NONSMOOTH ANALYSIS OF DOUBLY NONLINEAR EVOLUTION EQUATIONS

ALEXANDER MIELKE, RICCARDA ROSSI, AND GIUSEPPE SAVARÉ

Abstract. In this paper we analyze a broad class of abstract doubly nonlinear evolution equations in Banach spaces, driven by nonsmooth and nonconvex energies. We provide some general sufficient conditions, on the dissipation potential and the energy functional, for existence of solutions to the related Cauchy problem. We prove our main existence result by passing to the limit in a time-discretization scheme with variational techniques. Finally, we discuss an application to a material model in finite-strain elasticity.
AMS Subject Classification: 35A15, 35K50, 35K85 49Q20, 58E99.
Key words: doubly nonlinear equations, differential inclusions, generalized gradient flows, finite-strain elasticity.

1. Introduction

In this paper we investigate (the Cauchy problem for) the *doubly nonlinear evolution equation*

$$\partial\Psi(u'(t)) + \mathrm{F}_t(u(t)) \ni 0 \quad \text{in } V^* \quad \text{for a.a. } t \in (0,T). \tag{1.1}$$

Here,

$$V \text{ is a (separable) reflexive Banach space,}$$

and

$$\Psi : V \to [0,+\infty) \quad \text{is a convex potential with} \quad \Psi(0) = 0, \quad \lim_{\|v\|\uparrow+\infty} \frac{\Psi(v)}{\|v\|} = +\infty, \tag{1.2}$$

$\partial\Psi : V \rightrightarrows V^*$ is its usual (convex analysis) subdifferential, and $\mathrm{F} : [0,T] \times V \rightrightarrows V^*$ is a time-dependent family of multivalued maps which are induced by a suitable "(sub)differential" (with respect to the variable u), of a lower semicontinuous time-dependent

$$\text{energy functional} \quad \mathcal{E} : (t,u) \in [0,T] \times V \mapsto \mathcal{E}_t(u) \in (-\infty,+\infty].$$

The quadruple $(V,\mathcal{E},\Psi,\mathrm{F})$ indeed generates what will be later on referred to as *generalized gradient system*. The aim of this paper is to study existence, stability and approximation results for solutions to *generalized gradient systems*, for a large family of quadruples $(V,\mathcal{E},\Psi,\mathrm{F})$. Beside the generality of the convex dissipation potential Ψ (our main assumption is that it exhibits superlinear growth at infinity), we aim to tackle a class of multivalued operators F as broad as possible. Furthermore, we consider a general dependence of the energy functional $\mathcal{E}$ on time (we refer in particular to the properties of the map $t \mapsto \mathcal{E}_t(u)$ and the related notion $\partial_t\mathcal{E}$ of derivative with respect to time).

To highlight these issues, let us consider some motivating examples, in an increasing order of generality.

1. ***Finite dimensional ODE's.*** The simplest example of gradient system is provided by a finite-dimensional space $V = \mathbb{R}^d$ and an energy functional $\mathcal{E} \in C^1([0,T] \times V)$; in this case we take

$$\mathrm{F}_t(u) = \mathrm{D}\mathcal{E}_t(u), \text{ with } \mathrm{D}\mathcal{E}_t \text{ the standard differential of the energy } \ u \mapsto \mathcal{E}_t(u),$$

Date: May 21, 2011.
R.R. and G.S. have been partially supported by a MIUR-PRIN 2008 grant for the project "Optimal mass transportation, geometric and functional inequalities and applications"; A.M. was partially supported by DFG via FOR787MicroPlast under Mie459/5-2. R.R. acknowledges the kind hospitality of the Weierstraß-Institut, where part of this work was carried out.

and (1.1) reads

$$\partial\Psi(u'(t)) + \mathrm{D}\mathcal{E}_t(u(t)) \ni 0 \quad \text{for a.a. } t \in (0,T). \tag{1.3}$$

In the quadratic case $\Psi(v) := \frac{1}{2}|v|^2$, $|\cdot|$ being the usual Euclidean norm on $\mathbb{R}^d$, (1.3) is the Gradient Flow generated by $\mathcal{E}$

$$u'(t) + \mathrm{D}\mathcal{E}_t(u(t)) = 0 \quad \text{for a.a. } t \in (0,T). \tag{1.4}$$

2. ***λ-convex functionals, C^1-perturbation.*** More generally, one can consider energies of the form

$$\mathcal{E}_t(v) := E(v) - \langle \ell(t), v\rangle \quad \text{with } \ell : [0,T] \to V^* \text{ an external loading,} \tag{1.5}$$

and $E : V \to (-\infty, +\infty]$ a λ-convex functional for some $\lambda \in \mathbb{R}$, i.e. it satisfies

$$E((1-\theta)u_0 + \theta u_1) \le (1-\theta)E(u_0) + \theta E(u_1) - \frac{\lambda}{2}\theta(1-\theta)\|u_0 - u_1\|_V^2 \quad \text{for all } u_0, u_1 \in D. \tag{1.6}$$

In this case, F admits the representation

$$\mathrm{F}_t(u) = \mathrm{F}(u) - \ell(t), \quad \mathrm{F}(u) = \partial E(v) \quad \text{for all } u \in V, \tag{1.7}$$

with ∂E the *Fréchet* subdifferential of E, defined at $u \in D := \mathrm{dom}(E)$ by

$$\xi \in \mathrm{F}(u) = \partial E(u) \subset V^* \quad \Leftrightarrow \quad E(v) - E(u) - \langle \xi, v - u\rangle \ge o(\|v-u\|) \quad \text{as } v \to u \text{ in } V. \tag{1.8}$$

It is well known that for all $u \in D$ the (possibly empty) set $\partial E(u) \subset V^*$ is weakly* closed, and it reduces to the singleton $\{\mathrm{D}E(u)\}$ if the functional E is Gâteaux-differentiable at u. Furthermore, if E is convex, then $\partial E(u)$ coincides with the subdifferential of E in the sense of convex analysis. In such a framework, existence and approximation results for the generalized gradient system $(V, \mathcal{E}, \Psi, \partial\mathcal{E})$, with V a reflexive space and Ψ a general dissipation potential as in (1.2), have been proved by [16, 15], while in [52] the long-time behavior of the solutions to $(V, \mathcal{E}, \Psi, \partial\mathcal{E})$ has been addressed. Notice that, when E is a C^1-perturbation of a convex functional E_0, i.e.

$$E(u) := E_0(u) + E_1(u), \quad E_0 \text{ convex, } E_1 \text{ of class } C^1, \tag{1.9}$$

then one has the natural decomposition

$$\partial E(u) = \partial E_0(u) + \mathrm{D}E_1(u), \tag{1.10}$$

which has been exploited in [51] to prove well-posedness (for the Cauchy problem) for the gradient system $(V, \mathcal{E}, \Psi, \partial\mathcal{E})$, and existence of the global attractor for the related dynamical system.

It is worthwhile mentioning that, in cases 1–2, the pair (E, F) satisfies a crucial *closedness* property: the graph of the multivalued map $u \mapsto (E(u), \mathrm{F}(u))$, i.e. the set $\{(u, E(u), \xi) \,:\, u \in D,\ \xi \in \mathrm{F}(u)\} \subset V \times \mathbb{R} \times V^*$, is strongly-strongly-weakly closed, meaning that, if sequences $u_n \in V, \mathscr{E}_n \in \mathbb{R}, \xi_n \in V^*$ are given, then

$$\big(\xi_n \in \mathrm{F}(u_n), \quad \mathscr{E}_n = E(u_n), \quad u_n \to u, \quad \mathscr{E}_n \to \mathscr{E}, \quad \xi_n \rightharpoonup \xi\big) \;\Rightarrow\; \mathscr{E} = E(u), \quad \xi \in \mathrm{F}(u). \tag{1.11}$$

Let us also emphasize that, in cases 1–2, under standard conditions on the external loading ℓ as a function of time, the energy functional $\mathcal{E}_t(v) = E(v) - \langle \ell(t), v\rangle$ fulfills the following *chain rule*: for all $u \in \mathrm{AC}([0,T]; V)$ and $\xi \in L^1(0,T; V^*)$ with $\xi(t) \in \mathrm{F}_t(u(t))$ for almost all $t \in (0,T)$ (where $\mathrm{AC}([0,T]; V)$ denotes the space of absolutely continuous functions on $[0,T]$ with values in V), and such that $\int_0^T \Psi(u'(t))\,\mathrm{d}t < +\infty$, $\int_0^T \Psi^*(-\xi(t))\,\mathrm{d}t < +\infty$, and $\sup_{t\in[0,T]} |\mathcal{E}_t(u(t))| < +\infty$, then

$$\begin{gathered} \text{the map } t \mapsto \mathcal{E}_t(u(t)) \text{ is absolutely continuous and} \\ \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) = \langle \xi(t), u'(t)\rangle + \partial_t\mathcal{E}_t(u(t)) \quad \text{for a.a. } t \in (0,T)\,. \end{gathered} \tag{1.12}$$

3 ***Marginal functionals.*** There are examples when the Fréchet subdifferential does not satisfy the closedness property (1.11), see also [44]. A typical one, which we analyze in Section 3 in more detail, is given by the so-called *marginal functions*, which are defined via an infimum operation. Let us still consider a finite-dimensional case $V = \mathbb{R}^d$ and a functional

$$\mathcal{E}_t(u) = \min_{\eta \in \mathcal{C}} I_t(\eta, u), \tag{1.13}$$

where $\mathcal{C}$ is a compact topological space and $I \in C^0([0,T]\times\mathcal{C}\times V;\mathbb{R})$ is such that the functional $(t,u)\mapsto I_t(\eta,u)$ is of class C^1 for every $\eta\in\mathcal{C}$. Being $\mathcal{C}$ compact, for every $(t,u)\in[0,T]\times V$ the set

$$M(t,u) := \operatorname{Argmin} I_t(\cdot,u) = \Big\{\eta\in\mathcal{C} : \mathcal{E}_t(u) = I_t(\eta,u)\Big\} \tag{1.14}$$

is not empty. If in addition the map $(t,\eta,u)\mapsto \mathrm{D}_u I_t(\eta,u)$ is continuous on $[0,T]\times\mathcal{C}\times V$, it is not difficult to check that, if ξ belongs to the Fréchet subdifferential $\partial\mathcal{E}_t(u)$, then

$$\xi = \mathrm{D}_u I_t(\eta,u) \quad \text{for } \textit{all } \eta\in M(t,u). \tag{1.15}$$

On the other hand, simple examples show that a limit $\xi = \lim_{n\to\infty}\xi_n$ of sequences ξ_n satisfying (1.15) will only obey the relaxed property

$$\xi = \mathrm{D}_u I_t(\eta,u) \quad \text{for } \textit{some } \eta\in M(t,u). \tag{1.16}$$

In view of property (1.16), it appears that, for reduced functionals of the type (1.13), the appropriate subdifferential is

$$\widehat{\partial}\mathcal{E}_t(u) := \{\mathrm{D}_u I_t(\eta,u) \ : \ \eta\in M(t,u)\}, \tag{1.17}$$

which will be referred to as the *marginal subdifferential* of $\mathcal{E}$. We examine this notion with some detail in Section 3, with the help of significant examples. The latter also highlight that, in the case of marginal energies $\mathcal{E}$ like (1.13), smoothness of the function $t\mapsto\mathcal{E}_t(u)$ for $u\in V$ fixed is no longer to be expected. That is why, one has to recur to a surrogate for the partial derivative $\partial_t\mathcal{E}$, tailored to the marginal case (1.13). In Examples 3.3 and 3.4, we develop some heuristics for such a generalization of $\partial_t\mathcal{E}$, and motivate the fact that this object should be also *conditioned* to the (marginal) subdifferential of the energy with respect to the variable u, and therefore depend on the additional variable $\xi\in\widehat{\partial}\mathcal{E}_t(u)$. This leads to a *generalized derivative with respect to time* $\mathrm{P} = \mathrm{P}_t(u,\xi)$, where $\xi\in\widehat{\partial}\mathcal{E}_t(u)$. For the marginal functional in (1.13), P is defined by

$$\mathrm{P}_t(u,\xi) := \sup\left\{\partial_t I_t(\eta,u) \ : \ \eta\in M(t,u), \ \xi = \mathrm{D}_u I_t(\eta,u)\right\}.$$

4 ***General nonhomogeneous dissipation potentials.*** Last but not least, we emphasize that, beside tackling the above-mentioned nonsmoothness and nonconvexity of the energy, at the same time we treat *general* convex dissipation potentials.

First of all, we extend the existence results of [43], which also addressed doubly nonlinear evolution equations driven by nonconvex energies. Moving from the analysis of gradient systems in a metric setting, the latter paper examines the case of nonconvex energy functionals, albeit smoothly depending on time, but with dissipation potentials of the form $\Psi(v) = \psi(\|v\|)$, where $\psi : [0,+\infty)\to[0,+\infty)$ is convex, l.s.c., and with superlinear growth at infinity. However, in view of applications it is also interesting to deal with dissipation potentials like

$$\Psi(v) = c_1|v|^{p_1} + c_2\|v\|^{p_2}, \quad \text{with } p_1\in[1,\infty),\ p_2\in(1,\infty), \tag{1.18}$$

with $|\cdot|$ a second norm on V. In particular, dissipations of the type (1.18) arise in the vanishing viscosity approximation of rate-independent evolutions described by the doubly nonlinear equation

$$\partial\Psi_{\mathrm{hom}}(u'(t)) + \mathrm{F}_t(u(t)) \ni 0 \quad \text{in } V^* \quad \text{for a.a. } t\in(0,T), \tag{1.19}$$

featuring the 1-positively homogeneous dissipation potential $\Psi_{\mathrm{hom}}(v) = |v|$. The *natural* viscous approximation of (1.19) is indeed the gradient system

$$\partial\Psi_\varepsilon(u'(t)) + \mathrm{F}_t(u(t)) \ni 0 \quad \text{in } V^* \quad \text{for a.a. } t\in(0,T), \quad \text{with } \Psi_\varepsilon(v) = |v| + \frac{\varepsilon}{2}\|v\|^2. \tag{1.20}$$

We mention that the vanishing viscosity limit of (1.20) as $\varepsilon\searrow 0$ has been studied in [34] in the case of a *finite-dimensional* ambient space V. Moving from the existence results for *viscous* doubly nonlinear equations of the present paper, we are going to address the vanishing viscosity analysis of (1.20) in an infinite-dimensional context in the forthcoming paper [35].

Secondly, we consider dissipation potentials $\Psi = \Psi_u(v)$ also depending on the state variable u, hence address the doubly nonlinear equation

$$\partial\Psi_{u(t)}(u'(t)) + \mathrm{F}_t(u(t)) \ni 0 \quad \text{in } V^* \quad \text{for a.a. } t\in(0,T). \tag{1.21}$$

A significant example for potentials of this type will be provided in Section 5, focusing on a model in finite-strain elasticity. In fact, state-dependent dissipations naturally occur in various plasticity models, see for example [31, 30, 5].

The discussion developed throughout Examples 1–4 motivates the analysis of *generalized gradient systems* $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$ which is developed in this paper. As a main goal, we will prove an existence and approximation result for the Cauchy problem for (1.21), under suitable conditions on $\mathcal{E}, \Psi, \mathrm{F}, \mathrm{P}$. To be more precise, we will call a function $u : [0, T] \to V$ *solution* for the generalized gradient system $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$, if $u \in \mathrm{AC}([0, T]; V)$, and there exists $\xi \in L^1(0, T; V^*)$ such that

$$\partial\Psi_{u(t)}(u'(t)) + \xi(t) \ni 0 \qquad \text{for a.a.}\, t \in (0, T), \tag{1.22a}$$

$$\xi(t) \in \mathrm{F}_t(u(t)) \qquad \text{for a.a.}\, t \in (0, T), \tag{1.22b}$$

and (u, ξ) fulfill the *energy identity*

$$\int_0^T \Psi_{u(t)}(u'(t)) + \Psi^*_{u(t)}(-\xi(t))\,\mathrm{d}t + \mathcal{E}_T(u(T)) = \mathcal{E}_0(u(0)) + \int_0^T \mathrm{P}_t(u(t), \xi(t))\,\mathrm{d}t. \tag{1.22c}$$

Let us point out that the energy identity (1.22c) is a crucial item in our definition of solution to $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$. On the one hand, (1.22c) is a consequence of (1.22a)–(1.22b) and of the chain rule (1.12), as it can be checked by testing (1.22a) by $u'(t)$ and integrating on $(0, T)$. On the other hand, as mentioned below, for proving existence of solutions to (1.21), in fact we are going to first derive (1.22b) and (1.22c), and then combine them to obtain (1.22a).

The **plan of the paper** is as follows: in Section 2 we address the analysis of the doubly nonlinear evolution equation (1.1) in a simplified setting: the dissipation potential is *independent* of the variable u, and the energy $\mathcal{E} = \mathcal{E}_t(u)$ is possibly nonsmooth and nonconvex with respect to u, but *smoothly* depending on time. Thus, throughout Sec. 2, the multivalued map $\mathrm{F} : [0, T] \times V \rightrightarrows V^*$ is given by the Fréchet subdifferential of the energy, i.e. $\mathrm{F}_t(u) = \partial\mathcal{E}_t(u)$ for all $(t, u) \in [0, T] \times V$, while $\mathrm{P}_t(u, \xi)$ reduces to the usual partial time-derivative $\partial_t\mathcal{E}_t(u)$. Nonetheless, the analysis of this case still highlights the most significant difficulties arising for nonconvex energies. In such a context, we enucleate the main conditions on the energy functional for proving existence for (1.1). First, we require some suitable *coercivity* property, which amounts to asking that the sublevels of the energy are compact. Second, we impose that the energy $\mathcal{E} : [0, T] \times V \to \mathbb{R}$ and the Fréchet subdifferential $\partial\mathcal{E} : [0, T] \times V \rightrightarrows V^*$ fulfill a (joint) *closedness property*, cf. (1.11). Third, we require that a suitable form of the chain rule (1.12) holds.

Then, we state an existence result for the Cauchy problem for equation (1.1), and outline the steps of its proof, viz. approximation by time-discretization, a priori estimates on the approximate solutions, compactness arguments, and the final passage to the limit of the time-discrete scheme. In developing the proof, we highlight the role played at each step by the aforementioned conditions on the energy functional. Namely, the incremental minimization leads to a discretized version of (1.22a)–(1.22b) and, using the *variational interpolant* of the discrete solutions, we obtain a discrete upper energy estimate, corresponding to the inequality $\leq$ in (1.22c). Exploiting lower semicontinuity arguments and the closedness of the graph of the map $(t, u) \mapsto \{(\mathcal{E}_t(u), \xi, \partial_t\mathcal{E}_t(u)) : \xi \in \mathrm{F}_t(u)\}$, the passage to the limit yields (1.22b) and (1.22c) with $\leq$ instead of $=$. Hence, we employ a suitable *lower chain-rule estimate* to conclude that (1.22c) holds with equality. From this argument, we also have (1.22a).

In Section 3 we discuss finite-dimensional examples of marginal energy functionals. In this way, we motivate and develop some heuristics for new notions of subdifferential of the energy with respect to the variables t and to u, tailored to the case of marginal functionals, viz. the aforementioned marginal subdifferential $\widehat{\partial}\mathcal{E}$ and the generalized partial time-derivative P. We emphasize that, even in finite-dimensional cases, the nonsmoothness of $\mathcal{E}$ forces us to make $\mathrm{P}_t(u, \cdot)$ dependent on $\xi \in \mathrm{F}_t(u)$.

From Section 4 on, we examine the generalized gradient system $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$, with a *state-dependent* dissipation potential $\Psi = \Psi_u(v)$. In such a context we state our main existence and approximation result for (1.21). We also give an upper semicontinuity result for the set of solutions to (1.21) with respect to perturbations of the dissipation potential and of the energy functional.

In Section 5 we present an application of our existence theorem to a PDE system for material models with finite-strain elasticity. Indeed, we consider dissipative material models (also called *generalized standard materials*, cf. [25, 33]) with an internal variable $z : \Omega \to \mathrm{K} \subset \mathbb{R}^m$, while the elastic deformation $\varphi : \Omega \to \mathbb{R}^d$ is quasistatically minimized at each time instant. Thus, we are in the realm of marginal functionals

$$\mathcal{E}_t(z) = \min_{\varphi \in \mathcal{F}} I_t(\varphi, z), \quad \text{where } I_t(\varphi, z) = \mathcal{E}^1(z) + \int_\Omega W(\nabla\varphi, z)\,\mathrm{d}x - \langle \ell(t), \varphi \rangle,$$

with $\mathcal{E}^1$ a convex functional with compact sublevels. Here, $W(\cdot, z)$ is polyconvex to guarantee that the set of minimizers $M(t, z)$ is compact and nonempty. We use the technical assumption $|\mathrm{D}_z W(F, z)| \le \kappa_1 (W(F, z) + \kappa_2)^{1/2}$.

All the proofs of our abstract results are developed in Section 6, relying on some technical tools and auxiliary results collected in the Appendix.

Basic set-up and notation. Hereafter, we will set our analysis in the framework of

a reflexive separable Banach space V

with norm $\|\cdot\|$. We denote by $\langle \cdot, \cdot \rangle$ the duality pairing between V^* and V and by $\|\cdot\|_*$ the norm on V^*.

Our basic assumption on the energy functional $\mathcal{E} : [0, T] \times V \to (-\infty, +\infty]$ is that there exists $D \subset V$ such that

$$\begin{gathered} \mathrm{dom}(\mathcal{E}) = [0, T] \times D, \quad \text{the map } u \mapsto \mathcal{E}_t(u) \text{ is lower semicontinuous for all } t \in [0, T], \\ \exists\, C_0 > 0 \ \ \forall\, (t, u) \in [0, T] \times D \ : \ \mathcal{E}_t(u) \ge C_0. \end{gathered} \tag{E$_0$}$$

Indeed, if the functionals $\mathcal{E}_t$ are bounded from below by some constant independent of t, up to a translation it is not restrictive to assume such a constant to be strictly positive.

Hereafter, we will use the following notation

$$\mathcal{G}(u) := \sup_{t \in [0,T]} \mathcal{E}_t(u) \quad \text{for every } u \in D. \tag{1.23}$$

Furthermore, we will denote by $\mathrm{F} : [0, T] \times D \rightrightarrows V^*$ a time-dependent family of multivalued maps, such that for $t \in [0, T]$ the mapping F_t is a (suitable notion of) subdifferential of the functional $u \mapsto \mathcal{E}_t(u)$. We use the notation

$$\begin{aligned} \mathrm{dom}(\mathrm{F}) &= \{(t, u) \in [0, T] \times D : \ \ \mathrm{F}_t(u) \ne \emptyset\}, \\ \mathrm{graph}(\mathrm{F}) &= \{(t, u, \xi) \in [0, T] \times D \times V^* \ : \ \ \xi \in \mathrm{F}_t(u)\} \end{aligned}$$

for the domain and the graph of the multivalued mapping $\mathrm{F} : [0, T] \times D \rightrightarrows V^*$, respectively. The basic measurability requirement on F is that graph(F) is a Borel set of $[0, T] \times V \times V^*$.

In the framework of the space $\mathbb{R}^m$, we will denote by $|\cdot|$ the Euclidean norm and by $B_r(0)$ the ball centered at 0 and of radius r. The symbol $\rightharpoonup$ will indicate weak convergence both in V and in V^*. Finally, throughout the paper we will use the symbols C and C' for various positive constants depending only on known quantities.

2. Analysis in a simplified setting

In this section, we deal with a single dissipation potential Ψ, independent of the state variable, and an energy functional $\mathcal{E} : [0, T] \times V \to (-\infty, +\infty]$ as in (E$_0$) with a *smooth* time-dependence (see for instance [32, §3], [36] for analogous assumptions within the analysis of abstract doubly nonlinear and rate-independent problems). The focus of this section is on the nonsmoothness and nonconvexity of the map $u \mapsto \mathcal{E}_t(u)$. We leave the questions arising from nonsmooth time-dependence and state-dependent dissipation potentials to later sections.

In the present framework, it is natural to work with the Fréchet subdifferential of the functionals $\mathcal{E}_t : V \to (-\infty, +\infty]$, defined in (1.8). Hence, we address the Cauchy problem

$$\partial\Psi(u'(t)) + \partial\mathcal{E}_t(u(t)) \ni 0 \quad \text{in } V^* \quad \text{for a.a. } t \in (0, T); \qquad u(0) = u_0, \tag{2.1}$$

which is a particular case of (1.1), with $\mathrm{F}_t(u) = \partial\mathcal{E}_t(u)$ for all $(t,u) \in [0,T]\times D$. In Section 2.1 we enucleate the abstract assumptions for Theorem 2.2 below that yield existence for problem (2.1). Next, we give an outline of its proof, highlighting the role played by the aforementioned assumptions. Then, in Section 2.2 we discuss sufficient conditions for the latter. We conclude with some PDE applications in Section 2.3.

2.1. **An existence result.**

Assumptions on the dissipation potential Ψ. Throughout this section we will suppose that

$$\Psi : V \to [0,+\infty) \ \text{ is l.s.c. and convex,} \tag{2.Ψ_1}$$

$$\Psi(0) = 0, \quad \lim_{\|v\|\uparrow+\infty} \frac{\Psi(v)}{\|v\|} = +\infty, \ \lim_{\|\xi\|_*\uparrow+\infty} \frac{\Psi^*(\xi)}{\|\xi\|_*} = +\infty, \quad \text{and} \tag{2.Ψ_2}$$

$$\forall\, w_1,\, w_2 \in \partial\Psi(v) : \qquad \Psi^*(w_1) = \Psi^*(w_2), \tag{2.Ψ_3}$$

where Ψ^* denotes the Fenchel-Moreau conjugate of Ψ. Hereafter, we will call any $\Psi : V \to [0,+\infty)$ complying with (2.Ψ_1)–(2.Ψ_3) an *admissible dissipation potential*.

We emphasize that, in this paper we only consider dissipation potentials Ψ with $\mathrm{dom}(\Psi) = V$. From this it follows (see, e.g., [23, Chap. I, Cor. 2.5]), that Ψ is continuous on V, and that Ψ^* has superlinear growth at infinity. Hence, the third of (2.Ψ_2) could be omitted, and has been stated here just for the sake of analogy with condition (4.Ψ_2) later on, for *state-dependent* dissipation potentials $\Psi = \Psi_u(v)$.

In fact, our analysis can be extended to the case in which $\mathrm{dom}(\Psi)$ is an open subset of V (i.e., it contains one continuity point). However, this rules out dissipation potentials enforcing irreversible evolution, like for example in damage models, see e.g. [37].

Remark 2.1. (1) We point out that, since $\Psi(0) = 0$, we have

$$\Psi^*(\xi) \geq 0 \qquad \text{for all } \xi \in V^*. \tag{2.2}$$

Furthermore, it follows from the superlinear growth of Ψ and Ψ^* that

$$\partial\Psi : V \rightrightarrows V^* \text{ is a bounded operator, and } \partial\Psi(v) \neq \emptyset \text{ for all } v \in V. \tag{2.3}$$

(2) Let us now get some further insight into condition (2.Ψ_3): a lower semicontinuous and convex potential $\Psi : V \to [0,+\infty)$ satisfies (2.Ψ_3) if and only if

$$\text{the mapping } \lambda \mapsto \Psi(\lambda v) \text{ is differentiable at } \lambda = 1. \tag{2.4}$$

Indeed, let $v \in V$ be such that $\partial\Psi(v) \neq \emptyset$. The convexity of Ψ gives

$$\liminf_{\lambda\downarrow 0} \frac{\Psi(v+\lambda v) - \Psi(v)}{\lambda} \geq \langle w, v\rangle \geq \limsup_{\lambda\uparrow 0} \frac{\Psi(v+\lambda v) - \Psi(v)}{\lambda}$$

for all $w \in \partial\Psi(v)$. Hence, (2.4) holds if and only if $\langle w_1, v\rangle = \langle w_2, v\rangle$ for all $w_1,\, w_2 \in \partial\Psi(v)$, which is obviously equivalent to (2.Ψ_3).

Therefore, condition (2.Ψ_3) is satisfied for example when Ψ is a linear combination of (positively) homogeneous, or differentiable, convex potentials.

Assumptions on the energy functional $\mathcal{E}$.

Coercivity:

$$\exists\, \tau_o > 0 \quad \forall\, t \in [0,T] : \quad u \mapsto \mathcal{E}_t(u) + \tau_o\Psi(u/\tau_o) \quad \text{has compact sublevels.} \tag{2.E_1}$$

Variational sum rule: If for some $u_o \in V$ and $\tau > 0$ the point $\bar u$ is a minimizer of $u \mapsto \mathcal{E}_t(u) + \tau\Psi((u-u_o)/\tau)$, then $\bar u$ satisfies the Euler-Lagrange equation $\partial\Psi((\bar u - u_o)/\tau) + \partial\mathcal{E}_t(\bar u) \ni 0$, viz.

$$\exists\, \xi \in \partial\mathcal{E}_t(\bar u) : \quad -\xi \in \partial\Psi((\bar u - u_o)/\tau). \tag{2.E_2}$$

Time-dependence:

$$\begin{aligned} &\forall\, u \in D : \ (t \mapsto \mathcal{E}_t(u)) \ \text{ is differentiable on } (0,T), \ \text{ with derivative } \ \partial_t\mathcal{E}_t(u); \\ &\exists\, C_1 > 0 \ \ \forall\, (t,u) \in [0,T]\times D : \quad |\partial_t\mathcal{E}_t(u)| \leq C_1\mathcal{E}_t(u). \end{aligned} \tag{2.E_3}$$

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

$$\sup_{t\in[0,T]} |\mathcal{E}_t(u(t))| < +\infty, \quad \xi(t) \in \partial\mathcal{E}_t(u(t)) \text{ for a.a. } t \in (0,T), \text{ and}$$
$$\int_0^T \Psi(u'(t))\,\mathrm{d}t < +\infty, \qquad \int_0^T \Psi^*(-\xi(t))\,\mathrm{d}t < +\infty, \tag{2.5}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) = \langle \xi(t), u'(t)\rangle + \partial_t\mathcal{E}_t(u(t)) \quad \text{for a.a.}\, t \in (0,T)\,. \tag{2.E$_4$}$$

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

$$\begin{aligned} &\text{if } u_n \to u \text{ in } V,\ \xi_n \in \partial\mathcal{E}_t(u_n),\ \xi_n \rightharpoonup \xi \text{ in } V^*,\ \partial_t\mathcal{E}_t(u_n) \to p,\ \mathcal{E}_t(u_n) \to \mathscr{E} \text{ in } \mathbb{R},\\ &\qquad\qquad \text{then} \quad \xi \in \partial\mathcal{E}_t(u), \quad p \le \partial_t\mathcal{E}_t(u), \quad \mathscr{E} = \mathcal{E}_t(u). \end{aligned} \tag{2.E$_5$}$$

A few comments on the above abstract conditions are in order:

(1) in Proposition 4.2 we are going to show that the variational sum rule (2.E$_2$) is indeed a consequence of the closedness property (2.E$_5$);
(2) in Section 2.2 we discuss sufficient conditions for (2.E$_5$) and the chain rule (2.E$_4$), showing in particular that they are valid if the functionals $\mathcal{E}_t(\cdot)$ are λ-convex;
(3) (2.E$_3$) and the Gronwall Lemma yield the following estimate

$$\mathcal{E}_t(u) \le \exp(C_1|t-s|)\mathcal{E}_s(u) \quad \text{for all } t,s \in [0,T],\ u \in D, \tag{2.6}$$

whence, in particular,

$$\mathcal{G}(u) \le \exp(C_1 T)\mathcal{E}_t(u) \quad \text{for all } t \in [0,T]. \tag{2.7}$$

Existence theorem and outline of the proof. We are now in the position to state the main result of this section.

