GRADIENT FLOWS OF NON CONVEX FUNCTIONALS IN HILBERT SPACES AND APPLICATIONS [*]

RICCARDA ROSSI[1] AND GIUSEPPE SAVARÉ[1]

Abstract. This paper addresses the Cauchy problem for the gradient flow equation in a Hilbert space $\mathscr{H}$

$$\begin{cases} u'(t) + \partial_\ell \phi(u(t)) \ni f(t) & a.e. \text{ in } (0, T), \\ u(0) = u_0, \end{cases}$$

where $\phi : \mathscr{H} \to (-\infty, +\infty]$ is a proper, lower semicontinuous functional which is not supposed to be a (smooth perturbation of a) convex functional and $\partial_\ell \phi$ is (a suitable limiting version of) its subdifferential. We will present some new existence results for the solutions of the equation by exploiting a *variational approximation* technique, featuring some ideas from the theory of *Minimizing Movements* and of *Young measures*.

Our analysis is also motivated by some models describing phase transitions phenomena, leading to systems of evolutionary PDEs which have a *common underlying gradient flow structure*: in particular, we will focus on *quasistationary* models, which exhibit highly non convex Lyapunov functionals.

Mathematics Subject Classification. 35A15, 35K50, 35K85, 58D25, 80A22.

1. Introduction and main results

The aim of this paper is to study the existence and the approximation of strong solutions of the gradient flow equation

$$\begin{cases} u'(t) + \partial_\ell \phi(u(t)) \ni f(t) & a.e. \text{ in } (0, T), \\ u(0) = u_0, \end{cases} \tag{GF}$$

associated with the *limiting subdifferential* $\partial_\ell \phi : \mathscr{H} \to 2^{\mathscr{H}}$ of a proper and lower semicontinuous functional $\phi : \mathscr{H} \to (-\infty, +\infty]$ defined in a (separable) Hilbert space $\mathscr{H}$ with scalar product $\langle \cdot, \cdot \rangle$ and norm $|\cdot|$; $D(\phi) := \{v \in \mathscr{H} : \phi(v) < +\infty\}$ will denote the proper domain of ϕ.

The limiting subdifferential $\partial_\ell \phi$ is obtained by taking a sequential strong-weak closure in $\mathscr{H} \times \mathscr{H}$ of the graph of the *Fréchet subdifferential* $\partial \phi$, which is the (possibly multivalued) operator defined for every $v \in D(\phi)$ by

$$\partial \phi(v) := \Big\{ \xi \in \mathscr{H} : \phi(w) - \phi(v) - \langle \xi, w - v \rangle \geq o(|w - v|) \quad \text{as } w \to v \Big\}, \tag{1.1}$$

where the above Landau's notation should be understood as

$$\liminf_{w \to v} \frac{\phi(w) - \phi(v) - \langle \xi, w - v \rangle}{|w - v|} \geq 0. \tag{1.2}$$

Thus, a vector ξ belongs to the limiting subdifferential $\partial_\ell \phi$ at $v \in D(\phi)$ if there exist sequences

$$v_n, \xi_n \in \mathscr{H} \quad \text{such that} \quad \xi_n \in \partial \phi(v_n), \ v_n \to v, \ \xi_n \rightharpoonup \xi, \ \sup_{n \in \mathbb{N}} \phi(v_n) < +\infty \quad \text{as } n \uparrow +\infty. \tag{1.3}$$

As usual in multivalued analysis, we will denote by $D(\partial \phi), D(\partial_\ell \phi)$ the proper domains

$$D(\partial \phi) := \big\{ v \in \mathscr{H} : \partial \phi(v) \neq \emptyset \big\}, \quad D(\partial_\ell \phi) := \big\{ v \in \mathscr{H} : \partial_\ell \phi(v) \neq \emptyset \big\}. \tag{1.4}$$

Before discussing the motivations for introducing and studying this kind of subdifferential operators (see [25, 33], and the monograph [37], Chap. VIII, where analogous notions are introduced for different purposes) and the related evolution equations, let us first recall the well-established theories that cover some simpler situations.

The convex case: existence and regularity. When ϕ is *convex*, the Fréchet (and the limiting) subdifferential coincides with the usual subdifferential of convex analysis, and it can be characterized by

$$\xi \in \partial \phi(v) \quad \Leftrightarrow \quad v \in D(\phi), \quad \phi(w) - \phi(v) - \langle \xi, w - v \rangle \geq 0 \quad \forall w \in \mathscr{H}. \tag{1.5}$$

Since ϕ is lower semicontinuous (l.s.c., in the sequel), taking the *strong-weak* closure of the graph of $\partial \phi$ in $\mathscr{H} \times \mathscr{H}$ does not modify it, since

$$v_n \to v, \ \xi_n \rightharpoonup \xi, \ \xi_n \in \partial \phi(v_n) \ \Rightarrow \ \xi \in \partial \phi(v), \quad \text{so that} \quad \partial_\ell \phi \equiv \partial \phi. \tag{1.6}$$

It is well-known that $\partial \phi$ is a (possibly multivalued) *maximal monotone operator*; existence, uniqueness, and regularity of the solution of (GF) follow from the well-known theory developed by Komura [24], Crandall-Pazy [18], Brézis [11]: we refer to the monograph [12]. In particular, if

$$u_0 \in D(\phi), \quad f \in L^2(0, T; \mathscr{H}), \tag{DATA}$$

then the solution u belongs to $H^1(0, T; \mathscr{H})$, for *a.e.* $t \in (0, T)$ its derivative

$$u'(t) \text{ is } \textit{the projection of the origin on the affine hull } \text{aff}\big(f(t) - \partial \phi(u(t)) \big),$$
$$\text{thus realizes the } \textit{minimal section principle:} \quad u'(t) = \big(f(t) - \partial \phi(u(t)) \big)^\circ, \tag{1.7}$$

where for every subset $A \subset \mathscr{H}$ we set

$$\text{aff } A := \Big\{ \sum_i t_i a_i : a_i \in A, \ t_i \in \mathbb{R}, \ \sum_i t_i = 1 \Big\}, \tag{1.8}$$

$$|A^\circ| := \inf_{\xi \in A} |\xi|, \quad A^\circ := \big\{ \xi \in A : |\xi| = |A^\circ| \big\}; \tag{1.9}$$

moreover, u satisfies the *energy identity*

$$\phi(u(T)) + \int_0^T |u'(t)|^2 \, \mathrm{d}t = \phi(u_0) + \int_0^T \langle f(t), u'(t) \rangle \, \mathrm{d}t, \tag{1.10}$$

which is an immediate consequence of the *Chain Rule*

$$\begin{aligned}
&\text{if} \quad u \in H^1(0,T;\mathscr{H}), \ \xi \in L^2(0,T;\mathscr{H}), \ \xi(t) \in \partial\phi(u(t)) \text{ for } a.e. \ t \in (0,T) \\
&\text{then} \quad \phi \circ u \in AC(0,T), \quad \tfrac{\mathrm{d}}{\mathrm{d}t}\phi(u(t)) = \langle \xi(t), u'(t) \rangle \text{ for } a.e. \ t \in (0,T).
\end{aligned} \tag{1.11}$$

The convex case: approximation. Approximating (GF) by the *implicit Euler scheme*

$$U_\tau^0 := u_0, \quad \frac{U_\tau^n - U_\tau^{n-1}}{\tau} + \partial\phi(U_\tau^n) \ni F_\tau^n \quad n = 1, \ldots, N, \tag{1.12}$$

is one of the possible ways of proving the above results, and provides a useful constructive method which is interesting by itself: here $\tau > 0$ is the time step with $N\tau = T$, F_τ^n are suitable approximations of the values of f in the interval $((n-1)\tau, n\tau]$, *e.g.*,

$$F_\tau^n := \frac{1}{\tau} \int_{(n-1)\tau}^{n\tau} f(t) \, \mathrm{d}t, \tag{1.13}$$

and the sequence $\{U_\tau^n\}_{n=1}^N$ is uniquely determined by solving (1.12) recursively, starting from the assigned value of U_τ^0.

If we denote by $U_\tau(t)$ the piecewise linear interpolant taking the value U_τ^n at t_n, several kinds of more and more refined estimates of the error $|u(t) - U_\tau(t)|$ could be derived, starting from the pioneering ones of Crandall-Liggett [17]: we mention [12], Cor. 4.4, the optimal *a priori* estimates of [4, 40, 41], and the optimal *a posteriori* estimates of [34] for even non uniform meshes. In our case, assuming for simplicity $\phi \geq 0$, it is possible to prove the uniform Cauchy estimates (see [34])

$$\max_{t \in [0,T]} |U_\eta(t) - U_\tau(t)|^2 \leq 3(\tau + \eta)\left(\phi(u_0) + \int_0^T |f(t)|^2 \, \mathrm{d}t\right), \tag{1.14}$$

thus showing that the family of functions U_τ is uniformly convergent to a continuous function u as $\tau \downarrow 0$. Note that (1.14), the *discrete energy estimate* analogous to (1.10)

$$\phi(U_\tau^n) + \tau \sum_{k=1}^N \left|\frac{U_\tau^k - U_\tau^{k-1}}{\tau}\right|^2 \leq \phi(u_0) + \tau \sum_{k=1}^N \left\langle F_\tau^k, \frac{U_\tau^k - U_\tau^{k-1}}{\tau} \right\rangle, \tag{1.15}$$

and the *strong-weak* closure of the graph of $\partial\phi$ (1.6) are the main ingredients to show that the uniform limit u of U_τ as $\tau \downarrow 0$ belongs to $H^1(0,T;\mathscr{H})$ and is the (unique) solution of (GF).

Quadratic perturbations of convex functions. The previous results extend (up to an exponential factor, which modifies the constant in (1.14)) to λ-*convex* functionals, *i.e.*, *quadratic perturbations of convex functions* satisfying

$$\exists \lambda \geq 0: \quad v \mapsto \phi(v) + \frac{\lambda}{2}|v|^2 \quad \text{is convex.} \tag{1.16}$$

Of course, in this case the characterization (1.5) of the Fréchet subdifferential is affected by λ, namely

$$\xi \in \partial\phi(v) \quad \Leftrightarrow \quad v \in D(\phi), \quad \phi(w) - \phi(v) - \langle \xi, w - v \rangle \geq -\frac{\lambda}{2}|w - v|^2 \quad \forall w \in \mathscr{H},$$

which shows that in the λ-convex case we can always choose the infinitesimal term $o(r) := -\frac{\lambda}{2}r^2$ in (1.1); of course, this characterization degenerates as $\lambda \uparrow +\infty$.

It is not hard to check that any function ϕ which admits the representation

$$\phi(v) = \psi_1(v) - \psi_2(v), \quad \text{with} \quad \psi_1 \text{ convex}, \quad \psi_2 \in C^{1,1}(\mathscr{H}), \tag{1.17}$$

satisfies (1.16), simply by taking as λ the Lipschitz constant of the (Fréchet) differential $D\psi_2$; further, (1.17) yields the decomposition

$$\partial\phi(v) = \partial\psi_1(v) - D\psi_2(v) \quad \forall v \in D(\partial\phi). \tag{1.18}$$

C^1 **perturbation of convex functions.** When ϕ is not a quadratic perturbation of a convex function any more, things become remarkably more difficult. It was one of the main achievements of the theory of Curves of Maximal Slope, developed in a series of papers originating from [20] and culminating in [29] (but see also the more recent [15] and the presentations [1,3]), to partially extend the previous existence results to the case in which ϕ admits the decomposition (1.17), where ψ_2 is now *simply of class $C^1(\mathscr{H})$*, provided $f \equiv 0$ and ϕ satisfies the coercivity/compactness property

$$\exists \tau_* > 0 : \quad v \mapsto \phi(v) + \frac{1}{2\tau_*}|v|^2 \quad \text{has compact sublevels.} \tag{COMP}$$

In this case, the approximation algorithm (1.12) has to be rewritten in a variational form, observing that (1.12) is in fact the Euler equation associated with the functional

$$\Phi(\tau, F_\tau^n, U_\tau^{n-1}; V) := \frac{1}{2\tau}\left|V - U_\tau^{n-1}\right|^2 + \phi(V) - \langle F_\tau^n, V\rangle. \tag{1.19}$$

Then, (1.12) is replaced by the variational iterative scheme

$$\begin{cases} U_\tau^0 := u_0 \text{ is given; whenever } U_\tau^1, \ldots, U_\tau^{n-1} \text{ are known,} \\[2mm] \text{find } U_\tau^n \in \mathscr{H} : \ \Phi(\tau, F_\tau^n, U_\tau^{n-1}; U_\tau^n) \leq \Phi(\tau, F_\tau^n, U_\tau^{n-1}; V) \quad \forall V \in \mathscr{H}, \end{cases} \tag{1.20}$$

which *a fortiori* also yields a solution U_τ^n of (1.12), since for every functional ψ in $C^1(\mathscr{H})$, the Fréchet subdifferential obeys the usual calculus rules

$$\begin{aligned} \partial(\phi + \psi)(v) &= \partial\phi(v) + D\psi(v) \quad \forall v \in \mathscr{H}, \\ v_0 \text{ minimizes } \phi + \psi \text{ in } \mathscr{H} &\Rightarrow v_0 \in D(\partial\phi), \quad \partial\phi(v_0) + D\psi(v_0) \ni 0. \end{aligned} \tag{1.21}$$

This variational approach has been independently applied to different kinds of problems (see *e.g.* [22, 26, 30]), and has been proposed (in an even more general formulation) as a possible general method to study Gradient Flows in [19] (see also the lecture notes [1,3]).

Unlike the convex framework, solutions to (1.20) are not unique, in general: in any case, we call *discrete solution* any piecewise constant interpolant $\overline{U}_\tau$ of a sequence of discrete values solving (1.20), *i.e.*, $\overline{U}_\tau(t) := U_\tau^n$ if $t \in ((n-1)\tau, n\tau]$. Following [19], we say that

> u is a *generalized Minimizing Movement* associated with the scheme (1.20) if there exist a subsequence $\tau_k \downarrow 0$ and a corresponding family of discrete solutions $\overline{U}_{\tau_k}$ such that

$$\lim_{k\uparrow+\infty} \overline{U}_{\tau_k}(t) = u(t) \quad \text{in } \mathscr{H}, \quad \forall t \in [0, T]. \tag{1.22}$$

We denote by $\mathrm{GMM}(\Phi; u_0, f)$ the collection of the generalized Minimizing Movements starting from u_0 and with forcing term f.

In order to recover information on the time derivative of the limit functions, we will also often consider the *piecewise linear* interpolant U_τ of the values U_τ^n.

Although it is not easy to present a short overview of the wide and complex set of assumptions considered in [29], following [1] we observe that the crucial assumption of their approach is the *strong-weak closedness* of the graph of $(\partial\phi, \phi)$ in $\mathscr{H} \times \mathscr{H} \times \mathbb{R}$, *i.e.*,

$$\left.\begin{array}{l} \xi_n \in \partial\phi(v_n), \ r_n = \phi(v_n) \\ v_n \to v, \ \xi_n \rightharpoonup \xi, \ r_n \to r \end{array}\right\} \quad \Rightarrow \quad \xi \in \partial\phi(v), \ r = \phi(v), \tag{1.23}$$

which in particular yields $\partial_\ell\phi \equiv \partial\phi$. Under (1.23), they prove that for each choice of $u_0 \in D(\phi)$ satisfying (DATA), $\mathrm{GMM}(\Phi; u_0, 0)$ is not empty and that its elements are solutions of (GF).

Note that (1.23) yields three crucial properties for $\partial_\ell\phi$, which are somehow hidden in the proof of the existence result in [29]:

$\partial_\ell\phi$ is convex-valued:

$$\partial_\ell\phi(v) \quad \text{is a closed convex subset of } \mathscr{H} \quad \forall v \in D(\partial_\ell\phi). \tag{CONV}$$

Chain rule: If $v \in H^1(0, T; \mathscr{H})$, $\xi \in L^2(0, T; \mathscr{H})$ with $\xi(t) \in \partial_\ell\phi(v(t))$ for *a.e.* $t \in (0, T)$, and $\phi \circ v$ is *a.e.* equal to a function φ of bounded variation, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = \langle\xi(t), v'(t)\rangle \quad \text{for } a.e. \ t \in (0, T). \tag{CHAIN$_1$}$$

Continuity of ϕ along sequences with equibounded slope:

$$v_n \to v, \ \sup_n\left(|\partial^\circ\phi(v_n)|, \phi(v_n)\right) < +\infty \ \Rightarrow \ \phi(v_n) \to \phi(v) \quad \text{as } n \uparrow +\infty. \tag{CONT}$$

Remark 1.1. We stress once again that if ϕ is λ-convex, *i.e.*, it fulfills (1.16), then it also satisfies (1.23) and therefore (CONV, CHAIN$_1$, CONT); moreover, the chain rule holds in the stronger formulation (1.11).

A more general situation. In this paper, we will show that general existence results can be proved even when *only one* of the two assumptions (CONV) and (CHAIN$_1$) is supposed to hold; (CONT) will play a complementary role, which will be discussed in each situation. In particular, we can consider functionals whose Fréchet subdifferential is not strongly-weakly closed in the sense of (1.23). Actually, the (strong-weak) closedness of $\partial\phi$ may fail even for simple one-dimensional functionals, as in the case of (*cf.* also (2.14, 2.17) later on)

$$\mathscr{H} := \mathbb{R}, \quad \phi(x) := \begin{cases} \frac{1}{3}(x+1)^2 & x \leq 0, \\ \frac{1}{3}(x-1)^2 & x > 0, \end{cases} = \min\left\{\frac{1}{3}(x+1)^2, \frac{1}{3}(x-1)^2\right\} \tag{1.24}$$

(see Fig. 1). It is easy to check that $\partial\phi(x)$ is reduced to a singleton for $x \neq 0$, while $\partial\phi(0)$ is empty, so that obviously $\partial\phi$ differs from its closure $\partial_\ell\phi$ at $x = 0$, which turns out to be *non convex*. In turn, note that ϕ cannot be decomposed as in (1.17), even if ψ_2 is required to be only of class $C^1(\mathscr{H})$.

As we will see in Example 2, this real function provides the simplest one-dimensional caricature of infinite-dimensional Lyapunov functionals arising in quasistationary models for phase transitions. In spite of its triviality, it captures two main features, which are typical in those evolution models: the presence of anti-monotone jumps and non convex sections in the graph of $\partial_\ell\phi$ (see Fig. 2). This lack of convexity is a serious difficulty, since in this general infinite-dimensional setting only weak convergence properties are available for the time derivatives of any family of approximating solutions to (GF).
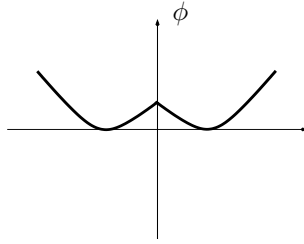
FIGURE 1. The potential ϕ of (1.24) is not subdifferentiable at $x = 0$.
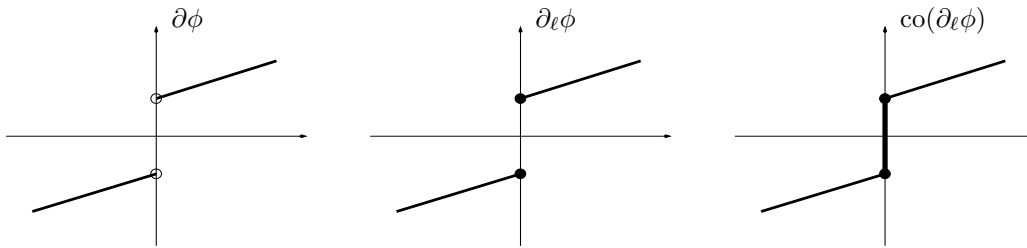

FIGURE 2. Fréchet, limiting, and convexified subdifferential of ϕ

We will see in Example 2 that, even in the finite-dimensional situation, the convexification of $\partial_\ell \phi$ (*e.g.*, in the present case, at $x = 0$) is often not acceptable, since solutions of the easier relaxed formulation do not solve, in general, the original one.

We postpone a detailed presentation of some other significant examples to the next section, while here we are going to present our main abstract results. First of all, we state some natural and general compactness conditions which guarantee that $\mathrm{GMM}(\Phi; u_0, f)$ is not empty. Recall that U_τ will denote the piecewise linear interpolant of the discrete values $\{U_\tau^n\}_{n=0}^N$.

Lemma 1.2 (Compactness)**.** *Let us assume that*

$$\phi : \mathscr{H} \to (-\infty, +\infty] \quad \text{is proper, lower semicontinuous, and}$$

$$\exists \tau_* > 0 : \quad v \mapsto \phi(v) + \frac{1}{2\tau_*}|v|^2 \quad \text{has compact sublevels,} \tag{COMP}$$

and the data satisfy

$$u_0 \in D(\phi), \quad f \in L^2(0, T; \mathscr{H}). \tag{DATA}$$

Then, there exists a constant $C > 0$ (independent of τ) such that

$$\|\overline{U}_\tau\|_{L^\infty(0,T;\mathscr{H})} + \|U_\tau\|_{L^\infty(0,T;\mathscr{H})} \le C, \tag{1.25}$$

$$\sup_{[0,T]} \phi(\overline{U}_\tau) \le C, \tag{1.26}$$

$$\|U_\tau'\|_{L^2(0,T;\mathscr{H})} \le C, \quad \|\overline{U}_\tau - U_\tau\|_{L^\infty(0,T;\mathscr{H})} \le C\sqrt{\tau}, \tag{1.27}$$

for every $0 < \tau \le \tau_*/10$. In particular, $\mathrm{GMM}(\Phi; u_0, f)$ is not empty and every $u \in \mathrm{GMM}(\Phi; u_0, f)$ belongs to $H^1(0, T; \mathscr{H})$.

The case of a convex-valued limiting subdifferential. Here the main assumption on ϕ, besides the compactness one (COMP), is the convexity of the values of $\partial_\ell \phi$ (CONV): indeed, conditions (COMP) and (CONV) are sufficient in order to prove that every $u \in \mathrm{GMM}(\Phi; u_0, f)$ is a solution of (GF) and satisfies a natural Lyapunov-like inequality, which holds in a stronger form when also (CONT) is verified.

Theorem 1 (Lyapunov solutions). *Let us suppose that $\phi : \mathscr{H} \to (-\infty, +\infty]$, u_0, and f satisfy the assumptions* (COMP) *and* (DATA) *of the compactness Lemma 1.2, so that* $\mathrm{GMM}(\Phi; u_0, f)$ *is not empty. If*

$$\partial_\ell \phi(v) \text{ is a convex closed } subset \text{ of } \mathscr{H} \text{ for every } v \in D(\partial_\ell \phi), \qquad \text{(CONV)}$$

then any $u \in \mathrm{GMM}(\Phi; u_0, f)$ is a solution to (GF) *satisfying the Lyapunov inequality*

$$\int_0^t \left(\frac{1}{2}|u'(\sigma)|^2 + \frac{1}{2}|(\partial_\ell \phi(u(\sigma)) - f(\sigma))^0|^2 \right) \mathrm{d}\sigma + \phi(u(t)) \leq \phi(u_0) + \int_0^t \langle f(\sigma), u'(\sigma) \rangle \, \mathrm{d}\sigma \quad \forall t \in (0, T). \quad (1.28)$$

Moreover, if ϕ complies with the additional continuity assumption (CONT), *then there exists a negligible set $\mathcal{N} \subset (0, T)$ such that*

$$\int_s^t \left(\frac{1}{2}|u'(\sigma)|^2 + \frac{1}{2}|(\partial_\ell \phi(u(\sigma)) - f(\sigma))^0|^2 \right) \mathrm{d}\sigma + \phi(u(t)) \quad \leq \quad \phi(u(s)) + \int_s^t \langle f(\sigma), u'(\sigma) \rangle \, \mathrm{d}\sigma \qquad (1.29)$$
$$\forall t \in (0, T), \; s \in (0, t) \setminus \mathcal{N}.$$

Remark 1.3 (Lyapunov inequality). By (1.29), $\phi \circ u$ satisfies the Lyapunov inequality in the distributional differential form

$$\frac{\mathrm{d}}{\mathrm{d}t} \phi(u(t)) \leq -\frac{1}{2}|u(t)|^2 - \frac{1}{2}\left|(\partial_\ell \phi(u(t)) - f(t))^\circ\right|^2 + \langle f(t), u'(t) \rangle \quad \text{in } \mathscr{D}'(0, T), \qquad (1.30)$$

and there exists a real function of bounded variation $\varphi \geq \phi \circ u$, which coincides with $\phi \circ u$ *a.e.* in $(0, T)$, satisfying (1.30) *a.e.* in $(0, T)$. When $f \equiv 0$, (1.30) reduces to

$$\frac{\mathrm{d}}{\mathrm{d}t} \phi(u(t)) \leq -\frac{1}{2}|u'(t)|^2 - \frac{1}{2}|\partial_\ell^\circ \phi(u(t))|^2 \leq -|u'(t)| \cdot |\partial_\ell^\circ \phi(u(t))| \quad \text{in } \mathscr{D}'(0, T),$$

which is the key point of the metric approach to gradient flows proposed by E. De Giorgi (see the discussion in [3], Chap. 2). In particular, the map $t \mapsto \varphi(t)$ is non-increasing on $(0, T)$. This fact justifies the name of Lyapunov solutions, introduced by S. Luckhaus [26] for a particular model which can be considered in this more general framework, see Example 5 later on.

