Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Weighted energy-dissipation functionals for gradient flows

Alexander Mielke[1,2] and Ulisse Stefanelli[3]

submitted: 6th February 2009

[1] Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstraße 39
10117 Berlin, Germany
E-Mail: mielke@wias-berlin.de
http://www.wias-berlin.de/people/mielke/

[2] Institut für Mathematik
Humboldt-Universität zu Berlin
Rudower Chaussee 25
12489 Berlin-Adlershof, Germany

[3] IMATI - CNR
v. Ferrata 1
I-27100 Pavia, Italy
E-Mail: ulisse.stefanelli@imati.cnr.it
http://www.imati.cnr.it/ulisse/

No. 1400
Berlin 2009

2000 *Mathematics Subject Classification.* 35K55.

Key words and phrases. Variational principle, gradient flow, convergence.

The authors wish to thank Sir John Ball for proposing the issues eventually developed in Subsection "Cauchy argument". A. Mielke was partially supported by the DFG Resarch Unit FOR 797 *Analysis and Computation of Microstructures in Finite Plasticity* under Mie 459/5-1. U. Stefanelli was partially supported by the FP7-IDEAS-ERC-StG Grant #200497 *BioSMA: Mathematics for Shape Memory Technologies in Biomechanics* and gratefully acknowledges the hospitality of WIAS..

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract. We investigate a global-in-time variational approach to abstract evolution by means of the *weighted energy-dissipation* functionals proposed by Mielke & Ortiz [MO08]. In particular, we focus on gradient flows in Hilbert spaces. The main result is the convergence of minimizers and approximate minimizers of these functionals to the unique solution of the gradient flow. Sharp convergence rates are provided and the convergence analysis is combined with time-discretization. Applications of the theory to various classes of parabolic PDE problems are presented. In particular, we focus on two examples of microstructure evolution from [CO08].

1. Introduction

Assume we are given a real Hilbert space H with scalar product $(\cdot,\cdot)$ and corresponding norm $|\cdot|$. Moreover, let the functional $\phi : H \to (-\infty,\infty]$ be proper, lower semicontinuous, bounded from below, and λ-convex for some $\lambda \in \mathbb{R}$, i.e., $u \mapsto \phi(u) - (\lambda/2)|u|^2$ is convex. Finally, let $f \in L^2(0,T;H)$, and $u_0 \in D(\phi) \doteq \{u \in H \ : \ \phi(u) < \infty\}$. This note is concerned with the classical *gradient flow*

$$u' + \partial\phi(u) \ni f \quad \text{a.e. in} \ \ (0,T), \quad u(0) = u_0. \tag{1.1}$$

Gradient flows can be regarded as the paradigm of dissipative evolution. They arise almost ubiquitously in connection with applications and have hence attracted a constant attention during the last four decades starting from the fundamental work by Kōmura [Kōm67], Crandall-Pazy[CP69], and Brezis [Bre71, Bre73b]. It is beyond our purposes to even attempt to review the huge existing literature on gradient flows. Let us however mention that, even restricting to the present quite classical setting [Bre73b], relation (1.1) stems in a variety of different applications such as heat conduction, the Stefan problem, the Hele-Shaw cell, porous media, parabolic variational inequalities, some classes of ODEs with obstacles, degenerate parabolic PDEs, and the mean curvature flow for Cartesian graphs, among many others [NSV00], see Section 7 below. More recently, following the pioneering work by Otto [Ott01], an even larger class of PDE problems have been translated into gradient flows by resorting to probability spaces endowed with the Wasserstein metric. The reader is referred to the recent monograph by Ambrosio, Gigli, & Savaré [AGS05] for a collection of results (let us however stress that the metric theory is beyond the reach of the analysis presented here).

The general gradient-flow theory, although quite developed, is however not yet providing a sound description of the evolution of nonlinear systems that develop evolving microstructures. For these systems, the *energy* ϕ is generally not lower semicontinuous and equilibrium states which minimize ϕ do not exist. At the stationary level, a classical solution to this obstruction is the *relaxation* of the functional ϕ. Namely, one changes ϕ with its lower semicontinuous envelope $\mathrm{sc}^-\phi$ and interprets the respective minimization as an *effective* or macroscopic problem. In the evolution case, the natural idea would be to introduce a functional on *entire trajectories* whose minimizers solve the gradient flow (1.1) and consider its relaxation. Moving from these considerations Mielke & Ortiz [MO08] introduced a variational reformulation of evolution problems as (limits of) minimizers of a class of global-in-time functionals. These functionals feature the sum of the (scaled) energy and the dissipation, integrated in time via an exponentially decaying weight. The resulting so-called *weighted energy-dissipation* (WED) functionals

$I_\varepsilon : H^1(0,T;H) \to (-\infty,\infty]$ read, in the case of the gradient flow (1.1), as

$$I_\varepsilon(v) \doteq \int_0^T \mathrm{e}^{-t/\varepsilon} \left(\frac{1}{2}|v'|^2 + \frac{1}{\varepsilon}\Big(\phi(v) - (f,v)\Big)\right) \mathrm{d}t. \tag{1.2}$$

We will check in Subsection 2.4 that, for all ε small, the functional I_ε admits a unique minimizer in the closed convex set $K(u_{0\varepsilon}) \doteq \{v \in H^1(0,T;H) \ : \ v(0) = u_{0\varepsilon}\}$ where $u_{0\varepsilon}$ is a suitable approximation of u_0 (see below).

The WED functional approach has been originally applied in [MO08] to the description of rate-independent evolution, which, roughly speaking, corresponds to replacing 2 by 1 in (1.2). Later on, the analysis of the rate-independent case has been extended and adapted to time-discretizations in [MS08].

As for the gradient flow situation, a discussion on a linear case is contained in [MO08] together with a first example of relaxation. More recently, two additional examples of relaxation related with micro-structure evolution have been provided by Conti & Ortiz [CO08], see Section 7. In the above-mentioned papers, the problem of proving the convergence $u_\varepsilon \to u$ is left open. This question is solved here and our main result reads as follows.

Theorem 1.1 (Convergence). *$u_\varepsilon \to u$ uniformly in H.*

In the easiest possible setting, namely the scalar and linear case of

$$H = \mathbb{R}, \quad \phi(u) = -u^2/2, \quad f = 0, \quad u_0 = 1, \quad T = 1, \tag{1.3}$$

the convergence result of Theorem 1.1 is illustrated in Figure 1.

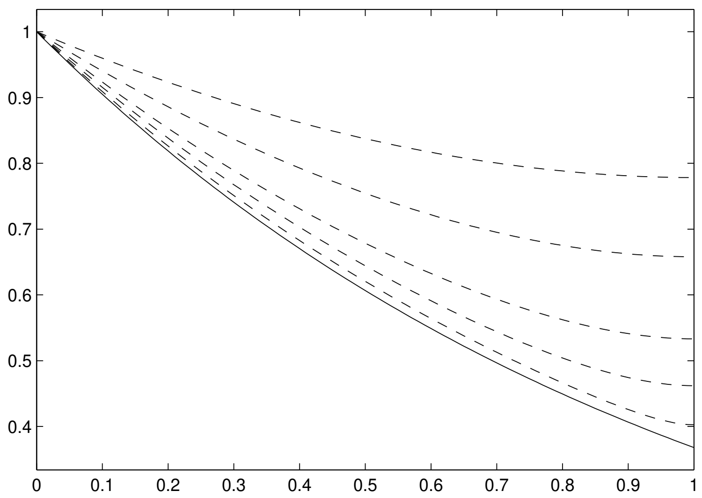

Figure 1: Convergence in the special case (1.3). As $\varepsilon \to 0$, the minimizers of I_ε (dashed lines) approach the solution of the gradient flow (solid line). Note that minimizers fulfill the *artificial* homogeneous Neumann boundary condition at T (see (1.4c)).

Besides the stated uniform convergence, much more is true for we are in the position of providing a quantitative statement, even in finer topologies (see Subsection 5.1 below). Moreover, the assumptions on u_0 can be substantially weakened (Subsection 5.2) and we obtain some novel regularity results as a by-product (Subsection 5.3). Furthermore, the convergence analysis can be extended to the case of sequences of approximate minimizers (Subsection 5.6) and combined with time-discretization (Section 6). Finally, some application of the abstract theory to a collection of examples of linear and nonlinear parabolic problems is provided in Section 7.

An important step toward the proof of Theorem 1.1 is the analysis of the Euler system for I_ε in $K(u_{0\varepsilon})$. In particular, we prove that the minimizer u_ε fulfills

$$-\varepsilon u'' + u' + \partial\phi(u) \ni f \quad \text{a.e. in} \quad (0,T), \tag{1.4a}$$

$$u(0) = u_{0\varepsilon}, \tag{1.4b}$$

$$u'(T) = 0. \tag{1.4c}$$

Namely, to minimize I_ε is equivalent to perform an elliptic-in-time regularization of the gradient flow (1.1). We shall stress that, at all levels $\varepsilon > 0$, causality is lost. Consequently, the convergence for $\varepsilon \to 0$ is generally referred to as the *causal limit* of (1.4). As the problem above is second order in time, an extra boundary condition (1.4c) at the final point T is needed and our choice for a homogeneous Neumann condition is motivated by simplicity. Other choices may be considered and a specific alternative, originally proposed in [MO08], is commented in Subsection 5.7.

Before moving on, let us recall that the idea of taking the causal limit in an elliptic-in-time regularization of a parabolic problem is not new. In the linear case, some results can be found in the classical monograph by LIONS & MAGENES [LM72]. As for the nonlinear case, this procedure has been followed by ILMANEN [Ilm94] for proving existence and partial regularity of the so-called Brakke mean curvature flow of varifolds. In [Gio96, Sect. 4] a conjecture suggests the weighted functional

$$u \mapsto \int_{\Omega\times(0,\infty)} \mathrm{e}^{-t/\varepsilon} \left[|u_{tt}|^2 + \frac{1}{\varepsilon^2}\Big(|\nabla u|^2 + u^{2k} \Big) \right] \mathrm{d}x\, \mathrm{d}t$$

as elliptic regularization for studying the wave equation $u_{tt} = \Delta u - k u^{2k-1}$.

Besides the WED functional approach here considered, a number of different variational principles have been proposed for characterizing entire trajectories of evolution systems. In the linear realm, we shall mention BIOT's work on irreversible Thermodynamics [Bio55] and GURTIN's principle for viscoelasticity and elastodynamics [Gur63, Gur64a, Gur64b] among many others (see also the survey by HLAVÁČEK [Hla69]). In the nonlinear setting, a crucial result is the BREZIS, EKELAND, & NAYROLES principle [BE76a, BE76b, Nay76a, Nay76b] which specifically focuses on the case of *convex* functionals ϕ. The literature on this principle is vast and the reader is referred to the recent monograph by GHOUSSOUB [Gho08] and the papers [Ste08a, Ste08c, Ste08b, Vis08] for additional information. Apart from the convex case, we shall record the variational principle from DE GIORGI, MARINO, & TOSQUES which actually paved the way to the analysis of gradient flow evolution in metric spaces (see [AGS05, MST89, RMS08, RSS08], for instance). Finally, we mention VISINTIN [Vis01], where generalized solutions are obtained as minimal elements of a certain partial-order relation on the trajectories.

2. Preliminaries

We shall collect here some notation, general assumptions, and a selection of classical results on λ-convex functions and the corresponding gradient flows.

2.1. **Convexity and λ-convexity.** Throughout the paper H is a real Hilbert space with scalar product $(\cdot,\cdot)$ and norm $|\cdot|$. Given the functional $\phi : H \to (-\infty,\infty]$ with effective domain $D(\phi) = \{u \in H \ : \ \phi(u) < \infty\}$, we recall that its *Fréchet subdifferential* $\partial\phi : H \to 2^H$ is defined as

$$v \in \partial\phi(u) \quad \text{iff} \quad u \in D(\phi) \ \text{ and } \ \liminf_{w\to u} \frac{\phi(w) - \phi(u) - (v, w-u)}{|w-u|} \geq 0.$$

We denote by $D(\partial\phi)$ the corresponding domain $D(\partial\phi) \doteq \{u \in H \ : \ \partial\phi(u) \neq \emptyset\}$. The functional ϕ is said to be *proper* if $D(\phi) \neq \emptyset$ and *λ-convex* for some given $\lambda \in \mathbb{R}$, if

$$v \mapsto \psi(v) = \phi(v) - \frac{\lambda}{2}|v|^2 \quad \text{is convex.}$$

Equivalently, ϕ is λ-convex if and only if

$$\phi(ru + (1-r)v) \leq r\phi(u) + (1-r)\phi(v) - \frac{\lambda}{2} r(1-r)|u-v|^2 \quad \forall u,\, v \in H,\ 0 \leq r \leq 1.$$

Let us explicitly remark that $D(\psi) = D(\phi)$, $D(\partial\psi) = D(\partial\phi)$, and $\partial\phi(v) = \partial\psi(v) + \lambda v$ for all $v \in D(\partial\phi)$. In particular, the set $\partial\phi(v)$ turns out to be convex and closed. Hence, it possesses a unique element of minimal norm which we indicate by $(\partial\phi(u))^\circ$.

A crucial tool in Convex Analysis is the *Moreau-Yosida* approximation $\psi_\delta : H \to \mathbb{R}$ of the proper, convex, and lower semicontinuous function $\psi : H \to (-\infty,\infty]$ given, for all $\delta > 0$, by

$$\psi_\delta(u) = \inf_{v\in H} \left(\frac{|v-u|^2}{2\delta} + \psi(v) \right) \quad \forall u \in H.$$

Recall that $\psi_\delta \in C^{1,1}(H)$ and that one has [Bre73b]

$$|\mathrm{D}\psi_\delta(u)| \leq |(\partial\psi(u))^\circ| \quad \text{and} \quad \mathrm{D}\psi_\delta(u) \to (\partial\psi(u))^\circ \ \ \forall u \in D(\partial\psi). \tag{2.1}$$

For any proper functional $\phi : H \to (-\infty,\infty]$ we denote by $\mathrm{sc}^-\phi$ the corresponding lower semicontinuous envelope or *relaxation*, classically defined by

$$\mathrm{sc}^-\phi(u) \doteq \inf\big\{ \liminf_{k\to\infty} \phi(u_k),\ u_k \to u \ \text{ strongly in } \ H \big\}.$$

2.2. **Function spaces.** Standard notation for spaces of vector-valued functions as $L^p(0,T;H)$, $C([0,T];H)$, $W^{1,p}(0,T;H)$, and $H^s(0,T;H)$ will be used throughout, cf. [LM72]. Moreover, we will consider the following characterizations of Besov spaces [BL76, Thm. 6.2.4, p. 142]

$$B^s_{p\,q}(0,T;H) \doteq (L^p(0,T;H), W^{1,p}(0,T;H))_{s,q} \quad 0 < s < 1,\ 1 \leq p,\, q \leq \infty,$$
$$B^{-s}_{p'\,q'}(0,T;H) \doteq (B^s_{p\,q}(0,T;H))' \quad 0 < s < 1,\ 1 \leq p,\, q < \infty$$

where p' and q' are conjugate to p and q, respectively; and $(X,Y)_{s,q}$ denotes L^q interpolation. Let us recall the identifications [Tri95, Rem. 4, p. 179], for all $0 < s < 1$,

$$H^s(0,T;H) = B^s_{2\,2}(0,T;H),$$
$$C^s([0,T];H) = (L^\infty(0,T;H), W^{1,\infty}(0,T;H))_{s,\infty},$$

where the latter is the space of Hölder continuous functions endowed with the norm

$$\|u\|_{C^s([0,T];H)} \doteq \|u\|_{C([0,T];H)} + \sup_{t\neq r} \frac{|u(t)-u(r)|}{|t-r|}.$$

2.3. **General assumptions and well-posedness for** (1.1)**.** Unless otherwise stated, throughout this analysis we shall assume the following

$$\phi : H \to (-\infty,\infty] \quad \text{is proper, lower semicontinuous, bounded from below}$$
$$\text{and } u \mapsto \psi(u) \doteq \phi(u) - \frac{\lambda}{2}|u|^2 \text{ is convex,} \tag{2.2a}$$
$$f \in L^2(0,T;H), \tag{2.2b}$$
$$u_0 \in D(\phi). \tag{2.2c}$$

Note that the lower-bound request for ϕ can be weakened and is here chosen for the sake of simplicity only. As for the λ-convexity assumption, note that any $C^{1,1}$ perturbation of a convex function turns out to be λ-convex (but see (5.6) below).

We assume from the very beginning that

$$\min \psi = \psi(0) = 0.$$

This can be achieved without loss of generality simply by replacing ψ (and hence ϕ), f, and u_0, by

$$\tilde\psi(u) \doteq \psi(u+v) - (\eta, u) - \psi(v), \quad \tilde f \doteq f - \eta, \quad \tilde u_0 \doteq u_0 - v$$

for some fixed $v \in D(\partial\psi)$ with $\eta \in \partial\psi(v)$.

Let us recall that, the well-posedness of the gradient flow (1.1) follows from the classical theory of [Kōm67, CP69, Bre71, Bre73b] (see also [AGS05]). Indeed, the assumption $u_0 \in D(\phi)$ can be weakened to $u_0 \in \overline{D(\partial\phi)}$. In this case as well, a strong solution $u \in H^1_{\rm loc}(0,T;H)$ of (1.1) uniquely exists.

2.4. **Well-posedness for the minimum problem.** In the convex case

$$\lambda^- \doteq \max\{0, -\lambda\} = 0$$

assumptions (2.2a)-(2.2b) guarantee that I_ε admits a (unique) minimizer in $K(w_0)$ for any $w_0 \in H$. As for the general λ-convex case, existence and uniqueness of minimizers follow by letting ε be small enough. More precisely, we have the following.

Proposition 2.1 (Well-posedness for the minimum problem)**.** *Let $\phi : H \to (-\infty,\infty]$ be λ-convex, $f \in L^2(0,T;H)$, and $w_0 \in H$. Letting ε be small enough, the functional I_ε is κ_ε-convex in $K(w_0)$ with respect to the metric of $H^1(0,T;H)$ for*

$$\kappa_\varepsilon \doteq \varepsilon^2 \mathrm{e}^{-T/\varepsilon}. \tag{2.3}$$

In particular, I_ε is uniformly convex in $K(w_0)$.

Additionally, if ϕ is lower semicontinuous, then I_ε admits a unique minimizer in $K(w_0)$.

