{\mathcal{A}}:(0,\sigma_*)\rightrightarrows M$ is a measurable multivalued mapping with non-empty and closed values, and we can apply the selection theorem as above, see also [AuF90, Thm. 8.1.4]. ■

We are now ready to carry out the proof of Theorem 4.12 which provides the De Giorgi simple and improved estimate for general dissipation potentials $\mathcal{R}$. For this, consider variational integrands for the regularized gBGS with dissipation potential $\mathcal{R}_\eta$.

Proof of Theorem 4.12.
Step 1: Regularization. For $\eta\in(0,1]$ we consider the dissipation potentials $\mathcal{R}_\eta$, which are radially differentiable. Hence, by Theorem 4.2 there exist variational interpolants $\sigma\mapsto\widetilde{u}^\eta_\sigma$ such that the De Giorgi identity holds

$$\forall\,\sigma>0:\quad\mathcal{E}(\widetilde{u}^\eta_\sigma)+\sigma\mathcal{R}_\eta\big(\frac{1}{\sigma}(\widetilde{u}^\eta_\sigma{-}u^\circ)\big)+\int_0^\sigma\mathcal{C}_{\mathcal{R}_\eta}(\rho;\widetilde{u}^\eta_\rho)\,\mathrm{d}\rho=\mathcal{E}(u^\circ)\,.\tag{4.29}$$

In particular, we have $\widetilde{u}^\eta_\sigma\in\boldsymbol{J}_\eta(\sigma)$, and $\mathcal{C}_{\mathcal{R}_{\eta_n}}(\cdot;\cdot)$ is defined in (3.13).
Step 2: A priori bounds. Clearly, we have that $\mathcal{E}(\widetilde{u}^\eta_\sigma)\leq\mathcal{E}(u^\circ)$ for all σ and η. We will now use the place-holder $v^\eta_\sigma:=\frac{1}{\sigma}(u_{\sigma,\eta}{-}u^\circ)$. Recalling that $\mathcal{E}$ is bounded below by some constant E_0, we deduce from (4.29) that

$$\sigma\mathcal{R}_\eta(v^\eta_\sigma)+E_0\leq\sigma\mathcal{R}_\eta(v^\eta_\sigma)+\mathcal{E}(\widetilde{u}^\eta_\sigma)\overset{(1)}{\leq}\mathcal{E}(u^\circ)\quad\text{for all }\sigma,\,\eta>0\,,\tag{4.30}$$

where (1) derives from the minimality of $\widetilde{u}^\eta_\sigma$. Thus, it follows from (4.20) applied with $S:=\mathcal{E}(u^\circ)$ that there exists a constant $C_S>0$ such that $\sigma\|v^\eta_\sigma\|\leq C_S$. Now, recalling that $v_\sigma^\eta = \frac{1}{\sigma}(\widetilde{u}_\sigma^\eta - u^\circ)$ we infer that $\|\widetilde{u}_\sigma^\eta\| \leq \sigma\|v_\sigma^\eta\| \ + \|u^\circ\| \leq C_S + \|u^\circ\|$. All in all, we proved that

$$\exists\, \overline{E} > 0 \ \ \forall\, \eta \in (0,1] \ \ \forall\, \sigma \in (0,\sigma_*) : \qquad \mathcal{E}(\widetilde{u}_\sigma^\eta) + \|\widetilde{u}_\sigma^\eta\| \leq \overline{E}. \tag{4.31}$$

Step 3: Measurable selection. Using the result of Step 2, we are able to apply Proposition 4.15 with the choices $M = B_{\overline{E}}(0) \cap S_{\overline{E}} \subset X$ equipped with the weak topology and $\mathcal{A} = \mathcal{S}_\mathcal{R}$ which is lsc, see (3.11b). Hence, we obtain a measurable function $\sigma \mapsto \widetilde{u}_\sigma \in X$ with $\widetilde{u}_\sigma^\eta \rightharpoonup \widetilde{u}_\sigma$ along a subsequence $\eta = \eta_k^\sigma \searrow 0$, where $(\eta_k^\sigma)_{k\in\mathbb{N}}$ may depend on σ. Moreover,

$$\mathcal{S}_\mathcal{R}(\widetilde{u}_\sigma) \leq \liminf_{\eta\to 0} \mathcal{S}_\mathcal{R}(\widetilde{u}_\sigma^\eta) \quad \text{for all } \sigma \in (0,\sigma_*). \tag{4.32}$$