The case of a limiting subdifferential satisfying the Chain Rule. Let us now assume that ϕ satisfies the Chain Rule condition (CHAIN$_1$), namely

$$\text{if } v \in H^1(0, T; \mathscr{H}), \; \xi \in L^2(0, T; \mathscr{H}), \; \xi \in \partial_\ell \phi(v) \; a.e. \text{ in } (0, T),$$
$$\text{and } \phi \circ v \text{ is } a.e. \text{ equal to a function } \varphi \text{ of bounded variation, then} \qquad \text{(CHAIN}_1\text{)}$$
$$\frac{\mathrm{d}}{\mathrm{d}t} \varphi(t) = \langle \xi(t), v'(t) \rangle \quad \text{for } a.e. \; t \in (0, T).$$

Theorem 2. *Let us suppose that $\phi : \mathscr{H} \to (-\infty, +\infty]$, u_0, f satisfy the assumptions* (COMP) *and* (DATA) *of the compactness Lemma 1.2, so that* $\mathrm{GMM}(\Phi; u_0, f)$ *is not empty. If the Chain Rule condition* (CHAIN$_1$) *and the continuity condition* (CONT) *are satisfied, then every $u \in \mathrm{GMM}(\Phi; u_0, f)$ is a solution in $H^1(0, T; \mathscr{H})$ of* (GF), *it satisfies for almost every $t \in (0, T)$*

$$u'(t) \text{ is the projection of the origin on the affine hull } \mathrm{aff}\big(f(t) - \partial_\ell \phi(u(t))\big) \qquad (1.31)$$

$$\text{and fulfills the minimal section principle} \quad u'(t) = (f(t) - \partial_\ell \phi(u(t)))^\circ, \qquad (1.32)$$

as well as

$$f(t) - u'(t) \text{ belongs to the strong closure of } \partial \phi(u(t)) \text{ in } \mathscr{H} \times \mathscr{H}. \qquad (1.33)$$

Finally, the energy inequality

$$\int_s^t |u'(\sigma)|^2 \, \mathrm{d}\sigma + \phi(u(t)) \leq \phi(u(s)) + \int_s^t \langle f(\sigma), u'(\sigma)\rangle \, \mathrm{d}\sigma \qquad (1.34)$$

holds $\forall t \in (0,T)$, $s \in (0,t) \setminus \mathcal{N}$, $\mathcal{N}$ *being a negligible subset of* $(0,T)$, *and* $\phi \circ u$ *coincides a.e. in* $(0,T)$ *with a function* $\varphi \geq \phi \circ u$ *of bounded variation satisfying*

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = -\big|u'(t)\big|^2 + \langle f(t), u'(t)\rangle \quad \text{a.e. in } (0,T). \qquad (1.35)$$

Remark 1.4 (Affine projection and minimal section). Notice that we have retrieved the minimal section principle (1.32) (in the even stronger formulation (1.31)) in this non convex case as well: even if $f(t) - \partial_\ell \phi(u(t))$ in general is not convex, $u'(t)$ is its *unique element* of minimal norm for *a.e.* $t \in (0,T)$.

In both the previous theorems, the Lyapunov/energy inequalities (1.29) and (1.34) hold almost everywhere, and even though the pointwise differential identity (1.35) holds, we cannot exclude that the distributional derivative of $t \mapsto \phi(u(t))$ is not absolutely continuous with respect to the Lebesgue measure. Thus, the functional ϕ can have essential negative jumps along the solution u. This phenomenon can be circumvented if we reinforce our Chain Rule condition (CHAIN$_1$) a little bit, mimicking the statement (1.11) for convex functionals. In this way, we can also avoid to assume the continuity condition (CONT).

Theorem 3 (Energy solutions). *Let us suppose that* $\phi : \mathscr{H} \to (-\infty, +\infty]$, u_0, f *satisfy the assumptions* (COMP) *and* (DATA) *of the compactness Lemma 1.2, so that* GMM$(\Phi; u_0, f)$ *is not empty.*
If ϕ *satisfies the following stronger Chain rule condition*

$$\begin{aligned} &\text{if } v \in H^1(0,T;\mathscr{H}), \; \xi \in L^2(0,T;\mathscr{H}), \; \xi \in \partial_\ell \phi(v) \text{ a.e. in } (0,T), \\ &\quad \text{and } \phi \circ v \text{ is bounded, then } \phi \circ v \in AC(0,T) \text{ and} \\ &\quad \frac{\mathrm{d}}{\mathrm{d}t}\phi(v(t)) = \langle \xi(t), v'(t)\rangle \quad \text{for a.e. } t \in (0,T), \end{aligned} \qquad (\text{CHAIN}_2)$$

then, every $u \in$ GMM$(\Phi; u_0, f)$ *is a solution in* $H^1(0,T;\mathscr{H})$ *to* (GF), *fulfilling the affine projection property* (1.31), *the minimal selection* (1.32), *the strong closure property* (1.33), *and the* energy identity

$$\int_s^t |u'(\sigma)|^2 \, \mathrm{d}\sigma + \phi(u(t)) = \phi(u(s)) + \int_s^t \langle f(\sigma), u'(\sigma)\rangle \, \mathrm{d}s \qquad \forall s, t: \quad 0 \leq s \leq t \leq T. \qquad (1.36)$$

Finally, if $\{\tau_k\}$ *is a decreasing sequence as in* (1.22), *we also have*

$$U'_{\tau_k} \to u' \quad \text{in } L^2(0,T;\mathscr{H}), \quad \phi(\overline{U}_{\tau_k}(t)) = \phi(u(t)) \quad \forall t \in [0,T], \qquad \text{as } k \uparrow \infty. \qquad (1.37)$$

Remark 1.5 ($W^{1,1}(0,T)$ *versus* $AC(0,T)$). If ϕ complies with a slightly weaker form of (CHAIN$_2$), ensuring that $\phi \circ v \in W^{1,1}(0,T)$ instead of $\phi \circ v \in AC(0,T)$, then the previous Theorem 3 still holds, but the energy identity (1.36) and the convergence (1.37) of $\phi(\overline{U}_{\tau_k}(t))$ hold for every $s, t \in [0,T] \setminus \mathcal{N}$, $\mathcal{N}$ being a negligible subset of $(0,T)$.

Remark 1.6 (The metric theory). In a purely metric setting, the role of a suitable chain rule has also been discussed in [3]: actually, the theory developed therein could be applied to our situation as well, in the case of a *constant* source term f. On the other hand, here we take advantage of the linear structure of $\mathscr{H}$ and we combine some ideas of the *Minimizing Movements* approach with the flexibility of *Young measures*, obtaining more precise information in the limit.

(Dominated) concave perturbations of convex functionals. We conclude this introductory section with a direct application of Theorems 2 and 3, which shows that even in the cases of (dominated) *concave* perturbations of a convex functional we can still prove an existence result for the associated Gradient Flow equation. This result extends the range of application of the theory of curves of maximal slope and, as it will be clear from the following Examples 4 and 5, this class of functionals allows for several interesting applications to phase transition problems.

In fact, we will prove that the class of functionals to which Theorems 2 and 3 apply is closed w.r.t. dominated concave perturbations. Thus, we will focus on functionals ϕ admitting the decomposition

$$\phi = \psi_1 - \psi_2 \text{ in } D(\phi), \quad \text{with}$$
$$\psi_1 : D(\phi) \to \mathbb{R} \text{ l.s.c. and satisfying } (\text{CHAIN}_{1,2}), \tag{1.38}$$
$$\psi_2 : \text{co}\left(D\phi\right) \to \mathbb{R} \text{ convex and l.s.c. in } D(\phi), \quad D(\partial_\ell \psi_1) \subset D(\partial \psi_2),$$

where $\text{co}\left(D\phi\right)$ denotes the convex hull of $D(\phi)$; we shall see that the (corresponding) chain rule $(\text{CHAIN}_{1,2})$ holds in this case as well, provided that the *concave* contribution of the term $-\psi_2$ is somehow controlled.

Theorem 4. *Let $\phi : \mathscr{H} \to (-\infty, +\infty]$ be a functional admitting the decomposition* (1.38) *and satisfying the compactness assumption* (COMP) *of Lemma* 1.2*. If*

$$\forall M \geq 0 \quad \exists \rho < 1, \gamma \geq 0 \quad \text{such that} \quad \sup_{\xi_2 \in \partial \psi_2(u)} |\xi_2| \leq \rho |\partial_\ell^\circ \psi_1(u)| + \gamma$$
$$\text{for every } u \in D(\partial_\ell \psi_1) \text{ with } \max(\phi(u), |u|) \leq M, \tag{1.39}$$

then ϕ satisfies the (corresponding) Chain Rule property $(\text{CHAIN}_{1,2})$.

We postpone the proof of the above results to Section 4.

Remark 1.7 (Kato's condition). Let us point out that condition (1.39) is analogous to a condition, proposed in [12], Chapter II, which ensures that the sum of two maximal monotone operators is still maximal. In the linear framework, this corresponds to the well-known Kato's condition, ensuring the closedness of a perturbation of a closed operator [23], IV, Theorem 1.1.

Remark 1.8 (Equivalent norms). The choice of the norm $|\cdot|_{\mathscr{H}}$ in (1.39) is not essential: Theorem 4 still holds if we consider another equivalent (even non hilbertian) norm in $\mathscr{H}$.

Remark 1.9 (ψ_1, ψ_2 convex). Conditions (1.38) and (1.39) get slightly simpler when ψ_1, ψ_2 are convex: in fact, (1.38) reads

$$\phi = \psi_1 - \psi_2 \text{ in } D(\phi) \quad \text{with}$$
$$\psi_1, \psi_2 : D(\phi) \to \mathbb{R} \text{ l.s.c. and convex}, \quad D(\partial \psi_1) \subset D(\partial \psi_2), \tag{1.38'}$$

whereas in (1.39) the limiting subdifferential of ψ_1 coincides with the Fréchet subdifferential $\partial \psi_1$. In this case, ϕ satisfies the chain rule (CHAIN_2) in the stronger formulation of Remark 1.5.

Plan of the paper. In the next section, we discuss some examples of ordinary differential inclusions and PDE's systems, which motivate our interest for the gradient flow equation (GF). In particular, Example 2, though finite-dimensional, will clearly illustrate the difficulties arising from the lack of convexity and, hence, the role of our variational approach.

The classical variational formulation of the Stefan problem (Ex. 3) is then briefly recalled, mainly to point out the role of the dual Sobolev space $H^{-1}(\Omega)$ and to introduce the basic structure of the functional ϕ, which is common to the other more complicated models, discussed in the last two examples. The fourth one presents new general results for quasistationary phase field models, extending previous contributions of [35] and [43]; the last example deals with the Stefan-Gibbs-Thomson problem and shows how to retrieve S. Luckhaus' theorem as a consequence of the abstract theory. One of the most interesting points, here, is the unifying approach which is provided by the gradient flow viewpoint (see also [39]).

The third section collects two technical tools, which will be essential in the proofs of the main theorems: the fundamental theorem for Young measures in Hilbert space with respect to weak topologies, and their interplay with more refined forms of the Chain Rule properties (CHAIN$_{1,2}$).

Section 4 is devoted to the proof of our main abstract results. It is divided in some parts, each one focusing on a particular aspect which is of independent interest. Refined estimates for the stationary problems (1.20) and for their evolutionary counterparts, the introduction of a new "variational interpolant" of the discrete values, the asymptotic description of the limits and their energy inequalities in terms of Young measures, and the Chain Rule are the main points of the argument.

Finally, in Section 5 we present new general results for diffusion problems with quasistationary non monotone relations, which cover all the examples discussed in Section 2 and can be applied to many different situations.

2. EXAMPLES

2.1. Finite-dimensional examples

Example 1 (Anti-monotone differential inclusions)**.** Let us consider the following Cauchy problem for an anti-monotone differential inclusion in $\mathscr{H} := \mathbb{R}^d$

$$\begin{cases} x'(t) - A(x(t)) \ni 0 & t > 0, \\ \qquad\qquad x(0) = x_0, \end{cases} \tag{2.1}$$

where $A : \mathbb{R}^d \to 2^{\mathbb{R}^d}$ is a cyclically monotone (multivalued) operator such that

$$A(x) \text{ is } \textit{non empty and compact} \text{ for every } x \in \mathbb{R}^d, \text{ and} \tag{2.2}$$

$$\xi_n \in A(x_n), \quad x_n \to x, \xi_n \to \xi \quad \Rightarrow \quad \xi \in A(x), \tag{2.3}$$

i.e., A has closed graph in $\mathbb{R}^d \times \mathbb{R}^d$. This problem was addressed in [10], where the local existence of a solution to (2.1) is obtained by means of an explicit discretization technique; if A is linearly growing at infinity, the solution is also global. Similar ideas have also been applied in a much more general context in [15].

A gradient flow approach. Let us assume this linear growth condition, and let us introduce a proper, l.s.c., and convex function $\psi : \mathbb{R}^d \to \mathbb{R}$ such that (see [36], Th. V.24.8)

$$\emptyset \neq A(x) \subset \partial\psi(x) \quad \forall x \in \mathbb{R}^d. \tag{2.4}$$

In particular, $D(\psi) = \mathbb{R}^d$, ψ is locally Lipschitz, and it grows at most quadratically as $|x| \uparrow +\infty$: thus, Theorems 3 and 4 (applied to the functional $\phi := -\psi$: note that the chain rule (CHAIN$_2$) trivially holds for $-\psi$, since ψ is convex and $\partial_\ell(-\psi) \subset -\partial\psi$), yield the existence of a global solution to the gradient flow equation (GF), which is also a solution to (2.1), since it is not difficult to check that

$$\partial_\ell(-\psi)(x) \subset -A(x) \quad \forall x \in \mathbb{R}^d. \tag{2.5}$$

Indeed, as A is closed in $\mathbb{R}^d \times \mathbb{R}^d$, (2.5) is a consequence of

$$\partial(-\psi)(x) \subset -A(x) \quad \forall x \in \mathbb{R}^d. \tag{2.6}$$

To this aim, we note that a vector ξ belongs to $\partial(-\psi)(x)$ iff ψ is differentiable at x and

$$-\partial\psi(x) = \{-D\psi(x)\} = \{\xi\}.$$

Now, (2.4) yields $A(x) = \{D\psi(x)\}$ for every $x \in D(\partial(-\psi))$, and thus (2.6) is trivially satisfied.

Remark 2.1. The previous argument also shows that $-\partial_\ell(-\psi)$ provides the *minimal closed multivalued map* among those satisfying

$$A(v) \subset \partial\psi(v) \quad \forall v \in \mathscr{H}; \quad D(A) \supset \{v : \psi \text{ is differentiable at } v\}. \tag{2.7}$$

Example 2 (differential equations and non monotone couplings). Let $F, G \in C^1(\mathbb{R}^d)$ be two given functions satisfying

$$\liminf_{|u|\uparrow+\infty} \frac{F(u)}{|u|^2} > -\infty, \quad \liminf_{|\chi|\uparrow+\infty} \frac{G(\chi)}{|\chi|} = +\infty.$$

We consider the following system in the unknowns $u, \chi : [0, +\infty) \to \mathbb{R}^d$, where an ODE is coupled with a (possibly) non monotone relation

$$\begin{cases} u'(t) + \nabla F(u(t)) = \chi(t) + f(t), \\ \quad\quad \nabla G(\chi(t)) = u(t). \end{cases} \tag{2.8}$$

It is interesting to note that (2.8) can be interpreted as the gradient flow equation associated with the functional

$$\phi(u) := \min_{\sigma \in \mathbb{R}^d} \left(F(u) + G(\sigma) - \langle u, \sigma \rangle \right) = F(u) - G^*(u), \tag{2.9}$$

where $G^*(u) := \sup_{\sigma \in \mathbb{R}^d} \langle u, \sigma \rangle - G(\sigma)$ is the (convex) Legendre-Fenchel-Moreau transform of G.

In order to show this fact, we first observe that the nonlinear relation $\nabla G(\chi) = u$ is the Euler equation associated with the minimization problem

$$\text{given } u \in \mathbb{R}^d, \text{ find } \chi \text{ which minimizes } \sigma \mapsto G(\sigma) - \langle \sigma, u \rangle; \tag{2.10}$$

since G has a superlinear growth, the set $M(u)$ of the solutions of (2.10) is not empty, and we can rewrite (2.8) in the more restrictive formulation

$$u'(t) + \nabla F(u(t)) \in M(u(t)) + f(t), \tag{2.11}$$

which is the differential inclusion associated with the operator $u \mapsto \nabla F(u) - M(u)$. It is immediate to check that the Fréchet subdifferential $\partial\phi(u)$ is single-valued in its domain and

$$\partial\phi(u) = \nabla F(u) - M(u) \quad \forall u \in D(\partial\phi).$$

Suppose in fact that $\xi \in \partial\phi(u)$: if $\chi \in M(u)$ and $v \in \mathbb{R}^d$ we have that

$$\begin{aligned} \langle \nabla F(u) - \chi, v - u \rangle &\geq F(v) - F(u) + o(|v - u|) - \langle \chi, v - u \rangle \\ &= F(v) + G(\chi) - \langle \chi, v \rangle - \left(F(u) + G(\chi) - \langle \chi, u \rangle \right) + o(|v - u|) \\ &\geq \phi(v) - \phi(u) + o(|v - u|) \geq \langle \xi, v - u \rangle + o(|v - u|) \end{aligned}$$

as $v \to u$, and therefore $\xi = \nabla F(u) - \chi$. Since the graph of M in $\mathbb{R}^d \times \mathbb{R}^d$ is closed, it is immediate to check that a weaker form of this relation extends to the limiting subdifferential of ϕ, *i.e.*

$$\xi \in \partial_\ell\phi(u) \quad \Rightarrow \quad \xi = \nabla F(u) - \chi \quad \text{for some } \chi \in M(u), \tag{2.12}$$

so that a solution of the gradient flow (GF) for the functional ϕ defined by (2.9) also solves (2.11) and (2.8).

Since $\partial G^*(u)$ is the closed convex hull $\overline{\text{co}}\big(M(u)\big)$ of $M(u)$, we have

$$\partial_\ell\phi(u) \subset \nabla F(u) - M(u) \subset \overline{\text{co}}(\nabla F(u) - M(u)) = \nabla F(u) - \partial G^*(u), \tag{2.13}$$
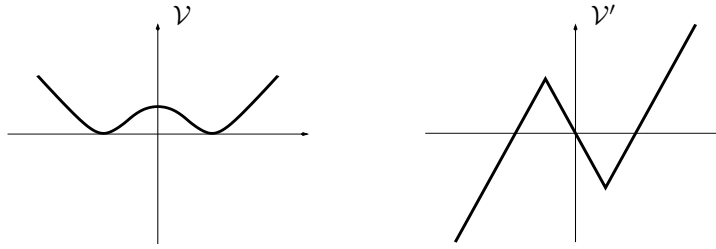
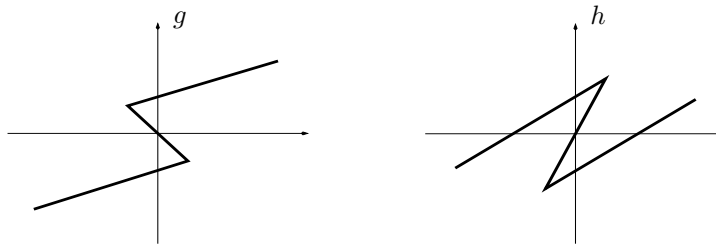FIGURE 3. The potential $\mathcal{V}$ and its derivative.

FIGURE 4. The graphs of the (multivalued) maps $g = (\mathcal{V}' + I)^{-1}$ and $h = I - g$.

which shows that $\partial_\ell \phi(u)$ contains all the extremal points of the convex set $\nabla F(u) - \partial G^*(u)$. On the other hand, since ϕ is given by the difference of the C^1 function F and the convex function G^*, with

$$\chi \in M(u) \quad \Rightarrow \quad \chi \in \partial G^*(u),$$

it is easy to check that ϕ satisfies the chain rule condition (CHAIN$_2$): therefore an application of Theorem 3 provides the existence of an energy solution of (2.11) and there is no need to convexifying the evolution operator.

We can check directly in the following particularly simple one-dimensional case that (2.8) corresponds to a non monotone differential inclusion, whose convexification would introduce spurious solutions which do not solve the original problem (2.8). For, let $F(u) := \frac{1}{2}u^2$, $u \in \mathbb{R}$ and $G(\chi) := \frac{1}{2}\chi^2 + \mathcal{V}(\chi)$, where $\mathcal{V} : \mathbb{R} \to \mathbb{R}$ is the (piecewise quadratic) double well potential

$$\mathcal{V}(\chi) := \begin{cases} (\chi+1)^2 & \chi < -\frac{1}{2}, \\ -\chi^2 + \frac{1}{2} & |\chi| \leq \frac{1}{2}, \\ (\chi-1)^2 & \chi > \frac{1}{2}, \end{cases} \quad \text{with derivative} \quad \mathcal{V}'(\chi) = \begin{cases} 2(\chi+1) & \chi < -\frac{1}{2}, \\ -2\chi & |\chi| < \frac{1}{2}, \\ 2(\chi-1) & \chi > \frac{1}{2} \end{cases} \quad (2.14)$$

(see Fig. 3). In this setting, (2.8) reads

$$\begin{cases} u'(t) + u(t) = \chi(t) + f(t), \\ \mathcal{V}'(\chi(t)) + \chi(t) = u(t), \end{cases} \quad (2.15)$$

which is a drastic one-dimensional caricature of the PDE model of Example 4 later on. Notice that the map $\chi \mapsto G'(\chi) = \mathcal{V}'(\chi) + \chi$ is not monotone. Thus, the inverse $g := (\mathcal{V}' + I)^{-1}$ is a multivalued function, and (2.15) is equivalent to the differential inclusion

$$u'(t) + u(t) - g(u(t)) \ni f(t), \quad (2.16)$$

associated with the *non monotone* multivalued map $u \mapsto h(u) := u - g(u)$ (see Fig. 4).
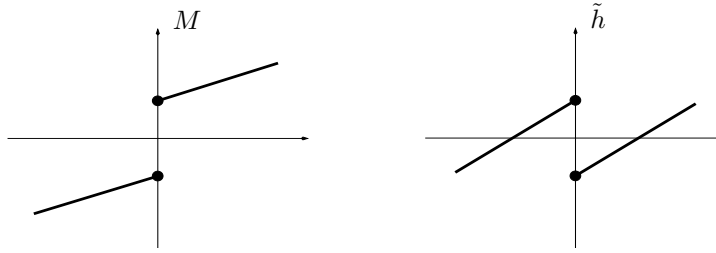
FIGURE 5. The graphs of $M(u)$ and $\tilde{h}(u) = u - M(u)$.

In this case, the functional ϕ is given by

$$\phi(u) = \min_{\sigma \in \mathbb{R}} \left(\frac{1}{2}u^2 + \frac{1}{2}\sigma^2 + \mathcal{V}(\sigma) - \sigma u \right) = \min_{\sigma \in \mathbb{R}} \left(\frac{1}{2}|u - \sigma|^2 + \mathcal{V}(\sigma) \right), \tag{2.17}$$

which has the analytical expression of (1.24) and provides a variational selection $M(u)$ of χ in $g(u)$, given by the minimization problem

$$\text{find } \chi \in \mathbb{R} \text{ which attains the minimum in (2.17),} \tag{2.18}$$

where u is considered as a given parameter.

Therefore, coupling (2.15) with (2.18) yields the differential inclusion

$$u'(t) + u(t) - M(u(t)) = u'(t) - \tilde{h}(u) \ni f(t), \tag{2.19}$$

(see Fig. 5), whose solutions also solve the previous (2.16), since the variational principle (2.18) has in fact selected suitable branches of g and h. Instead, the convexification of the values of $\tilde{h}$ (by adding the vertical segment $[-2/3, 2/3]$ at $u = 0$), would destroy this property.

Remark 2.2. Systems like (2.8) arise naturally from the *formal* limit as $\varepsilon \downarrow 0$ of a time relaxation in the second equation, *e.g.* $\varepsilon\chi'(t) + \nabla G(\chi(t)) = u(t)$. In the non monotone case, one should expect that hysteresis occurs in the limit (see *e.g.* the discussion of [46], Chap. X, and the approaches proposed in [30, 48]). Here, we are neglecting these non local effects: our variational selection principle (2.10) always forces χ to jump towards an absolute minimum of the map $v \mapsto G(v) - \langle u, v \rangle$.

The next examples, arising from some models for phase transitions, exhibit a similar structure in infinite-dimensional spaces.

2.2. The gradient flow structure of some quasistationary models for phase transitions

Before developing the main applications of our results to quasistationary phase field models, let us first examine the well-known example of the H^{-1} formulation of the classical Stefan problem, see [11, 14].

Example 3 (The Stefan problem). We consider the boundary value problem for the evolution PDE

$$\begin{cases} \partial_t u - \Delta\beta(u) = f & \text{in } \Omega \times (0, T), \\ \beta(u) = 0 & \text{on } \partial\Omega \times (0, T), \\ u(x, 0) = u_0(x) & \text{in } \Omega, \end{cases} \tag{2.20}$$

where Ω is an open bounded and connected subset of $\mathbb{R}^m$, $f : \Omega \times (0, T) \to \mathbb{R}$, and

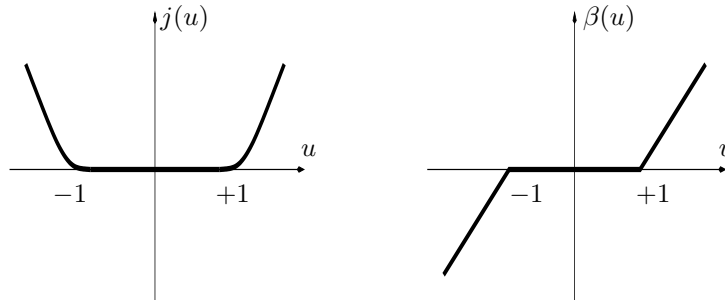$$\beta(u) := (u - 1)^+ - (u + 1)^-, \quad u \in \mathbb{R}.$$

FIGURE 6. The potential j and its derivative β

By adopting the usual convention of identifying a real function $(x,t) \mapsto v(x,t)$ defined in $\Omega \times (0,T)$ with the time dependent function $t \mapsto v_t = v(\cdot,t)$ with values in some function space defined on Ω, (2.20) can be interpreted as the gradient flow in the space $\mathscr{H} := H^{-1}(\Omega)$ of the convex functional

$$\phi(u) := \begin{cases} \int_\Omega j(u(x)) \mathrm{d}x & \text{if } u \in L^2(\Omega), \\ +\infty & \text{if } u \in H^{-1}(\Omega) \setminus L^2(\Omega), \end{cases} \tag{2.21}$$

where j is the primitive of β, i.e.,

$$j(u) := \frac{1}{2}((|u| - 1)^+)^2, \quad u \in \mathbb{R},$$

(see Fig. 6). In order to understand the role of the space $H^{-1}(\Omega)$ and to highlight the gradient flow structure of (2.20), we first consider the realization of $-\Delta$ with homogeneous Dirichlet boundary conditions as unbounded operator in $L^2(\Omega)$ and, by inverting it, we rewrite the Stefan equation as

$$(-\Delta)^{-1} \partial_t u + \frac{\delta \phi}{\delta u} = 0 \quad \text{in } \Omega \times (0,T),$$

where $\frac{\delta \phi}{\delta u} = \beta(u)$ is the first variation of the integral functional ϕ. It is then natural to introduce the scalar product

$$\langle u,v \rangle := \left((-\Delta)^{-1} u, v\right)_{L^2(\Omega)} = \left(u, (-\Delta)^{-1} v\right)_{L^2(\Omega)},$$
$$\text{with the induced norm} \quad |u|^2 := \left((-\Delta)^{-1} u, u\right)_{L^2(\Omega)}; \tag{2.22}$$

it is well-known that we can identify the completion of $L^2(\Omega)$ w.r.t. this norm with the space $\mathscr{H} = H^{-1}(\Omega)$, and the operator $-\Delta$ can be extended by continuity to an unbounded operator $A : D(A) \subset \mathscr{H} \to \mathscr{H}$ with $D(A) := H_0^1(\Omega)$.