Proof. Let us start by decomposing I_ε into the sum of a quadratic part Q_ε and a convex remainder R_ε as follows.

$$\begin{aligned} I_\varepsilon(u) &= \left(\int_0^T \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}|u'|^2 - \frac{\lambda^-}{2\varepsilon}|u|^2\right)\right) + \left(\int_0^T \frac{1}{\varepsilon}\mathrm{e}^{-t/\varepsilon}\big(\psi(u) - (f,u)\big)\right) \\ &\doteq Q_\varepsilon(u) + R_\varepsilon(u). \end{aligned} \tag{2.4}$$

In order to handle Q_ε, we will exploit the auxiliary function $v(t) \doteq \mathrm{e}^{-t/(2\varepsilon)}u(t)$. As we readily have that

$$\mathrm{e}^{-t/(2\varepsilon)}u'(t) = v'(t) + \frac{1}{2\varepsilon}v(t), \tag{2.5}$$

the value $Q_\varepsilon(u)$ can be rewritten in terms of v as

$$\begin{aligned} Q_\varepsilon(u) &= \int_0^T \left(\frac{1}{2}|v'|^2 + \frac{1}{2\varepsilon}(v',v) + \frac{1-4\varepsilon\lambda^-}{8\varepsilon^2}|v|^2\right) \\ &= \int_0^T \left(\frac{1}{2}|v'|^2 + \frac{1-4\varepsilon\lambda^-}{8\varepsilon^2}|v|^2\right) + \frac{1}{4\varepsilon}\mathrm{e}^{-T/\varepsilon}|u(T)|^2 - \frac{1}{4\varepsilon}|u(0)|^2 \\ &\doteq V_\varepsilon(v) + \frac{1}{4\varepsilon}\mathrm{e}^{-T/\varepsilon}|u(T)|^2 - \frac{1}{4\varepsilon}|u(0)|^2. \end{aligned} \tag{2.6}$$

Moreover, by possibly letting ε be small, standard computations lead to

$$\mathrm{e}^{-T/\varepsilon}\|u\|^2_{L^2(0,T;H)} \le \|v\|^2_{L^2(0,T;H)} \le \|u\|^2_{L^2(0,T;H)}, \tag{2.7}$$

$$\varepsilon^2\mathrm{e}^{-T/\varepsilon}\|u\|^2_{H^1(0,T;H)} \le \|v\|^2_{H^1(0,T;H)} \le \varepsilon^{-2}\|u\|^2_{H^1(0,T;H)}. \tag{2.8}$$

Let now $\theta \in [0,1]$ and u_1, $u_2 \in K(w_0)$ be given. Moreover, define $v_i(t) \doteq \mathrm{e}^{-t/2\varepsilon}u_i(t)$ for $i = 1,2$. Arguing as in (2.6), for all ε small enough one deduces that

$$\begin{aligned} &Q_\varepsilon\big(\theta u_1 + (1-\theta)u_2\big) \\ &= V_\varepsilon(\theta v_1 + (1-\theta)v_2) + \frac{1}{4\varepsilon}\mathrm{e}^{-T/\varepsilon}|\theta u_1(T) + (1-\theta)u_2(T)|^2 - \frac{1}{4\varepsilon}|w_0|^2 \\ &\le \theta V_\varepsilon(v_1) + (1-\theta)V_\varepsilon(v_2) - \frac{\theta(1-\theta)}{2}\int_0^T \left(|v_1' - v_2'|^2 + \frac{1-4\varepsilon\lambda^-}{4\varepsilon^2}|v_1 - v_2|^2\right) \\ &+ \frac{\theta}{4\varepsilon}\mathrm{e}^{-T/\varepsilon}|u_1(T)|^2 + \frac{1-\theta}{4\varepsilon}\mathrm{e}^{-T/\varepsilon}|u_2(T)|^2 - \frac{1}{4\varepsilon}|w_0|^2 \\ &= \theta Q_\varepsilon(u_1) + (1-\theta)Q_\varepsilon(u_2) - \frac{\theta(1-\theta)}{2}\int_0^T \left(|v_1' - v_2'|^2 + \frac{1-4\varepsilon\lambda^-}{4\varepsilon^2}|v_1 - v_2|^2\right) \\ &\le \theta Q_\varepsilon(u_1) + (1-\theta)Q_\varepsilon(u_2) - \frac{\theta(1-\theta)}{2}\|v_1 - v_2\|^2_{H^1(0,T;H)}. \end{aligned}$$

By exploiting the first estimate in (2.8), we have proved that Q_ε is κ_ε-convex in $K(w_0)$ with respect to the metric of $H^1(0,T;H)$. As $I_\varepsilon = Q_\varepsilon + R_\varepsilon$ and R_ε is convex, the κ_ε-convexity of I_ε follows as well.

Once the uniform convexity of I_ε in $K(w_0)$ is established, the existence of a unique minimizer is a consequence of the Direct Method whenever lower semicontinuity is assumed. $\square$

The proof of Proposition 2.1 entails the existence of $\varepsilon^* > 0$, possibly depending on λ^- only, such that, for all $\varepsilon \in (0, \varepsilon^*)$, the functional I_ε has a unique minimizer in $K(w_0)$. This can be seen as a manifestation of the fact that, for small ε, we are close to the (causal) initial-value problem, where we can expect existence and uniqueness. In the following, the parameter ε will be assumed to fulfill $\varepsilon \in (0, \varepsilon^*)$ throughout.

Note that, for large values of the parameter ε, existence of minimizers may fail. Let us give an example for this fact. In order to keep the presentation simple, we shall consider a scalar example, i.e. $H = \mathbb{R}$, by dropping the lower boundedness assumption on ϕ. We consider

$$\phi(u) = -\frac{u^2}{4}, \quad f = 0, \quad w_0 = 0.$$

By fixing $\varepsilon = 1$, for simplicity, the corresponding WED functional reads

$$I_1(u) \doteq \int_0^T \mathrm{e}^{-t} \left(\frac{|u'|^2}{2} - \frac{u^2}{4} \right)$$

and we readily check that I_1 is 2-homogeneous, namely $I_1(\alpha u) = \alpha^2 I_1(u)$.

Let us firstly prove that $\inf I_1 = -\infty$ in $K(0)$, in particular no global minimizer exists. To this aim it suffices to consider $v(t) \doteq \mathrm{e}^{t/2} - 1$ and compute

$$I_1(v) = -\frac{T}{8} + \frac{1}{4}(\mathrm{e}^{-T} - 1) + (1 - \mathrm{e}^{-T/2})$$

so that, for T suitably large, $I_1(v) < 0$. Then, by homogeneity, we have that $I_1(\alpha v) \to -\infty$ as $\alpha \to \infty$.

We now turn our attention to local minimizers. The Euler equation for I_1 is

$$-u'' + u' - \frac{u}{2} = 0$$

which, letting $u(0) = 0$, is solved by $u_\alpha(t) \doteq \alpha \mathrm{e}^{t/2} \sin(t/2)$ for all $\alpha \in \mathbb{R}$.

If $T \neq (3/2 + 2k)\pi$, no choice of $\alpha \neq 0$ fulfills $u'_\alpha(T) = 0$. Namely, u_α is not a local minimizer for $\alpha \neq 0$. Moreover, the trajectory δv (with v as above and $\delta > 0$ small) is an admissible perturbation of the trivial solution and $I_1(\delta v) = \delta^2 I_1(v) < 0 = I_1(0)$. Namely, $u = 0$ is not a local minimizer either.

If $T = (3/2 + 2k)\pi$, all $\alpha \in \mathbb{R}$ give rise to a solution of the Euler system and one has that $I_1(u_\alpha) = 0$. Still, exactly as for $u = 0$ (see above), the functions u_α are not local minimizers as

$$\begin{aligned} I_1(u_\alpha + \alpha\delta v) &= \alpha^2 I_1(u_1 + \delta v) = \alpha^2 \left(I_1(u_1) + \delta^2 I_1(v) + \delta \int_0^T \mathrm{e}^{-t} \left(u'v' - \frac{1}{2} uv \right) \right) \\ &= \alpha^2 \delta^2 I_1(v) + \frac{\alpha^2 \delta}{2} \int_0^T (\mathrm{e}^{-t/2} - 1) \sin(t/2) < \alpha^2 \delta^2 I_1(v) < 0 = I_1(u_\alpha) \end{aligned}$$

and $u_\alpha + \alpha\delta v$ is a strict competitor of u_α, for δ small.

Uniqueness of minimizers directly follows by uniform convexity if ϕ is convex or ε is small (see above). In the general λ-convex case a uniqueness result for large ε is however not to be expected. Indeed, by letting

$$\phi(u) = I_B(u) - \frac{u^2}{4}, \quad f = 0, \quad w_0 = 0,$$

where I_B is the indicator function of the interval $B \doteq [-e^{T/2}, e^{T/2}]$, as the trajectory v is such that $I_1(v) < 0$ (for T large) and the functional is even, we have that I_1 has two symmetric minimizers (global).

2.5. **Approximation of the initial datum.** As we have already mentioned in the Introduction, the initial datum u_0 of the gradient flow (1.1) is approximated here by a sequence $u_{0\varepsilon}$ and the minimization of I_ε will take place in $K(u_{0\varepsilon})$. Following Brezis [Bré73a] (see also [BS94, Bre75]), we introduce the *interpolation sets* $D_{r,p} \subset H$ for $0 < r < 1$, $1 \le p \le \infty$ as

$$D_{r,p} = \{u \in \overline{D(\partial\psi)} \ : \ \varepsilon \mapsto \varepsilon^{-r}|u - J_\varepsilon u| \in L^p_*(0,1)\}$$

where $J_\varepsilon = (\mathrm{id} + \varepsilon\partial\psi)^{-1}$ is the standard *resolvent* operator and $L^p_*(0,1)$ is the L^p space endowed with the *Haar measure* $\mathrm{d}\varepsilon/\varepsilon$. We will use the equivalence [Bré73a, Thm. 2]

$$u_0 \in D_{r,p} \quad \text{iff} \quad \begin{cases} \exists \varepsilon \in [0,1] \mapsto v(\varepsilon) \ : \ v \in W^{1,1}_{\mathrm{loc}}(0,1], \\ \text{continuous in } [0,1],\ v(0) = u_0,\ v(\varepsilon) \in D(\partial\psi) \text{ a.e., and} \\ \varepsilon^{1-r}\Big(|(\partial\psi(v(\varepsilon)))^\circ| + |v'(\varepsilon)|\Big) \in L^p_*(0,1). \end{cases}$$

As we have that $u_0 \in D(\phi) \equiv D(\psi) \equiv D_{1/2,2}$ and $D_{1/2,2} \subset D_{1/2,\infty}$ [Bre75, Thm. 6], we fix from the very beginning the sequence $u_{0\varepsilon} \doteq v(\varepsilon) \to u_0$ in H in such a way that

$$\varepsilon^{-1/2}|u_0 - u_{0\varepsilon}| + \varepsilon^{1/2}|(\partial\phi(u_{0\varepsilon}))^\circ| \le c_0, \tag{2.9}$$

for some fixed $c_0 > 0$ (recall that $(\partial\phi(u))^\circ = (\partial\psi(u))^\circ + \lambda u$). Note that the first term in the left-hand side above is under control as

$$\begin{aligned} |u_0 - u_{0\varepsilon}| &\le \int_0^\varepsilon |v'(e)|\,\mathrm{d}e \le \varepsilon^{1/2}\left(\int_0^\varepsilon \big(|v'(e)|e^{1/2}\big)^2 \frac{\mathrm{d}e}{e}\right)^{1/2} \\ &\le \varepsilon^{1/2}\|e \mapsto e^{1/2}v(e)\|_{L^2_*(0,1)}. \end{aligned}$$

In particular, we will use the fact that

$$\phi(u_{0\varepsilon}) = \phi(u_0) + ((\partial\phi(u_{0\varepsilon}))^\circ, u_{0\varepsilon} - u_0) \le \phi(u_0) + c_0^2. \tag{2.10}$$

Note that, as we shall comment below, in case $u_0 \in D(\partial\phi)$ no approximation $u_{0\varepsilon}$ is actually needed and the minimization of I_ε could be considered in the fixed $K(u_0)$ as well. A concrete example of sets $D_{r,p}$ is provided in Subsection 7.1.

2.6. **Time-discretization.** In the following, we shall also be considering the classical time-discretization of the gradient flow (1.1) by means of the so-called implicit Euler scheme which, given $n \in \mathbb{N}$ and the constant time-step $\tau = T/n$, consists in the system

$$u^0 = u_0 \quad \text{and} \quad \frac{u^i - u^{i-1}}{\tau} + \partial\phi(u^i) \ni f^i \quad \text{for} \quad i = 1, \ldots, n. \tag{2.11}$$

Whenever a suitable approximation $(f^1, \dots, f^n) \in H^n$ of f is given, the latter system turns out to admit a unique solution $(u^0, u^1, \dots, u^n) \in H^{n+1}$ for τ small. In fact, (2.11) is equivalent to the successive minimization problems

$$(2.12)\quad u^0 = u_0 \quad \text{and} \quad u^i = \operatorname*{Arg\,min}_{u \in H} \left(\frac{|u - u^{i-1}|^2}{2\tau} + \phi(u) - (f^i, u) \right) \quad \text{for} \quad i = 1, \dots, n,$$

where all of the functionals above are uniformly convex (for small τ) and lower semicontinous.

Given any vector $(v^0, \dots, v^n) \in V^{n+1}$ ($V = H, \mathbb{R}$), we will denote by $\overline{v}_\tau : (0, T] \to V$ and $v_\tau : [0, T] \to V$ the corresponding backward piecewise constant and piecewise affine interpolants on the time-partition. Namely, we have

$$\overline{v}_\tau(t) = v^i, \quad v_\tau(0) = v^0, \quad v_\tau(t) = \alpha_i(t) v^i + (1 - \alpha_i(t)) v^{i-1}$$
$$\text{for} \quad t \in ((i-1)\tau, i\tau], \ i = 1, \dots, n,$$

where $\alpha_i(t) = (t - (i-1)\tau)/\tau$, for $i = 1, \dots, n$. Finally, we will also set $\delta v^i = (v^i - v^{i-1})/\tau$, so that, in particular, $\overline{\delta v}_\tau = v'_\tau$. A basic convergence result for (2.11) is combined with the error analysis by Ambrosio, Gigli, & Savaré [AGS05] (see also [NSV00]) in the following.

Lemma 2.2 (Convergence of time-discretizations)**.** *Let $(f^1_\tau, \dots, f^n_\tau)$ be such that $\overline{f}_\tau \to f$ strongly in $L^2(0, T; H)$ and $(u^0_\tau, \dots, u^n_\tau)$ solve* (2.11)*. Then $u_\tau \to u$ strongly in $H^1(0, T; H)$ where u solves* (1.1)*.*

By letting $f \equiv 0$ and τ small enough (in particular $\lambda\tau > -1$), we have that

$$(2.13)\qquad |(u - u_\tau)(t)| \le c_1 \sqrt{\tau} \phi(u_0) \mathrm{e}^{-2\lambda_\tau t} \quad \textit{where} \quad \lambda_\tau \doteq \ln\left(\frac{1 + \lambda\tau}{\tau} \right)$$

where c_1 depends solely on λ. Moreover, if $u_0 \in D(\partial\phi)$ we also have

$$(2.14)\qquad |(u - u_\tau)(t)| \le c_2 \tau |(\partial\phi(u_0))^\circ| \mathrm{e}^{-2\lambda_\tau t}$$

where c_2 depends solely on λ.

Note that the factor $\mathrm{e}^{-2\lambda_\tau t}$ in (2.13)-(2.14) essentially plays the role of the exponential $\mathrm{e}^{-2\lambda t}$. In particular, if $\lambda > 0$ the error constant decays whereas if $\lambda < 0$ it deteriorates exponentially with time. Although we restrict here to the error control for $f \equiv 0$ for the sake of simplicity, the non-homogeneous case can be considered as well. The reader is referred to [NSV00] for some results in this direction.

3. Euler equation

As already mentioned in the Introduction, our analysis relies on the specific structure of the Euler equation for I_ε, namely its linearity with respect to the time-derivatives. The aim of this section is to provide some detail on the Euler system and we shall start form the following.

Theorem 3.1 (Euler equation)**.** *Let u_ε minimize I_ε in $K(u_{0\varepsilon})$. Then, $u_\varepsilon \in H^2(0,T;H)$ and there exists a function $\xi_\varepsilon \in L^2(0,T;H)$ such that*

$$-\varepsilon u_\varepsilon'' + u_\varepsilon' + \xi_\varepsilon = f \quad a.e. \ in \ \ (0,T), \tag{3.1}$$

$$u_\varepsilon(0) = u_{0\varepsilon}, \tag{3.2}$$

$$u_\varepsilon'(T) = 0, \tag{3.3}$$

$$\xi_\varepsilon \in \partial\phi(u_\varepsilon) \quad a.e. \ in \ \ (0,T). \tag{3.4}$$

3.1. **Analysis of a regularized convex problem.** For the sake of proving Theorem 3.1, we focus on a regularized problem first. Let ψ_δ be the Yosida approximation of ψ at level $\delta > 0$. We have the following.

Lemma 3.2. *There exists a unique $u_\delta \in H^2(0,T;H)$ such that*

$$-\varepsilon u_\delta'' + u_\delta' + \mathrm{D}\psi_\delta(u_\delta) = f \quad a.e. \ in \ \ (0,T), \tag{3.5}$$

$$u_\delta(0) = u_{0\varepsilon}, \tag{3.6}$$

$$u_\delta'(T) = 0. \tag{3.7}$$

Proof. By possibly redefining $\mathrm{D}\psi_\delta$ as $\mathrm{D}\psi_\delta(\cdot + u_{0\varepsilon})$, we assume with no loss of generality that $u_{0\varepsilon} = 0$. Let $V = \{u \in H^1(0,T;H) \ : \ u(0) = 0\}$ and denote by V' the corresponding dual. A weak formulation of (3.5)-(3.7) is provided by the equation $Au + Bu = \ell$, where A, $B : V \to V'$ and $\ell \in V'$ are given, for all $v \in V$, by

$$\langle Au, v\rangle \doteq \varepsilon \int_0^T (u', v') + \int_0^T (u', v),$$

$$\langle Bu, v\rangle \doteq \int_0^T (\mathrm{D}\psi_\delta(u), v),$$

$$\langle \ell, v\rangle \doteq \int_0^T (f, v)$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between V' and V. The linear operator A is coercive as

$$\langle Au, u\rangle = \varepsilon \int_0^T |u'|^2 + \frac{1}{2}|u(T)|^2 \quad \forall u \in V.$$

On the other hand, B is clearly monotone and continuous. Hence, $A + B$ is maximal monotone and coercive [Bar76, Cor. 1.1, p. 39]. Namely, $Au + Bu = \ell$ admits at least a solution $u \in V$ [Bar76, Cor. 1.3, p. 48]. Finally, as A is strongly monotone, this solution is unique.

