universität wien

DISSERTATION | DOCTORAL THESIS

Titel | Title

Variational models and methods in material science and evolution equations

verfasst von | submitted by

Dott. mag. Riccardo Voso

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 796 605 405
Dissertationsgebiet lt. Studienblatt | Field of study as it appears on the student record sheet:	Mathematik
Betreut von | Supervisor:	Univ.-Prof. Ulisse Stefanelli PhD

Abstract

Variational principles are an important tool to analyze and describe natural phenomena. The *Weighted Energy-Dissipation* (WED) variational principle offers interesting mathematical perspectives, particularly when dealing with partial differential equations (PDEs). A possible approach to proving the existence of solutions for PDEs involves minimizing a functional whose Euler-Lagrange equation corresponds to the given PDE. If such a functional can be found, the problem of existence for the PDE is translated into seeking stationary points for the identified functional. In some cases, finding stationary points of the functional is often simpler than directly demonstrating the existence of solutions for the PDE. Despite its effectiveness, defining an appropriate functional is not always straightforward or trivial. The WED variational principle presents an alternative approach compared to classical variational principles. In this method, the goal is to find a functional whose Euler-Lagrange equation is a regularized version of the considered PDE. This regularization introduces greater flexibility in applications, allowing both conservative and dissipative systems to be described in a variational fashion. Nevertheless, this advantage of applicability comes at the expense of proving the so-called *causal limit*, namely the removal of the regularization.

In this thesis, we tackle three problems by means of the WED approach. We first introduce a novel approach to studying semilinear gradient flows with state-dependent dissipation. This extension has deep impact on the technical level. Most notably, the structure of the Euler-Lagrange equation features specific state-dependent dissipative terms showing minimal integrability in time. Then, we move to a general optimal control problem in the setting of gradient flows. We propose two approximations of the problem, both relying on the variational formulation of the gradient-flow dynamics via the WED approach. Eventually, we present a global-in-time variational approach to the dynamics of hyperelastic materials via the WED variational theory.

Contents

Chapter 1

Introduction

1.1 Rationale

Variational principles constitute a fundamental tool for the description of natural phenomena. Typically, variational principles are connected with conservative systems and therefore they are not suited for the analysis of dissipative systems. The most well-known variational principle is *Hamilton's principle* or *principle of least action*. This is based on finding stationary points of the *action* S defined as the time-integral of the *Lagrangian* L of the system, namely,

$$S = \int_{t_i}^{t_f} L(u, \dot{u}, t)\,\mathrm{d}t,$$

where t_i and t_f indicate the initial and the final time of the system's evolution, respectively, under the constraint that the trajectory u takes given initial and final values. Here, u and $\dot{u}$ are usually referred to as generalized coordinates and velocity, respectively. The Lagrangian L uniquely defines the state of a physical system. Indeed, according to Hamilton's principle, trajectories u which make S stationary are solutions of the Euler-Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{u}} - \frac{\partial L}{\partial u} = 0.$$

L is usually defined as the difference between the kinetic and the potential energy of the system. For example, the general form of the Lagrangian for a particle with mass m moving in an external potential $V(u)$ is

$$L = \frac{1}{2}m\dot{u}^2 - V(u),$$

and the corresponding Euler-Lagrange equation is *Newton's equation* $m\ddot{u} = -\partial V/\partial u$.

The application of Hamilton's principle is restricted to those cases where it is possible to define a potential energy, i.e., when only conservative forces are acting on the system. Nevertheless, there exist other variational principles that deal with dissipative settings. In particular, we are interested in the *Weighted Energy-Dissipation* (WED) variational principle. In the context of partial differential equations (PDEs), establishing well-posedness involves defining an action whose corresponding Euler-Lagrange equation coincides with the given PDE. Consequently, finding a stationary point for the associated functional (action) becomes equivalent to demonstrating the existence of a solution for the PDE. The WED method adopts a different approach, seeking a functional (action) that yields a regularized version of the target PDE as its Euler-Lagrange equation. This chosen functional incorporates a positive real parameter, allowing the removal of the regularization by letting the parameter approach zero. In this limit, the original PDE is recovered. The WED approach has been proved to be applicable to a number of classical PDE settings. The purpose of this thesis is to test the WED approach in three less standard settings. In particular, this thesis is dedicated to the study of a field theory with state-dependent dissipation, optimal control theory, and elastodynamics.

1.2 Weighted Energy-Dissipation approach

The Weighted Energy-Dissipation approach is a global-in-time variational method, hinging on the minimization of the ε-dependent ($\varepsilon > 0$) functionals W^ε. The functionals W^ε feature an ε-dependent exponentially decaying *weight*, usually given by the (weighted) sum of the potential *energy* and the *dissipation*. As such, they are usually referred to as *Weighted Energy-Dissipation* functionals. This approach contributes a variational approximation of the limiting differential problem. As such, it paves the way to the application of the general methods of the calculus of variations. The existence of the approximations/minimizers is ascertained by the Direct Method. This can be especially beneficial in actually computing such approximations, for one can profit from the vast toolbox of optimization methods. In addition, combinations with *other* approximations (additional parameters, time and space discretization, data approximations) can be treated at the $\varepsilon > 0$ level within the general frame of Γ-convergence theories.