It is useful to rephrase (2.20) by introducing the function

$$\chi := u - \beta(u);$$

in the applications, u is the internal energy of a physical system (occupying the region Ω), undergoing a solid-liquid phase transition in the time interval $(0,T)$, while χ is an order parameter, yielding the local proportion of the two phases (here we are normalizing all the relevant physical constants to 1).

By inverting the Laplace operator in this setting, (2.20) becomes

$$A^{-1} \partial_t u + u = \chi + A^{-1} f \qquad\qquad \text{in } \Omega \times (0,T), \tag{2.23}$$

$$\chi + \text{sign}^{-1}(\chi) \ni u \qquad\qquad \text{in } \Omega \times (0,T), \tag{2.24}$$

whose formal structure looks like (2.8); here, the (multivalued) map sign^{-1} is defined as

$$\mathrm{sign}^{-1}(x) := \begin{cases} (-\infty, 0] & \text{if } x = -1, \\ 0 & \text{if } -1 < x < 1, \\ [0, +\infty) & \text{if } x = 1, \end{cases} \quad x \in [-1, 1],$$

and it is the (convex) subdifferential of the indicator function $I_{[-1,1]}$,

$$I_{[-1,1]}(x) := \begin{cases} 0 & \text{if } -1 \le x \le 1, \\ +\infty & \text{otherwise.} \end{cases}$$

We may easily check that the functional ϕ (2.21) admits the variational representation

$$\phi(u) := \min_{\sigma \in L^2(\Omega)} \mathscr{F}(u, \sigma), \quad \mathscr{F}(u, \sigma) := \int_\Omega \left(\frac{1}{2} |u - \sigma|^2 + I_{[-1,1]}(\sigma) \right) \mathrm{d}x. \tag{2.25}$$

In this case, the map $\chi \mapsto \chi + \mathrm{sign}^{-1}(\chi)$ is monotone and the minimization problem (2.25) is convex in σ, so that there is no difference between (2.24) and $\chi \in \mathrm{argmin}_{L^2(\Omega)} \mathscr{F}(u, \cdot)$. From the mathematical point of view, the interest of the less direct representation (2.23)–(2.24) of (2.20) is that several quasistationary phase field models are obtained simply by replacing the indicator function $I_{[-1,1]}$ in (2.25) by more complex non convex functionals, which in some sense force χ to stay near the extreme points -1 and 1 of the interval.

Example 4 (The quasistationary phase field model). An alternative model for solid-liquid transitions is the *quasistationary phase field* system:

$$\partial_t u - \Delta(u - \chi) = f \qquad\qquad \text{in } \Omega \times (0, T), \tag{2.26}$$

$$-\varepsilon \Delta \chi + \frac{1}{\varepsilon}(\chi^3 - \chi) = u - \chi \qquad\qquad \text{in } \Omega \times (0, T). \tag{2.27}$$

(2.27) is the Euler-Lagrange equation for the minima of the Landau-Ginzburg free energy potential

$$\mathscr{F}_\varepsilon(u, \chi) := \int_\Omega \left(\frac{\varepsilon}{2} |\nabla \chi|^2 + \frac{1}{4\varepsilon}(\chi^2 - 1)^2 + \frac{1}{2}|\chi - u|^2 \right) \mathrm{d}x \tag{2.28}$$

at constant u.

The existence of solutions for the system (2.26)–(2.27), supplemented with the initial and Dirichlet/Neumann boundary conditions ($\boldsymbol{n}$ denotes the outer unit normal to $\partial\Omega$)

$$u(x, 0) = u_0(x) \quad \text{in } \Omega; \qquad u - \chi = 0, \ \partial_{\boldsymbol{n}} \chi = 0 \quad \text{on } \partial\Omega \times (0, T), \tag{2.29}$$

was proved in [35] for the space dimensions $m \le 3$, by means of a compactness method and a unique continuation result, thus *heavily relying on the precise form of the equation* (2.27), in particular on the presence of the Laplacian and of the *analytic potential*

$$W(\chi) := \frac{1}{4\varepsilon}(\chi^2 - 1)^2, \quad W'(\chi) = \frac{1}{\varepsilon}(\chi^3 - \chi).$$

In [43], Schätzle proved an existence result for the initial-boundary value problem obtained by supplementing (2.26)–(2.27) with the initial and homogeneous Neumann conditions

$$u(x, 0) = u_0(x) \quad \text{in } \Omega, \quad \partial_{\boldsymbol{n}}(u - \chi) = \partial_{\boldsymbol{n}} \chi = 0 \quad \text{on } \partial\Omega \times (0, T). \tag{2.30}$$

Here the approach of [35] is no more possible and it is replaced by refined spectral analysis arguments, still based on the specific form of W (the proof involves its analyticity), of the elliptic operator in (2.27), and on the dimension $m \leq 3$.

In analogy with the discussion developed in Example 3 for the Stefan Problem, we will adopt a gradient flow approach to (2.26)–(2.27), also obtaining an existence result for more general differential problems, both with Dirichlet (2.29) or Neumann (2.30) boundary conditions in arbitrary space dimension m.

The gradient flow approach. We consider the system

$$\partial_t u - \operatorname{div} \mathsf{A}_1 \nabla(u - \chi) = f \qquad\qquad \text{in } \Omega \times (0, T), \qquad\qquad (2.31)$$

$$-\varepsilon \operatorname{div} \mathsf{A}_2 \nabla\chi + \frac{1}{\varepsilon} W'(\chi) = u - \chi \qquad\qquad \text{in } \Omega \times (0, T), \qquad\qquad (2.32)$$

where $\mathsf{A}_1, \mathsf{A}_2 : \Omega \to \mathbb{M}^{m \times m}$ are symmetric matrices, with measurable coefficients, satisfying the usual uniform ellipticity condition

$$\mathsf{a}^{-1} \geq \mathsf{A}_{1,2}(\mathsf{x})\eta \cdot \eta \geq \mathsf{a} > 0 \quad \forall \mathsf{x} \in \Omega, \ \eta \in \mathbb{R}^{\mathsf{m}}, \ |\eta| = 1, \qquad\qquad (2.33)$$

and W is *an arbitrary C^1 real function with superlinear growth*; extended real valued semiconvex functions could also be considered (thus allowing convex constraints on χ), simply replacing the equation (2.32) by the corresponding differential inclusion involving ∂W instead of W'.

In accordance with the analysis developed in Examples 2 and 3, in the case of Dirichlet boundary conditions (2.29) we will endow the Hilbert space $\mathscr{H} := H^{-1}(\Omega)$ with the scalar product induced by the differential operator $A_1 := -\operatorname{div}(\mathsf{A}_1 \nabla \cdot)$ as in (2.22), and we will consider the functionals

$$\mathscr{F}_\varepsilon(u, \chi) := \int_\Omega \left(\frac{1}{2}|u - \chi|^2 + \frac{\varepsilon}{2}\mathsf{A}_2(x)\nabla\chi \cdot \nabla\chi + \frac{1}{\varepsilon}W(\chi) \right) \mathrm{d}x, \qquad\qquad (2.34)$$

$$\phi_\varepsilon(u) := \min_{\sigma \in H^1(\Omega)} \mathscr{F}_\varepsilon(u, \sigma), \quad M_\varepsilon(u) := \left\{ \chi \in H^1(\Omega) : \mathscr{F}_\varepsilon(u, \chi) = \phi_\varepsilon(u) \right\}. \qquad\qquad (2.35)$$

We will investigate the gradient flow equation (GF) for the functional ϕ_ε, with the source term f. The following result is a particular case of a general existence and convergence result (Th. 5.8 in Sect. 5) for quasistationary phase field models.

Theorem 2.3. *Given $u_0 \in L^2(\Omega)$ and $f \in L^2(0, T; H^{-1}(\Omega))$, there exist*

$$u \in H^1(0, T; H^{-1}(\Omega)) \cap L^\infty(0, T; L^2(\Omega)), \quad \chi \in L^\infty(0, T; H^1(\Omega)),$$
$$\vartheta := u - \chi \in L^2(0, T; H_0^1(\Omega)), \qquad\qquad\qquad (2.36)$$

such that the pair (u, χ) fulfills (2.31), coupled with (here $\chi_t := \chi(\cdot, t), u_t := u(\cdot, t)$)

$$\chi_t \in M_\varepsilon(u_t) = \operatorname*{argmin}_{\sigma \in H^1(\Omega)} \mathscr{F}_\varepsilon(u_t, \sigma) \quad \text{a.e. in } (0, T), \qquad\qquad (2.37)$$

(yielding in particular (2.32)), and the initial and boundary conditions (2.29).

In fact, any generalized Minimizing Movement in $H^{-1}(\Omega)$ is an Energy solution of the gradient flow (GF) generated by ϕ_ε, u_0, f (cf. Th. 3) and solves (2.31), (2.37), (2.29).

Finally, a completely analogous existence result holds in the case of Neumann boundary conditions and $f \in L^2(0, T; L^2(\Omega))$, by simply replacing $H^{-1}(\Omega)$ and $H_0^1(\Omega)$ with $\left(H^1(\Omega)\right)'$ and $H^1(\Omega)$ in (2.36).

Remark 2.4. [43] adopts a discretization scheme which is different from the Minimizing Movement one (1.19, 1.20) induced by the functional (2.35) in $\left(H^1(\Omega)\right)'$. Correspondingly, the solutions χ_t of [43] are

only μ-minimizers of (2.37) for an (arbitrarily small) constant $\mu > 0$ and they satisfy the weak Lyapunov-type condition

$$\phi_\varepsilon(u_s) + \int_0^s \int_\Omega |\nabla\vartheta|^2 \, \mathrm{d}x \, \mathrm{d}t \le \phi_\varepsilon(u_0) + \int_0^s \int_\Omega f\vartheta \, \mathrm{d}x \, \mathrm{d}t, \quad \text{for } a.e. \ s \in (0,T).$$

Besides (2.37), Theorem 2.3 (*cf.* (5.46)) shows that every Minimizing Movement solution u is in fact an energy solution and therefore satisfies the stronger energy identity

$$\phi_\varepsilon(u_s) + \int_0^s \int_\Omega |\nabla\vartheta|^2 \, \mathrm{d}x \, \mathrm{d}t = \phi_\varepsilon(u_0) + \int_0^s \int_\Omega f\vartheta \, \mathrm{d}x \, \mathrm{d}t, \quad \forall s \in [0,T]. \tag{2.38}$$

In the particular case $f \equiv 0$, (2.38) shows that the potential ϕ_ε is non increasing along the solution u.

Example 5 (the Stefan-Gibbs-Thomson problem). If we formally pass to the limit as $\varepsilon \downarrow 0$ in (2.26, 2.37), we get a system where equation (2.26)

$$\partial_t u - \Delta(u - \chi) = f \quad \text{in } \Omega \times (0,T), \tag{2.39}$$

is coupled with the minimum condition for $\chi_t = \chi(\cdot, t)$

$$\chi_t \in \operatorname*{argmin}_\sigma \mathscr{F}_0(u(t), \sigma) = M_0(u(t)) \quad \text{in } (0,T), \tag{2.40}$$

for the functional

$$\mathscr{F}_0(u, \chi) := \int_\Omega \left(\frac{1}{2}|u - \chi|^2 + I_{\{-1,1\}}(\chi) \right) \mathrm{d}x + \alpha \int_\Omega |D\chi|.$$

Here, $\alpha = \int_{-1}^1 \sqrt{2\varepsilon W(\rho)} \, \mathrm{d}\rho = \frac{2\sqrt{2}}{3}$, $I_{\{-1,1\}}$ is the indicator function of the *non convex* set $\{-1,1\}$, and $D\chi$ is the distributional gradient of χ: it is a Borel vector measure, whose total variation $|D\chi|$ is defined on every open set $A \subset \Omega$ by

$$|D\chi|(A) = \int_A |D\chi| := \sup \left\{ \int_A \chi \operatorname{div}\boldsymbol{\zeta} \, \mathrm{d}x : \boldsymbol{\zeta} \in C_0^1(A; \mathbb{R}^m), \quad \sup_\Omega |\boldsymbol{\zeta}| \le 1 \right\}. \tag{2.41}$$

$\mathscr{F}_0(u, \chi)$ is the functional obtained by taking the Γ-limit of $\mathscr{F}_\varepsilon$, for u fixed in $L^2(\Omega)$ [31, 32] (see also [9] in the case (2.34) of space-dependent coefficients). Denoting by $BV(\Omega; \{-1,1\})$ the subset of all the functions of bounded variation taking their values in the set $\{-1,1\}$, (2.40) may be also rephrased for every time $t \in (0,T)$ as

$$\chi_t \in BV(\Omega; \{-1,1\}), \quad \text{and}$$
$$\int_\Omega |D\chi_t| - \int_\Omega u_t\chi_t \, \mathrm{d}x \le \int_\Omega |D\sigma| - \int_\Omega u_t\sigma \, \mathrm{d}x \quad \forall \sigma \in BV(\Omega; \{-1,1\}). \tag{2.42}$$

At each time $t \in (0,T)$ we can associate with the characteristic function χ_t the phases $E_t^\pm$ and their common essential boundary S_t (see [2], Def. 3.60):

$$E_t^\pm := \left\{ x \in \Omega : \chi(x,t) = \pm 1 \right\}, \quad S_t = \partial^* E_t^+ = \partial^* E_t^- ; \tag{2.43}$$

The De Giorgi-Federer Theorems (see [2], Ths. 3.59, 3.61), show that S_t is countably $(m-1)$-rectifiable and

$$\frac{1}{2}|D\chi_t| = \mathcal{H}^{m-1}\lfloor S_t, \quad D\chi_t = \boldsymbol{\nu}_t \cdot |D\chi_t| = \boldsymbol{\nu}_t D\chi_t = 2\boldsymbol{\nu}_t \cdot \mathcal{H}^{m-1}\lfloor S_t, \tag{2.44}$$

where the (Borel) vector field $\boldsymbol{\nu}_t : S_t \to \mathscr{S}^{m-1}$ is the inner measure theoretic normal to E_t^+.

As shown by [35, 43] following the argument of [28], the limit of (2.27) yields the Gibbs-Thomson condition

$$\boldsymbol{H} = \vartheta_t \boldsymbol{\nu}_t \quad \text{on } S_t$$

for the mean curvature vector $\boldsymbol{H}$ at the evolving phase interface S_t: its weak formulation reads [27, 43]

$$\alpha \int_\Omega \left(\text{div } \boldsymbol{\zeta} - \boldsymbol{\nu}_t^T D\boldsymbol{\zeta} \, \boldsymbol{\nu}_t \right) \text{d}|D\chi_t| = \int_\Omega \text{div}(\vartheta_t \boldsymbol{\zeta})\chi_t \, \text{d}x, \quad \vartheta_t := u_t - \chi_t,$$
$$\forall \boldsymbol{\zeta} \in C^2(\overline{\Omega}; \mathbb{R}^m), \quad \boldsymbol{\zeta} \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega. \tag{2.45}$$

We refer to (2.39, 2.40, 2.45) as the Stefan-Gibbs-Thomson problem; its formulation as a gradient flow was suggested by A. Visintin and existence of (Lyapunov) solutions has been first proved by S. Luckhaus in the pioneering paper [26] (see also [27]) and then further investigated in [47], Chap. VIII. The proof is based on a time discretization technique (whose link with the present Minimizing Movement scheme (1.19, 1.20) will be discussed in the next Rem. 2.7) and a clever passage to the limit, which relies on careful capacity type estimates for Sobolev functions defined on Ω; abstract versions of this argument have been further proposed and investigated in [35, 38, 42]. The convergence of the quasistationary phase field model (2.26, 2.27) to the Stefan-Gibbs-Thomson problem has been proved in [35, 43].

As in the previous example, we will obtain a solution of the system (2.26, 2.40) supplemented with the conditions

$$u(x, 0) = u_0(x) \quad \text{in } \Omega, \quad u - \chi = 0 \quad \text{on } \partial\Omega \times (0, T), \tag{2.46}$$

by solving the gradient flow equation (GF) in the Hilbert space $\mathscr{H} := H^{-1}(\Omega)$ for the functional

$$\phi_0(u) := \inf_{\sigma \in BV(\Omega)} \mathscr{F}_0(u, \sigma) \tag{2.47}$$

with domain $D(\phi_0) = L^2(\Omega) \subset \mathscr{H}$. By adopting this gradient flow approach, we are able to retrieve Luckhaus's existence result [26], and we get some insight into the Lyapunov inequality satisfied by u.

Theorem 2.5. *Let Ω be a C^1 connected open set. For every $u_0 \in L^2(\Omega)$ and $f \in L^2(0, T; H^{-1}(\Omega))$ the gradient flow equation* (GF) *for the functional* (2.47) *in $\mathscr{H} := H^{-1}(\Omega)$ has a (Lyapunov, according to Theorem 1) solution $u \in H^1(0, T; H^{-1}(\Omega))$ which solves the Stefan-Gibbs-Thomson problem, i.e., there exists (a measurable selection) $\chi_t = \chi(\cdot, t) \in M_0(u_t)$, $t \in (0, T)$, such that*

$$\vartheta = u - \chi \in L^2(0, T; H_0^1(\Omega) \cap L^\infty(0, T; L^2(\Omega)), \tag{2.48}$$

and the pair (u, χ) solves the initial-boundary value problem (2.39, 2.40, 2.45, 2.46).

Moreover, u and χ fulfill for a.e. $s, t \in (0, T)$, $s \le t$, the Lyapunov inequality

$$\int_s^t \int_\Omega |\nabla\vartheta(x, r)|^2 \, \text{d}x \, \text{d}r + \mathscr{F}(u_t, \chi_t) \le \mathscr{F}(u_s, \chi_s) + \int_s^t \langle f_r, \vartheta_r \rangle \, \text{d}r. \tag{2.49}$$

We postpone the proof of this result to Section 5: here we only recall the crucial link with the abstract theory of the previous section: if $u \in D(\partial_\ell\phi_0) \subset L^2(\Omega)$ then the limiting subdifferential $\partial_\ell\phi(u)$ contains a unique element ξ and it satisfies

$$\xi = -\Delta\vartheta, \quad \vartheta = u - \chi \in H_0^1(\Omega), \quad \chi \in M_0(u), \quad (2.45) \text{ holds.} \tag{2.50}$$

Remark 2.6. As detailed in [26], p. 106, since for $m \le 3$ the Sobolev imbedding theorem yields $u_t = \vartheta_t + \chi_t \in L^p(\Omega)$ for a.e. $t \in (0, T)$ with $p = 2m/(m-2) > m$, the minimality condition (2.42) satisfied by χ_t and the regularity results for minimal surfaces yield that the interface S_t can be locally represented as the graph of a

$C^{1,1/4}$ function g. Namely, after a change of coordinates at each point $x \in \Omega$ there exists an open ball $B_\rho(x)$ such that

$$E_t^- \cap B_\rho(x) = \left\{ (x', x_m) \in \mathbb{R}^{m-1} \times \mathbb{R} : x_m > g(x') \right\} \cap B_\rho(x). \tag{2.51}$$

In this new reference system, (2.45) becomes (see [2], 7.33, [27], p. 4)

$$-\alpha \operatorname{div} \left(\frac{\nabla g(x')}{\sqrt{1 + |\nabla g(x')|^2}} \right) = \theta(x', g(x')), \tag{2.52}$$

which is the original condition of [26], p. 102.

Remark 2.7. According to the Minimizing Movement approach (1.19, 1.20), the solution u of the Stefan-Gibbs-Thomson problem is approximated by the sequence $\{U_\tau^n\}_{n=0}^N$, $\tau = T/N$, whose elements recursively minimize

$$\Phi_\tau^n(U) := \frac{1}{2\tau} \|U - U_\tau^{n-1}\|_{H^{-1}(\Omega)}^2 + \phi_0(U) \tag{2.53}$$

(for the sake of simplicity, here we consider the case $f \equiv 0$ and we omit to indicate the explicit dependence on $\tau, U_\tau^{n-1}, X_\tau^{n-1}$ in the various functionals). Taking into account (2.47), this problem is equivalent finding a couple (U_τ^n, X_τ^n) which minimizes

$$\tilde{\Phi}_\tau^n(U, X) := \frac{1}{2\tau} \|U - U_\tau^{n-1}\|_{H^{-1}(\Omega)}^2 + \mathscr{F}_0(U, X). \tag{2.54}$$

The discretization algorithm introduced by Luckhaus in [27] is equivalent to (2.54), but it gives a distinguished role to the variable X instead of U. More precisely, by introducing the new variable $\Theta := U - X$ and writing $\tilde{\Phi}_\tau^n$ in terms of the couple (Θ, X) as

$$\Psi_\tau^n(\Theta, X) := \frac{\|\Theta - \Theta_\tau^{n-1} + X - X_\tau^{n-1}\|_{H^{-1}(\Omega)}^2}{2\tau} + \alpha \int_\Omega |DX| + \int_\Omega \left(\frac{1}{2} |\Theta|^2 + I_{\{-1,1\}}(X) \right) dx,$$

the minimum problem for Ψ_τ^n can be split into two iterated minima:

$$\min_{(\Theta, X)} \Psi_\tau^n(\Theta, X) = \min_X \psi_\tau^n(X), \quad \text{where} \quad \psi_\tau^n(X) := \min_\Theta \Psi_\tau^n(\Theta, X).$$

It is easy to check that for a fixed X the minimum $\psi_\tau^n(X)$ of $\Psi_\tau^n(X, \cdot)$ is attained at a unique $\Theta = \Theta(X) \in H_0^1(\Omega)$, which is determined by the Euler equation

$$\Theta - \tau \Delta \Theta = \Theta_\tau^{n-1} + X_\tau^{n-1} - X, \quad \Theta(X) = K_\tau(\Theta_\tau^{n-1} + X_\tau^{n-1} - X), \tag{2.55}$$

where $K_\tau := (I - \tau \Delta)^{-1}$. We thus find

$$\frac{\|\Theta - \Theta_\tau^{n-1} + X - X_\tau^{n-1}\|_{H^{-1}(\Omega)}^2}{2\tau} + \frac{1}{2} \int_\Omega |\Theta|^2 \, dx = \frac{1}{2} \left(\tau \|\Delta \Theta\|_{H^{-1}(\Omega)}^2 + \|\Theta\|_{L^2(\Omega)}^2 \right)$$

$$= \frac{1}{2} \left(\tau \|\Theta\|_{H_0^1(\Omega)}^2 + \|\Theta\|_{L^2(\Omega)}^2 \right) = \frac{1}{2} {}_{H^{-1}(\Omega)} \langle \Theta - \tau \Delta \Theta, \Theta \rangle_{H_0^1(\Omega)}$$

$$= \frac{1}{2} \int_\Omega \left(\Theta_\tau^{n-1} + X_\tau^{n-1} - X \right) K_\tau \left(\Theta_\tau^{n-1} + X_\tau^{n-1} - X \right) dx$$

$$= \frac{1}{2} \int_\Omega \left(X - X_\tau^{n-1} \right) K_\tau \left(X - X_\tau^{n-1} \right) dx - \int_\Omega X K_\tau \left(\Theta_\tau^{n-1} \right) dx + C_\tau^{n-1}$$

where $C_\tau^{n-1} = C(\Theta_\tau^{n-1}, X_\tau^{n-1})$ is independent of X. It follows that

$$\psi_\tau^n(X) = \alpha \int_\Omega |D\chi| + \int_\Omega I_{\{-1,1\}}(X)\mathrm{d}x + \frac{1}{2}\int_\Omega \left(X - X_\tau^{n-1}\right)K_\tau\left(X - X_\tau^{n-1}\right)\mathrm{d}x - \int_\Omega X\, K_\tau\left(\Theta_\tau^{n-1}\right)\mathrm{d}x + C_\tau^{n-1}$$

and the algorithm

$$given \quad (\Theta_\tau^{n-1}, X_\tau^{n-1}) \quad find \quad X_\tau^n \quad which\ minimizes\ \psi_\tau^n, \quad \Theta_\tau^n := K_\tau(X_\tau^{n-1} + \Theta_\tau^{n-1} - X_\tau^n),$$

coincides with the algorithm $(2h)$ introduced, in [27], p. 12.

3. HILBERT SPACE VALUED YOUNG MEASURES AND THE CHAIN RULE

In this section we mainly discuss two technical tools, which we will extensively use in the sequel.

The first one is concerned with parametrized measures with values in a Hilbert space: we shall deduce from this well-established theory a version of the fundamental theorem of Young measure theory (see [5], Th. 1, and also [8]), in the context of weak topologies.

Then, by means of a measurable selection result, we will study the relations between parametrized measures and the Chain Rule conditions we have presented in the Introduction.

3.1. **Parametrized young measures**

First of all, we fix some notation and we recall the notion of (time dependent) *parametrized measures.*

Notation. Let E be a separable metric space: we denote by $\mathcal{B}(E)$ its Borel σ-algebra, while $\mathcal{L}$ is the σ-algebra of the Lebesgue measurable subsets of $(0,T)$, and $\mathcal{L}\otimes\mathcal{B}(E)$ is the product σ-algebra on $(0,T)\times E$. A $\mathcal{L}\otimes\mathcal{B}(E)$-measurable function $h : (0,T)\times E \to (-\infty, +\infty]$ is a *normal integrand* if

$$v \mapsto h_t(v) := h(t,v) \quad \text{is l.s.c. on } E \text{ for } a.e.\ t \in (0,T). \tag{3.1}$$

When $E = \mathscr{H}$ is a Hilbert space, we say that a $\mathcal{L}\otimes\mathcal{B}(\mathscr{H})$-measurable functional $h : (0,T)\times\mathscr{H} \to (-\infty, +\infty]$ is a *weakly normal integrand* if

$$v \mapsto h_t(v) = h(t,v) \text{ is sequentially weakly l.s.c. for } a.e.\ t \in (0,T). \tag{3.2}$$

Definition 3.1 ((Time dependent) parametrized measures)**.** A *parametrized measure* in the separable metric space E is a family $\boldsymbol{\nu} := \{\nu_t\}_{t\in(0,T)}$ of Borel probability measures on E such that

$$t \in (0,T) \mapsto \nu_t(B) \quad \text{is} \quad \mathcal{L}\text{-measurable} \quad \forall B \in \mathscr{B}(E). \tag{3.3}$$

We denote by $\mathscr{Y}(0,T;E)$ the set of all parametrized measures.

A version of Fubini's theorem [21], p. 20-II, states that for every parametrized measure $\boldsymbol{\nu} = \{\nu_t\}_{t\in(0,T)}$, there exists a unique measure ν on $\mathcal{L}\otimes\mathscr{B}(E)$ defined by

$$\nu(I\times A) = \int_I \nu_t(A)\,\mathrm{d}t \quad \forall I \in \mathcal{L},\ A \in \mathscr{B}(E).$$

Moreover, for every $\mathcal{L}\otimes\mathscr{B}(E)$-measurable function $h : (0,T)\times E \to [0, +\infty]$, the function

$$t \in (0,T) \mapsto \int_E h(t,\xi)\,\mathrm{d}\nu_t(\xi) \quad \text{is } \mathcal{L}\text{-measurable,}$$

and the Fubini's integral representation formula holds:

$$\int_{(0,T)\times E} h(t,\xi)\,\mathrm{d}\nu(t,\xi) = \int_0^T \left(\int_E h(t,\xi)\mathrm{d}\nu_t(\xi)\right)\mathrm{d}t. \tag{3.4}$$

Note that (3.4) holds even for ν-integrable real valued functions.