Equation $Au + Bu = \ell$ reads

$$\varepsilon \int_0^T (u', v') = \int_0^T (-u' - \mathrm{D}\psi_\delta(u) + f, v) \ \ \forall v \in V \tag{3.8}$$

By choosing $v \in V$ such that $v(T) = 0$ we recover $u \in H^2(0,T;H)$ and that relation (3.5) holds. Hence, again from (3.8), by using the already established (3.5) one also has that $\varepsilon\big(-u'(T), v(T)\big) = 0$ for all $v \in V$ and (3.7) follows. ☐

The forthcoming discussion of Subsection 4.1 will in particular entail the validity of the following estimate.

Lemma 3.3 (Estimate on u_δ)**.** *Let u_δ solve* (3.5)-(3.7). *Then*

$$\|u_\delta\|_{H^2(0,T;H)} \le c \tag{3.9}$$

where $c > 0$ depends on $\|f\|_{L^2(0,T;H)}$, $|u_0|$, c_0, and ε but not on δ.

3.2. **Proof of Theorem 3.1.** Let us assume with no loss of generality $\lambda = -\lambda^- \le 0$, decompose the functional I_ε into its convex and its non-convex part as $I_\varepsilon = C_\varepsilon + N_\varepsilon$, and extend it to the whole $L^2(0,T;H)$, namely, for $v \in L^2(0,T;H)$, we let

$$C_\varepsilon(v) \doteq \int_0^T \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}|v'|^2 + \frac{1}{\varepsilon}\Big(\psi(v) - (f,v)\Big)\right) \quad \text{for } v \in K(u_{0\varepsilon}) \text{ and } \infty \text{ otherwise,}$$
$$N_\varepsilon(v) \doteq -\frac{\lambda^-}{2\varepsilon}\int_0^T \mathrm{e}^{-t/\varepsilon}|v|^2.$$

We shall now compute subdifferentials in the weighted space $L^2(0,T,\mathrm{e}^{-t/\varepsilon}\mathrm{d}t;H)$. As N_ε is clearly C^1, one has that

$$\partial I_\varepsilon = \partial C_\varepsilon + \mathrm{D}N_\varepsilon \quad \text{in} \quad L^2(0,T,\mathrm{e}^{-t/\varepsilon}\mathrm{d}t;H).$$

As the minimality of u implies that $0 \in \partial I_\varepsilon(u)$, what is now needed is a description of the set $\partial C_\varepsilon(u)$ as, clearly, $\mathrm{D}N_\varepsilon(u) = -\lambda^- u/\varepsilon$. We shall prove that $\partial C_\varepsilon(u) = A_\varepsilon(u)$ where the possibly multivalued operator A_ε is defined on $D(A_\varepsilon) \doteq \{v \in H^2(0,T;H) \cap K(u_{0\varepsilon}) : v'(T) = 0\}$ as

$$A_\varepsilon(u) \doteq \frac{1}{\varepsilon}\Big(-\varepsilon u'' + u' + \partial\Psi_\varepsilon(u) - f\Big).$$

In the latter, the integral functional $\Psi_\varepsilon : L^2(0,T;H) \to (-\infty,\infty]$ is given by

$$\Psi_\varepsilon(u) \doteq \begin{cases} \displaystyle\int_0^T \mathrm{e}^{-t/\varepsilon}\psi(u)\,\mathrm{d}t & \text{if } t \mapsto \psi(u(t)) \in L^1(0,T), \\ \infty & \text{else,} \end{cases}$$

and the subdifferential $\partial\Psi_\varepsilon$ is again taken in $L^2(0,T,\mathrm{e}^{-t/\varepsilon}\mathrm{d}t;H)$.

Let us firstly check that $A_\varepsilon(u) \subset \partial C_\varepsilon(u)$. Let $\eta \in L^2(0,T;H)$ such that $\eta \in \partial\Psi_\varepsilon(u)$, namely $\eta \in \partial\psi(u)$ almost everywhere. For all $w \in K(u_{0\varepsilon})$ we compute that

$$\begin{aligned}
&\frac{1}{\varepsilon}\int_0^T \mathrm{e}^{-t/\varepsilon}\Big(-\varepsilon u'' + u' + \eta - f, w - u\Big) \\
&= \int_0^T \Big((-\mathrm{e}^{-t/\varepsilon}u')', w - u\Big) + \frac{1}{\varepsilon}\int_0^T \mathrm{e}^{-t/\varepsilon}\big(\eta - f, w - u\big) \\
&= \int_0^T \mathrm{e}^{-t/\varepsilon}\big(u', w' - u'\big) + \frac{1}{\varepsilon}\int_0^T \mathrm{e}^{-t/\varepsilon}\big(\eta - f, w - u\big) \\
&= \frac{1}{2}\int_0^T \mathrm{e}^{-t/\varepsilon}\big(|w'|^2 - |u' - w'|^2 - |u'|^2\big) + \frac{1}{\varepsilon}\int_0^T \mathrm{e}^{-t/\varepsilon}\big(\xi - f, w - u\big) \\
&\overset{\eta\in\partial\Psi(u)}{\le} \frac{1}{2}\int_0^T \mathrm{e}^{-t/\varepsilon}\big(|w'|^2 - |u'|^2\big) + \frac{1}{\varepsilon}\Big(\Psi_\varepsilon(w) - \Psi_\varepsilon(u)\Big) - \frac{1}{\varepsilon}\int_0^T \mathrm{e}^{-t/\varepsilon}\big(f, w - u\big) \\
&= C_\varepsilon(w) - C_\varepsilon(u).
\end{aligned}$$

In order to prove the converse inclusion $\partial C_\varepsilon(u) \subset A_\varepsilon(u)$ we shall check that the monotone operator A_ε is *maximal* [Bre73b], namely that, for all $g \in L^2(0,T;H)$, the problem

$$(\mathrm{id} + A_\varepsilon)(u_\varepsilon) \ni g$$

admits a (unique) solution u_ε. We proceed by regularization and passage to the limit. Let ψ_δ be the Yosida approximation of ψ at level $\delta > 0$. Let now u_δ solve (3.5)-(3.7) with $\psi_\delta(\cdot)$ replaced by $\psi_\delta(\cdot) + \varepsilon|\cdot|^2/2$ and f replaced by $f + \varepsilon g$. Namely, we have that

$$-\varepsilon u_\delta'' + u_\delta' + \mathrm{D}\psi_\delta(u_\delta) + \varepsilon u_\delta = f + \varepsilon g \quad \text{a.e. in } \ (0,T) \tag{3.10}$$

The bound (3.9) still holds, independently of δ (but depending on g) and we can extract subsequences, without relabeling, in such a way that

$$u_\delta \to u_\varepsilon \quad \text{weakly in } \ H^2(0,T;H), \tag{3.11}$$

$$\mathrm{D}\psi_\delta(u_\delta) \to \eta_\varepsilon \quad \text{weakly in } \ L^2(0,T;H), \tag{3.12}$$

pass to the limit for $\delta \to 0$ in (3.10) and (3.7), and get

$$-\varepsilon u_\varepsilon'' + u_\varepsilon' + \eta_\varepsilon + \varepsilon u_\varepsilon = f + \varepsilon g \quad \text{a.e. in } \ (0,T) \tag{3.13}$$

and (3.3), respectively. As the initial condition (3.2) is clearly satisfied, one is left with the proof of the inclusion (3.4). To this aim, let us test the regularized equation (3.10) by u_δ and pass to the lim sup as $\delta \to 0$. We obtain by lower semicontinuity that

$$\begin{aligned}
&\limsup_{\delta\to 0} \int_0^T (\mathrm{D}\psi_\delta(u_\delta), u_\delta) \\
&= \limsup_{\delta \to 0} \Bigg(-\varepsilon \int_0^T |u_\delta'|^2 - \varepsilon(u_\delta'(0), u_{0\varepsilon}) - \frac{1}{2}|u_\delta(T)|^2 + \frac{1}{2}|u_{0\varepsilon}|^2 \\
&-\varepsilon \int_0^T |u_\delta|^2 + \int_0^T (f + \varepsilon g, u_\delta) \Bigg) \\
&\leq -\varepsilon \int_0^T |u_\varepsilon'|^2 - \varepsilon(u_\varepsilon'(0), u_{0\varepsilon}) - \frac{1}{2}|u_\varepsilon(T)|^2 + \frac{1}{2}|u_{0\varepsilon}|^2 - \varepsilon \int_0^T |u_\varepsilon|^2 + \int_0^T (f + \varepsilon g, u_\varepsilon) \\
&\overset{(3.13)}{=} \int_0^T (\eta_\varepsilon, u_\varepsilon).
\end{aligned}$$

The above lim sup estimate is sufficient for identifying the limit η_ε [Bre73b, Prop. 2.5, p. 27]. In particular, we have proved that u_ε solves $(\mathrm{id} + A_\varepsilon)(u_\varepsilon) \ni g$ and the assertion of the Theorem follows.

4. Proof of Theorem 1.1

4.1. **Key estimate.** Given the minimizer u_ε of I_ε in $K(u_{0\varepsilon})$ we have checked that u_ε solves (3.1)-(3.4). The proof of Theorem 1.1 consists in a direct control of the distance between u_ε and the solution u of the gradient flow (1.1). This check is performed in Subsection 4.2. The key step in the computation is the validity of some estimates on u_ε which are independent of ε. Let us state this crucial point in the following lemma.

Lemma 4.1 (Key estimate)**.** *Let u_ε minimize I_ε in $K(u_{0\varepsilon})$. For all ε small there exists a constant $c>0$ depending on $\|f\|_{L^2(0,T;H)}$, $|u_0|$, and c_0, but independent of ε such that*

$$\varepsilon\,\|u_\varepsilon''\|_{L^2(0,T;H)}+\varepsilon^{1/2}\,\|u_\varepsilon'\|_{L^\infty(0,T;H)}+\|u_\varepsilon'\|_{L^2(0,T;H)}+\|\xi_\varepsilon\|_{L^2(0,T;H)}\le c, \tag{4.1}$$

where ξ_ε is defined in Theorem 3.1.

The full proof of this result will be achieved by means of a time-discretization technique and is postponed to Subsection 6.5. Let us however provide here a simplified argument in case we have

$$\xi_\varepsilon\in W^{1,1}(0,T;H)\quad\text{and}\quad \xi_\varepsilon(0)=(\partial\phi(u_{0\varepsilon}))^\circ. \tag{4.2}$$

Note that the latter, being false in general, directly follows from $u_\varepsilon\in H^2(0,T;H)$ as soon as ϕ is smooth, say $\phi\in C^{1,1}$.

Henceforth the symbol c will denote a positive constant, possibly varying from line to line and depending on $\|f\|_{L^2(0,T;H)}$, $|u_0|$, and c_0 but independent of ε.

From equation (3.1) we clearly have that $-\varepsilon u_\varepsilon''+u_\varepsilon'+\xi_\varepsilon$ is in $L^2(0,T;H)$. Our aim is now to deduce separate bounds for the three terms above. We argue as follows

$$\begin{aligned}
&\int_0^T|\varepsilon u_\varepsilon''|^2+\int_0^T|u_\varepsilon'|^2+\int_0^T|\xi_\varepsilon|^2\\
&=\int_0^T|-\varepsilon u_\varepsilon''+u_\varepsilon'+\xi_\varepsilon|^2+2\int_0^T(\varepsilon u_\varepsilon'',u_\varepsilon')-2\int_0^T(u_\varepsilon',\xi_\varepsilon)+2\int_0^T(\varepsilon u_\varepsilon'',\xi_\varepsilon)\\
&=\int_0^T|f|^2+2\int_0^T(\varepsilon u_\varepsilon'',u_\varepsilon')-2\int_0^T(u_\varepsilon',\xi_\varepsilon)+2\int_0^T(\varepsilon u_\varepsilon'',\xi_\varepsilon)\\
&=\int_0^T|f|^2+\varepsilon|u_\varepsilon'(T)|^2-\varepsilon|u_\varepsilon'(0)|^2-2\phi(u_\varepsilon(T))+2\phi(u_{0\varepsilon})+2\int_0^T(\varepsilon u_\varepsilon'',\xi_\varepsilon).
\end{aligned}$$

The last term above may be controlled by virtue of (4.2) as

$$\begin{aligned}
2\int_0^T(\varepsilon u_\varepsilon'',\xi_\varepsilon)&=-2\varepsilon\big(u_\varepsilon'(0),\xi_\varepsilon(0)\big)-2\varepsilon\int_0^T(u_\varepsilon',\xi_\varepsilon')\\
&\le\varepsilon|u_\varepsilon'(0)|^2+\varepsilon|\xi_\varepsilon(0)|^2-2\varepsilon\lambda\int_0^T|u_\varepsilon'|^2 \qquad (4.3)
\end{aligned}$$

where we have used $u'(T)=0$ and the λ-convexity of ϕ. Hence, by collecting these computations we have that

$$\begin{aligned}
&\frac12\int_0^T|\varepsilon u_\varepsilon''|^2+\frac{1+2\varepsilon\lambda}{2}\int_0^T|u_\varepsilon'|^2+\frac12\int_0^T|\xi_\varepsilon|^2+\phi(u_\varepsilon(T))\\
&\le\phi(u_{0\varepsilon})+\frac{\varepsilon}{2}|\xi_\varepsilon(0)|^2+\frac12\int_0^T|f|^2\le c+c\varepsilon|\xi_\varepsilon(0)|^2,
\end{aligned}$$

where, in the last inequality, we have used (2.10). Now, by taking ε small with respect to λ in such a way that

$$2\varepsilon\lambda^-\le 1/2, \tag{4.4}$$

we conclude that

$$\varepsilon\|u_\varepsilon''\|_{L^2(0,T;H)}+\|u_\varepsilon'\|_{L^2(0,T;H)}+\|\xi_\varepsilon\|_{L^2(0,T;H)}\le c+c\varepsilon^{1/2}|(\partial\phi(u_{0\varepsilon}))^\circ|. \tag{4.5}$$

Classical interpolation between $L^2(0,T;H)$ and $H^1(0,T;H)$ (cf. [LM72, BL76] or equivalently by Gagliardo-Nirenberg) we obtain

$$\begin{aligned} \|u'_\varepsilon\|_{C^0([0,T];H)} &\le c\|u'_\varepsilon\|^{1/2}_{L^2(0,T;H)}\|u'_\varepsilon\|^{1/2}_{H^1(0,T;H)} \\ &= c\Big(\|u'_\varepsilon\|_{L^2(0,T;H)} + \|u'_\varepsilon\|^{1/2}_{L^2(0,T;H)}\|u''_\varepsilon\|^{1/2}_{L^2(0,T;H)}\Big) \\ &\le c\Big(1 + |(\partial\phi(u_{0\varepsilon}))^\circ|\Big), \end{aligned} \tag{4.6}$$

so that (4.1) follows from (2.9).

Besides the regular case $\phi \in C^{1,1}$, the above argument can be easily adapted to the situation where $\partial\phi$ is *single-valued*. This can be done my means of a nested approximation argument via Moreau-Yosida approximations in (4.3).

4.2. **Proof of Theorem 1.1.** The strategy of this proof is elementary. We shall directly compare the minimizer u_ε of I_ε and the unique solution u of the gradient flow (1.1). In particular, take the difference between (1.1) and the Euler equation (3.1), test it on $w_\varepsilon \doteq u - u_\varepsilon$, and integrate in time getting

$$\begin{aligned} &\varepsilon\int_0^t |w'_\varepsilon|^2 + \frac{1}{2}|w_\varepsilon(t)|^2 + \int_0^t (\xi - \xi_\varepsilon, w_\varepsilon) \\ &= \frac{1}{2}|u_0 - u_{0\varepsilon}|^2 + \varepsilon\int_0^t (u', w'_\varepsilon) - \varepsilon(u'_\varepsilon(t), w_\varepsilon(t)) + \varepsilon(u'_\varepsilon(0), u_0 - u_{0\varepsilon}), \end{aligned} \tag{4.7}$$

where $\xi \in \partial\phi(u)$ almost everywhere. Using λ-convexity we find

$$\begin{aligned} &\varepsilon\int_0^t |w'_\varepsilon|^2 + \frac{1}{2}|w_\varepsilon(t)|^2 + \lambda\int_0^t |w_\varepsilon|^2 \\ &= |u_0 - u_{0\varepsilon}|^2 + \varepsilon\int_0^t \Big(|u'|^2 + |w'_\varepsilon|^2\Big) + \varepsilon^2|u'_\varepsilon(t)|^2 + \frac{1}{4}|w_\varepsilon(t)|^2 + \frac{\varepsilon^2}{2}|u'_\varepsilon(0)|^2. \end{aligned}$$

Owing to Lemma 4.1 and applying Gronwall's Lemma, we readily compute that

$$\begin{aligned} &\frac{\varepsilon}{2}\int_0^t |w'_\varepsilon|^2 + \frac{1}{4}|w_\varepsilon(t)|^2 \\ &\le c\left(|u_0 - u_{0\varepsilon}|^2 + \varepsilon\int_0^t |u'|^2 + \varepsilon^2|u'_\varepsilon(t)|^2 + \varepsilon^2|u'_\varepsilon(0)|^2\right) \;\le\; c\varepsilon, \end{aligned} \tag{4.8}$$

where now c depends on λ^- as well. The strong convergence $u_\varepsilon \to u$ in $C([0,T];H)$ follows. Let us observe that, by inspecting the proof of Lemma 4.1, in case $u_0 \in D(\partial\phi)$ one realizes that no approximation of the initial datum is actually needed and the convergence result holds for minimizers of I_ε in $K(u_0)$ as well.

5. Extensions and comments

5.1. **Sharper statements.** The proof of Theorem 1.1 can be made precise in two different directions. Firstly, the convergence proof is quantitative for we have obtained an explicit convergence rate. Secondly, we can exploit real interpolation in order to check convergence in some finer topology as well.

Let us refer to [BL76] for notation and results on real interpolation between Banach spaces, in particular for the definition of $(C([0,T];H),H^1(0,T;H))_{\eta,1}$ which is used in the following result.

Theorem 5.1 (Sharper convergence result)**.** *For $0<\eta<1$ we have that*

$$\|u-u_\varepsilon\|_{(C([0,T];H),H^1(0,T;H))_{\eta,1}} \le c\varepsilon^{(1-\eta)/2}, \tag{5.1}$$

where $c>0$ depends on $\|f\|_{L^2(0,T;H)}$, $|u_0|$, c_0, T, λ^-, and η, but not on ε.