Theorem 2.2. *Let us assume that* $(V, \mathcal{E}, \Psi, \partial\mathcal{E}, \partial_t\mathcal{E})$ *comply with* (2.Ψ_1)–(2.Ψ_3) *and* (E$_0$), (2.E$_1$)–(2.E$_4$). *Then, for every* $u_0 \in D$ *there exists a curve* $u \in \mathrm{AC}([0,T];V)$ *solving the Cauchy problem* (2.1). *In fact, there exists a function* $\xi \in L^1(0,T;V^*)$ *fulfilling*

$$\xi(t) \in \partial\mathcal{E}_t(u(t)) \cap (-\partial\Psi(u'(t))) \quad \textit{for a.a.}\, t \in (0,T),$$

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

$$\int_s^t \Big(\Psi(u'(r)) + \Psi^*(-\xi(r))\Big)\,\mathrm{d}r + \mathcal{E}_t(u(t)) = \mathcal{E}_s(u(s)) + \int_s^t \partial_t\mathcal{E}_r(u(r))\,\mathrm{d}r. \tag{2.8}$$

Theorem 2.2 is a direct consequence of the more general Theorem 4.4, which is proved in Section 6. Nonetheless, in order to provide some heuristics for conditions (2.E$_1$)–(2.E$_4$), in the following lines we enucleate the main steps of the proof, heavily simplifying most of the technical points and referring to Section 6 for all details.

Sketch of the proof. We split the proof in four steps.

Time-discretization. Following a well-established routine for gradient flows (cf., e.g., [17, 12, 6, 1, 47, 48, 41, 2, 44]), and in general doubly nonlinear equations [16, 15, 52, 51, 36], we approximate (2.1) with the *implicit Euler scheme*

$$U_\tau^0 := u_0, \quad \partial\Psi\left(\frac{U_\tau^n - U_\tau^{n-1}}{\tau}\right) + \partial\mathcal{E}_{t_n}(U_\tau^n) \ni 0 \quad n = 1, \ldots, N, \tag{2.9}$$

where $\tau = T/N$ is the time step, inducing a partition of $[0,T]$ with nodes $(t_n := n\tau)_{n=0}^N$. Since (2.9) is the Euler-Lagrange equation for the minimum problem

$$U_\tau^n \in \mathop{\mathrm{Argmin}}_{U\in D}\left\{\tau\Psi\left(\frac{U - U_\tau^{n-1}}{\tau}\right) + \mathcal{E}_{t_n}(U)\right\} \quad n = 1, \cdots, N,$$

we look for $(U_\tau^n)_{n=1}^N$ solving the above family of variational problems. Assumption (2.E$_1$) yields, via the *direct method* in the Calculus of Variations, the existence of solutions $(U_\tau^n)_{n=1}^N$. The variational sum rule ($2.\mathrm{E}_2$) ensures that for every U_τ^n fulfills (2.9) for all $n = 1, \dots, N$. Hence, we construct approximate solutions to (2.1) by introducing the left-continuous piecewise constant $(\overline{U}_\tau)_\tau$ and the piecewise linear $(U_\tau)_\tau$ interpolants of the discrete solutions $(U_\tau^n)_{n=1}^N$ (cf. Sec. 4.2), which clearly fulfill the approximate equation

$$\partial\Psi\left(U_\tau'(t)\right) + \partial\mathcal{E}_{\bar{\mathsf{t}}_\tau(t)}(\overline{U}_\tau(t)) \ni 0 \quad \text{for a.a. } t \in (0,T), \tag{2.10}$$

where $\bar{\mathsf{t}}_\tau$ is the left-continuous piecewise constant interpolant associated with the partition $(n\tau)_{n=0}^N$ of $(0,T)$.

Approximate energy inequality and a priori estimates. In the present nonconvex setting, (2.10) does not yield sufficient information to pass to the limit and conclude existence for (2.1). One needs the finer information provided by the *approximate energy inequality* involving the Fenchel-Moreau conjugate Ψ^* of Ψ, namely

$$\begin{aligned}\int_0^T \Psi\left(U_\tau'(t)\right) \mathrm{d}t + \int_0^T \Psi^*\big(-\widetilde{\xi}_\tau(t)\big) \mathrm{d}t + \mathcal{E}_T(\overline{U}_\tau(T)) \leq \mathcal{E}_0(u_0) + \int_0^T \partial_t\mathcal{E}_t(\widetilde{U}_\tau(t))\,\mathrm{d}t,\\ \text{with } \ \widetilde{\xi}_\tau(t) \in \partial\mathcal{E}_t(\widetilde{U}_\tau(t))) \ \text{ for a.a. } t \in (0,T).\end{aligned} \tag{2.11}$$

Here $\widetilde{U}_\tau$ is the so-called *De Giorgi variational interpolant* of the discrete solutions $(U_\tau^n)_{n=1}^N$ (see (4.12) for its definition, and Lemma 6.1 for its properties). Relying on the positivity of Ψ and of Ψ^* (cf. (2.2)), on ($2.\Psi_1$)–($2.\Psi_3$), on ($2.\mathrm{E}_1$), and on estimate ($2.\mathrm{E}_3$), from inequality (2.11) it is possible to deduce suitable a priori estimates on the sequences $(\overline{U}_\tau)$, (U_τ), $(\widetilde{U}_\tau)$, and $(\widetilde{\xi}_\tau)$. Then, one infers that, there exist $u \in \mathrm{AC}([0,T];V)$ and $\xi \in L^1(0,T;V^*)$ such that, up to a subsequence,

$$\begin{array}{ll}\overline{U}_\tau,\, U_\tau,\, \widetilde{U}_\tau \to u & \text{in } L^\infty(0,T;V),\\ U_\tau' \rightharpoonup u' & \text{in } L^1(0,T;V),\\ \widetilde{\xi}_\tau \rightharpoonup \xi & \text{in } L^1(0,T;V^*).\end{array}$$

Passage to the limit and proof of the upper energy estimate. Using that Ψ and Ψ^* are convex, it is possible to pass to the limit by lower semicontinuity in (2.11) and conclude that the functions u and ξ fulfill the *upper energy estimate*

$$\int_0^T \Psi(u'(t))\,\mathrm{d}t + \int_0^T \Psi^*(-\xi(t))\,\mathrm{d}t + \mathcal{E}_T(u(T)) \leq \mathcal{E}_0(u_0) + \int_0^T \partial_t\mathcal{E}_t(u(t))\,\mathrm{d}t\,. \tag{2.12}$$

Furthermore, the closedness property ($2.\mathrm{E}_5$) and an argument combining Young measures and measurable selection tools yield that there exists $\xi \in L^1(0,T;V^*)$ such that

$$\xi(t) \in \partial\mathcal{E}_t(u(t)) \ \text{ for a.a. } t \in (0,T). \tag{2.13}$$

Proof of the energy identity and conclusion. The chain rule ($2.\mathrm{E}_4$) entails

$$\mathcal{E}_0(u_0) + \int_0^T \partial_t\mathcal{E}_t(u(t))\,\mathrm{d}t = \mathcal{E}_T(u(T)) + \int_0^T \langle -\xi(t), u'(t)\rangle\,\mathrm{d}t.$$

Combining this with (2.12), we ultimately deduce

$$\int_0^T \big(\Psi(u'(t)) + \Psi^*(-\xi(t)) - \langle -\xi(t), u'(t)\rangle\big)\,\mathrm{d}t \leq 0.$$

Using the Fenchel-Young inequality $\Psi(v) + \Psi^*(w) \geq \langle w, v\rangle$, we arrive at

$$\Psi(u'(t)) + \Psi^*(-\xi(t)) - \langle -\xi(t), u'(t)\rangle = 0 \quad \text{for a.a. } t \in (0,T), \tag{2.14}$$

whence $-\xi(t) \in \partial\Psi(u'(t))$ for almost all $t \in (0,T)$. Combining this with (2.13), we conclude that u is a solution of (2.1). ∎

Remark 2.3 (Chain-rule inequality). In view of the discussion developed in Section 4, let us anticipate that the chain-rule condition ($2.\mathrm{E}_4$) could be weakened. In fact, a close perusal at the argument for the proof of Theorem 2.2 reveals that, it is sufficient to require the chain-rule inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t))\rangle \geq \langle \xi(t), u'(t)\rangle + \partial_t\mathcal{E}_t(u(t)) \quad \text{for a.a. } \ t \in (0,T)\,. \tag{2.15}$$

Indeed, (2.15) yields the *lower energy estimate*

$$\mathcal{E}_T(u(T)) - \mathcal{E}_0(u_0) \geq \int_0^T \langle \xi(t), u'(t) \rangle \,\mathrm{d}t + \int_0^T \partial_t \mathcal{E}_t(u(t)) \,\mathrm{d}t,$$

which, combined with the *upper energy estimate* (2.12), ultimately yields (2.14). In turn, the *upper energy estimate* (2.12) is a consequence of the time-discretization scheme (in particular, of the approximate energy inequality (2.11)), and of classical lower semicontinuity results.

2.2. **Sufficient conditions for closedness and chain rule.** In this section we revisit the abstract assumptions of Theorem 2.2, and in particular we provide sufficient conditions of λ-convexity type on the energy functional $\mathcal{E}$ for the closedness property (2.E$_5$) and for the chain rule (2.E$_4$).

Throughout the following discussion, we will suppose that $\mathcal{E}$ complies with the time-dependence condition (2.E$_3$).

Uniformly (Fréchet-)subdifferentiable functionals. A first sufficient condition for (2.E$_5$) and (2.E$_4$) is some sort of *uniform subdifferentiability* of the functionals $\mathcal{E}_t$ on their sublevels, cf. (2.16) below.

For every $R > 0$ we set

$$D_R = \{u \in D \,:\, \mathcal{G}(u) \leq R\} \,.$$

In view of (2.7), every $u \in D$ satisfies $u \in D_R$ for some $R > 0$.

Proposition 2.4. *Let* $\mathcal{E} : [0, T] \times V \to (-\infty, +\infty]$ *be a family of time-dependent functionals as in* (E$_0$)*, complying with* (2.E$_3$)*. Moreover assume that for all* $R > 0$ *there exists a* modulus of subdifferentiability $\omega^R : [0, T] \times D_R \times D_R \to [0, +\infty)$ *such that for all* $t \in [0, T]$*:*

$$\begin{gathered} \omega_t^R(u, u) = 0 \textit{ for every } u \in D_R, \\ \textit{the map } (t, u, v) \mapsto \omega_t^R(u, v) \textit{ is upper semicontinuous, and} \\ \mathcal{E}_t(v) - \mathcal{E}_t(u) - \langle \xi, v - u \rangle \geq -\omega_t^R(u, v)\|v - u\| \quad \textit{for all } u,\, v \in D_R \textit{ and } \xi \in \partial\mathcal{E}_t(u). \end{gathered} \tag{2.16}$$

Then, $\mathcal{E}$ *complies with the closedness condition* (2.E$_5$) *and with the chain rule* (2.E$_4$)*.*

Proof: Ad (2.E$_5$)**.** Let $v \in D$ be fixed, and let (u_n), u, (ξ_n) and ξ be like in (2.E$_5$). It follows from estimate (2.7) that there exists some $R > 0$ such that v, $u_n \in D_R$ for all $n \in \mathbb{N}$. Thanks to (2.16), we have

$$\mathcal{E}_t(v) - \mathcal{E}_t(u_n) - \langle \xi_n, v - u_n \rangle \geq -\omega_t^R(u_n, v)\|v - u_n\| \quad \text{for all } n \in \mathbb{N}. \tag{2.17}$$

Since the functionals $\mathcal{E}_t$ and ω_t^R are respectively lower and upper semicontinuous, we can pass to the limit in (2.17), obtaining

$$\mathcal{E}_t(v) - \mathcal{E}_t(u) - \langle \xi, v - u \rangle \geq -\omega_t^R(u, v)\|v - u\| \,.$$

It is not difficult to check that this inequality yields $\xi \in \partial\mathcal{E}_t(u)$: indeed, notice that, in the definition (1.8) of Fréchet subdifferential, it is not restrictive to consider sequences $(v_k)_k$ converging to u, such that $\limsup_{k\to\infty} \mathcal{E}_t(v_k) \leq \mathcal{E}_t(u)$. Furthermore, choosing $v = u$ in (2.17) (notice that $u \in D_R$ by lower semicontinuity), and exploiting the properties of ω_t^R, we have the following chain of inequalities

$$0 \leq \limsup_{n\to\infty} \left(\mathcal{E}_t(u_n) - \mathcal{E}_t(u)\right) \leq \lim_{n\to\infty} \langle \xi_n, u_n - u \rangle + \limsup_{n\to\infty} \omega_t^R(u_n, u)\|u - u_n\| = 0,$$

whence $\mathcal{E}_t(u_n) \to \mathcal{E}_t(u)$, which concludes the proof of (2.E$_5$).

Ad (2.E$_4$)**.** Let $u \in \mathrm{AC}([0, T]; V)$ and $\xi \in L^1(0, T; V^*)$ fulfill (2.5). Up to a suitable reparametrization (cf. [2, Lemma 1.1.4]), it is possible to assume that the curve u is 1-Lipschitz. Furthermore, due to $\sup_{t\in[0,T]} \mathcal{E}_t(u(t)) < +\infty$ there exists $R > 0$ such that $u(t) \in D_R$ for all $t \in [0, T]$. In order to show the absolute continuity of the map $t \mapsto \mathcal{E}_t(u(t))$, we estimate for $0 \leq s \leq t \leq T$ the difference

$$\mathcal{E}_t(u(t)) - \mathcal{E}_s(u(s)) = \mathcal{E}_t(u(t)) - \mathcal{E}_s(u(t)) + \mathcal{E}_s(u(t)) - \mathcal{E}_s(u(s)).$$

First, due to (2.E$_3$), for all $0 \leq s \leq t \leq T$ there holds

$$\mathcal{E}_t(u(t)) - \mathcal{E}_s(u(t)) = \int_s^t \partial_t \mathcal{E}_r(u(t)) \,\mathrm{d}r \leq C_1 \mathcal{G}(u(t))(t - s) \leq C_1 R(t - s). \tag{2.18}$$

Second, in view of (2.16) we have

$$\mathcal{E}_s(u(t)) - \mathcal{E}_s(u(s)) \geq \langle \xi(s), u(t) - u(s) \rangle - \omega_s^R(u(s), u(t)) \|u(t) - u(s)\|. \tag{2.19}$$

Combining (2.18) and (2.19), inverting the role of s and t, and using the 1-Lipschitz continuity of u, we conclude

$$|\mathcal{E}_t(u(t)) - \mathcal{E}_s(u(s))| \leq (C_1 R + \|\xi(t)\|_* + \|\xi(s)\|_* + \omega_t^R(u(s), u(t)) + \omega_s^R(u(s), u(t)))|t - s|. \tag{2.20}$$

Now, the upper semicontinuity of $(t, u, v) \mapsto \omega_t^R(u, v)$, joint with the fact that $u \in \mathrm{C}^0([0,T]; V)$, yield that there exists $C > 0$ such that there holds $\omega_t^R(u(s), u(t))$, $\omega_s^R(u(s), u(t)) \leq C$ for all s, $t \in [0,T]$. Therefore, arguing as in [2, Thm. 1.2.5], we conclude that $t \mapsto \mathcal{E}_t(u(t))$ is absolutely continuous. Let us now fix a point $t \in (0,T)$ such that $u'(t)$, $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t))$ exist: arguing throughout (2.18)–(2.19), it is not difficult to deduce that

$$\mathcal{E}_{t+h}(u(t+h)) - \mathcal{E}_t(u(t)) \geq \int_t^{t+h} \partial_t \mathcal{E}_r(u(t))\,\mathrm{d}r + \langle \xi(t), u(t+h) - u(t) \rangle - \omega_t^R(u(t), u(t+h)) \|u(t+h) - u(t)\|.$$

Dividing by $h > 0$ and $h < 0$ and taking the limit as $h \downarrow 0$ and $h \uparrow 0$, we prove the chain rule (2.E$_4$). ■

Remark 2.5 (λ-convex functionals). A sufficient condition yielding (2.16) in the case of an energy functional $\mathcal{E} : [0,T] \times V \to (-\infty, +\infty]$ complying with (E$_0$) and (2.E$_3$), is that the map $u \mapsto \mathcal{E}_t(u)$ is λ-*convex* uniformly in $t \in [0,T]$, namely

$$\begin{aligned} &\exists\, \lambda \in \mathbb{R} \quad \forall\, t \in [0,T]\ \forall\, u_0,\ u_1 \in D\ \forall\, \theta \in [0,1]\ : \\ &\qquad \mathcal{E}_t((1-\theta)u_0 + \theta u_1) \leq (1-\theta)\mathcal{E}_t(u_0) + \theta \mathcal{E}_t(u_1) - \frac{\lambda}{2}\theta(1-\theta)\|u_0 - u_1\|^2 . \end{aligned} \tag{2.21}$$

Indeed, given u, $v \in D$, (2.21) and the very definition (1.8) of Fréchet subdifferential yield for any $\xi \in \partial\mathcal{E}_t(u)$ and $\theta \downarrow 0$

$$\begin{aligned} \theta\left(\mathcal{E}_t(v) - \mathcal{E}_t(u)\right) &\geq \frac{\lambda}{2}\theta(1-\theta)\|v-u\|^2 + \mathcal{E}_t((1-\theta)u + \theta v) - \mathcal{E}_t(u) \\ &\geq \theta\left(\frac{\lambda}{2}(1-\theta)\|v-u\|^2 + \langle \xi, v - u \rangle + o(1)\right). \end{aligned}$$

Upon diving by θ, we conclude inequality (2.16) with the choice $\omega_t(u, v) := \frac{\lambda^-}{2}\|u - v\|$.

Remark 2.6 (Perturbations of λ-convex functionals). In [43] a broad family of time-dependent energies, which for instance encompasses λ-convex functionals, was tackled. However, as hinted in the Introduction, [43] focuses on the analysis (from a metric viewpoint) of doubly nonlinear equations driven by a less general class of dissipation potentials than those considered in the present paper. While referring to [43] for details, here we recall that the energies therein considered are given by the sum of two time-dependent functionals $\mathcal{E}^1$, $\mathcal{E}^2 : [0,T] \times V \to (-\infty, +\infty]$, such that the functionals $\mathcal{E}_t^1$ are λ-convex, uniformly with respect to $t \in [0,T]$, and the functionals $\mathcal{E}_t^2$ are *dominated concave perturbations* of $\mathcal{E}_t^1$ (cf. [43, Definitions 5.4, 5.10], as well as [44] for an analogous class of functionals). In [43, Propositions 5.6, 5.10] it was shown that, under the above conditions, the energy $\mathcal{E} = \mathcal{E}^1 + \mathcal{E}^2$ fulfills the closedness condition (2.E$_5$) and the chain rule (2.E$_4$).

2.3. Examples and applications.

Example 2.7 (A model in ferro-magnetism)**.** We take $\Omega \subset \mathbb{R}^3$ a bounded sufficiently smooth domain, and let

$$V = L^2(\Omega; \mathbb{R}^3) \qquad \text{and} \quad \Psi(v) = \int_\Omega |v| + \frac{1}{2}|v|^2\,\mathrm{d}x = \|v\|_{L^1(\Omega;\mathbb{R}^3)} + \frac{1}{2}\|v\|^2_{L^2(\Omega;\mathbb{R}^3)}.$$

We consider a simplified model for ferro-magnetism, in which the interplay between the elastic and the magnetic effects is neglected (see [32, Sec. 7.4] for a rate-independent model accounting for both features). In this framework, the relevant energy functional $\mathcal{E} : [0,T] \times L^2(\Omega;\mathbb{R}^3) \to (-\infty,+\infty]$ is given by

$$\mathcal{E}_t(m) = \begin{cases} \int_\Omega \left(\frac{\alpha}{2}|\nabla m|^2 + W(m)\right) \mathrm{d}x + \int_{\mathbb{R}^3} |\nabla \Phi_m|^2 \,\mathrm{d}x - \langle H_{\mathrm{ext}}(t), m\rangle_{H^1} & \text{if } m \in H^1(\Omega;\mathbb{R}^3), \\ +\infty & \text{otherwise.} \end{cases} \tag{2.22}$$

Here, $\langle \cdot,\cdot\rangle_{H^1}$ is a short-hand notation for the duality pairing between $H^1(\Omega;\mathbb{R}^3)^*$ and $H^1(\Omega;\mathbb{R}^3)$, $m : \Omega \to \mathbb{R}^3$ is the magnetization,

$$\begin{gathered} W \in \mathrm{C}^1(\mathbb{R}^3;\mathbb{R}) \text{ a } \lambda_W\text{-convex potential for some } \lambda_W \in \mathbb{R}, \text{ such that} \\ \exists\, c_W,\, C_W > 0 \;\forall\, m \in \mathbb{R}^3 \;:\; \; W(r) \geq c_W|m|^2 - C_W \end{gathered} \tag{2.23}$$

(e.g., $W(m) = (1-|m|^2)^2$), and the external magnetic field H_{ext} fulfills

$$H_{\mathrm{ext}} \in \mathrm{C}^1([0,T]; H^1(\Omega;\mathbb{R}^3)^*), \qquad \mathrm{div}(H_{\mathrm{ext}}(t)) \equiv 0, \tag{2.24}$$

the latter equation meaning that $\int_\Omega H_{\mathrm{ext}}(t)\cdot \nabla v \,\mathrm{d}x = 0$ for all $v \in H^1(\Omega;\mathbb{R}^3)$. In (2.22), the potential Φ_m describes the field induced by the magnetization inside the body. Hence, the magnetic flux is $J = (H_{\mathrm{ext}} - \nabla\Phi_m + \mathbb{E}_\Omega(m))$, $\mathbb{E}_\Omega(m)$ denoting the trivial extension of m to all of Ω by 0. Thus, $\mathrm{div}\, J = 0$ and (2.24) yield that Φ_m is the solution of

$$\mathrm{div}\,(-\nabla\Phi_m + \mathbb{E}_\Omega(m)) = 0 \quad \text{in } \mathbb{R}^3.$$

Note that the operator $\mathcal{J} : L^2(\Omega;\mathbb{R}^3) \to L^2(\mathbb{R}^3;\mathbb{R}^3)$ mapping $m \mapsto \nabla\Phi_m$ is bounded and linear; it was proved in [20] that $m \mapsto \mathcal{J}(m)_{|\Omega}$ is an orthogonal projection on $L^2(\Omega;\mathbb{R}^3)$, satisfying

$$\int_{\mathbb{R}^3} |\nabla\Phi_m|^2 \,\mathrm{d}x = \int_\Omega m \cdot \mathcal{J}(m)_{|\Omega} \,\mathrm{d}x \qquad \text{for all } m \in L^2(\Omega;\mathbb{R}^3)\,. \tag{2.25}$$

One can see that for all $(t,m) \in [0,T] \times H^1(\Omega;\mathbb{R}^3)$ such that $\partial\mathcal{E}_t(m) \neq \emptyset$ there holds

$$\partial\mathcal{E}_t(m) = \left\{-\Delta m + \mathrm{D}W(m) + \mathcal{J}(m)_{|\Omega}\right\}\,.$$

Therefore, with the present choices of Ψ and $\mathcal{E}$, the Cauchy problem (2.1) translates into

$$\mathrm{Sign}(\dot m) + \dot m - \alpha\Delta m + \mathrm{D}W(m) + \mathcal{J}(m)_{|\Omega} = H_{\mathrm{ext}} \quad \text{a.e. in } \Omega\times(0,T); \;\; m(0) = m_0, \tag{2.26}$$

with variational boundary conditions; here

$$\mathrm{Sign} : \mathbb{R}^3 \rightrightarrows \mathbb{R}^3 \quad \text{is given by} \quad \mathrm{Sign}(v) = \begin{cases} \dfrac{v}{|v|} & \text{if } v \neq 0, \\ \overline{B_1(0)} & \text{if } v = 0. \end{cases}$$

Now, combining (2.23)–(2.24), it is easy to see that $\mathcal{E}$ complies with (2.E$_1$) and (2.E$_3$). Also using (2.25) and arguing as in [43, Sec. 7.2], we further check that for some $\lambda \in \mathbb{R}$ the energy $\mathcal{E}$ is λ-convex (uniformly in $t \in [0,T]$), with respect to the $L^2(\Omega;\mathbb{R}^3)$-norm. Therefore, the closedness property (2.E$_5$) and, a fortiori, the variational sum rule (2.E$_2$) (cf. Proposition 4.2) hold, as well as the chain rule (2.E$_4$). Thus, it follows from Theorem 2.2 that the Cauchy problem (2.26) admits a solution.

Example 2.8 (Doubly nonlinear evolutions of Allen-Cahn type)**.** Let us consider the following class of evolution equations

$$\varrho\mathrm{Sign}(\dot u) + |\dot u|^{p-2}\dot u - \mathrm{div}(\beta(\nabla u)) + W'(u) = \ell \quad \text{in } \Omega\times(0,T), \tag{2.27}$$

with $\varrho > 0$, $1 < p < \infty$, and $u : [0,T]\times\Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^d$, $d \geq 1$, is a sufficiently smooth bounded domain. In (2.27), $\beta : \mathbb{R}^d \to \mathbb{R}^d$ is the gradient of some smooth function j on $\mathbb{R}^d$, $W : \mathbb{R} \to \mathbb{R}$ a differentiable function and $\ell : \Omega\times(0,T) \to \mathbb{R}$ some source term. To fix ideas (cf. [43, Sec. 8.2] for the precise statement of the assumptions on j and W), we may think of the case in which $j(\zeta) = \frac{1}{q}|\zeta|^q$ for some $q > 1$ (hence $\beta(\zeta) = |\zeta|^{q-2}\zeta$ and the elliptic operator in (2.27) is indeed the q-Laplacian), and W is given by the sum of a convex function, perturbed by a nonconvex nonlinearity which complies with suitable growth conditions (like for instance in the classical, double-well potential case $W(u) := (u^2-1)^2/4$).

We supplement equation (2.27) with homogeneous Dirichlet boundary conditions, and notice that this boundary-value problem can be written in the abstract form (2.1), in the framework of the ambient space

$$V = L^p(\Omega), \text{ with the dissipation potential } \ \Psi(v) = \varrho\,\|v\|_{L^1(\Omega)} + \frac{1}{p}\|v\|^p_{L^p(\Omega)} \tag{2.28}$$

and the driving energy functional

$$\mathcal{E}_t(u) = \begin{cases} \int_\Omega \left(j(\nabla u(x)) + W(u(x))\right) \mathrm{d}x - \ \langle \ell(t), u\rangle_{W^{1,q}_0} & \text{if } u \in W^{1,q}_0(\Omega),\ W(u) \in L^1(\Omega), \\ +\infty & \text{otherwise,} \end{cases} \tag{2.29}$$

where $\langle \cdot, \cdot \rangle_{W^{1,q}_0}$ stands for the duality pairing between $W^{-1,q'}(\Omega)$ and $W^{1,q}_0(\Omega)$, with $q' = q/(q-1)$.

In [43, Sec. 8.2], under suitable conditions on j and W the existence of a solution for the initial-boundary value problem for (2.27) was proved for $\varrho = 0$. Namely, in [43] only the case of a dissipation potential Ψ induced by the *single norm* $\|\cdot\|_{L^p(\Omega)}$ was considered, which does not include the more physical form (2.28).

Relying on the analysis of [43], it can be checked that, if $\ell \in \mathrm{C}^1([0,T]; W^{-1,q'}(\Omega))$, then the energy functional $\mathcal{E}$ (2.29) complies with (2.E$_1$)–(2.E$_5$). Hence, Theorem 2.2 applies, yielding the existence of a solution $u \in L^\infty(0,T; W^{1,q}_0(\Omega)) \cap W^{1,p}(0,T; L^p(\Omega))$ to the initial-boundary value problem for (2.27).

3. Motivating examples for marginal subdifferentials

In this section, we restrict to a finite-dimensional setting and give an outlook to a twofold generalization of the set-up considered in Section 2. Such an extension is motivated by the analysis of abstract evolutionary systems of the form

$$\left.\begin{array}{ll} \partial\Psi(u'(t)) + \partial\mathcal{E}^1(u(t)) + \mathrm{D}_u I_t(\eta(t), u(t)) \ni \ell(t) & \text{in } V^* \\ \mathrm{D}_\eta I_t(\eta(t), u(t)) = 0 & \text{in } X^* \end{array}\right\} \quad \text{for a.a.}\, t \in (0,T), \tag{3.1}$$

where $\mathcal{E}^1 : V \to (-\infty, +\infty]$ is a convex energy, perturbed by some smooth functional $I : [0,T] \times X \times V \to \mathbb{R}$ (where X is a second Banach space), and $\ell : [0,T] \to V^*$ is the external loading. Couplings like (3.1) arise in the modeling of physical systems described in terms of two variables (η, u), such that energy dissipation only occurs through the *internal variable* u, and η fulfills some stationary law. PDE systems of the type (3.1) typically arise in connection with rate-independent behavior (cf. [32] and the references therein). Nonetheless, they can also occur in the modeling of *rate-dependent* evolutions, like for instance in the case of *quasistationary phase-field models*, cf. [29, 50, 55, 44]. In Section 5 later on, we analyze a PDE system of the type (3.1) in finite-strain elasticity.

Let us observe that the second stationary relation in (3.1) is the Euler-Lagrange equation for the minimum problem $\inf_{\eta \in X} I_t(u(t), \eta)$, and suppose that

$$M(t,u) := \operatorname*{Argmin}_{\eta \in X} I_t(\eta, u) \neq \emptyset \quad \text{for all } (t,u) \in [0,T] \times V.$$

Hence, we introduce the *reduced energy* functionals $\mathcal{E}^2$, $\mathcal{E} : [0,T] \times V \to (-\infty, +\infty]$

$$\begin{cases} \mathcal{E}^2_t(u) := \min_{\eta \in X} I_t(\eta, u), \\ \mathcal{E}_t(u) := \mathcal{E}^1(u) + \mathcal{E}^2_t(u) - \langle \ell(t), u \rangle = \min_{\eta \in X} \left(\mathcal{E}^1(u) + I_t(\eta, u) - \langle \ell(t), u\rangle \right). \end{cases} \tag{3.2}$$

In this setting, it is natural to introduce the following subdifferential notion for energy $\mathcal{E}^2$, tailored to its reduced form.

Definition 3.1 (Marginal subdifferential)**.** The *marginal subdifferential* (with respect to the variable u) of the reduced functional $\mathcal{E}^2 : [0,T] \times V \to (-\infty, +\infty]$ at $(t,u) \in [0,T] \times V$ is

$$\widehat{\partial}\mathcal{E}^2_t(u) := \{\mathrm{D}_u I_t(\eta, u) \ : \ \eta \in M(t,u)\}. \tag{3.3}$$

Hereafter, we will address the doubly nonlinear evolution equation

$$\partial\Psi(u'(t)) + \partial\mathcal{E}^1(u(t)) + \widehat{\partial}\mathcal{E}_t^2(u(t)) \ni \ell(t) \quad \text{in } V^* \quad \text{for a.a.}\, t \in (0,T). \tag{3.4}$$

Clearly, solutions to (3.4) are in fact solutions to the quasi-stationary evolution system (3.1).

Notice that (3.4) may be viewed as a generalization of the doubly nonlinear equations, featuring the Fréchet subdifferential, examined in Section 2. Indeed, under quite standard assumptions there holds

$$\partial\mathcal{E}_t(u) \subset \partial\mathcal{E}^1(u) + \widehat{\partial}\mathcal{E}_t^2(u) - \ell(t) \quad \text{for all } (t,u) \in [0,T]\times V, \tag{3.5}$$

while the converse inclusion is not true, in general.