If ν is concentrated on the graph of a measurable function $u : (0,T) \to E$, then $\nu_t = \delta_{u(t)}$ for *a.e.* $t \in (0,T)$, where $\delta_{u(t)}$ denotes the Dirac's measure carried by $\{u(t)\}$. In this case, by (3.4)

$$\int_{(0,T)\times E} h(t,\xi)\,\mathrm{d}\nu(t,\xi) = \int_0^T h(t,u(t))\,\mathrm{d}t$$

for every $\mathcal{L} \otimes \mathscr{B}(E)$-measurable and nonnegative function h.

The following theorem is a direct consequence of [5], Theorem 1 (see also [6], Th. 2.2, [7], Th. 4.2, [45], Th. 16).

Theorem 3.2 (The fundamental theorem for weak topologies). *Let $\{v_n\}_{n\in\mathbb{N}}$ be a bounded sequence in $L^p(0,T;\mathscr{H})$, for some $p > 1$. Then there exists a subsequence $k \mapsto v_{n_k}$ and a parametrized measure $\boldsymbol{\nu} = \{\nu_t\}_{t\in(0,T)} \in \mathcal{Y}(0,T;\mathscr{H})$ such that for a.e. $t \in (0,T)$*

$$\limsup_{k\uparrow+\infty} |v_{n_k}(t)| < +\infty, \quad \nu_t \text{ is concentrated on the set } L(t) := \bigcap_{p=1}^{\infty} \overline{\{v_{n_k}(t) : k \geq p\}}^w \tag{3.5}$$

of the weak limit points *of $\{v_{n_k}(t)\}$, and*

$$\liminf_{k\to\infty} \int_0^T h(t,v_{n_k}(t))\,\mathrm{d}t \geq \int_0^T \left(\int_{\mathscr{H}} h(t,\xi)\,\mathrm{d}\nu_t(\xi)\right)\mathrm{d}t \tag{3.6}$$

for every weakly normal *integrand h such that $h^-(\cdot,v_{n_k}(\cdot))$ is uniformly integrable. In particular,*

$$\int_0^T \left(\int_{\mathscr{H}} |\xi|^p\,\mathrm{d}\nu_t(\xi)\right)\mathrm{d}t \leq \liminf_{k\to\infty} \int_0^T |v_{n_k}(t)|^p\,\mathrm{d}t < +\infty, \tag{3.7}$$

and, setting

$$v(t) := \int_{\mathscr{H}} \xi\,\mathrm{d}\nu_t(\xi), \quad \text{we have} \quad v_{n_k} \rightharpoonup v \text{ in } L^p(0,T;\mathscr{H}). \tag{3.8}$$

Finally, if $\nu_t = \delta_{v(t)}$ for a.e. $t \in (0,T)$, then

$$\langle v_{n_k}, w\rangle \to \langle v, w\rangle \quad \text{in } L^1(0,T) \quad \forall w \in L^q(0,T;\mathscr{H}), \quad \frac{1}{q} + \frac{1}{p} = 1. \tag{3.9}$$

and, up to an extraction of a further subsequence independent of t (still denoted by v_{n_k})

$$v_{n_k}(t) \rightharpoonup v(t) \quad \text{for a.e. } t \in (0,T). \tag{3.10}$$

Proof. We cannot apply directly [5], Theorem 1, since $\mathscr{H}$, endowed with its weak topology, is not a metrizable space: we circumvent this difficulty by introducing an even weaker metric on $\mathscr{H}$, which induces the usual weak convergence on every bounded set, and by considering the new sequence $\boldsymbol{v}_n(t) := (v_n(t), |v_n(t)|)$, with values in the metric space

$$E := \left\{\boldsymbol{v} = (v,w) \in \mathscr{H} \times \mathbb{R} : |v| \leq w\right\} \subset \mathscr{H} \times \mathbb{R}. \tag{3.11}$$

This construction is well-known (see *e.g.* [13]), and even easier in Hilbert spaces: indeed, we fix an orthonormal basis $\{e_m\}_{m \in \mathbb{N}}$ in $\mathscr{H}$, and we define

$$\|\|v\|\|^2 := \sum_{m=0}^{\infty} 2^{-m} |\langle e_m, v \rangle|^2 \quad \forall v \in \mathscr{H}. \tag{3.12}$$

Then, we consider the following distance on E

$$d(\boldsymbol{v}_1, \boldsymbol{v}_2) := \|\|v_1 - v_2\|\| + |w_1 - w_2|, \quad \text{if } \boldsymbol{v}_i = (v_i, w_i) \in E, \quad i = 1, 2, \tag{3.13}$$

observing that

$$\boldsymbol{v}_n = (v_n, w_n) \in E, \quad \boldsymbol{v}_n \to \boldsymbol{v} = (v, w) \text{ in } E \quad \Leftrightarrow \quad \begin{cases} v_n \rightharpoonup v \text{ weakly in } \mathscr{H}, \\ w_n \to w \text{ in } \mathbb{R}. \end{cases} \tag{3.14}$$

It is immediate to check that E is separable and complete with the distance (3.13); moreover, bounded weakly closed subsets of E are compact with respect to this new topology. In particular, any intersection of closed balls of $\mathscr{H} \times \mathbb{R}$ with E is a Borel subset of E. Therefore, for any Borel subset B of $\mathscr{H} \times \mathbb{R}$ we have

$$B \in \mathscr{B}(\mathscr{H} \times \mathbb{R}) \quad \Rightarrow \quad B \cap E \in \mathscr{B}(E);$$

thus, any Borel (probability) measure μ on E can be trivially extended to a Borel (probability) measure on $\mathscr{H} \times \mathbb{R}$.

We can now apply Balder's theorem ([5], Th. 1) to the sequence $\boldsymbol{v}_n := (v_n, |v_n|)$ in E, and we thus find a subsequence $\boldsymbol{v}_{n_k}$ and a parametrized measure $\boldsymbol{\mu} = \{\mu_t\}_{t \in (0,T)} \in \mathcal{Y}(0, T; E)$ such that for *a.e.* $t \in (0, T)$

$$\mu_t \text{ is concentrated on the set } \boldsymbol{L}(t) := \bigcap_{p=1}^{\infty} \overline{\{\boldsymbol{v}_{n_k}(t) : k \geq p\}}^E, \quad \boldsymbol{v}_{n_k}(t) \text{ is bounded} \tag{3.15}$$

(*i.e.*, $\boldsymbol{L}(t)$ is the set of the *E-limit points* of $\{\boldsymbol{v}_{n_k}(t)\}$; the boundedness of $\boldsymbol{v}_{n_k}(t)$ follows, *e.g.*, by [42], Th. 2), and

$$\liminf_{k \to \infty} \int_0^T g(t, v_{n_k}(t), |v_{n_k}(t)|) \, \mathrm{d}t \geq \int_0^T \left(\int_E g(t, v, w) \, \mathrm{d}\mu_t(v, w) \right) \mathrm{d}t \tag{3.16}$$

for every E-normal integrand g such that $g^-(\cdot, v_{n_k}(\cdot), |v_{n_k}(\cdot)|)$ is uniformly integrable.

Setting

$$\nu_t(A) := \mu_t\big(E \cap (A \times [0, +\infty))\big) \quad \forall A \in \mathscr{B}(\mathscr{H}), \tag{3.17}$$

we obtain a parametrized measure $\boldsymbol{\nu} := \{\nu_t\}_{t \in (0,T)}$ which satisfies (3.5) and (3.6). Indeed, by (3.15) the sequences $\{v_{n_k}(t)\}_{k \in \mathbb{N}}$ are bounded in $\mathscr{H}$ for *a.e.* $t \in (0, T)$, and we note that for *a.e.* $t \in (0, T)$

$$\boldsymbol{L}(t) \subset L(t) \times [0, +\infty) \quad \text{and, from (3.15),} \quad \mu_t(E \setminus \boldsymbol{L}(t)) = 0.$$

Then, taking into account (3.17), we conclude that

$$\nu_t(\mathscr{H} \setminus L(t)) = \mu_t(E \setminus (L(t) \times [0, +\infty))) \leq \mu_t(E \setminus \boldsymbol{L}(t)) = 0,$$

which yields (3.5). In the end, (3.6) follows from (3.16) simply by choosing $g(t, v, w) := h(t, v)$ and observing that

$$\int_E h(t, v) \, \mathrm{d}\mu_t(v, w) = \int_{\mathscr{H}} h(t, v) \, \mathrm{d}\nu_t(v),$$

whereas we deduce (3.8) from (3.7) by choosing the family of weakly normal integrands

$$h(t, v) := \langle w(t), v \rangle \quad \text{with } w \in L^q(0, T; \mathscr{H}).$$

Finally (3.9) follows by the same argument, by putting

$$h(t, v) := |\langle v - v(t), w(t) \rangle| \quad \text{with } w \in L^q(0, T; \mathscr{H}).$$

(3.10) ensues from the boundedness of $v_{n_k}(t)$ (3.5) and (3.16), by choosing $g(t, v, \rho) := -\|v - v(t)\|$. $\qquad\square$

3.2. **Young measures and the Chain Rule**

We have stated the Chain Rule conditions ($\mathrm{CHAIN}_{1,2}$) in a sort of "global" formulation: roughly speaking, whenever we know that a curve $v \in H^1(0, T; \mathscr{H})$ admits a global selection $\xi \in L^2(0, T; \mathscr{H})$ in the limiting subdifferential $\partial_\ell \phi \circ v$, then we are able to evaluate the time derivative of $\phi \circ v$ in terms of v' and of that particular selection ξ outside a negligible set $\mathscr{N} \subset (0, T)$, which in principle depends on ξ.

We have adopted this point of view, since these kinds of conditions are easier to check in several concrete cases (we will see an important example in Prop. 5.7 later on); on the other hand, it would often be useful to know if the following two stronger properties, valid *e.g.* in the convex case, hold too:

(1) the chain rule holds outside a negligible set $\mathscr{N}$ which does not depend on the particular selection ξ;
(2) if $t \in (0, T) \setminus \mathscr{N}$, we can choose an arbitrary element of $\partial_\ell \phi(v(t))$ in order to evaluate the time derivative of $\phi \circ v$.

We are going to show now that the Chain Rule conditions ($\mathrm{CHAIN}_{1,2}$) imply the two properties above, which are also suitable to deal with Young measures.

Before stating this result, we recall that aff A (resp. $\overline{\text{aff}}\, A$) denotes the *affine hull* (resp. its closure) of a subset $A \subset \mathscr{H}$ (1.8), and we set $|\partial_\ell^\circ \phi(v)| := \inf_{\xi \in \partial_\ell \phi(v)} |\xi|$ (1.9), with the convention $\inf \varnothing = +\infty$.

We will see in Lemma 3.4 below that if $v \in C^0(0, T; \mathscr{H})$ with $v(t) \in D(\partial_\ell \phi)$ for *a.e.* $t \in (0, T)$, then the map $t \mapsto |\partial_\ell^\circ \phi(v(t))|$ is measurable and

$$\int_0^T |\partial_\ell^\circ \phi(v(t))|^2 \, \mathrm{d}t < +\infty \quad \Leftrightarrow \quad \begin{cases} \exists \xi \in L^2(0, T; \mathscr{H}) : \\ \xi(t) \in \partial_\ell \phi(v(t)) \quad \text{for } a.e. \ t \in (0, T). \end{cases} \tag{3.18}$$

Theorem 3.3. *Let us suppose that ϕ satisfies the Chain Rule condition* (CHAIN_1)*, let $v \in H^1(0, T; \mathscr{H})$ be such that $\phi \circ v$ is a.e. equal to a function φ of bounded variation and $v(t) \in D(\partial_\ell \phi)$ for a.e. $t \in (0, T)$.*

(1) *If*

$$\int_0^T |\partial_\ell^\circ \phi(v(t))|^2 \, \mathrm{d}t < +\infty, \tag{3.19}$$

then

$$\varphi'(t) = \langle \xi, v'(t) \rangle \quad \forall \xi \in \overline{\text{aff}}\big(\partial_\ell \phi(v(t))\big) \quad \text{for a.e. } t \in (0, T). \tag{3.20}$$

(2) *If $\boldsymbol{\mu} = \{\mu_t\}_{t \in (0,T)}$ is a Young measure in $\mathscr{H}$ satisfying*

$$\int_0^T \int_{\mathscr{H}} |\xi|^2 \, \mathrm{d}\mu_t(\xi) \, \mathrm{d}t < +\infty, \quad \mu_t(\mathscr{H} \setminus \partial_\ell \phi(v(t))) = 0 \quad \text{for a.e. } t \in (0, T), \tag{3.21}$$

then

$$\varphi'(t) = \int_{\mathscr{H}} \langle \xi, v'(t) \rangle \, \mathrm{d}\mu_t(\xi) \quad \text{for a.e. } t \in (0, T). \tag{3.22}$$

Proof. (1). The next lemma shows that we can find a sequence $(\zeta_n)_{n \in \mathbb{N}} \subset L^2(0, T; \mathscr{H})$ and a Borel set $\mathcal{D} \subset (0, T)$ with full measure, *i.e.* $|(0, T) \setminus \mathcal{D}| = 0$, such that

$$\{\zeta_n(t) : \ n \in \mathbb{N}\} \subset \partial_\ell \phi(v(t)) \subset \overline{\{\zeta_n(t) : \ n \in \mathbb{N}\}} \quad \forall t \in \mathcal{D}. \tag{3.23}$$

Applying the Chain Rule (CHAIN$_1$) to each map ζ_n, we can find negligible sets $\mathcal{N}_n$ such that

$$\varphi'(t) = \langle \zeta_n(t), v'(t) \rangle \quad \forall t \in (0, T) \setminus \mathcal{N}_n.$$

Therefore, setting $\mathcal{D}_0 := \mathcal{D} \setminus \bigcup \mathcal{N}_n$, we have

$$\varphi'(t) = \langle \zeta_n(t), v'(t) \rangle \quad \forall n \in \mathbb{N}, \ \forall t \in \mathcal{D}_0,$$

and this relation extends to $\overline{\mathrm{aff}}\{\zeta_n(t) : n \in \mathbb{N}\}$, which coincides with $\overline{\mathrm{aff}}(\partial_\ell \phi(v(t)))$.

(2). Condition (3.21) yields that $\partial_\ell \phi(v(t)) \neq \varnothing$ for *a.e.* $t \in (0, T)$ and (3.19) is satisfied, since

$$|\partial_\ell^\circ \phi(v(t))|^2 \leq \int_{\mathscr{H}} |\xi|^2 \, \mathrm{d}\mu_t(\xi) \quad \text{for } a.e. \ t \in (0, T).$$

Then, by the previous claim, (3.20) holds: since μ_t is *a.e.* concentrated on $\partial_\ell \phi(v(t))$, integrating (3.20) in $\mathscr{H}$ with respect to μ_t yields (3.22). $\qquad\square$

We conclude this section with the following measurable selection result, which is the technical crucial point of the proof of Theorem 3.3.

Lemma 3.4. *Let $v : (0, T) \to \mathscr{H}$ be a Borel map and $\mathcal{D} \subset (0, T)$ be a Borel set such that*

$$(0, T) \setminus \mathcal{D} \quad \text{is negligible}, \quad \partial_\ell \phi(v(t)) \neq \emptyset \quad \text{for } t \in \mathcal{D}.$$

Then, there exists a sequence of (strongly) measurable maps $\zeta_n : (0, T) \to \mathscr{H}$ such that

$$\{\zeta_n(t) : \ n \in \mathbb{N}\} \subset \partial_\ell \phi(v(t)) \subset \overline{\{\zeta_n(t) : \ n \in \mathbb{N}\}} \quad \text{for } t \in \mathcal{D}. \tag{3.24}$$

In particular, the map $t \mapsto |\partial^\circ \phi(v(t))| = \inf_{n \in \mathbb{N}} |\zeta_n(t)|$ is measurable; if (3.19) holds, too, then we can choose the family $\{\zeta_n\}_{n \in \mathbb{N}}$ so that $\zeta_n \in L^2(0, T; \mathscr{H})$ for every $n \in \mathbb{N}$.

Proof. On behalf of [16], Theorem III.22, (3.24) follows once we prove that the graph of the multivalued function $t \in \mathcal{D} \mapsto \partial_\ell \phi(v(t))$, *i.e.*, the set

$$\mathcal{G} := \{(t, \xi) \in \mathcal{D} \times \mathscr{H} : \ \xi \in \partial_\ell \phi(v(t))\},$$

is a Borel subset of $\mathcal{D} \times \mathscr{H}$.

To check this, it suffices to note that the set

$$G := \mathrm{graph}(\partial_\ell \phi) = \big\{(v, \xi) \in \mathscr{H} \times \mathscr{H} : \ \xi \in \partial_\ell \phi(v)\big\}$$

admits the representation $G = \bigcup_{m \in \mathbb{N}} G_m$, where G_m is the set of all $(v, \xi) \in \mathscr{H} \times \mathscr{H}$ satisfying

$$\exists \xi_n \in \partial \phi(v_n) \quad \text{s.t.} \quad v_n \to v, \ \xi_n \rightharpoonup \xi, \ \sup_n \big\{\phi(v_n), |\xi_n|\big\} \leq m.$$

In fact, G is a Borel subset of $\mathscr{H} \times \mathscr{H}$ since G_m is a *closed* subset of $\mathscr{H} \times \mathscr{H}$ for every $m \in \mathbb{N}$: indeed, let $(v^k, \xi^k) \in G_m$ and $(v, \xi) \in \mathscr{H}$ such that

$$\lim_{k \uparrow +\infty} \big(|v^k - v| + |\xi^k - \xi|\big) = 0.$$

Recalling (3.12) and (3.14), we thus find points $\xi_n^k \in \partial \phi(v_n^k)$, $n \in \mathbb{N}$, such that

$$\varepsilon_n^k := |v_n^k - v^k| + |\|\xi_n^k - \xi^k\|| \to 0 \text{ as } n \uparrow +\infty, \quad \sup_{n,k} \big\{\phi(v_n^k), |\xi_n^k|\big\} \leq m.$$

Choosing the integer $n = n_k$ such that $\varepsilon_{n_k}^k = k^{-1}$, we get $v_{n_k}^k \to v$, $\xi_{n_k}^k \rightharpoonup \xi$, since the $\mathscr{H}$-norms of ξ_n^k are uniformly bounded.

Finally, $\mathcal{G}$ may be represented as

$$\mathcal{G} = \{(t, \xi) \in \mathcal{D} \times \mathscr{H} \ : (v(t), \xi) \in G\},$$

and it is a Borel set, too, being v Borel on $(0, T)$ and $\mathscr{H}$ separable.

In order to prove the second part of the lemma, let $\{\zeta_n\}_{n \in \mathbb{N}}$ be a family of measurable maps satisfying (3.24), and let us introduce the measurable sets which are recursively defined by

$$A_0 := \varnothing, \quad A_k := \big\{t \in \mathcal{D} : |\zeta_k(t)| < |\partial_\ell^\circ \phi(v(t))| + 1\big\} \setminus \bigcup_{j=0}^{k-1} A_j.$$

By construction, the family $\{A_k\}_{k \in \mathbb{N}}$ is disjoint, $\bigcup_{k \in \mathbb{N}} A_k = \mathcal{D}$, and $|\zeta_k(t)| \le |\partial_\ell^\circ \phi(v(t))| + 1$ on A_k. The map

$$\zeta(t) := \sum_{k=1}^{+\infty} \zeta_k(t) \chi_{A_k}(t), \tag{3.25}$$

is a measurable selection of $\partial_\ell \phi(v(t))$ and belongs to $L^2(0, T; \mathscr{H})$ since $|\zeta(t)| \le |\partial_\ell^\circ \phi(v(t))| + 1$ for every $t \in \mathcal{D}$.

Finally, we use ζ to construct a new countable family of functions

$$\zeta_{n,k}(t) := \begin{cases} \zeta_n(t) & \text{if } |\zeta_n(t)| \le k, \\ \zeta(t) & \text{otherwise}, \end{cases} \tag{3.26}$$

which belong to $L^2(0, T; \mathscr{H})$, and satisfy

$$\zeta_{n,k}(t) \in \partial_\ell \phi(v(t)), \quad \{\zeta_n(t) : n \in \mathbb{N}\} \subset \{\zeta_{n,k}(t) : n, k \in \mathbb{N}\} \quad \text{if } t \in \mathcal{D}. \qquad \square$$

4. Proof of the abstract results

In this section, we collect the proof of the main abstract theorems we have stated in the Introduction. We split the presentation in four steps:

(1) in the first preliminary paragraph, we study in some detail the "stationary" estimates which concern each single step of the minimization scheme (1.20): here, we adapt to the presence of the discrete source term F_τ^n some well-known estimates, which are related to the celebrated Moreau-Yosida approximation of ϕ. The crucial lemma 4.2 will provide the basic discrete energy estimate for the so-called De Giorgi's *variational interpolant* of the discrete values U_τ^n.

(2) In the second step, we will introduce this new kind of interpolating families, and we will briefly derive the discrete equations satisfied by the approximate solutions.

(3) The third step is devoted to the basic *a priori* estimates, which yield enough compactness to extract a limit curve $u \in GMM(\Phi; u_0, f)$ from the family of the discrete solutions; Proposition 4.7 provides a fine description, in terms of Young measures, of the asymptotic inequalities satisfied by this limit.

(4) Finally, in the fourth step we show that this limit curve is in fact a solution to (GF) if $\partial_\ell \phi$ satisfies the convexity condition (CONV) or one of the Chain Rule properties (CHAIN$_{1,2}$).

The last paragraph discusses the validity of the Chain Rule for dominated concave perturbations and contains the proof of Theorem 4.

Throughout this section, we will always assume ϕ to fulfill (COMP), which in particular ensures that there exists some positive constant S such that

$$\phi(v) + \frac{1}{2\tau_*}|v|^2 \geq -S \quad \forall v \in \mathscr{H}. \tag{4.1}$$

In particular, since $|w - v|^2 \geq \frac{1}{2}|w|^2 - |v|^2$, we get

$$\frac{|w - v|^2}{\tau} + \phi(w) \geq -S - \frac{|v|^2}{\tau_*} \quad \forall \tau \leq \tau_*, \quad \forall v, w \in \mathscr{H}. \tag{4.2}$$

4.1. Estimates for the Moreau-Yosida approximation

Here we collect some preliminary results on the minimization problem (1.20): in particular, we introduce the *linearly perturbed* functionals

$$\phi(v; g) := \phi(v) - \langle g, v \rangle \quad \forall v, g \in \mathscr{H},$$

their *Moreau-Yosida* approximation

$$\begin{cases} \Phi(\sigma, g, v; w) := \dfrac{|w - v|^2}{2\sigma} + \phi(w; g), \\ \phi_\sigma(v; g) := \inf_{w \in \mathscr{H}} \Phi(\sigma, g, v; w) \quad v, g \in \mathscr{H}, \ \sigma > 0, \end{cases} \tag{4.3}$$

and the set $J_\sigma(v; g)$ where the infimum in (4.3) is attained, *i.e.*,

$$J_\sigma(v; g) := \operatorname*{argmin}_{w \in \mathscr{H}} \Phi(\sigma, g, v; w), \quad v, g \in \mathscr{H}. \tag{4.4}$$

Let us point out that $J_\sigma(v; g)$ is non empty for every $\sigma \in (0, \tau_*)$. Denoting by v_σ the generic element of $J_\sigma(v; g)$, we also set

$$\mathsf{d}_\sigma^+(v; g) := \sup_{v_\sigma \in J_\sigma(v; g)} |v_\sigma - v|, \quad \mathsf{d}_\sigma^-(v; g) := \inf_{v_\sigma \in J_\sigma(v; g)} |v_\sigma - v|. \tag{4.5}$$

Let us collect some basic properties of ϕ_σ.

Lemma 4.1. *For every $v, g \in \mathscr{H}$ and for every $0 < \sigma_1 < \sigma_2$ there holds*

$$\phi_{\sigma_2}(v; g) \leq \phi_{\sigma_1}(v; g) \leq \phi(v; g), \qquad |v_{\sigma_1} - v| \leq |v_{\sigma_2} - v|,$$
$$\mathsf{d}_{\sigma_1}^+(v; g) \leq \mathsf{d}_{\sigma_2}^-(v; g) \leq \mathsf{d}_{\sigma_2}^+(v; g). \tag{4.6}$$

In particular, for every $v, g \in \mathscr{H}$ there exists an (at most) countable set $\mathscr{N}_{v,g} \subset (0, \tau_*)$ such that

$$\mathsf{d}_\sigma^+(v; g) = \mathsf{d}_\sigma^-(v; g) \quad \forall \sigma \in (0, \tau_*) \setminus \mathscr{N}_{v,g}. \tag{4.7}$$

Finally, for every $v \in D(\phi)$ we have

$$\lim_{\sigma \downarrow 0} \mathsf{d}_\sigma^+(v; g) = 0, \qquad \lim_{\sigma \downarrow 0} \phi_\sigma(v; g) = \lim_{\sigma \downarrow 0} \inf_{v_\sigma \in J_\sigma(v; g)} \phi(v_\sigma) = \phi(v; g),$$
$$g \in \partial\phi(v) \ \Rightarrow \ \lim_{\sigma \downarrow 0} \frac{\mathsf{d}_\sigma^+(v; g)}{\sigma} = 0. \tag{4.8}$$

Proof. It is straightforward to check the first chain of inequalities in (4.6), yielding that the map $\sigma \mapsto \phi_\sigma(v; g)$ is non-increasing. As for the second inequality, we notice that

$$
\frac{1}{2\sigma_1}|v_{\sigma_1} - v|^2 + \phi(v_{\sigma_1}) - \langle g, v_{\sigma_1} \rangle \leq \frac{1}{2\sigma_1}|v_{\sigma_2} - v|^2 + \phi(v_{\sigma_2}) - \langle g, v_{\sigma_2} \rangle
$$

$$
= \left(\frac{1}{2\sigma_1} - \frac{1}{2\sigma_2} \right)|v_{\sigma_2} - v|^2 + \frac{1}{2\sigma_2}|v_{\sigma_2} - v|^2 + \phi(v_{\sigma_2}) - \langle g, v_{\sigma_2} \rangle
$$

$$
\leq \left(\frac{1}{2\sigma_1} - \frac{1}{2\sigma_2} \right)|v_{\sigma_2} - v|^2 + \frac{1}{2\sigma_2}|v_{\sigma_1} - v|^2 + \phi(v_{\sigma_1}) - \langle g, v_{\sigma_1} \rangle,
$$

whence

$$
\left(\frac{1}{2\sigma_1} - \frac{1}{2\sigma_2} \right)|v_{\sigma_1} - v|^2 \leq \left(\frac{1}{2\sigma_1} - \frac{1}{2\sigma_2} \right)|v_{\sigma_2} - v|^2,
$$

and the third chain of inequalities in (4.6) follows. Therefore, for every $v, g \in \mathscr{H}$ the functions $\sigma \mapsto \mathsf{d}_\sigma^+(v; g)$, $v \mapsto \mathsf{d}_\sigma^-(v; g)$ are non decreasing. Let $\mathscr{N}_{v,g} \subset (0, \tau_*)$ be the (at most countable) set of the discontinuity points of $\sigma \mapsto \mathsf{d}_\sigma^\pm(v; g)$: then, for every $\sigma \in (0, \tau_*) \setminus \mathscr{N}_{v,g}$ we have by (4.6) and continuity (here we omit the dependence on g)

$$
\liminf_{s \uparrow \sigma} \mathsf{d}_s^-(v) \leq \limsup_{s \uparrow \sigma} \mathsf{d}_s^+(v) \leq \mathsf{d}_\sigma^-(v) \leq \mathsf{d}_\sigma^+(v) \leq \liminf_{t \downarrow \sigma} \mathsf{d}_t^-(v) \leq \limsup_{t \downarrow \sigma} \mathsf{d}_t^+(v),
$$

$$
\mathsf{d}_\sigma^-(v) = \lim_{s \uparrow \sigma} \mathsf{d}_s^-(v) = \lim_{t \downarrow \sigma} \mathsf{d}_t^-(v), \quad \mathsf{d}_\sigma^+(v) = \lim_{s \uparrow \sigma} \mathsf{d}_s^+(v) = \lim_{t \downarrow \sigma} \mathsf{d}_t^+(v),
$$

$$
\tag{4.9}
$$

whence (4.7).