Step 4: Convergence of Φ_σ. Recall that $\Phi_\sigma^\eta(u) = \mathcal{E}(u) + \sigma\mathcal{R}_\eta\big(\frac{1}{\sigma}(u - u^\circ)\big)$ and that the family $\big(\Phi_\sigma^\eta(\cdot)\big)_\eta$ Mosco-converges to $\Phi_\sigma(\cdot)$. Moreover, the functions $\sigma \mapsto \phi^\eta(\sigma) = \min\big\{\, \Phi_\sigma^\eta(u) \big| u \in X \,\big\}$ are given independently of our curves $\widetilde{u}_\eta^\sigma$ and $\widetilde{u}_\sigma$.

From $\widetilde{u}_\eta^\sigma \in \boldsymbol{J}_\eta(\sigma)$ we have $\phi^\eta(\sigma) = \Phi_\sigma^\eta(\widetilde{u}_\sigma^\eta)$. Using the convergence $\widetilde{u}_\sigma^{\eta_k^\sigma} \rightharpoonup \widetilde{u}_\sigma$ as $k \to \infty$ and the Mosco-convergence of Φ_σ^η we obtain (cf. the argument for (4.25))

$$\phi(\sigma) = \Phi_\sigma(\widetilde{u}_\sigma) = \lim_{k\to\infty} \phi^{\eta_k^\sigma}(\sigma) = \lim_{k\to\infty} \Phi_\sigma^{\eta_k^\sigma}\big(\widetilde{u}_\sigma^{\eta_k^\sigma}\big).$$

Step 5: The simple De Giorgi estimate. Here we use the trivial lower bound $\mathcal{C}_{\mathcal{R}_\eta}(\rho; w) \geq \mathcal{S}_{\mathcal{R}_\eta}(w) \geq \mathcal{S}_\mathcal{R}(w)$ because of $\mathcal{R}_\eta^* \geq \mathcal{R}^*$. We start from (4.29) and obtain for all fixed $\sigma \in (0,\sigma_*)$ and all $\eta \in (0,1)$ the estimates

$$\Phi_\sigma^\eta\big(\widetilde{u}_\eta^\sigma\big) + \int_0^\sigma \mathcal{S}_\mathcal{R}(\widetilde{u}_\rho^\eta)\,\mathrm{d}\rho \leq \mathcal{E}(u^\circ).$$

In Step 4 we have shown that the first term converges to the desired limit. For the integral it suffices to use Fatou's lemma and the liminf estimate (4.32). Thus, we obtain

$$\Phi_\sigma\big(\widetilde{u}_\sigma\big) + \int_0^\sigma \mathcal{S}_\mathcal{R}(\widetilde{u}_\rho)\,\mathrm{d}\rho \leq \mathcal{E}(u^\circ),$$

which is the desired simple De Giorgi estimate.

Step 6: The improved De Giorgi estimate. We proceed as in the previous steps, but apply the selection theorem to the enlarged family $(g^\eta)_{\eta\in(0,1]}$ with

$$g^\eta : (0,\sigma_*) \to M_2 \quad \text{with } \ g_\sigma^\eta := (\widetilde{u}_\sigma^\eta, \eta) \in M_2 := (B_{\overline{E}}(0) \cap S_{\overline{E}}) \times [0,1] \subset X \times \mathbb{R}.$$

The measurable function $\mathcal{A}$ is defined via the conditioned $\mathcal{R}_\eta$ slope as

$$\mathcal{A}\big(\sigma, (u,\eta)\big) := \begin{cases} \mathcal{C}_{\mathcal{R}_\eta}(\sigma; u) & \text{for } u \in \boldsymbol{J}_\eta(\sigma), \\ \infty & \text{otherwise,} \end{cases}$$

where for $\eta = 0$ we use the notations $\mathcal{R}_0 = \mathcal{R}$ and $\boldsymbol{J}_0(\sigma) = \boldsymbol{J}(\sigma)$. The additional Hypothesis (4.17) is now exactly tailored in such a way that $\mathcal{A}$ is lower semi-continuous on $(0,\sigma_*) \to M_2$. Hence, Proposition 4.15 provides a selection $\widetilde{u}$ such that

$$\widetilde{u}_\sigma^\eta \rightharpoonup \widetilde{u}_\sigma \text{ and } \ \mathcal{C}_\mathcal{R}(\sigma; \widetilde{u}_\sigma) \leq \liminf_{\eta\to 0} \mathcal{C}_{\mathcal{R}_\eta}(\sigma; \widetilde{u}_\sigma^\eta) \quad \text{for all } \sigma \in (0,\sigma_*).$$