This fact is illustrated in Example 3.2: for a specific choice of the functional $\mathcal{E}^1$ and for a *time-independent* marginal functional $\mathcal{E}^2$, it is shown that the Fréchet subdifferential of the energy $\mathcal{E}_t(u) = \mathcal{E}^1(u) + \mathcal{E}^2(u) - \langle \ell(t), u\rangle$ in (3.2) does not comply with the closedness property (2.E$_5$). Furthermore, the closure of $\partial\mathcal{E}$ in the sense of graphs coincides for all $(t,u) \in [0,T]\times V$ with the set $\partial\mathcal{E}^1(u) + \widehat{\partial}\mathcal{E}^2(u) - \ell(t)$.

Another important feature which sets aside reduced energy functionals from the class of energies examined in Section 2 is that, even if the function $t \mapsto I_t(\eta,u)$ is smooth, the resulting reduced functionals $\mathcal{E}^2$ and $\mathcal{E}$ (cf. (3.2)) may be nonsmooth with respect to time, see Examples 3.3 and 3.4. Therein, we suggest the usage of a *generalized time-derivative*, defined in such a way as to comply with a suitable chain-rule inequality.

Example 3.2. For simplicity we restrict to the one-dimensional case $V = X = \mathbb{R}$, and to gradient flows, hence taking $\Psi(v) = \frac{1}{2}|v|^2$. As in [44, Ex. 2], we choose

$$\mathcal{E}^1(u) = \frac{1}{2}|u|^2, \quad I_t(\eta,u) = I(\eta,u) = \frac{1}{2}|\eta|^2 - u\eta + \mathcal{W}(\eta),$$

where $\mathcal{W} : \mathbb{R} \to \mathbb{R}$ is the (piecewise quadratic) double well potential

$$\mathcal{W}(\eta) := \begin{cases} (\eta+1)^2 & \eta < -\frac{1}{2}, \\ -\eta^2 + \frac{1}{2} & |\eta| \le \frac{1}{2}, \\ (\eta-1)^2 & \eta > \frac{1}{2}, \end{cases} \quad \text{with derivative } \mathcal{W}'(\eta) = \begin{cases} 2(\eta+1) & \eta < -\frac{1}{2}, \\ -2\eta & |\eta| < \frac{1}{2}, \\ 2(\eta-1) & \eta > \frac{1}{2}. \end{cases} \tag{3.6}$$

In this setting, given some smooth external loading $\ell : [0,T] \to \mathbb{R}$, the coupled system (3.1) reads

$$\begin{cases} u'(t) + u(t) - \eta(t) = \ell(t), \\ \mathcal{W}'(\eta(t)) + \eta(t) = u(t) \end{cases} \quad \text{for a.a.}\, t \in (0,T),$$

which may be viewed as the one-dimensional caricature of the *quasistationary phase-field system* (cf. [29, 50, 55]).

It was observed in [44, Sec. 2.1] that the Fréchet subdifferential of the energy functional $\mathcal{E}$ defined in (3.2) satisfies (3.5). More precisely,

$$\partial\mathcal{E}_t(u) \neq \emptyset \;\Rightarrow\; \begin{cases} \partial\mathcal{E}_t(u) \text{ and } M(t,u) \text{ reduce to a singleton, and} \\ \partial\mathcal{E}_t(u) = \partial\mathcal{E}^1(u) + \widehat{\partial}\mathcal{E}^2(u) - \ell(t) = \partial\mathcal{E}^1(u) - M(t,u) - \ell(t). \end{cases} \tag{3.7}$$

Now, since the subdifferential mapping $\partial\mathcal{E}_t : \mathbb{R} \rightrightarrows \mathbb{R}$ is not closed in the sense of graphs, it is natural to introduce its closure, i.e. the *limiting subdifferential* (cf. [39, 40], and [44, 45, 46] for some analysis of gradient flow and doubly nonlinear equations featuring such a notion of subdifferential), defined by

$$\partial_{\lim}\mathcal{E}_t(u) := \{\xi \in \mathbb{R} \,:\, \exists\, (u_n),\, (\xi_n) \subset \mathbb{R},\; u_n \to u,\; \xi_n \to \xi,\; \mathcal{E}_t(u_n) \to \mathcal{E}_t(u)\}.$$

From the closedness of the graph of the multivalued mapping $M(t,\cdot) : \mathbb{R} \rightrightarrows \mathbb{R}$ we infer that a weaker form of (3.7) passes to the limit, i.e.

$$\partial_{\lim}\mathcal{E}_t(u) \subset \partial\mathcal{E}^1(u) - M(t,u) - \ell(t) = \partial\mathcal{E}^1(u) + \widehat{\partial}\mathcal{E}^2(u) - \ell(t) \quad \text{for every } (t,u) \in [0,T]\times\mathbb{R}. \tag{3.8}$$

In fact, in the case of (3.6) we even have $\partial_{\lim}\mathcal{E}_t(u) = \partial\mathcal{E}^1(u) + \widehat{\partial}\mathcal{E}^2(u) - \ell(t)$. Relations (3.7) and (3.8) suggest the choice of the subdifferential notion $\mathrm{F}_t(u) := \partial\mathcal{E}^1(u) + \widehat{\partial}\mathcal{E}^2(u) - \ell(t)$ for reduced energies of the type (3.2). We explore this viewpoint in Section 5.

Example 3.3. We take $V = \mathbb{R}$, $\Psi(v) = \frac{1}{2}|v|^2$ and, we set

$$\mathcal{E}_t(u) = -\alpha|u - \beta t| \qquad \text{for all } u \in \mathbb{R},\ t \in (0,T), \text{ with } \alpha > \beta > 0,\ \beta < 1\,.$$

Note that $\mathcal{E}$ is a marginal function: indeed,

$$\mathcal{E}_t(u) = \min_{\eta \in \{0,1\}} I_t(\eta, u), \quad \text{with} \quad I_t(\eta,u) = \begin{cases} -\alpha u + \alpha\beta t & \text{if } \eta = 0, \\ \alpha u - \alpha\beta t & \text{if } \eta = 1. \end{cases}$$

In this case, $\mathcal{E}$ does not comply with the smooth time-dependence condition (2.E_3), and it is only Lipschitz continuous with respect to both variables t and u. It is then natural to consider the *Clarke* subdifferentials of the energy $\mathcal{E}$ with respect to u and t, which are easily calculated:

$$\partial_u^{\mathrm{Clarke}}\mathcal{E}_t(u) = \begin{cases} \{-\alpha\} & \text{if } u > \beta t, \\ [-\alpha, \alpha] & \text{if } u = \beta t, \\ \{\alpha\} & \text{if } u < \beta t, \end{cases} \qquad \partial_t^{\mathrm{Clarke}}\mathcal{E}_t(u) = \begin{cases} \{\alpha\beta\} & \text{if } u > \beta t, \\ [-\alpha\beta, \alpha\beta] & \text{if } u = \beta t, \\ \{-\alpha\beta\} & \text{if } u < \beta t. \end{cases} \tag{3.9}$$

Notice that $\partial\mathcal{E}_t(u) \subset \partial_u^{\mathrm{Clarke}}\mathcal{E}_t(u)$ for all $(t,u) \in [0,T] \times \mathbb{R}$. Furthermore, the multivalued mapping $\partial_u^{\mathrm{Clarke}}\mathcal{E}_t : \mathbb{R} \rightrightarrows \mathbb{R}$ is closed in the sense of graphs. We may choose $\mathrm{F}_t(u) := \partial_u^{\mathrm{Clarke}}\mathcal{E}_t(u)$ and consider the gradient flow

$$u'(t) + \partial_u^{\mathrm{Clarke}}\mathcal{E}_t(u(t)) \ni 0 \qquad \text{for a.a. } t \in (0,T)\,. \tag{3.10}$$

We immediately verify that the curve $\bar{u} : [0,T] \to \mathbb{R}$ defined by $\bar{u}(t) = \beta t$ is a solution of (3.10). Now, we aim to get some insight into a possible surrogate notion of chain rule in this nonsmooth setting. Imposing that the chain-rule inequality (2.15) holds along the curve $\bar{u}$, with the Clarke subdifferentials (3.9), we arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(\bar{u}(t)) \geq \xi \bar{u}'(t) + p \ \text{ for all } \xi \in F_t(\bar{u}(t)),\ p \in \partial_t^{\mathrm{Clarke}}\mathcal{E}_t(\bar{u}(t)), \text{ and for a.a. } t \in (0,T). \tag{3.11}$$

Since $\mathcal{E}_t(\bar{u}(t)) \equiv 0$ and $\bar{u}'(t) \equiv \beta$, this amounts to checking if

$$0 \geq \xi\beta + p \qquad \text{for all } \xi \in [-\alpha, \alpha],\ p \in [-\alpha\beta, \alpha\beta],$$

which does not hold.

However, it is true that for a fixed $\xi \in F_t(\bar{u}(t))$ there exists a set $\widehat{\mathrm{P}}_t(\bar{u}(t), \xi)$ such that inequality (3.11) holds for ξ and all elements $p \in \widehat{\mathrm{P}}_t(\bar{u}(t), \xi)$, namely

$$\widehat{\mathrm{P}}_t(\bar{u}(t), \xi) = [-\alpha\beta, -\xi\beta]\,. \tag{3.12}$$

Finally, we may observe that, if we ask for an equality sign in (3.11) for a fixed $\xi \in F_t(\bar{u}(t))$, then the corresponding set $\widehat{\mathrm{P}}_t(\bar{u}(t), \xi)$ reduces to the singleton $\{-\xi\beta\}$, cf. also Remarks 4.1 and 4.5 for further related comments.

Example 3.4. We reconsider the triple $(V, \mathcal{E}_t(u), \Psi) = (\mathbb{R}, \min_{\eta \in \{0,1\}} I_t(\eta, u), \frac{1}{2}|\cdot|^2)$ of Example 3.3, but choose for F the marginal subdifferential of $\mathcal{E}$ (cf. Definition 3.1), viz.

$$\mathrm{F}_t(u) = \widehat{\partial}\mathcal{E}_t(u) = \begin{cases} \{-\alpha\} & \text{if } u > \beta t \\ \{-\alpha, \alpha\} & \text{if } u = \beta t \\ \{\alpha\} & \text{if } u < \beta t. \end{cases}$$

Notice that, in this case, the curve $\bar{u} : [0,T] \to \mathbb{R}$ defined by $\bar{u}(t) = \beta t$ is not a solution of the gradient flow $u'(t) + \widehat{\partial}\mathcal{E}_t(u(t)) \ni 0$ on $(0,T)$. Imposing that the chain-rule inequality (2.15) holds along the curve $\bar{u}$, for the marginal subdifferential with respect to u and the Clarke subdifferential with respect to t, yields

$$0 \geq \xi\beta + p \qquad \text{for all } \xi \in \{-\alpha, \alpha\},\ p \in [-\alpha\beta, \alpha\beta].$$

Thus, referring to notation (3.12), we conclude that, in this case,

$$\widehat{\mathrm{P}}_t(\bar{u}(t), \xi) = [-\alpha\beta, -\xi\beta] = \begin{cases} \{-\alpha\beta\} & \text{if } \xi = \alpha \\ [-\alpha\beta, \alpha\beta] & \text{if } \xi = -\alpha. \end{cases}$$

Examples 3.3 and 3.4 seem to suggest that, to deal with marginal functions, one should use a notion of time-derivative P *conditioned, via the chain rule, to elements* ξ *of the subdifferential.* This means that, in addition to the (t,u)-dependence, such a notion P also depends on the elements $\xi \in \mathrm{F}_t(u)$. This is the point of view we are going to adopt in what follows.

4. Main results

4.1. **Assumptions.** We recall that V is a reflexive separable Banach space. Below we enlist our general assumptions on the *state-dependent* dissipation $\Psi = \Psi_u(v)$, and on the energy functional $\mathcal{E} : [0,T] \times V \to (-\infty,+\infty]$, with domain $[0,T] \times D$. We emphasize that the conditions on $\mathcal{E}$ involve *both* its subdifferential $\mathrm{F} : [0,T] \times D \rightrightarrows V^*$ (with domain and graph dom(F) and graph(F), respectively), *and* its *generalized partial time-derivative* $\mathrm{P} = \mathrm{P}_t(u,\xi)$, for $(t,u,\xi) \in$ graph(F), since we encompass a nonsmooth dependence of the energy $\mathcal{E}$ on the time variable.

A (Finsler) family of dissipation potentials. We consider a family

$$\Psi_u : V \to [0,+\infty),\ u \in D, \text{ of } \textit{admissible dissipation potentials} \tag{4.Ψ_1}$$

i.e. Ψ_u complies with (2.Ψ_1)–(2.Ψ_3) for all $u \in D$. We now require that the potentials $(\Psi_u)_{u\in D}$ and $(\Psi_u^*)_{u\in D}$ have a superlinear growth, uniformly with respect to u in sublevels of the energy $\mathcal{E}$, viz.

$$\forall\, R > 0 \ : \quad \begin{cases} \displaystyle\lim_{\|v\|\to+\infty} \frac{1}{\|v\|} \inf_{\mathcal{G}(u)\leq R} \Psi_u(v) = +\infty, \\ \displaystyle\lim_{\|\xi\|_*\to+\infty} \frac{1}{\|\xi\|_*} \inf_{\mathcal{G}(u)\leq R} \Psi_u^*(\xi) = +\infty, \end{cases} \tag{4.Ψ_2}$$

where we have used the notation $\mathcal{G}(u) = \sup_{t\in[0,T]} \mathcal{E}_t(u)$. Furthermore, we require that the dependence $u \mapsto \Psi_u$ is continuous, on sublevels of the energy, in the sense of Mosco-convergence (see, e.g, [3, § 3.3, p. 295]), i.e.

$$\begin{array}{llll} \forall\, R > 0 : & u_n \to u, \quad \mathcal{G}(u_n) \leq R, \quad v_n \rightharpoonup v \text{ in } V & \Rightarrow & \liminf_{n\to\infty} \Psi_{u_n}(v_n) \geq \Psi_u(v) \\ \forall\, R > 0 : & u_n \to u, \quad \mathcal{G}(u_n) \leq R, \quad v \in V & \Rightarrow & \begin{cases} \exists\, v_n \to v : \\ \lim_{n\to\infty} \Psi_{u_n}(v_n) = \Psi_u(v). \end{cases} \end{array} \tag{4.Ψ_3}$$

For later use, we explicitly remark that assumption (4.Ψ_2) means that

$$\forall\, R > 0,\ M > 0$$

$$\begin{cases} \exists K > 0 \quad \forall\, u \in D \text{ with } \mathcal{G}(u) \leq R \quad \forall\, v \in V : \ \|v\| \geq K \quad \Rightarrow \quad \Psi_u(v) \geq M\|v\|, \\ \exists K^* > 0 \quad \forall\, u \in D \text{ with } \mathcal{G}(u) \leq R \quad \forall\, \xi \in V : \ \|\xi\|_* \geq K^* \quad \Rightarrow \quad \Psi_u^*(\xi) \geq M\|\xi\|_*. \end{cases} \tag{4.1}$$

We also recall an important consequence of assumption (4.Ψ_3) (see [3, Chap. 3]): for all $R > 0$

$$u_n \to u \quad \text{in } V, \quad \mathcal{G}(u_n) \leq R, \qquad \xi_n \rightharpoonup \xi \quad \text{in } V^* \quad \Rightarrow \quad \liminf_{n\to\infty} \Psi_{u_n}^*(\xi_n) \geq \Psi_u^*(\xi). \tag{4.2}$$

Indeed, it has been proved in [53, Lemma 4.1] that the first condition in (4.Ψ_3), combined with (4.2), is in fact *equivalent* to (4.Ψ_3).

Assumptions on the energy functional. We now formulate our assumptions on the functional $\mathcal{E}$. We recall the basic condition

$$\begin{gathered} u \mapsto \mathcal{E}_t(u) \text{ is l.s.c. for all } t \in [0,T], \quad \exists\, C_0 > 0 \ \forall\, (t,u) \in [0,T] \times D : \ \mathcal{E}_t(u) \geq C_0 \ \text{ and} \\ \text{graph(F) is a Borel set of } [0,T] \times V \times V^*. \end{gathered} \tag{4.E_0}$$

Coercivity: For all $t \in [0,T]$

$$\text{the map} \quad u \mapsto \mathcal{E}_t(u) \quad \text{has compact sublevels.} \tag{4.E_1}$$

Variational sum rule: If for some $u_o \in V$ and $\tau > 0$ the point $\bar{u}$ is a minimizer of $u \mapsto \mathcal{E}_t(u) + \tau\Psi_{u_o}((u-u_o)/\tau)$, then $\bar{u}$ fulfills the Euler-Lagrange equation

$$\exists\, \xi \in \mathrm{F}_t(\bar{u}) : \quad -\xi \in \partial\Psi_{u_o}((\bar{u}-u_o)/\tau). \tag{4.E_2}$$

Lipschitz continuity:

$$\exists\, C_1 > 0 \ \forall\, u \in D \ \forall\, t,s \in [0,T] : \quad |\mathcal{E}_t(u) - \mathcal{E}_s(u)| \leq C_1 \mathcal{E}_t(u)|t-s|. \tag{4.E_3}$$

Conditioned one-sided time-differentiability:
there exists a Borel function $\mathrm{P} : \mathrm{graph}(\mathrm{F}) \to \mathbb{R}$ and a constant $C_2 > 0$ such that

$$\forall\, (t,u,\xi) \in \mathrm{graph}(\mathrm{F}) : \quad \liminf_{h \downarrow 0} \frac{\mathcal{E}_{t+h}(u) - \mathcal{E}_t(u)}{h} \le \mathrm{P}_t(u,\xi) \le C_2 \mathcal{G}(u)\,. \tag{4.E$_4$}$$

Chain-rule inequality: $\mathcal{E}$ satisfies the *chain-rule inequality with respect to the triple* $(\Psi, \mathrm{F}, \mathrm{P})$, i.e. for every absolutely continuous curve $u \in \mathrm{AC}([0,T];V)$ and for all $\xi \in L^1(0,T;V^*)$ such that

$$\sup_{t \in (0,T)} |\mathcal{E}_t(u(t))| < +\infty, \qquad \xi(t) \in \mathrm{F}_t(u(t)) \ \text{ for a.a. } t \in (0,T), \tag{4.3}$$

$$\int_0^T \Psi_{u(t)}(u'(t))\,\mathrm{d}t < +\infty, \qquad \text{and} \qquad \int_0^T \Psi^*_{u(t)}(-\xi(t))\,\mathrm{d}t < +\infty, \tag{4.4}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) \ge \langle \xi(t), u'(t) \rangle + \mathrm{P}_t(u(t),\xi(t)) \quad \text{for a.a. } t \in (0,T). \tag{4.E$_5$}$$

Weak closedness of $(\mathcal{E}, \mathrm{F}, \mathrm{P})$**:** For all $t \in [0,T]$ and for all sequences $\{u_n\} \subset V$, $\xi_n \in \mathrm{F}_t(u_n)$, $\mathscr{E}_n = \mathcal{E}_t(u_n)$, $p_n = \mathrm{P}_t(u_n, \xi_n)$ fulfilling

$$u_n \to u \ \text{ in } V, \quad \xi_n \rightharpoonup \xi \ \text{ weakly in } V^*, \quad p_n \to p \ \text{ and } \ \mathscr{E}_n \to \mathscr{E} \ \text{ in } \mathbb{R},$$

there holds

$$(t,u) \in \mathrm{dom}(F), \quad \xi \in \mathrm{F}_t(u), \quad p \le \mathrm{P}_t(u,\xi), \quad \mathscr{E} = \mathcal{E}_t(u). \tag{4.E$_6$}$$

For later use, we point out that (4.E$_3$) yields the following estimate

$$\exists\, C_3 > 0 \ \ \forall\, u \in D \,:\, \mathcal{G}(u) \le C_3 \inf_{t \in [0,T]} \mathcal{E}_t(u)\,. \tag{4.5}$$

Notice that, under the above conditions (cf. (4.E$_3$)), for fixed $u \in D$ the function $t \mapsto \mathcal{E}_t(u)$ is Lipschitz continuous, hence a.e. differentiable. Still, it may happen that, along some curve $u \in \mathrm{AC}([0,T];V)$, the energy $\mathcal{E}_t(u)$ is not differentiable at $(t, u(t))$, for any $t \in [0,T]$, cf. e.g. Example 3.3. Hence, one needs to recur to the generalized notion P.

Example 4.1 (Example 3.3 revisited)**.** Let us refer to the setting of Example 3.3, and to the subdifferential $\mathrm{F}_t(u) = \partial_u^{\mathrm{Clarke}} \mathcal{E}_t(u)$, explicitly calculated in (3.9) (analogous considerations can be developed in the case $\mathrm{F}_t(u) = \widehat{\partial}\mathcal{E}_t(u)$ examined in Example 3.4). Since $\partial \mathcal{E}_t(u) \subset \partial_u^{\mathrm{Clarke}} \mathcal{E}_t(u)$, in view of the forthcoming Proposition 4.2, condition (4.E$_2$) is satisfied. As for the choice of the function $\mathrm{P} : \mathrm{graph}(\mathrm{F}) \to \mathbb{R}$ in such a way that chain-rule inequality holds, it follows from (3.9) that

$$\begin{cases} u > \beta t & \Rightarrow \quad \mathrm{F}_t(u) = \{-\alpha\}, \quad \mathrm{P}_t(u,-\alpha) = \alpha\beta, \\ u < \beta t & \Rightarrow \quad \mathrm{F}_t(u) = \{\alpha\}, \quad\ \ \mathrm{P}_t(u,\alpha) = -\alpha\beta. \end{cases}$$

Asking for the chain-rule inequality (3.11) along the curve $\bar{u}(t) = \beta t$ only needs, for every $\xi \in \mathrm{F}_t(\bar{u}(t)) = [-\alpha, \alpha]$, that $\mathrm{P}_t(\bar{u}(t), \xi)$ is a selection in the set $[-\alpha\beta, -\xi\beta]$ of the admissible p's. However, the closedness condition (4.E$_6$) is fulfilled only for the choice $\mathrm{P}_t(\bar{u}(t), \xi) = -\xi\beta$.

We conclude this section with a result providing sufficient conditions for the variational sum rule (4.E$_2$). As in the case of sum rules for convex functionals, we use that Ψ_{u_o} is locally Lipschitz, since its domain is the whole space V. Our proof relies on [40, Lemma 2.32].

Proposition 4.2. *Let* $\{\Psi_u\}$ *be a family of admissible dissipation potentials on the reflexive space* V*, and* $\mathcal{E} : [0,T] \times V \to (-\infty, +\infty]$ *an energy functional complying with* (4.E$_0$)*, with subdifferential* $\mathrm{F} : [0,T] \times V \rightrightarrows V^*$*. Suppose that*

$$\partial \mathcal{E}_t(u) \subset \mathrm{F}_t(u) \quad \textit{for every } (t,u) \in [0,T] \times D, \tag{4.6}$$

and that $(\mathcal{E}, \mathrm{F})$ *comply with the weak closedness condition* (4.E$_6$)*.*

Then, the variational sum rule (4.E$_2$) *holds.*

Proof: Let $\bar{u}$ be a minimizer of $u \mapsto \mathcal{E}_t(u) + \tau\Psi_{u_o}((u-u_o)/\tau)$. It follows from [40, Lemma 2.32, p. 214] that

$$\forall\,\eta > 0 \;\; \exists\, u^1_\eta \in V,\; u^2_\eta \in D \;:\quad \begin{cases} \|u^1_\eta - \bar{u}\| + \|u^2_\eta - \bar{u}\| & \leq \eta\,, \\ |\Psi_{u_o}\big(\frac{u^1_\eta - u_o}{\tau}\big) - \Psi_{u_o}\big(\frac{\bar{u}-u_o}{\tau}\big)| & \leq \eta\,, \\ |\mathcal{E}_t(u^2_\eta) - \mathcal{E}_t(\bar{u})| & \leq \eta\,, \end{cases} \tag{4.7}$$

and

$$\exists\, w_\eta \in \partial\Psi_{u_o}\left(\frac{u^1_\eta - u_o}{\tau}\right),\; \xi_\eta \in \partial\mathcal{E}_t(u^2_\eta),\; \zeta_\eta \in V^* \;:\quad \|\zeta_\eta\|_* \leq \eta, \;\text{ and }\; w_\eta + \xi_\eta + \zeta_\eta = 0\,. \tag{4.8}$$

Due to (4.6), we have $\xi_\eta \in \mathrm{F}_t(u^2_\eta)$. Choosing $\eta = 1/n$, we find sequences (u^1_n), (u^2_n), (w_n), (ξ_n), and (ζ_n) such that $\zeta_n \to 0$ in V^*, $u^1_n \to \bar{u}$ and $u^2_n \to \bar{u}$ in V, with $\mathcal{E}_t(u^2_n) \to \mathcal{E}_t(\bar{u})$. Since $w_n \in \partial\Psi_{u_o}\left(\frac{u^1_n - u_o}{\tau}\right)$ and $\partial\Psi_{u_o} : V \rightrightarrows V^*$ is a bounded operator, we also deduce that $\sup_n \|w_n\|_* < +\infty$. Hence, in view of (4.8), we ultimately have that (ξ_n) is bounded in V^*. Thus, there exists $\xi \in V^*$ such that, up to a (not relabeled) subsequence, $\xi_n \rightharpoonup \xi$ in V^*. Due to (4.E$_6$), we conclude that $\xi \in \mathrm{F}_t(\bar{u})$. On the other hand, passing to the limit in (4.8) and using the well-known strong-weak closedness property of $\partial\Psi_{u_o}$ gives $-\xi \in \partial\Psi_{u_o}((\bar{u}-u_o)/\tau)$, and (4.E$_2$) ensues. ∎

4.2. **Approximation.** For a fixed initial datum $u_0 \in D$ and a time step $\tau > 0$, we consider a uniform partition $\mathscr{P}_\tau = \{0 = t_0 < t_1 < t_2 < \cdots < t_{N-1} < T \leq t_N\}$ with $t_n := n\tau$, $U^0_\tau = u_0$, and construct a sequence $(U^n_\tau)^N_{n=1}$ by recursively solving

$$U^n_\tau \in \operatorname*{Argmin}_{U \in D} \left\{ \tau\Psi_{U^{n-1}_\tau}\left(\frac{U - U^{n-1}_\tau}{\tau}\right) + \mathcal{E}_{t_n}(U) \right\} \qquad n = 1, \cdots, N. \tag{4.9}$$

Using the *direct method* in the Calculus of Variations and exploiting assumption (4.E$_1$), one sees (cf. Lemma 6.1) that for all $u_0 \in D$ and $\tau \in (0, \tau_o)$ there exists at least one solution $(U^n_\tau)^N_{n=1}$ to the time-incremental minimization problem (4.9). We denote by $\overline{U}_\tau$ and $\underline{U}_\tau$, respectively, the left-continuous and right-continuous piecewise constant interpolants of the values $(U^n_\tau)^N_{n=1}$, i.e.,

$$\overline{U}_\tau(t) := U^n_\tau \quad \text{for } t \in (t_{n-1}, t_n], \qquad \underline{U}_\tau(t) := U^{n-1}_\tau \quad \text{for } t \in [t_{n-1}, t_n), \quad n = 1, \ldots, N, \tag{4.10}$$

and by U_τ the piecewise linear interpolant

$$U_\tau(t) := \frac{t - t_{n-1}}{\tau} U^n_\tau + \frac{t_n - t}{\tau} U^{n-1}_\tau \quad \text{for } t \in [t_{n-1}, t_n), \quad n = 1, \ldots, N. \tag{4.11}$$

Thanks to (4.E$_2$), for all $n = 1, \ldots, N$ there exists $\xi^n_\tau \in \mathrm{F}_{t_n}(U^n_\tau) \cap (-\partial\Psi_{U^{n-1}_\tau}(U^n_\tau - U^{n-1}_\tau/\tau))$. We denote by $\overline{\xi}_\tau$ the (left-continuous) piecewise constant interpolant of the family $(\xi^n_\tau)^N_{n=1} \subset V^*$.

Furthermore, we also consider the *variational interpolant* $\widetilde{U}_\tau$ of the discrete values $(U^n_\tau)^N_{n=1}$, which was first introduced by E. De Giorgi within the *Minimizing Movements* theory (see [19, 18, 1]). It is defined in the following way: the map $t \mapsto \widetilde{U}_\tau(t)$ is Lebesgue measurable in $(0, T)$ and satisfies

$$\begin{cases} \widetilde{U}_\tau(0) = u_0, \quad \text{and, for } t = t_{n-1} + r \in (t_{n-1}, t_n], \\ \\ \widetilde{U}_\tau(t) \in \operatorname{Argmin}_{U \in D} \left\{ r\Psi_{U^{n-1}_\tau}\left(\frac{U - U^{n-1}_\tau}{r}\right) + \mathcal{E}_t(U) \right\}, \end{cases} \tag{4.12}$$

The existence of such a measurable selection is ensured by [14, Cor. III.3, Thm. III.6], see also [43, Rem. 3.4]. When $t = t_n$, the minimization problems (4.9) and (4.12) coincide, so that we may assume

$$\overline{U}_\tau(t_n) = \underline{U}_\tau(t_n) = U_\tau(t_n) = \widetilde{U}_\tau(t_n), \quad \text{for every } n = 1, \ldots, N. \tag{4.13}$$

Then, $\widetilde{U}_\tau$ contains all the information on U_τ, $\overline{U}_\tau$, and $\underline{U}_\tau$. Furthermore, again by (4.E$_2$) and the measurable selection result [4, Thm. 8.2.9], with $\widetilde{U}_\tau$ we can associate a measurable function

$\widetilde{\xi}_\tau : (0,T) \to V^*$ fulfilling the Euler equation for the minimization problem (4.12), i.e.

$$\widetilde{\xi}_\tau(t) \in \mathrm{F}_t(\widetilde{U}_\tau(t)) \cap \left(-\partial\Psi_{\underline{U}_\tau(t)} \left(\frac{\widetilde{U}_\tau(t) - \underline{U}_\tau(t)}{t - t_{n-1}} \right) \right) \qquad \forall\, t \in [t_{n-1}, t_n), \quad n = 1, \ldots, N. \tag{4.14}$$

For later notational convenience, we also introduce the piecewise constants interpolants $\bar{\mathsf{t}}_\tau$ and $\underline{\mathsf{t}}_\tau$ associated with the partition $\mathscr{P}_\tau$, namely

$$\bar{\mathsf{t}}_\tau(0) = \underline{\mathsf{t}}_\tau(0) := 0, \quad \bar{\mathsf{t}}_\tau(t) := t_k \quad \text{for} \quad t \in (t_{k-1}, t_k], \quad \underline{\mathsf{t}}_\tau(t) := t_{k-1} \quad \text{for} \quad t \in [t_{k-1}, t_k). \tag{4.15}$$

Of course, for every $t \in [0,T]$ we have $\bar{\mathsf{t}}_\tau(t) \downarrow t$ and $\underline{\mathsf{t}}_\tau(t) \uparrow t$ as $\tau \downarrow 0$.

4.3. **Main existence result.** Before stating our main existence result, let us first specify the notion of solution we are interested in.

Definition 4.3. Let $\mathcal{E} : [0,T] \times V \to (-\infty, +\infty]$ be an energy functional fulfilling (4.E_0) and $\{\Psi_u\}_{u \in D}$ a family of *admissible dissipation potentials*. Suppose that $\mathcal{E}$ complies with the chain rule (4.E_5). We say that $(u, \xi) \in \mathrm{AC}([0,T]; V) \times L^1(0,T; V^*)$ is a *solution pair* to the generalized gradient system $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$ if

(1) (u, ξ) fulfills the doubly nonlinear equation

$$\xi(t) \in \mathrm{F}_t(u(t)), \quad \partial\Psi_{u(t)}(u'(t)) + \xi(t) \ni 0 \quad \text{for a.a.}\, t \in (0,T), \tag{4.16}$$

(2) (u, ξ) complies with the energy identity

$$\int_s^t \big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}(-\xi(r))\big)\,\mathrm{d}r + \mathcal{E}_t(u(t)) = \mathcal{E}_s(u(s)) + \int_s^t \mathrm{P}_r(u(r), \xi(r))\,\mathrm{d}r \tag{4.17}$$

for every $0 \le s \le t \le T$.