As for (4.8), recalling (4.2) let us preliminarily note that for every $w \in \mathscr{H}$ and $\sigma < \tau_*/2$

$$
\phi(w) + \frac{|w - v|^2}{2\sigma} = \phi(w) + \frac{\tau_* - 2\sigma}{2\sigma(\tau_* + 2\sigma)}|w - v|^2 + \frac{2}{\tau_* + 2\sigma}|w - v|^2 \geq \frac{\tau_* - 2\sigma}{4\sigma\tau_*}|w - v|^2 - S - \tau_*^{-1}|v|^2.
$$

In particular, choosing $w = v_\sigma \in J_\sigma(v; g)$ we obtain

$$
\frac{\tau_* - 2\sigma}{4\tau_*} \frac{|v_\sigma - v|^2}{\sigma} \leq \frac{|v_\sigma - v|^2}{2\sigma} + \phi(v_\sigma) + \frac{|v|^2}{\tau_*} + S.
\tag{4.10}
$$

On the other hand, the minimization scheme defining ϕ_σ yields

$$
\frac{|v_\sigma - v|^2}{2\sigma} + \phi(v_\sigma) \leq \phi(v) - \langle g, v - v_\sigma \rangle \leq \phi(v) + 2\frac{\sigma\tau_*}{\tau_* - 2\sigma}|g|^2 + \frac{\tau_* - 2\sigma}{8\sigma\tau_*}|v_\sigma - v|^2.
\tag{4.11}
$$

Combining (4.10) and (4.11), we infer the existence of a constant C depending only on S and τ_* such that

$$
|v_\sigma - v|^2 \leq 8\sigma\frac{\tau_*}{\tau_* - 2\sigma} \left(S + \tau_*^{-1}|v|^2 + \phi(v) + 2\frac{\sigma\tau_*}{\tau_* - \sigma}|g|^2 \right) \leq C\sigma(1 + |v|^2 + \phi(v) + \sigma|g|^2),
\tag{4.12}
$$

supposing that, *e.g.*, $\tau_*/(\tau_* - 2\sigma) < 2$. Hence, since v_σ is arbitrary in $J_\sigma(v; g)$, we have proved the first assertion in (4.8).

To check the second one, it suffices to remark that

$$
\phi(v; g) \geq \limsup_{\sigma \downarrow 0} \phi_\sigma(v; g) \geq \liminf_{\sigma \downarrow 0} \phi_\sigma(v; g) \geq \liminf_{\sigma \downarrow 0} \inf_{v_\sigma \in J_\sigma(v; g)} \phi(v_\sigma; g) \geq \phi(v; g),
$$

where we have used (4.6) and the lower semicontinuity of the functional $v \mapsto \phi(v; g)$.

Finally, the last limit of (4.8) follows from the first inequality of (4.11), since by (1.1)

$$\frac{|v_\sigma - v|^2}{2\sigma} \leq \phi(v) - \phi(v_\sigma) - \langle g, v - v_\sigma \rangle \leq |v_\sigma - v| \cdot o(1) \quad \text{as } \sigma \to 0. \qquad \square$$

We point out that (4.6) yields that for every $0 < \sigma_1 < \sigma_2 < \tau_*$ the maps $\sigma \mapsto \sigma^{-1} \mathsf{d}_\sigma^\pm$ have finite pointwise variation on (σ_1, σ_2).

The next lemma, though simple, contains the crucial estimate for the De Giorgi variational interpolants we will introduce in the next paragraph (see *e.g.* [3], Lem. 4.1.2, Th. 4.1.4, and [1], Lem. 2.5).

Lemma 4.2. *Under the present assumptions, we have that for every $v \in D(\phi)$ and $g \in \mathscr{H}$ the map $\sigma \mapsto \phi_\sigma(v; g)$ is locally Lipschitz on $(0, \tau_*)$ and*

$$\frac{\mathrm{d}}{\mathrm{d}\sigma} \phi_\sigma(v; g) = -\frac{(\mathsf{d}_\sigma^+(v; g))^2}{2\sigma^2} = -\frac{(\mathsf{d}_\sigma^-(v; g))^2}{2\sigma^2} \quad \forall \sigma \in (0, \tau_*) \setminus \mathscr{N}_{v,g}, \tag{4.13}$$

($\mathscr{N}_{v,g}$ being as in (4.7)).
In particular, we have

$$\frac{|v_{\sigma_0} - v|^2}{2\sigma_0} + \frac{1}{2} \int_0^{\sigma_0} \frac{(\mathsf{d}_\sigma^\pm(v; g))^2}{\sigma^2} \, \mathrm{d}\sigma = \phi(v) - \phi(v_{\sigma_0}) - \langle g, v - v_{\sigma_0} \rangle \tag{4.14}$$

for every $0 < \sigma_0 < \tau_*$ and $v_{\sigma_0} \in J_{\sigma_0}(v; g)$.

Proof. A simple computation yields

$$\phi_{\sigma_1}(v; g) - \phi_{\sigma_2}(v; g) \leq \phi(v_{\sigma_2}; g) + \frac{|v_{\sigma_2} - v|^2}{2\sigma_1} - \phi(v_{\sigma_2}; g) - \frac{|v_{\sigma_2} - v|^2}{2\sigma_2} = \frac{(\sigma_2 - \sigma_1)|v_{\sigma_2} - v|^2}{2\sigma_1 \sigma_2} \tag{4.15}$$

for every $\sigma_1, \sigma_2 \in (0, \tau_*)$ and $v_{\sigma_i} \in J_{\sigma_i}(v; g)$, $i = 1, 2$. Changing sign to the inequality and interchanging σ_1 and σ_2, we obtain

$$\phi_{\sigma_1}(v; g) - \phi_{\sigma_2}(v; g) \geq \frac{(\sigma_2 - \sigma_1)|v_{\sigma_1} - v|^2}{2\sigma_1 \sigma_2}.$$

Then, since the real map $\sigma \mapsto \phi_\sigma(v; g)$ is non increasing, we deduce

$$0 \leq \frac{(\mathsf{d}_{\sigma_1}^+(v; g))^2}{2\sigma_1 \sigma_2} \leq \frac{\phi_{\sigma_1}(v; g) - \phi_{\sigma_2}(v; g)}{\sigma_2 - \sigma_1} \leq \frac{(\mathsf{d}_{\sigma_2}^-(v; g))^2}{2\sigma_1 \sigma_2} \quad \forall \, 0 < \sigma_1 < \sigma_2, \tag{4.16}$$

so that $\sigma \mapsto \phi_\sigma(v; g)$ is locally Lipschitz on $(0, \tau_*)$. Passing to the limit in (4.16) as $\sigma_1 \uparrow \sigma$ and $\sigma_2 \downarrow \sigma$ and recalling (4.9), we conclude that (4.13) holds at every $\sigma \in (0, \tau_*) \setminus \mathscr{N}_{v,g}$.

Integrating (4.13) on $(0, \sigma_0)$ and using (4.8), we deduce that

$$\frac{|v_{\sigma_0} - v|^2}{2\sigma_0} + \phi(v_{\sigma_0}; g) - \phi(v; g) = \phi_{\sigma_0}(v; g) - \lim_{\sigma \downarrow 0} \phi_\sigma(v; g) = \int_0^{\sigma_0} \frac{\mathrm{d}}{\mathrm{d}\sigma} \phi_\sigma(v; g) \, \mathrm{d}\sigma = -\int_0^{\sigma_0} \frac{(\mathsf{d}_\sigma^\pm(v; g))^2}{2\sigma^2} \, \mathrm{d}\sigma,$$

i.e., (4.14). $\qquad \square$

4.2. Discrete equations and variational interpolants

Let us first briefly recall the approximation algorithm for (GF) we sketched in Section 1 (*cf.* (1.19)–(1.20)). At each time step $\tau > 0$ there corresponds a partition of the interval $(0, T)$

$$\mathscr{P}_\tau := \{t_0 = 0 < t_1 < \cdots < t_n < \cdots < t_{N-1} < T \leq t_N\}, \quad t_n := n\tau, \quad N \in \mathbb{N}.$$

For $f \in L^2(0, T; \mathscr{H})$ (which we trivially extend to 0 outside the time interval $(0, T)$), we denote by $\overline{F}_\tau$ the (left-continuous) piecewise constant interpolant taking the value F_τ^n on $(t_{n-1}, t_n]$:

$$\overline{F}_\tau(t) := F_\tau^n = \frac{1}{\tau} \int_{t_{n-1}}^{t_n} f(s) \, ds \quad \text{for } t \in (t_{n-1}, t_n], \quad n = 1, \dots, N. \tag{4.17}$$

Note that

$$\tau |\overline{F}_\tau(t)|^2 \leq \|f\|_{L^2(t_{n-1}, t_n; \mathscr{H})}^2 \quad \forall t \in (t_{n-1}, t_n], \tag{4.18}$$

$$\|\overline{F}_\tau\|_{L^2(t_m, t_n; \mathscr{H})}^2 \leq \|f\|_{L^2(t_m, t_n; \mathscr{H})}^2 \quad \forall 1 \leq m < n \leq N. \tag{4.19}$$

$$\overline{F}_\tau \to f \quad \text{as } \tau \downarrow 0 \text{ strongly in } L^2(0, T; \mathscr{H}). \tag{4.20}$$

We consider an approximate solution $\{U_\tau^n\}_{n=0}^N$ of the iterative scheme (1.20), which, taking into account (1.19) and the notation (4.4), may also be rephrased as

$$\begin{cases} U_\tau^0 := u_0, \\ U_\tau^n \in J_\tau(U_\tau^{n-1}; F_\tau^n) = \operatorname{argmin}_{V \in \mathscr{H}} \Phi(\tau, F_\tau^n, U_\tau^{n-1}; V), \end{cases} \tag{4.21}$$

for $n = 1, \dots, N$. For $\tau < \tau_*$ (COMP) guarantees that this minimization scheme always admits a solution. We have already denoted by $\overline{U}_\tau$ the piecewise constant interpolant of the values $\{U_\tau^n\}_{n=0}^n$, and by U_τ the corresponding piecewise linear interpolant, defined by

$$\overline{U}_\tau(t) := U_\tau^n, \quad U_\tau(t) := \frac{t_n - t}{\tau} U_\tau^{n-1} + \frac{t - t_{n-1}}{\tau} U_\tau^n, \quad t \in (t_{n-1}, t_n]. \tag{4.22}$$

We also introduce the following *variational interpolant*, due to E. De Giorgi, for which the local *energy inequality* (4.14) plays a key role.

Definition 4.3 (De Giorgi variational interpolants). We denote by $\widetilde{U}_\tau$ any interpolant of the discrete values $\{U_\tau^n\}_{n=0}^N$ given by

$$\begin{cases} \widetilde{U}_\tau(0) = u_0, \quad \text{and, for } t = t_{n-1} + \sigma \in (t_{n-1}, t_n], \\ \widetilde{U}_\tau(t) \in J_\sigma(U_\tau^{n-1}; F_\tau^n) = \operatorname{argmin}_{V \in \mathscr{H}} \Phi(\sigma, F_\tau^n, U_\tau^{n-1}; V), \end{cases} \tag{4.23}$$

such that the map $t \mapsto \widetilde{U}_\tau(t)$ is Lebesgue measurable in $(0, T)$.

Remark 4.4 (measurability of $\widetilde{U}_\tau$). Since the map $\sigma \mapsto J_\sigma(U_\tau^{n-1}; F_\tau^n)$ is compact valued and upper semicontinuous, the existence of a measurable selection $\widetilde{U}_\tau(t_{n-1} + \sigma) \in J_\sigma(U_\tau^{n-1}; F_\tau^n)$, $\sigma \in (t_{n-1}, t_n]$, is ensured by [16], Corollary III.3, Theorem III.6.

Note that when $t = t_n$, the minimization scheme in (4.23) coincides with the one in (4.21), so that we can always assume that

$$\widetilde{U}_\tau(t_n) = U_\tau(t_n) = \overline{U}_\tau(t_n) = U_\tau^n, \quad \text{for every } n = 1, \dots, N. \tag{4.24}$$

Further, (4.23) and the standard calculus properties (1.21) of the Fréchet subdifferential yield the discrete inclusion

$$\frac{\widetilde{U}_\tau(t) - U_\tau^{n-1}}{t - t_{n-1}} + \partial \phi(\widetilde{U}_\tau(t)) \ni F_\tau^n \quad \forall t \in (t_{n-1}, t_n], \quad n = 1, \dots, N, \tag{4.25}$$

which we also write as

$$\widetilde{\Theta}_\tau(t) + \partial \phi(\widetilde{U}_\tau(t)) \ni \overline{F}_\tau(t) \quad \forall t \in (0, T), \tag{4.26}$$

where

$$\widetilde{\Theta}_\tau(t) := \frac{\widetilde{U}_\tau(t) - U_\tau^{n-1}}{t - t_{n-1}} \quad \text{for } t \in (t_{n-1}, t_n].$$ (4.27)

Note that (4.25) reduces to

$$\frac{U_\tau^n - U_\tau^{n-1}}{\tau} + \partial\phi(U_\tau^n) \ni F_\tau^n,$$ (4.28)

at the nodes t_n of the partition, which can be equivalently rephrased as

$$U_\tau'(t) + \partial\phi(\overline{U}_\tau(t)) \ni \overline{F}_\tau(t) \quad \text{for } a.e. \ t \in (0, T).$$ (4.29)

Finally, recalling (4.5), we introduce the real function G_τ by

$$G_\tau(t) := \frac{\mathsf{d}_\sigma^+(U_\tau^{n-1}; F_\tau^n)}{\sigma}, \quad \text{for } t = t_{n-1} + \sigma \in (t_{n-1}, t_n].$$

Note that

$$G_\tau(t) \geq \frac{|\widetilde{U}_\tau(t) - U_\tau^{n-1}|}{t - t_{n-1}} = |\widetilde{\Theta}_\tau(t)| \quad \text{for } t = t_{n-1} + \sigma \in (t_{n-1}, t_n].$$ (4.30)

4.3. *A priori* estimates and compactness of the approximate solutions

Preliminarily, we recall the following well-known *Discrete Gronwall lemma*:

Lemma 4.5. *Let B, b, and κ be positive constants fulfilling $1 - b \geq \frac{1}{\kappa} > 0$ and let $\{a_n\} \subset [0, +\infty)$ be a sequence satisfying*

$$a_n \leq B + b\sum_{k=1}^n a_k \quad \forall n \in \mathbb{N}.$$

Then, $\{a_n\}$ can be bounded by

$$a_n \leq \kappa B \mathrm{e}^{\kappa b n} \quad \forall n \in \mathbb{N}.$$ (4.31)

Proof. (4.31) can be easily checked by induction: for $n = 1$ we have

$$a_1 \leq B + b a_1 \quad \Rightarrow \quad a_1 \leq \kappa B \leq \kappa B \mathrm{e}^{\kappa b};$$

assuming (4.31) for all $n < \bar{n}$, we get

$$\kappa^{-1} a_{\bar{n}} \leq (1 - b) a_{\bar{n}} \leq B + b\sum_{n=1}^{\bar{n}-1} a_n \leq B + b\kappa B \sum_{n=1}^{\bar{n}-1} \mathrm{e}^{\kappa b n} = B + b\kappa B \frac{\mathrm{e}^{\kappa b \bar{n}} - 1}{\mathrm{e}^{\kappa b} - 1}$$

$$\leq B + b\kappa B \frac{\mathrm{e}^{\kappa b \bar{n}} - 1}{\kappa b} \leq B + B\mathrm{e}^{\kappa b \bar{n}} - B \leq B\mathrm{e}^{\kappa b \bar{n}} \qquad \square$$

Proposition 4.6 (*A priori* estimates). *For $0 < \tau < \tau_*$ let $\overline{U}_\tau$, U_τ, and $\widetilde{U}_\tau$ be the families of interpolant defined by (4.22) and (4.23), respectively, in terms of the discrete solutions of the minimization scheme (4.21). Then, the* discrete energy equality

$$\frac{1}{2}\int_s^t \left(|U_\tau'(\sigma)|^2 + |G_\tau(\sigma)|^2\right) \mathrm{d}\sigma + \phi(\overline{U}_\tau(t)) = \phi(\overline{U}_\tau(s)) + \int_s^t \langle \overline{F}_\tau(\sigma), U_\tau'(\sigma) \rangle \mathrm{d}\sigma$$ (4.32)

holds for every pair of nodes $s < t \in \mathscr{P}_\tau$.

Moreover, there exists a positive constant C, only depending on the data u_0, f, and on the parameters τ_*, S, such that

$$\|\overline{U}_\tau\|_{L^\infty(0,T;\mathscr{H})} + \|U_\tau\|_{L^\infty(0,T;\mathscr{H})} + \|\widetilde{U}_\tau\|_{L^\infty(0,T;\mathscr{H})} \le C, \tag{4.33}$$

$$\sup_{[0,T]} \left(\phi(\overline{U}_\tau), \phi(\widetilde{U}_\tau) \right) \le C, \tag{4.34}$$

$$\|U_\tau'\|_{L^2(0,T;\mathscr{H})} + \|\widetilde{\Theta}_\tau\|_{L^2(0,T;\mathscr{H})} \le C, \tag{4.35}$$

$$\|\overline{U}_\tau - U_\tau\|_{L^\infty(0,T;\mathscr{H})} = C\tau^{1/2}, \quad \|\widetilde{U}_\tau - U_\tau\|_{L^\infty(0,T;\mathscr{H})} = C\tau^{1/2}, \tag{4.36}$$

for every $0 < \tau \le \tau_*/10$.

Proof. First of all, starting from (4.14) and recalling (4.23), and (4.27), if we choose $t \in (t_{j-1}, t_j]$, $\sigma_0 := t_j - t_{j-1} = \tau$, $g = F_\tau^j$, $v = U_\tau^{j-1}$, $v_\sigma := \widetilde{U}_\tau(t_{j-1} + \sigma)$, $0 < \sigma < \tau$, and $u_{\sigma_0} = U_\tau^j$, we get

$$\frac{|U_\tau^j - U_\tau^{j-1}|^2}{2\tau} + \frac{1}{2}\int_{t_{j-1}}^{t_j} |G_\tau(s)|^2\,\mathrm{d}s + \phi(U_\tau^j) = \phi(U_\tau^{j-1}) + \langle F_\tau^j, U_\tau^j - U_\tau^{j-1}\rangle. \tag{4.37}$$

In particular, (4.37) yields (4.32) for $s = t_{j-1}, t = t_j$; the general case of (4.32) trivially follows by adding up the contributions (4.37) of each subinterval of the partition.

Also, (4.37) implies

$$\frac{|U_\tau^j - U_\tau^{j-1}|^2}{4\tau} \le \phi(U_\tau^{j-1}) - \phi(U_\tau^j) + \tau|F_\tau^j|^2. \tag{4.38}$$

Turning to the proof of (4.33), we may first of all note that, for every $0 < \tau < \tau_*$ and every $n = 1, \ldots, N$, we have

$$\begin{aligned}
\frac{1}{2}|U_\tau^n|^2 - \frac{1}{2}|u_0|^2 &= \sum_{k=1}^n \left(\frac{1}{2}|U_\tau^k|^2 - \frac{1}{2}|U_\tau^{k-1}|^2 \right) \le \sum_{k=1}^n \left(|U_\tau^k|^2 - |U_\tau^k||U_\tau^{k-1}| \right) \le \sum_{k=1}^n |U_\tau^k||U_\tau^k - U_\tau^{k-1}| \\
&\le \eta \sum_{k=1}^n \frac{|U_\tau^k - U_\tau^{k-1}|^2}{2\tau} + \frac{1}{2\eta}\sum_{k=1}^n \tau|U_\tau^k|^2 \le \sum_{k=1}^n 2\eta\left(\phi(U_\tau^{k-1}) - \phi(U_\tau^k) + \tau|F_\tau^k|^2 \right) + \frac{1}{2\eta}\tau|U_\tau^k|^2 \\
&\le 2\eta\left(\phi(u_0) - \phi(U_\tau^n) + \|f\|_{L^2(0,t_n;\mathscr{H})}^2 \right) + \frac{\tau}{2\eta}\sum_{k=1}^n |U_\tau^k|^2 \\
&\le \frac{\eta}{\tau_*}|U_\tau^n|^2 + 2\eta\left(\phi(u_0) + S + \|f\|_{L^2(0,t_n;\mathscr{H})}^2 \right) + \frac{\tau}{2\eta}\sum_{k=1}^n |U_\tau^k|^2,
\end{aligned}$$

where we have used Young's inequality for some $\eta > 0$, the estimate (4.18), and also (4.1) in the last inequality. Then, multiplying the previous inequality by 4 and choosing $\eta := \tau_*/4$, we get

$$|U_\tau^n|^2 \le 2|u_0|^2 + 2\tau_*\left(\phi(u_0) + S + \|f\|_{L^2(0,t_n;\mathscr{H})}^2 \right) + \frac{8\tau}{\tau^*}\sum_{k=1}^n |U_\tau^k|^2. \tag{4.39}$$

We can apply Lemma 4.5 with $\kappa := 5$, corresponding to $\tau \le \tau_*/10$, and we easily conclude the L^∞ bounds for $\overline{U}_\tau$ and U_τ in (4.33).

As for (4.35), using Young's inequality and the estimate (4.19) we deduce from (4.32)

$$\frac{1}{4}\int_0^{t_n} |U_\tau'(s)|^2\,\mathrm{d}s + \frac{1}{2}\int_0^{t_n} |G_\tau(s)|^2\,\mathrm{d}s + \phi(U_\tau(t_n)) \le \phi(u_0) + \|f\|_{L^2(0,T;\mathscr{H})}^2. \tag{4.40}$$

Taking into account (4.1) and the estimate (4.33) for $\overline{U}_\tau$ as well, we easily infer from (4.40) the L^∞ bound for $\phi(\overline{U}_\tau)$, and the L^2 estimates for U'_τ and $\widetilde{\Theta}_\tau$, also in view of (4.30).

Further, the first of (4.36) ensues from

$$|U^n_\tau - U_\tau(t)|^2 \leq \tau \int_{t_{n-1}}^{t_n} |U'_\tau(s)|^2 \, \mathrm{d}s \leq C\tau, \tag{4.41}$$

on account of (4.35).

On the other hand, (4.12) yields

$$|\widetilde{U}_\tau(t) - U^{n-1}_\tau|^2 \leq C\tau \left(1 + \phi(U^{n-1}_\tau) + \tau |F^n_\tau|^2 \right) \leq C\tau,$$

where the last inequality follows from (4.18), as well as the estimate (4.33) for $\overline{U}_\tau$ and (4.34) for $\phi(\overline{U}_\tau)$; we thus get the L^∞ bound for $\widetilde{U}_\tau$ of (4.33) and the second estimate of (4.36). Finally, we have by the definition of $\widetilde{U}_\tau$

$$\phi(\widetilde{U}_\tau(t)) \leq \phi(U^{n-1}_\tau) + \langle F^n_\tau, \widetilde{U}_\tau(t) - U^{n-1}_\tau \rangle \quad \text{if } t \in (t_{n-1}, t_n],$$

which yields the uniform bound for $\phi(\widetilde{U}_\tau)$ in (4.34). $\qquad\qquad\square$

Lemma 1.2 is a consequence of the following more refined result.