As in Step 4 and 5 we can pass to the liminf in (4.29) by combining the last estimate and Fatou's lemma, and the improved De Giorgi estimate follows.

Finally, the improved estimate with selection 4.19b directly follows via Lemma 4.3. ∎

4.5 An alternative route to equality via the chain rule

In this section, we extend an argument sketched in the introduction by showing that, even without radial differentiability for the dissipation potential $\mathcal{R}$, the improved De Giorgi estimate with force selection (4.19b) can be enhanced to an identity along a curve $\sigma \mapsto \widetilde{u}_\sigma$

1. with suitable regularity
2. such that the chain rule holds for $\sigma \mapsto \mathcal{E}(\widetilde{u}_\sigma)$.

Thus, the motivation for this section is to support our conjecture that the De Giorgi identity is valid in more general situations than those understood by now.

We substantiate the above requirements in the following hypothesis and highlight that, for the variational interpolant it is in fact sufficient to have (piecewise) absolute continuity, whereas the results from Section 4.2 even granted, under additional assumptions, (local) Lipschitz continuity of the curve $\sigma \mapsto \widetilde{u}_\sigma$.

Hypothesis 4.16 (Regularity of the variational interpolant & chain rule) *Let $\sigma_* > 0$. We consider a curve $(0,\sigma_*) \ni \sigma \mapsto \widetilde{u}_\sigma \in J_\sigma(u^\circ)$ such that*

1. *$\widetilde{u}$ is* piecewise absolutely continuous *on $(0,\sigma_*)$, namely*

$$\exists\, p \geq 1\ \exists \text{ a partition } \{\tau_j\}_{j=0}^J \text{ of } [0,\sigma_*)\ \forall\, j \in \{1,..,J\}:\ \widetilde{u} \in \mathrm{AC}^p\big((\tau_{j-1},\tau_j);X\big)\,. \tag{4.33}$$

As a consequence, for all $\{1,\dots,J\}$ the one-sided limits $\widetilde{u}_{\tau_j}^- := \lim_{\sigma\to\tau_j^-} \widetilde{u}_\sigma$ and $\widetilde{u}_{\tau_{j-1}}^+ := \lim_{\sigma\to\tau_{j-1}^+} \widetilde{u}_\sigma$ exist;

2. *there exists a measurable selection $\widetilde{\xi} : (0,\sigma_*) \to X^*$ with $\widetilde{\xi}_\sigma \in \partial\mathcal{E}(\widetilde{u}_\sigma)$ for a.a. $\sigma \in (0,\sigma_*)$ such that*

$$\begin{aligned} &\forall\, j \in \{1,\dots,J\}:\ \sigma \mapsto \mathcal{E}(\widetilde{u}_\sigma) \text{ is absolutely continuous on } (\tau_{j-1},\tau_j),\\ &\qquad\qquad \text{and } \frac{\mathrm{d}}{\mathrm{d}\sigma}\mathcal{E}(\widetilde{u}_\sigma) = \langle \widetilde{\xi}_\sigma, \widetilde{u}_\sigma' \rangle \quad \text{for a.a. } \sigma \in (\tau_{j-1},\tau_j)\,. \end{aligned} \tag{4.34}$$