We shortly say that $u \in \mathrm{AC}([0,T]; V)$ is a solution to the generalized gradient system $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$, if there exists $\xi \in L^1(0,T; V^*)$ such that (u, ξ) is a solution pair to $(V, \mathcal{E}, \mathrm{F}, \mathrm{P}, \Psi)$.

Theorem 4.4 (Existence)**.** *Assume that* $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$ *comply with* (4.Ψ_1)–(4.Ψ_3) *and with* (4.E_0)–(4.E_6).

Then, for every $u_0 \in D$ *there exists a solution* $u \in \mathrm{AC}([0,T]; V)$ *to the doubly nonlinear equation* (1.21)*, fulfilling the initial condition* $u(0) = u_0$.

In fact, for any family of approximate solutions $(\widetilde{U}_\tau, \widetilde{\xi}_\tau)_{\tau > 0}$ *there exist a sequence* $\tau_k \downarrow 0$ *as* $k \to \infty$*, and* $\xi \in L^1(0,T; V^*)$ *such that the following convergences hold as* $k \to \infty$

$$\overline{U}_{\tau_k},\ U_{\tau_k},\ \widetilde{U}_{\tau_k} \to u \quad \text{in } L^\infty(0,T; V), \tag{4.18a}$$

$$U_{\tau_k} \rightharpoonup u \quad \text{in } W^{1,1}(0,T; V), \tag{4.18b}$$

$$\mathcal{E}_t(\overline{U}_{\tau_k}(t)) \to \mathcal{E}_t(u(t)) \quad \text{for all } t \in [0,T], \tag{4.18c}$$

$$\int_s^t \Psi_{\underline{U}_{\tau_k}(r)}(U'_{\tau_k}(r))\,\mathrm{d}r \to \int_s^t \Psi_{u(r)}(u'(r))\,\mathrm{d}r \quad \text{for all } 0 \le s \le t \le T, \tag{4.18d}$$

$$\int_s^t \Psi^*_{\underline{U}_{\tau_k}(r)}(-\widetilde{\xi}_{\tau_k}(r))\,\mathrm{d}r \to \int_s^t \Psi^*_{u(r)}(-\xi(r))\,\mathrm{d}r \quad \text{for all } 0 \le s \le t \le T, \tag{4.18e}$$

and (u, ξ) *is a solution pair to the generalized gradient system* $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$.

Furthermore, if

$$\Psi^*_u \text{ is strictly convex for all } u \in V, \tag{4.19}$$

we have the additional convergence

$$\widetilde{\xi}_{\tau_k} \rightharpoonup \xi \quad \text{in } L^1(0,T; V^*). \tag{4.20}$$

The proof of Theorem 4.4 is developed throughout Section 6.

Remark 4.5. The considerations set forth in Remark 2.3 for energies smoothly depending on time extend to the present setting. Namely, the proof of Theorem 4.4 reveals that the *one-sided* chain-rule inequality (4.E_5) is sufficient to conclude the existence of solutions to the Cauchy problem for (1.21), in that it is combined with the upper energy estimate following from the discretization scheme.

Clearly, in order to enforce the energy identity (4.17) for *any* solution to (1.21), it would be necessary to impose (4.E_5) as an *equality*. As shown by Example 3.3, this may lead to restrictions on the admissible functions P.

Weakened assumptions. The two ensuing remarks explore the possibility of refining our requirements on the chain rule (4.E_5), and on the properties of the dissipation potentials.

Remark 4.6 (A weaker chain rule). Like in the gradient flow case (cf. [44, Thm. 2], and [2, Thm. 2.3.1] in the metric setting), it is possible to state our existence result for the Cauchy problem for (1.21) under a (slightly) weaker form of the chain rule (4.E_5), which requires that for every absolutely continuous curve $u \in \mathrm{AC}([0,T];V)$ and $\xi \in L^1(0,T;V^*)$ satisfying (4.3), as well as

$$\int_0^T \Psi_{u(t)}(u'(t))\,\mathrm{d}t < +\infty, \qquad \int_0^T \Psi^*_{u(t)}(-\xi(t))\,\mathrm{d}t < +\infty, \tag{4.21}$$

and such that the map $t \mapsto \mathcal{E}_t(u(t))$ is a.e. equal to a function $\mathscr{E}$ of bounded variation,

there holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{E}(t) \geq \langle \xi(t), u'(t)\rangle + \mathrm{P}_t(u(t),\xi(t)) \quad \text{for a.a. } t \in (0,T). \tag{CHAIN$_{\mathrm{weak}}$}$$

In this case, suitably adapting the proof of [44, Thm. 2] to the doubly nonlinear case, one obtains that there exist $u \in \mathrm{AC}([0,T];V)$ and $\xi \in L^1(0,T;V^*)$ fulfilling the differential inclusion (4.16) and the following *energy inequality* (compare with the energy identity (4.17) under the chain rule (4.E_5))

$$\begin{aligned}\int_s^t \big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}(-\xi(r))\big)\,\mathrm{d}r + \mathcal{E}_t(u(t)) \leq \mathcal{E}_s(u(s)) + \int_s^t \mathrm{P}_r(u(r),\xi(r))\,\mathrm{d}r \\ \text{for all } t \in [0,T] \text{ and almost all } s \in (0,t).\end{aligned} \tag{4.22}$$

Remark 4.7 (Weaker conditions on the dissipation). In fact, condition (2.Ψ_3) in the definition of the dissipation potentials $(\Psi_u)_{u\in V}$, is only used in the proof of the forthcoming Lemma 6.1, which is the crucial technical result for the a priori estimates on the approximate solutions $(\overline{U}_\tau)$, (U_τ) and $(\widetilde{U}_\tau)$. As shown in Remark 6.2 later on, it is possible to dispense with (2.Ψ_3) if the following condition, involving the energy $\mathcal{E} : [0,T] \times V \to (-\infty,+\infty]$ and its subdifferential mapping $\mathrm{F} : [0,T] \times V \rightrightarrows V^*$, holds:

$$\begin{aligned}\text{for all } u,\, v \in V \text{ the directional derivative } \delta\mathcal{E}_t(u;v) := \lim_{h\downarrow 0} \frac{1}{h}\left(\mathcal{E}_t(u+hv) - \mathcal{E}_t(u)\right) \text{ exists,} \\ \text{and } \langle \xi, v\rangle \geq \delta\mathcal{E}_t(u;v) \quad \text{for all } \xi \in \mathrm{F}_t(u) \text{ and } v \in V.\end{aligned} \tag{4.23}$$

Condition (4.23) has to be coupled with a strengthened version of the first inequality in (4.E_4), namely

$$\begin{aligned}&\text{for every } (t,u,\xi) \in \mathrm{graph}(\mathrm{F}) \text{ and } (u_h) \subset V \text{ such that } u_h \to u \text{ as } h \downarrow 0, \\ &\text{there holds } \liminf_{h\downarrow 0} \frac{\mathcal{E}_{t+h}(u_h) - \mathcal{E}_t(u_h)}{h} \leq \mathrm{P}_t(u,\xi).\end{aligned} \tag{4.24}$$

Notice that (4.23) holds for marginal functionals which are λ-concave.

4.4. **Upper semicontinuity of the set of solutions.** We now address the issue of *upper semicontinuity* of the set of solutions to the Cauchy problem for (1.21), with respect to convergence of the initial data and (a suitable kind of) variational convergence for the driving energy functionals.

We consider sequences $(V, \mathcal{E}^n, \Psi^n, \mathrm{F}^n, \mathrm{P}^n)$ of generalized gradient systems, and impose the following.

Assumption (H1). Let $(\mathcal{E}^n)_{n\in\mathbb{N}}$ be a sequence of lower semicontinuous energy functionals $\mathcal{E}^n : [0,T]\times V \to (-\infty,+\infty]$, with domains $\mathrm{dom}(\mathcal{E}^n) = [0,T]\times D_n$ for some $D_n \subset V$, and with subdifferentials $\mathrm{F}^n : [0,T]\times D_n \rightrightarrows V^*$; we use the notation $\mathcal{G}^n(u) := \sup_{t\in[0,T]} \mathcal{E}^n_t(u)$ for $u \in D_n$. We suppose that the functionals $(\mathcal{E}^n)_{n\in\mathbb{N}}$ comply with (4.E$_0$), (4.E$_1$), (4.E$_3$), and (4.E$_4$), with constants *uniform* with respect to $n \in \mathbb{N}$. We also require that there exists a generalized gradient system $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$, such that the energy

$$\mathcal{E} : [0,T]\times V \to (-\infty,+\infty] \ \text{ complies with (4.E}_0\text{) and the chain rule (4.E}_5\text{)},$$

and the functionals $(\mathcal{E}^n)_n$ converge to $\mathcal{E}$ in the following sense: for all $t \in [0,T]$ and for all sequences $\{u_n\} \subset V,\ \xi_n \in \mathrm{F}^n_t(u_n)$, fulfilling

$$u_n \to u \ \text{ in } V, \quad \xi_n \rightharpoonup \xi \ \text{ weakly in } V^*, \quad \mathrm{P}^n_t(u_n,\xi_n) \to p,$$

there holds

$$\begin{gathered} (t,u) \in \mathrm{dom}(F), \quad \xi \in \mathrm{F}_t(u), \quad p \le \mathrm{P}_t(u,\xi), \\ \text{and, if } \mathcal{E}^n_t(u_n) \text{ converges to some } \mathscr{E} \in \mathbb{R}, \text{ then } \mathscr{E} = \mathcal{E}_t(u). \end{gathered} \tag{4.25}$$

Assumption (H2). Let $\{\Psi^n_u\}_{u\in D_n}$ be a family of *admissible* dissipation potentials, satisfying conditions of superlinear growth on sublevels of the energies $\mathcal{E}^n$, uniformly with respect to n (i.e., (4.1) holds for constants *independent* of n). We also suppose that the potentials $(\Psi^n_u)_{u\in D_n}$ Mosco converge on sublevels of the energies to a family $(\Psi_u)_{u\in D}$ of admissible potentials, viz.

$$\begin{aligned} &\forall\, R>0: \quad u_n \to u, \quad \sup_{n\in\mathbb{N}} \mathcal{G}^n(u_n) \le R, \quad v_n \rightharpoonup v \text{ in } V \quad \Rightarrow \quad \liminf_{n\to\infty} \Psi^n_{u_n}(v_n) \ge \Psi_u(v) \\ &\forall\, R>0: \quad u_n \to u, \quad \sup_{n\in\mathbb{N}} \mathcal{G}^n(u_n) \le R, \quad v \in V \quad \Rightarrow \quad \begin{cases} \exists\, v_n \to v: \\ \lim_{n\to\infty} \Psi^n_{u_n}(v_n) = \Psi_u(v). \end{cases} \end{aligned} \tag{4.26}$$

Theorem 4.8 (Upper semicontinuity)**.** *Let $(V, \mathcal{E}_n, \Psi_n, \mathrm{F}_n, \mathrm{P}_n)$ be a family of generalized gradient systems complying with **Assumption (H1)** and **Assumption (H2)**. Let $(u^n_0)_n$ be a sequence of initial data, with $u^n_0 \in D_n$ for all $n \in \mathbb{N}$, such that*

$$u^n_0 \rightharpoonup u_0 \quad \text{in } V \text{ and } \quad \mathcal{E}^n_0(u^n_0) \to \mathcal{E}_0(u_0), \tag{4.27}$$

and let $(u_n,\xi_n)_{n\in\mathbb{N}}$ a sequence of solution pairs *to the Cauchy problems*

$$\partial\Psi_{u(t)}(u'(t)) + \mathrm{F}^n_t(u(t)) \ni 0 \quad \text{in } V^*, \quad \text{for a.a. } t \in (0,T); \qquad u(0) = u^n_0 \tag{4.28}$$

(in particular, complying with the energy identity (4.17) for all $n \in \mathbb{N}$). Then, there exist a subsequence $(u_{n_k},\xi_{n_k})_{k\in\mathbb{N}}$ and functions $(u,\xi) \in \mathrm{AC}([0,T];V)\times L^1(0,T;V^*)$ such that (u,ξ) is a solution pair of the Cauchy problem for (1.21), and the following convergences hold as $k \to \infty$

$$u_{n_k} \to u \text{ in } \mathrm{C}^0([0,T];V), \quad u_{n_k} \rightharpoonup u \text{ in } W^{1,1}(0,T;V), \tag{4.29a}$$

$$\mathcal{E}^{n_k}_t(u_{n_k}(t)) \to \mathcal{E}_t(u(t)) \text{ for all } t \in [0,T], \tag{4.29b}$$

$$\begin{cases} \int_s^t \Psi^{n_k}_{u_{n_k}(r)}(u'_{n_k}(r))\,\mathrm{d}r \to \int_s^t \Psi_{u(r)}(u'(r))\,\mathrm{d}r, \\ \int_s^t (\Psi^{n_k}_{u_{n_k}(r)})^*(-\xi_{n_k}(r))\,\mathrm{d}r \to \int_s^t \Psi^*_{u(r)}(-\xi(r))\,\mathrm{d}r \end{cases} \quad \text{for all } 0 \le s \le t \le T. \tag{4.29c}$$

The proof of this result is outlined at the end of Section 6.

Remark 4.9. Suppose that the energy functionals $\mathcal{E}^n$ have the special form

$$\mathcal{E}^n_t(u) = E^n(u) - \langle \ell^n(t), u\rangle, \quad \text{with } E^n : V \to (-\infty,+\infty] \text{ convex functionals and } (\ell^n) \subset \mathrm{C}^1([0,T];V^*).$$

Hence, if the functionals (E^n) Mosco-converge to some convex functional $E : V \to (-\infty,+\infty]$, and if the functions (ℓ_n) suitably converge to some $\ell \in \mathrm{C}^1([0,T];V^*)$, then the energies $(\mathcal{E}^n)$ converge to $\mathcal{E}_t(u) := E(u) - \langle \ell(t), u\rangle$ in the sense specified by **Assumption (H1)**. Indeed, Theorem 4.8 might be viewed as an extension, to the doubly nonlinear case, of the result on stability of gradient flows (with V a Hilbert space and $\Psi(u) = \frac{1}{2}\|u\|^2_V$), with respect to Mosco-convergence of the (convex) energies, stated in [3, Thm. 3.74(2), p. 388]. The reader may also consult [53] and the references therein.

5. Application: evolutions driven by marginal functionals in finite-strain elasticity

In this section we examine a mechanical model for finite-strain elasticity, described in terms of the elastic deformation and of some internal, dissipative variable z. Its analysis has already been developed in [24], in the case of a *rate-independent* evolution for z. Therein, existence of *energetic solutions* to the (Cauchy problem for the) related PDE system has been proved. Here, we address the case in which the evolution of z is driven by *viscous* dissipation.

5.1. **Problem set-up and existence result.** We consider an elastic body occupying a bounded domain $\Omega \subset \mathbb{R}^d$, $d \geq 1$, with Lipschitz boundary Γ. We denote by $\phi : \Omega \to \mathbb{R}^d$ the elastic deformation field, and assume that the inelasticity of Ω is described by an internal variable $z : \Omega \to \mathbb{R}^m$, $m \geq 1$, which we may envisage as a mesoscopic averaged phase variable.
Energy functional. The stored energy $I = I_t(\phi, z)$ has the form

$$I_t(\phi, z) = \mathcal{E}^1(z) + I_t^2(\phi, z) + E_0, \tag{5.1}$$

with $E_0 \in \mathbb{R}$ to be precised later on (cf. Lemma 5.5).
In (5.1), $\mathcal{E}^1 : V \to (-\infty, +\infty]$ is the *convex* functional

$$\mathcal{E}^1(z) := \begin{cases} \int_\Omega \left(\frac{1}{q}|\nabla z|^q + I_{\mathrm{K}}(z)\right) \mathrm{d}x & \text{if } z \in W^{1,q}(\Omega;\mathbb{R}^m), \\ +\infty & \text{otherwise,} \end{cases} \tag{5.2}$$
$$\text{where } q > d, \text{and K is a compact subset of } \mathbb{R}^m,$$

and we consider the Fröbenius norm $|\nabla z| = \left(\sum_{j=1}^d \sum_{i=1}^m |\partial_{x_j} z^i|^2\right)^{1/2}$ of the matrix ∇z.
The *nonconvex* contribution $I^2 : [0,T] \times W^{1,p}(\Omega;\mathbb{R}^d) \times L^2(\Omega;\mathbb{R}^m) \to (-\infty, +\infty]$ is given by

$$I_t^2(\phi, z) := \int_\Omega W(\nabla\phi(x), z(x))\, \mathrm{d}x - \langle \ell(t), \phi \rangle_{W^{1,p}},$$

where $p > d$, $\langle \cdot, \cdot \rangle_{W^{1,p}}$ denotes the duality pairing between $W^{1,p}(\Omega;\mathbb{R}^d)^*$ and $W^{1,p}(\Omega;\mathbb{R}^d)$, and we suppose that

$$\ell \in \mathrm{C}^1([0,T]; W^{1,p}(\Omega;\mathbb{R}^d)^*). \tag{5.3}$$

The stored energy density $W : \mathbb{R}^{d\times d} \times \mathrm{K}' \to (-\infty, +\infty]$ has domain $\mathrm{dom}(W) = D_W \times \mathrm{K}'$, where K' is a compact subset of $\mathbb{R}^m$ containing K. We neglect the dependence of W on the variable x for the sake of simplicity and with no loss of generality. We impose the following conditions on W:

$$\exists\, \kappa_1,\, \kappa_2 > 0\ \forall\, (F, z) \in \mathbb{R}^{d\times d} \times \mathrm{K}' : \quad W(F, z) \geq \kappa_1 |F|^p - \kappa_2 \quad \text{with } p > d; \tag{W$_1$}$$

$$\begin{aligned} &\exists\, \mathbb{W} : \mathbb{R}^{\mu_d} \times \mathrm{K}' \to (-\infty, +\infty] \text{ such that} \\ &\text{(i) } \mathbb{W} \text{ is lower semicontinuous,} \\ &\text{(ii) } \forall\, (F, z) \in \mathbb{R}^{d\times d} \times \mathrm{K}' : \quad W(F, z) = \mathbb{W}(\mathbb{M}(F), z), \\ &\text{(iii) } \forall z \in \mathrm{K}' : \quad \mathbb{W}(\cdot, z) : \mathbb{R}^{\mu_d} \to (-\infty, +\infty] \text{ is convex;} \end{aligned} \tag{W$_2$}$$

$$\begin{aligned} &\text{for all } F \in D_W \text{ the map } W(F, \cdot) \text{ is continuous and Gâteau-differentiable on } \mathrm{K}', \text{ and} \\ &\text{(i) } \exists\, \kappa_3,\, \kappa_4 > 0\ \forall\, (F, z) \in D_W \times \mathrm{K}' : \quad |\mathrm{D}_z W(F, z)| \leq \kappa_3 (W(F, z) + \kappa_4)^{1/2}; \\ &\text{(ii)} \exists\, \kappa_5,\, \kappa_6 > 0,\, \exists\, \alpha \in (0,1] \quad \forall\, F \in D_W\ \forall\, z_1,\, z_2 \in \mathrm{K}' : \\ &\qquad |\mathrm{D}_z W(F, z_1) - \mathrm{D}_z W(F, z_2)| \leq \kappa_5 |z_1 - z_2|^\alpha (W(F, z_1) + \kappa_6)^{1/2}. \end{aligned} \tag{W$_3$}$$

In (W$_2$), we have used the notation $\mu_d = \sum_{s=1}^d \binom{d}{s}^2$, and $\mathbb{M} : \mathbb{R}^{d\times d} \to \mathbb{R}^{\mu_d}$ is the function which maps a matrix to all its minors (subdeterminants). Hence, (W$_2$) states that for all $z \in \mathrm{K}'$ the map $W(\cdot, z)$ is *polyconvex.*

Dissipation. We consider a measurable (dissipation density) function $\psi : \mathrm{K} \times \mathbb{R}^m \to [0,+\infty)$ (again, we omit the dependence of ψ on the variable x with no loss of generality), fulfilling

$$\psi : K \times \mathbb{R}^m \to [0,+\infty) \quad \text{is continuous;} \tag{ψ_1}$$

$$\begin{aligned}&\forall\, z \in \mathrm{K}: \quad \psi(z,\cdot) : \mathbb{R}^m \to [0,+\infty) \ \text{ is convex, with } \ \psi(z,0) = 0, \ \text{ and}\\ &\psi^*(z,w_1) = \psi^*(z,w_2) \ \text{ for all } w_1,\, w_2 \in \partial_v \psi(z,v) \text{ and all } v \in \mathbb{R}^m;\end{aligned} \tag{ψ_2}$$

$$\exists\, \kappa_7,\, \kappa_8,\, \kappa_9 > 0 \ \ \forall\, z \in \mathrm{K}\ \forall\, v,\, \xi \in \mathbb{R}^m : \quad \begin{cases} \psi(z,v) \geq \kappa_7 |v|^2 - \kappa_9, \\ \psi^*(z,\xi) \geq \kappa_8 |\xi|^2 - \kappa_9. \end{cases} \tag{ψ_3}$$

In (ψ_2) the symbols $\partial_v\psi$ and ψ^* respectively denote the subdifferential and the Fenchel-Moreau conjugate of the function $\psi(z,\cdot)$. Let us point out that there is a crucial interplay between the exponent 1/2 in (W_3)(ii), and the exponents 2 in (ψ_3), see also Remark 5.4 later on.

PDE system and existence theorem. Within this setting, we address the analysis of the doubly nonlinear evolution equation

$$\begin{aligned}\partial_v\psi(z(t,x),\dot z(t,x)) - \Delta_q z(t,x) + \partial I_{\mathrm{K}}(z(t,x)) + \mathrm{D}_z W(\nabla\phi(t,x), z(t,x)) \ni 0 \\ \text{for a.a. } (t,x) \in (0,T)\times\Omega,\end{aligned} \tag{5.4a}$$

(where $\Delta_q z = \mathrm{div}(|\nabla z|^{q-2}\nabla z)$), supplemented with homogeneous Neumann boundary conditions, and coupled with the minimum problem

$$\phi(t,x) \in \mathrm{Argmin}\{I_t(\phi, z(t,x)) \ : \ \phi \in \mathcal{F}\} \quad \text{for a.a. } (t,x) \in (0,T)\times\Omega, \tag{5.4b}$$

where $\mathcal{F}$ denotes the set of the kinematically admissible deformation fields, viz.

$$\mathcal{F} = \left\{\phi \in W^{1,p}(\Omega;\mathbb{R}^d) \ : \ \phi = \phi_{\mathrm{Dir}} \text{ on } \Gamma_{\mathrm{Dir}}\right\},$$

for some $\Gamma_{\mathrm{Dir}} \subset \Gamma$, $\Gamma_{\mathrm{Dir}} \neq \emptyset$ with positive Hausdorff measure, and

$$\phi_{\mathrm{Dir}} \in W^{1,p}(\Omega;\mathbb{R}^d), \ \text{such that the map } x \mapsto \max_{z\in\mathrm{K}} W(\nabla\phi_{\mathrm{Dir}}(x), z) \text{ is in } L^1(\Omega). \tag{5.5}$$

Theorem 5.1. *Under assumptions* (5.3), (W_1)–(W_3), (ψ_1)–(ψ_3), *and* (5.5), *for every*

$$z_0 \in W^{1,q}(\Omega;\mathbb{R}^m) \quad \textit{with} \quad z_0(x) \in \mathrm{K} \quad \textit{for all } x \in \Omega, \tag{5.6}$$

there exist functions $z \in L^\infty(0,T;W^{1,q}(\Omega;\mathbb{R}^m)) \cap H^1(0,T;L^2(\Omega;\mathbb{R}^m))$ *and* $\phi \in L^\infty(0,T;W^{1,p}(\Omega;\mathbb{R}^d))$ *fulfilling* (5.4a)*, supplemented with homogeneous Neumann boundary conditions and the initial condition* $z(0,x) = z_0(x)$ *for a.a.* $x \in \Omega$*, and* (5.4b)*. In particular, there exists* $\xi \in L^2(0,T;L^2(\Omega;\mathbb{R}^m))$ *satisfying, for almost all* $(t,x) \in (0,T)\times\Omega$*, the inclusions*

$$\begin{cases} \partial_v\psi(z(t,x),\dot z(t,x)) + \xi(t,x) \ni 0, \\ \xi(t,x) \in -\Delta_q z(t,x) + \partial I_{\mathrm{K}}(z(t,x)) + \mathrm{D}_z W(\nabla\phi(t,x), z(t,x)) \end{cases} \tag{5.7}$$

and such that (z,ϕ,ξ) *fulfill the energy identity for all* $0 \leq s \leq t \leq T$

$$\begin{aligned}\int_s^t \int_\Omega \big(\psi(z(r,x),\dot z(r,x)) + \psi^*(z(r,x),-\xi(r,x))\big)\,\mathrm{d}x\,\mathrm{d}r + I_t(\phi(t),z(t)) \\ = I_s(\phi(s),z(s)) - \int_s^t \langle \ell'(r), \phi(r)\rangle_{W^{1,p}}\,\mathrm{d}r.\end{aligned} \tag{5.8}$$

Example 5.2. In finite-strain elasticity there are two main conditions, namely (i) frame indifference and (ii) local invertibility:

$$\begin{aligned} &i) \quad W(RF,z) = W(F,z) \quad \text{for all } R \in SO(d),\ F \in D_W,\ z \in \mathrm{K}', \\ &ii) \quad\qquad W(F,z) = \infty \quad \text{for all } \det(F) \leq 0. \end{aligned}$$

These conditions are compatible with polyconvexity, e.g. by choosing functions of the type

$$W(F,z) = C|F|^p + w_{\mathrm{co}}(F,z) + h(\det(F)),$$

where $h : \mathbb{R} \to (-\infty,+\infty]$ is continuous, convex, and satisfies $h(y) = \infty$ for $y \leq 0$. Thus, $D_W = \{F \ : \ \det(F) > 0\} \subset \mathbb{R}^{d\times d}$ is the nonconvex domain. Recall that the Fröbenius norm

$|F| = (\mathrm{tr}(F^T F))^{1/2}$ satisfies $|RF| = |F|$. Conditions (W_2)–(W_3) can be now satisfied if the coupling energy w_{co} satisfies $w_{\mathrm{co}} \in \mathrm{C}^1(\mathbb{R}^{d\times d} \times \mathrm{K}'; \mathbb{R})$, and

$$
\begin{aligned}
&w_{\mathrm{co}}(F, z) \geq 0, \quad w_{\mathrm{co}}(RF, z) = w_{\mathrm{co}}(F, z), \quad w_{\mathrm{co}}(\cdot, z) \text{ is polyconvex,}\\
&|\mathrm{D}_z w_{\mathrm{co}}(F, z)| \leq \kappa(|F| + 1)^{p/2},\\
&|\mathrm{D}_z w_{\mathrm{co}}(F, z_1) - \mathrm{D}_z w_{\mathrm{co}}(F, z_2)| \leq \kappa |z_1 - z_2|^{\alpha}(|F| + 1)^{p/2}
\end{aligned}
$$

for all arguments.

In magnetism (see [33]), z denotes the magnetization (with respect to material coordinates), and we have $z \in \mathbb{R}^d$ and $\mathrm{K} = \{z \in \mathbb{R}^d \,:\, |z| \leq z_{\mathrm{sat}}\}$, where the subscript "sat" stands for *saturation.* A choice for the coupling energy for $p \geq 4$ is $w_{\mathrm{co}}(F, z) = C|Fz|^2 + \tilde{w}(z)$, where $\tilde{w} \in \mathrm{C}^2(\mathrm{K}')$ gives the anisotropy of magnetization, as well as the saturation term $\frac{1}{4\delta}(|z|^2 - z_{\mathrm{sat}}^2)^2$, with $\delta > 0$.

For shape memory alloys, z may denote volume fractions of m different phases, such that $\mathrm{K} = \{z = (z_1, z_2, \ldots, z_m) \in \mathbb{R}^m \,:\, z_j \leq 0,\ \sum_{k=1}^m z_k = 1\}$. Denoting by $\mathrm{cof}F \in \mathbb{R}^{d\times d}$ the cofactor matrix $\det(F)F^{-T}$ (which is contained in $\mathbb{M}(F)$), and by $C_n(z)$, $n = 1, \ldots, N$ the z-dependent, effective transformation Cauchy strains, we may use

$$
w_{\mathrm{co}}(F, z) = \sum_{n=1}^{N} \alpha_n |F C_n(z)^{-1} - \mathrm{cof}F|^2 + w_{\mathrm{mix}}(z),
$$

with $\alpha_1, \ldots, \alpha_N > 0$ and a mixture energy $w_{\mathrm{mix}} \in \mathrm{C}^2(\mathrm{K}')$, see [28, 26]. Here we need $p \geq 2d$, because $|\mathrm{D}_z w_{\mathrm{co}}(F, z)| \leq C(|F| + 1)^d$, as the highest power $|\mathrm{cof}F|^2 \sim O(|F|^{2d-2})$ is independent of z. Note that we follow the ideas in [26, 38], where $W(\cdot, z)$ is considered to be a polyconvex relaxation, under given volume fractions of the different phases.

Example 5.3. Most commonly, the dissipation potentials ψ are assumed to be independent of the state z, i.e. $\psi(z, v) = \psi(v)$, which simplifies the analysis considerably. However, there are cases where ψ must depend on z, like in finite-strain elasticity where the internal variable is the plastic tensor $P \in \mathrm{SL}(d) = \{P \in \mathbb{R}^{d\times d} \,:\, \det(P) = 1\}$, and $\psi_P(\dot{P}) = \hat{\psi}(\dot{P}P^{-1})$. In the framework of the modeling for magnetization illustrated in Example 5.2, we may consider $\psi : \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty)$ of the form

$$
\psi(z, v) = \psi_{\mathrm{rad}}(z \cdot v) + \psi_{\mathrm{tang}}\left(\left(\mathrm{I} - \tfrac{1}{|z|^2} z \otimes z\right) v\right),
$$

to account for different dissipations for enlarging the magnetization or changing its orientation.

5.2. **Proof of Theorem 5.1.**

Outline of the proof. We follow an abstract approach to the analysis of (5.4), by rephrasing it as a doubly nonlinear equation of the type (1.21), generated by the generalized gradient system $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$ specified in the following lines.

Space V: We choose

$$
\text{the ambient space } \ V = L^2(\Omega; \mathbb{R}^m).
$$

Energy $\mathcal{E}$: We consider the *reduced functional* $\mathcal{E} : [0, T] \times W^{1,q}(\Omega; \mathbb{R}^m) \to (-\infty, +\infty]$ obtained by minimizing out the displacements from I, i.e.

$$
\begin{aligned}
\mathcal{E}_t(z) := \inf\{I_t(\phi, z) \,:\, \phi \in \mathcal{F}\} \ \text{ with domain } [0, T] \times D, \text{ where}\\
D = \{z \in W^{1,q}(\Omega; \mathbb{R}^m) \,:\, z(x) \in \mathrm{K} \text{ for a.a. } x \in \Omega\}.
\end{aligned}
\tag{5.9}
$$

We often use the decomposition of $\mathcal{E}$ as a sum of a convex and of a nonconvex, reduced functional

$$
\mathcal{E}_t(z) = \mathcal{E}^1(z) + \inf\{I_t^2(\,,\phi) \,:\, \phi \in \mathcal{F}\} \doteq \mathcal{E}^1(z) + \mathcal{E}_t^2(z) \quad \text{for } (t, z) \in [0, T] \times D. \tag{5.10}
$$

Indeed, in Lemma 5.5 below we prove that for all $(t, z) \in [0, T] \times D$, the set

$$
M(t, z) := \operatorname*{Argmin}_{\phi \in \mathcal{F}} \{I_t(\phi, z)\} \ \text{ is nonempty and weakly compact in } W^{1,p}(\Omega; \mathbb{R}^d). \tag{5.11}
$$

Subdifferential F: Reflecting (5.10), we use the following subdifferential notion

$$\mathrm{F}_t(z) := \partial\mathcal{E}^1(z) + \widehat{\partial}\mathcal{E}_t^2(z) \quad \text{for all } (t,z) \in [0,T] \times D, \tag{5.12}$$

where, as in Section 3, $\widehat{\partial}\mathcal{E}_t^2(z)$ is the *marginal subdifferential* of the reduced energy $\mathcal{E}^2$, viz.

$$\widehat{\partial}\mathcal{E}_t^2(z) = \left\{ \mathrm{D}_z I_t^2(\phi, z) \,:\, \phi \in M(t,z) \right\},$$

with $\mathrm{D}_z I_t^2(\phi,\cdot)$ the Gâteau derivative of the functional $I_t^2(\phi,\cdot)$.