There are numerous noteworthy examples of this approach. Here, we introduce and describe only a few that we believe are particularly relevant to this thesis. By examining these chosen cases, we can achieve two objectives: first, to illustrate how this method typically works, and second, to understand the specific role and significance of this thesis within the wider landscape of this approach.

- *Gradient flows* (Mielke – Stefanelli [45])

This paper addresses gradient flows of the type

$$u_t + \partial\phi(u) \ni f, \quad \text{in } H, \text{ a.e. in } (0,T), \tag{1.2.1}$$

where H is an Hilbert space and ∂ denotes the subdifferential from H to H. Here, f is assumed to be a function in $L^2(0,T;H)$ and $\phi : (-\infty,\infty]$ is a proper, lower semicontinuous, bounded from below, and λ-convex function, namely $u \mapsto \phi(u) - (\lambda/2)\|u\|_H^2$ is convex for some $\lambda \in \mathbb{R}$. A family of ε-dependent functionals $W^\varepsilon : H^1(0,T;H) \to (-\infty,\infty]$ is introduced as

$$W^\varepsilon(u) = \int_0^T e^{-t/\varepsilon}\left(\frac{\varepsilon}{2}\|u_t\|_H^2 + \phi(u) - (f,u)_H\right) \mathrm{d}t,$$

over the space of trajectories u attaining initial value u^0. Then, the existence of minimizers u^ε for W^ε is ascertained by means of the Direct Method. Then, by computing the Euler-Lagrange equation, i.e.,

$$-\varepsilon u_{tt}^\varepsilon + u_t^\varepsilon + \partial\phi(u^\varepsilon) \ni f, \quad \text{in } H, \text{ a.e. in } (0,T),$$

an elliptic-in-time regularization of the gradient flow (1.2.1) is obtained. Eventually, a uniform convergence $u^\varepsilon \to u$ in H, with u solution of (1.2.1) is proved. This contributes a new existence proof for the classical problem (1.2.1) with initial condition $u(0) = u_0$.

- *Generalized gradient flows* (Akagi – Melchionna – Stefanelli [3])

A general doubly nonlinear flow on the time half line $t > 0$ is considered, namely,

$$\mathrm{d}_V\psi(u_t) + \partial\phi(u) \ni 0, \quad \text{in } V^*, \text{ a.e. } t > 0, \tag{1.2.2}$$

where $\psi, \phi : V \to (-\infty, \infty]$ are convex functionals acting on a Banach space V with dual V^*. Here, $\mathrm{d}_V : V \to V^*$ indicates the Gâteaux derivative and $\partial : V \to 2^{V^*}$ stands for the subdifferential in the sense of convex analysis. The dissipative (pesudo-)potential ψ is assumed to have p-growth, for any $p > 1$, e.g., $\psi(v) = \|v\|_V^p/p$ and ϕ is such that its sublevel sets are compact in an Banach space X compactly and densely embedded into V. The corresponding family of WED functionals $W^\varepsilon : L^p(\mathbb{R}_+, e^{-t/\varepsilon}\mathrm{d}t; V) \to (-\infty, \infty]$ is defined as

$$W^\varepsilon(u) = \int_0^\infty e^{-t/\varepsilon}\Big(\varepsilon\psi(u_t) + \phi(u)\Big)\,\mathrm{d}t.$$

After proving the existence of minimizer via the Direct Method, the Euler-Lagrange equation is computed as

$$-\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{d}_V\psi(u_t^\varepsilon) + \mathrm{d}_V\psi(u^\varepsilon) + \partial\phi(u^\varepsilon) \ni 0, \quad \text{in } X^*, \text{ a.e. } t > 0.$$

Finally, the limit $\varepsilon \to 0$ is considered and the proof of the convergence of u^ε (up to subsequences) to the solution u of (1.2.2) is given.

- *Semilinear wave equations* (Stefanelli [56])

In this paper, the *De Giorgi's conjecture on elliptic regularization* is proved on a finite time interval $(0, T)$. For any $p > 1$ and for any $\varepsilon > 0$ the WED functional $W^\varepsilon : H^2(0, T; L^2(\mathbb{R}^d)) \cap L^2(0, T; H^1(\mathbb{R}^d)) \cap L^p((0, T) \times \mathbb{R}^d) \to [0, \infty]$ is defined by

$$W^\varepsilon(u) = \int_0^T \int_{\mathbb{R}^d} e^{-t/\varepsilon}\Big(\frac{\varepsilon^2}{2}|u_{tt}|^2 + \frac{1}{2}|\nabla u|^2 + |u|^p\Big)\,\mathrm{d}x\,\mathrm{d}t.$$

The existence of minimizers u^ε and the convergence $u^\varepsilon \to u$ (up to subsequences) is shown where u is the solution of the semilinear wave equation, namely,

$$u_{tt} - \Delta u + p|u|^{p-2}u = 0,$$

is eventually given. A proof of the De Giorgi conjecture on the semiline $(0, \infty)$, that is, in its original formulation, has been given by Serra & Tilli in [53].

1.2.1 Review of the literature

In this subsection, we provide a concise overview on the Weighted Energy-Dissipation variational principle.