Let us dwell on the chain-rule condition in Hypothesis 4.16. First of all, we mention that, in the proof of Theorem 4.17 we shall in fact apply (4.34) to a measurable force selection $\sigma \mapsto \widetilde{\xi}_\sigma \in \partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma)$. We also emphasize the chain rule for $\mathcal{E}$ *evaluated along* the (assumedly) absolutely continuous curve $\sigma \mapsto \widetilde{u}_\sigma$ is required, only. Nonetheless, it is natural to wonder for which classes of energies the chain rule holds *in general.* Some sufficient conditions for its validity, among which λ-convexity of $\mathcal{E}$ for some $\lambda \in \mathbb{R}$, were provided in [MRS13, Prop. 2.4], [MiR23, Prop. A.1]. There, it was shown that for *any* pair $(\mathsf{u},\xi) \in \mathrm{AC}([a,b];X)\times\mathrm{L}^1([a,b];X^*)$ satisfying

$$\sup_{s\in[a,b]} \mathcal{E}(\mathsf{u}(s)) < \infty \quad \text{and} \quad \int_a^b \|\xi(s)\|_* \,\|\mathsf{u}'(s)\|\,\mathrm{d}s < \infty\,, \tag{4.35}$$

the energy $s \mapsto \mathcal{E}(\mathsf{u}(s))$ is absolutely continuous and the chain rule formula $\frac{\mathrm{d}}{\mathrm{d}s}(\mathcal{E}\circ\mathsf{u}) = \langle \xi, \mathsf{u}' \rangle$ holds a.e. in (a,b). For instance, the above chain rule holds for the energy functional from Example 3.5, which is indeed λ-convex for $\lambda = -C_W$, cf. [MRS23, Sec. 4.2].

Now, in the upcoming Theorem 4.17 we will consider curves $\sigma \mapsto \widetilde{u}_\sigma \in J_\sigma(u^\circ)$ and force selections $\sigma \mapsto \widetilde{\xi}_\sigma \in \partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma)$ as in Hypothesis 4.16 *and* satisfying the improved

De Giorgi estimate (4.19b) (see Theorem 4.12 for the existence of such selections). Then, we will clearly have the energy bound $\sup_{\sigma\in]0,\sigma_*)}\mathcal{E}(\widetilde{u}_\sigma)<\infty$, as well as

$$\int_0^{\sigma_*}\mathcal{R}^*(-\widetilde{\xi}_\rho)\,\mathrm{d}\rho<\infty\,.$$

Therefore, *if* the dual dissipation potential $\xi\mapsto\mathcal{R}^*(\xi)$ controls $\|\xi\|_*^{p'}$ with p' conjugate to p, then combining the above estimate with the condition $u\in\mathrm{AC}^p\big((\tau_j,\tau_{j+1});X\big)$, we conclude that for the pair $(\tilde{u},\tilde{\xi})$ also the second estimate in (4.35) holds.

We are now in a position to state the our last result that the De Giorgi identity holds even without radial differentiability, if we have some regularity of the variational interpolant $\sigma\mapsto(\widetilde{u}_\sigma,\widetilde{\xi}_\sigma)$.

Theorem 4.17 (From the De Giorgi estimate to the identity) *Let the* gBGS $(X,\mathcal{E},\mathcal{R})$ *satisfy Hypotheses 3.1. and let* $u^\circ\in\mathrm{dom}(\mathcal{E})$ *be fixed. Let* $(0,\sigma_*)\ni\sigma\mapsto\widetilde{u}_\sigma\in J_\sigma$ *and* $(0,\sigma_*)\ni\sigma\mapsto\widetilde{\xi}_\sigma\in\partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma)$ *satisfy Hypothesis 4.16, as well as the improved De Giorgi estimate with selection* (4.19b) *on* $(0,\sigma_*)$.

Then, (4.19b) *holds as an equality, i.e.*

$$\mathcal{E}(\widetilde{u}_{\sigma_*})+\sigma_*\mathcal{R}\big(\frac{1}{\sigma_*}(\widetilde{u}_{\sigma_*}-u^\circ)\big)+\int_0^{\sigma_*}\mathcal{R}^*(-\widetilde{\xi}_\rho)\,\mathrm{d}\rho=\mathcal{E}(u^\circ) \tag{4.36}$$

and, a fortiori, we have

$$\widetilde{\xi}_\sigma\in\mathfrak{A}_{\mathcal{R}}(\sigma,\widetilde{u}_\sigma)=\operatorname{argmin}\big\{\,\mathcal{R}^*(-\xi)\,\big|\,\xi\in\partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma)\,\big\}\quad\textit{for a.a.}\,\sigma\in(0,\sigma_*)\,, \tag{4.37}$$

hence the De Giorgi identity (3.17) *holds.*

Proof. Step 1: Let us fix $j\in\{1,\ldots,J{-}1\}$. It is immediate to check that $\widetilde{u}_{\tau_j}^{\pm}\in J_{\tau_j}$, and thus $\Phi_{\tau_j}(\widetilde{u}_{\tau_j}^-)=\phi(\tau_j)=\Phi_{\tau_j}(\widetilde{u}_{\tau_j}^+)$, i.e.