Generalized time-derivative P: We set

$$R(t,z,\xi) := \{\varphi \in M(t,z) \,:\, \xi \in \partial\mathcal{E}^1(z) + \mathrm{D}_z I_t^2(\varphi, z)\} \quad \text{for all } (t,z,\xi) \in \operatorname{graph}(\mathrm{F})$$

and define

$$\mathrm{P}_t(z,\xi) := \max_{\varphi \in R(t,z,\xi)} \langle -\ell'(t), \varphi\rangle_{W^{1,p}}\,. \tag{5.13}$$

Dissipation potential Ψ: We consider the Finsler family $(\Psi_z)_{z\in D}$ of dissipation potentials

$$\Psi_z : V \to [0,+\infty) \ \text{ defined by } \ \Psi_z(v) := \int_\Omega \psi(z(x), v(x))\,\mathrm{d}x. \tag{5.14}$$

In what follows, throughout Lemmas 5.5–5.10 we check that the above generalized gradient system $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$ complies with the abstract assumptions $(4.\Psi_1)$–$(4.\Psi_3)$, $(4.\mathrm{E}_0)$–$(4.\mathrm{E}_6)$ of Theorem 4.4. The latter result yields the existence of a pair (u,ξ) fulfilling the Cauchy problem for (4.16), and the energy identity (4.17). The forthcoming calculations show that, in the present setting, (4.16) and (4.17) entail (5.4) and (5.8).

Remark 5.4. As it will be clear from the ensuing calculations, it is possible to generalize the theory to the case where, in place of (ψ_3), we have for some $r \in (1,\infty)$

$$\psi(z,v) \geq \kappa_7 |v|^r - \kappa_9, \quad \psi(z,v) \geq \kappa_8 |v|^{r'} - \kappa_9$$

(where $r' = r/(r-1)$ is the conjugate exponent of r), and the growth conditions in (W_3) are replaced by

$$|\mathrm{D}_z W(F,z)| \leq \kappa_3 (W(F,z) + \kappa_4)^{1-1/r},$$
$$|\mathrm{D}_z W(F,z_1) - \mathrm{D}_z W(F,z_2)| \leq \kappa_5 |z_1 - z_2|^\alpha (W(F,z_1) + \kappa_6)^{1-1/r}.$$

Under these assumptions, it is again possible to develop the abstract approach of Section 4. The natural ambient space is now $V = L^r(\Omega;\mathbb{R}^m)$ and, as in Theorem 5.1, one concludes the existence of a pair (z,ξ) fulfilling $\dot z \in L^r(0,T;L^r(\Omega;\mathbb{R}^m))$, $\xi \in L^{r'}(0,T;L^{r'}(\Omega;\mathbb{R}^m))$, and satisfying (5.7) and (5.8).

Coercivity and time-dependence of $\mathcal{E}$.

Lemma 5.5. *Assume* (5.3), (W_1)–(W_3), *and* (5.5). *Then* (5.11) *holds. Moreover, there exist positive constants $c_1, \ldots, c_6 > 0$ such that for all $(t,z) \in [0,T] \times D$ and all $\varphi \in M(t,z)$ we have*

$$c_1 \|\varphi\|^p_{W^{1,p}(\Omega;\mathbb{R}^d)} - c_2 \leq \mathcal{E}_t^2(z) \leq c_3, \tag{5.15}$$

$$\int_\Omega W(\nabla\varphi(x), z(x))\,\mathrm{d}x \leq c_4, \tag{5.16}$$

$$\mathcal{E}_t(z) \geq c_5 \|z\|^q_{W^{1,q}(\Omega;\mathbb{R}^m)} - c_6\,. \tag{5.17}$$

Further, for a sufficiently large constant E_0 (cf. (5.1)*), the energy functional $\mathcal{E}$ is bounded from below by a positive constant, it complies with* $(4.\mathrm{E}_0)$ *and* $(4.\mathrm{E}_1)$*, and for every (z_n), $z \subset L^2(\Omega;\mathbb{R}^m)$ we have*

$$\left(z_n \rightharpoonup z \ \text{ in } L^2(\Omega;\mathbb{R}^m) \text{ and } \sup_n \mathcal{E}^1(z_n) < +\infty \right) \Rightarrow \ z_n \to z \quad \text{in } \mathrm{C}^0(\overline{\Omega};\mathbb{R}^m). \tag{5.18}$$

Moreover,

$$\exists\, c_7 > 0 \ \ \forall\, t,\, s \in [0,T] \ \ \forall\, z \in D \,:\ \ |\mathcal{E}_t(z) - \mathcal{E}_s(z)| \leq c_7 |t-s|. \tag{5.19}$$

Hence, $\mathcal{E}$ fulfills $(4.\mathrm{E}_3)$.

Proof: We have for every $(t,\phi,z) \in [0,T]\times \mathcal{F} \times L^2(\Omega;\mathbb{R}^m)$

$$\begin{aligned} I_t(\phi,z) &\geq \kappa_1 \int_\Omega |\nabla\phi(x)|^p \,\mathrm{d}x - \kappa_2|\Omega| - \|\ell(t)\|_{W^{1,p}(\Omega;\mathbb{R}^d)^*}\, \|\phi\|_{W^{1,p}(\Omega;\mathbb{R}^d)} \\ &\geq c_1\|\phi\|^p_{W^{1,p}(\Omega;\mathbb{R}^d)} - \kappa_2|\Omega| - C'\|\ell\|^{p'}_{L^\infty(0,T;W^{1,p}(\Omega;\mathbb{R}^d)^*}, \end{aligned} \tag{5.20}$$

where the first inequality follows from the positivity of the functional $\mathcal{E}^1$ and from (W_2), and the second one from Poincaré's and Young's inequalities. Taking into account (5.3), we deduce the lower estimate in (5.15). Hence, it is sufficient to choose $E_0 := 2c_2$ in order to have $\mathcal{E}$ bounded from below by a positive constant.

Next, we remark that the functional $I_t(\cdot,z)$ is (sequentially) lower semicontinuous with respect to the weak topology of $W^{1,p}(\Omega;\mathbb{R}^d)$. Indeed, let $(\phi_k)_k$ weakly converge to some $\phi\in W^{1,p}(\Omega;\mathbb{R}^d)$ as $k\to\infty$. Then, by the weak continuity of minors of gradients (cf. [10, 42]), $\mathbb{M}(\nabla\phi_k) \rightharpoonup \mathbb{M}(\nabla\phi)$ in $L^{p/d}(\Omega;\mathbb{R}^{\mu_d})$. Taking into account the polyconvexity assumption (W_2), we ultimately have $\liminf_{k\to\infty} I_t(\phi_k,z) \geq I_t(\phi,z)$. We combine this weak lower semicontinuity property with the coercivity estimate (5.20), and thus we conclude that the set of minimizers (5.11) is not empty via the direct method of the calculus of variations.

Secondly, we observe that

$$\begin{aligned} \mathcal{E}^2_t(z) = \min_{\phi\in\mathcal{F}} I^2_t(\phi,z) &\leq \int_\Omega W(\nabla\phi_{\mathrm{Dir}}(x), z(x))\,\mathrm{d}x - \langle \ell(t),\phi_{\mathrm{Dir}}\rangle_{W^{1,p}} \\ &\leq \int_\Omega \max_{z\in\mathrm{K}} W(\nabla\phi_{\mathrm{Dir}}(x),z)\,\mathrm{d}x + C \doteq c_3 \end{aligned}$$

where the last inequality follows from (5.3) and (5.5). Hence, the upper estimate in (5.15) ensues. Then, (5.16) follows from

$$\begin{aligned} \int_\Omega W(\nabla\varphi,z)\,\mathrm{d}x &\leq \mathcal{E}^2_t(z) + \kappa_2|\Omega| + C\|\ell\|_{L^\infty(0,T;W^{1,p}(\Omega;\mathbb{R}^d)^*)}\|\varphi\|_{W^{1,p}(\Omega;\mathbb{R}^d)} \\ &\leq c_3 + \kappa_2|\Omega| + \left(\frac{c_3+c_2}{c_1}\right)^{1/p} \|\ell\|_{L^\infty(0,T;W^{1,p}(\Omega;\mathbb{R}^d)^*)}, \end{aligned}$$

where the first inequality is due to (5.20) and (5.3), and the second one to (5.15).

Next, in view of (5.20) we have for all $(t,z)\in[0,T]\times L^2(\Omega;\mathbb{R}^m)$ and for every $\varphi\in M(t,z)$

$$\mathcal{E}_t(z) \geq \frac{1}{q}\|\nabla z\|^q_{L^q(\Omega)} + \int_\Omega I_{\mathrm{K}}(z(x))\,\mathrm{d}x + c_1\|\varphi\|^p_{W^{1,p}(\Omega;\mathbb{R}^d)} - C.$$

Then, (5.17) ensues from the Poincaré inequality, and (5.18) follows from (5.17) and the fact that $q>d$, hence $W^{1,q}(\Omega;\mathbb{R}^m) \Subset C^0(\overline{\Omega};\mathbb{R}^m)$.

To prove $(4.\mathrm{E}_3)$, we observe that for all $z\in D$, for every $0\leq s\leq t\leq T$ and every $\varphi_t\in M(t,z)$ and $\varphi_s\in M(s,z)$ there holds

$$\begin{aligned} \mathcal{E}_t(z) - \mathcal{E}_s(z) = \mathcal{E}^2_t(z) - \mathcal{E}^2_s(z) &= I^2_t(\varphi_t,z) - I^2_s(\varphi_s,z) \\ &\leq I^2_t(\varphi_s,z) - I^2_s(\varphi_s,z) \\ &\leq -\langle \ell(t)-\ell(s),\varphi_s\rangle_{W^{1,p}} \\ &\leq C\|\ell(t)-\ell(s)\|_{W^{1,p}(\Omega;\mathbb{R}^d)^*}\|\varphi_s\|_{W^{1,p}(\Omega;\mathbb{R}^d)} \\ &\leq C\|\ell'\|_{L^\infty(0,T;W^{1,p}(\Omega;\mathbb{R}^d)^*)}\,|t-s|\,c_1^{-1/p}(\mathcal{E}^2_s(z)+c_2)^{1/p} \leq c_7|t-s| \end{aligned}$$

where we have used (5.3) and (5.15). Exchanging the roles of s and t, we infer (5.19). ∎

Properties of the dissipation potentials $(\Psi_z)_{z\in D}$.

Lemma 5.6. *Assume* (5.3), (W_1)–(W_3), *and* (ψ_1)–(ψ_3). *Then, the dissipation potentials* $(\Psi_z)_{z\in D}$ *satisfy* $(4.\Psi_1)$, $(4.\Psi_3)$, *and for every* $z\in D$ *we have*

$$\begin{cases} \Psi_z(v) \geq \kappa_7\|v\|^2_{L^2(\Omega;\mathbb{R}^m)} - \kappa_9|\Omega| \\ \Psi^*_z(w) \geq \kappa_8\|w\|^2_{L^2(\Omega;\mathbb{R}^m)} - \kappa_9|\Omega| \end{cases} \quad \text{for every } v,\, w\in L^2(\Omega;\mathbb{R}^m), \tag{5.21}$$

where κ_7, κ_8 *and* κ_9 *are the same constants as in* (ψ_3). *Thus,* $(4.\Psi_2)$ *is fulfilled.*

Proof: It follows from [12, Prop. 2.16, p. 47] that, for every $z \in D$ the subdifferential and conjugate of the potential Ψ_z are given for all v, $w \in L^2(\Omega;\mathbb{R}^m)$ by

$$\begin{cases} w \in \partial\Psi_z(v) \;\Leftrightarrow\; w(x) \in \partial_v\psi(z(x),v(x)) \;\text{ for a.a. } x \in \Omega, \\ \Psi_z^*(w) = \int_\Omega \psi^*(z(x),w(x))\,\mathrm{d}x. \end{cases} \tag{5.22}$$

Hence, (ψ_2) yields that $(\Psi_z)_{z\in D}$ is a family of *admissible dissipation potentials* on $L^2(\Omega;\mathbb{R}^m)$ in the sense of $(2.\Psi_1)$–$(2.\Psi_3)$, and (ψ_3) obviously implies (5.21). Finally, exploiting (ψ_1), (ψ_2), (5.18), and relying on Ioffe's theorem [27], it is not difficult to check that the first of $(4.\Psi_3)$ and (4.2) are fulfilled. This implies $(4.\Psi_3)$. ∎

Closedness and variational sum rule. We need the following preliminary result.

Lemma 5.7 (Subdifferentials)**.** *Assume* (5.3)*,* (W_1)*–*(W_3)*, and* (5.5)*. Then:*

(1) the subdifferential of $\mathcal{E}^1$ *is*

$$\partial\mathcal{E}^1(z) = -\Delta_q z + \partial I_{\mathrm{K}}(z) \quad \textit{for all } z \in \mathrm{dom}(\partial\mathcal{E}^1), \tag{5.23}$$

with domain described by the following conditions

$$z \in \mathrm{dom}(\partial\mathcal{E}^1) \;\Leftrightarrow\; \begin{cases} z \in D \subset W^{1,q}(\Omega;\mathbb{R}^m), \\ -\Delta_q z \in L^2(\Omega;\mathbb{R}^m), \end{cases} \tag{5.24}$$

hence

$$\mathrm{dom}(\partial\mathcal{E}^1) \subset W^{\nu,q}(\Omega;\mathbb{R}^m) \quad \textit{for all } \nu \in \Big[1, 1+\tfrac{1}{q}\Big). \tag{5.25}$$

(2) There exists a constant c_8 *such that for every* $(t,z) \in [0,T]\times D$ *and* $\varphi \in M(t,z)$ *we have* $\mathrm{D}_z W(\nabla\varphi, z) \in L^2(\Omega;\mathbb{R}^m)$*, with*

$$\|\mathrm{D}_z W(\nabla\varphi,z)\|_{L^2(\Omega;\mathbb{R}^m)} \le c_8. \tag{5.26}$$

Hence the marginal *subdifferential*

$$\widehat{\partial}\mathcal{E}^2_t(z) = \{\mathrm{D}_z W(\nabla\varphi,z) \;:\; \varphi \in M(t,z)\} \qquad \textit{is bounded in } L^2(\Omega;\mathbb{R}^m). \tag{5.27}$$

(3) For all $t \in [0,T]$ *and all* z_1*,* $z_2 \in D$ *with* $z_1(x)$*,* $z_2(x) \in \mathrm{K}'$ *for all* $x \in \Omega$*, and for every* $\varphi_1 \in M(t,z_1)$ *there holds*

$$\begin{aligned} &\mathcal{E}^2_t(z_2) - \mathcal{E}^2_t(z_1) - \int_\Omega \mathrm{D}_z W(\nabla\varphi_1, z_1)(z_2-z_1)\,\mathrm{d}x \\ &\qquad \le \kappa_5 \|z_1 - z_2\|^\alpha_{L^\infty(\Omega;\mathbb{R}^m)} \left(\int_\Omega W(\nabla\varphi_1,z_1)\,\mathrm{d}x + \kappa_6|\Omega|\right)^{1/2} \|z_1-z_2\|_{L^2(\Omega;\mathbb{R}^m)} \end{aligned} \tag{5.28}$$

where κ_5*,* κ_6*, and* α *are the same constants as in* (W_3)*.*

(4) For every $(t,z) \in [0,T]\times D$ *the Fréchet subdifferential* $\partial\mathcal{E}_t$ *satisfies*

$$\partial\mathcal{E}_t(z) \subset \mathrm{F}_t(z) = -\Delta_q z + \partial I_{\mathrm{K}}(z) + \{\mathrm{D}_z W(\nabla\varphi,z) \;:\; \varphi \in M(t,z)\}\,. \tag{5.29}$$

Proof: Formulae (5.23) and (5.24) can be obtained by adapting the proof of [51, Lemma 2.4], see also [12, Prop. 2.17]. Notice that (5.25) ensues from (5.24) and the regularity results in [49], cf. also [22]. We conclude (5.26) combining condition (W_3)(i) with estimate (5.16), and then (5.27) follows from trivial calculations.

Estimate (5.28) is a consequence of the following chain of inequalities

$$\begin{aligned} \mathcal{E}^2_t(z_2) - \mathcal{E}^2_t(z_1) &\le \int_\Omega \left(W(\nabla\varphi_1,z_2) - W(\nabla\varphi_1,z_1)\right)\mathrm{d}x \\ &= \int_\Omega\int_0^1 \mathrm{D}_z W(\nabla\varphi_1,(1-\theta)z_1+\theta z_2)(z_2-z_1)\,\mathrm{d}\theta\,\mathrm{d}x \\ &\le I + \int_\Omega \mathrm{D}_z W(\nabla\varphi_1,z_1)(z_2-z_1)\,\mathrm{d}x \end{aligned} \tag{5.30}$$

where, relying on condition (W$_3$)(ii), we estimate

$$\begin{aligned} I &= \int_0^1 \int_\Omega |\mathrm{D}_z W(\nabla\varphi_1, (1-\theta)z_1 + \theta z_2) - \mathrm{D}_z W(\nabla\varphi_1, z_1)||z_2 - z_1| \,\mathrm{d}x \,\mathrm{d}\theta \\ &\leq \int_0^1 \int_\Omega \kappa_5 \theta^\alpha |z_2 - z_1|^\alpha \left(W(\nabla\varphi_1, z_1) + \kappa_6\right)^{1/2} |z_2 - z_1| \,\mathrm{d}x \,\mathrm{d}\theta. \end{aligned} \tag{5.31}$$

Then, (5.28) follows upon using Hölder's inequality.

Finally, we prove (5.29), in fact in the following stronger form

$$\text{if } \partial\mathcal{E}_t(z) \neq \emptyset, \text{ then } \ \xi - \mathrm{D}_z W(\nabla\varphi, z) \in \partial\mathcal{E}^1(z) \text{ for every } \xi \in \partial\mathcal{E}_t(z) \text{ and } \varphi \in M(t,z).$$

which in particular yields (5.29). Indeed, we show that for every $\xi \in \partial\mathcal{E}_t(z)$ and $\varphi \in M(t,z)$, and for every $z_n \to z$ in $L^2(\Omega;\mathbb{R}^m)$, there holds

$$0 \leq \liminf_{n\to\infty} \frac{\mathcal{E}^1(z_n) - \mathcal{E}^1(z) - \int_\Omega (\xi - \mathrm{D}_z W(\nabla\varphi, z))(z_n - z)\,\mathrm{d}x}{\|z_n - z\|_{L^2(\Omega;\mathbb{R}^m)}} \doteq \Lambda. \tag{5.32}$$

To this aim, we observe that

$$\begin{aligned} \Lambda \geq \liminf_{n\to\infty} & \frac{\mathcal{E}_t(z_n) - \mathcal{E}_t(z) - \int_\Omega \xi(z_n - z)\,\mathrm{d}x}{\|z_n - z\|_{L^2(\Omega;\mathbb{R}^m)}} \\ & + \liminf_{n\to\infty} \frac{\mathcal{E}^2_t(z) - \mathcal{E}^2_t(z_n) + \int_\Omega \mathrm{D}_z W(\nabla\varphi, z)(z_n - z)\,\mathrm{d}x}{\|z_n - z\|_{L^2(\Omega;\mathbb{R}^m)}}. \end{aligned} \tag{5.33}$$

Since the first summand on the right-hand side of the above inequality is nonnegative by definition of the Fréchet subdifferential $\partial\mathcal{E}_t(z)$, it remains to prove that the second term is nonnegative. Now, it is not restrictive to suppose for the sequence (z_n) in (5.32) that $\sup \mathcal{E}^1(z_n) < +\infty$. Then, $z_n(x) \in \mathrm{K}$ for all $x \in \Omega$ and $n \in \mathbb{N}$. Hence, estimate (5.28) with the choices $z_1 = z$ and $z_2 = z_n$ yields

$$\begin{aligned} & \liminf_{n\to\infty} \frac{\mathcal{E}^2_t(z) - \mathcal{E}^2_t(z_n) + \int_\Omega \mathrm{D}_z W(\nabla\varphi, z)(z_n - z)\,\mathrm{d}x}{\|z_n - z\|_{L^2(\Omega;\mathbb{R}^m)}} \\ & \geq -\kappa_5 \lim_{n\to\infty} \frac{\|z_n - z\|^\alpha_{L^\infty(\Omega;\mathbb{R}^m)} \left(\int_\Omega W(\nabla\varphi, z)\,\mathrm{d}x + \kappa_6\right)^{1/2} \|z_n - z\|_{L^2(\Omega;\mathbb{R}^m)}}{\|z_n - z\|_{L^2(\Omega;\mathbb{R}^m)}} = 0, \end{aligned}$$

and the last limit follows from (5.18) and the bound (5.16). Ultimately, (5.32) ensues. ∎

Lemma 5.8 (Closedness)**.** *Assume* (5.3), (W$_1$)–(W$_3$), *and* (5.5). *Then, for all* $\{t_n\} \subset [0,T]$, $\{z_n\} \subset L^2(\Omega;\mathbb{R}^m)$, *and* $\{\xi_n\} \subset L^2(\Omega;\mathbb{R}^m)$ *with* $\xi_n \in \mathrm{F}_{t_n}(z_n)$ *for all* $n \in \mathbb{N}$, *we have*

$$\begin{aligned} & \big(t_n \to t, \quad z_n \rightharpoonup z \text{ in } L^2(\Omega;\mathbb{R}^m), \quad \xi_n \rightharpoonup \xi \text{ in } L^2(\Omega;\mathbb{R}^m), \quad \mathcal{E}_{t_n}(z_n) \to \mathscr{E} \quad \text{as } n \to \infty\big) \\ & \Longrightarrow \ \xi \in \mathrm{F}_t(z) \quad \text{and} \quad \mathscr{E} = \mathcal{E}_t(z). \end{aligned} \tag{5.34}$$

In particular, graph(F) *is a Borel set of* $[0,T] \times L^2(\Omega;\mathbb{R}^m) \times L^2(\Omega;\mathbb{R}^m)$.

Proof: From $\sup_n \mathcal{E}_{t_n}(z_n) < +\infty$ and from (5.15), (5.17), and (5.18), we deduce that $z_n \rightharpoonup z$ in $W^{1,q}(\Omega;\mathbb{R}^m)$ and $z_n \to z$ in $\mathrm{C}^0(\overline{\Omega};\mathbb{R}^m)$, and that there exist $\phi \in \mathcal{F}$ and a (not relabeled) subsequence (φ_n) such that $\varphi_n \rightharpoonup \phi$ in $W^{1,p}(\Omega;\mathbb{R}^d)$ as $n \to \infty$. Hence, we argue in the same way as in the proof of Lemma 5.5: combining the polyconvexity assumption (W$_2$) with the continuity of the map $z \mapsto W(F,z)$, we apply Ioffe's theorem [27] to find $\liminf_{n\to\infty} \int_\Omega W(\nabla\varphi_n, z_n)\,\mathrm{d}x \geq \int_\Omega W(\nabla\phi, z)\,\mathrm{d}x$. Therefore, in view of (5.3) we have

$$\begin{aligned} \liminf_{n\to\infty} \mathcal{E}^2_{t_n}(z_n) &= \liminf_{n\to\infty} \left(\int_\Omega W(\nabla\varphi_n, z_n)\,\mathrm{d}x - \ \langle \ell(t_n), \varphi_n \rangle_{W^{1,p}} \right) \\ &\geq \int_\Omega W(\nabla\phi, z)\,\mathrm{d}x - \ \langle \ell(t), \phi \rangle_{W^{1,p}} \geq \mathcal{E}^2_t(z). \end{aligned} \tag{5.35}$$

On the other hand, $\varphi_n \in M(t_n, z_n)$ gives

$$\mathcal{E}^2_{t_n}(z_n) = I^2_{t_n}(\varphi_n, z_n) \leq I^2_{t_n}(\varphi, z_n) \to I^2_t(\varphi, z) = \mathcal{E}^2_t(z), \tag{5.36}$$

where φ is any element in $M(t,z)$, and we have exploited (5.3) to take the limit as $n \to \infty$. Combining (5.35) and (5.36), we ultimately have $\mathcal{E}^2_{t_n}(z_n) \to \mathcal{E}^2_t(z)$, and the weak limit ϕ of the sequence (φ_n) is in fact an element in $M(t,z)$, which we will hereafter denote with φ.

Now, it follows from (5.12), (5.23), and (5.27) that the sequence $\xi_n \in \mathrm{F}_{t_n}(z_n)$ in (5.34) is given, for every $n \in \mathbb{N}$, by $\xi_n = -\Delta_q z_n + \zeta_n + \mathrm{D}_z W(\nabla\varphi_n, z_n)$, for some $\zeta_n \in \partial I_{\mathrm{K}}(z_n)$ and $\varphi_n \in M(t_n, z_n)$. Arguing by comparison and relying on the aforementioned [12, Prop. 2.17], from the boundedness of ξ_n in $L^2(\Omega;\mathbb{R}^m)$ we infer that

$$\sup_n \big(\|\Delta_q z_n\|_{L^2(\Omega;\mathbb{R}^m)} + \|\zeta_n\|_{L^2(\Omega;\mathbb{R}^m)}\big) < +\infty. \tag{5.37}$$

Relying on [49], we find that for every $\nu \in [1, 1+1/q)$ there holds $\sup_n \|z_n\|_{W^{\nu,q}(\Omega;\mathbb{R}^m)} < +\infty$. Since $W^{\nu,q}(\Omega;\mathbb{R}^m) \Subset W^{1,q}(\Omega;\mathbb{R}^m)$ for all $\nu \in (1, 1+1/q)$, we conclude that, indeed, the weak convergence of (z_n) in $L^2(\Omega;\mathbb{R}^m)$ improves to

$$z_n \to z \quad \text{in } W^{1,q}(\Omega;\mathbb{R}^m). \tag{5.38}$$

Therefore, $\mathcal{E}^1(z_n) \to \mathcal{E}^1(z)$. On account of the previously proved convergence of $\mathcal{E}^2_{t_n}(z_n)$, we obtain $\mathcal{E}_{t_n}(z_n) \to \mathcal{E}_t(z)$. Finally, combining estimate (5.37) with (5.38), and exploiting the monotonicity of the operator $-\Delta_q$ (cf. [12]), we find

$$-\Delta_q z_n \rightharpoonup -\Delta_q z \quad \text{in } L^2(\Omega;\mathbb{R}^m). \tag{5.39}$$

Furthermore, from (5.37) we also deduce that, up to a not relabeled subsequence,

$$\zeta_n \rightharpoonup \zeta \quad \text{in } L^2(\Omega;\mathbb{R}^m), \text{ with } \ \zeta \in \partial I_{\mathrm{K}}(z) \tag{5.40}$$

(the latter fact follows from the strong-weak closedness of the graph of ∂I_{K} in $L^2(\Omega;\mathbb{R}^m) \times L^2(\Omega;\mathbb{R}^m)$).

Estimate (5.26) yields

$$\sup_n \|\mathrm{D}_z W(\nabla\varphi_n, z_n)\|_{L^2(\Omega;\mathbb{R}^m)} \le c_8. \tag{5.41}$$

Then, along some (not relabeled) subsequence, the sequence $(\mathrm{D}_z W(\nabla\varphi_n, z_n))_n$ is weakly converging in $L^2(\Omega;\mathbb{R}^m)$. It remains to prove that

$$\mathrm{D}_z W(\nabla\varphi_n, z_n) \rightharpoonup \mathrm{D}_z W(\nabla\varphi, z) \quad \text{in } L^2(\Omega;\mathbb{R}^m). \tag{5.42}$$

To this aim, we mimick the argument in the proof of [24, Prop. 3.3]. We fix $\eta \in W^{1,q}(\Omega;\mathbb{R}^m)$ and $h > 0$, and apply (W_3)(ii) with the choices $z_1 = z_n$ and $z_2 = z_n + h\eta$. Indeed, $z_n(x) \in \mathrm{K}$ and (5.38) ensure that, for sufficiently large n and sufficiently small h, we have $z_n,\, z_n + h\eta \in \mathrm{K}'$. Arguing like in (5.30)–(5.31) and exploiting estimate (5.16), there exist constants $C > 0$ and $\alpha \in (0,1]$ such that for every $n \in \mathbb{N}$

$$\begin{aligned} \Big|\frac{1}{h}\int_\Omega \big(W(\nabla\varphi_n, z_n \pm h\eta) - W(\nabla\varphi_n, z_n) \mp h\mathrm{D}_z W(\nabla\varphi_n, z_n)\eta\big)\,\mathrm{d}x\Big| \\ \le C h^\alpha \|\eta\|^\alpha_{L^\infty(\Omega;\mathbb{R}^m)} \|\eta\|_{L^2(\Omega;\mathbb{R}^m)} \doteq \omega(h). \end{aligned} \tag{5.43}$$

On the other hand, again combining (5.38) and the weak convergence of φ_n with Ioffe's theorem, we conclude that for (sufficiently small) $h > 0$ there holds

$$\liminf_{n\to\infty} \int_\Omega \big(W(\nabla\varphi_n, z_n \pm h\eta) - W(\nabla\varphi_n, z_n)\big)\,\mathrm{d}x \ge \frac{1}{h}\int_\Omega \big(W(\nabla\varphi, z \pm h\eta) - W(\nabla\varphi, z)\big)\,\mathrm{d}x.$$

Estimate (5.43) and the above inequality yield

$$\begin{aligned} &\limsup_{n\to\infty} \int_\Omega \mathrm{D}_z W(\nabla\varphi_n, z_n)\eta\,\mathrm{d}x \\ &\quad\le \limsup_{n\to\infty} \frac{1}{h}\int_\Omega \big(W(\nabla\varphi_n, z_n) - W(\nabla\varphi_n, z_n - h\eta)\big)\,\mathrm{d}x + \omega(h) \\ &\quad\le -\frac{1}{h}\int_\Omega \big(W(\nabla\varphi, z - h\eta) - W(\nabla\varphi, z)\big)\,\mathrm{d}x + \omega(h) \le \int_\Omega \mathrm{D}_z W(\nabla\varphi, z)\eta\,\mathrm{d}x + 2\omega(h), \end{aligned} \tag{5.44}$$

where the last inequality follows from (5.43) written for $(\nabla\varphi, z)$. Analogously, we infer that

$$\liminf_{n\to\infty} \int_\Omega \mathrm{D}_z W(\nabla\varphi_n, z_n)\eta \,\mathrm{d}x \geq \int_\Omega \mathrm{D}_z W(\nabla\varphi, z)\eta \,\mathrm{d}x - 2\omega(h).$$

Since $h > 0$ is arbitrary, we conclude that

$$\lim_{n\to\infty} \int_\Omega \mathrm{D}_z W(\nabla\varphi_n, z_n)\eta \,\mathrm{d}x = \int_\Omega \mathrm{D}_z W(\nabla\varphi, z)\eta \,\mathrm{d}x \quad \text{for every } \eta \in W^{1,q}(\Omega;\mathbb{R}^m).$$

In view of (5.41), (5.42) follows. Thus, (5.39), (5.40) and (5.42) entail that the weak limit ξ of (ξ_n) fulfills $\xi \in \mathrm{F}_t(z)$, and (5.34) ensues.