Proposition 4.7 (Compactness of the approximate solutions). *Suppose that ϕ satisfies* (COMP). *Then, given any vanishing sequence $k \mapsto \tau_k \downarrow 0$ of time steps, we can find a further subsequence (still labelled τ_k), a limit generalized Minimizing Movement $u \in H^1(0, T; \mathscr{H})$, a non-increasing function $\varphi : [0, T] \to \mathbb{R}$, and two parametrized Young measures $\boldsymbol{\mu} = \{\mu_t\}_{t \in (0,T)}$, $\boldsymbol{\nu} = \{\nu_t\}_{t \in (0,T)}$, such that as $k \uparrow +\infty$*

$$\overline{U}_{\tau_k}, U_{\tau_k}, \widetilde{U}_{\tau_k} \to u \quad \text{in } L^\infty(0, T; \mathscr{H}), \tag{4.42}$$

$$U'_{\tau_k} \rightharpoonup u' \quad \text{weakly in } L^2(0, T; \mathscr{H}), \tag{4.43}$$

$$\begin{cases} \phi(\overline{U}_{\tau_k}(t)) \to \varphi(t) \geq \phi(u(t)) \quad \forall t \in [0, T], \\ \qquad\qquad \varphi(0) = \phi(u_0), \end{cases} \tag{4.44}$$

$$\boldsymbol{\mu} \text{ (resp. } \boldsymbol{\nu}\text{) is the limit Young measure associated with } U'_{\tau_k} \text{ (resp. } \widetilde{\Theta}_{\tau_k}\text{)},$$
$$\text{and } \mu_t, \nu_t \text{ are concentrated on } f(t) - \partial_\ell \phi(u(t)) \quad \text{for a.e. } t \in (0, T), \tag{4.45}$$

$$\frac{1}{2} \int_s^t \left(\int_{\mathscr{H}} |\xi|^2 \, \mathrm{d}\mu_r(\xi) + \int_{\mathscr{H}} |\xi|^2 \, \mathrm{d}\nu_r(\xi) \right) \mathrm{d}r + \varphi(t) \leq \varphi(s) + \int_s^t \langle f(r), u'(r) \rangle \, \mathrm{d}r, \quad \forall 0 \leq s < t \leq T, \tag{4.46}$$

$$\frac{1}{2} \int_s^t \left(\liminf_{k \uparrow +\infty} |U'_{\tau_k}(r)|^2 + \int_{\mathscr{H}} |\xi|^2 \, \mathrm{d}\nu_r(\xi) \right) \mathrm{d}r + \varphi(t) \leq \varphi(s) + \int_s^t \langle f(r), u'(r) \rangle \, \mathrm{d}r, \quad \forall 0 \leq s < t \leq T. \tag{4.47}$$

Finally, if ϕ also satisfies the continuity assumption (CONT), *then we have $\phi(u(t)) = \varphi(t)$ for a.e. $t \in (0, T)$.*

Proof. It is easy to infer from (4.35) and from the elementary inequality

$$|U_\tau(t) - U_\tau(s)| \leq (t - s)^{\frac{1}{2}} \|U'_\tau\|_{L^2(0,T;\mathscr{H})},$$

that $\{U_\tau\}$ is *equicontinuous* on $\mathscr{H}$, for $0 < \tau \leq \tau_*/10$. Further, (4.33) and (4.34) yield that $\{U_\tau(t)\}_\tau$ is contained in some sublevel of the function $u \mapsto \phi(u) + \frac{1}{2\tau_*}|u|^2$: recalling (COMP), we conclude that $\{U_\tau(t)\}_\tau$ is relatively compact in $\mathscr{H}$ for every $t \in [0, T]$. Thanks to the equicontinuity property, the Ascoli compactness theorem yields that $\{U_\tau\}_\tau$ is relatively compact in $C^0([0, T]; \mathscr{H})$.

Furthermore, always in view of (4.33)–(4.36) and of well-known weak compactness results, we deduce that there exists a function $u \in H^1(0,T;\mathscr{H})$ for which the convergences (4.42, 4.43) hold (note that $\overline{U}_{\tau_k}, U_{\tau_k}$, and $\widetilde{U}_{\tau_k}$ converge to the same limit thanks to (4.36)). In particular, $u \in GMM(\Phi; u_0, f)$.

Then, (4.44) is a consequence of Helly's theorem: indeed, let us introduce the piecewise constant functions

$$\mathsf{t}_\tau(t) := \min\{n\tau : t \le n\tau\}, \quad \text{so that} \quad \mathsf{t}_\tau(t) - \tau < t \le \mathsf{t}_\tau(t), \tag{4.48}$$

and let

$$\psi_\tau(t) := \int_0^{\mathsf{t}_\tau(t)} \langle \overline{F}_\tau(s), U'_\tau(s) \rangle \, \mathrm{d}s.$$

Let us observe that the map $t \mapsto \zeta_\tau(t) := \phi(\overline{U}_\tau(t)) - \psi_\tau(t)$ is nonincreasing; therefore, by Helly's Theorem we can suppose that, up to the extraction of a suitable subsequence,

$$\zeta(t) := \lim_{k\uparrow+\infty} \zeta_{\tau_k}(t) \quad \text{exists for every } t \in [0,T],$$

and defines a non increasing function ζ. On the other hand, since $\lim_{\tau\downarrow 0} \mathsf{t}_\tau(t) = t$, (4.43) and (4.20) yield

$$\psi_{\tau_k}(t) \to \psi(t) := \int_0^t \langle f(s), u'(s) \rangle \, \mathrm{d}s \quad \forall t \in [0,T];$$

it follows that

$$\varphi(t) := \lim_{k\uparrow+\infty} \phi(\overline{U}_{\tau_k}(t)) = \lim_{k\uparrow+\infty} \zeta_{\tau_k}(t) - \psi_{\tau_k}(t) = \zeta(t) - \psi(t) \quad \text{exists for all } t \in [0,T],$$

and defines a function of bounded variation, which satisfies $\varphi(t) \ge \phi(u(t))$ in $[0,T]$ thanks to the lower semi-continuity of ϕ.

As for the last assertion in the statement, (4.35) and Fatou's lemma yield

$$\liminf_{k\uparrow+\infty} |U'_{\tau_k}(t) - \overline{F}_{\tau_k}(t)| < +\infty \quad \text{for } a.e. \ t \in (0,T);$$

if (CONT) holds, this bound and the inclusion (4.29) entail that $\liminf_{k\uparrow+\infty} \phi(\overline{U}_{\tau_k}(t)) = \phi(u(t))$ for $a.e.$ $t \in (0,T)$; since $\phi(\overline{U}_{\tau_k}(t)) \to \varphi(t)$, we get $\varphi(t) = \phi(u(t))$ for $a.e.$ $t \in (0,T)$.

Since we can assume that $\overline{F}_{\tau_k}$ converges to f pointwise $a.e.$, the a priori estimates (4.34), (4.35), as well as the inclusions (4.26) and (4.29), yield (4.45), *via* the fundamental theorem for Young measures 3.2.

Finally, (4.32) yields

$$\frac{1}{2} \int_{\mathsf{t}_\tau(s)}^t \left(|U'_\tau(\sigma)|^2 + |\widetilde{\Theta}_\tau(\sigma)|^2 \right) \mathrm{d}\sigma + \phi(\overline{U}_\tau(t)) \le \phi(\overline{U}_\tau(s)) + \int_{\mathsf{t}_\tau(s)}^{\mathsf{t}_\tau(t)} \langle \overline{F}_\tau(\sigma), U'_\tau(\sigma) \rangle \, \mathrm{d}\sigma$$

for *every choice* $0 \le s < t \le T$.

Since $\mathsf{t}_\tau(s) \downarrow s$ as $\tau \downarrow 0$, we can pass to the limit in the previous inequality thanks to (3.7), thus obtaining for every $0 \le s < s' < t \le T$

$$\frac{1}{2} \int_{s'}^t \left(\int_\mathscr{H} |\xi|^2 \, \mathrm{d}\mu_r(\xi) + \int_\mathscr{H} |\xi|^2 \, \mathrm{d}\nu_r(\xi) \right) \mathrm{d}r + \varphi(t) \le \varphi(s) + \int_s^t \langle f(r), u'(r) \rangle \, \mathrm{d}r.$$

Since s' is arbitrary, letting $s' \downarrow s$ we obtain (4.46). Finally, (4.47) follows by the same argument and Fatou's lemma.   □

4.4. **Convergence results**

Proof of Theorem 1. Now, we are going to show that the limit function $u \in GMM(\Phi; u_0, f)$ of Proposition 4.7 is a solution to (GF), if ϕ satisfies (CONV). It is immediate to check that u is a solution of (GF): (3.8) of the Fundamental Theorem for Young Measures show that

$$u'(t) = \int_{\mathscr{H}} \xi \, \mathrm{d}\mu_t(\xi) \quad \text{for } a.e.\ t \in (0, T), \tag{4.49}$$

and (4.45) ensures that for *a.e.* $t \in (0, T)$ μ_t is concentrated on $f(t) - \partial_\ell \phi(u(t))$; since $\partial_\ell \phi(u(t))$ is a closed convex set of $\mathscr{H}$, (4.49) yields that $u'(t) \in f(t) - \partial_\ell \phi(u(t))$.

Moreover, (1.28) is a direct consequence of (4.46) (for $s = 0$) and (4.44), since

$$|u'(t)|^2 \leq \int_{\mathscr{H}} |\xi|^2 \, \mathrm{d}\mu_t(\xi), \quad \left|\left(\partial_\ell \phi(u(t)) - f(t)\right)^{\circ}\right|^2 \leq \int_{\mathscr{H}} |\xi|^2 \, \mathrm{d}\nu_t(\xi), \quad \varphi(0) = \phi(u_0).$$

Finally, if (CONT) holds, we have $\phi(u(s)) = \varphi(s)$ for *a.e.* $s \in (0, T)$, and (1.29) still follows from (4.46) and the previous inequality. □

Proof of Theorem 2. Inequality (4.46) and the Lebesgue differentiation theorem yield for *a.e.* $s \in (0, T)$

$$\frac{1}{2} \int_{\mathscr{H}} |\eta|^2 \, \mathrm{d}\mu_s(\eta) + \frac{1}{2} \int_{\mathscr{H}} |\xi|^2 \, \mathrm{d}\nu_s(\xi) \leq -\varphi'(s) + \langle f(s), u'(s) \rangle. \tag{4.50}$$

If (CHAIN$_1$) holds, since we know that $\phi \circ u = \varphi$ *a.e.* in $(0, T)$ by (CONT), then Theorem 3.3 and (4.45) ensure that

$$\varphi'(s) = \int_{\mathscr{H}} \langle f(s) - \xi, u'(s) \rangle \, \mathrm{d}\nu_s(\xi) = \langle f(s) - \theta(s), u'(s) \rangle \quad \text{for } a.e.\ s \in (0, T), \tag{4.51}$$

where we have denoted by $\theta(s)$ the barycenter of ν_s, *i.e.* $\theta(s) := \int_{\mathscr{H}} \xi \, \mathrm{d}\nu_s(\xi)$. Combining (4.50) and (4.51), we get

$$\frac{1}{2} \int_{\mathscr{H}} |\eta|^2 \, \mathrm{d}\mu_s(\eta) + \frac{1}{2} \int_{\mathscr{H}} |\xi|^2 \, \mathrm{d}\nu_s(\xi) \leq \langle \theta(s), u'(s) \rangle \quad \text{for } a.e.\ s \in (0, T). \tag{4.52}$$

Since $u'(s) = \int_{\mathscr{H}} \eta \, \mathrm{d}\mu_s(\eta)$, and

$$\int_{\mathscr{H}} |\xi - \theta(s)|^2 \, \mathrm{d}\nu_s(\xi) = \int_{\mathscr{H}} |\xi|^2 \, \mathrm{d}\nu_s(\xi) - |\theta(s)|^2, \qquad \int_{\mathscr{H}} |\eta - u'(s)|^2 \, \mathrm{d}\mu_s(\eta) = \int_{\mathscr{H}} |\eta|^2 \, \mathrm{d}\mu_s(\eta) - |u'(s)|^2,$$

from (4.52) we obtain

$$\frac{1}{2} \int_{\mathscr{H}} |\eta - u'(s)|^2 \, \mathrm{d}\mu_s(\eta) + \frac{1}{2} \int_{\mathscr{H}} |\xi - \theta|^2 \, \mathrm{d}\nu_s(\xi) + \frac{1}{2} |u'(s) - \theta(s)|^2 \leq 0, \tag{4.53}$$

i.e.,

$$u'(s) = \theta(s) \in f(s) - \partial_\ell \phi(u(s)), \quad \mu_s = \nu_s = \delta_{u'(s)} \quad \text{for } a.e.\ s \in (0, T), \tag{4.54}$$

since μ_s, ν_s are concentrated on $f(s) - \partial_\ell \phi(u(s))$ for *a.e.* $s \in (0, T)$. In particular, u is a strong solution of (GF) and (1.34) holds.

Applying (3.20), we get

$$\varphi'(s) = -\langle f(s) - u'(s), u'(s)\rangle = \langle \xi, u'(s)\rangle \quad \forall \xi \in \overline{\mathrm{aff}}\big(\partial_\ell \phi(u(s))\big) \quad \text{for } a.e. \ s \in (0,T),$$

i.e., denoting by K_s the closed affine hull $\overline{\mathrm{aff}}\big(f(s) - \partial_\ell \phi(u(s))\big)$, $u'(s)$ satisfies the system

$$u'(s) \in K_s, \quad \langle u'(s), u'(s) - \eta \rangle = 0 \quad \forall \eta \in K_s, \tag{4.55}$$

which yields (1.31).

Finally, observe that (1.33) follows from (4.29) if we show that

$$\liminf_{k\uparrow+\infty} |U'_{\tau_k}(s) - u'(s)|^2 = 0 \quad \text{for } a.e. \ s \in (0,T). \tag{4.56}$$

Choosing $w := u'$ in (3.9) and possibly extracting a further subsequence (still denoted by τ_k), we get

$$\lim_{k\uparrow+\infty} \langle U'_{\tau_k}(s) - u'(s), u'(s)\rangle = 0 \quad \text{for } a.e. \ s \in (0,T). \tag{4.57}$$

Inequality (4.47), the Lebesgue differentiation theorem and the previous computations yield

$$\frac{1}{2}\liminf_{k\uparrow+\infty} |U'_{\tau_k}(s)|^2 + \frac{1}{2}|u'(s)|^2 \le \langle \theta(s), u'(s)\rangle = |u'(s)|^2,$$

i.e.,

$$\liminf_{k\uparrow+\infty} |U'_{\tau_k}(s)|^2 \le |u'(s)|^2 \quad \text{for } a.e. \ s \in (0,T).$$

Combining this relation with (4.57), we end up with

$$\liminf_{k\uparrow+\infty} |U'_{\tau_k}(s) - u'(s)|^2 = \liminf_{k\uparrow+\infty} \Big(|U'_{\tau_k}(s)|^2 - |u'(s)|^2 + 2\langle u'(s) - U'_{\tau_k}(s), u'(s)\rangle \Big)$$

$$\le \liminf_{k\uparrow+\infty} \Big(|U'_{\tau_k}(s)| - |u'(s)|^2 \Big) \le 0. \qquad \square$$

Proof of Theorem 3. We first consider the "$W^{1,1}(0,T)$" case discussed in Remark 1.5. By (CHAIN$_2$), the map $\phi \circ u$ belongs to the Sobolev space $W^{1,1}(0,T)$, and it coincides almost everywhere with its continuous representative; moreover, (4.46) with $s = 0$, (4.44), and the lower semicontinuity of ϕ yield

$$\lim_{t\downarrow 0} \phi(u(t)) = \phi(u_0),$$

so that 0 is a continuity point for $\phi \circ u$. Let $\mathcal{T} \subset [0,T]$ be the Lebesgue set of $\phi \circ u$, thus satisfying $|(0,T)\setminus\mathcal{T}| = 0$; by (3.22) we have

$$\phi(u(t)) = \phi(u_0) + \int_0^t \int_{\mathscr{H}} \langle f(r) - \xi, u'(r)\rangle \, d\nu_r(\xi) \, dr \quad \forall t \in \mathcal{T}. \tag{4.58}$$

Inserting this identity in (4.46) for $s := 0, t \in \mathcal{T}$ and arguing as in the proof of Theorem 2, we obtain the integrated form of (4.53)

$$\int_0^t \left(\int_{\mathscr{H}} |\eta - u'(r)|^2 \, d\mu_r(\eta) + \int_{\mathscr{H}} |\xi - \theta(r)|^2 \, d\nu_r(\xi) + |u'(r) - \theta(r)|^2 \right) dr \le 0, \tag{4.59}$$

which yields (4.54) and the energy identity (1.36).

Then, (1.31) can be proved in the same way as in the previous proof; passing to the limit in (4.32) for $s = 0$ and taking into account (4.58), we get for $t \in \mathcal{T}$

$$\limsup_{k\uparrow\infty} \int_0^{\mathbf{t}_{\tau_k}(t)} \left(\frac{1}{2}|U'_{\tau_k}(s)|^2 + \frac{1}{2}|\widetilde{\Theta}_{\tau_k}(s)|^2 \right) \mathrm{d}s + \phi(\overline{U}_{\tau_k}(t)) \le \phi(u_0) + \int_0^t \langle f(s), u'(s) \rangle \, \mathrm{d}s = \int_0^t |u'(s)|^2 \, \mathrm{d}s + \phi(u(t)).$$

Since each term is lower semicontinuous, *i.e.*,

$$\liminf_{k\uparrow+\infty} \int_0^{\mathbf{t}_{\tau_k}(t)} |U'_{\tau_k}|^2 \, \mathrm{d}s \ge \int_0^t |u'|^2 \, \mathrm{d}s, \quad \liminf_{k\uparrow+\infty} \int_0^{\mathbf{t}_{\tau_k}(t)} |\widetilde{\Theta}_{\tau_k}|^2 \, \mathrm{d}s \ge \int_0^t |u'|^2 \, \mathrm{d}s,$$

and $\liminf_{k\uparrow+\infty} \phi(\overline{U}_{\tau_k}(t)) \ge \phi(u(t))$, by comparison we infer for every $t \in \mathcal{T}$

$$\lim_{k\uparrow\infty} \phi(\overline{U}_{\tau_k}(t_k)) = \phi(u(t)), \quad \limsup_{k\uparrow\infty} \int_0^t |U'_{\tau_k}(s)|^2 \mathrm{d}s \le \int_0^t |u'(s)|^2 \, \mathrm{d}s, \tag{4.60}$$

which entails the strong convergence (1.37) and thus that (up to the extraction of a further subsequence) $U'_{\tau_k}(t) \to u'(t)$ for *a.e.* $t \in (0, T)$. Finally, (1.33) follows by passing to the limit in (4.29), also recalling (4.20). When $\phi \circ u$ is absolutely continuous, too, then the set $\mathcal{T}$ coincides with $[0, T]$, and we recover the complete statement of Theorem 3. $\qquad\square$

4.5. The chain rule and concave perturbations: proof of Theorem 4

We split the *proof* in two lemmata; observe that the following assumption (4.61) is slightly weaker than (1.39).

Lemma 4.8 (Subdifferential decomposition). *Let $\phi : \mathscr{H} \to (-\infty, +\infty]$ be a function satisfying* (COMP)*, let $\psi_1 : D(\phi) \to \mathbb{R}$ be l.s.c., and let $\psi_2 : \mathrm{co}\left(D(\phi)\right) \to \mathbb{R}$ be convex and l.s.c. in $D(\phi)$. If ϕ admits the decomposition $\phi = \psi_1 - \psi_2$ in $D(\phi)$, and*

$$\begin{aligned} &\forall M \ge 0 \quad \exists \rho < 1, \, \gamma \ge 0 \quad \text{such that} \quad |\partial^\circ \psi_2(u)| \le \rho|\partial^\circ \psi_1(u)| + \gamma \\ &\text{for every } u \in D(\partial\psi_1) \text{ with } \max(\phi(u), |u|) \le M, \end{aligned} \tag{4.61}$$

then every $g \in \partial\phi(u)$ with $\max(\phi(u), |u|) < M$ can be decomposed as

$$g = \lambda^1 - \lambda^2, \quad \lambda^i \in \partial\psi_i(u), \quad |\lambda^i| \le (1 - \rho)^{-1}\left(|g| + \gamma\right), \quad i = 1, 2, \tag{4.62}$$

where ρ, γ are given in terms of M by (4.61).

Proof. Let us formally extend ψ_1, ψ_2 to $+\infty$ outside $D(\phi)$ and $\mathrm{co}\left(D(\phi)\right)$ respectively; we introduce the Moreau-Yosida approximation of the convex function ψ_2:

$$\psi_{2,\varepsilon}(v) := \inf_{z \in \mathrm{co}(D(\phi))} \left(\frac{|z - v|^2}{2\varepsilon} + \psi(z) \right) \quad \forall v \in \mathscr{H}, \tag{4.63}$$

and its l.s.c. envelope

$$\bar{\psi}_2(v) := \sup_{\varepsilon > 0} \psi_{2,\varepsilon}(v) = \lim_{\varepsilon \downarrow 0} \psi_{2,\varepsilon}(v), \quad \text{with} \quad \bar{\psi}_2(v) = \psi_2(v) \quad \text{if } v \in D(\phi). \tag{4.64}$$

In particular, (4.64) yields that

$$\partial\psi_2(v) = \partial\bar{\psi}_2(v) \quad \forall v \in D(\phi). \tag{4.65}$$

Let us denote by J_ε the related resolvent operator associated with $\partial\bar\psi_2$, *i.e.* $J_\varepsilon := (I + \varepsilon\partial\bar\psi_2)^{-1}$, which is a contraction in $\mathscr{H}$. Since ψ_2 is convex, it is well-known (see [12], Props. 2.6, 2.11), that $\psi_{2,\varepsilon}$ is a $C^{1,1}$ convex function whose Fréchet differential $D\psi_{2,\varepsilon}$ satisfies

$$D\psi_{2,\varepsilon}(u) = \frac{u - J_\varepsilon(u)}{\varepsilon}, \quad D\psi_{2,\varepsilon}(u) \subset \partial\bar\psi_2(J_\varepsilon(u)), \quad |D\psi_{2,\varepsilon}(u)| \le |\partial^\circ\bar\psi_2(u)|. \tag{4.66}$$

For a given $u \in D(\partial\phi)$ with $\max\{\phi(u), |u|\} < M$, and for every $0 < \tau < \tau_*/8$ and $0 < \varepsilon < \varepsilon_*$, we consider the minimization problem

$$\text{find } w \in \mathscr{H} \text{ such that} \quad w \in \operatorname*{argmin}_{z\in\mathscr{H}} \left\{ \frac{|z - u|^2}{2\tau} + \psi_1(z; g) - \psi_{2,\varepsilon}(z) \right\}; \tag{4.67}$$

since $\psi_1(\cdot; g) - \psi_{2,\varepsilon_*} \ge \psi_1(\cdot; g) - \psi_{2,\varepsilon} \ge \phi(\cdot; g)$, assumption (COMP) ensures that the minimum problem (4.67) admits at least a solution $w_{\varepsilon,\tau}$, which fulfills by construction

$$\frac{|w_{\varepsilon,\tau} - u|^2}{2\tau} + \phi(w_{\varepsilon,\tau}; g) \le \frac{|w_{\varepsilon,\tau} - u|^2}{2\tau} + \psi_1(w_{\varepsilon,\tau}; g) - \psi_{2,\varepsilon}(w_{\varepsilon,\tau})$$
$$\le \psi_1(u; g) - \psi_{2,\varepsilon_*}(u).$$

Applying the estimate (4.12) of Lemma 4.1 to the functional $\psi_1(\cdot; g) - \psi_{2,\varepsilon}$, we deduce that for ε, τ sufficiently small we have

$$\max\{\phi(w_{\varepsilon,\tau}), |w_{\varepsilon,\tau}|\} \le M, \quad \frac{|w_{\varepsilon,\tau} - u|}{\sqrt\tau} \le M', \tag{4.68}$$

where $M' = C(1 + M + |g|^2)$. In particular, the sequence $\{w_{\varepsilon,\tau}\}_{\varepsilon,\tau}$ is relatively compact in $\mathscr{H}$ by (COMP). Moreover, since $w_{\varepsilon,\tau}$ complies with (4.67), there exists $\lambda^1_{\varepsilon,\tau} \in \partial\psi_1(w_{\varepsilon,\tau})$ fulfilling

$$\frac{w_{\varepsilon,\tau} - u}{\tau} + \lambda^1_{\varepsilon,\tau} - D\psi_{2,\varepsilon}(w_{\varepsilon,\tau}) = g, \tag{4.69}$$

whence $\partial\psi_1(w_{\varepsilon,\tau})$ is non empty. Then, we deduce from (4.61) that $\partial\psi_2(w_{\varepsilon,\tau}) = \partial\bar\psi_2(w_{\varepsilon,\tau}) \ne \emptyset$; setting

$$\lambda^2_{\varepsilon,\tau} := D\psi_{2,\varepsilon}(w_{\varepsilon,\tau}) \in \partial\bar\psi_2(J_\varepsilon(w_{\varepsilon,\tau})),$$

we get by (4.66) and (4.61)

$$|\lambda^2_{\varepsilon,\tau}| = |D\psi_{2,\varepsilon}(w_{\varepsilon,\tau})| \le |\partial^\circ\psi_2(w_{\varepsilon,\tau})| \le \rho|\lambda^1_{\varepsilon,\tau}| + \gamma. \tag{4.70}$$

Combining (4.69) and (4.70), we obtain

$$|\lambda^1_{\varepsilon,\tau}| \le \frac{|w_{\varepsilon,\tau} - u|}{\tau} + |D\psi_{2,\varepsilon}(w_{\varepsilon,\tau})| + |g| \le \frac{|w_{\varepsilon,\tau} - u|}{\tau} + |g| + \rho|\lambda^1_{\varepsilon,\tau}| + \gamma,$$

i.e.,

$$|\lambda^1_{\varepsilon,\tau}| \le (1 - \rho)^{-1}\Big(\frac{|w_{\varepsilon,\tau} - u|}{\tau} + |g| + \gamma \Big). \tag{4.71}$$

Taking into account (4.68), (4.70), and (4.71), we deduce that, for every $0 < \tau < \tau_*/8$, there exist w_τ, λ^1_τ, and $\lambda^2_\tau \in \mathscr{H}$ and a suitable subsequence $k \mapsto \varepsilon_k \downarrow 0$ such that

$$w_{\varepsilon_k,\tau} \to w_\tau, \quad \lambda^1_{\varepsilon_k,\tau} \rightharpoonup \lambda^1_\tau, \quad \lambda^2_{\varepsilon_k,\tau} \rightharpoonup \lambda^2_\tau \quad \text{as } k \uparrow +\infty.$$

We notice that $\lambda_\tau^2 \in \partial\bar{\psi}_2(w_\tau) = \partial\psi_2(w_\tau)$ by the strong-weak closedness of $\partial\bar{\psi}_2$ and by (4.65), that $\bar{\psi}_2(w_\tau) = \psi_2(w_\tau)$, and therefore w_τ is a minimizer of

$$v \mapsto \frac{1}{2\tau}|v - u|^2 + \phi(v; g) \quad v \in D(\phi), \tag{4.72}$$

so that

$$g - \tau^{-1}(w_\tau - u) \in \partial\phi(w_\tau).$$

Passing to the limit in (4.69) as $\varepsilon \downarrow 0$, we find

$$\frac{w_\tau - u}{\tau} + \lambda_\tau^1 - \lambda_\tau^2 = g, \tag{4.73}$$

so that λ_τ^1 is the sum of a vector in $\partial\phi(w_\tau)$ and a vector in $\partial\psi_2(w_\tau)$; since $\psi_1 = \phi + \psi_2$ in $D(\phi)$, we deduce that $\lambda_\tau^1 \in \partial\psi_1(w_\tau)$, too. The estimates (4.70)–(4.71) and (4.73) yield

$$\begin{cases} |\lambda_\tau^1| \leq (1 - \rho)^{-1}\left(\dfrac{|w_\tau - u|}{\tau} + |g| + \gamma\right), \\[2mm] |\lambda_\tau^2| \leq \rho(1 - \rho)^{-1}\left(\dfrac{|w_\tau - u|}{\tau} + |g| + \gamma\right) + \gamma. \end{cases} \tag{4.74}$$

Recalling (4.8), as $g \in \partial\phi(u)$ we have $\lim_{\tau \downarrow 0} \frac{w_\tau - u}{\tau} = 0$; therefore (4.74) yield, up to some subsequence,

$$\lambda_\tau^1 \rightharpoonup \lambda^1, \quad \lambda_\tau^2 \rightharpoonup \lambda^2 \in \partial\bar{\psi}_2(u) = \partial\psi_2(u), \quad \text{as } \tau \downarrow 0, \quad \lambda^1 - \lambda^2 = g \in \partial\phi(u) \tag{4.75}$$

with

$$|\lambda^1| \leq (1 - \rho)^{-1}(|g| + \gamma), \quad |\lambda^2| \leq \rho(1 - \rho)^{-1}(|g| + \gamma) + \gamma \leq (1 - \rho)^{-1}(|g| + \gamma) \tag{4.76}$$

Arguing as in the previous lines, (4.75) yields $\lambda^1 \in \partial\psi_1(u)$, and (4.62) follows. $\qquad\square$

Lemma 4.9. *Let us assume that $\phi : \mathscr{H} \to (-\infty, +\infty]$ admits the decomposition $\phi = \psi_1 - \psi_2$ as in Lemma 4.8, and let us suppose that (1.39) holds. Then for every $M \geq 0$ there exists a constant $C > 0$ such that every $g \in \partial_\ell\phi(u)$ with $\max(\phi(u), |u|) \leq M$ can be decomposed as*

$$g = \lambda^1 - \lambda^2, \quad \lambda^1 \in \partial_\ell\psi_1(u), \ \lambda^2 \in \partial\psi_2(u), \quad |\lambda^i| \leq C\left(1 + |g|\right), \ i = 1, 2. \tag{4.77}$$

In particular, if ψ_1 satisfies one of the Chain Rule conditions (CHAIN$_{1,2}$), then ϕ satisfies the same chain rule property, too.