Proof. By interpolation we have that

$$\begin{aligned}\|u-u_\varepsilon\|_{(C([0,T];H),H^1(0,T;H))_{\eta,1}} &\le c\|u-u_\varepsilon\|^{1-\eta}_{C([0,T];H)}\|u-u_\varepsilon\|^{\eta}_{H^1(0,T;H)}\\ &\overset{(4.8)}{\le} c\varepsilon^{(1-\eta)/2}\varepsilon^0 = c\varepsilon^{(1-\eta)/2},\end{aligned} \tag{5.2}$$

and the result is established. □

Let us make concrete this discussion in the Hilbert scale $H^s(0,T;H)$. Recalling that

$$\begin{aligned}(C([0,T];H),H^1(0,T;H))_{\eta,1} &\subset (L^2(0,T;H),H^1(0,T;H))_{\eta,2}\\ &= B^{\eta}_{2\,2}(0,T;H) = H^\eta(0,T;H),\end{aligned}$$

we get the following.

Corollary 5.2 (Strong convergence in $H^\eta(0,T;H)$)**.** *For $0<\eta<1$ we have that*

$$\|u-u_\varepsilon\|_{H^\eta(0,T;H)} \le c\varepsilon^{(1-\eta)/2} \tag{5.3}$$

where $c>0$ depends on $\|f\|_{L^2(0,T;H)}$, $|u_0|$, c_0, T, λ^-, and η, but not on ε.

5.2. **Weaker assumptions.** The above results can be easily extended to the case when

$$u_0 \in D_{r,\infty} \quad \text{for some } \ 0<r<1. \tag{5.4}$$

Let us ask for a sequence $u_{0\varepsilon}\in D(\partial\phi)$ such that $u_{0\varepsilon}\to u_0$ strongly in H and (see (2.9))

$$\varepsilon^{-r}|u_0-u_{0\varepsilon}| + \varepsilon^{1-r}|(\partial\phi(u_{0\varepsilon}))^\circ| \le c_0, \tag{5.5}$$

for some $c_0>0$. The arguments leading to the key estimate (4.1) still holds (note that (2.10) is fulfilled) and we deduce that

$$\varepsilon\|u''_\varepsilon\|_{L^2(0,T;H)} + \varepsilon^{1/2}\|u'_\varepsilon\|_{L^\infty(0,T;H)} + \|u'_\varepsilon\|_{L^2(0,T;H)} + \|\xi_\varepsilon\|_{L^2(0,T;H)} \le c\varepsilon^{r-1/2}.$$

In particular, estimate (4.8) turns out to be

$$\frac{\varepsilon}{2}\int_0^t |w'_\varepsilon|^2 + \frac14|w_\varepsilon(t)|^2 \le c\left(|u_0-u_{0\varepsilon}|^2 + \varepsilon\int_0^t |u'|^2 + \varepsilon^2|u'_\varepsilon(t)|^2\right) \le c\varepsilon^{2r},$$

and uniform convergence holds for all $r>0$. Of course, the convergence rates of Theorem 5.1 are to be modified as follows

$$\|u-u_\varepsilon\|_{(C([0,T];H),H^1(0,T;H))_{\eta,1}} \le c\|u-u_\varepsilon\|_{H^\eta(0,T;H)} \le c\varepsilon^{r-\eta/2}.$$

Some concretization of this construction in the frame of linear parabolic PDEs is given in Subsection 7.1.

The λ-convexity assumption on the functional can be relaxed to

$$\exists\lambda:[0,\infty)\to\mathbb{R} \ \text{ such that } \phi \text{ is } \lambda(r)\text{-convex on } \{|u|\le r\} \text{ for all } r\ge 0. \tag{5.6}$$

This assumption includes the case of a C^2 functional which is not $C^{1,1}$. Indeed, by testing (3.1) by u'_ε and taking the integral on $(0,T)$, one has that

$$\frac{\varepsilon}{2}|u'(0)|^2 + \int_0^T |u'_\varepsilon|^2 + \phi(u_\varepsilon(T)) = \phi(u_{0\varepsilon}) + \int_0^T (f, u'_\varepsilon).$$

In particular, a bound in $H^1(0,T;H)$ for u_ε, independent of ε, follows and it suffices to fix $r \doteq \sup_{\varepsilon,t}|u_\varepsilon(t)|$ in (5.6) and repeat the argument of Lemma 4.1 with $\lambda = \lambda(r)$ fixed.

5.3. **Regularity result.** A regularity theory for the gradient flow (1.1) in the Hölder scale $C^s([0,T];H)$ has been outlined by Savaré in [Sav96] where he proves that

$$\begin{aligned} u_0 \in \overline{D(\partial\phi)} \ \text{ and } \ f \in B^{-1/2}_{2\,1}(0,T;H) \quad &\Rightarrow \quad u \in C([0,T];H), \\ u_0 \in D(\partial\phi) \ \text{ and } \ f \in B^{1/2}_{2\,1}(0,T;H) \quad &\Rightarrow \quad u \in W^{1,\infty}(0,T;H). \end{aligned}$$

Although classical nonlinear interpolation [Tar70, Tar72] does not directly apply to the present situation, some intermediate regularity is expected. At level 1/2, we readily have that

$$\begin{aligned} D_{1/2,2} &= D(\phi), \\ (B^{-1/2}_{2\,1}(0,T;H), B^{1/2}_{2\,1}(0,T;H))_{1/2,2} &= L^2(0,T;H), \\ (C([0,T];H), W^{1,\infty}(0,T;H))_{1/2,\infty} &= C^{1/2}([0,T];H) \end{aligned}$$

and nothing has to be proved for the intermediate regularity

$$u_0 \in D(\phi) \ \text{ and } \ f \in L^2(0,T;H) \quad \Rightarrow \quad u \in C^{1/2}([0,T];H)$$

follows at once from $H^1(0,T;H) \subset C^{1/2}([0,T];H)$.

On the other hand, we are in the position of completing this regularity theory for weaker assumptions on the initial data u_0 (but keeping $f \in L^2(0,T;H)$ fixed). Indeed, we have that the following regularity result, which is, to our knowledge, new even in the classical convex setting for ϕ.

Lemma 5.3 (Regularity)**.**

$$u_0 \in D_{r,\infty}, \quad f \in L^2(0,T;H) \quad \Longrightarrow \quad u \in C^r([0,T];H).$$

The result follows easily from the fact that, in case $u_0 \in D_{r,\infty}$, one has

$$\varepsilon^{1-r}\|u_\varepsilon\|_{W^{1,\infty}(0,T;H)} + \varepsilon^{-r}\|u - u_\varepsilon\|_{C([0,T];H)} \le c,$$

as the latter entails in particular that

$$u \in (C([0,T];H), W^{1,\infty}(0,T;H))_{s,\infty} = C^s([0,T];H).$$

5.4. **Sharpness of the convergence rates.** Although specific situations (see below) exhibit a stronger convergence rate, in general the above proved error bounds are *sharp* as the estimates (recall (4.8))

$$\|u - u_\varepsilon\|_{C([0,T];H)} \le c\varepsilon^{1/2+\delta} \tag{5.7}$$

$$\|u - u_\varepsilon\|_{H^1(0,T;H)} \le c\varepsilon^{\delta} \tag{5.8}$$

are false for all $\delta > 0$.

We shall prove this fact by contradicting the maximal regularity $u \in H^1(0,T;H)$ via interpolation. In particular, assume (5.7). From (4.8) we have that, for all $0 < \eta < 1$,

$$\|u-u_\varepsilon\|_{C^{\eta/2}([0,T];H)} \le c\|u-u_\varepsilon\|_{(C([0,T];H),C^{1/2}([0,T];H))_{\eta,\infty}}$$
$$\le c\|u-u_\varepsilon\|_{(C([0,T];H),H^1(0,T;H))_{\eta,1}} \le c\,\varepsilon^{(1-\eta)(1/2+\delta)}.$$

Choosing η such that

$$1/2 = (1-\eta)(1/2+\delta) \tag{5.9}$$

and recalling (4.1) we get that

$$\varepsilon^{1/2}\|u_\varepsilon\|_{W^{1,\infty}(0,T;H)} + \varepsilon^{-1/2}\|u-u_\varepsilon\|_{C^{\eta/2}([0,T];H)} \le c.$$

Hence, by interpolation we have that

$$u \in (C^{\eta/2}([0,T];H), W^{1,\infty}(0,T;H))_{1/2,\infty} = C^s([0,T];H) \ \text{ for } s = \frac{1}{2}\frac{\eta}{2} + \frac{1}{2}1 > \frac{1}{2}.$$

On the other hand, as we surely have that, for any $s > 1/2$, there exist functions in $H^1(0,T;H)$ which do not belong to $C^s([0,T];H)$, this clearly amounts to a contradiction.

A similar (easier) argument proves the sharpness in $H^1(0,T;H)$. Indeed, assume (5.8). Then, estimate (4.1) ensures that

$$\varepsilon^{1/2}\|u_\varepsilon\|_{W^{1,\infty}(0,T;H)} + \varepsilon^{-\delta}\|u-u_\varepsilon\|_{H^1(0,T;H)} \le c.$$

Choosing the interpolation level $0 < \delta < 1$ we obtain

$$\begin{aligned} u &\in (H^1(0,T;H), W^{1,\infty}(0,T;H))_{\delta,\infty} \\ &\subset (C^{1/2}([0,T];H), W^{1,\infty}(0,T;H))_{\delta,\infty} \\ &= C^r([0,T];H) \quad \text{with } r = \delta + \frac{1-\delta}{2} = \frac{1+\delta}{2} > \frac{1}{2}, \end{aligned}$$

which again is contradicting the maximal regularity $u \in H^1(0,T;H)$.

Note that the above proofs rely on the choice of a general datum $f \in L^2(0,T;H)$ and a more regular setting could give rise to better convergence rates. Let us stress that we do not presently know if strong convergence holds in $H^1(0,T;H)$. On the other hand, we have just proved that no rate in $H^1(0,T;H)$ can be expected.

5.5. **Special case of** (1.3)**.** In the specific situation of the scalar and linear case of (1.3), some improved convergence rate of u_ε is available. In particular, one can explicitly prove that

$$|(u-u_\varepsilon)(t)| = \varepsilon\left(\mathrm{e}^{(t-1)/\varepsilon-1} - \mathrm{e}^{-1/\varepsilon-t-1}\right) \le \frac{2\varepsilon}{\mathrm{e}},$$

so that a linear convergence rate is achieved in $C([0,T];H)$, see Figure 2.

Moreover, strong convergence in $H^1(0,1)$ holds with rate $1/2$ as we have that

$$\|u-u_\varepsilon\|_{H^1(0,1)} \sim \sqrt{\varepsilon}\left(\frac{1}{2\mathrm{e}^2} + \frac{2}{\mathrm{e}}\right).$$

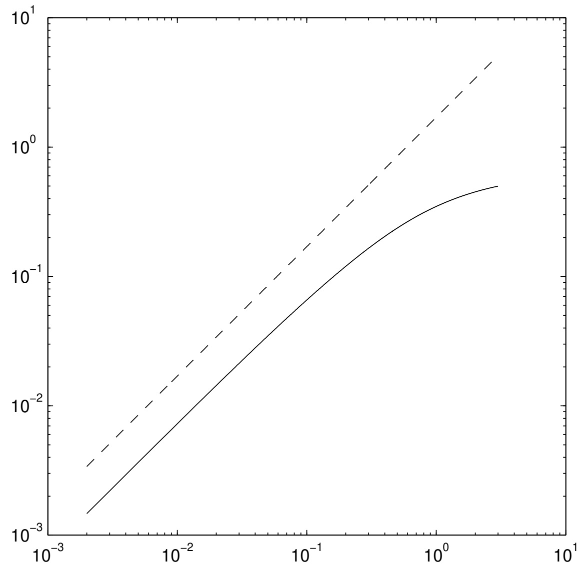

Figure 2: The convergence rate in $C[0,T]$ in the special case (1.3). The solid line is the function $\varepsilon \mapsto \max_{[0,T]} |u - u_\varepsilon|$ and the dashed line is linear in ε (log-log scale).

5.6. **Approximate minimizers, relaxation.** The convergence result of Theorem 1.1 can be extended to the case of qualified sequences of approximate minimizers of the functional I_ε.

Theorem 5.4 (Convergence for approximate minimizers)**.** *Let $v_\varepsilon \in K(u_{0\varepsilon})$ be such that*

$$I_\varepsilon(v_\varepsilon) \le \inf_{K(u_{0\varepsilon})} I_\varepsilon + \alpha_\varepsilon, \qquad \alpha_\varepsilon = \mathrm{o}(\varepsilon^2 \mathrm{e}^{-T/\varepsilon}) \quad \textit{as} \quad \varepsilon \to 0. \tag{5.10}$$

Then $v_\varepsilon \to u$ in $C([0,T];H)$.

Note that the above statement can be generalized in the many directions commented above. In particular, a convergence rate in $C([0,T];H)$ can be derived and the requirement on α_ε can be weakened in case ϕ is convex.

Proof. Let v_ε fulfill (5.10) and ε be small enough. Moreover, let u_ε denote the minimizer of I_ε in $K(u_{0\varepsilon})$. By using the κ_ε-convexity of I_ε from Proposition 2.1 we readily obtain that, for all $\theta \in [0,1)$,

$$\begin{aligned} I_\varepsilon(u_\varepsilon) &\le I_\varepsilon\big(\theta u_\varepsilon + (1-\theta)v_\varepsilon\big) \\ &\le \theta I_\varepsilon(u_\varepsilon) + (1-\theta) I_\varepsilon(v_\varepsilon) - \frac{\theta(1-\theta)}{2}\kappa_\varepsilon \|u_\varepsilon - v_\varepsilon\|^2_{H^1(0,T;H)}. \end{aligned}$$

Dividing by $1-\theta$ and taking $\theta \to 1$ we get that

$$\frac{\kappa_\varepsilon}{2}\|u_\varepsilon - v_\varepsilon\|^2_{H^1(0,T;H)} \le I_\varepsilon(v_\varepsilon) - \min_{K(u_{0\varepsilon})} I_\varepsilon \le \alpha_\varepsilon.$$

As $\alpha_\varepsilon = \mathrm{o}(\kappa_\varepsilon)$ for $\varepsilon \to 0$, we have checked that $u_\varepsilon - v_\varepsilon \to 0$ in $H^1(0,T;H)$ and the assertion follows from Theorem 1.1. $\square$

The convergence result of Theorem 5.4 may be extended in the direction of relaxation. In particular, sequences of approximate minimizers converge even if ϕ is not λ-convex nor lower semicontinuous, provided that $\mathrm{sc}^- I_\varepsilon$ is itself a WED functional for a λ-convex and lower semicontinuous potential. This is the case, for instance, for the two relaxation examples of Subsections 7.5-7.6 below.

Corollary 5.5 (Convergence without convexity and lower semicontinuity)**.** *Assume that* $\mathrm{sc}^- I_\varepsilon$ *is a* WED *functional fulfilling* (2.2a)-(2.2b). *Moreover, let* $v_\varepsilon \in K(u_{0\varepsilon})$ *be such that*

$$I_\varepsilon(v_\varepsilon) \le \inf_{K(u_{0\varepsilon})} I_\varepsilon + \alpha_\varepsilon, \qquad \alpha_\varepsilon = \mathrm{o}(\varepsilon^2 \mathrm{e}^{-T/\varepsilon}) \quad as \quad \varepsilon \to 0.$$

Then $v_\varepsilon \to u$ *in* $C([0,T];H)$.

Proof. Let u_ε be the unique minimizer of $\mathrm{sc}^- I_\varepsilon$ in $K(u_{0\varepsilon})$. As we clearly have that

$$\mathrm{sc}^- I_\varepsilon(v_\varepsilon) \le I_\varepsilon(v_\varepsilon) \le \inf_{K(u_{0\varepsilon})} I_\varepsilon + \alpha_\varepsilon = \mathrm{sc}^- I_\varepsilon(u_\varepsilon) + \alpha_\varepsilon,$$

we are in the position of applying directly Theorem 5.4 to the functional $\mathrm{sc}^- I_\varepsilon$ and conclude. □

5.7. **Another choice for the artificial boundary condition in T.** The choice of the homogeneous Neumann boundary condition in T for (1.4a) is just motivated by the sake of simplicity and one may wonder if other possibilities would give rise to better convergence results. We shall not discuss here this issue in full generality but rather consider the original setting by Mielke & Ortiz [MO08] where the functional $\bar I_\varepsilon : H^1(0,T;H) \to (-\infty, \infty]$ given by

$$\bar I_\varepsilon(v) \doteq \int_0^T \mathrm{e}^{-t/\varepsilon} \left(\frac{1}{2}|v'|^2 + \frac{1}{\varepsilon}\Big(\phi(v) - (f,v)\Big)\right) \mathrm{d}t + \mathrm{e}^{-T/\varepsilon}\Big(\phi(v(T)) - (f_T, v(T))\Big),$$

for a given $f_T \in H$, are considered instead. The corresponding Euler system includes (1.4a)-(1.4b) along with the boundary condition

$$u'(T) + \partial\phi(u(T)) \ni f_T. \tag{5.11}$$

By choosing $f_T = f(T)$ for f regular, the above condition is enforcing, independently of ε, the attainment of the gradient flow equation (1.1) at the final time T.

The results of this paper can be equivalently stated for minimizers v_ε of $\bar I_\varepsilon$ in $K(u_{0\varepsilon})$ and the corresponding proofs just follow from the (sometimes technical) adaptation of the present ones to that case. In particular, the convergence $v_\varepsilon \to u$ in $C([0,T];H)$ holds. The difference in considering v_ε may be related to the fact that we impose no artificial constraint on the first time-derivative in T. On the other hand, by asking for (5.11) we are (formally) imposing $v_\varepsilon''(T) = 0$.