In the gradient-flow case of ψ quadratic, the use of the WED approach has to be traced back at least to the study of Brakke mean-curvature flow of varifolds in [31]. Existence of periodic solutions to gradient flows by the WED approach has also been obtained [32]. The general theory for the case of a λ-convex energy ϕ is in [45]. Note that the linear case is also mentioned in the classical PDE textbook by Evans [27, Problem 3, p. 487]. Nonconvex problems are discussed in [7], see also [40] for nonpotential perturbations, [35] for a stability result via Γ-convergence, and [39] for an application to the study of symmetries of solutions. Existence of variational solutions to the equation $u_t - \nabla \cdot f(x, u, \nabla u) + \partial_u f(x, u, \nabla u) = 0$ where the field f is convex in $(u, \nabla u)$ has been tackled in [13]. Applications to stochastic PDE [50, 51] and to curves of maximal slope in metric spaces [48, 49] are also available. We also mention the applications to micro-structure evolution [21], to mean-curvature evolution of Cartesian surfaces [55], to the incompressible Navier–Stokes system [10, 47], and to the non-newtonian Navier-Stokes system [34].

The WED approach has been also considered in the nonquadratic case, namely, for dissipations $v \mapsto \psi(v)$ of q growth, for $1 \leq q \neq 2$. The rate-independent case $q = 1$ [43] is tackled in [42], see also [44] for a discrete counterpart. Applications to dynamic fracture [23, 33] and dynamic plasticity are available [24]. The doubly nonlinear setting $1 < q \neq 2$ has been studied in [3, 4, 5, 6] under different assumptions for the energy ϕ, see also [2] for the case of nonpotential perturbations.

The WED approach has been applied to hyperbolic and mixed parabolic-hyperbolic problems, as well. Inspired by a conjecture by De Giorgi [25], the case of semilinear waves ($\psi = 0$) is discussed in [53, 56]. Extensions to other hyperbolic problems [36, 52], also including forcings [38, 57, 58, 59] or dissipative terms ($\psi \neq 0$) [1, 37, 52], have already been obtained.

1.3 Plan of the thesis

In this section, we provide a schematic outline of this thesis.

Chapter 2 introduces the Weighted Energy-Dissipation approach applied to semilinear gradient flows with state-dependent dissipation. This chapter is based on a submitted manuscript coauthored with Goro Akagi and Ulisse Stefanelli [8]. A concise introduction can be found in Subsection 1.3.1.

Chapter 3 explores an optimal control problem for gradient flows using the Weighted Energy-Dissipation method. This is another submitted work in collaboration with Takeshi Fukao and Ulisse Stefanelli [29]. An introduction is provided in Subsection 1.3.2.

Chapter 4 records my ongoing research with Martin Kružík and Ulisse Stefanelli on a global variational approach to the dynamics of elastic bodies. An introduction is in Subsection 1.3.3.

1.3.1 Weighted Energy-Dissipation approach to semilinear gradient flows with state-dependent dissipation

In this subsection, we introduce the problem tackled in Chapter 2.

This work is motivated by the study a gradient flow-type dynamic in the presence of state-dependent dissipation, namely a dynamic of the type

$$\mathrm{d}_2\psi(u, u_t) + \partial\phi(u) \ni 0 \quad \text{in } H, \text{ a.e. in } (0, T), \tag{1.3.1}$$

where $u : [0, T] \to H$ is a trajectory from the finite time interval $[0, T]$ to the Hilbert space H. Here, d_2 denotes the Fréchet differential with respect to the second variable. We assume that the smooth functional $\psi : H \times H \to [0, \infty)$ is such that $v \mapsto \psi(u, v)$ is convex and differentiable for any given $u \in H$ and that the monotone map $v \mapsto \mathrm{d}_2\psi(u, v)$ is nondegenerate and linearly bounded, uniformly w.r.t. u. The potential ϕ is assumed to be given by the sum $\phi(u) = \langle Au, u\rangle_X + \phi^2(u)$ where the linear, continuous, and self-adjoint operator $A : X \to X^*$ maps a second Hilbert space X, compact and dense in H, to its dual. Here, $\phi^2 : H \to [0, \infty]$ is a convex functional. Moreover, ∂ denotes the subdifferential in the sense of convex analysis.

In the concrete case (see Section 2.5), we define Ω as an open $C^{1,1}$ bounded domain in $\mathbb{R}^d$ for $d \in \mathbb{N}$ and we let $H = L^2(\Omega)$ and $X = H^1(\Omega)$. Then, we consider dissipative (pseudo-)potentials of the form

$$\psi(u, u_t) = \int_\Omega g(k * u)\frac{|u_t|^2}{2}\,\mathrm{d}x,$$

with $g \in C^{2,1}_{\mathrm{loc}}(\mathbb{R})$ uniformly positive. Here, the kernel k belongs to $L^2(\mathbb{R}^d)$ and $k * u$ stands for the convolution of k with the trivial extension $\bar{u}$ of u out of Ω. The potential is taken as

$$\phi(u) = \int_\Omega \left(\frac{1}{2}|\nabla u|^2 + \hat{\beta}(u) \right) \mathrm{d}x,$$

where $\hat{\beta} : \mathbb{R} \to [0, \infty]$ is a proper, convex, and lower semicontinuous function such that $\hat{\beta}(0) = 0$.