$$\mathcal{E}(\widetilde{u}_{\tau_j}^-)+\tau_j\mathcal{R}\Big(\frac{1}{\tau_j}(\widetilde{u}_{\tau_j}^-{-}u^\circ)\Big)=\mathcal{E}(\widetilde{u}_{\tau_j}^+)+\tau_j\mathcal{R}\Big(\frac{1}{\tau_j}(\widetilde{u}_{\tau_j}^+{-}u^\circ)\Big)\,. \tag{4.38}$$

Let now $\sigma\mapsto\widetilde{\xi}_\sigma\in\partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma)$ be as in the statement. Hypothesis 4.16 enables us to apply the chain rule on any interval $[s_*,s^*]\subset(\tau_j,\tau_{j+1})$, thus concluding that

$$\lim_{\sigma\to\tau_{j+1}^-}\mathcal{E}(\widetilde{u}_\sigma)-\lim_{\sigma\to\tau_j^+}\mathcal{E}(\widetilde{u}_\sigma)=\int_{\tau_j}^{\tau_{j+1}}\langle\widetilde{\xi}_\rho,\widetilde{u}'_\rho\rangle\,\mathrm{d}\rho\,.$$

Now, we have that

$$\lim_{\sigma\to\tau_j^+}\mathcal{E}(\widetilde{u}_\sigma)=\lim_{\sigma\to\tau_j^+}\bigg(\phi(\sigma){-}\sigma\mathcal{R}\Big(\frac{1}{\sigma}(\widetilde{u}_\sigma{-}u^\circ)\Big)\bigg)=\phi(\tau_j){-}\tau_j\mathcal{R}\Big(\frac{1}{\tau_j}(\widetilde{u}_{\tau_j}{-}u^\circ)\Big)=\mathcal{E}(\widetilde{u}_{\tau_j}^+)$$

by the continuity of ϕ (recall Proposition 3.17) and of $\mathcal{R}$. Analogously, $\lim_{\sigma\to\tau_{j+1}^-}\mathcal{E}(\widetilde{u}_\sigma)=\mathcal{E}(\widetilde{u}_{\tau_{j+1}}^-)$. Therefore,

$$\mathcal{E}(\widetilde{u}_{\tau_{j+1}}^-)-\mathcal{E}(\widetilde{u}_{\tau_j}^+)=\int_{\tau_j}^{\tau_{j+1}}\langle\widetilde{\xi}_\rho,\widetilde{u}'_\rho\rangle\,\mathrm{d}\rho\,. \tag{4.39}$$

In turn, since $\widetilde{\xi}_\sigma \in \partial_{\mathcal{R}}\mathcal{E}(\sigma;\widetilde{u}_\sigma)$ implies $-\widetilde{\xi}_\sigma \in \partial\mathcal{R}\left(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\right)$, we may apply the chain rule for the functional $\sigma \mapsto \sigma\mathcal{R}\left(\frac{1}{\sigma}(\widetilde{u}_\sigma - u^\circ)\right)$ and find that

$$\begin{aligned} &\tau_{j+1}\mathcal{R}\Big(\frac{1}{\tau_{j+1}}(\widetilde{u}^-_{\tau_{j+1}} - u^\circ)\Big) - \tau_j\mathcal{R}\Big(\frac{1}{\tau_j}(\widetilde{u}^+_{\tau_j} - u^\circ)\Big) \\ &= \int_{\tau_j}^{\tau_{j+1}} \Big(\mathcal{R}\big(\frac{1}{\rho}(\widetilde{u}_\rho - u^\circ)\big) + \rho\big\langle -\widetilde{\xi}_\rho, \frac{1}{\rho}\widetilde{u}'_\rho - \frac{1}{\rho^2}(\widetilde{u}_\rho - u^\circ)\big\rangle\Big)\,\mathrm{d}\rho\,. \end{aligned} \tag{4.40}$$