Despite the fact that the very same analytical results are available for the two different choices of boundary conditions in T (and that the *same* sharpness of convergence rates can be checked, see Subsection 5.4), the use of $\bar I_\varepsilon$ instead of I_ε could show some advantage in some situation. In the very specific scalar and linear case of (1.3) an illustration of the uniform convergence of v_ε is given in Figure 3. The plots in Figures 1 and 3 are produced by the same choices of ε. In particular, it is evident that that the trajectory v_ε are closer

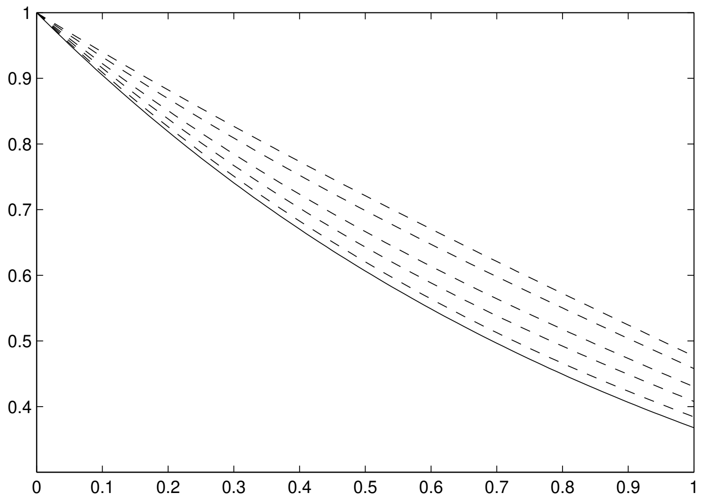

Figure 3: The convergence result in the special case (1.3). As $\varepsilon \to 0$, the minimizers of $\bar{I}_\varepsilon$ (dashed lines) approach the solution of the gradient flow (solid line).

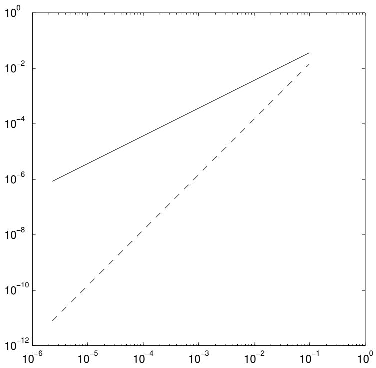

Figure 4: The functions $\varepsilon \mapsto \|u - u_\varepsilon\|_{C[0,T]}$ (solid) and $\varepsilon \mapsto \|u - v_\varepsilon\|_{C[0,T]}$ (dashed) in a log-log scale.

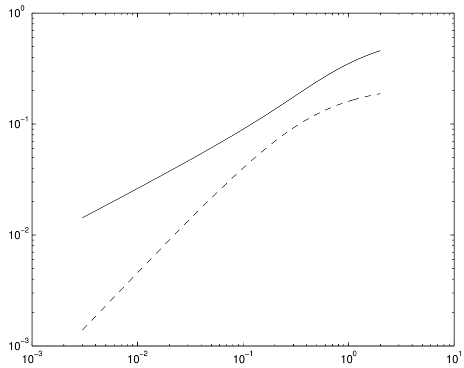

Figure 5: The functions $\varepsilon \mapsto \|u - u_\varepsilon\|_{H^1(0,T)}$ (solid) and $\varepsilon \mapsto \|u - v_\varepsilon\|_{H^1(0,T)}$ (dashed) in a log-log scale.

to u than the former u_ε. Explicit convergence rates can be easily computed for v_ε in the specific case of (1.3) from

$$|(u - v_\varepsilon)(t)| \sim 4\varepsilon^2 \left(\mathrm{e}^{(t-1)/\varepsilon - 1} - \mathrm{e}^{-1/\varepsilon - t - 1}\right) \le \frac{4\varepsilon^2}{\mathrm{e}},$$

$$\|u - v_\varepsilon\|_{H^1(0,1)} \sim \frac{2\sqrt{2}\varepsilon}{\mathrm{e}}.$$

The comparison between the convergence rates for u_ε and v_ε are reported in Figures 4-5.

5.8. **Cauchy argument.** An alternative strategy for the proof of Theorem 1.1 is that of directly checking that u_ε is a Cauchy sequence in $C([0,T];H)$. By taking the difference between the Euler equation (3.1) at level ε and the same equation at level μ, testing it on $w \doteq u_\varepsilon - u_\mu$, and integrating in time one gets that

$$\varepsilon\int_0^t |w'|^2 + \frac{1}{2}|w(t)|^2 + \int_0^t (\xi_\varepsilon - \xi_\mu, w)$$
$$= \varepsilon(w'(t), w(t)) - \varepsilon(w'(0), w(0)) + \frac{1}{2}|w(0)|^2 + (\varepsilon - \mu)\int_0^t (u''_\mu, w).$$

Let us now exploit λ-convexity and integrate by parts the last term in the above right-hand side obtaining

$$\varepsilon\int_0^t |w'|^2 + \frac{1}{2}|w(t)|^2 + \lambda\int_0^t |w|^2$$
$$\leq \varepsilon(w'(t), w(t)) - \varepsilon(w'(0), w(0)) + \frac{1}{2}|w(0)|^2$$
$$+ (\varepsilon-\mu)(u'_\mu(t), w(t)) - (\varepsilon-\mu)(u'_\mu(0), w(0)) - (\varepsilon-\mu)\int_0^t (u'_\mu, w')$$
$$= \frac{1}{2}|w(0)|^2 + (\varepsilon u'_\varepsilon(t) - \mu u'_\mu(t), w(t)) - (\varepsilon u'_\varepsilon(0) - \mu u'_\mu(0), w(0)) - (\varepsilon-\mu)\int_0^t (u'_\mu, w')$$
$$\leq |w(0)|^2 + \frac{1}{4}|w(t)|^2 + c\varepsilon^2\|u'_\varepsilon\|^2_{C([0,T],H)} + c\mu^2\|u'_\mu\|^2_{C([0,T],H)}$$
$$+ (\varepsilon+\mu)\|u'_\mu\|_{L^2(0,T;H)}\|w'\|_{L^2(0,T;H)}$$

where we have exploited Lemma 4.1. In particular, we have that

$$\varepsilon\int_0^t |w'|^2 + \frac{1}{4}|w(t)|^2 + \lambda\int_0^t |w|^2 \leq c(\varepsilon+\mu)$$

and the Cauchy character in $C([0,T];H)$ follows by Gronwall's Lemma. Once u_ε is proved to admit a strong limit u it is standard to check that indeed u solves (1.1).

The advantage of this argument with respect to the former proof of Theorem 1.1 is that it does not rely on the well-posedness of the limiting gradient flow (1.1). This fact allows us to state a modification of Theorem 1.1 as follows.

Proposition 5.6 (Convergence without lower semicontinuity)**.** *Let ϕ be proper, bounded below, λ-convex but not necessarily lower semicontinuous. Moreover let $f \in L^2(0,T;H)$, $u_0 \in D(\phi)$, $u_{0\varepsilon}$ fulfill* (2.9)*, and u_ε solve the Euler system* (3.1)-(3.3)*. Then, $u_\varepsilon \to u$ strongly in $C([0,T];H)$ and weakly in $H^1(0,T;H)$ where u is the only solution of the gradient flow*

$$u' + \overline{\partial\phi}(u) \ni f \quad a.e.\ in\ \ (0,T), \quad u(0) = u_0, \tag{5.12}$$

where $\overline{\partial\phi}$ is the strong $\times$ weak closure of $\partial\phi$ in $H \times H$, namely

$$\overline{\partial\phi}(u) \doteq \{\xi \in H \ :\ \exists (u_k, \xi_k) \to (u,\xi) \ \ strongly \times weakly\ in\ \ H\times H \ \ and\ \ \xi_k \in \partial\phi(u_k)\}.$$

Proof. As the compactness of the sequence u_ε in $C([0,T];H)$ (has well as its boundedness in $H^1(0,T;H)$) has been already established, owing to Lemma 4.1 and by passing to the limit in (3.1) we get the assertion. ∎

For the sake of illustrating the above result, let us remark that

$$\overline{\partial\phi} \subset \partial(\mathrm{sc}^-\psi), \tag{5.13}$$

the inclusion being strict. First of all, we have that $\overline{\partial\phi} = \overline{\partial\psi} + \lambda\,\mathrm{id}$, where $\overline{\partial\psi}$ is the corresponding closure of $\partial\psi$ (note that $\partial\psi$ does not coincide with $\overline{\partial\psi}$ as ψ may be not lower semicontinuous).

On the one hand, by exploiting the very definition of subdifferential and relaxation we readily get that $\overline{\partial\psi} \subset \partial(\mathrm{sc}^-\psi)$. Indeed, let $\eta \in \overline{\partial\psi}(u)$. Then there exists $(u_k, \eta_k) \to (u, \eta)$ strongly $\times$ weakly such that

$$(\eta_k, w_k - u_k) \leq \psi(w_k) - \psi(u_k) \leq \psi(w_k) - \mathrm{sc}^-\psi(u_k) \quad \forall w_k \in H.$$

Fix now $w \in H$ and choose $w_k \to w$ to be such that $\psi(w_k) \to \mathrm{sc}^-\psi(w)$. By passing to the lim inf in the above inequality we get that $\eta \in \partial(\mathrm{sc}^-\psi)(u)$.

On the other hand, let $H = \mathbb{R}$ and ϕ be defined by

$$\phi(u) \doteq \begin{cases} 0 & \text{for } \ u < 0 \\ 1 & \text{for } \ u = 0 \\ \infty & \text{otherwise} \end{cases}$$

so that we immediately compute the relaxation

$$\mathrm{sc}^-\phi(u) = \begin{cases} 0 & \text{for } \ u \leq 0 \\ \infty & \text{otherwise.} \end{cases}$$

The corresponding subdifferentials read

$$\partial\phi(u) = \begin{cases} 0 & \text{for } \ u < 0 \\ \emptyset & \text{otherwise,} \end{cases}$$

$$\overline{\partial\phi}(u) = \begin{cases} 0 & \text{for } \ u \leq 0 \\ \emptyset & \text{otherwise,} \end{cases}$$

$$\partial(\mathrm{sc}^-\phi)(u) = \begin{cases} 0 & \text{for } \ u \leq 0 \\ [0,\infty) & \text{for } \ u = 0 \\ \emptyset & \text{otherwise.} \end{cases}$$

In particular, the inclusion in (5.13) is strict.

Note that, from the one hand, Proposition 5.6 is more general than Theorem 1.1 as the lower semicontinuity assumption on ϕ is dropped. This would in principle open the way to relaxation. On the other hand, Proposition 5.6 directly assumes the existence of solutions to the Euler system (3.1)-(3.3), a circumstance that we check for lower semicontinuous functionals only (see Theorem 3.1).

6. Time-discretization

The convergence result of Theorem 1.1 can be efficiently combined with time-discretization which, in turn, provides a sound frame for the proof of Lemma 4.1 out of the regular case of (4.2).

We start by recalling the notation for the constant time-step $\tau = T/n$ and introducing the functional $I_{\varepsilon\tau}$ defined on discrete trajectories $(v^0, \dots, v^n) \in H^{n+1}$ as

$$I_{\varepsilon\tau}(v^0, \dots, v^n) = \sum_{i=1}^{n} \rho^i_{\varepsilon\tau} \frac{\tau}{2} \left| \frac{v^i - v^{i-1}}{\tau} \right|^2 + \sum_{i=1}^{n-1} (\rho^i_{\varepsilon\tau} - \rho^{i+1}_{\varepsilon\tau}) \Big(\phi(v^i) - (f^i, v^i) \Big).$$

Here, the weights $(\rho^1_{\varepsilon\tau}, \dots, \rho^n_{\varepsilon\tau})$ are given by

$$\rho^i_{\varepsilon\tau} = \left(\frac{\varepsilon}{\tau + \varepsilon} \right)^i \quad \text{for} \quad i = 1, \dots, n. \tag{6.1}$$

In particular, $(\rho^1_{\varepsilon\tau}, \dots, \rho^n_{\varepsilon\tau})$ is nothing but the solution of the constant time-step implicit Euler discretization of the problem $\rho' + \rho/\varepsilon = 0$ with initial condition $\rho(0) = 1$. As this choice ensures that, for $i = 1, \dots, n-1$,

$$\rho^i_{\varepsilon\tau} - \rho^{i+1}_{\varepsilon\tau} = \frac{\tau}{\varepsilon} \rho^{i+1}_{\varepsilon\tau} > 0, \tag{6.2}$$

and we have by Lemma 2.2 that $\rho_{\varepsilon\tau}(t) \to \mathrm{e}^{-t/\varepsilon}$ uniformly as $\tau \to 0$ for $\varepsilon > 0$ fixed, the functional $I_{\varepsilon\tau}$ may be regarded as a quadrature of the time-continuous functional I_ε.

Before moving on, let us motivate our specific choice for the functional $I_{\varepsilon\tau}$. First of all, we recall that the incremental minimization scheme of (2.12) in equivalent to

$$u^0 = u_0 \quad \text{and} \quad u^i = \operatorname*{Arg\,min}_{u \in H} \left(\frac{1}{2} \left| \frac{u - u^{i-1}}{\tau} \right|^2 + \frac{\phi(u) - (f^i, u)}{\tau} - \frac{\phi(u^{i-1}) - (f^{i-1}, u^{i-1})}{\tau} \right)$$
$$\text{for} \quad i = 1, \dots, n. \tag{6.3}$$

Indeed, the latter is nothing but (2.12) where, at each level i, we have added the inconsequential term $-(\phi(u^{i-1}) - (f^{i-1}, u^{i-1}))/\tau$. Here, the point $f^0 \in H$ is assumed to be given (its actual value being irrelevant).

The latter minimization problems are usually solved *sequentially*. On the other hand, a direct computation shows that

$$\begin{aligned} & I_{\varepsilon\tau}(v^0, \dots, v^n) \\ & = \sum_{i=1}^{n} \rho^i_{\varepsilon\tau} \tau \left(\frac{1}{2} \left| \frac{v^i - v^{i-1}}{\tau} \right|^2 + \frac{\phi(v^i) - (f^i, v^i)}{\tau} - \frac{\phi(v^{i-1}) - (f^{i-1}, v^{i-1})}{\tau} \right) \\ & \quad - \rho^n_{\varepsilon\tau} \Big(\phi(v^n) - (f^n, v^n) \Big) + \rho^1_{\varepsilon\tau} \Big(\phi(v^0) - (f^0, v^0) \Big). \end{aligned} \tag{6.4}$$

Hence, the minimization of $I_{\varepsilon\tau}$ in $K_\tau(u_{0\varepsilon}) \doteq \{(v^0, \dots, v^n) \in H^{n+1} \ : \ v^0 = u_{0\varepsilon}\}$ roughly corresponds to collect all the minimization problems in (6.3) in a single constrained minimization problem for the entire discrete trajectory $(u^0, \dots, u^n)$. This in particular motivates our reference to the values $\rho^i_{\varepsilon\tau}$ as *Pareto weights* in analogy with the corresponding notion in multi-objective optimization [Cla90]. More specifically, as $\rho^1_{\varepsilon\tau} \gg \rho^2_{\varepsilon\tau} \gg \dots \gg \rho^n_{\varepsilon\tau}$ for $\varepsilon \to 0$, it turns out that, by minimizing $I_{\varepsilon\tau}$, a much larger priority is accorded to the first minimum problem in (6.3) with respect to the second, to the second with respect to the third, and so on. Hence, the limit $\varepsilon \to 0$ again formally corresponds to causality restoring, see also [MS08].

Exactly as in the time-continuous situation, in case ϕ is convex, the functional $I_{\varepsilon\tau}$ turns out to be uniformly convex for all ε. In particular, a unique minimizer of $I_{\varepsilon\tau}$ in $K_\tau(w_0)$

exists for all $w_0 \in H$. The same holds true for general λ-convex functionals whenever ε and τ are chosen to be small enough. Indeed, we have the following.

Proposition 6.1 (Well-posedness of the discrete minimum problem)**.** *For ε and τ small and all $w_0 \in H$, the functional $I_{\varepsilon\tau}$ admits a unique minimizer in $K_\tau(w_0)$.*

Proof. This argument is the discrete analogue of the proof of Proposition 2.1. In particular, we start by decomposing $I_{\varepsilon\tau}$ into a quadratic part $Q_{\varepsilon\tau}$ and a convex remainder $R_{\varepsilon\tau}$ as

$$
\begin{aligned}
&I_{\varepsilon\tau}(u^0,\ldots,u^n)\\
&= \left(\sum_{i=1}^{n} \frac{\tau}{2}\rho^i_{\varepsilon\tau}\left|\delta u^i\right|^2 - \sum_{i=1}^{n-1}(\rho^i_{\varepsilon\tau}-\rho^{i+1}_{\varepsilon\tau})\frac{\lambda^-}{2}|u^i|^2\right) + \left(\sum_{i=1}^{n-1}(\rho^i_{\varepsilon\tau}-\rho^{i+1}_{\varepsilon\tau})\Big(\psi(u^i)-(f^i,u^i)\Big)\right)\\
&\doteq Q_{\varepsilon\tau}(u^0,\ldots,u^n) + R_{\varepsilon\tau}(u^0,\ldots,u^n).
\end{aligned}
$$

The result follows by checking that, for small ε and τ, the functional $Q_{\varepsilon\tau}$ is uniformly convex on $K_\tau(w_0)$. To this aim, for all $(u^0,\ldots,u^n) \in K(w_0)$ let $(v^0,\ldots,v^n)$ be defined as $v^i \doteq \sqrt{\rho^i_{\varepsilon\tau}}u^i$. Then, we compute that (see (2.5))

$$
\delta u^i = \frac{1}{\sqrt{\rho^{i-1}_{\varepsilon\tau}}}\delta v^i + v^i\delta\left(\frac{1}{\sqrt{\rho^i_{\varepsilon\tau}}}\right) = \frac{1}{\sqrt{\rho^i_{\varepsilon\tau}}}\left(r_{\varepsilon\tau}\delta v^i + \frac{1-r_{\varepsilon\tau}}{\tau}v^i\right)
$$

where $r_{\varepsilon\tau} \doteq \sqrt{\varepsilon/(\varepsilon+\tau)}$. By observing that $\rho^i_{\varepsilon\tau}-\rho^{i+1}_{\varepsilon\tau} = \rho^i_{\varepsilon\tau}(1-r^2_{\varepsilon\tau})$, the value $Q_{\varepsilon\tau}(u^0,\ldots,u^n)$ can hence be rewritten as

$$
\begin{aligned}
Q_{\varepsilon\tau}(u^0,\ldots,u^n) &= \sum_{i=1}^{n}\frac{\tau}{2}\left(r^2_{\varepsilon\tau}|\delta v^i|^2 + \frac{(1-r_{\varepsilon\tau})^2}{\tau^2}|v^i|^2 + \frac{2r_{\varepsilon\tau}(1-r_{\varepsilon\tau})}{\tau}(\delta v^i,v^i)\right)\\
&\quad - \sum_{i=1}^{n-1}\frac{(1-r^2_{\varepsilon\tau})\lambda^-}{2}|v^i|^2\\
&= \sum_{i=1}^{n}\frac{\tau r^2_{\varepsilon\tau}}{2}|\delta v^i|^2 + \frac{(1-r_{\varepsilon\tau})^2}{2\tau}|v^n|^2 + \sum_{i=1}^{n-1}\left[\frac{(1-r_{\varepsilon\tau})^2}{2\tau} - \frac{(1-r^2_{\varepsilon\tau})\lambda^-}{2}\right]|v^i|^2\\
&\quad + \frac{r_{\varepsilon\tau}(1-r_{\varepsilon\tau})}{\tau}\left(\frac{1}{2}|v^n|^2 + \frac{1}{2}\sum_{i=1}^{n}|v^i-v^{i-1}|^2 - \frac{1}{2}|w_0|^2\right).
\end{aligned}
$$

As we readily check that

$$
\left[\frac{(1-r_{\varepsilon\tau})^2}{2\tau} - \frac{(1-r^2_{\varepsilon\tau})\lambda^-}{2}\right] \;\to\; \frac{1}{2\tau} - \frac{\lambda^-}{2} \quad \text{as} \quad \varepsilon \to 0,
$$

for all ε and τ small (depending on λ^- only), the functional $Q_{\varepsilon\tau}$ turns out to be uniformly convex in $K(w_0)$. ☐

The main result of this section is the convergence of minimizers of the time-discrete functional $I_{\varepsilon\tau}$ to solutions of the gradient flow (1.1) as the time-step τ and the causal parameter ε go to 0. To this aim, we assume for the very beginning that $\overline{f}_\tau \to f$ strongly in $L^2(0,T;H)$. This convergence holds, for instance, if $\overline{f}_\tau$ is built on local means. We have the following.