In Chapter 2, we provide a variational formulation of an elliptic-in-time regularization of the gradient flow (1.3.1) via the Weighted Energy-Dissipation variational principle. For every $\varepsilon > 0$, we introduce the WED functionals $W^\varepsilon : L^2(0, T; H) \to [0, \infty]$ as

$$W^\varepsilon(u) = \int_0^T e^{-t/\varepsilon} (\varepsilon\psi(u(t), u_t(t)) + \phi(u(t))) \, \mathrm{d}t,$$

over the trajectories u attaining initial value u^0. Our main result (Theorem 2.2.5) proves that for all $\varepsilon > 0$, there exists a global minimizer u^ε of W^ε and that, up to subsequences, u^ε converge to a solution of (1.3.1) as $\varepsilon \to 0$. By computing the Euler–Lagrange problem corresponding to the minimization of W^ε, namely,

$$-\varepsilon \frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{d}_2\psi(u, u_t)) + \varepsilon \mathrm{d}_1\psi(u, u_t) + \mathrm{d}_2\psi(u, u_t) + \partial\phi(u) \ni 0 \quad \text{in } H, \text{ a.e. in } (0, T),$$

is revealed that the minimization of W^ε corresponds to the elliptic-in-time regularization of the gradient flow (1.3.1). Indeed, by formally taking the limit $\varepsilon \to 0$ in the latter, one would recover the gradient flow (1.3.1).

1.3.2 Optimal control of gradient flows via the Weighted Energy-Dissipation method

In this subsection, we provide the synopsis of Chapter 3, where we introduce two approximation techniques for solving a general optimal control problem.

We analyze optimal control problems of the form

$$\min\{J(y, u) \ : \ u \in U \ , y = S(u)\}, \tag{1.3.2}$$

where J is a target functional of the control $u : (0, T) \to H$ and of the solution $y = S(u)$ to the following gradient flow

$$\dot{y} + \partial\phi(y) \ni u \quad \text{in } H, \text{ a.e. in } (0, T). \tag{1.3.3}$$

Here, H is a Hilbert space, the dot indicates a time derivative, and ∂ indicates the subdifferential from H to H in the sense of convex analysis. The set of admissible controls U is assumed to be a compact subset of $L^2(0, T; H)$ and J is asked to be lower semicontinuous with respect to the weak $\times$ strong topology in $H^1(0, T; H) \times L^2(0, T; H)$. The functional $\phi : H \to (-\infty, \infty]$ is assumed to be proper, bounded from below, lower semicontinuous, and κ-convex for some $\kappa \in \mathbb{R}$, i.e., $y \mapsto \phi(y) - (\kappa/2)\|y\|^2$ is convex.

We prove the existence of a solution for Problem (1.3.2) and investigate two approximations based on the WED functionals $W_\varepsilon : H^1(0, T; H) \times U \to (-\infty, \infty]$ defined by

$$W_\varepsilon(y, u) = \int_0^T e^{-t/\varepsilon} \left(\frac{\varepsilon}{2}\|\dot{y}\|^2_H + \phi(y) - (u, y) \right) \mathrm{d}t,$$

over trajectories y attaining initial value y^0. At first, we approximate the constraint $y = S(u)$ by considering the following optimal control problem

$$\min\{J(y,u) \ : \ u \in U \ , \ y = \arg\min W_\varepsilon(\cdot, u)\}. \tag{1.3.4}$$

This essentially amounts to replacing the gradient-flow constraint (1.3.3) by its elliptic regularization

$$\begin{aligned} &- \varepsilon\ddot{y} + \dot{y} + \partial\phi(y) \ni u \quad \text{in } H, \ \text{a.e. in } (0,T), \\ &y(0) = y^0, \quad \varepsilon\dot{y}(T) = 0, \end{aligned}$$

namely, the Euler-Lagrange problem corresponding to $W_\varepsilon(\cdot, u)$. Secondly, we relax even more, by turning to an unconstrained optimal control problem, namely,

$$\min\left\{J(y,u) + \frac{1}{\lambda}(W_\varepsilon(y,u) - m_\varepsilon(u))\right\}, \tag{1.3.5}$$

where $\lambda > 0$ and $m_\varepsilon(u)$ is the minimum value of $W_\varepsilon(\cdot, u)$, given u. In particular, after establishing existence for Problems (1.3.2), (1.3.4), and (1.3.5), we consider limits as ε and λ go to 0. As a result of a Γ-convergence argument, we prove that

$$P_{\varepsilon\lambda} \xrightarrow{\Gamma} P_\varepsilon \quad \text{and} \quad P_\varepsilon \xrightarrow{\Gamma} P$$

in the sense of Γ-convergence in $H^1(0,T;H) \times L^2(0,T;H)$ endowed with the weak $\times$ strong topology, where we have defined the functionals P, P_ε, and $P_{\varepsilon\lambda}$ as

$$\begin{aligned} P(y,u) &= \begin{cases} J(y,u) & \text{if } y = S(u), \\ \infty & \text{else,} \end{cases} \\ P_\varepsilon(y,u) &= \begin{cases} J(y,u) & \text{if } y \in \arg\min W_\varepsilon(\cdot,u), \\ \infty & \text{else.} \end{cases} \\ P_{\varepsilon\lambda}(y,u) &= J(y,u) + \frac{1}{\lambda}(W_\varepsilon(y,u) - m_\varepsilon(u)). \end{aligned}$$

By specifying a joint speed of convergence, we may also pass to the simultaneous limit $(\varepsilon, \lambda) \to (0,0)$. In particular, we prove that

$$\text{for } \lambda = \lambda(\varepsilon) \text{ s.t. } \limsup_{\varepsilon\to 0} \lambda(\varepsilon)\varepsilon^{-3}e^{T/\varepsilon} = 0 \text{ one has that } P_{\varepsilon\lambda} \xrightarrow{\Gamma} P.$$

1.3.3 A global variational approach to the dynamics of elastic bodies

This subsection aims at giving an introduction to Chapter 4.