Finally, let us observe that $\mathrm{graph}(\mathrm{F}) = \cup_{m\in\mathbb{N}} G^m$, with

$$G^m = \left\{(t,u,\xi) \in [0,T] \times L^2(\Omega;\mathbb{R}^m) \times L^2(\Omega;\mathbb{R}^m) \, : \, \xi \in \mathrm{F}_t(u), \ |\mathcal{E}_t(u)| \leq m\right\}.$$

Now, it follows from the closedness property (5.34) that every G_m is a closed, hence Borelian, set. Hence, graph(F) is a Borel set. ∎

Corollary 5.9 (Variational sum rule)**.** *Assume* (5.3)*,* (W_1)*–*(W_3)*,* (ψ_1)*–*(ψ_3)*, and* (5.5)*. Then, the dissipation potentials* $(\Psi_z)_{z\in D}$ *and the reduced energy functional* $\mathcal{E}$ *comply with the variational sum rule* $(4.\mathrm{E}_2)$*.*

Proof: This follows from Lemmas 5.6, 5.7, and 5.7, combined with Proposition 4.2. ∎

Chain rule.

Lemma 5.10 (Chain rule)**.** *Assume* (5.3)*,* (W_1)*–*(W_3)*,* (ψ_1)*–*(ψ_3)*, and* (5.5)*. Then, the function* $\mathrm{P} : \mathrm{graph}(\mathrm{F}) \to \mathbb{R}$ *defined in* (5.13) *complies with* $(4.\mathrm{E}_4)$*. Moreover, the system* $(V, \mathcal{E}, \Psi, \mathrm{F}, \mathrm{P})$ *fulfills the closedness condition* $(4.\mathrm{E}_6)$*, and the chain-rule inequality* $(4.\mathrm{E}_5)$*.*

Proof: We first observe that

$$R(t,z,\xi) \ \text{ is weakly sequentially compact in } W^{1,p}(\Omega;\mathbb{R}^d) \text{ for every } (t,z,\xi) \in \mathrm{graph}(\mathrm{F}). \tag{5.45}$$

Indeed, every sequence $(\varphi_n)_n \subset R(t,z,\xi)$ is bounded in $W^{1,p}(\Omega;\mathbb{R}^d)$ thanks to (5.15). Hence, up to a subsequence it converges to some φ. From the arguments in Lemma 5.8 it follows that $\varphi \in R(t,z,\xi)$. Thus, it is immediate to see that the maximum in formula (5.13) is attained.

For every $(t,z) \in [0,T] \times W^{1,q}(\Omega)$, $h \in (0, T-t]$, and $\varphi(t) \in M(t,z)$ there holds

$$\frac{\mathcal{E}_{t+h}(z) - \mathcal{E}_t(z)}{h} = \frac{\mathcal{E}^2_{t+h}(z) - \mathcal{E}^2_t(z)}{h} \leq \frac{1}{h}\left\langle -\ell(t+h) + \ell(t), \varphi(t)\right\rangle_{W^{1,p}},$$

whence $\limsup_{h\downarrow 0} \frac{\mathcal{E}_{t+h}(z)-\mathcal{E}_t(z)}{h} \leq \mathrm{P}_t(z,\xi)$. On the other hand, it follows from (5.3) and (5.15) that $|\mathrm{P}_t(z,\xi)| \leq \|\ell'(t)\|_{W^{1,p}(\Omega;\mathbb{R}^d)^*} \cdot \sup_{\varphi\in M(t,z)} \|\varphi\|_{W^{1,p}(\Omega;\mathbb{R}^d)} \leq C$. Therefore, $(4.\mathrm{E}_4)$ is fulfilled.

Combining the previously proved closedness property (5.34) with arguments analogous to those developed for (5.45), it is possible to check that $(4.\mathrm{E}_6)$ holds in a slightly stronger form, viz.

$$\begin{aligned}
&\big(t_n \to t, \ u_n \to u \text{ in } V, \ \mathrm{F}_{t_n}(u_n) \ni \xi_n \rightharpoonup \xi \text{ in } V^*, \ \ \mathrm{P}_{t_n}(u_n,\xi_n) \to p, \ \mathcal{E}_{t_n}(u_n) \to \mathscr{E}\big) \\
&\implies (t,u) \in \mathrm{dom}(F), \quad \xi \in \mathrm{F}_t(u), \quad p \leq \mathrm{P}_t(u,\xi), \quad \mathscr{E} = \mathcal{E}_t(u).
\end{aligned}$$

Hence, mimicking the argument at the end of the proof of Lemma 5.8, it is possible to check that for every $\lambda \in \mathbb{R}$, the set $\mathrm{P}^{-1}([\lambda,+\infty))$ is a Borel set of $[0,T]\times V \times V^*$. Therefore, $\mathrm{P} : \mathrm{graph}(\mathrm{F}) \to \mathbb{R}$ defined by (5.13) is a Borel function.

Finally, in order to prove that the chain rule $(4.\mathrm{E}_5)$ is fulfilled, let us fix a curve $z \in \mathrm{AC}([0,T];L^2(\Omega))$ and a function $\xi \in L^1(0,T;L^2(\Omega))$ fulfilling (4.3) and (4.4). Taking into account (5.17) and (5.21), we have a fortiori that

$$\begin{gathered}
z \in L^\infty(0,T;W^{1,q}(\Omega;\mathbb{R}^m)) \cap H^1(0,T;L^2(\Omega;\mathbb{R}^m)) \subset \mathrm{C}^0([0,T];L^\infty(\Omega;\mathbb{R}^m)), \\
\xi \in L^2(0,T;L^2(\Omega;\mathbb{R}^m)).
\end{gathered} \tag{5.46}$$

Furthermore, there exist measurable selections $t \mapsto \zeta(t) \in \partial I_{\mathrm{K}}(z(t))$ and $t \mapsto \varphi(t) \in M(t,z(t))$ such that

$$\xi(t) = -\Delta_q z(t) + \zeta(t) + \mathrm{D}_z W(\nabla\varphi(t), z(t)) \quad \text{for a.a. } t \in (0,T). \tag{5.47}$$

Arguing as in the proof of Lemma 5.8, from $\xi \in L^2(0,T;L^2(\Omega;\mathbb{R}^m))$ we deduce that

$$\|\Delta_q z(t)\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^m))} + \|\zeta(t)\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^m))} + \|\mathrm{D}_z W(\nabla\varphi(t), z(t))\|_{L^\infty(0,T;L^2(\Omega;\mathbb{R}^m))} \leq C, \tag{5.48}$$

where the latter estimate follows from (5.26). Thus, the chain rule for the *convex* functional $\mathcal{E}^1$ (see [12]) yields that

$$\begin{aligned} &\text{the map } t \mapsto \mathcal{E}^1(z(t)) \text{ is absolutely continuous, and} \\ &\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}^1(z(t)) = \int_\Omega (-\Delta_q z(t) + \zeta(t)) z'(t)\,\mathrm{d}x \quad \text{for a.a. } t \in (0,T). \end{aligned} \tag{5.49}$$

As for the map $t \mapsto \mathcal{E}^2_t(z(t))$, there exist constants $C > 0$, $\alpha \in (0,1]$ such that for every $0 \leq s \leq t \leq T$ we have

$$\begin{aligned} &\mathcal{E}^2_t(z(t)) - \mathcal{E}^2_s(z(s)) \\ &= \mathcal{E}^2_t(z(t)) - \mathcal{E}^2_t(z(s)) + \mathcal{E}^2_t(z(s)) - \mathcal{E}^2_s(z(s)) \\ &\leq \mathcal{E}^2_t(z(t)) - \mathcal{E}^2_t(z(s)) + I^2_t(\varphi(s), z(s)) - I^2_s(\varphi(s), z(s)) \\ &\leq C\|z(t)-z(s)\|^\alpha_{L^\infty(\Omega;\mathbb{R}^m)}\|z(t)-z(s)\|_{L^2(\Omega;\mathbb{R}^m)} \\ &\quad + \int_\Omega \mathrm{D}_z W(\nabla\varphi(s)), z(s))(z(t)-z(s))\,\mathrm{d}x - \langle \ell(t)-\ell(s), \varphi(s)\rangle_{W^{1,p}}. \end{aligned}$$

where the second inequality follows from estimate (5.28) with $z_1 = z(s)$ and $z_2 = z(t)$, also taking into account (5.16). Exchanging the role of s and t, we thus conclude for every $0 \leq s \leq t \leq T$

$$\begin{aligned} &|\mathcal{E}^2_t(z(t)) - \mathcal{E}^2_s(z(s))| \\ &\leq C\|z(t)-z(s)\|_{L^2(\Omega;\mathbb{R}^m)}\big(\|z(t)-z(s)\|^\alpha_{L^\infty(\Omega;\mathbb{R}^m)} + 2\sup_{t\in(0,T)}\|\mathrm{D}_z W(\nabla\varphi(t), z(t))\|_{L^2(\Omega;\mathbb{R}^m)}\big) \\ &\quad + \|\ell(t)-\ell(s)\|_{W^{1,p}(\Omega;\mathbb{R}^d)^*}(\|\varphi(t)\|_{W^{1,p}(\Omega;\mathbb{R}^d)} + \|\varphi(s)\|_{W^{1,p}(\Omega;\mathbb{R}^d)}) \\ &\leq C'\left(\|z(t)-z(s)\|_{L^2(\Omega;\mathbb{R}^m)} + |t-s|\right), \end{aligned}$$

where the second inequality follows from (5.46), (5.3), and estimates (5.15) and (5.48). Thus, the map $t \mapsto \mathcal{E}^2_t(z(t))$ is absolutely continuous.

Finally, let us fix $t \in (0,T)$, such that formula (5.49) for $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}^1(z(t))$ holds, $\frac{z(t+h)-z(t)}{h} \to z'(t)$ in $L^2(\Omega)$, $\ell(t+h)-\ell(t) \to \ell'(t)$ in $W^{1,p}(\Omega;\mathbb{R}^d)^*$, and $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}^2_t(z(t))$ exists (the set of such t's has full measure). Now, in view of (5.28), and again taking into account (5.46) and (5.16), for all $h \in (-t,0]$ and $\tilde\varphi(t) \in R(t,z(t),\xi(t))$ there holds

$$\begin{aligned} &\frac{1}{h}(\mathcal{E}^2_{t+h}(z(t+h)) - \mathcal{E}^2_t(z(t))) \\ &\geq C\|z(t+h)-z(t)\|^\alpha_{L^\infty(\Omega;\mathbb{R}^m)}\frac{\|z(t+h)-z(t)\|_{L^2(\Omega;\mathbb{R}^m)}}{h} \\ &\quad + \frac{1}{h}\int_\Omega \mathrm{D}_z W(\nabla\tilde\varphi(t), z(t))(z(t+h)-z(t))\,\mathrm{d}x - \frac{1}{h}\langle \ell(t+h)-\ell(t), \tilde\varphi(t)\rangle_{W^{1,p}}. \end{aligned}$$

Taking the $\lim_{h\uparrow 0}$ in the above inequality and using that $z \in \mathrm{C}^0([0,T];L^\infty(\Omega;\mathbb{R}^m))$ by (5.46), we conclude that for every $\tilde\varphi(t) \in R(t,z(t),\xi(t))$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}^2_t(z(t)) \geq \int_\Omega \mathrm{D}_z W(\nabla\tilde\varphi(t), z(t))z'(t)\,\mathrm{d}x - \langle \ell'(t), \tilde\varphi(t)\rangle_{W^{1,p}}. \tag{5.50}$$

Now, from the definition of $R(t,z(t),\xi(t))$ it follows that, in correspondence to the map $t \mapsto \tilde\varphi(t)$, there exists a selection $t \mapsto \tilde\zeta(t) \in \partial I_{\mathrm{K}}(z(t))$ such that $\tilde\zeta(t) + \mathrm{D}_z W(\nabla\tilde\varphi(t), z(t)) = \zeta(t) + \mathrm{D}_z W(\nabla\varphi(t), z(t))$ for almost all $t \in (0,T)$ (where ζ and φ are the selections in (5.47)). Thus, using the chain rule for I_{K}, we have

$$\int_\Omega \mathrm{D}_z W(\nabla\tilde\varphi(t), z(t))z'(t)\,\mathrm{d}x = \int_\Omega \mathrm{D}_z W(\nabla\varphi(t), z(t))z'(t)\,\mathrm{d}x + \underbrace{\int_\Omega \zeta(t)z'(t)\,\mathrm{d}x - \int_\Omega \tilde{\zeta(t)}z'(t)\,\mathrm{d}x}_{=0}.$$

Since the selection $t \mapsto \tilde{\varphi}(t) \in R(t, z(t), \xi(t))$ in (5.50) is arbitrary, from the above equality we ultimately conclude

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t^2(z(t)) \geq \int_\Omega \mathrm{D}_z W(\nabla\varphi(t), z(t))z'(t)\,\mathrm{d}x + \mathrm{P}_t(z(t), \xi(t)) \quad \text{for a.a. } t \in (0,T). \tag{5.51}$$

Combining (5.49) and (5.51), we obtain (4.E$_5$). ■

Thus, we have shown that all the abstract assumptions of Section 4.1 are fulfilled, which implies that Theorem 5.1 follows from Theorem 4.4.

6. Proofs

Plan of the proof of Theorem 4.4. First, in Section 6.1 we provide some "stationary estimates" on every single step of the incremental minimization scheme. In particular, in Lemma 6.1 we prove the crucial energy inequality (6.7), which is the starting point for the a priori estimates on the approximate solutions. We prove the latter estimates in Proposition 6.3. Hence we deduce in Proposition 6.4 that, along some subsequence, the approximate solutions converge to a curve $u \in \mathrm{AC}([0,T];V)$. In Section 6.3 we conclude the proof of Theorem 4.4, showing that u is in fact a solution of the Cauchy problem for (1.21). In doing so, we rely on some technical results proved in the Appendix.

6.1. **Discrete energy inequality.** In the following, we gain further insight into problems (4.9) and (4.12) (which give rise to approximate solutions), by fixing some crucial properties of the general minimization problem

$$\mathcal{I}_{\mathsf{t},r}(u) := \inf_{v\in D}\left\{ r\Psi_u\left(\frac{v-u}{r}\right) + \mathcal{E}_{\mathsf{t}+r}(v)\right\} \qquad \text{for given } \mathsf{t} \in [0,T],\ u \in D,\ 0 < r < T-\mathsf{t}. \tag{6.1}$$

The following result is the Banach-space counterpart to [43, Lemmas 4.4,4.5] (see also [1, 2, 44]).

Lemma 6.1. *Assume* (4.Ψ_1), *and* (4.E$_0$)–(4.E$_4$). *Then, for every* $\mathsf{t} \in [0,T]$, $u \in D$, *and for all* $0 < r < T - \mathsf{t}$

$$\textit{the set } \ \mathrm{A}_{\mathsf{t},r}(u) := \operatorname*{Argmin}_{v\in D}\left\{ r\Psi_u\left(\frac{v-u}{r}\right) + \mathcal{E}_{\mathsf{t}+r}(v)\right\} \quad \textit{is nonempty.} \tag{6.2}$$

Moreover, for all $\mathsf{t} \in [0,T]$ *there exists a measurable selection* $r \mapsto u_r \in \mathrm{A}_{\mathsf{t},r}(u)$ *such that*

$$0 \in \partial\Psi_u\left(\frac{u_r-u}{r}\right) + \mathrm{F}_{\mathsf{t}+r}(u_r). \tag{6.3}$$

Further, there holds

$$\forall\, \mathsf{t} \in [0,T],\ u \in D,\ 0 < r < T-\mathsf{t},\ u_r \in \mathrm{A}_{\mathsf{t},r}(u): \quad \mathcal{G}(u_r) \leq C_3\mathcal{G}(u), \tag{6.4}$$

$$\lim_{r\downarrow 0}\ \sup_{u_r\in \mathrm{A}_{\mathsf{t},r}(u)} \|u_r - u\| = 0, \qquad \lim_{r\downarrow 0}\mathcal{I}_{\mathsf{t},r}(u) = \mathcal{E}_{\mathsf{t}}(u) \quad \textit{for all } \mathsf{t} \in [0,T],\ u \in D\,, \tag{6.5}$$

with C_3 *the constant in* (4.5). *Finally,*

$$\textit{the map } (0, T-\mathsf{t}) \ni r \mapsto \mathcal{I}_{\mathsf{t},r}(u) \textit{ is a.e. differentiable in } (0,T-\mathsf{t}) \tag{6.6}$$

and for every $r_0 \in (0, T-\mathsf{t})$ *and every measurable selection* $r \in (0, r_0] \mapsto u_r \in \mathrm{A}_{\mathsf{t},r}(u)$ *there holds*

$$r_0\Psi_u\left(\frac{u_{r_0}-u}{r_0}\right) + \int_0^{r_0}\Psi_u^*(-\xi_r)\,\mathrm{d}r + \mathcal{E}_{\mathsf{t}+r_0}(u_{r_0}) \leq \mathcal{E}_{\mathsf{t}}(u) + \int_0^{r_0}\mathrm{P}_{\mathsf{t}+r}(u_r,\xi_r)\,\mathrm{d}r \tag{6.7}$$

where ξ_r *is* any *selection in* $\mathrm{F}_{\mathsf{t}+r}(u_r) \cap \big(-\partial\Psi_u(\frac{u_r-u}{r})\big)$.

Proof: The *direct method of the calculus of variations* gives (6.2) because of the coercivity condition (4.E$_1$). Further, [14, Cor. III.3, Thm. III.6] guarantee the existence of a measurable selection $(0,+\infty) \ni r \mapsto u_r \in \mathrm{A}_{\mathsf{t},r}(u)$, which complies with the Euler equation (6.3) thanks to (4.E$_2$).

Estimate (6.4) follows from the chain of inequalities

$$\mathcal{G}(u) \geq r\Psi_u\left(\frac{v-u}{r}\right) + \mathcal{E}_{\mathsf{t}+r}(u_r) \geq \mathcal{E}_{\mathsf{t}+r}(u_r) \geq \frac{1}{C_3}\mathcal{G}(u_r), \tag{6.8}$$

where the first one is due to the minimality of u_r, and the third one to (4.5). We refer to [43, Lemma 4.4] for the proof of (6.5), only pointing out that the first limit in (6.5) follows from the superlinear growth of Ψ_u.

Then, to check (6.6) we fix $0 < r_1 < r_2$ and remark that

$$\begin{aligned}
&\mathcal{I}_{\mathsf{t},r_2}(u) - \mathcal{I}_{\mathsf{t},r_1}(u) - (\mathcal{E}_{\mathsf{t}+r_2}(u_{r_1}) - \mathcal{E}_{\mathsf{t}+r_1}(u_{r_1}))\\
&\quad\leq r_2\Psi_u\left(\frac{u_{r_1}-u}{r_2}\right) - r_1\Psi_u\left(\frac{u_{r_1}-u}{r_1}\right)\\
&\quad\leq (r_2-r_1)\Psi_u\left(\frac{u_{r_1}-u}{r_2}\right) + r_1\left(\Psi_u\left(\frac{u_{r_1}-u}{r_2}\right) - \Psi_u\left(\frac{u_{r_1}-u}{r_1}\right)\right)\\
&\quad\leq (r_2-r_1)\left(\Psi_u\left(\frac{u_{r_1}-u}{r_2}\right) - \left\langle w_2^1, \frac{u_{r_1}-u}{r_2}\right\rangle\right) = -(r_2-r_1)\Psi_u^*(w_2^1)
\end{aligned} \tag{6.9}$$

where the first inequality follows from (6.1), the second one from algebraic manipulations, the third one by choosing some $w_2^1 \in \partial\Psi_u((u_{r_1}-u)/r_2)$ (which is nonempty, cf. (2.3)), and the last passage from an elementary convex analysis identity. Since $-(r_2-r_1)\Psi_u^*(w_2^1) \leq 0$ by (2.2), we conclude that

$$\begin{aligned}
\mathcal{I}_{\mathsf{t},r_2}(u) &\leq \mathcal{I}_{\mathsf{t},r_1}(u) + \mathcal{E}_{\mathsf{t}+r_2}(u_{r_1}) - \mathcal{E}_{\mathsf{t}+r_1}(u_{r_1})\\
&\leq \mathcal{I}_{\mathsf{t},r_1}(u) + C_1(r_2-r_1)\mathcal{G}(u_{r_1}) \leq \mathcal{I}_{\mathsf{t},r_1}(u) + C_1(r_2-r_1)C_3\mathcal{G}(u)\,,
\end{aligned} \tag{6.10}$$

the second inequality thanks to (4.E$_3$), and the third one to (6.4). Therefore, the map $r \mapsto \mathcal{I}_{\mathsf{t},r}(u)$ is given by the sum of a nonincreasing and of an absolutely continuous function, whence we deduce that it is almost everywhere differentiable, viz. (6.6). In order to conclude (6.7), we fix $r \in (0, T-t)$, outside a negligible set, such that r is a differentiability point of the map $r \mapsto \mathcal{I}_{\mathsf{t},r}(u)$, and we consider a selection $w_h^r \in \partial\Psi_u((u_r-u)/(r+h))$ for $h > 0$ sufficiently small. We also fix a sequence $h_k \downarrow 0$ such that

$$\liminf_{h_k\downarrow 0} \frac{\mathcal{E}_{\mathsf{t}+r+h_k}(u_r) - \mathcal{E}_{\mathsf{t}+r}(u_r)}{h_k} = \liminf_{h\downarrow 0} \frac{\mathcal{E}_{\mathsf{t}+r+h}(u_r) - \mathcal{E}_{\mathsf{t}+r}(u_r)}{h}\,. \tag{6.11}$$

Since $\partial\Psi_u : V \rightrightarrows V^*$ is a bounded operator, from (6.5) we easily deduce that $\|w_{h_k}^r\|_* \leq C$, so that there exist $w_r \in \partial\Psi((u_r-u)/r)$ and a subsequence such that $w_{h_j}^r \rightharpoonup w_r$ in V^*. Then, we find that

$$\begin{aligned}
&\Psi_u^*(w_r) \leq \liminf_{h_j\downarrow 0} \Psi_u^*(w_{h_j}^r) \leq \limsup_{h_j\downarrow 0} \Psi_u^*(w_{h_j}^r)\\
&\leq \lim_{h_j\downarrow 0}\left\langle w_{h_j}^r, \frac{u_r-u}{r+h_j}\right\rangle - \liminf_{h_j\downarrow 0}\Psi_u\left(\frac{u_r-u}{r+h_j}\right) \leq \left\langle w_r, \frac{u_r-u}{r}\right\rangle - \Psi_u\left(\frac{u_r-u}{r}\right) = \Psi_u^*(w_r)\,,
\end{aligned}$$

using an elementary convex analysis identity and that both Ψ_u^* is weakly lower semicontinuous on V^* and Ψ_u weakly lower semicontinuous on V. Therefore $\lim_j \Psi_u^*(w_{h_j}^r) = \Psi_u^*(w_r)$. Since the limit is independent of the subsequence, we conclude that, for the whole sequence $w_{h_k}^r$ there holds

$$\lim_{h_k\downarrow 0}\Psi_u^*(w_{h_k}^r) = \Psi_u^*(w_r) = \Psi_u^*(-\xi_r) \quad \text{for all } \xi_r \in \mathrm{F}_{\mathsf{t}+r}(u_r) \cap \left(-\partial\Psi_u\left(\frac{u_r-u}{r}\right)\right)\,,$$

the last identity thanks to condition (2.Ψ_3). Then, from (6.9) we deduce

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{d}r}\mathcal{I}_{\mathsf{t},r}(u)|_{r=r} + \Psi_u^*(-\xi_r) &= \lim_{h_k\downarrow 0}\left(\frac{\mathcal{I}_{\mathsf{t},r+h_k}(u) - \mathcal{I}_{\mathsf{t},r}(u)}{h_k} + \Psi_u^*(w_{h_k}^r)\right)\\
&\leq \liminf_{h_k\downarrow 0}\frac{\mathcal{E}_{\mathsf{t}+r+h_k}(u_r) - \mathcal{E}_{\mathsf{t}+r}(u_r)}{h_k} \leq \mathrm{P}_{\mathsf{t+r}}(u_r, \xi_r)\,,
\end{aligned}$$

the latter inequality due to (6.11) and (4.E$_4$). Since r is arbitrary, we ultimately find

$$\tfrac{\mathrm{d}}{\mathrm{d}r}\mathcal{I}_{\mathsf{t},r}(u)|_{r=r} + \Psi_u^*(-\xi_r) \leq \mathrm{P}_{\mathsf{t}+r}(u_r, \xi_r) \qquad \text{for a.a. } r \in (0, T-t)\,. \tag{6.12}$$

Hence, (6.7) follows from integrating (6.12) on the interval $(0, r_0)$, also using the second of (6.5). ∎

Remark 6.2. Under assumption (4.23) as a replacement of (2.Ψ_3), it is possible to prove inequality (6.12) in the following way. We obtain the differentiability property (6.6) in the same way as throughout (6.9)–(6.11) and then we observe that, for a fixed $r \in (0, T-t)$ outside a negligible set, such that r is a differentiability point of the map $r \mapsto \mathcal{J}_{\mathsf{t},r}(u)$, we have the following chain of inequalities for all $h > 0$ (in which we have set $\tilde u_{r,h} = u + \frac{r+h}{r}(u_r - u)$):

$$\begin{aligned}\mathcal{J}_{\mathsf{t},r+h}(u) - \mathcal{J}_{\mathsf{t},r}(u) &\leq \mathcal{E}_{\mathsf{t}+r+h}(\tilde u_{r,h}) - \mathcal{E}_{\mathsf{t}+r}(u_r) + (r+h)\Psi_u\left(\frac{\tilde u_{r,h} - u}{r+h}\right) - r\Psi_u\left(\frac{u_r - u}{r}\right) \\ &= \mathcal{E}_{\mathsf{t}+r+h}(\tilde u_{r,h}) - \mathcal{E}_{\mathsf{t}+r}(\tilde u_{r,h}) + \mathcal{E}_{\mathsf{t}+r}(\tilde u_{r,h}) - \mathcal{E}_{\mathsf{t}+r}(u_r) + h\Psi_u\left(\frac{u_r - u}{r}\right),\end{aligned}$$

where in the second passage we have used that $\Psi_u(\frac{\tilde u_{r,h}-u}{r+h}) = \Psi_u(\frac{u_r-u}{r})$. Then, upon dividing the above inequality by $h > 0$ and taking the $\limsup$ as $h \downarrow 0$, (4.23) and (4.24) yield (recall that $\delta\mathcal{E}_t(u;v) = \lim_{h\downarrow 0}\frac{1}{h}(\mathcal{E}_t(u+hv) - \mathcal{E}_t(u))$)

$$\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{d}r}\mathcal{J}_{\mathsf{t},r}(u)|_{r=r} &\leq \liminf_{h\downarrow 0}\frac{1}{h}\left(\mathcal{E}_{\mathsf{t}+r+h}(\tilde u_{r,h}) - \mathcal{E}_{\mathsf{t}+r}(\tilde u_{r,h})\right) + \delta\mathcal{E}_{\mathsf{t}+r}\left(u_r;\frac{u_r-u}{r}\right) + \Psi_u\left(\frac{u_r-u}{r}\right) \\ &\leq \mathrm{P}_{\mathsf{t+r}}(u_r,\xi_r) + \left\langle -\xi_r, \frac{u_r-u}{r}\right\rangle + \Psi_u\left(\frac{u_r-u}{r}\right) \\ &= \mathrm{P}_{\mathsf{t+r}}(u_r,\xi_r) - \Psi_u^*(-\xi_r) \quad \text{for all } \ \xi_r \in \mathrm{F}_{\mathsf{t}+r}(u_r) \cap \left(-\partial\Psi_u\left(\frac{u_r-u}{r}\right)\right).\end{aligned}$$

6.2. A priori estimates and compactness for the approximate solutions.

Proposition 6.3 (A priori estimates). *Assume* (4.Ψ_1)–(4.Ψ_2), *and* (4.E_0)–(4.E_4) *for the generalized gradient system* $(V,\mathcal{E},\Psi,\mathrm{F},\mathrm{P})$. *Let* $\overline{U}_\tau$, $\underline{U}_\tau$, U_τ, $\widetilde{U}_\tau$, *and* $\widetilde{\xi}_\tau$ *be the interpolants defined by* (4.10)–(4.12) *and* (4.14). *Then, the* discrete upper energy estimate

$$\begin{aligned}\int_{\bar{\mathsf{t}}_\tau(s)}^{\bar{\mathsf{t}}_\tau(t)} \Psi_{\underline{U}_\tau(r)}(U_\tau'(r))\,\mathrm{d}r + \int_{\bar{\mathsf{t}}_\tau(s)}^{\bar{\mathsf{t}}_\tau(t)} \Psi^*_{\underline{U}_\tau(r)}(-\widetilde{\xi}_\tau(r))\,\mathrm{d}r + \mathcal{E}_{\bar{\mathsf{t}}_\tau(t)}(\overline{U}_\tau(t)) \\ \leq \mathcal{E}_{\bar{\mathsf{t}}_\tau(s)}(\overline{U}_\tau(s)) + \int_{\bar{\mathsf{t}}_\tau(s)}^{\bar{\mathsf{t}}_\tau(t)} \mathrm{P}_r(\widetilde{U}_\tau(r),\widetilde{\xi}_\tau(r))\,\mathrm{d}r\end{aligned} \tag{6.13}$$

holds for every $0 \leq s \leq t \leq T$. *Moreover, there exists a positive constant S such that the following estimates are valid for every* $\tau > 0$*:*

$$\sup_{t\in(0,T)}|\mathcal{E}_t(\overline{U}_\tau(t))| \leq S, \qquad \sup_{t\in(0,T)}|\mathcal{E}_t(\widetilde{U}_\tau(t))| \leq S, \qquad \sup_{t\in(0,T)}\left|\mathrm{P}_t(\widetilde{U}_\tau(t),\widetilde{\xi}_\tau(t))\right| \leq S, \tag{6.14}$$

$$\int_0^T \Psi_{\underline{U}_\tau(s)}(U_\tau'(s))\,\mathrm{d}s \leq S, \qquad \int_0^T \Psi^*_{\underline{U}_\tau(s)}(-\widetilde{\xi}_\tau(s))\,\mathrm{d}s \leq S, \tag{6.15}$$

the families $(U_\tau') \subset L^1(0,T;V)$ *and* $(\widetilde{\xi}_\tau) \subset L^1(0,T;V^*)$ *are uniformly integrable, and* (6.16)

$$\sup_{t\in(0,T)}\|\overline{U}_\tau(t) - \underline{U}_\tau(t)\| + \sup_{t\in(0,T)}\|U_\tau(t) - \overline{U}_\tau(t)\| + \sup_{t\in(0,T)}\|\widetilde{U}_\tau(t) - \underline{U}_\tau(t)\| = o(1) \tag{6.17}$$

as $\tau \downarrow 0$.

Proof: The proof of Proposition 6.3 closely follows the argument for [43, Prop. 4.7]. For the reader's convenience we just outline its main steps here, referring to [43] for the details.