Proof. By the previous lemma and the definition of limiting subdifferential, if $g \in \partial_\ell\phi(u)$, then we can find sequences $u_k \in D(\partial\phi)$, $g_k = \lambda_k^1 - \lambda_k^2 \in \partial\phi(u_k)$, $\lambda_k^i \in \partial\psi_i(u_k)$, $i = 1, 2$, $k \in \mathbb{N}$, such that

$$\begin{gathered} u_k \to u, \quad g_k = \lambda_k^1 - \lambda_k^2 \rightharpoonup g, \\ M' = \sup_k(\phi(u_k), |u_k|) < +\infty, \quad |\lambda_k^i| < C(M')(1 + |g_k|). \end{gathered} \tag{4.78}$$

Possibly extracting a subsequence, we thus have

$$\lambda_k^i \rightharpoonup \lambda^i, \ i = 1, 2, \quad \lambda^2 \in \partial\psi_2(u), \quad g = \lambda^1 - \lambda^2. \tag{4.79}$$

Moreover,

$$\limsup_{k \uparrow +\infty} \psi_2(u_k) \leq \limsup_{k \uparrow +\infty} \langle \lambda_k^2, u_k - u \rangle + \psi_2(u) = \psi_2(u) < +\infty$$

and therefore
$$\limsup_{k\uparrow+\infty} \psi_1(u_k) = \limsup_{k\uparrow+\infty} \big(\phi(u_k) + \psi_2(u_k)\big) \le M' + \psi_2(u) < +\infty.$$
It follows that $\lambda^1 \in \partial_\ell \psi_1(u)$ and (1.39) yields
$$|\lambda^1| \le |g| + |\lambda^2| \le |g| + \rho|\lambda^1| + \gamma, \quad i.e. \quad |\lambda^1| \le (1-\rho)^{-1}\big(|g| + \gamma\big), \tag{4.80}$$
which proves (4.77).

Let us now suppose that ψ_1 satisfies (CHAIN$_1$); in order to check that ϕ satisfies the same property, let us consider a curve $v \in H^1(0,T;\mathscr{H})$ such that $\phi(v(t)) = \varphi(t)$ for $a.e.$ $t \in (0,T)$, φ being of bounded variation, and a selection $\xi \in L^2(0,T;\mathscr{H})$ with $\xi(t) \in \partial_\ell \phi(v(t))$ for $a.e.$ $t \in (0,T)$. By (4.77), we get
$$\int_0^T |\partial_\ell^\circ \psi_1(v(t))|^2 \, \mathrm{d}t + \int_0^T |\partial^\circ \psi_2(v(t))|^2 \, \mathrm{d}t < +\infty. \tag{4.81}$$

Being ψ_2 convex, (4.81) shows that $\psi_2 \circ v$ is absolutely continuous with
$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_2(v(t)) = \langle \lambda^2, v'(t)\rangle \quad \forall \lambda^2 \in \partial \psi_2(v(t)) \quad \text{for } a.e. \ t \in (0,T). \tag{4.82}$$

Thus, $\varphi + \psi_2 \circ v$ is of bounded variation, and $\psi_1(v(t)) = \varphi(t) + \psi_2(v(t))$ for $a.e.$ $t \in (0,T)$; applying Theorem 3.3 we get
$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\varphi(t) + \psi_2(v(t))\Big) = \langle \lambda^1, v'(t)\rangle \quad \forall \lambda^1 \in \partial_\ell \psi_1(v(t)) \quad \text{for } a.e. \ t \in (0,T), \tag{4.83}$$
and therefore
$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = \langle \lambda^1 - \lambda^2, v'(t)\rangle \quad \forall \lambda^1 \in \partial_\ell \psi_1(v(t)), \ \lambda^2 \in \partial \psi_2(v(t)) \quad \text{for } a.e. \ t \in (0,T).$$

Applying (4.77) again we conclude; when ψ_1 satisfies (CHAIN$_2$), too, it is immediate to check that $\phi \circ u$ belongs to $AC(0,T)$. $\qquad\square$

5. Coupling diffusion equations with quasistationary conditions

5.1. An abstract setting

Let us consider a standard Hilbert triplet
$$V \subset H \equiv H' \subset V', \quad \text{the inclusions being dense and compact.} \tag{5.1}$$

Since we have adopted the usual convention of identifying H with its dual, the duality pairing between V and V' is the (unique) extension by continuity of the H-scalar product $(\cdot,\cdot)_H$, $i.e.$
$$_{V'}\langle u, v\rangle_V = (u,v)_H \quad \forall u \in H, v \in V. \tag{5.2}$$

We are given a *nonnegative, symmetric, and continuous* bilinear form $a : V \times V \to \mathbb{R}$, thus satisfying for some $\gamma > 0$
$$0 \le a(u) := a(u,u) \le \gamma\|u\|_V^2 \quad \forall u \in V, \tag{5.3}$$
and a proper functional $F : H \to [0,+\infty]$ whose sublevels
$$\big\{\chi \in H : F(\chi) \le s\big\} \quad \text{are } strongly \text{ compact in } H. \tag{5.4}$$

We denote by $A : V \to V'$ the continuous linear operator associated with a

$$_{V'}\langle Au, v \rangle_V = a(u, v) \quad \forall u, v, \in V, \tag{5.5}$$

and by $\Lambda : H \to 2^H$ the H-subdifferential of F, *i.e.*

$$\vartheta \in \Lambda\chi \quad \Leftrightarrow \quad \chi \in D(F) \subset H, \quad \liminf_{\|\eta - \chi\|_H \to 0} \frac{F(\eta) - F(\chi) - (\vartheta, \eta - \chi)_H}{\|\eta - \chi\|_H} \geq 0. \tag{5.6}$$

In this section, we aim at studying evolution problems of the following type

Problem 5.1 (Quasistationary evolution systems)**.** Given $u_0 \in H$ and $f : (0, T) \to V'$, find a pair $u, \chi : (0, T) \to H$, with $u - \chi \in V$, which satisfies at *a.e.* $t \in (0, T)$ the system

$$\begin{cases} u'(t) + A(u(t) - \chi(t)) = f(t) & \text{in } V', \\ \chi(t) + \Lambda\chi(t) \ni u(t) & \text{in } H, \\ u(0) = u_0. \end{cases} \tag{5.7}$$

We will distinguish two cases:

(1) The form a is *coercive, i.e.*, there exists $\mathsf{a} > 0$ s.t.

$$a(u) \geq \mathsf{a}\|u\|_V^2 \quad \forall u \in V; \tag{5.8}$$

(2) The form a is only *weakly* coercive, *i.e.*, there exists $\lambda, \mathsf{a} > 0$ such that

$$a(u) + \lambda\|u\|_H^2 \geq \mathsf{a}\|u\|_V^2 \quad \forall u \in V. \tag{5.9}$$

In fact, since the quadratic form $a(\cdot)$ is nonnegative, for every $\lambda > 0$ we can find $\mathsf{a} > 0$ such that (5.9) holds.

1) *The coercive case*

If (5.8) holds, we will denote by $\mathscr{H}$ the dual space

$$\mathscr{H} := V', \quad \text{endowed with the scalar product}$$
$$\langle u, v \rangle_{\mathscr{H}} := a(A^{-1}u, A^{-1}v) = {}_{V'}\langle u, A^{-1}v \rangle_V = {}_{V'}\langle v, A^{-1}u \rangle_V, \tag{5.10}$$

and we want to show that the system (5.7) is associated with the Gradient Flow in $\mathscr{H}$ of the functional ϕ defined by

$$\phi(u) := \inf_{\chi \in H} \mathscr{F}(u, \chi), \quad \mathscr{F}(u, \chi) := \begin{cases} \frac{1}{2}\|u - \chi\|_H^2 + F(\chi) & \text{if } u, \chi \in H, \\ +\infty & \text{otherwise.} \end{cases} \tag{5.11}$$

Remark 5.2 ($\mathscr{H}$ or V'?)**.** The spaces $\mathscr{H}$ and V' are linearly and topologically isomorphic, the sole possible difference being the choice of their scalar products. On the other hand, the notion of Gradient Flow is intrinsically related to this choice, and in the applications the scalar product (5.10) could be different from the one usually adopted in V'. Therefore, we will use the letter $\mathscr{H}$ when we want to stress the role of the particular scalar product and highlight the link with the general theory of Gradient Flows.

The connection with the previous theory is provided by the following preliminary result:

Proposition 5.3 (The subdifferential of the marginal function ϕ). *The functional $\phi : \mathscr{H} \to [0, +\infty]$ defined by* (5.11) *has domain $D(\phi) = H$ and satisfies* (COMP) *and* (CONT). *For every $u \in H$, the infimum in* (5.11) *is attained, and therefore for every $u \in H$ the set*

$$M(u) := \big\{ \chi \in H : \mathscr{F}(u, \chi) = \phi(u) \big\} \quad \text{is not empty,} \tag{5.12}$$

and satisfies

$$\chi \in M(u) \quad \Rightarrow \quad \chi + \Lambda\chi \ni u. \tag{5.13}$$

If $u \in D(\partial\phi)$ (the Fréchet subdifferential of ϕ in $\mathscr{H}$), $\partial\phi(u)$ is single-valued, and the set $M(u)$ contains a unique element χ, which fulfils

$$u - \chi \in V, \quad \partial\phi(u) = A(u - \chi) \in \mathscr{H} = V'. \tag{5.14}$$

For every $u \in D(\partial_\ell\phi)$ (the domain of the limiting subdifferential of ϕ in $\mathscr{H}$)

$$\xi \in \partial_\ell\phi(u) \quad \Rightarrow \quad \exists\chi \in M(u): \quad u - \chi \in V, \quad \xi = A(u - \chi). \tag{5.15}$$

Proof. Let us point out that for every positive constant C the sublevels

$$\{(u, \chi) \in \mathscr{H} \times H \ : \mathscr{F}(u, \chi) \leq C\}$$

are compact $\mathscr{H} \times H$, whence we easily get the existence of a minimizer $\chi \in M(u)$ of (5.11) for every $u \in H$. Since $\phi(u) = \mathscr{F}(u, \chi)$ for every $\chi \in M(u)$, it is immediate to check that the sublevels of ϕ are closed and bounded in H, hence compact in $\mathscr{H}$, so that ϕ satisfies (COMP). Then, (5.13) is simply the Euler equation satisfied by the minimizer of (5.11).

Let us now suppose that $u \in H, \xi \in \partial\phi(u)$, and let χ be *any* element in $M(u)$: observe that the minimality of χ yields

$$\phi(w) = \min_{\eta \in H} \frac{1}{2}\|\eta - w\|_H^2 + F(\eta) \leq \frac{1}{2}\|\chi - w\|_H^2 + F(\chi) = \frac{1}{2}\|\chi - w\|_H^2 - \frac{1}{2}\|\chi - u\|_H^2 + \phi(u),$$

so that for every $w \in H$ we have, by the very definition of subdifferential,

$$0 \leq \liminf_{w \to u} \frac{\phi(w) - \phi(u) - \langle \xi, w - u \rangle_{\mathscr{H}}}{|w - u|_{\mathscr{H}}}$$
$$\leq \liminf_{w \to u} \frac{\frac{1}{2}\|w - \chi\|_H^2 - \frac{1}{2}\|u - \chi\|_H^2 - \langle \xi, w - u \rangle_{\mathscr{H}}}{|w - u|_{\mathscr{H}}}.$$

Choosing $w := u + hz$, $h \in \mathbb{R}$, $0 \neq z \in H$, we get

$$0 \leq \liminf_{h \to 0} \frac{\frac{1}{2}h^2\|z\|_H^2 + h(u - \chi, z)_H + \frac{1}{2}\|u - \chi\|_H^2 - \frac{1}{2}\|u - \chi\|_H^2 - h\langle \xi, z \rangle_{\mathscr{H}}}{|h|\,|z|_{\mathscr{H}}}$$
$$= |z|_{\mathscr{H}}^{-1} \liminf_{h \to 0} \left(\frac{1}{2}|h|\,\|z\|_H^2 + \frac{h}{|h|}\big((u - \chi, z)_H - \langle \xi, z \rangle_{\mathscr{H}} \big) \right)$$
$$= -|z|_{\mathscr{H}}^{-1}\Big|(u - \chi, z)_H - \langle \xi, z \rangle_{\mathscr{H}}\Big|.$$

Being $z \in H$ and $\chi \in M(u)$ arbitrary, recalling (5.2) and (5.10), we get (5.14).

Finally, if $\xi \in \partial_\ell \phi(u)$, then the definition of limiting subdifferential and (5.14) show that there exist a constant C and sequences $u_k, \chi_k \in H,\ u_k - \chi_k \in V,\ \xi_k \in \mathscr{H}$ such that

$$\chi_k \in M(u_k), \quad \xi_k = A(u_k - \chi_k) = \partial\phi(u_k),$$
$$u_k \to u, \quad \xi_k \rightharpoonup \xi \quad \text{in } \mathscr{H}, \quad \phi(u_k) = \mathscr{F}(u_k, \chi_k) \leq C. \tag{5.16}$$

It follows that $u_k - \chi_k$ converges to $A^{-1}\xi$ weakly in V and strongly in H; being the sequence $\{\chi_k\}$ relatively compact in H, we get that $u_k \to u$ and $\chi_k \to \chi := u - A^{-1}\xi$ strongly in H. Then, we can pass to the limit in the family of inequalities

$$\mathscr{F}(u_k, \chi_k) \leq \mathscr{F}(u_k, \eta) \quad \forall \eta \in H,$$

obtaining

$$\mathscr{F}(u, \chi) \leq \mathscr{F}(u, \eta) \quad \forall \eta \in H, \qquad i.e.,\ \chi \in M(u).$$

Finally, the above inequalities yield

$$\limsup_{k \to +\infty} \phi(u_k) = \limsup_{k \to +\infty} \mathscr{F}(u_k, \chi_k) \leq \limsup_{k \to +\infty} \mathscr{F}(u_k, \chi) = \mathscr{F}(u, \chi) = \phi(u), \tag{5.17}$$

which shows that ϕ satisfies (CONT) and the approximating family χ_k satisfies

$$\lim_{k \to +\infty} F(\chi_k) = F(\chi). \tag{5.18}$$

$\square$

Corollary 5.4 (Gradient flows solve the system). *If $u \in H^1(0, T; \mathscr{H})$ is a solution of the Gradient Flow equation* (GF) *in $\mathscr{H}$ for the functional ϕ defined by (5.11) and the data $u_0 \in H, f \in L^2(0, T; \mathscr{H})$, then $u \in L^\infty(0, T; H)$ and there exists $\chi \in L^\infty(0, T; H)$ with*

$$u - \chi \in L^2(0, T; V), \quad \chi(t) \in M(u(t)) \quad for\ a.e.\ t \in (0, T), \tag{5.19}$$

such that the pair (u, χ) solves the system (5.7).

Applying Theorem 1, we readily get the following existence result:

Theorem 5.5. *Under the previous assumptions, let us suppose that for every $u \in H$ the set*

$$\left\{ \chi \in H : \chi \in M(u), \quad u - \chi \in V \right\} \quad \text{is convex.} \tag{5.20}$$

Then, Theorem 1 can be applied, and every generalized Minimizing Movement $u \in GMM(\Phi; u_0, f)$ with $u_0 \in H,\ f \in L^2(0, T; \mathscr{H})$ is a $H^1(0, T; \mathscr{H})$-solution of Problem 5.1, fulfilling the properties stated in Corollary 5.4.

Now, we investigate the possibility to prove a chain rule for the functional ϕ. Let us denote by G the functional

$$G(\chi) := \frac{1}{2}\|\chi\|_H^2 + F(\chi), \quad \chi \in H, \tag{5.21}$$

and by G^* its Legendre-Fenchel-Moreau conjugate (in H), *i.e.*,

$$G^*(u) := \sup_{\chi \in H} (u, \chi)_H - G(\chi) = \sup_{\chi \in H} (u, \chi)_H - \left(\frac{1}{2}\|\chi\|_H^2 + F(\chi) \right). \tag{5.22}$$

Of course, G^* is a convex functional defined in $D(\phi) = H$, and

$$G^*(u) \geq (u, \chi)_H - G(\chi) \quad \forall \chi \in H,$$
$$G^*(u) = (u, \chi)_H - G(\chi) \quad \text{iff} \quad \chi \in M(u). \tag{5.23}$$

Since ϕ admits the decomposition (*cf.* Rem. 1.9)

$$\phi(u) = \frac{1}{2}\|u\|_H^2 - G^*(u), \tag{5.24}$$

it is natural to check if Theorem 4 can be applied. The easiest case is provided by the following result:

Proposition 5.6. *Suppose that for every $M \geq 0$ there exist constants $\rho < 1, \gamma \geq 0$ such that the following* a priori estimate *holds:*

$$\left.\begin{array}{c} u \in V, \quad \chi \in M(u), \\ \max(\|u\|_H, F(\chi)) \leq M \end{array}\right\} \quad \Rightarrow \quad \chi \in V, \quad \|A\chi\|_{V'} \leq \rho\|Au\|_{V'} + \gamma. \tag{5.25}$$

Then, the decomposition (5.25) provides a dominated perturbation of $\psi_1 := \frac{1}{2}\|\cdot\|_H^2$. Thus, ϕ satisfies the Chain rule condition (CHAIN$_2$), *which in this case yields that $t \mapsto \mathscr{F}(u(t), \chi(t))$ is absolutely continuous with*

$$\frac{\mathrm{d}}{\mathrm{d}t}(\phi \circ u) = \frac{\mathrm{d}}{\mathrm{d}t}\mathscr{F}(u, \chi) = {}_{V'}\langle u', u - \chi \rangle_V \quad \textit{a.e. in } (0, T), \tag{5.26}$$

whenever

$$u \in H^1(0, T; V'), \quad \chi \in M(u), \quad \sup_{t \in [0,T]} \mathcal{F}(u, \chi) < +\infty, \quad u - \chi \in L^2(0, T; V). \tag{5.27}$$

Proof. Let us first observe that

$$\lambda^1 \in \partial\psi_1(u) \quad \Leftrightarrow \quad u \in V, \quad \lambda^1 = Au. \tag{5.28}$$

Thus, if $u \in D(\partial\psi_1)$, then also $u \in D(\partial\psi_2)$, since it is easy to check by (5.25) that

$$\chi \in M(u) \cap V \quad \Rightarrow \quad A\chi \in \partial G^*(u). \tag{5.29}$$

To see this, it is sufficient to notice that by (5.23) we have

$$\begin{aligned} G^*(v) - G^*(u) &\geq (v, \chi)_H - G(\chi) - \big((u, \chi)_H - G(\chi)\big) \\ &= (v - u, \chi)_H = \langle v - u, A\chi \rangle_{\mathscr{H}} \quad \forall v \in H. \end{aligned} \tag{5.30}$$

We want to show now that if $u \in V$, then

$$\lambda^2 \in \partial G^*(u) \quad \Rightarrow \quad \xi := A^{-1}\lambda^2 \in \overline{\mathrm{co}}\big(M(u)\big). \tag{5.31}$$

Indeed, recall that if $\lambda^2 \in \partial G^*(u)$, then

$$G^*(v) - G^*(u) \geq \langle \lambda^2, v - u \rangle_{\mathscr{H}} = (\xi, v - u)_H \quad \forall v \in H,$$

whence, denoting by G^{**} the lower semicontinuous and convex envelope of G in H, we get

$$(\xi, u)_H - G^*(u) \geq \sup_v (\xi, v)_H - G^*(v) = G^{**}(\xi). \tag{5.32}$$

Since G has compact sublevels in H and it has superlinear growth, there exists a Borel probability measure μ in H such that

$$\int_H \lambda \, \mathrm{d}\mu(\lambda) = \xi, \quad G^{**}(\xi) = \int_H G(\lambda) \, \mathrm{d}\mu(\lambda) \leq G(\xi) \tag{5.33}$$

(*cf.* [38], Th. 3.6 and (3.38), p. 419). Integrating (5.32), we obtain

$$G^*(u) \leq (\xi, u)_H - G^{**}(\xi) = \int_H ((\lambda, u)_H - G(\lambda)) \, \mathrm{d}\mu(\lambda). \tag{5.34}$$

Since $(\lambda, u)_H - G(\lambda) \leq G^*(u)$ for every $\lambda \in H$, the previous inequality shows that μ is concentrated on the set $M(u)$, which yields (5.31) in view of (5.33).

Further, (5.25) shows that

$$\|\lambda^2\|_{V'} = \|A\xi\|_{V'} \leq \int_{M(u)} \|A\chi\|_{V'} \, \mathrm{d}\mu(\chi) \leq \rho\|Au\|_{V'} + \gamma,$$

i.e., ϕ satisfies the assumptions of Theorem 4, and therefore the Chain Rule condition (CHAIN$_2$) (see Rems. 1.8 and 1.9). $\qquad\square$

The crucial assumption (5.25) in the previous proposition allows to obtain a separate control on u and χ in V from the estimate of $u - \chi$ in the same space. The simplest case in which this is possible is when $F(\chi)$ controls the norm of χ in V.

In some circumstances this does not occur: the next proposition shows that it is sufficient to control the norm of u and χ in a bigger space $W \supset V$, such that H lies "in the middle" between W and V'. This can be formalized in terms of Interpolation Theory, by asking that $(W, V')_{1/2,2} \subset H$ as in (5.37) (since $(V, V')_{1/2,2} = H$ (5.37) is in fact an identity, up to equivalent norms). In fact, the next result shows that for the "$W^{1,1}(0,T)$" chain rule the weaker assumption $(W, V')_{1/2,1} \subset H$ is sufficient.

Proposition 5.7. *Let us suppose that W is an intermediate Banach space between V and H (i.e. $V \subset W \subset H$, the inclusions being continuous), which satisfies*

$$(W, V')_{1/2,1} \subset H, \qquad \textit{e.g., for a suitable constant } K > 0 \qquad \|v\|_H \leq K \|v\|_W^{1/2} \|v\|_{V'}^{1/2} \quad \forall v \in W. \tag{5.35}$$

If for every constant $M \geq 0$ there exists a constant $C > 0$ such that the following a priori *estimate holds*

$$\left. \begin{array}{c} u - \chi \in V, \quad \chi \in M(u) \\ \max(\|u\|_H, F(\chi)) \leq M \end{array} \right\} \quad \Rightarrow \quad \|\chi\|_W \leq C \left(1 + \|A(u - \chi)\|_{V'}\right), \tag{5.36}$$

then ϕ satisfies the Chain rule condition (5.26), (5.27), which corresponds to the "$W^{1,1}(0,T)$" case of Remark 1.5. If, moreover, H satisfies the stronger interpolation property

$$(W, V')_{1/2,2} \subset H \tag{5.37}$$

then ϕ satisfies the Chain rule condition (CHAIN$_2$) in its stronger "$AC(0,T)$" form.

Proof. Let us suppose that $u \in H^1(0,T;V')$, $\xi \in L^2(0,T;V')$ with $\xi(t) \in \partial_\ell \phi(v(t))$ for *a.e.* $t \in (0,T)$ and $\sup_{t \in [0,T]} \phi(u(t)) < +\infty$; it follows from (5.15) that there exists $\chi \in L^\infty(0,T;H)$ with $\chi(t) \in M(u(t))$ and $\xi(t) = A(u(t) - \chi(t))$ for *a.e.* $t \in (0,T)$; let us first show that

$$\phi \circ u \in W^{1,1}(0,T), \quad \frac{\mathrm{d}}{\mathrm{d}t}(\phi \circ u) = \langle A(u - \chi), u' \rangle_{\mathscr{H}} \quad \textit{a.e. in } (0,T). \tag{5.38}$$

For $h \in \mathbb{R}, t + h \in (0,T)$, let $u_h := u(\cdot + h), \chi_h := \chi(\cdot + h)$: a simple computation shows that

$$\begin{aligned} \phi(u_h) - \phi(u) = \mathscr{F}(u_h, \chi_h) - \mathscr{F}(u, \chi) &\leq \mathscr{F}(u_h, \chi) - \mathscr{F}(u, \chi) = \frac{1}{2}\|u_h - \chi\|_H^2 - \frac{1}{2}\|u - \chi\|_H^2 \\ &= \frac{1}{2}(u_h - u, u + u_h - 2\chi)_H = (u_h - u, u - \chi)_H + \frac{1}{2}\|u_h - u\|_H^2, \end{aligned} \tag{5.39}$$

and, writing the same formula at the time $t - h$, we get

$$\phi(u) - \phi(u_{-h}) \leq (u - u_{-h}, u_{-h} - \chi_{-h})_H + \frac{1}{2}\|u - u_{-h}\|_H^2;$$

replacing $-h$ with h and inverting the sign, we end up with

$$\phi(u_h) - \phi(u) \geq (u_h - u, u_h - \chi_h)_H - \frac{1}{2}\|u_h - u\|_H^2.$$

If we subtract the term $h\langle A(u - \chi), u'\rangle_{\mathscr{H}}$ from both the previous inequalities and recall that $u - \chi, u_h - \chi_h \in V$, we obtain

$$\begin{aligned}
\langle u_h - u, A(u_h - \chi_h)\rangle_{\mathscr{H}} &- h\langle u', A(u - \chi)\rangle_{\mathscr{H}} - \frac{1}{2}\|u_h - u\|_H^2 \\
&\leq \phi(u_h) - \phi(u) - h\langle u', A(u - \chi)\rangle_{\mathscr{H}} \\
&\leq \langle u_h - u, A(u - \chi)\rangle_{\mathscr{H}} - h\langle u', A(u - \chi)\rangle_{\mathscr{H}} + \frac{1}{2}\|u_h - u\|_H^2,
\end{aligned} \tag{5.40}$$

hence

$$\begin{aligned}
\Big|\phi(u_h) - \phi(u) - h\langle u', A(u - \chi)\rangle_{\mathscr{H}}\Big| &\leq \Big|\langle u_h - u, A(u_h - \chi_h)\rangle_{\mathscr{H}} - h\langle u', A(u - \chi)\rangle_{\mathscr{H}}\Big| \\
&+ \Big|\langle u_h - u, A(u - \chi)\rangle_{\mathscr{H}} - h\langle u', A(u - \chi)\rangle_{\mathscr{H}}\Big| + \|u_h - u\|_H^2.
\end{aligned} \tag{5.41}$$

Now we choose $h > 0$, we divide by h, and we integrate on an interval $(0, T - \delta)$, $\delta > h$. Recalling that as $h \downarrow 0$

$$\frac{u_h - u}{h} \to u', \quad u_h - \chi_h \to u - \chi \quad \text{strongly in } L^2(0, T - \delta; \mathscr{H}),$$

we get

$$\limsup_{h\downarrow 0} \int_0^{T-\delta} \Big|\frac{\phi(u_h) - \phi(u)}{h} - \langle u', A(u - \chi)\rangle_{\mathscr{H}}\Big| \, \mathrm{d}t \leq \limsup_{h\downarrow 0} \frac{1}{h}\int_0^{T-\delta} \|u_h - u\|_H^2 \, \mathrm{d}t. \tag{5.42}$$

Since $u - \chi \in V$, $\chi \in M(u)$ a.e. in $(0, T)$ and $\|u\|_H, F(\chi)$ are uniformly bounded on $(0, T)$, from (5.36) we get $\chi \in L^2(0, T; W)$, and therefore $u \in L^2(0, T; W)$ as well, being V continuously imbedded in W.