Adding (4.39) and (4.40), observing the cancellation of the term $\int_{\tau_j}^{\tau_{j+1}}\langle\widetilde{\xi}_\rho, \widetilde{u}'_\rho\rangle\,\mathrm{d}\rho$ and rearranging the remaining integral terms, we find

$$\begin{aligned} &\mathcal{E}(\widetilde{u}^-_{\tau_{j+1}}) + \tau_{j+1}\mathcal{R}\Big(\frac{1}{\tau_{j+1}}(\widetilde{u}^-_{\tau_{j+1}} - u^\circ)\Big) + \int_{\tau_j}^{\tau_{j+1}}\Big(\big\langle -\widetilde{\xi}_\rho, \tfrac{1}{\rho}(\widetilde{u}_\rho - u^\circ)\big\rangle - \mathcal{R}\big(\frac{1}{\rho}(\widetilde{u}_\rho - u^\circ)\big)\Big)\,\mathrm{d}\rho \\ &= \mathcal{E}(\widetilde{u}^+_{\tau_j}) + \tau_j\mathcal{R}\Big(\frac{1}{\tau_j}(\widetilde{u}^+_{\tau_j} - u^\circ)\Big)\,. \end{aligned}$$

Now, since $-\widetilde{\xi}_\rho \in \partial\mathcal{R}\big(\frac{1}{\rho}(\widetilde{u}_\rho - u^\circ)\big)$, the integrand in the third term on the left-hand side equals $\mathcal{R}^*(-\widetilde{\xi}_\rho)$. In turn, by (4.38), the right-hand side equals $\Phi_{\tau_j}(\widetilde{u}^-_{\tau_j})$. All in all, we conclude that

$$\begin{aligned} &\mathcal{E}(\widetilde{u}^-_{\tau_{j+1}}) + \tau_{j+1}\mathcal{R}\Big(\frac{1}{\tau_{j+1}}(\widetilde{u}^-_{\tau_j} - u^\circ)\Big) + \int_{\tau_j}^{\tau_{j+1}} \mathcal{R}^*(-\widetilde{\xi}_\rho)\,\mathrm{d}\rho \\ &\quad = \mathcal{E}(\widetilde{u}^-_{\tau_j}) + \tau_j\mathcal{R}\Big(\frac{1}{\tau_j}(\widetilde{u}^-_{\tau_j} - u^\circ)\Big) \quad \text{for all } j \in \{1, \ldots, J-1\}\,. \end{aligned} \tag{4.41}$$

<u>*Step 2:*</u> It remains to derive the analogue of (4.41) for $j = 0$. With this aim, we consider the first interval $[0, \tau_1]$ of the partition and fix $\mu \in (0, \tau_1]$. Repeating the arguments from Step 1 we find

$$\begin{aligned} &\mathcal{E}(\widetilde{u}^-_{\tau_1}) + \tau_1\mathcal{R}\Big(\frac{1}{\tau_1}(\widetilde{u}^-_{\tau_1} - u^\circ)\Big) + \int_{\mu}^{\tau_{j+1}} \mathcal{R}^*(-\widetilde{\xi}_\rho)\,\mathrm{d}\rho \\ &\quad = \mathcal{E}(\widetilde{u}_\mu) + \mu\mathcal{R}\Big(\frac{1}{\mu}(\widetilde{u}_\mu - u^\circ)\Big) = \Phi_\mu(\widetilde{u}_\mu) = \phi(\mu) \to \mathcal{E}(u^\circ)\,, \end{aligned}$$

where the last identity follows because $\widetilde{u}_\mu$ is a minimizer and ϕ is the value function, and the convergence stems from (3.6). We thus conclude

$$\mathcal{E}(\widetilde{u}^-_{\tau_1}) + \tau_1\mathcal{R}\Big(\frac{1}{\tau_1}(\widetilde{u}^-_{\tau_1} - u^\circ)\Big) + \int_0^{\tau_{j+1}} \mathcal{R}^*(-\widetilde{\xi}_\rho)\,\mathrm{d}\rho = \mathcal{E}(u^\circ)\,. \tag{4.42}$$