Theorem 6.2 (Convergence + discretization)**.** *$u_{\varepsilon\tau} \to u$ in $C([0,T];H)$ as $\varepsilon + \tau \to 0$.*

6.1. **Discrete Euler equation.** The functional $I_{\varepsilon\tau}$ is the quadratic perturbation of a convex functional. Hence, its Fréchet subdifferential is readily computed and, letting $(u_\varepsilon^0, \dots, u_\varepsilon^n)$ be the minimizer of $I_{\varepsilon\tau}$ in $K_\tau(u_{0\varepsilon})$, from $0 \in \partial I_{\varepsilon\tau}(u_\varepsilon^0, \dots, u_\varepsilon^n)$ we have that there exist $\xi_\varepsilon^i \in \partial\phi(u_\varepsilon^i)$, $i = 1, \dots, n-1$, such that

$$0 \in \sum_{i=1}^{n} \rho_{\varepsilon\tau}^i \tau(\delta u_\varepsilon^i, \delta v^i) + \sum_{i=1}^{n-1} (\rho_{\varepsilon\tau}^i - \rho_{\varepsilon\tau}^{i+1})(\xi_\varepsilon^i - f^i, v^i) \quad \forall (v^0, \dots, v^n) \in K_\tau(0).$$

The first term in the above right-hand side reads (recall that $v^0 = 0$)

$$\sum_{i=1}^{n} \rho_{\varepsilon\tau}^i \tau(\delta u_\varepsilon^i, \delta v^i) = \sum_{i=1}^{n-1} (\rho_{\varepsilon\tau}^i \delta u_\varepsilon^i - \rho_{\varepsilon\tau}^{i+1} \delta u_\varepsilon^{i+1}, v^i) + \rho_{\varepsilon\tau}^n (\delta u_\varepsilon^n, v^n),$$

and, by using (6.2), we have that

$$\begin{aligned} \rho_{\varepsilon\tau}^i \delta u_\varepsilon^i - \rho_{\varepsilon\tau}^{i+1} \delta u_\varepsilon^{i+1} &= (\rho_{\varepsilon\tau}^i - \rho_{\varepsilon\tau}^{i+1}) \delta u_\varepsilon^i + \rho_\varepsilon^{i+1} (\delta u_\varepsilon^i - \delta u_\varepsilon^{i+1}) \\ &= \frac{\tau}{\varepsilon} \rho_{\varepsilon\tau}^{i+1} \delta u_\varepsilon^i + \rho_{\varepsilon\tau}^{i+1} (\delta u_\varepsilon^i - \delta u_\varepsilon^{i+1}). \end{aligned}$$

On the other hand, again form (6.2) we have that

$$\sum_{i=1}^{n-1} (\rho_{\varepsilon\tau}^i - \rho_{\varepsilon\tau}^{i+1})(\xi_\varepsilon^i - f^i, v^i) = \sum_{i=1}^{n-1} \frac{\tau}{\varepsilon} \rho_\varepsilon^{i+1} (\xi_\varepsilon^i - f^i, v^i).$$

Hence, the minimizer $(u_\varepsilon^0, \dots, u_\varepsilon^n)$ of $I_{\varepsilon\tau}$ in $K_\tau(u_{0\varepsilon})$ fulfills the discrete Euler equation (see (3.1)-(3.4))

$$-\varepsilon \frac{\delta u_\varepsilon^{i+1} - \delta u_\varepsilon^i}{\tau} + \delta u_\varepsilon^i + \xi_\varepsilon^i = f^i \quad \text{for } \ i = 1, \dots, n-1, \tag{6.5a}$$

$$u_\varepsilon^0 = u_{0\varepsilon}, \tag{6.5b}$$

$$\delta u_\varepsilon^n = 0, \tag{6.5c}$$

$$\xi_\varepsilon^i \in \partial\phi(u_\varepsilon^i) \quad \text{for } \ i = 1, \dots, n-1. \tag{6.5d}$$

6.2. **Key estimate at the discrete level.** Our next aim is to reproduce at the discrete level the key estimate (4.1). Note that the regularity in time in (4.2) is not needed here as integration by parts is here replaced by summation.

Let us fix with no loss of generality $\xi_\varepsilon^0 = (\partial\phi(u_{0\varepsilon}))^\circ$. For the sake of notational simplicity, we let $v_\varepsilon^i = \delta u_\varepsilon^{i+1}$ for $i = 1, \dots, n-1$. The Euler equation (6.5a) ensures that

$$\sum_{i=1}^{n-1} \tau \left| -\varepsilon \delta v_\varepsilon^i + \delta u_\varepsilon^i + \xi_\varepsilon^i \right|^2 = \sum_{i=1}^{n-1} \tau |f^i|^2$$

and we shall now proceed to the proof of separate bounds on the three terms in the left-hand side above. In particular, we have that

$$
\begin{aligned}
&\sum_{i=1}^{n-1}\tau\Big(\varepsilon^2|\delta v_\varepsilon^i|^2+|\delta u_\varepsilon^i|^2+|\xi_\varepsilon^i|^2\Big)\\
&=\sum_{i=1}^{n-1}\tau|-\varepsilon\delta v_\varepsilon^i+\delta u_\varepsilon^i+\xi_\varepsilon^i|^2+2\varepsilon\sum_{i=1}^{n-1}\tau(\delta v_\varepsilon^i,\delta u_\varepsilon^i)-2\sum_{i=1}^{n-1}\tau(\delta u_\varepsilon^i,\xi_\varepsilon^i)+2\varepsilon\sum_{i=1}^{n-1}\tau(\delta v_\varepsilon^i,\xi_\varepsilon^i)\\
&=\sum_{i=1}^{n-1}\tau|f^i|^2+2\varepsilon\sum_{i=1}^{n-1}\tau(\delta v_\varepsilon^i,\delta u_\varepsilon^i)-2\sum_{i=1}^{n-1}\tau(\delta u_\varepsilon^i,\xi_\varepsilon^i)+2\varepsilon\sum_{i=1}^{n-1}\tau(\delta v_\varepsilon^i,\xi_\varepsilon^i).
\end{aligned}
\tag{6.6}
$$

We now aim at controlling the last three terms in the above right-hand side. We have that

$$
2\varepsilon\sum_{i=1}^{n-1}\tau(\delta v_\varepsilon^i,\delta u_\varepsilon^i)=-\varepsilon|\delta u_\varepsilon^1|^2-\varepsilon\sum_{i=1}^{n-1}|v_\varepsilon^i-v_\varepsilon^{i-1}|.
\tag{6.7}
$$

Moreover, by exploiting λ-convexity, one computes that

$$
\begin{aligned}
-2\sum_{i=1}^{n-1}\tau(\delta u_\varepsilon^i,\xi_\varepsilon^i)&=2\sum_{i=1}^{n-1}(u_\varepsilon^{i-1}-u_\varepsilon^i,\xi_\varepsilon^i)\\
&\leq 2\sum_{i=1}^{n-1}\Big(\psi(u_\varepsilon^{i-1})-\psi(u_\varepsilon^i)-\frac{\lambda}{2}|u_\varepsilon^i|^2+\frac{\lambda}{2}|u_\varepsilon^{i-1}|^2-\frac{\lambda}{2}|u_\varepsilon^i-u_\varepsilon^{i-1}|^2\Big)\\
&=2\phi(u_{0\varepsilon})-2\phi(u_\varepsilon^{n-1})-\lambda\sum_{i=1}^{n-1}|u_\varepsilon^i-u_\varepsilon^{i-1}|^2\\
&=2\phi(u_{0\varepsilon})-2\phi(u_\varepsilon^{n-1})-\lambda\tau\sum_{i=1}^{n-1}\tau|\delta u_\varepsilon^i|^2.
\end{aligned}
\tag{6.8}
$$

Finally, again λ-convexity ensures that

$$
\begin{aligned}
2\varepsilon\sum_{i=1}^{n-1}\tau(\delta v_\varepsilon^i,\xi_\varepsilon^i)&=-2\varepsilon(\delta u_\varepsilon^1,\xi_\varepsilon^1)-2\varepsilon\sum_{i=2}^{n-1}\tau(\delta u_\varepsilon^i,\delta\xi_\varepsilon^i)\\
&\leq -2\varepsilon(\delta u_\varepsilon^1,\xi_\varepsilon^0)-2\lambda\varepsilon\sum_{i=1}^{n-1}\tau|\delta u_\varepsilon^i|^2.
\end{aligned}
\tag{6.9}
$$

Now, taking into account (6.7)-(6.8), estimate (6.6) becomes

$$
\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{n-1}\tau\Big(\varepsilon^2|\delta v_\varepsilon^i|^2+\big(1+\lambda(2\varepsilon+\tau)\big)|\delta u_\varepsilon^i|^2+|\xi_\varepsilon^i|^2\Big)\\
&+\frac{\varepsilon}{2}|\delta u_\varepsilon^1|^2+\frac{\varepsilon}{2}\sum_{i=1}^{n-1}|v_\varepsilon^i-v_\varepsilon^{i-1}|+\phi(u_\varepsilon^{n-1})\\
&\leq\frac{1}{2}\sum_{i=1}^{n-1}\tau|f^i|^2-\varepsilon(\delta u_\varepsilon^1,\xi_\varepsilon^0)+\phi(u_{0\varepsilon}).
\end{aligned}
\tag{6.10}
$$

In particular, as soon as $\lambda^-(2\varepsilon+\tau)\le 1/2$ (see (4.4)), we recall (2.9)-(2.10) and conclude that

$$\varepsilon^2\sum_{i=1}^{n-1}\tau\left|\delta v_\varepsilon^i\right|^2+\sum_{i=1}^{n}\tau|\delta u_\varepsilon^i|^2+\sum_{i=1}^{n-1}\tau|\xi_\varepsilon^i|^2\le c, \tag{6.11}$$

where c depends on $\|f\|_{L^2(0,T;H)}$, $|u_0|$, and c_0.

We now aim at reproducing estimate (4.6). Let for brevity $z_\varepsilon^i=\delta u_\varepsilon^i=v_\varepsilon^{i-1}$. Estimate (6.11) yields that both $\overline{z}_{\varepsilon\tau}$ and $\varepsilon v'_{\varepsilon\tau}$ are bounded in $L^2(0,T-\tau;H)$ independently of ε and τ. Let us now handle the difference $\overline{z}_{\varepsilon\tau}-v_{\varepsilon\tau}$ as follows.

$$\begin{aligned}\|\overline{z}_{\varepsilon\tau}-v_{\varepsilon\tau}\|^2_{L^2(0,T-\tau;H)}&=\sum_{i=1}^{n-1}\int_{(i-1)\tau}^{i\tau}|z_\varepsilon^i-(\alpha^i(t)z_\varepsilon^{i+1}+(1-\alpha^i(t))z_\varepsilon^i)|^2\mathrm{d}t\\&=\sum_{i=1}^{n-1}\left(\int_{(i-1)\tau}^{i\tau}(\alpha^i(t))^2\mathrm{d}t\right)|z_\varepsilon^{i+1}-z_\varepsilon^i|^2=\frac{\tau^2}{3}\sum_{i=1}^{n-1}\tau\left|\frac{z_\varepsilon^{i+1}-z_\varepsilon^i}{\tau}\right|^2\\&=\frac{\tau^2}{3}\sum_{i=1}^{n-1}\tau|\delta v^i|^2=\frac{\tau^2}{3}\|v'_{\varepsilon\tau}\|^2_{L^2(0,T-\tau;H)}.\end{aligned}\tag{6.12}$$

In particular, we have that

$$\|v_{\varepsilon\tau}\|_{L^2(0,T-\tau;H)}\le\|\overline{z}_{\varepsilon\tau}\|_{L^2(0,T-\tau;H)}+\frac{\tau}{\sqrt{3}}\|v'_{\varepsilon\tau}\|_{L^2(0,T-\tau;H)}\le c\left(1+\frac{\tau}{\varepsilon}\right)$$

Hence, we have

$$\begin{aligned}\|v_{\varepsilon\tau}\|_{C^0([0,T-\tau];H)}&\le c\|v_{\varepsilon\tau}\|^{1/2}_{L^2(0,T-\tau;H)}\|v_{\varepsilon\tau}\|^{1/2}_{H^1(0,T-\tau;H)}\\&\le c\left(1+\frac{\tau}{\varepsilon}\right)^{1/2}\left(\left(1+\frac{\tau}{\varepsilon}\right)^2+\frac{1}{\varepsilon^2}\right)^{1/4}\\&\le c\left(1+\frac{\tau^{1/2}}{\varepsilon^{1/2}}\right)\left(1+\frac{\tau^{1/2}}{\varepsilon^{1/2}}+\frac{1}{\varepsilon^{1/2}}\right)\ \le\ c\left(1+\frac{1}{\varepsilon^{1/2}}+\frac{\tau^{1/2}}{\varepsilon}\right).\end{aligned}\tag{6.13}$$

This bound is the discrete counterpart to (4.6) (recall (2.9)).

6.3. **Proof of Theorem 6.2.** This argument is nothing but the discrete analogue of the proof of Theorem 1.1. Let $(u^0,u^1,\dots,u^n)$ solve the implicit Euler scheme (2.11) and $(u_\varepsilon^0,u_\varepsilon^1,\dots,u_\varepsilon^n)$ minimize $I_{\varepsilon\tau}$ locally in $K_\tau(u_{0\varepsilon})$. Test (2.11) by $w_\varepsilon^i=u^i-u_\varepsilon^i$ getting

$$(\delta u^i,w_\varepsilon^i)+\phi(u^i)+\frac{\lambda}{2}|w_\varepsilon^i|\le\phi(u_\varepsilon^i)+(f^i,w_\varepsilon^i)\quad\text{for}\ \ i=1,\dots,n-1.$$

Test now (6.5a) by $-w_\varepsilon^i$ and obtain that

$$\begin{aligned}&-\frac{\varepsilon}{\tau}(\delta u_\varepsilon^{i+1}-\delta u_\varepsilon^i,-w_\varepsilon^i)+(\delta u_\varepsilon^i,-w_\varepsilon^i)+\phi(u_\varepsilon^i)+\frac{\lambda}{2}|w_\varepsilon^i|\le\phi(u^i)-(f^i,w_\varepsilon^i)\\&\qquad\text{for}\ \ i=1,\dots,n-1.\end{aligned}\tag{6.14}$$

Take the sum of the last two inequalities, multiply it by τ, and sum for $i=1,\dots,m\le n-1$ getting

$$\varepsilon\sum_{i=1}^{m}(\delta u_\varepsilon^{i+1}-\delta u_\varepsilon^i,w_\varepsilon^i)+\sum_{i=1}^{m}\tau(\delta w_\varepsilon^i,w_\varepsilon^i)+\lambda\sum_{i=1}^{m}\tau|w_\varepsilon^i|^2\le 0.\tag{6.15}$$

We easily handle the first term above by computing

$$\varepsilon\sum_{i=1}^{m}(\delta u_\varepsilon^{i+1}-\delta u_\varepsilon^i,w_\varepsilon^i)=-\varepsilon\sum_{i=1}^{m}\tau(\delta u_\varepsilon^i,\delta w_\varepsilon^i)+\varepsilon(\delta u_\varepsilon^{m+1},w_\varepsilon^m)-\varepsilon(\delta u_\varepsilon^1,w_\varepsilon^0).$$

Hence, (6.15) entails that

$$\begin{aligned}&\varepsilon\sum_{i=1}^{m}\tau|\delta w_\varepsilon^i|^2+\frac{1}{2}|w_\varepsilon^m|^2\\&\leq\frac{1}{2}|w_\varepsilon^0|^2+\lambda^-\sum_{i=1}^{m}\tau|w_\varepsilon^i|^2+\varepsilon\sum_{i=1}^{m}\tau(\delta u^i,\delta w_\varepsilon^i)-\varepsilon(\delta u_\varepsilon^{m+1},w_\varepsilon^m)+\varepsilon(\delta u_\varepsilon^1,w_\varepsilon^0).\end{aligned}$$

In particular, by letting $4\lambda^-\tau<1$, as u_τ' is bounded in $L^2(0,T;H)$ independently of τ, we have proved by the discrete Gronwall Lemma that

$$|w_\varepsilon^m|^2\leq c\big(\varepsilon+\varepsilon^2|\delta u_{\varepsilon\tau}^{m+1}|^2+\varepsilon^2|\delta u_\varepsilon^1|^2\big)$$

for some constant $c>0$ depending also on λ^- but independent of ε and τ. Recall now the bound (6.13) and obtain that

$$|w_\varepsilon^m|^2\leq c\left(\varepsilon+\varepsilon^2\left(1+\varepsilon^{-1/2}+\frac{\tau^{1/2}}{\varepsilon}\right)^2\right)\leq c(\varepsilon+\tau).$$

Hence, we have checked that $\max_{i=1,\ldots,n-1}|w_\varepsilon^i|\leq c(\varepsilon+\tau)$. In fact, this bound can be extended to $i=n$ as $w_\varepsilon^n=w_\varepsilon^{n-1}$. In particular, we have proved that $w_{\varepsilon\tau}\to 0$ in $C([0,T];H)$ as $\varepsilon+\tau\to 0$. Finally, the strong convergence $u_{\varepsilon\tau}\to u$ in $C([0,T];H)$ follows from Lemma 2.2.