Let t_{n-1}, t_n be two consecutive nodes of the partition $\mathscr{P}_\tau$ and let $t \in (t_{n-1}, t_n]$: applying inequality (6.7) with the choices $\mathsf{t} = t_{n-1}$, $u = U_\tau^{n-1}$, $r_0 = t - t_{n-1}$, $u_{r_0} = \widetilde{U}_\tau(t)$, $u_r = \widetilde{U}_\tau(r)$ and $\xi_r = \widetilde{\xi}_\tau(r)$ for $r \in (t_{n-1}, t)$ (where $\widetilde{U}_\tau$ and $\widetilde{\xi}_\tau$ are defined by (4.12) and (4.14), respectively), we easily obtain

$$
\begin{aligned}
(t-t_{n-1})\Psi_{\underline{U}_\tau(t)}\left(\frac{\widetilde{U}_\tau(t)-\underline{U}_\tau(t)}{t-t_{n-1}}\right)+\int_{t_{n-1}}^{t}\Psi^*_{\underline{U}_\tau(r)}(-\widetilde{\xi}_\tau(r))\,\mathrm{d}r \\
+\mathcal{E}_t(\widetilde{U}_\tau(t))\leq \mathcal{E}_{t_{n-1}}(\overline{U}_\tau(t_{n-1}))+\int_{t_{n-1}}^{t}\mathrm{P}_r(\widetilde{U}_\tau(r),\widetilde{\xi}_\tau(r))\,\mathrm{d}r\,.
\end{aligned}
\tag{6.18}
$$

Writing (6.18) for $t=t_n$ yields

$$
\begin{aligned}
\int_{t_{n-1}}^{t_n}\Psi_{\underline{U}_\tau(r)}\,(U_\tau'(r))\,\mathrm{d}r+\int_{t_{n-1}}^{t_n}\Psi^*_{\underline{U}_\tau(r)}(-\widetilde{\xi}_\tau(r))\,\mathrm{d}r+\mathcal{E}_{t_n}(\overline{U}_\tau(t_n)) \\
\leq \mathcal{E}_{t_{n-1}}(\overline{U}_\tau(t_{n-1}))+\int_{t_{n-1}}^{t_n}\mathrm{P}_r(\widetilde{U}_\tau(r),\widetilde{\xi}_\tau(r))\,\mathrm{d}r\,.
\end{aligned}
\tag{6.19}
$$

Upon summing up on the subintervals of the partition, we obtain (6.13). Now, we estimate the right-hand side of (6.19) via

$$
\begin{aligned}
\mathcal{E}_{t_{n-1}}(\overline{U}_\tau(t_{n-1}))+\int_{t_{n-1}}^{t_n}\mathrm{P}_s(\widetilde{U}_\tau(s),\widetilde{\xi}_\tau(s))\,\mathrm{d}s &\leq \mathcal{E}_{t_{n-1}}(\overline{U}_\tau(t_{n-1}))+C_2\int_{t_{n-1}}^{t_n}\mathcal{G}(\widetilde{U}_\tau(s))\,\mathrm{d}s \\
&\leq \mathcal{E}_{t_{n-1}}(\overline{U}_\tau(t_{n-1}))+C_2C_3\int_{t_{n-1}}^{t_n}\mathcal{G}(\underline{U}_\tau(s))\,\mathrm{d}s\,,
\end{aligned}
$$

the first inequality due to (4.E$_4$) and the second one to (6.4). On the other hand, condition (4.5) yields $\mathcal{E}_{t_n}(\overline{U}_\tau(t_n))\geq C_3^{-1}\mathcal{G}(\overline{U}_\tau(t_n))$. Taking into account the positivity of the two other integral terms on the left-hand side of (6.19) (cf. (2.2)), and summing it up on the intervals of the partition, we obtain the following inequality

$$
\mathcal{G}(\overline{U}_\tau(t_k))\leq C\left(\mathcal{E}_0(u_0)+\int_0^{t_k}\mathcal{G}(\underline{U}_\tau(s))\,\mathrm{d}s+1\right)\,. \tag{6.20}
$$

Then, the first estimate in (6.14) follows from applying to (6.20) a discrete version of the Gronwall lemma (see, e.g., [44, Lemma 4.5]), and the second of (6.14) is a consequence of (6.4). The bound in (6.14) for the sequence $\{\mathrm{P}_t(\widetilde{U}_\tau(t),\widetilde{\xi}_\tau(t))\}$ again follows from the estimate for $\mathcal{E}_t(\widetilde{U}_\tau(t))$, via (4.E$_4$).

Ultimately, the right-hand side in the discrete energy inequality (6.13) is bounded. Thus, we conclude (6.15). From (6.18) we also deduce

$$
\sup_{t\in[0,T]}(t-\underline{\mathrm{t}}_\tau(t))\Psi_{\underline{U}_\tau(t)}\left(\frac{\widetilde{U}_\tau(t)-\underline{U}_\tau(t)}{t-\underline{\mathrm{t}}_\tau(t)}\right)\leq C \tag{6.21}
$$

Now, combining this information with (6.14) and (4.Ψ_2) (cf. (4.1)), we infer that

$\forall\, M>0\ \exists\, S>0\ \forall\,\tau>0,\ t\in[0,T]\ :$

$$
\|\widetilde{U}_\tau(t)-\underline{U}_\tau(t)\|\leq (t-\underline{\mathrm{t}}_\tau(t))S+\frac{1}{M}(t-\underline{\mathrm{t}}_\tau(t))\Psi_{\underline{U}_\tau(t)}\left(\frac{\widetilde{U}_\tau(t)-\underline{U}_\tau(t)}{t-\underline{\mathrm{t}}_\tau(t)}\right).
$$

Estimates (6.15) and, again, the superlinear growth condition (4.Ψ_2), yield the uniform integrability of $(\widetilde{\xi}_\tau)$ and (U_τ'), and the latter in turn implies (6.17). ∎

Hereafter, we will use the short-hand notation

$$
P_\tau(t):=\mathrm{P}_t(\widetilde{U}_\tau(t),\widetilde{\xi}_\tau(t)). \tag{6.22}
$$

The following result subsumes all compactness information on the approximate solutions. Some of the convergences below are stated in terms of a (limit) *Young measure* associated with the family $(U_\tau',\widetilde{\xi}_\tau,P_\tau)_\tau\subset V\times V^*\times\mathbb{R}$, the latter space endowed with the *weak* topology. The definition of Young measure, and related results, are recalled in Appendix A. Without going into details, we may just mention here that the aforementioned limit Young measure allows to express the limit as $\tau\downarrow 0$ of the sequence $(\mathcal{J}(U_\tau',\widetilde{\xi}_\tau,P_\tau))_\tau$ for any *weakly continuous* functional $\mathcal{J}$ on $V\times V^*\times\mathbb{R}$

(and the lim inf as $\tau \downarrow 0$ of the sequence $(\mathcal{H}(U'_\tau, \widetilde{\xi}_\tau, P_\tau))_\tau$ for any *weakly lower semicontinuous* functional $\mathcal{H}$ on $V \times V^* \times \mathbb{R}$).

Proposition 6.4 (Compactness). *Assume* (4.Ψ_1)–(4.Ψ_3)*, and* (4.E_0)–(4.E_6)*. Then, for every vanishing sequence (τ_k) of time-steps there exist a (not relabeled) subsequence, a curve $u \in \mathrm{AC}([0,T];V)$, a function $\mathscr{E} : [0,T] \to \mathbb{R}$ of bounded variation, and a time-dependent Young measure $\boldsymbol{\mu} = \{\mu_t\}_{t\in(0,T)} \in \mathscr{Y}(0,T; V \times V^* \times \mathbb{R})$, such that as $k \uparrow +\infty$*

$$\overline{U}_{\tau_k}, \underline{U}_{\tau_k}, U_{\tau_k}, \widetilde{U}_{\tau_k} \to u \quad \textit{in } L^\infty(0,T;V), \tag{6.23}$$

$$U'_{\tau_k} \rightharpoonup u' \quad \textit{weakly in } L^1(0,T;V), \tag{6.24}$$

$$\begin{cases} \mathcal{E}_t(\overline{U}_\tau(t)) \to \mathscr{E}(t) & \textit{for all } t \in [0,T], \qquad \mathscr{E}(0) = \mathcal{E}_0(u_0), \\ \mathscr{E}(t) \geq \mathcal{E}_t(u(t)) & \textit{for all } t \in [0,T], \\ \mathscr{E}(t) = \mathcal{E}_t(u(t)) & \textit{for a.a. } t \in (0,T), \end{cases} \tag{6.25}$$

and, moreover, $\boldsymbol{\mu}$ is the limit Young measure associated with $(U'_{\tau_k}, \widetilde{\xi}_{\tau_k}, P_{\tau_k})$ in the space $V \times V^* \times \mathbb{R}$ (endowed with the weak topology), which implies

$$u'(t) = \int_{V\times V^*\times\mathbb{R}} v \, \mathrm{d}\mu_t(v,\zeta,p) \quad \textit{for a.a. } t \in (0,T), \tag{6.26a}$$

$$\widetilde{\xi}_{\tau_k} \rightharpoonup \tilde{\xi} \quad \textit{in } L^1(0,T;V^*) \quad \textit{with } \ \tilde{\xi}(t) := \int_{V\times V^*\times\mathbb{R}} \zeta \, \mathrm{d}\mu_t(v,\zeta,p) \quad \textit{for a.a. } t \in (0,T), \tag{6.26b}$$

$$\begin{aligned} &P_{\tau_k} \rightharpoonup^* P \quad \textit{in } L^\infty(0,T) \textit{ with} \\ &P(t) := \int_{V\times V^*\times\mathbb{R}} p \, \mathrm{d}\mu_t(v,\zeta,p) \leq \int_{V\times V^*\times\mathbb{R}} \mathrm{P}_t(u(t),\zeta) \, \mathrm{d}\mu_t(v,\zeta,p) \quad \textit{for a.a. } t \in (0,T). \end{aligned} \tag{6.26c}$$

Finally, the following energy inequality holds for all $0 \leq s \leq t \leq T$:

$$\begin{aligned} &\int_s^t \int_{V\times V^*\times\mathbb{R}} \Big(\Psi_{u(r)}(v) + \Psi^*_{u(r)}(-\zeta) \Big) \, \mathrm{d}\mu_r(v,\zeta,p) \, \mathrm{d}r + \mathscr{E}(t) \\ &\leq \mathscr{E}(s) + \int_s^t P(r) \, \mathrm{d}r \leq \mathscr{E}(s) + \int_s^t \int_{V\times V^*\times\mathbb{R}} \mathrm{P}_r(u(r),\zeta) \, \mathrm{d}\mu_r(v,\zeta,p) \, \mathrm{d}r \, . \end{aligned} \tag{6.27}$$

Proof: Let (τ_k) be a vanishing sequence of time-steps. It follows from the uniform integrability (6.16) of the sequence (U'_{τ_k}) that (U_{τ_k}) is equicontinuous on V. Furthermore, (6.14) and assumption (4.E_1) give that $\widetilde{U}_{\tau_k}$ is contained in some compact subset of V. Hence U_{τ_k} is contained in its convex hull, which is also compact. Therefore, with the Arzelà-Ascoli theorem we conclude that there exists $u \in \mathrm{C}^0([0,T];V)$ such that, up to a subsequence,

$$U_{\tau_k} \to u \quad \text{in } \mathrm{C}^0([0,T];V). \tag{6.28}$$

Combining this with (6.17), we conclude convergences (6.23). Next, (6.24) ensues from the aforementioned uniform integrability of (U'_{τ_k}) via the Dunford-Pettis criterion (see, e.g., [21, Cor. IV.8.11]).

Secondly, from the third of (6.14) we have that, up to a further subsequence,

$$P_{\tau_k} \text{ converges weakly}^* \text{ in } L^\infty(0,T) \text{ to some } P \in L^\infty(0,T). \tag{6.29}$$

Thus, to prove (6.25) we proceed in the same way as for [44, Prop. 4.7], viz. we deduce from the discrete energy inequality (6.13) that the map

$$t \mapsto \eta_\tau(t) := \mathcal{E}_{\bar{\mathrm{t}}_\tau(t)}(\overline{U}_\tau(t)) - \int_0^{\bar{\mathrm{t}}_\tau(t)} P_\tau(r) \, \mathrm{d}r \qquad \text{is nonincreasing on } [0,T].$$

Therefore by Helly's theorem there exists $\eta : [0,T] \to \mathbb{R}$, nonincreasing, such that, up to a subsequence, $\eta_{\tau_k}(t) \to \eta(t)$ for all $t \in \mathbb{R}$. In view of (6.29), we conclude that

$$\mathcal{E}_{\bar{\mathrm{t}}_{\tau_k}(t)}(\overline{U}_{\tau_k}(t)) \to \mathscr{E}(t) := \eta(t) + \int_0^t P(r) \, \mathrm{d}r \qquad \forall\, t \in [0,T]. \tag{6.30}$$

This ultimately yields the first of (6.25) via (4.E$_3$) and (6.14), which give

$$\left|\mathcal{E}_{\bar{\mathfrak{t}}_\tau(t)}(\overline{U}_\tau(t)){-}\mathcal{E}_t(\overline{U}_\tau(t))\right| \le C_1|\bar{\mathfrak{t}}_\tau(t){-}t|\,\mathcal{G}(\overline{U}_\tau(t)) \le SC_1\,|\bar{\mathfrak{t}}_\tau(t){-}t| \to 0 \quad \text{as } \tau\downarrow 0. \tag{6.31}$$

Then, the second of (6.25) is a straightforward consequence of the lower semicontinuity of $\mathcal{E}_t(\cdot)$, while the third of (6.25) follows from assumption (4.E$_6$).

In view of estimates (6.14) and (6.15) (which imply the uniform integrability of the sequence $\{\widetilde{\xi}_\tau\}$ in $L^1(0,T;V^*)$ as well), we are in the position of applying the Young measure result in Theorem A.2 to the sequence $(U'_{\tau_k},\widetilde{\xi}_{\tau_k},P_{\tau_k})$, with which we associate a limit Young measure $\boldsymbol{\mu}=\{\mu_t\}_{t\in(0,T)}$ such that for a.a. $t\in(0,T)$

$$\begin{gathered}\mu_t \ \text{ is concentrated on the set } L(t) \text{ of the limit points of } (U'_{\tau_k}(t),\widetilde{\xi}_{\tau_k}(t),P_{\tau_k}(t))\\ \text{with respect to the weak-weak-strong topology of } V\times V^*\times\mathbb{R},\end{gathered} \tag{6.32}$$

(cf. (A.4)), and there hold (A.5) and (A.6). Note that the latter relations imply (6.26a), (6.26b), and (6.26c). Then, from Jensen's inequality we have

$$\left\{\begin{array}{l}\int_{V\times V^*\times\mathbb{R}}\Psi_{u(t)}(v)\,\mathrm{d}\mu_t(v,\zeta,p)\ge\Psi_{u(t)}(u'(t)),\\ \int_{V\times V^*\times\mathbb{R}}\Psi^*_{u(t)}(-\zeta)\,\mathrm{d}\mu_t(v,\zeta,p)\ge\Psi^*_{u(t)}(-\tilde{\xi}(t))\end{array}\right. \qquad \text{for a.a. } t\in(0,T)\,. \tag{6.33}$$

Passing to the limit in the Euler equation (4.14), we deduce from (4.E$_6$), from convergence (6.23) for $(\widetilde{U}_{\tau_k})$, and from the first of (6.25), that for a.a. $t\in(0,T)$ the set $L(t)$ has the following property

$$\text{for all } (v,\zeta,p)\in L(t) \text{ there holds } \ \zeta\in\mathrm{F}_t(u(t)),\quad p\le\mathrm{P}_t(u(t),\zeta). \tag{6.34}$$

Hence, from the latter inequality and (A.6) we also deduce the inequality in (6.26c). Furthermore, we apply the Γ-lim inf inequality (A.5) with the choice $\mathcal{H}_k(t,v,\zeta,p)=\Psi_{\underline{U}_{\tau_k}(t)}(v)$ (notice that (A.3) is fulfilled in view of assumption (4.Ψ_3), of (6.14), and of (6.23)). Thus, we obtain for all $0\le s\le t\le T$

$$\liminf_{k\to\infty}\int_{\bar{\mathfrak{t}}_{\tau_k}(s)}^{\bar{\mathfrak{t}}_{\tau_k}(t)}\Psi_{\underline{U}_{\tau_k}(r)}\left(U'_{\tau_k}(r)\right)\mathrm{d}r\ge\int_s^t\int_{V\times V^*\times\mathbb{R}}\Psi_{u(r)}(v)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r \tag{6.35}$$

The choice $\mathcal{H}_k(t,v,\zeta,p)=\Psi^*_{\underline{U}_\tau(t)}(\zeta)$ (which complies with (A.3) thanks to (4.2) and again (6.14)) obviously gives

$$\liminf_{k\to\infty}\int_{\bar{\mathfrak{t}}_{\tau_k}(s)}^{\bar{\mathfrak{t}}_{\tau_k}(t)}\Psi^*_{\underline{U}_{\tau_k}(r)}(-\widetilde{\xi}_{\tau_k}(r))\,\mathrm{d}r\ge\int_s^t\int_{V\times V^*\times\mathbb{R}}\Psi^*_{u(r)}(-\zeta)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r. \tag{6.36}$$

Therefore, we pass to the limit in the discrete energy inequality (6.13). Using (6.25), (6.26c), (6.29), (6.35), and (6.36), we conclude inequality (6.27). This completes the proof. ■

6.3. **Proof of Theorem 4.4.**

Step 1: a Young measure argument. It follows from the a priori estimates (6.14)–(6.15) and from (6.25), (6.32)–(6.35) that the curve $u\in\mathrm{AC}([0,T;V)$ and the Young measure $\{\mu_t\}_{t\in(0,T)}$ comply with assumptions (B.1)–(B.4) of Theorem B.1. Therefore, the map $t\mapsto\mathcal{E}_t(u(t))$ is absolutely continuous and we have the following chain of inequalities

$$\begin{aligned}&\int_0^t\int_{V\times V^*\times\mathbb{R}}\Big(\Psi_{u(r)}(v){+}\Psi^*_{u(r)}(-\zeta)\Big)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r+\mathcal{E}_t(u(t))\\ &\quad\le\mathcal{E}_0(u(0))+\int_0^t P(r)\,\mathrm{d}r\le\mathcal{E}_t(u(t))+\int_0^t\int_{V\times V^*\times\mathbb{R}}\langle-\zeta,u'(r)\rangle\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r\end{aligned} \tag{6.37}$$

where the first inequality follows from (6.27) (written for $t\in(0,T]$ and $s=0$) and from the second of (6.25), while the second inequality is a consequence of the *Young measure* chain-rule inequality (B.5). Taking into account inequality (6.33), we thus conclude

$$\int_0^t\int_{V\times V^*\times\mathbb{R}}\Big(\Psi_{u(r)}(u'(r)){+}\Psi^*_{u(r)}(-\zeta){-}\langle-\zeta,u'(r)\rangle\Big)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r\le 0.$$

Since the integrand is nonnegative, we find

$$\int_{V\times V^*\times\mathbb{R}} \Big(\Psi_{u(t)}(u'(t))+\Psi^*_{u(t)}(-\zeta)-\langle-\zeta,u'(t)\rangle\Big)\,\mathrm{d}\mu_t(v,\zeta,p)=0 \quad \text{for a.a.}\, t\in(0,T). \tag{6.38}$$

Now, it follows from the above discussion that all inequalities in (6.37) indeed hold as equalities. Again using the chain-rule inequality (B.5), it is easy to deduce that, for almost all $t \in (0,T)$, we have

$$\begin{aligned}\int_{V\times V^*\times\mathbb{R}} &\Big(\Psi_{u(t)}(u'(t))+\Psi^*_{u(t)}(-\zeta)-p\Big)\,\mathrm{d}\mu_t(v,\zeta,p)\\ &= \int_{V\times V^*\times\mathbb{R}} (\langle-\zeta,u'(t)\rangle-p)\,\mathrm{d}\mu_t(v,\zeta,p) = -\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t))\end{aligned} \tag{6.39}$$

Thus, we conclude for every $0\le s\le t\le T$ the energy identity

$$\int_s^t\int_{V\times V^*\times\mathbb{R}} \Big(\Psi_{u(r)}(v)+\Psi^*_{u(r)}(-\zeta)-p\Big)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r = \mathcal{E}_s(u(s))-\mathcal{E}_t(u(t))\,. \tag{6.40}$$

Step 2: a measurable selection. Let us consider the measure $\nu_t := (\pi_{2,3})_\#(\mu_t)$, i.e. the marginal of μ_t with respect to the (ζ,p)-component, defined by $(\pi_{2,3})_\#(\mu_t)(B)=\mu_t(\pi_{2,3}^{-1}(B))$ for all $B\in\mathscr{B}(V^*\times\mathbb{R})$ (the Borel σ-algebra of $V^*\times\mathbb{R}$). As a consequence of (6.34) and (6.38), for almost all $t\in(0,T)$ the measure ν_t is concentrated on the set

$$\mathcal{S}(t,u(t),u'(t)) := \big\{(\zeta,p)\in V^*\times\mathbb{R}\,:\,\zeta\in \mathrm{F}_t(u(t))\cap(-\partial\Psi_{u(t)}(u'(t)),\ p\le \mathrm{P}_t(u(t),\zeta)\big\}\,, \tag{6.41}$$

namely

$$\nu_t((V^*\times\mathbb{R})\setminus\mathcal{S}(t,u(t),u'(t)))=0 \quad \text{for a.a.}\, t\in(0,T). \tag{6.42}$$

In particular, for almost all $t\in(0,T)$ the set $\mathcal{S}(t,u(t),u'(t))$ is nonempty. Then, Lemma B.2 in the appendix below guarantees that there exists a measurable selection $t\in(0,T)\mapsto(\xi(t),p(t))\in\mathcal{S}(t,u(t),u'(t))$ such that

$$\Psi^*_{u(t)}(-\xi(t))-p(t) = \min_{(\zeta,p)\in\mathcal{S}(t,u(t),u'(t))}\{\Psi^*_{u(t)}(-\zeta)-p\} \doteq \mathcal{M}^*(t) \quad \text{for a.a.}\, t\in(0,T). \tag{6.43}$$

In particular, ξ satisfies equation (4.16), hence we conclude that u solves the Cauchy problem for (1.21). In fact, we have

$$\int_0^T \Psi^*_{u(t)}(-\xi(t))\,\mathrm{d}t<+\infty, \tag{6.44}$$

which in particular yields $\xi\in L^1(0,T;V^*)$ via (4.Ψ_2). To check (6.44), it is sufficient to observe that

$$\begin{aligned}\int_0^T \Psi^*_{u(t)}(-\xi(t))\,\mathrm{d}t &\le \int_0^T \mathcal{M}^*(t)+p(t)\,\mathrm{d}t\\ &\le \int_0^T\int_{V^*\times\mathbb{R}}\Big(\Psi^*_{u(r)}(-\zeta)-p\Big)\,\mathrm{d}\nu_r(\zeta,p)\,\mathrm{d}r+\int_0^T \mathrm{P}_r(u(r),\xi(r))\,\mathrm{d}r\\ &\le \int_0^T\int_{V\times V^*\times\mathbb{R}}\Big(\Psi^*_{u(r)}(-\zeta)-p\Big)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r+C_2\int_0^T \mathcal{G}(u(r))\,\mathrm{d}r\le C\end{aligned}$$

where the second inequality ensues from (6.42) and the fact that $p(t)\le \mathrm{P}_t(u(t),\xi(t))$ for almost all $t\in(0,T)$, the third inequality from (4.E_4), and the last one from (6.40) and the fact that $\sup_{t\in(0,T)}\mathcal{G}(u(t))<+\infty$.

Step 3: proof of the energy identity (4.17)**.** On the one hand, we observe that for every $0\le s\le t\le T$ there holds

$$\begin{aligned}\int_s^t &\Big(\Psi_{u(r)}(u'(r))+\Psi^*_{u(r)}(-\xi(r))-\mathrm{P}_r(u(r),\xi(r))\Big)\,\mathrm{d}r\\ &\le \int_s^t \Psi_{u(r)}(u'(r))+\Psi^*_{u(r)}(-\xi(r))-p(r)\,\mathrm{d}r \le \mathcal{E}_s(u(s))-\mathcal{E}_t(u(t)),\end{aligned} \tag{6.45}$$

where the first estimate follows from $p(t) \le \mathrm{P}_t(u(t),\xi(t))$ for almost all $t \in (0,T)$ by definition of $\mathcal{S}(t,u(t),u'(t))$, and the second estimate is due to (6.40), combined with (6.42) and (6.43). On the other hand, applying the chain-rule inequality (4.E$_5$) to the pair (u,ξ) we have

$$\mathcal{E}_t(u(t)) - \mathcal{E}_s(u(s)) \ge \int_s^t (\langle \xi(r), u'(r)\rangle + \mathrm{P}_r(u(r),\xi(r)))\,\mathrm{d}r \quad \text{for every } 0 \le s \le t \le T. \tag{6.46}$$

Combining (6.45) and (6.46) and arguing in the same way as throughout (6.37)–(6.38), we obtain that all inequalities in (6.45) ultimately hold as equalities; in particular, $p(t) = \mathrm{P}_t(u(t),\xi(t))$ for almost all $t \in (0,T)$. We have thus proved that the pair (u,ξ) satisfies the energy identity (4.17). A comparison between the latter and the *Young-measure energy identity* (6.40) also reveals that, for almost all $t \in (0,T)$,

$$\Psi_{u(t)}(u'(t)) = \int_{V\times V^*\times\mathbb{R}} \Psi_{u(t)}(v)\,\mathrm{d}\mu_t(v,\zeta,p), \tag{6.47}$$

$$\begin{aligned}\Psi^*_{u(t)}(-\xi(t)) - \mathrm{P}_t(u(t),\xi(t)) &= \min_{(\zeta,p)\in\mathcal{S}(t,u(t),u'(t))} \{\Psi^*_{u(t)}(-\zeta) - p\} \\ &= \Psi^*_{u(t)}(-\zeta) - p \quad \text{for } \nu_t\text{-almost all } (\zeta,p) \in V^*\times\mathbb{R}.\end{aligned} \tag{6.48}$$

Taking into account that $\Psi^*_{u(t)}(-\zeta) = \Psi^*_{u(t)}(-\xi(t))$ due to condition (2.Ψ_3), we thus conclude the *maximum selection principle*

$$\mathrm{P}_t(u(t),\xi(t)) = \max\{p \,:\, (\zeta,p) \in \mathcal{S}(t,u(t),u'(t))\}. \tag{6.49}$$

Step 4: enhanced convergences. Convergences (4.18c)–(4.18d) and (4.18e) are proved by passing to the limit in (6.13), written for $s=0$ and $t\in[0,T]$. We use the short-hand notation (6.22), as well as

$$A_k(t) = \int_0^{\bar{\mathrm{t}}_{\tau_k}(t)} \Psi_{\underline{U}_{\tau_k}(r)}\left(U'_{\tau_k}(r)\right)\mathrm{d}r,\ \ B_k(t) = \int_0^{\bar{\mathrm{t}}_{\tau_k}(t)} \Psi^*_{\underline{U}_{\tau_k}(r)}(-\widetilde{\xi}_{\tau_k}(r))\,\mathrm{d}r,\ \ C_k(t) = \mathcal{E}_{\bar{\mathrm{t}}_{\tau_k}(t)}(\overline{U}_{\tau_k}(t)).$$

For all $t \in [0,T]$ we find

$$\begin{aligned}&\textstyle\int_0^t \Psi_{u(r)}(u'(r))\,\mathrm{d}r + \int_0^t \Psi^*_{u(r)}(-\xi(r))\,\mathrm{d}r \ + \mathcal{E}_t(u(t))\\ &= \textstyle\int_0^t \Psi_{u(r)}(u'(r))\,\mathrm{d}r + \int_0^t \int_{V\times V^*\times\mathbb{R}} \Psi^*_{u(r)}(-\zeta)\,d\nu_r(v,\zeta,p)\,\mathrm{d}r \ + \mathcal{E}_t(u(t))\\ &\le \liminf_{k\to\infty} A_k(t) + \liminf_{k\to\infty} B_k(t) + \liminf_{k\to\infty} C_k(t)\\ &\le \limsup_{k\to\infty} (A_k(t)+B_k(t)+C_k(t))\\ &\le \limsup_{k\to\infty} \mathcal{E}_0(\overline{U}_{\tau_k}(0)) + \limsup_{k\to\infty} \textstyle\int_0^{\bar{\mathrm{t}}_{\tau_k}(t)} P_\tau(r)\,\mathrm{d}r\\ &\le \mathcal{E}_0(u(0)) + \textstyle\int_0^t \mathrm{P}_r(u(r),\xi(r))\,\mathrm{d}r\\ &= \textstyle\int_0^t \Psi_{u(r)}(u'(r))\,\mathrm{d}r + \int_0^t \Psi^*_{u(r)}(-\xi(r))\,\mathrm{d}r \ + \mathcal{E}_t(u(t))\end{aligned} \tag{6.50}$$

where the first identity follows from (6.48), the second estimate from (6.25), (6.30)–(6.31), and (6.35)–(6.36), the third estimate is trivial and the fourth one ensues from inequality (6.13), whereas the fifth estimate is a consequence of (6.26c), and the sixth identity is due to (4.17). Altogether, all inequalities in (6.50) turn out to be equalities, and with an elementary argument we conclude (4.18c), (4.18d), as well as (4.18e).

Step 5: the strictly convex case. Finally, if we further assume (4.19), from (6.47), (6.48), and the strict convexity of $\Psi^*_{u(t)}(\cdot)$ we infer

$$(\pi_2)_\#(\mu_t) = \delta_{\xi(t)} \qquad \text{for a.a. } t \in (0,T). \tag{6.51}$$

Hence, from (6.26b) we deduce convergence (4.20). This concludes the proof of Theorem 4.4. ■

Remark 6.5. Notice that, if in addition we assume Ψ_u to be strictly convex for all u, then we also have $(\pi_1)_\#(\mu_t) = \delta_{u'(t)}$. The latter relation, joint with (6.51), yields

$$U'_{\tau_k}(t) \rightharpoonup u'(t), \qquad \widetilde{\xi}_{\tau_k}(t) \rightharpoonup \xi(t) \qquad \text{for a.a. } t \in (0,T).$$

Sketch of the proof of Theorem 4.8. For every $n \in \mathbb{N}$, the solution pair (u_n, ξ_n) fulfills the energy identity associated with the Cauchy problem (4.28), namely there holds

$$\int_0^t \left(\Psi^n_{u_n(r)}(u_n'(r)) + (\Psi^n_{u_n(r)})^*(-\xi_n(r))\right) \mathrm{d}r + \mathcal{E}^n_t(u_n(t)) = \mathcal{E}^n_0(u^n_0) + \int_0^t \mathrm{P}^n_r(u_n(r), \xi_n(r))\,\mathrm{d}r \quad (6.52)$$

for all $t \in [0,T]$. From (6.52) we deduce all the a priori estimates on the sequence (u_n, ξ_n), with the very same arguments as in the proof of Proposition 6.3. Indeed, we exploit condition (4.27) on $\mathcal{E}^n_0(u^n_0)$, and use (4.E$_4$) (for a constant uniform with respect to $n \in \mathbb{N}$), to estimate the terms on the right-hand side of (6.52). Then, all of the terms on the left-hand side are estimated as well. Combining this with the coercivity properties of the potentials (Ψ^n_u), viz.

$\forall\, R > 0,\ M > 0$

$$\begin{cases} \exists K > 0 \quad \forall\, u \in D \text{ with } \sup_{n\in\mathbb{N}} \mathcal{G}^n(u) \le R \quad \forall\, v \in V \quad : \quad \|v\| \ge K \quad \Rightarrow \quad \Psi^n_u(v) \ge M\|v\|, \\ \exists K^* > 0 \quad \forall\, u \in D \text{ with } \sup_{n\in\mathbb{N}} \mathcal{G}^n(u) \le R \quad \forall\, \xi \in V^* \quad : \quad \|\xi\|_* \ge K^* \quad \Rightarrow \quad (\Psi^n_u)^*(\xi) \ge M\|\xi\|_*, \end{cases}$$

we have that the sequence $(u_n') \subset L^1(0,T;V)$ is uniformly integrable. Furthermore, the estimate $\sup_{n\in\mathbb{N}} \sup_{t\in[0,T]} \mathcal{E}^n_t(u_n(t))$ yields compactness which, combined with uniform integrability, ensures convergences (4.29a), along a subsequence, to some curve $u \in \mathrm{AC}([0,T];V)$. Like in Proposition 6.4, up to a subsequence we also find some limit Young measure for the sequence (u_n', ξ_n, P_n), with $P_n(t) := \mathrm{P}^n_t(u_n(t), \xi_n(t))$.