<u>*Step 3: Conclusion of the proof.*</u> Adding (4.41) for $j = 1, \ldots, J-1$ and (4.42) we obtain

$$\mathcal{E}(\widetilde{u}^-_{\sigma_*}) + \sigma_*\mathcal{R}\Big(\frac{1}{\sigma_*}(\widetilde{u}^-_{\sigma_*} - u^\circ)\Big) + \int_0^{\sigma_*} \mathcal{R}^*(-\widetilde{\xi}_\rho)\,\mathrm{d}\rho = \mathcal{E}(u^\circ)\,.$$

In analogy to (4.38) for $\tau_J = \sigma_*$ we also have $\Phi_{\sigma_*}(\widetilde{u}^-_{\tau_J}) = \phi(\sigma_*) = \Phi_{\sigma_*}(\widetilde{u}_{\sigma_*})$, such that (4.36) is established. ∎

Acknowledgments. The authors are grateful to an anonymous reviewer for several improvements and many helpful remarks. The research of A.M. has been partially supported by Deutsche Forschungsgemeinschaft (DFG) through the Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689) subproject “DistFell”. R.R. acknowledges the support of GNAMPA (INDAM), and of the MIUR-PRIN Grant 2020F3NCPX “Mathematics for industry 4.0 (Math4I4)”.

References

[AGS05] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, first ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005.

[Amb95] L. Ambrosio: *Minimizing movements.* Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5) **19** (1995) 1773–1799.

[AmT04] L. Ambrosio and P. Tilli, *Topics on analysis in metric spaces*, Oxford Lecture Series in Mathematics and its Applications, vol. 25, Oxford University Press, Oxford, 2004.

[Asp68] E. Asplund: *Fréchet differentiability of convex functions.* Acta Math. **121** (1968) 31–47.

[Att84] H. Attouch, *Variational convergence for functions and operators*, Applicable Mathematics Series, Pitman (Advanced Publishing Program), Boston, MA, 1984.

[AuF90] J.-P. Aubin and H. Frankowska, *Set-valued analysis*, Birkhäuser 1990.

[Bac21] A. Bacho, *On the nonsmooth analysis of nonlinear evolution inclusions of first and second order with applications*, Ph.D. thesis, Technischen Universität Berlin, 2021.

[BaP12] V. Barbu and T. Precupanu, *Convexity and optimization in Banach spaces*, fourth ed., Springer Monographs in Mathematics, Springer, Dordrecht, 2012.

[Bre73] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Mathematics Studies, No. 5, North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973.

[BBI01] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, vol. 33, American Mathematical Society, Providence, RI, 2001.

[CaV77] C. Castaing and M. Valadier, *Convex analysis and measurable multifunctions*, Lectures Notes in Mathematics, Vol. 580, Springer-Verlag, Berlin-New York, 1977.

[De93] E. De Giorgi, *New problems on minimizing movements*, Boundary value problems for partial differential equations and applications, RMA Res. Notes Appl. Math., vol. 29, Masson, Paris, 1993, pp. 81–98.

[EkT99] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Classics In Applied Mathematics, vol. 28, Society for Industrial and Applied Mathematics (SIAM), Philadelphia,1999.

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

[Mie23] A. Mielke: *An introduction to the analysis of gradients systems.* WIAS-Preprint 3022 (2023), arXiv:2306.05026.

[MiR23] A. Mielke and R. Rossi: *Balanced-viscosity solutions to infinite-dimensional multi-rate systems.* Arch. Ration. Mech. Anal. **247**:53 (2023) 1–100.

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

[MRS23] A. Mielke, R. Rossi, and A. Stephan: *On time-splitting methods for gradient flows with two dissipation mechanisms*. Preprint arXiv:2307.16137 (2023).

[Mor06] B. S. Mordukhovich, *Variational analysis and generalized differentiation. I*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 2006.

[RMS08] R. Rossi, A. Mielke, and G. Savaré: *A metric approach to a class of doubly nonlinear evolution equations and applications*. Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **7**:1 (2008) 97–169.

[RoSa06] R. Rossi and G. Savaré: *Gradient flows of non convex functionals in Hilbert spaces and applications*. ESAIM Control Optim. Calc. Var. **12**:3 (2006) 564–614.