In Chapter 4, we focus on the dynamics of hyperelastic materials of the form

$$\rho y_{tt} - \nabla \cdot \mathrm{D}W(\nabla y) + \delta\nabla^4 : \nabla^4 y = 0 \quad \text{in } (0,\infty) \times \Omega, \tag{1.3.6}$$

where the reference configuration Ω is a bounded domain in $\mathbb{R}^3$ of class C^4, $y : \Omega \to \mathbb{R}^3$ is the deformation, $\rho > 0$ is the referential material density, and δ is a length-scale parameter. Here, D denotes the Gâteaux derivative, W is the stored elastic energy density, and the fourth-order differential operator ∇^4 is given by $\nabla\otimes\nabla\otimes\nabla\otimes\nabla$. The nonlinear map $W : \mathbb{R}^{3\times 3} \to [0,\infty]$ is taken as the sum of a convex $C^{1,1}_{\text{loc}}$ functional and a term inversely proportional to the determinant of its argument, so that $W(F) \to \infty$ as $\det F \to 0^+$. The existence of solutions of the equation (1.3.6), subject to the initial conditions $y(0,x) = y^0$, $y_t(0,x) = y^1$ is proven via the Weighted Energy-Dissipation method. The WED functionals $G_\varepsilon : L^2(\mathbb{R}_+ \times \Omega) \to [0,\infty]$ are defined via

$$G_\varepsilon(y) = \int_0^\infty \int_\Omega e^{-t/\varepsilon} \left(\frac{\varepsilon^2 \rho}{2} |y_{tt}|^2 + W(\nabla y) + \frac{\delta}{2} |\nabla^4 y|^2 \right) \mathrm{d}x\, \mathrm{d}t,$$

over the locally orientation-preserving trajectories attaining the initial conditions. Our main result proves that

(i) for all $\varepsilon > 0$, there exists a global minimizer y^ε of G_ε and

(ii) by taking the limit $\varepsilon \to 0$, such y^ε converge up to subsequences to a solution of (1.3.6).

Chapter 2

Weighted Energy-Dissipation approach to semilinear gradient flows with state-dependent dissipation

This chapter consists of my submitted manuscript with Goro Akagi and Ulisse Stefanelli [8].

Abstract

We investigate the Weighted Energy-Dissipation variational approach to semilinear gradient flows with state-dependent dissipation. A family of parameter-dependent functionals defined over entire trajectories is introduced and proved to admit global minimizers. These global minimizers correspond to solutions of elliptic-in-time regularizations of the limiting causal problem. By passing to the limit in the parameter we prove that such global minimizers converge, up to subsequences, to a solution of the gradient flow.

2.1 Introduction

We are interested in a global-in-time variational approach to abstract gradient flows of the form

$$\mathrm{d}_2\psi(u,\dot{u}) + Au(t) + \partial\phi^2(u) \ni 0 \quad \text{in } H, \text{ a.e. in } (0,T). \tag{2.1.1}$$

Here, $u : [0,T] \to H$ is a trajectory from the finite time interval $[0,T]$ to the Hilbert space H and the dot denotes time differentiation. In relation (2.1.1), the smooth functional $(u,v) \in H \times H \mapsto \psi(u,v) \in [0,\infty)$ is assumed to be convex in the *rate* v, the symbol d_2 indicates the Fréchet differential with respect to the second variable, and the monotone map $v \mapsto \mathrm{d}_2\psi(u,v)$ is assumed to be nondegenerate and linearly bounded, uniformly for u. The linear, continuous, and self-adjoint operator $A : X \to X^*$ maps a second Hilbert space X, compact and dense in H, to its dual X^*. The functional $\phi^2 : H \to [0,\infty]$ is convex and $\partial\phi^2$ indicates its subdifferential. The doubly nonlinear inclusion (2.1.1) hence corresponds to the gradient flow of the *energy* $\phi(u) = \langle Au, u\rangle_X/2 + \phi^2(u)$, where $\langle\cdot,\cdot\rangle_X$ denotes the duality pairing between X^* and X, with respect to the *state-dependent dissipation* ψ. Note in particular that the nonlinearity $A + \partial\phi^2$ is tailored to the variational formulation of a semilinear elliptic operator and that we implicitly ask $Au \in H$. Relation (2.1.1) hence corresponds to a *strong* formulation of a semilinear parabolic problem with state-dependent dynamics, see Section 2.5.

This paper is concerned with the analysis of the *Weighted Energy-Dissipation* (WED) approach to the gradient flow (2.1.1). This is a global-in-time variational approach, hinging on the minimization of the ε-dependent ($\varepsilon > 0$) functionals $W^\varepsilon : L^2(0,T;H) \to [0,\infty]$ given by

$$W^\varepsilon(u) = \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon}\Big(\varepsilon\psi(u(t),\dot u(t)) + \phi(u(t))\Big)\,\mathrm{d}t & \text{if } u \in K(u_0), \\ \infty & \text{else,} \end{cases} \tag{2.1.2}$$

where the convex domain $K(u_0)$ depends on the initial condition $u_0 \in H$ and is defined as

$$K(u_0) = \{u \in L^2(0,T;H) \ : \ u(0) = u_0 \text{ and } \phi(u),\ \psi(u,\dot u) \in L^1(0,T)\}. \tag{2.1.3}$$

The functionals W^ε feature an ε-dependent exponentially decaying *weight* multiplying the (weighted) sum of the *energy* ϕ and the *dissipation* ψ. As such, they are usually referred to as Weighted Energy-Dissipation functionals. Note that the parameter $\varepsilon > 0$ plays the role of a relaxation time, which justifies the factor ε in front of the dissipation ψ.