More specifically, in case ϕ is lower semicontinuous and $f\equiv 0$, by exploiting the error control in (2.13)-(2.14), we have proved the joint convergence rates

$$\|u-u_{\varepsilon\tau}\|_{C([0,T];H)}\leq c\big(\varepsilon+\tau\big)^{1/2}.$$

Note that, in this case, the sub-optimality of the rate $\tau^{1/2}$ is already expected for the Euler scheme (recall (2.13)). Namely, the present functional approach is not deteriorating convergence with respect to the time-step size. Let us mention that the above joint convergence result can be specialized for establishing quantitative convergence in interpolation spaces and allowing for less-regular initial data in the spirit of Subsections 5.1 and 5.2.

6.4. **Limit $\tau\to 0$ for $\varepsilon>0$: convergence to the Euler equation.** By letting $\lambda\geq 0$ and $\varepsilon>0$ be fixed and passing to the limit in the time-step τ we can prove the following.

Theorem 6.3 ($\tau\to 0$ for $\varepsilon>0$)**.** *Let $\lambda\geq 0$ and $(u_\varepsilon^0,\ldots,u_\varepsilon^n)$, $(\xi_\varepsilon^1,\ldots,\xi_\varepsilon^{n-1})$ solve* (6.5)*. Then, there exists non-relabeled subsequences such that $u_{\varepsilon\tau}\to u_\varepsilon$ weakly in $H^1(0,T;H)$ and $\overline{\xi}_{\varepsilon\tau}\to\xi$ weakly in $L^2(0,T;H)$ where $(u_\varepsilon,\xi_\varepsilon)$ solves* (3.1)-(3.4)*.*

Sketch of the proof. Let $(u_\varepsilon^i,\xi_\varepsilon^i)\in H^{N+1}\times H^{N-1}$ solve (6.5) and define $v_\varepsilon^i=\delta u_\varepsilon^{i+1}$ for $i=1,\ldots,n-1$. Our first aim is to pass to the limit in the discrete equations (6.5a), (6.5c)-(6.5d) written in the compact form

$$-\varepsilon v'_{\varepsilon\tau} + u'_{\varepsilon\tau} + \overline{\xi}_{\varepsilon\tau} = \overline{f}_{\varepsilon\tau} \quad \text{a.e. in } (0, T-\tau), \tag{6.16}$$

$$v_{\varepsilon\tau}(T-\tau) = 0, \tag{6.17}$$

$$\overline{\xi}_{\varepsilon\tau} \in \partial\phi(\overline{u}_{\varepsilon\tau}) \quad \text{a.e. in } (0, T-\tau). \tag{6.18}$$

Owing to estimates (6.11)-(6.12) we find a pair $(u_\varepsilon, \xi_\varepsilon)$ such that, by extracting not relabeled subsequences (and possibly considering standard projections for $t > T - \tau$),

$$u_{\varepsilon\tau} \to u_\varepsilon \quad \text{weakly in } H^1(0,T;H),$$
$$v_{\varepsilon\tau} \to u'_\varepsilon \quad \text{weakly in } H^1(0,T;H),$$
$$\overline{\xi}_{\varepsilon\tau} \to \xi_\varepsilon \quad \text{weakly in } L^2(0,T;H).$$

The above convergences suffice for ensuring that equations (3.1)-(3.3) hold. Moreover, we have

$$\begin{aligned}
&\limsup_{\tau\to0} \int_0^{T-\tau} (\overline{\xi}_{\varepsilon\tau}, \overline{u}_{\varepsilon\tau}) \\
&\leq \limsup_{\tau\to0} \Big(-\varepsilon \int_0^{T-\tau} |u'_{\varepsilon\tau}|^2 - \varepsilon(v_{\varepsilon\tau}(0), u_{0\varepsilon}) - \frac{1}{2}|u_{\varepsilon\tau}(T-\tau)|^2 + \frac{1}{2}|u_{0\varepsilon}|^2 + \int_0^{T-\tau} (\overline{f}_\tau, \overline{u}_{\varepsilon\tau}) \Big) \\
&\leq -\varepsilon \int_0^T |u'|^2 - \varepsilon(u'(0), u_{0\varepsilon}) - \frac{1}{2}|u(T)|^2 + \frac{1}{2}|u_{0\varepsilon}|^2 + \int_0^T (f, u) \\
&= \int_0^T (\xi, u),
\end{aligned} \tag{6.19}$$

and the inclusion (3.4) follows again from the classical [Bre73b, Prop. 2.5, p. 27]. □

Note that the above proof can be adapted to the non-convex case $\lambda < 0$ by additionally requiring some compactness on the sublevels of ϕ. Hence, the extracted sequences would fulfill the strong convergence [Sim87, Cor. 4], namely

$$u_{\varepsilon\tau} \to u_\varepsilon \quad \text{strongly in } C([0,T];H).$$

This convergence suffices in order to pass to the limit in

$$\overline{\xi}_{\varepsilon\tau} - \lambda\overline{u}_{\varepsilon\tau} \in \partial\psi(\overline{u}_{\varepsilon\tau}) \quad \text{a.e. in } (0,T)$$

and get that

$$\xi_\varepsilon - \lambda u_\varepsilon \in \partial\psi(u_\varepsilon) \quad \text{a.e. in } (0,T).$$

Namely, inclusion (3.4) holds.

6.5. **Proof of the key estimate.** Let us finally come to the proof of Lemma 4.1. By reconsidering the argument of Subsection 6.2 and Theorem 6.3 we readily have that, given $g \in L^2(0,T;H)$, the solution $(u_\varepsilon, \eta_\varepsilon) \in H^2(0,T;H) \times L^2(0,T;H)$ of

$$-\varepsilon u''_\varepsilon + u'_\varepsilon + \eta_\varepsilon = g \quad \text{a.e. in } (0,T), \tag{6.20}$$

$$u_\varepsilon(0) = u_{0\varepsilon}, \tag{6.21}$$

$$u'_\varepsilon(T) = 0, \tag{6.22}$$

$$\eta_\varepsilon \in \partial\psi(u_\varepsilon) \quad \text{a.e. in } (0,T) \tag{6.23}$$

is the limit (the component u_ε being unique) of a discrete problem which in turn fulfills the expected estimates. In particular, by passing to the limit we find that there exists a positive constant $c > 0$ depending on $|g|_{L^2(0,T;H)}$, $|u_0|$, and c_0 such that

$$\varepsilon\|u''_\varepsilon\|_{L^2(0,T;H)} + \|u'_\varepsilon\|_{L^2(0,T;H)} + \|\eta_\varepsilon\|_{L^2(0,T;H)} \le c. \tag{6.24}$$

Moreover, arguing exactly as in Subsection 4.1, we also have that

$$\varepsilon^{1/2}\,\|u'_\varepsilon\|_{L^\infty(0,T;H)} \le c. \tag{6.25}$$

Take now u_ε to be the minimizer of I_ε on $K(u_{0\varepsilon})$. Owing to Theorem 3.1 we have that, indeed, u_ε solves (6.20)-(6.23) (along with the associated selection $\eta_\varepsilon = \xi_\varepsilon - \lambda u_\varepsilon$) with the datum g replaced by $f - \lambda u_\varepsilon$. Hence, in order to conclude for Lemma 4.1, what we are actually left to prove is that the norm $\|u_\varepsilon\|_{L^2(0,T;H)}$ is uniformly bounded in terms of data for all minimizers. This is however a standard estimation argument. Test (3.1) by $u_\varepsilon + \alpha u'_\varepsilon$ ($\alpha \ge 0$ to be determined later) and integrate in time getting

$$\begin{aligned}&\frac{\alpha\varepsilon}{2}|u'_\varepsilon(0)|^2 + (\varepsilon+\alpha)\int_0^T |u'_\varepsilon|^2 + \frac{1}{2}|u_\varepsilon(T)|^2 - \lambda^-\int_0^T |u_\varepsilon|^2 + \alpha\phi(u_\varepsilon(T))\\ &\quad \le -\varepsilon(u'_\varepsilon(0), u_{0\varepsilon}) + \int_0^T (f, u_\varepsilon + \alpha u'_\varepsilon) + \frac{1}{2}|u_{0\varepsilon}|^2 + \alpha\phi(u_{0\varepsilon}).\end{aligned} \tag{6.26}$$

By taking α large enough (precisely, by taking α/λ^- (for $\lambda \neq 0$) strictly larger than the first eigenvalue of the one-dimensional Laplacian in $(0,T)$ with non-homogeneous Dirichlet and homogeneous Neumann conditions in 0 and T, respectively) we conclude for

$$\|u_\varepsilon\|_{H^1(0,T;H)} \le c$$

where now $c > 0$ depends on $|f|_{L^2(0,T;H)}$, $|u_0|$, c_0, and λ^-.

7. Applications

7.1. **Linear parabolic PDEs.** Let the bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$ be given and $f \in L^2(\Omega\times(0,T))$ and $u_0 \in H^2(\Omega)\cap H^1_0(\Omega)$. Then, the minimizers u_ε in $K(u_0)$ of the WED functionals given by

$$u \mapsto \begin{cases} \displaystyle\int_0^T\!\!\int_\Omega \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}u_t^2 + \frac{1}{2\varepsilon}|\nabla u|^2 - \frac{1}{\varepsilon}fu\right) & \text{for } \ u \in L^2(0,T;H^1_0(\Omega))\\ \infty & \text{otherwise}\end{cases}$$

converge to the solution of the *heat equation*

$$u_t - \Delta u = f \quad \text{a.e. in } \ \Omega\times(0,T) \tag{7.1}$$

supplemented with the initial condition and with homogeneous Dirichlet conditions (other boundary conditions can be considered as well) in the following sense

$$\max_{t\in[0,T]}\|u(t) - u_\varepsilon(t)\|_{L^2(\Omega)} \le c\varepsilon^{1/2}, \tag{7.2}$$

$$\|u - u_\varepsilon\|_{H^\eta(0,T;L^2(\Omega))} \le c\varepsilon^{(1-\eta)/2} \qquad \text{for all } \ 0<\eta<1. \tag{7.3}$$

Note that, given (7.2), convergence (7.3) is equivalent to

$$\left(\int_0^T\!\!\int_0^T \frac{\|(u-u_\varepsilon)(t) - (u-u_\varepsilon)(s)\|^2_{L^2(\Omega)}}{|t-s|^{1+2\eta}}\mathrm{d}t\,\mathrm{d}s\right)^{1/2} \le c\varepsilon^{(1-\eta)/2} \qquad \text{for all } \ 0<\eta<1.$$

Analogous conclusions hold for more general initial data u_0. Define ϕ to be the Dirichlet integral

$$\phi(u) \doteq \frac{1}{2}\int_\Omega |\nabla u|^2, \qquad D(\phi) \doteq H^1_0(\Omega).$$

We readily characterize the corresponding *interpolation set* $D_{r,2}$ for $0 < r < 1$. Indeed, one has that [Bre75, Thm. 2] $u_0 \in D_{r,2}$ iff there exists $\varepsilon \mapsto v(\varepsilon) \in H^2(\Omega) \cap H^1_0(\Omega)$ such that $\varepsilon \mapsto \varepsilon^{1-r}\|\Delta v(\varepsilon)\|_{L^2(\Omega)} \in L^2_*(0,1)$, and $\varepsilon \mapsto \varepsilon^{-r}\|u_0 - v(\varepsilon)\|_{L^2(\Omega)} \in L^2_*(0,1)$. This precisely amounts to say that

$$u_0 \in \big(L^2(\Omega), H^2(\Omega)\cap H^1_0(\Omega)\big)_{r,2} \equiv \begin{cases} H^{2r}(\Omega) & \text{for } \ 0<r<1/4 \\ H^{1/2}_{00}(\Omega) & \text{for } \ r=1/4 \\ H^{2r}_0(\Omega) & \text{for } \ 1/4 \le r < 1 \end{cases} \tag{7.4}$$

where $H^s_0(\Omega)$, $1/2 < s < 2$ and $H^{1/2}_{00}(\Omega)$ classically denote the spaces of functions whose trivial extension to $\mathbb{R}^n$ belongs to $H^s(\mathbb{R}^n)$ and $H^{1/2}(\mathbb{R}^n)$, respectively [LM72] (note that $H^s_0(\Omega)$, $1/2 < s < 2$ is the closure in $H^s(\Omega)$ of the space of compactly supported smooth functions whereas $H^{1/2}_{00}(\Omega)$ is not).

Choose now u_0 fulfilling (7.4) for some $0 < r < 1$ and let $\varepsilon \mapsto u_{0\varepsilon} \in H^2(\Omega)\cap H^1_0(\Omega)$ be such that $\varepsilon \mapsto \varepsilon^{1-r}\|u_{0\varepsilon}\|_{H^2(\Omega)}$, $\varepsilon \mapsto \varepsilon^{-r}\|u_0 - u_{0\varepsilon}\|_{L^2(\Omega)} \in L^2_*(0,1)$. Then, the unique minimizers u_ε of the WED functionals over $K(u_{0\varepsilon})$ fulfill

$$\max_{t\in[0,T]} \|u(t) - u_\varepsilon(t)\|_{L^2(\Omega)} \le c\varepsilon^r$$

and quantitative convergence in $H^\eta(0,T;L^2(\Omega))$ holds as well. Obvious modifications lead to the more general *linear parabolic equation* $u_t - \mathrm{div}(A\nabla u) = f$ where the bounded function $A : \Omega \to \mathbb{R}^{n\times n}$ takes symmetric and uniformly positive definite values.

Let now Ω be $C^{1,1}$ or convex, $u_0 \in H^2(\Omega)\cap H^1_0(\Omega)$, and $u_{0\varepsilon}$ be suitable approximations in the same spirit above. Define

$$u \mapsto \begin{cases} \displaystyle\int_0^T\!\!\int_\Omega \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}u_t^2 + \frac{1}{2\varepsilon}|\Delta u|^2 - \frac{1}{\varepsilon}fu\right) & \text{for } \ u \in L^2(0,T;H^2(\Omega)\cap H^1_0(\Omega)) \\ \infty & \text{otherwise.} \end{cases}$$

The minimizers to the latter, constrained to fulfill $u_\varepsilon(\cdot,0) = u_{0\varepsilon}$ almost everywhere in Ω, fulfill (7.2)-(7.3) where u is the solution of the *biharmonic equation*

$$u_t + \Delta^2 u = f \quad \text{a.e. in } \ \Omega\times(0,T)$$

subject to the initial condition, homogeneous Dirichlet conditions on u and homogeneous Neumann conditions on Δu (again other boundary conditions may be considered).

We may recollect the above examples (as well as a variety of other symmetric parabolic problems of order $2k$) in the following abstract setting. Let the Hilbert spaces H and V be given with the injection $V \subset H$ being dense. Moreover, let the bilinear and symmetric form $a : V\times V \to \mathbb{R}$ be coercive and continuous and define

$$u \mapsto \begin{cases} \displaystyle\int_0^T \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}|u'|^2 + \frac{1}{2\varepsilon}a(u,u) - \frac{1}{\varepsilon}(f,u)\right) & \text{for } \ u \in L^2(0,T;V) \\ \infty & \text{otherwise.} \end{cases}$$

Then, the minimizers of the above WED functionals (suitably constrained to fulfill initial conditions) converge in H, uniformly with respect to time, to a solution of the *abstract* *linear equation*

$$u' + Au = f \quad \text{a.e. in } \ (0,T)$$

where the linear operator $A : H \to H$ is defined by $(Au, v) \doteq a(u,v)$ for all $v \in V$ and $u \in D(A) \doteq \{v \in V \ : \ \sup_{|z|=1} a(v,z) < \infty\}$. Indeed, in the same spirit of (7.2)-(7.3), much more is true as we have that

$$\max_{t\in[0,T]} \|u(t) - u_\varepsilon(t)\|_H \leq c\varepsilon^{1/2}, \tag{7.5}$$

$$\|u - u_\varepsilon\|_{H^\eta(0,T;H)} \leq c\varepsilon^{(1-\eta)/2} \qquad \text{for all } \ 0 < \eta < 1. \tag{7.6}$$

7.2. **Parabolic variational inequalities.** Under the above assumptions, let now $g \in H^1(\Omega)$ be given with $g \leq 0$ on $\partial\Omega$ and consider the WED functionals

$$u \mapsto \begin{cases} \displaystyle\int_0^T\!\!\int_\Omega \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}u_t^2 + \frac{1}{2\varepsilon}|\nabla u|^2 - \frac{1}{\varepsilon}fu\right) & \text{for } \ u \in L^2(0,T;H_0^1(\Omega)) \\ & \text{with } \ u(\cdot,t) \geq g(\cdot) \ \text{ a.e.} \\ \infty & \text{otherwise.} \end{cases}$$

Then, (suitably constrained) minimizers converge in $C([0,T];H)$ to a solution of the *parabolic obstacle problem*

$$\int_\Omega u_t(u-v) + \int_\Omega \nabla u \cdot \nabla(u-v) \leq \int_\Omega f(u-v) \quad \forall v \in K, \quad \text{a.e. in } \ (0,T)$$

where the convex set K is defined by $K \doteq \{v \in H_0^1(\Omega) \ : \ v \geq g \ \text{ a.e.}\}$. More precisely, the error estimates (7.2)-(7.3) hold. Within the abstract setting introduced in the previous subsection, a variety of other constraints can be discussed as well.