The interpolation inequality (5.35) thus yields

$$\begin{aligned}
\frac{1}{h}\int_0^{T-\delta} \|u - u_h\|_H^2 \, \mathrm{d}t &\leq K^2 \int_0^{T-\delta} \Big\|\frac{u - u_h}{h}\Big\|_{V'} \|u - u_h\|_W \, \mathrm{d}t \\
&\leq C\|u'\|_{L^2(0,T;V')} \|u - u_h\|_{L^2(0,T;W)},
\end{aligned} \tag{5.43}$$

which goes to 0 as $h \downarrow 0$. Then, (5.38) readily follows from (5.42) and (5.43).

Finally, if (5.37) holds, then from $u \in L^2(0, T; W)$ with $u' \in L^2(0, T; V')$ we deduce $u \in C^0(0, T; H)$; it is easy to check from (5.39) and the lower semicontinuity of ϕ that the map $\phi \circ u$ is continuous and therefore it belongs to $AC(0, T)$. □

Theorem 5.8 (An existence result). *If either assumption* (5.25) *or assumptions* (5.36), (5.37) *of the above propositions are verified, then for every $u_0 \in H$, $f \in L^2(0, T; V')$ Problem 5.1 admits an "energy solution"*

$u \in H^1(0,T;V') \cap C^0([0,T];H)$, $\chi \in L^\infty(0,T;H)$ with $\vartheta = u - \chi \in L^2(0,T;V)$, which satisfies a.e. in $(0,T)$ (here $a(v) := a(v,v)$)

$$\chi \in M(u), \quad \text{which yields } \chi + \Lambda\chi \ni u, \tag{5.44}$$

$$a(u - \chi - A^{-1}f) = \min_{\sigma \in M(u)} a(u - \sigma - A^{-1}f), \tag{5.45}$$

and the energy identity

$$\int_s^t a(\vartheta(r))\, dr + \mathscr{F}(u(t),\chi(t)) = \mathscr{F}(u(s),\chi(s)) + \int_s^t {}_{V'}\langle f(r),\vartheta(r)\rangle_V\, dr \quad \forall s,t \in [0,T]. \tag{5.46}$$

Moreover, $u \in C^0([0,T];H)$.

The *proof* follows directly from Theorem 3 by elementary computations.

2) *The weakly coercive case*

When the operator A is no more injective and only (5.9) holds true, we can argue by approximation and recover exactly the same result as before.

Theorem 5.9. *Theorem 5.8 also holds (except for (5.45)) in the weakly coercive case.*

Proof. Let us consider the bilinear forms

$$a_\lambda(u,v) := a(u,v) + \lambda(u,v)_H \quad \forall u,v \in V$$

which are coercive for every $\lambda > 0$; Theorem 5.8 yields the existence of an energy solution $(u_\lambda, \chi_\lambda)$ to Problem 5.1; since

$$\mathscr{F}(u_\lambda, \chi_\lambda) \geq \frac{1}{2}\|u_\lambda - \chi_\lambda\|_H^2,$$

we obtain the *a priori bounds*

$$\|u_\lambda - \chi_\lambda\|_{L^2(0,T;V)} + \|u_\lambda'\|_{L^2(0,T;V')} + \sup_{[0,T]} \mathscr{F}(u_\lambda, \chi_\lambda) \leq C$$

with C independent of $\lambda > 0$. Thus we can pass to the limit (up to extracting a suitable vanishing subsequence λ_k) as $\lambda_k \downarrow 0$, obtaining

$$u_{\lambda_k} \rightharpoonup u \quad \text{in } H^1(0,T;V'), \quad u_{\lambda_k} - \chi_{\lambda_k} \rightharpoonup u - \bar\chi \quad \text{in } L^2(0,T;V)$$

$$\chi_{\lambda_k} \rightharpoonup^* \bar\chi \quad \text{in } L^\infty(0,T;H), \quad u_{\lambda_k}(t) \rightharpoonup u(t) \quad \text{in } H \quad \forall t \in [0,T],$$

$$u' + A(u - \bar\chi) = f \quad \text{a.e. in } (0,T).$$

In order to show the strong convergence of $\{A(u_{\lambda_k} - \chi_{\lambda_k})\}$ in $L^2(0,T;V')$, we argue as in the proof of Theorem 3, trying to get more information from the energy inequality. In fact, (5.46) yields

$$\int_0^t a(u(r) - \bar\chi(r))\, dr + \phi(u(t)) \leq \limsup_{k\uparrow+\infty} \int_0^t a_{\lambda_k}(u_{\lambda_k}(r) - \chi_{\lambda_k}(r))\, dr + \phi(u_{\lambda_k}(t))$$

$$\leq \phi(u_0) + \int_0^t {}_{V'}\langle f(r), u(r) - \bar\chi(r)\rangle_V\, dr. \tag{5.47}$$

On the other hand, arguing as in Theorem 3.3, the Chain rule (5.26) shows that $t \mapsto \phi(u(t))$ is in $AC(0,T)$ and

$$\frac{d}{dt}(\phi \circ u) = {}_{V'}\langle u', u - \bar\chi\rangle_V = {}_{V'}\langle f - A(u - \bar\chi), u - \bar\chi\rangle_V,$$

so that

$$\phi(u_0) = \phi(u(t)) - \int_0^t {}_{V'}\langle f(r) - A(u(r) - \bar{\chi}(r)), u(r) - \bar{\chi}(r)\rangle_V \, \mathrm{d}r$$

at each point $t \in [0, T]$. Substituting the identity in (5.47) and using the lower semicontinuity of each term of the sum, we obtain

$$A(u_{\lambda_k} - \chi_{\lambda_k}) \to A(u - \bar{\chi}) \quad \text{strongly in } L^2(0, T; V'), \qquad \phi(u_{\lambda_k}(t)) \to \phi(u(t)) \quad \forall t \in [0, T]. \tag{5.48}$$

Therefore, up to the further extraction of a subsequence (still denoted by λ_k), there exists a Borel negligible set $\mathcal{N} \subset (0, T)$ such that $A(u_{\lambda_k} - \chi_{\lambda_k})$ converges to $A(u - \bar{\chi})$ in V' in $(0, T) \setminus \mathcal{N}$.

Let now $\boldsymbol{L}(t)$ be the (not empty) set of all (strong) limit points of χ_{λ_k} in H

$$\boldsymbol{L}(t) := \bigcap_{p=1}^{\infty} \overline{\{\chi_{\lambda_k}(t) : k \geq p\}}^H, \tag{5.49}$$

and let us choose $t \in (0, T) \setminus \mathcal{N}$ and $\chi \in \boldsymbol{L}(t)$; there exists a vanishing subsequence $\lambda_{k'} \downarrow 0$ (possibily depending on t) such that $\chi_{\lambda_{k'}}(t) \to \chi$ in H. From the convergence of the energies (5.48) it is immediate to check that $u_{\lambda_{k'}}(t) - \chi_{\lambda_{k'}}(t) \to u(t) - \chi$ in H and $\chi \in M(u(t))$. Further, as $A(u_{\lambda_{k'}}(t) - \chi_{\lambda_{k'}}(t))$ converges in V', the weak coercivity yields $u_{\lambda_{k'}}(t) - \chi_{\lambda_{k'}}(t) \to u(t) - \chi$ in V. We deduce that $A(u(t) - \bar{\chi}(t)) = A(u(t) - \chi)$. Operating a measurable selection $\chi(t) \in \boldsymbol{L}(t)$, we conclude. $\qquad \square$

5.2. Applications

We conclude this paper by briefly showing that the PDE Examples discussed in Section 2 are particular cases of Problem 5.1: in each situation, the application fits in the framework we have proposed in the first part of this section.

The Stefan-Gibbs-Thomson problem. Let us first consider Example 5. We have already discussed the functional setting: $V' = H^{-1}(\Omega)$, $A := -\Delta$, $V := H_0^1(\Omega)$ and

$$F_0(\chi) := \alpha \int_\Omega |D\chi| + \int_\Omega I_{\{-1,1\}}(\chi) \, \mathrm{d}x, \quad \mathscr{F}_0(u, \chi) := \frac{1}{2}\|u - \chi\|_{L^2(\Omega)}^2 + F_0(\chi),$$
$$\phi_0(u) := \min_\chi \mathscr{F}_0(u, \chi).$$

We want to show that in this situation we can apply Theorem 5.5.

Here the crucial remark, which is somehow hidden in the original proof by Luckhaus and was also used in a different form by [35], is that for every $u \in H$ the set

$$\left\{\chi \in H : \chi \in M_0(u), \quad u - \chi \in H_0^1(\Omega)\right\} \quad \text{is a singleton,}$$

so that (5.20) trivially holds.

In fact, if $\chi_i \in M(u)$ and $u - \chi_i \in H_0^1(\Omega)$, $i = 1, 2$, we have in particular $\chi_1 - \chi_2 \in H_0^1(\Omega)$; on the other hand, since $\chi_i(x) \in \{-1, 1\}$ for *a.e.* $x \in \Omega$, the range of values of $\chi_2 - \chi_1$ is $\{0, 2, -2\}$. Hence, $\chi_2 - \chi_1$ is necessarily constant (an H^1-function on a connected domain cannot have jumps), and null at the boundary, thus $\chi_1 = \chi_2$.

Finally, we have to check that the elements in the limiting subdifferential of ϕ_0 satisfy the Gibbs-Thomson condition in the weak form (2.45): this important property is stated by the following lemma:

Lemma 5.10 (The Gibbs-Thomson condition induced by $\partial_\ell \phi_0$). *Let us suppose that*

$$\xi = -\Delta\vartheta \in \partial_\ell \phi_0(u) \quad \text{with} \quad \vartheta = u - \chi \in H_0^1(\Omega), \quad \chi \in M_0(u). \tag{5.50}$$

Then on the essential boundary $S = \partial^* E^{\pm}$ separating the two phases $E^{\pm} := \{x \in \Omega : \chi(x) = \pm 1\}$ the Gibbs-Thomson condition $\boldsymbol{H} = \vartheta\boldsymbol{\nu}$ holds in the weak form

$$\alpha \int_\Omega \left(\operatorname{div} \boldsymbol{\zeta} - \boldsymbol{\nu}^T D\boldsymbol{\zeta}\,\boldsymbol{\nu} \right) \mathrm{d}|D\chi| = \int_\Omega \operatorname{div}(\vartheta\boldsymbol{\zeta})\chi \,\mathrm{d}x, \qquad \forall \boldsymbol{\zeta} \in C^2(\overline{\Omega}; \mathbb{R}^m) \quad \boldsymbol{\zeta} \cdot \boldsymbol{n} = 0 \quad on\ \partial\Omega, \tag{5.51}$$

where the Radon-Nikodym derivative $\boldsymbol{\nu} = \frac{\mathrm{d}(D\chi)}{\mathrm{d}|D\chi|}$ is the measure theoretic inner normal to $\partial^* E^+$.

Proof. Let us first suppose that ξ in (5.50) belongs to the *Fréchet subdifferential* $\partial\phi_0(u)$. We introduce the flow $\boldsymbol{X}_s : \Omega \to \Omega$ associated with $\boldsymbol{\zeta}$, *i.e.* the family of diffeomeomorphisms satisfying the system of ODE for $(s, x) \in \mathbb{R} \times \Omega$

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{X}_s(x) = \boldsymbol{\zeta}(\boldsymbol{X}_s(x)) \\ \boldsymbol{X}_0(x) = x. \end{cases} \tag{5.52}$$

When Ω is sufficiently regular, condition $\boldsymbol{\zeta} \cdot \boldsymbol{n} = 0$ ensures that Ω is an invariant region for $\boldsymbol{X}_s$, *i.e.*

$$\boldsymbol{X}_s(\Omega) = \Omega \quad \forall s \in \mathbb{R}, \tag{5.53}$$

and the map $x \mapsto \boldsymbol{X}_s(x)$ is a C^2 diffeomeomorphism with inverse $\boldsymbol{X}_{-s}$; setting

$$\mathsf{D}_s(x) := D_x \boldsymbol{X}_s(x), \quad J_s(x) := \det \mathsf{D}_s(x) \tag{5.54}$$

we have

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}s}\mathsf{D}_s(x) = D\boldsymbol{\zeta}(\boldsymbol{X}_s(x))\mathsf{D}_s(x), \\ D_0(x) = I, \end{cases} \qquad \begin{cases} \frac{\mathrm{d}}{\mathrm{d}s}J_s(x) = \operatorname{div}\boldsymbol{\zeta}(\boldsymbol{X}_s(x))J_s(x), \\ J_0(x) = 1. \end{cases} \tag{5.55}$$

We consider a perturbation of (u, ϑ, χ) given by

$$\chi_s(x) := \chi(\boldsymbol{X}_{-s}(x)), \quad \vartheta_s(x) := \vartheta(x), \quad u_s := \vartheta + \chi_s; \tag{5.56}$$

since χ_s still belongs to $BV(\Omega; \{-1, 1\})$ and $\frac{1}{2}|D\chi_s|$ coincides with the $(m-1)$-dimensional Hausdorff measure restricted to $S_s = \boldsymbol{X}_s(S)$, the first variation formula for the area functional (see [44], Chap. 16 and [2], Th. 7.31), yields

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\int_\Omega |D\chi_s| \right)_{s=0} = \int_\Omega \left(\operatorname{div}\boldsymbol{\zeta} - \boldsymbol{\nu}^T D\boldsymbol{\zeta}\,\boldsymbol{\nu} \right) \mathrm{d}|D\chi|. \tag{5.57}$$

Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\int_\Omega \vartheta(x)\,u_s(x)\,\mathrm{d}x \right)_{s=0} = \frac{\mathrm{d}}{\mathrm{d}s}\left(\int_\Omega \vartheta(x)\,\chi_s(x)\,\mathrm{d}x \right)_{s=0} = \int_\Omega \operatorname{div}\left(\vartheta\,\boldsymbol{\zeta} \right)\chi \,\mathrm{d}x \tag{5.58}$$

since $u_s - u = \chi_s - \chi$ and the change of variable formula yields

$$\int_\Omega \vartheta(x)\,\chi_s(x)\,\mathrm{d}x = \int_\Omega \vartheta(\boldsymbol{X}_s(y))J_s(y)\,\chi(y)\,\mathrm{d}y.$$

By the very definition of subdifferential, we have

$$\phi_0(u_s) - \phi_0(u) - \int_\Omega \vartheta\,(u_s - u)\,\mathrm{d}x = \phi_0(u_s) - \phi_0(u) - \left\langle \xi, |u_s - u| \right\rangle_{H^{-1}(\Omega)}$$
$$\geq o\left(\|u_s - u\|_{H^{-1}(\Omega)} \right) = o\left(\|\chi_s - \chi\|_{H^{-1}(\Omega)} \right) = o(s) \tag{5.59}$$

where according to Landau's notation, $o(s)$ denotes a function depending on s such that $o(s)/s \to 0$ as $s \to 0$.

On the other hand, since $u_s - \chi_s = u - \chi = \vartheta$, we have

$$\alpha \int_\Omega |D\chi_s| - \alpha \int_\Omega |D\chi| = \mathscr{F}_0(u_s, \chi_s) - \mathscr{F}_0(u, \chi) \geq \phi_0(u_s) - \phi_0(u);$$

dividing (5.59) by s and passing to the limit first as $s \downarrow 0_+$ and then as $s \uparrow 0_-$ we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s}\Big(\alpha \int_\Omega |D\chi_s|\Big)_{s=0} - \int_\Omega \vartheta(x)\,\chi_s(x)\,\mathrm{d}x\Big)_{s=0} = 0,$$

which yields (5.51).

Finally, we have to show that (5.51) holds even if ξ belongs to the *limiting* subdifferential of ϕ_0 at u. In this case, by Proposition 5.3 and (5.18), we know that there exists a sequence $(u_k, \vartheta_k, \chi_k)_{k\in\mathbb{N}}$ which satisfies (5.50), $-\Delta\vartheta_k = \partial\phi_0(u_k)$, and (5.51), such that

$$\vartheta_k \rightharpoonup \vartheta \text{ in } H_0^1(\Omega), \quad \chi_k \to \chi \text{ in } L^2(\Omega), \quad \lim_{k\to\infty} \alpha \int_\Omega |D\chi_k| = \alpha \int_\Omega |D\chi|. \tag{5.60}$$

Since $D\chi_k \rightharpoonup^* D\chi$ and $|D\chi_k|(\Omega) \to |D\chi|(\Omega)$, we can pass to the limit in (5.51) as $k \to +\infty$ thanks to the Reshetnyak theorem [2], Theorem 2.39. $\qquad\square$

Quasistationary phase field with Dirichlet/Neumann boundary conditions. Here, we consider the system (2.31, 2.32) introduced in Example 4, with the boundary conditions

$$u - \chi = 0, \quad A_2\nabla\chi \cdot \boldsymbol{n} = 0 \quad \text{in } \partial\Omega \times (0, T). \tag{5.61}$$

Again $V := H_0^1(\Omega)$, $H := L^2(\Omega)$, $V' := H^{-1}(\Omega)$, A is the elliptic operator $-\operatorname{div}(\mathsf{A}_1\nabla\cdot)$, and $\mathscr{F}, \phi$ are given by (2.34) and (2.35), respectively.

In this case, we can apply Theorem 5.7 with the choice $W := H^1(\Omega)$. Observe that (5.61) is a well-known interpolation estimate of $L^2(\Omega)$ between $H^{-1}(\Omega)$ and $H^1(\Omega)$, whereas (5.36) is immediate since the functional F itself controls the $H^1(\Omega)$- norm of χ.

Quasistationary phase field with Neumann boundary conditions. Finally, the existence of the solutions to the system (2.31, 2.32), coupled with boundary conditions of variational type

$$\mathsf{A}_1\nabla(u - \chi) \cdot \boldsymbol{n} = 0, \quad \mathsf{A}_2\nabla\chi \cdot \boldsymbol{n} = 0 \qquad\qquad \text{in } \partial\Omega \times (0, T), \tag{5.62}$$

for $u_0 \in L^2(\Omega)$ and $f \in L^2(\Omega \times (0, T))$, follows directly from Theorem 5.9 with the choice $H := L^2(\Omega)$, $V := H^1(\Omega)$; the operator A is defined for every $u, v \in H^1(\Omega)$ as

$$_{(H^1(\Omega))'}\langle Au, v\rangle_{H^1(\Omega)} := \int_\Omega \Big(A_1(x)\nabla u(x) \cdot \nabla v(x)\Big)\,\mathrm{d}x. \tag{5.63}$$

REFERENCES

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

[2] L. Ambrosio, N. Fusco and D. Pallara, *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs, Clarendon Press, Oxford (2000).

[3] 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 (2005).

[4] C. Baiocchi, *Discretization of evolution variational inequalities*, 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) 59–92.

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

[6] E.J. Balder, An extension of Prohorov's theorem for transition probabilities with applications to infinite-dimensional lower closure problems. *Rend. Circ. Mat. Palermo* **34** (1985) 427–447.

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

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

[9] G. Bouchitté, Singular perturbations of variational problems arising from a two-phase transition model. *Appl. Math. Optim.* **21** (1990) 289–314.

[10] A. Bressan, A. Cellina and G. Colombo, Upper semicontinuous differential inclusions without convexity. *Proc. Amer. Math. Soc.* **106** (1989) 771–775.

[11] H. Brézis, *Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations*, Contribution to Nonlinear Functional Analysis, in *Proc. Sympos. Math. Res. Center, Univ. Wisconsin, Madison, 1971*. Academic Press, New York (1971) 101–156.

[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] H. Brézis, *Analyse fonctionnelle - Théorie et applications*. Masson, Paris (1983).

[14] H. Brézis, *On some degenerate nonlinear parabolic equations*, Nonlinear Functional Analysis (*Proc. Sympos. Pure Math., Vol. XVIII, Part 1, Chicago, Ill., 1968*), Providence, R.I., *Amer. Math. Soc.* (1970) 28–38.

[15] T. Cardinali, G. Colombo, F. Papalini and M. Tosques, On a class of evolution equations without convexity. *Nonlinear Anal.* **28** (1997) 217–234.

[16] C. Castaing and M. Valadier, *Convex analysis and measurable multifunctions*. Springer, Berlin-New York (1977).

[17] M.G. Crandall and T.M. Liggett, Generation of semi-groups of nonlinear transformations on general Banach spaces. *Amer. J. Math.* **93** (1971) 265–298.

[18] M.G. Crandall and A. Pazy, Semi-groups of nonlinear contractions and dissipative sets. *J. Functional Anal.* **3** (1969) 376–418.

[19] E. De Giorgi, *New problems on minimizing movements*, Boundary Value Problems for PDE and Applications, C. Baiocchi and J.L. Lions, Eds., Masson (1993) 81–98.

[20] 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) 180–187.

[21] C. Dellacherie and P.A. Meyer, *Probabilities and potential*. North-Holland Publishing Co., Amsterdam (1978).

[22] R. Jordan, D. Kinderlehrer and F. Otto, The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29** (1998) 1–17 (electronic).

[23] T. Kato, *Perturbation theory for linear operators*. Springer, Berlin (1976).

[24] Y. Kōmura, Nonlinear semi-groups in Hilbert space. *J. Math. Soc. Japan* **19** (1967) 493–507.

[25] A.Ja. Kruger and B.Sh. Mordukhovich, Extremal points and the Euler equation in nonsmooth optimization problems. *Dokl. Akad. Nauk BSSR* **24** (1980) 684–687, 763.

[26] S. Luckhaus, Solutions for the two-phase Stefan problem with the Gibbs-Thomson Law for the melting temperature. *Euro. J. Appl. Math.* **1** (1990) 101–111.

[27] S. Luckhaus, *The Stefan Problem with the Gibbs-Thomson law*. Preprint No. 591 Università di Pisa (1991) 1–21.

[28] S. Luckhaus, The Gibbs-Thompson relation within the gradient theory of phase transitions. *Arch. Rational Mech. Anal.* **107** (1989) 71–83.

[29] A. Marino, C. Saccon and M. Tosques, Curves of maximal slope and parabolic variational inequalities on nonconvex constraints. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **16** (1989) 281–330.

[30] A. Mielke, F. Theil and V.I. Levitas, A variational formulation of rate-independent phase transformations using an extremum principle. *Arch. Ration. Mech. Anal.* **162** (2002) 137–177.

[31] L. Modica, Gradient theory of phase transitions and minimal interface criterion. *Arch. Rational Mech. Anal.* **98** (1986) 123–142.

[32] L. Modica and S. Mortola, Un esempio di Γ-convergenza. *Boll. Un. Mat. Ital. B* **14** (1977) 285–299.

[33] B.Sh. Mordukhovich, Nonsmooth analysis with nonconvex generalized differentials and conjugate mappings. *Dokl. Akad. Nauk BSSR* **28** (1984) 976–979.

[34] 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) 525–589.

[35] P.I. Plotnikov and V.N. Starovoitov, The Stefan problem with surface tension as the limit of a phase field model. *Differential Equations* **29** (1993) 395–404.

[36] R.T. Rockafellar, *Convex analysis*. Princeton University Press, Princeton (1970).

[37] R.T. Rockafellar and R.J.B. Wets, *Variational analysis*. Springer-Verlag, Berlin (1998).

[38] R. Rossi and G. Savaré, Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Sup., Pisa* **2** (2003) 395–431.

[39] R. Rossi and G. Savaré, Existence and approximation results for gradient flows. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (9) Mat. Appl.* **15** (2004) 183–196.

[40] J. Rulla, Error analysis for implicit approximations to solutions to Cauchy problems. *SIAM J. Numer. Anal.* **33** (1996) 68–87.

[41] G. Savaré, Weak solutions and maximal regularity for abstract evolution inequalities. *Adv. Math. Sci. Appl.* **6** (1996) 377–418.

[42] G. Savaré, Compactness properties for families of quasistationary solutions of some evolution equations. *Trans. Amer. Math. Soc.* **354** (2002) 3703–3722.

[43] R. Schätzle, The quasistationary phase field equations with Neumann boundary conditions. *J. Differential Equations* **162** (2000) 473–503.

[44] L. Simon, Lectures on geometric measure theory, in *Proc. Centre for Math. Anal., Australian Nat. Univ.* **3** (1983).

[45] M. Valadier, *Young measures*, Methods of nonconvex analysis (Varenna, 1989). Springer, Berlin (1990) 152–188.

[46] A. Visintin, Differential models of hysteresis. *Appl. Math. Sci.* **111**, Springer-Verlag, Berlin (1994).

[47] A. Visintin, Models of phase transitions. *Progress in Nonlinear Differential Equations and Their Applications* **28**, Birkhäuser, Boston (1996).

[48] A. Visintin, Forward-backward parabolic equations and hysteresis. *Calc. Var. Partial Differential Equations* **15** (2002) 115–132.