By means of the functionals W^ε one can approximate the differential problem (2.1.1) on purely variational terms. The focus of this paper is on providing a rigorous analysis of such an approximation. Our main result (Theorem 2.2.5) proves that

(i) for all $\varepsilon > 0$, there exists a global minimizer u^ε of W^ε,

(ii) u^ε solves the Euler–Lagrange equation strongly,

(iii) up to subsequences, u^ε converge to a solution of (2.1.1) as $\varepsilon \to 0$.

In fact, the minimization of W^ε corresponds to an elliptic-in-time regularization of the gradient flow (2.1.1). This is revealed by computing the Euler–Lagrange problem corresponding to the minimization of W^ε, namely,

$$-\varepsilon\frac{\mathrm{d}}{dt}(\mathrm{d}_2\psi(u,\dot u)) + \varepsilon\mathrm{d}_1\psi(u,\dot u) + \mathrm{d}_2\psi(u,\dot u) + Au + \partial\phi^2(u) \ni 0 \quad \text{in } H, \tag{2.1.4}$$

$$u(0) = u_0, \quad \varepsilon\mathrm{d}_2\psi(u(T),\dot u(T)) = 0, \tag{2.1.5}$$

to be solved almost everywhere in $(0,T)$. Here, d_1 denotes the Fréchet differential with respect to the first variable and a natural Neumann condition is imposed at the final time T. Due to ellipticity in time, the global minimizers u^ε are more regular than the solution to the limiting problem (2.1.1). On the other hand, causality is lost at level $\varepsilon > 0$ and is restored just in the limit $\varepsilon \to 0$. Motivated by this fact, the limit $\varepsilon \to 0$ is usually referred to as *causal limit*.

The WED approach contributes a variational approximation of the limiting differential problem. As such, it paves the way to the application of the general methods of the calculus of variations. The existence of the approximations u^ε is ascertained by the Direct Method. This can be especially beneficial in actually computing such approximations, for one can profit from the vast toolbox of optimization methods. In addition, combinations with *other* approximations (additional parameters, time and space discretization, data approximations) can be treated at the $\varepsilon > 0$ level within the general frame of Γ-convergence [22].

The results of this paper can be compared with the ones of [45]. There, the dissipation is assumed to be independent of the state and quadratic, namely, $\psi(v) = \|v\|_H^2/2$, corresponding indeed to the classical gradient-flow case. The novelty here is in considering a *state-dependent* dissipation instead, which has the potential of considerably broaden the range of application of the theory. This extension has deep impact on the technical level, asking for the development of a quite different analytical approach. Most notably, the structure of the Euler–Lagrange equation features specific state-dependent dissipative terms showing minimal integrability in time. Note nonetheless that the simpler dissipation setting of [45] eventually allows more general choices for the energy ϕ.

Before moving on, let us provide a brief review of the the vast literature on the WED approach. We however stress that all available results to date exclusively deal with the case $v \mapsto \psi(v)$, namely, with a dissipation which is independent of the state.

In the gradient-flow case of ψ quadratic, the use of the WED approach has to be traced back at least to the study of Brakke mean-curvature flow of varifolds in [31]. Existence of periodic solutions to gradient flows by the WED approach has also been obtained [32]. The general theory for the case of a λ-convex energy ϕ is in [45]. Note that the linear case is also mentioned in the classical PDE textbook by Evans [27, Problem 3, p. 487]. Nonconvex problems are discussed in [7], see also [40] for nonpotential perturbations, [35] for a stability result via Γ-convergence, and [39] for an application to the study of symmetries of solutions. Existence of variational solutions to the equation $u_t - \nabla \cdot f(x,u,\nabla u) + \partial_u f(x,u,\nabla u) = 0$ where the field f is convex in $(u, \nabla u)$ has been tackled in [13]. Applications to stochastic PDE [50, 51] and to curves of maximal slope in metric spaces [48, 49] are also available. We also mention the applications to micro-structure evolution [21], to mean-curvature evolution of Cartesian surfaces [55], and to the incompressible Navier–Stokes system [10, 47].

The WED approach has been also considered in the nonquadratic case, namely, for dissipations $v \mapsto \psi(v)$ of q growth, for $1 \leq q \neq 2$. The rate-independent case $q = 1$ [43] is tackled in [42], see also [44] for a discrete counterpart. Applications to dynamic fracture [23, 33] and dynamic plasticity are available [24]. The doubly nonlinear setting $1 < q \neq 2$ has been studied in [3, 4, 5, 6] under different assumptions for the energy ϕ, see also [2] for the case of nonpotential perturbations.