Finally, in order to pass to the limit as $n \to \infty$, we reproduce on the time-continuous level the arguments developed in Steps 1–4 of the proof of Theorem 4.4. Namely, combining semicontinuity arguments with properties (4.25) and (4.26), we take the limit as $n \to \infty$ of (6.52), and deduce that the curve u fulfills the upper energy estimate. We obtain the lower energy estimate from the chain rule, and in this way we conclude that u is a solution to the Cauchy problem for (1.21). ∎

Appendix A. Young measure tools

In this section, we collect some results on parametrized (or Young) measures with values in infinite-dimensional spaces, see e.g. [7, 8, 11, 9, 13, 54]. In particular, we shall focus on Young measures with values in a reflexive Banach space $\mathcal{V}$. The definitions and results we are going to recall below, apply in Section 6.2 (cf. Proposition 6.4), to the space $\mathcal{V} = V \times V^* \times \mathbb{R}$.

Notation. Given an interval $I \subset \mathbb{R}$, we denote by $\mathscr{L}_I$ the σ-algebra of the Lebesgue measurable subsets of I and, given a reflexive Banach space $\mathcal{V}$, by $\mathscr{B}(\mathcal{V})$ its Borel σ-algebra. We use the symbol $\otimes$ for product σ-algebrae. We recall that a $\mathscr{L}_I \otimes \mathscr{B}(\mathcal{V})$-measurable function $h : I \times \mathcal{V} \to (-\infty, +\infty]$ is a *normal integrand* if for a.a. $t \in (0,T)$ the map $x \mapsto h_t(x) = h(t,x)$ is lower semicontinuous on $\mathcal{V}$.

We consider the space $\mathcal{V}$ endowed with the *weak* topology, and say that a $\mathscr{L}_{(0,T)} \otimes \mathscr{B}(\mathcal{V})$–measurable functional $\mathcal{H} : (0,T) \times \mathcal{V} \to (-\infty, +\infty]$ is a *weakly-normal integrand* if for a.a. $t \in (0,T)$ the map

$$w \mapsto h(t,w) \text{ is sequentially lower semicontinuous on } \mathcal{V} \text{ w.r.t. the weak topology.} \quad (A.1)$$

We denote by $\mathscr{M}(0,T;\mathcal{V})$ the set of all $\mathscr{L}_{(0,T)}$-measurable functions $y : (0,T) \to \mathcal{V}$. A sequence $(w_n) \subset \mathscr{M}(0,T;\mathcal{V})$ is said to be *weakly-tight* if there exists a weakly-normal integrand $\mathcal{H} : (0,T) \times \mathcal{V} \to (-\infty, +\infty]$ such that the map

$$w \mapsto \mathcal{H}_t(w) \text{ has compact sublevels w.r.t. the weak topology of } \mathcal{V}, \text{ and}$$

$$\sup_n \int_0^T \mathcal{H}(t, w_n(t))\,\mathrm{d}t < \infty.$$

Definition A.1 ((Time-dependent) Young measures). A *Young measure* in the space $\mathcal{V}$ is a family $\boldsymbol{\mu} := \{\mu_t\}_{t\in(0,T)}$ of Borel probability measures on $\mathcal{V}$ such that the map on $(0,T)$

$$t \mapsto \mu_t(B) \quad \text{is} \quad \mathscr{L}_{(0,T)}\text{-measurable} \quad \text{for all } B \in \mathscr{B}(\mathcal{V}). \quad (A.2)$$

We denote by $\mathscr{Y}(0,T;\mathcal{V})$ the set of all Young measures in $\mathcal{V}$.

The following Γ-$\liminf$ result is a straightforward consequence of [53, Thm. 4.2].

Theorem A.2. *Let* $\{\mathcal{H}_n\}$, $\mathcal{H} : (0,T) \times \mathcal{V} \to (-\infty, +\infty]$ *be weakly-normal integrands such that for all* $w \in \mathcal{V}$ *and for a.a.* $t \in (0,T)$

$$\mathcal{H}(t,w) \leq \inf\left\{ \liminf_{n\to\infty} \mathcal{H}_n(t,w_n) \, : \, w_n \rightharpoonup w \text{ in } \mathcal{V} \right\}. \tag{A.3}$$

Let $(w_n) \subset \mathscr{M}(0,T;\mathcal{V})$ *be a weakly-tight sequence. Then, there exist a subsequence* (w_{n_k}) *and a Young measure* $\boldsymbol{\mu} = \{\mu_t\}_{t\in(0,T)}$ *such that for a.a.* $t \in (0,T)$

$$\mu_t \text{ is concentrated on the set } L(t) := \bigcap\nolimits_{p=1}^{\infty} \overline{\{w_{n_k}(t) \, : \, k \geq p\}}^{\mathrm{w}} \tag{A.4}$$

of the limit points of the sequence $(w_{n_k}(t))$ *with respect to the weak topology of* $\mathcal{V}$ *and, if the sequence* $t \mapsto \mathcal{H}^-_{n_k}(t, w_{n_k}(t))$ *is uniformly integrable, there holds*

$$\liminf_{k\to+\infty} \int_0^T \mathcal{H}_{n_k}(t, w_{n_k}(t)) \,\mathrm{d}t \geq \int_0^T \int_{\mathcal{V}} \mathcal{H}(t,w) \,\mathrm{d}\mu_t(w) \,\mathrm{d}t \,. \tag{A.5}$$

As a corollary (the reader is referred to the discussion in [53] for more details), we have a generalization of the so-called *Fundamental Theorem of Young measures*, see [44, Thm. 3.2] for the case of the weak topology in Hilbert spaces, and the classical results [7, Thm. 1], [8, Thm. 2.2], [9, Thm. 4.2], [54, Thm. 16].

Theorem A.3 (**The Fundamental Theorem for strong-weak-weak topologies**). *Let* $1 \leq p \leq \infty$ *and let* $(w_n) \subset L^p(0,T;\mathcal{V})$ *be a bounded sequence. If* $p = 1$, *suppose further that* (w_n) *is uniformly integrable in* $L^1(0,T;\mathcal{V})$. *Then, there exists a subsequence* (w_{n_k}) *and a Young measure* $\boldsymbol{\mu} = \{\mu_t\}_{t\in(0,T)} \in \mathscr{Y}(0,T;\mathcal{V})$ *such that for a.a.* $t \in (0,T)$ *relation* (A.4) *holds and, setting*

$$\mathrm{w}(t) := \int_{\mathcal{V}} w \,\mathrm{d}\mu_t(w) \qquad \text{for a.a. } t \in (0,T)\,,$$

there holds

$$w_{n_k} \rightharpoonup \mathrm{w} \quad \text{in } L^p(0,T;V), \tag{A.6}$$

with $\rightharpoonup$ *replaced by* $\rightharpoonup^*$ *if* $p = \infty$.

In fact, in Section 6.2 (cf. Proposition 6.4), Theorem A.3 applies to the sequence $w_k(t) := (U'_{\tau_k}(t), \widetilde{\xi}_{\tau_k}(t), P_{\tau_k}(t)) \subset V \times V^* \times \mathbb{R}$.

Appendix B. Extension of the chain rule to Young measures

From now on, we work with Young measures valued in the reflexive space $\mathcal{V} := V \times V^* \times \mathbb{R}$, whose elements are denoted by (v, ζ, p). The main result of this section is a *Young measure version* of the chain-rule inequality (4.E$_5$).

Proposition B.1. *In the framework of* (4.Ψ_1)–(4.Ψ_2), *suppose that* $\mathcal{E}$ *complies with* (4.E$_0$)–(4.E$_6$). *Let* $u \in \mathrm{AC}([0,T];V)$ *be an absolutely continuous curve such that*

$$(t, u(t)) \in \mathrm{dom}(\mathrm{F}) \quad \text{for a.a. } t \in (0,T), \quad \text{and} \quad \sup_{t\in(0,T)} \mathcal{E}_t(u(t)) < +\infty. \tag{B.1}$$

Let $\boldsymbol{\mu} = \{\mu_t\}_{t\in(0,T)}$ *be a Young measure in* $V \times V^* \times \mathbb{R}$ *such that*

$$\int_0^T \int_{V\times V^*\times\mathbb{R}} \Big(\Psi_{u(t)}(v) + \Psi^*_{u(t)}(-\zeta) \Big) \,\mathrm{d}\mu_t(v,\zeta,p) \,\mathrm{d}t < +\infty, \tag{B.2}$$

$$\text{for a.a. } t \in (0,T) \quad u'(t) = \int_{V\times V^*\times\mathbb{R}} v \,\mathrm{d}\mu_t(v,\zeta,p), \tag{B.3}$$

$$\text{for a.a. } t \in (0,T) \quad \text{for all } (v,\xi,p) \in \mathrm{supp}(\mu_t) \text{ there holds } \ \xi \in \mathrm{F}_t(u(t)), \ p \leq \mathrm{P}_t(u(t),\xi). \tag{B.4}$$

Then,

$$\begin{aligned} &\text{the map } \ t \mapsto \mathcal{E}_t(u(t)) \ \text{ is absolutely continuous on } (0,T), \text{ and} \\ &\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}_t(u(t)) \geq \int_{V\times V^*\times\mathbb{R}} (\langle u'(t), \zeta \rangle + p) \,\mathrm{d}\mu_t(v,\zeta,p) \quad \text{for a.a. } t \in (0,T). \end{aligned} \tag{B.5}$$

The proof of this result closely follows the argument for [44, Thm. 3.3], to which we constantly refer the reader.

Proof: We split the argument in three claims.

Claim 1: *let us set for almost all $t \in (0,T)$*

$$\mathcal{K}(t,u(t)) := \{(\xi,p) \in V^* \times \mathbb{R} \,:\, \xi \in \mathrm{F}_t(u(t)),\ p \le \mathrm{P}_t(u(t),\xi)\}\,. \tag{B.6}$$

There exists a sequence of strongly measurable maps $(\xi_n,p_n) : (0,T) \to V^* \times \mathbb{R}$ such that

$$\{(\xi_n(t),p_n(t)) \,:\, n \in \mathbb{N}\} \subset \mathcal{K}(t,u(t)) \subset \overline{\{(\xi_n(t),p_n(t)) \,:\, n \in \mathbb{N}\}}, \tag{B.7}$$

($\overline{A}$ denoting the closure, with respect to the topology of $V^* \times \mathbb{R}$, of a subset $A \subset V^* \times \mathbb{R}$).

First of all, let us observe that the set

$$\begin{aligned} K := \{(t,u,\xi,p) \in [0,T] \times V \times V^* \times \mathbb{R} \,:\, \xi \in \mathrm{F}_t(u),\ p \le \mathrm{P}_t(u,\xi)\} \\ \text{is a Borel set of } [0,T] \times V \times V^* \times \mathbb{R}. \end{aligned} \tag{B.8}$$

This follows from the fact that graph(F) is a Borel set of $[0,T] \times V \times V^*$, and $\mathrm{P} : \mathrm{graph}(\mathrm{F}) \to \mathbb{R}$ a Borel function. Now, it follows from (B.1) and (B.3)–(B.4) that there exists a subset $\mathcal{T} \subset (0,T)$ of full measure such that $\mathcal{K}(t,u(t)) \neq \emptyset$ for all $t \in [0,T]$. Let us then consider the graph of the multivalued function $t \in \mathcal{T} \mapsto \mathcal{K}(t,u(t)) \subset V^* \times \mathbb{R}$, i.e. the set

$$\mathscr{K} = \{(t,\xi,p) \in \mathcal{T} \times V^* \times \mathbb{R} \,:\, (\xi,p) \in \mathcal{K}(t,u(t))\} = \{(t,\xi,p) \in \mathcal{T} \times V^* \times \mathbb{R} \,:\, (t,u(t),\xi,p) \in K\}\,.$$

Due to the latter representation, to (B.8) and to the fact that the function $u : [0,T] \to V$ is Borelian, we can conclude that $\mathscr{K}$ is a Borel set of $\mathcal{T} \times V^* \times \mathbb{R}$. Therefore, (B.7) ensues from [14, Thm. III.22].

Claim 2: *it is possible to choose the measurable maps $\xi_n : (0,T) \to V^*$ fulfilling* (B.7) *in such a way that*

$$\xi_n \in L^1(0,T;V^*) \ \text{ for every } n \in \mathbb{N} \text{ and } \ \sup_n \int_0^T \Psi^*_{u(t)}(-\xi_n(t))\,\mathrm{d}t < +\infty. \tag{B.9}$$

We set

$$\mathcal{M}_*(t) := \min_{(\xi,p)\in\mathcal{K}(t,u(t))} \Psi^*_{u(t)}(-\xi) \quad \text{for almost all } t \in (0,T).$$

It follows from (B.7) that

$$\text{the map } \ t \mapsto \mathcal{M}_*(t) = \min_n \Psi^*_{u(t)}(-\xi_n(t)) \ \text{ is measurable on } (0,T). \tag{B.10}$$

Moreover, (B.2) and (B.3) yield that

$$\int_0^T \mathcal{M}_*(t)\,\mathrm{d}t \le \int_0^T \int_{V^*} \Psi^*_{u(t)}(-\zeta)\,\mathrm{d}\mu_t(v,\zeta,p)\,\mathrm{d}t < +\infty. \tag{B.11}$$

Arguing as the proof of [44, Lemma 3.4], we recursively define the following family of subsets of $\mathcal{T}$, i.e.

$$A_0 := \emptyset, \quad A_k := \Big\{t \in \mathcal{T} \,:\, \Psi^*_{u(t)}(-\xi_k(t)) \le \mathcal{M}_*(t) + 1\Big\} \setminus \bigcup_{j=0}^{k-1} A_j.$$

Due to (B.10), for every $k \in \mathbb{N}$ the set A_k is measurable and, by construction, the family $\{A_k\}_{k\in\mathbb{N}}$ is disjoint with $\bigcup_{k\in\mathbb{N}} A_k = \mathcal{T}$. Hence, we set

$$\bar\xi(t) := \sum_{k=1}^{+\infty} \xi_k(t)\chi_{A_k}(t) \quad \bar p(t) := \sum_{k=1}^{+\infty} p_k(t)\chi_{A_k}(t) \quad \text{for all } t \in \mathcal{T}.$$

Notice that the map $t \mapsto (\bar\xi(t),\bar p(t))$ is a measurable selection of $\mathcal{K}(t,u(t))$, and that

$$\Psi^*_{u(t)}(-\bar\xi(t)) \le \mathcal{M}_*(t) + 1 \qquad \text{for every } t \in \mathcal{T}. \tag{B.12}$$

In particular, it follows from (4.Ψ_2) that $\bar\xi \in L^1(0,T;V^*)$. Then, we use $(\bar\xi,\bar p)$ to construct a new countable family of functions, setting

$$(\xi_{n,k}(t),p_{n_k}(t)) := \begin{cases} (\xi_n(t),p_n(t)) & \text{if } \Psi^*_{u(t)}(-\xi_n(t)) \le k, \\ (\bar\xi(t),\bar p(t)) & \text{otherwise,} \end{cases}$$

such that the functions $\xi_{n,k}$ belong to $L^1(0,T;V^*)$, and the pairs $(\xi_{n,k}, p_{n_k})$ satisfy

$$(\xi_{n,k}(t), p_{n_k}(t)) \in \mathcal{K}(t,u(t)), \quad \{(\xi_n(t), p_n(t)) : n \in \mathbb{N}\} \subset \{(\xi_{n,k}(t), p_{n_k}(t)) : n,k \in \mathbb{N}\} \quad \text{if } t \in T, \quad \square$$

as well as estimate (B.9), in view of (B.11).

Claim 3: *inequality* (B.5) *holds.*

Indeed, it follows from Claims 1 and 2 that we can apply the chain rule (4.E$_5$) to the pairs (u,ξ_n) for all $n \in \mathbb{N}$ (indeed, estimate (B.2) and the first of (B.3) yield $\int_0^T \Psi_{u(t)}(u'(t))\,\mathrm{d}t < +\infty$). Therefore, we conclude for all $n \in \mathbb{N}$ that there exists a set of full measure $\mathcal{T}_n \subset \mathcal{T}$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) \geq \langle \xi_n(t), u'(t)\rangle + \mathrm{P}_t(u(t),\xi_n(t)) \geq \langle \xi_n(t), u'(t)\rangle + p_n(t) \quad \text{for all } t \in \mathcal{T}_n.$$

Thus, we infer

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) \geq \langle \xi, u'(t)\rangle + p \quad \text{for all } (\xi,p) \in \overline{\mathrm{conv}}(\mathcal{K}(t,u(t))), \text{ for all } t \in \mathcal{T}_\infty, \tag{B.13}$$

with $\overline{\mathrm{conv}}(\mathcal{K}(t,u(t)))$ the closed convex hull of $\mathcal{K}(t,u(t))$ and $\mathcal{T}_\infty = \bigcap_{n\in\mathbb{N}} \mathcal{T}_n$ (note that $\mathcal{T}_\infty$ is a subset of $(0,T)$ of full measure, too). Then, (B.5) follows upon integrating (B.13) with respect to the measure μ_t, again taking into account (B.4) and (B.3). ∎

We conclude with the following

Lemma B.2 (Measurable selection)**.** *In the framework of* (4.Ψ_1)–(4.Ψ_2)*, suppose that* $\mathcal{E}$ *complies with* (4.E$_0$)–(4.E$_6$)*. Let* $u \in \mathrm{AC}([0,T];V)$ *be an absolutely continuous curve complying with* (B.1)*, and suppose that the set*

$$\mathcal{S}(t,u(t),u'(t)) := \left\{(\zeta,p) \in V^* \times \mathbb{R} \,:\, \zeta \in \mathrm{F}_t(u(t)) \cap (-\partial\Psi_{u(t)}(u'(t))),\ p \leq \mathrm{P}_t(u(t),\zeta)\right\}, \quad \textit{is nonempty for almost all } t \in (0,T). \tag{B.14}$$

Then, there exists measurable functions $\xi : (0,T) \to V^*$*,* $p : (0,T) \to \mathbb{R}$ *such that*

$$(\xi(t),p(t)) \in \mathrm{Argmin}\{\Psi^*_{u(t)}(-\zeta) - p \,:\, (\zeta,p) \in \mathcal{S}(t,u(t),u'(t))\} \quad \textit{for a.a. } t \in (0,T). \tag{B.15}$$

Proof: First of all, let us observe that

$$\mathrm{Argmin}\{\Psi^*_{u(t)}(-\zeta) - p \,:\, (\zeta,p) \in \mathcal{S}(t,u(t),u'(t))\} \neq \emptyset \quad \text{for a.a. } t \in (0,T). \tag{B.16}$$

For, let $(\xi_n,p_n) \subset \mathcal{S}(t,u(t),u'(t))$ be a infimizing sequence: then, there exist some positive constants C and C' such that

$$\Psi^*_{u(t)}(-\xi_n) \leq C + p_n \leq C + \mathrm{P}_t(u(t),\xi_n) \leq C + C_2\mathcal{G}(u(t)) \leq C' \quad \text{for every } n \in \mathbb{N}, \tag{B.17}$$

where the first inequality trivially follows from the fact that (ξ_n,p_n) is infimizing, the second one from the definition of $\mathcal{S}(t,u(t),u'(t))$, the third one from (4.E$_4$), and the last one from (4.5) and assumption (B.1). In view of the latter, and of the superlinear growth condition (4.Ψ_2), from the bound for $\Psi^*_{u(t)}(-\xi_n)$ we infer that $\sup_n \|\xi_n\|_* < +\infty$. It is also clear from (B.17) that $\sup_n |p_n| < +\infty$, therefore there exist (ξ_*,p_*) such that, up to a subsequence, $\xi_n \rightharpoonup \xi_*$ in V^* and $p_n \to p_*$. Exploiting the closedness condition (4.E$_6$), we infer that $(\xi_*,p_*) \in \mathcal{S}(t,u(t),u'(t))$, and by lower semicontinuity

$$\Psi^*_{u(t)}(-\xi_*) - p_* \leq \liminf_{n\to\infty}\Big(\Psi^*_{u(t)}(-\xi_n) - p_n\Big) = \inf_{(\zeta,p)\in\mathcal{S}(t,u(t),u'(t))}\{\Psi^*_{u(t)}(-\zeta) - p\},$$

whence (B.16).

In order to prove (B.15), we first observe that for all $t \in [0,T]$ the set

$$S := \{(t,u,v,\xi,p) \in [0,T] \times V \times V \times V^* \times \mathbb{R} \,:\, \xi \in \mathrm{F}_t(u) \cap (-\partial\Psi_u(v)),\ p \leq \mathrm{P}_t(u,\xi)\} \quad \text{is a Borel set of } [0,T] \times V \times V \times V^* \times \mathbb{R}. \tag{B.18}$$

This follows from the same arguments as for (B.8) in Proposition B.1. Now, due to (B.14) there exists a subset $\mathcal{T}' \subset (0,T)$ of full measure such that $\mathcal{S}(t,u(t),u'(t)) \neq \emptyset$ for all $t \in \mathcal{T}'$. We thus consider the graph of the multivalued function $t \in \mathcal{T}' \mapsto \mathcal{S}(t,u(t),u'(t)) \subset V^* \times \mathbb{R}$, i.e. the set

$$\begin{aligned} \mathscr{S} &:= \{(t,\xi,p) \in \mathcal{T}' \times V^* \times \mathbb{R} \, : \, (\xi,p) \in \mathcal{S}(t,u(t),u'(t))\} \\ &= \{(t,\xi,p) \in \mathcal{T}' \times V^* \times \mathbb{R} \, : \, (t,u(t),u'(t),\xi,p) \in S\}\,. \end{aligned}$$

Then, we combine the latter representation of $\mathscr{S}$ with (B.18), and the fact that the functions $u : (0,T) \to V$ and $u' : (0,T) \to V$ are Borelian up to choosing a suitable representative for u'. Thus, we conclude that $\mathscr{S}$ is a Borel set of $\mathcal{T}' \times V^* \times \mathbb{R}$. Hence, the existence of a measurable selection (ξ,p) as in (B.15) is a consequence of [14, Cor. III.3, Thm. III.6], cf. also Filippov's theorem, see e.g. [4, Thm. 8.2.11]. ∎

References

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

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

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

[4] J.-P. Aubin and H. Frankowska, *Set-valued Analysis*, Birkhäuser, Boston, 1990.

[5] J.-F. Babadjian, G. Francfort, and M. Mora, *Quasistatic evolution in non-associative plasticity - the cap model.* Submitted, 2011.

[6] C. Baiocchi, *Discretization of evolution variational inequalities*, in Partial differential equations and the calculus of variations, Vol. I, F. Colombini, A. Marino, L. Modica, and S. Spagnolo, eds., Birkhäuser Boston, Boston, MA, 1989, pp. 59–92.

[7] E. J. Balder, *A general approach to lower semicontinuity and lower closure in optimal control theory*, SIAM J. Control Optim., 22 (1984), pp. 570–598.

[8] ———, *An extension of Prohorov's theorem for transition probabilities with applications to infinite-dimensional lower closure problems*, Rend. Circ. Mat. Palermo (2), 34 (1985), pp. 427–447 (1986).

[9] ———, *Lectures on Young measure theory and its applications in economics*, Rend. Istit. Mat. Univ. Trieste, 31 (2000), pp. 1–69. Workshop on Measure Theory and Real Analysis (Italian) (Grado, 1997).

[10] J. M. Ball, *Convexity conditions and existence theorems in nonlinear elasticity*, Arch. Rational Mech. Anal., 63 (1976/77), pp. 337–403.

[11] J. M. Ball, *A version of the fundamental theorem for Young measures*, in PDEs and continuum models of phase transitions (Nice, 1988), vol. 344 of Lecture Notes in Phys., Springer, Berlin, 1989, pp. 207–215.

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

[13] C. Castaing, P. Raynaud de Fitte, and M. Valadier, *Young measures on topological spaces*, vol. 571 of Mathematics and its Applications, Kluwer Academic Publisher, Dordrecht, 2004.

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

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

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

[17] M. G. Crandall and T. M. Liggett, *Generation of semi-groups of nonlinear transformations on general Banach spaces*, Amer. J. Math., 93 (1971), pp. 265–298.

[18] E. De Giorgi, *New problems on minimizing movements*, in Boundary Value Problems for PDE and Applications, C. Baiocchi and J. L. Lions, eds., Masson, 1993, pp. 81–98.

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

[20] A. DeSimone, *Energy minimizers for large ferromagnetic bodies*, Arch. Rational Mech. Anal., 125 (1993), pp. 99–143.

[21] N. Dunford and J. T. Schwartz, *Linear Operators. Part I*, Interscience Publishers, a division of John Wiley & Sons, Inc., New York, 1958.

[22] C. Ebmeyer and J. Frehse, *Mixed boundary value problems for non-linear elliptic equations in multidimensional non-smooth domains*, Math. Nachrichten, 203 (1999), pp. 47–74.

[23] I. Ekeland and R. Temam, *Analyse Convexe et Problèmes Variationnels*, Dunod, Gauthier-Villars, Paris, 1974.

[24] G. Francfort and A. Mielke, *Existence results for a class of rate-independent material models with nonconvex elastic energies*, J. reine angew. Math., 595 (2006), pp. 55–91.

[25] M. Frémond, *Non-smooth Thermomechanics*, Springer-Verlag, Berlin, 2002.
[26] S. Govindjee, A. Mielke, and G. Hall, *The free–energy of mixing for n–variant martensitic phase transformations using quasi–convex analysis*, J. Mech. Physics Solids, 50 (2002), pp. 1897–1922. Erratum and Correct Reprinting: 51(4) 2003, pp. 763 & I-XXVI.
[27] A. D. Ioffe, *On lower semicontinuity of integral functionals. I*, SIAM J. Control Optimization, 15 (1977), pp. 521–538.
[28] M. Kružík, A. Mielke, and T. Roubíček, *Modelling of microstructure and its evolution in shape-memory-alloy single-crystals, in particular in CuAlNi*, Meccanica, 40 (2005), pp. 389–418.
[29] S. Luckhaus, *Solutions for the two-phase Stefan problem with the Gibbs-Thomson Law for the melting temperature*, Euro. Jnl. of Applied Mathematics, 1 (1990), pp. 101–111.
[30] A. Mainik and A. Mielke, *Global existence for rate-independent gradient plasticity at finite strain*, J. Nonlinear Sci., 19 (2009), pp. 221–248.
[31] A. Mielke, *Energetic formulation of multiplicative elasto–plasticity using dissipation distances*, Cont. Mech. Thermodynamics, 15 (2003), pp. 351–382.
[32] A. Mielke, *Evolution in rate-independent systems (Ch. 6)*, in Handbook of Differential Equations, Evolutionary Equations, vol. 2, C. Dafermos and E. Feireisl, eds., Elsevier B.V., Amsterdam, 2005, pp. 461–559.
[33] ———, *A mathematical framework for generalized standard materials in the rate-independent case*, in Multifield problems in solid and fluid mechanics, vol. 28 of Lect. Notes Appl. Comput. Mech., Springer, Berlin, 2006, pp. 399–428.
[34] A. Mielke, R. Rossi, and G. Savaré, *Bv solutions and viscosity approximations of rate- independent systems*, ESAIM Control Optim. Calc. Var. Published online since 23.12.2010.
[35] ———, *Bv solutions to infinite-dimensional rate-independent systems*. In preparation, 2011.
[36] A. Mielke, R. Rossi, and G. Savaré, *Modeling solutions with jumps for rate-independent systems on metric spaces*, Discrete Contin. Dyn. Syst., 25 (2009), pp. 585–615.
[37] A. Mielke and T. Roubíček, *Rate-independent damage processes in nonlinear elasticity*, M^3AS Math. Models Methods Appl. Sci., 16 (2006), pp. 177–209.
[38] A. Mielke, F. Theil, and V. I. Levitas, *A variational formulation of rate–independent phase transformations using an extremum principle*, Arch. Rational Mech. Anal., 162 (2002), pp. 137–177. (Essential Science Indicator: Emerging Research Front, August 2006).
[39] B. S. Mordukhovich, *Nonsmooth analysis with nonconvex generalized differentials and conjugate mappings*, Dokl. Akad. Nauk BSSR, 28 (1984), pp. 976–979.
[40] B. S. Mordukhovich, *Variational analysis and generalized differentiation. I*, vol. 330 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 2006. Basic theory.
[41] R. H. Nochetto, G. Savaré, and C. Verdi, *A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations*, Comm. Pure Appl. Math., 53 (2000), pp. 525–589.
[42] Y. Reshetnyak, *On the stability of conformal maps in multidimensional spaces*, Siberian Math. J., 8 (1967), pp. 69–85.
[43] 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), VII (2008), pp. 97–169.
[44] R. Rossi and G. Savaré, *Gradient flows of non convex functionals in Hilbert spaces and applications*, ESAIM Control Optim. Calc. Var., 12 (2006), pp. 564–614 (electronic).
[45] R. Rossi, A. Segatti, and U. Stefanelli, *Attractors for gradient flows of non convex functionals and applications to quasistationary phase field models*, Arch. Ration. Mech. Anal., 187 (2008), pp. 91–135.
[46] ———, *Global attractors for gradient flows in metric spaces*, J. Math. Pures Appl., 95 (2011), pp. 205–244.
[47] J. Rulla, *Error analysis for implicit approximations to solutions to Cauchy problems*, SIAM J. Numer. Anal., 33 (1996), pp. 68–87.
[48] G. Savaré, *Weak solutions and maximal regularity for abstract evolution inequalities*, Adv. Math. Sci. Appl., 6 (1996), pp. 377–418.
[49] G. Savaré, *Regularity results for elliptic equations in Lipschitz domains*, J. Funct. Anal., 152 (1998), pp. 176–201.
[50] R. Schätzle, *The quasistationary phase field equations with Neumann boundary conditions*, J. Differential Equations, 162 (2000), pp. 473–503.
[51] G. Schimperna, A. Segatti, and U. Stefanelli, *Well-posedness and long-time behavior for a class of doubly nonlinear equations*, Discrete Contin. Dyn. Syst., 18 (2007), pp. 15–38.
[52] A. Segatti, *Global attractor for a class of doubly nonlinear abstract evolution equations*, Discrete Contin. Dyn. Syst., 14 (2006), pp. 801–820.
[53] U. Stefanelli, *The Brezis-Ekeland principle for doubly nonlinear equations*, SIAM J. Control Optim., 47 (2008), pp. 1615–1642.
[54] M. Valadier, *Young measures*, in Methods of nonconvex analysis (Varenna, 1989), Springer, Berlin, 1990, pp. 152–188.
[55] A. Visintin, *Models of Phase Transitions*, vol. 28 of Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser, Boston, 1996.

Weierstrass-Institut, Mohrenstrasse 39, 10117 D–Berlin and Institut für Mathematik, Humboldt-Universität zu Berlin, Rudower Chaussee 25, D–12489 Berlin (Adlershof), Germany.
E-mail address: mielke@wias-berlin.de

Dipartimento di Matematica, Università di Brescia, via Valotti 9, I–25133 Brescia, Italy.
E-mail address: riccarda.rossi@ing.unibs.it

Dipartimento di Matematica "F. Casorati", Università di Pavia. Via Ferrata, 1 – 27100 Pavia, Italy.
E-mail address: giuseppe.savare@unipv.it