Next, let $W : \mathbb{R} \to \mathbb{R}$ be a λ-convex and smooth function. Then, (suitably constrained) minimizers of

$$u \mapsto \begin{cases} \displaystyle\int_0^T\!\!\int_\Omega \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}u_t^2 + \frac{1}{2\varepsilon}|\nabla u|^2 + \frac{1}{2\varepsilon}W(u) - \frac{1}{\varepsilon}fu\right) & \text{for } \ u \in L^2(0,T;H_0^1(\Omega)) \\ \infty & \text{otherwise} \end{cases}$$

converge in the sense of (7.2)-(7.3) to solutions of the *reaction-diffusion equation*

$$u_t - \Delta u + W'(u) = f \quad \text{a.e. in } \ \Omega \times (0,T).$$

The choice $W(u) = (u^2-1)^2$ corresponds to the so-called *Allen-Cahn equation.*

7.3. **Quasi-linear parabolic PDEs.** Let $F : \Omega \times \mathbb{R}^n \to [0,+\infty)$ be such that:

$$F(x,\cdot) \in C^1(\mathbb{R}^n) \quad \text{for a.e. } \ x \in \Omega, \tag{7.7}$$

$$F(x,\cdot) \quad \text{is convex and } \ F(x,0) = 0 \ \text{ for a.e. } \ x \in \Omega, \tag{7.8}$$

$$F(\cdot,\xi) \quad \text{is measurable for all } \ \xi \in \mathbb{R}^n. \tag{7.9}$$

Then, we can set $b \doteq \nabla_\xi F : \Omega \times \mathbb{R}^n \to \mathbb{R}^n$. We assume that, for a given $p > 1$, F satisfies the growth conditions

$$\exists c,\, C > 0 \quad \text{such that } \ F(x,\xi) \geq c|\xi|^p - C,$$

$$|b(x,\xi)| \leq C(1 + |\xi|^{p-1}) \quad \text{for a.e. } \ x \in \Omega \ \text{ and all } \ \xi \in \mathbb{R}^n. \tag{7.10}$$

Let us now consider the WED functionals

$$u \mapsto \begin{cases} \displaystyle\int_0^T\!\!\int_\Omega \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}u_t^2 + \frac{1}{\varepsilon}F(\cdot,\nabla u) - \frac{1}{\varepsilon}fu\right) & \text{for } \ u \in L^2(0,T;L^2(\Omega)\cap W_0^{1,p}(\Omega)) \\ \infty & \text{otherwise.} \end{cases}$$

In the latter, homogeneous Dirichlet conditions are considered, other choices being possible. The present analysis ensures that minimizers of the above functionals, suitably constrained as for initial values, converge in the sense of (7.2)-(7.3) to a solution of the *quasilinear equation*

$$u_t - \operatorname{div} b(\cdot,\nabla u) = f \quad \text{a.e. in } \ \Omega\times(0,T).$$

In particular, the choice $F(x,\xi) \doteq |\xi|^p/p$ gives rise to the so-called *p-Laplacian equation*, whereas the choice $F(x,\xi) \doteq (1+|\xi|^2)^{1/2}$ corresponds to the *mean curvature flow for Cartesian surfaces* (note however that the latter does not directly fit into this theory because of a lack of lower semicontinuity).

7.4. **Degenerate parabolic PDEs.** Assume we are given $\beta:\mathbb{R}\to\mathbb{R}$ monotone and continuous with $\beta(0)=0$ and superlinear growth at infinity [Bre73b]. Define j to be the only convex function such that $\beta = j'$ and $j(0)=0$. We now introduce the WED functionals on $H^1(0,T;H^{-1}(\Omega))$ given by

$$u \mapsto \begin{cases} \displaystyle\int_0^T \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}\|u_t\|^2_{H^{-1}(\Omega)} + \frac{1}{\varepsilon}\int_\Omega \big(j(u)-fu\big)\right) & \text{for } \ u\in L^2(0,T;L^2(\Omega)) \\ & \text{with } \ j(u)\in L^1(\Omega\times(0,T)) \\ \infty & \text{otherwise.} \end{cases}$$

Qualified minimizers of the latter functional converge in $H^{-1}(\Omega)$, uniformly in time, to a solution of the following *degenerate parabolic equation*

$$u_t - \Delta\beta(u) = f \quad \text{in } \ \Omega\times(0,T)$$

in a distributional sense, along with homogeneous Dirichlet boundary conditions for $\beta(u)$. More precisely, we have that (7.5)-(7.6) hold for $H = H^{-1}(\Omega)$. In particular, the choice $\beta(u) \doteq (u-1)^+ - u^-$ corresponds to the classical *two-phase Stefan problem*, $\beta(u)\doteq |u|^{m-2}u$ for $m>2$ leads to the *porous medium equation*. The multivalued case $\beta(u) = \partial I_{[0,1]}$ (subdifferential of the indicator function of the interval $[0,1]$), related to the *Hele-Shaw cell equation*, can be handled as well.

7.5. **Evolution of microstructure in a bistable bar.** In [CO08] Conti & Ortiz consider the WED functionals

$$F_\varepsilon(u) \doteq \begin{cases} \displaystyle\int_0^T\!\!\int_\Omega \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}u_t^2 - \frac{1}{\varepsilon}fu\right) & \text{if } \ |u_x| = 1 \ \text{ a.e.} \\ \infty & \text{else} \end{cases}$$

suggested by a modelization of branching in martensite in a one-dimensional bar occupying the reference domain $\Omega = (0,1)$. The function $u:\Omega\times(0,T)\to\mathbb{R}$ represents the bar displacement, the system is constrained in the two *phases* $u_x = 1$ and $u_x = -1$, no contribution from the interfacial energy is considered, and f stands for an applied body force (see [KM94]).

For fixed $\varepsilon > 0$, the functional F_ε fails to be lower semicontinuous with respect to the weak topology of $H^1(0,T;L^2(\Omega))$. The argument in [CO08, Thm. 3.1] entails that

$$\mathrm{sc}^- F_\varepsilon(u) = \int_0^T \int_\Omega \mathrm{e}^{-t/\varepsilon} \left(\frac{1}{2} u_t^2 + I_{[-1,1]}(u_x) - \frac{1}{\varepsilon} f u \right)$$

where the relaxation is taken with respect to the weak topology in $H^1(0,T;L^2(\Omega))$ (note that the actual proof in [CO08] is concerned with the weak topology in $H^1(\Omega \times (0,T))$ instead) and $I_{[-1,1]}$ is the indicator function of the interval $[-1,1]$. This computation is by no means trivial as the interplay between energy and dissipation has to be carefully taken into account. In this specific case, $\mathrm{sc}^- F_\varepsilon$ coincides with the convexification of F_ε. Note however that this is not the case in general, see [MO08, Sec. 5.1].

From this computation, CONTI & ORTIZ conjecture that the WED formalism can be of some use for describing microstructure evolution. In particular, *at a fixed level* $\varepsilon > 0$, the net effect of relaxation is that of allowing solutions u with $|u_x| < 1$ which may therefore be interpreted as the weak limit of a fine evolving microstructure.

The analysis in [CO08] left open the issue of considering $\varepsilon \to 0$, namely of extending the above interpretation to the causal limit. We are in the position of filling this gap. Fix an initial condition (say $u_0 = 0$ as in [CO08], for simplicity) and homogeneous Dirichlet boundary conditions. Note that $\mathrm{sc}^- F_\varepsilon$ is coercive with respect to the weak topology in $H^1(0,T;L^2(\Omega))$ on $K(0)$ (see again [CO08, Thm. 3.1]). Hence, at each level ε, the functional $\mathrm{sc}^- F_\varepsilon$ admits a unique minimizer u_ε in $K(0)$ (along with homogeneous Dirichlet boundary conditions) and, by applying our results, u_ε converges uniformly in $L^2(\Omega)$ and weakly in $H^1(0,T;L^2(\Omega))$ to a function u solving

$$u' + \partial I_C(u) \ni f \quad \text{a.e. in} \quad (0,T), \quad u(0) = 0. \tag{7.11}$$

Here I_C is the indicator function of the nonempty convex and closed set

$$C \doteq \{u \in H_0^1(\Omega) \ : \ |u_x| \leq 1 \ \text{ a.e.}\}.$$

In particular, the latter entails that $u : [0,T] \to H_0^1(\Omega)$ fulfills

$$\int_0^1 \left(\int_0^x (u_t - f) \right) (p - u_x) \leq 0 \quad \forall p \in L^2(\Omega) \ \text{ with } \ |p| \leq 1 \ \text{ a.e. in } \ \Omega \ \text{ and } \ \int_\Omega p = 0,$$
$$\text{a.e. in} \ (0,T), \ \ u(\cdot,0) = 0 \ \text{ a.e. in } \ \Omega.$$

Moreover, our convergence analysis may be extended to the case of approximate minimizers of the original unrelaxed functional F_ε. In particular, as $\mathrm{sc}^- F_\varepsilon$ turns out to be the WED functional corresponding to the convex and lower semicontinuous potential

$$u \mapsto \int_\Omega \left(I_{[-1,1]}(u_x) - \frac{1}{\varepsilon} f u \right),$$

we are in the position of applying Corollary 5.5 and deduce that all (qualified) sequences of approximate minimizers of F_ε converge to the unique solution of (7.11) in the sense of (7.2)-(7.3).
An illustration of this solution for a constant body force f is given in [CO08, Fig.4].

7.6. **Surface roughening by island growth.** A second example of relaxation in [CO08] concerns the WED functionals

$$F_\varepsilon(u) \doteq \begin{cases} \displaystyle\int_0^T\!\!\int_\Omega \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}u_t^2 - \frac{1}{\varepsilon}fu\right) & \text{if } \nabla u \in K \text{ a.e.} \\ \infty & \text{else} \end{cases}$$

where $\Omega \doteq [0,1]^2$ and

$$K \doteq \{(0,\pm 1),\ (\pm 1, 0)\}.$$

These functionals are considered in connection with the phenomenon of *island growth and coarsening* during the epitaxial growth of thin films. In particular, $u : \Omega \to \mathbb{R}$ represents the height of the thin film surface, f is a given deposition rate, and K is the set of preferred slopes (see [ORS99])

For fixed $\varepsilon > 0$ the relaxation of F_ε with respect to the weak topology of $H^1(0,T;L^2(\Omega))$ reads [CO08, Thm. 4.1]

$$\mathrm{sc}^- F_\varepsilon(u) \doteq \int_0^T\!\!\int_\Omega \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}u_t^2 + I_{\mathrm{co}\,K}(\nabla u) - \frac{1}{\varepsilon}fu\right)$$

where $I_{\mathrm{co}\,K}$ is the indicator function of the convex hull $\mathrm{co}\,K$ of K, namely

$$\mathrm{co}\,K = \{(x,y) \in \mathbb{R}^2 \ :\ |x| + |y| \leq 1\}.$$

By inspecting the specific form of $\mathrm{sc}^- F_\varepsilon$, in [CO08] the macroscopic behavior of the evolving thin film in the causal limit $\varepsilon \to 0$ is conjectured to corresponds to the gradient flow along with the choice

$$\phi(u) \doteq -\int_\Omega fu \quad \text{if } \nabla u \in \mathrm{co}\,K \quad \text{and} \quad \phi(u) \doteq \infty \quad \text{else},$$

the effect of the microstructure being that of relaxing the original constraint $\nabla u \in K$ to the weaker $\nabla u \in \mathrm{co}\,K$ (in particular, solutions with $\nabla u \in \mathrm{int}\,\mathrm{co}\,K$ are interpreted as weak limits of evolving microstructures).

This fact is confirmed by our convergence result. Indeed, the functionals $\mathrm{sc}^- F_\varepsilon$ are (convex and) lower semicontinuous. Hence, they admit unique minimizers u_ε in $K(0)$ (the initial condition 0 is chosen for simplicity and reference with [CO08]) and the sequence u_ε converges uniformly in $L^2(\Omega)$ to the unique gradient flow

$$u' + \partial I_M(u) \ni f \quad \text{a.e. in } (0,T), \quad u(0) = 0 \tag{7.12}$$

where $M \doteq \{v \in H^1_0(\Omega) \ :\ \nabla v \in \mathrm{co}\,K\}$.

Convergence also holds for approximate minimizers of the original unrelaxed functional F_ε. Indeed, as the relaxation $\mathrm{sc}^- F_\varepsilon$ is the WED functional related to the convex and lower semicontinuous potential

$$u \mapsto \int_\Omega \left(I_{\mathrm{co}\,K}(\nabla u) - \frac{1}{\varepsilon}fu\right),$$

Corollary 5.4 ensures that all (qualified) sequences of approximate minimizers of F_ε converge to the unique solution of (7.12) in the sense of (7.2)-(7.3). The reader is referred to [CO08, Fig. 5] for an illustration of a thin film evolution developing island growth under a constant deposition rate.

Acknowledgment. The authors wish to thank Sir John Ball for proposing the issues eventually developed in Subsection 5.8. A. Mielke was partially supported by the DFG Resarch Unit FOR 797 *Analysis and Computation of Microstructures in Finite Plasticity* under Mie 459/5-1. U. Stefanelli was partially supported by the FP7-IDEAS-ERC-StG Grant #200497 *BioSMA: Mathematics for Shape Memory Technologies in Biomechanics* and gratefully acknowledges the hospitality of WIAS.

References

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

[Bar76] V. Barbu. *Nonlinear semigroups and differential equations in Banach spaces.* Noordhoff International Publishing, Leyden, 1976.

[BE76a] H. Brezis and I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas dépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282(20):Ai, A1197–A1198, 1976.

[BE76b] H. Brezis and I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282(17):Aii, A971–A974, 1976.

[Bio55] M. A. Biot. Variational principles in irreversible thermodynamics with application to viscoelasticity. *Phys. Rev. (2)*, 97:1463–1469, 1955.

[BL76] J. Bergh and J. Löfström. *Interpolation spaces. An introduction.* Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[Bre71] H. Brezis. Monotonicity methods in Hilbert spaces and some application to nonlinear partial differential equations. In *Contrib. to nonlin. functional analysis. Proc. Sympos. Univ. Wisconsin, Madison*, pages 101–156. Academic Press, New York, 1971.

[Bré73a] D. Brézis. Classes d'interpolation associées à un opérateur monotone. *C. R. Acad. Sci. Paris Sér. A-B*, 276:A1553–A1556, 1973.

[Bre73b] H. Brezis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* Number 5 in North Holland Math. Studies. North-Holland, Amsterdam, 1973.

[Bre75] H. Brezis. Interpolation classes for monotone operators. In *Partial differential equations and related topics (Program, Tulane Univ., New Orleans, La., 1974)*, pages 65–74. Lecture Notes in Math., Vol. 446. Springer, Berlin, 1975.

[BS94] C. Baiocchi and G. Savaré. Singular perturbation and interpolation. *Math. Models Methods Appl. Sci.*, 4(4):557–570, 1994.

[Cla90] F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

[CO08] S. Conti and M. Ortiz. Minimum principles for the trajectories of systems governed by rate problems. *J. Mech. Phys. Solids*, 56:1885–1904, 2008.

[CP69] M. G. Crandall and A. Pazy. Semi-groups of nonlinear contractions and dissipative sets. *J. Funct. Anal.*, 3:376–418, 1969.

[Gho08] N. Ghoussoub. *Selfdual partial differential systems and their variational principles.* Springer Monographs in Mathematics, Springer New York, 2009.

[Gio96] E. De Giorgi. Conjectures concerning some evolution problems. *Duke Math. J.*, 81(2):255–268, 1996. A celebration of John F. Nash, Jr.

[Gur63] M. E. Gurtin. Variational principles in the linear theory of viscoelasticity. *Arch. Rational Mech. Anal.*, 13:179–191, 1963.

[Gur64a] M. E. Gurtin. Variational principles for linear elastodynamics. *Arch. Rational Mech. Anal.*, 16:34–50, 1964.

[Gur64b] M. E. Gurtin. Variational principles for linear initial value problems. *Quart. Appl. Math.*, 22:252–256, 1964.

[Hla69] I. Hlaváček. Variational principles for parabolic equations. *Apl. Mat.*, 14:278–297, 1969.

[Ilm94] T. Ilmanen. Elliptic regularization and partial regularity for motion by mean curvature. *Mem. Amer. Math. Soc.*, 108(520):x+90, 1994.

[KM94] R. V. Kohn and S. Müller. Surface energy and microstructure in coherent phase transitions. *Comm. Pure Appl. Math.*, 47(4):405–435, 1994.

[Kōm67] Y. Kōmura. Nonlinear semi-groups in Hilbert space. *J. Math. Soc. Japan*, 19:493–507, 1967.

[LM72] J.-L. Lions and E. Magenes. *Non-homogeneus boundary value problems and applications*, volume 1. Springer-Verlag, New York - Heidelberg, 1972.

[MO08] A. Mielke and M. Ortiz. A class of minimum principles for characterizing the trajectories of dissipative systems. *ESAIM Control Optim. Calc. Var.*, 14:494–516, 2008.

[MS08] A. Mielke and U. Stefanelli. A discrete variational principle for rate-independent evolution. *Adv. Calc. Var.*, 1(4):399–431, 2008.

[MST89] 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(2):281–330, 1989.

[Nay76a] B. Nayroles. Deux théorèmes de minimum pour certains systèmes dissipatifs. *C. R. Acad. Sci. Paris Sér. A-B*, 282(17):Aiv, A1035–A1038, 1976.

[Nay76b] B. Nayroles. Un théorème de minimum pour certains systèmes dissipatifs. Variante hilbertienne. *Travaux Sém. Anal. Convexe*, 6(Exp. 2):22, 1976.

[NSV00] R. Nochetto, G. Savaré, and C. Verdi. A posteriori error estimates for variable time-step discretization of nonlinear evolution equations. *Comm. Pure Appl. Math.*, 53(5):525–589, 2000.

[ORS99] M. Ortiz, E. A. Repetto, and H. Si. A continuum model of kinetic roughening and coarsening in thin films. *J. Mech. Phys. Solids*, 47(4):697–730, 1999.

[Ott01] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

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

[RSS08] R. Rossi, A. Segatti, and U. Stefanelli. Attractors for gradient flows of non convex functionals and applications. *Arch. Ration. Anal. Mech.*, 187(1):91–135, 2008.

[Sav96] G. Savaré. Weak solutions and maximal regularity for abstract evolution inequalities. *Adv. Math. Sci. Appl.*, (6), 1996.

[Sim87] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.

[Ste08a] U. Stefanelli. The Brezis-Ekeland principle for doubly nonlinear equations. *SIAM J. Control Optim.*, (8):1615–1642, 2008.

[Ste08b] U. Stefanelli. A variational principle for hardening elasto-plasticity. *SIAM J. Math. Anal.*, 40(2):623–652, 2008.

[Ste08c] U. Stefanelli. The discrete Brezis-Ekeland principle. *J. Convex Anal.*, to appear, 2008. Preprint available at: `http://www.imati.cnr.it/ulisse/pubbl.html`.

[Tar70] L. Tartar. Théorème d'interpolation non linéaire et applications. *C. R. Acad. Sci. Paris Sér. A-B*, 270:A1729–A1731, 1970.

[Tar72] L. Tartar. Interpolation non linéaire et régularité. *J. Functional Analysis*, 9:469–489, 1972.

[Tri95] H. Triebel. *Interpolation theory, function spaces, differential operators.* Johann Ambrosius Barth, Heidelberg, second edition, 1995.

[Vis01] A. Visintin. A new approach to evolution. *C. R. Acad. Sci. Paris Sér. I Math.*, 332(3):233–238, 2001.

[Vis08] A. Visintin. An extension of the Brezis-Ekeland-Nayroles principle to monotone operators. *Adv. Math. Sci. Appl.*, 18:633–650, 2008.