The WED approach has been applied to hyperbolic and mixed parabolic-hyperbolic problems, as well. Inspired by a conjecture by De Giorgi [25], the case of semilinear waves ($\psi = 0$) is discussed in [53, 56]. Extensions to other hyperbolic problems [36, 52], also including forcings [57, 58, 59] or dissipative terms ($\psi \neq 0$) [1, 37, 52], have already been obtained.

The paper is organized as follows. In Section 2.2, we introduce notation, present some preliminary material and assumptions, and state our main result, namely, Theorem 2.2.5. Sections 2.3 and 2.4 are devoted to the proof of the Theorem 2.2.5. In particular, in Section 2.3 we prove that the WED functional W^ε admits global minimizers and that each global minimizer is a strong solution of the Euler–Lagrange problem (2.1.4)–(2.1.5). Then, in Section 2.4 we consider the causal limit $\varepsilon \to 0$ and we prove that the solutions of the Euler–Lagrange problem (2.1.4)–(2.1.5) converge to the solutions of the gradient flow (2.1.1). Eventually, in Section 2.5 we present a PDE example illustrating the abstract theory.

2.2 Main result

In this section, we introduce the general functional setting (Subsection 2.2.1), record some preliminary material (Subsection 2.2.2), present our assumptions (Subsection 2.2.3), and state our main result (Subsection 2.2.4).

2.2.1 Notation

We assume X, H to be two real and separable Hilbert spaces with $X \subset H$ densely and compactly, so that, by indicating by X^* the dual of X, (X, H, X^*) is a classical Hilbert triplet. We indicate by $(\cdot, \cdot)$ the scalar product in H. The symbols $\|\cdot\|_E$ and $\langle \cdot, \cdot \rangle_E$ will be used to indicate the norm in the general Banach space E and the duality pairing between the dual E^* and E, respectively.

Given any convex, proper, and lower semicontinuous functional $f : H \to (-\infty, \infty]$, we indicate by $D(f) := \{u \in H \ : \ f(u) < \infty\}$ its *effective domain* and by $\partial f(u)$ its *subdifferential* at a point $u \in D(f)$ as the (possibly empty) set

$$\partial f(u) := \{w \in H \ : \ (w, v - u) \leq f(v) - f(u) \ \ \forall v \in H\}.$$

The map $\partial f : H \to 2^H$ (parts of H) is a maximal monotone operator with domain $D(\partial f) := \{u \in D(f) \ : \ \partial f(u) \neq \emptyset\}$. The *Legendre–Fenchel conjugate* f^* of f is classically defined by $f^*(w) := \sup_{u \in H}((w, u) - f(u))$ for $w \in H$ and is convex, proper, and lower semicontinuous. Moreover, one has that $f^*(w) + f(u) \geq (w, u)$ for all u, $w \in H$, as well as

$$w \in \partial f(u) \ \Leftrightarrow \ u \in \partial f^*(w) \ \Leftrightarrow \ f^*(w) + f(u) = (w, u). \tag{2.2.1}$$

For all convex, proper, and lower semicontinuous functionals $g : X \to (-\infty, \infty]$, we define the subdifferential $\partial_X g : X \to 2^{X^*}$ at a point $u \in D(g)$ as the (possibly empty) set

$$\partial_X g(u) := \{w \in X^* \ : \ \langle w, v - u \rangle_X \leq g(v) - g(u) \ \forall v \in X\}.$$

Given a sequence $(f_\lambda)_\lambda$ and f with f_λ, $f : H \to [0, \infty)$ convex, we say that f_λ *converge to* f *in the sense of Mosco* [9] if

$$\begin{aligned} \forall u_\lambda \rightharpoonup u \ \text{ in } H \quad &\Rightarrow \quad f(u) \leq \liminf_{\lambda \to 0} f_\lambda(u_\lambda), \\ \forall v \in H \ \exists v_\lambda \to v : \quad &\limsup_{\lambda \to 0} f_\lambda(v_\lambda) = f(v). \end{aligned}$$

In this context, $(v_\lambda)_\lambda$ is said to be a *recovery sequence* for v.

We indicate by $\mathrm{d}f : H \to H$ the Fréchet differential of a Fréchet-differentiable functional $f : H \to \mathbb{R}$. In the case of $\psi : H \times H \to \mathbb{R}$, the symbols d_1 and d_2 denote the partial Fréchet differentials in the first and in the second variable, respectively.

2.2.2 Preliminaries

We record here two preliminary lemmas, for later use.

Lemma 2.2.1 (Aubin–Lions–Simon, [54, Thm. 3])**.** *The embedding*

$$L^1(0, T; X) \cap H^1(0, T; H) \hookrightarrow C([0, T]; H)$$

is compact.

Lemma 2.2.2 (Integration by parts)**.** *Let $u \in H^1(0, T; H) \cap L^2(0, T; X)$ and $\xi \in L^2(0, T; H)$ be such that $\dot{\xi} \in (H^1(0, T; H))^* + L^2(0, T; X^*)$. Let t_1, t_2 be Lebesgue points of $t \mapsto (\xi(t), u(t))$. Then, it holds that*

$$\langle \dot{\xi}, u \rangle_{H^1(t_1, t_2; H) \cap L^2(t_1, t_2; X)} = (\xi(t_2), u(t_2)) - (\xi(t_1), u(t_1)) - \int_{t_1}^{t_2} (\xi, \dot{u}).$$

The latter is an extension of [5, Prop. 2.3]. The proof follows by adapting the argument there, which was originally developed in $L^r(t_1,t_2;X^*)$, and using the density of $L^r(t_1,t_2;X^*)$ into $(H^1(t_1,t_2;X))^*$.

For the reader's convenience, we record here the generalized Dominated Convergence Theorem from [60, Rem. (19a), p. 1015].

Lemma 2.2.3. *Let $f_n, g_n \in L^1(\Omega)$ with $|f_n| \leq g_n$ a.e., $f_n \to f$ pointwise a.e., $g_n \to g$ in $L^1(\Omega)$. Then $f_n \to f$ in $L^1(\Omega)$.*

Proof. As $g_n \to g$ a.e., by passing to the limit in $|f_n| \leq g_n$ one has that $|f| \leq g$ a.e., so that $f \in L^1(\Omega)$. Since $0 \leq g + g_n - |f - f_n|$ and $\liminf_n (g + g_n - |f - f_n|) = 2g$ a.e. one can apply the Fatou's Lemma and get

$$\int_\Omega 2g \leq \liminf_n \int_\Omega (g + g_n - |f - f_n|) = \int_\Omega 2g - \limsup_n \int_\Omega |f - f_n|.$$

We hence have that

$$\limsup_n \int_\Omega |f - f_n| = 0. \quad \square$$

2.2.3 Assumptions

In addition to the general functional setting specified in Subsection 2.2.1, we introduce here our assumptions on data. We ask for the following.

(A1) $\phi = \phi^1 + \phi^2 : H \to \mathbb{R}_+$, with $\phi^1(u) = \langle Au, u\rangle_X/2$ if $u \in X$ and $\phi^1(u) = \infty$ if $u \in H \setminus X$ for some $A : X \to X^*$ linear, continuous, and self-adjoint. There exist c_1, $c_2 > 0$ such that

$$\|u\|_X^2 \leq c_1(1 + \phi^1(u)) \quad \forall u \in X, \tag{2.2.2}$$

$$\|Au\|_{X^*} \leq c_2\|u\|_X \qquad \forall u \in X. \tag{2.2.3}$$

The functional $\phi^2 : H \to \mathbb{R}_+$ is a proper, convex, and lower semicontinuous functional and there exists $c_3 > 0$ such that

$$\|\eta^2\|_H^r \leq c_3(1 + \phi^2(u)) \ \text{ for every } u \in D(\partial\phi^2) \text{ and } \eta^2 \in \partial\phi^2(u), \tag{2.2.4}$$

for some $r \geq 2$. Moreover, $u_0 \in D(\phi)$.

(A2) $\psi \in C^2(H \times H; \mathbb{R}_+)$, $\mathbb{R}_+ := [0, \infty)$, is such that $\psi(u, \cdot)$ is convex and $\psi(u, 0) = 0$ for all $u \in H$. Moreover, there exists a positive constant c_4 such that

$$\|v\|_H^2 \leq c_4\psi(u, v) \quad \forall u, v \in H. \tag{2.2.5}$$

Furthermore, for all $R > 0$, there exist positive constants $c_{5,R}$, $c_{6,R}$, $c_{7,R}$, $c_{8,R}$, $c_{9,R}$, and $c_{10,R}$ such that the following conditions hold:

$$\|\mathrm{d}_1\psi(u, v)\|_H \leq c_{5,R}(1 + \|v\|_H^2) \quad \forall u, v \in H \text{ such that } \|u\|_H \leq R, \tag{2.2.6}$$

$$\|\mathrm{d}_1\psi(u_1, v) - \mathrm{d}_1\psi(u_2, v)\|_H \leq c_{6,R}\|u_1 - u_2\|_H\|v\|_H^2$$
$$\forall u_1, u_2, v \in H \text{ such that } \|u_1\|_H,\ \|u_2\|_H \leq R, \tag{2.2.7}$$

$$\|\mathrm{d}_2\psi(u, v)\|_H^2 \leq c_{7,R}(1 + \|v\|_H^2) \quad \forall u, v \in H \text{ such that } \|u\|_H \leq R, \tag{2.2.8}$$

$$\|\mathrm{d}_2\psi(u_1, v) - \mathrm{d}_2\psi(u_2, v)\|_H \leq c_{8,R}\|u_1 - u_2\|_H\|v\|_H$$
$$\forall u_1, u_2, v \in H \text{ such that } \|u_1\|_H,\ \|u_2\|_H \leq R, \tag{2.2.9}$$

$$\|\mathrm{d}_{21}\psi(u, v)w\|_H \leq c_{9,R}\|v\|_H\|w\|_H \quad \forall u, v, w \in H, \text{ such that } \|u\|_H \leq R, \tag{2.2.10}$$

$$(\mathrm{d}_{22}\psi(u, v)w, w) \geq c_{10,R}\|w\|_H^2 \quad \forall u, v, w \in H, \text{ such that } \|u\|_H \leq R. \tag{2.2.11}$$

Here, we have used the notation $\mathrm{d}_{21} = \mathrm{d}_2\mathrm{d}_1$ and $\mathrm{d}_{22} = \mathrm{d}_2\mathrm{d}_2$.

Before moving on, note that (2.2.5) and (2.2.8) imply that

$$\frac{1}{c_4}\|v\|_H^2 \leq \psi(u,v) \leq (\mathrm{d}_2\psi(u,v),v) \leq c_{11,R}(1+\|v\|_H^2)$$
$$\forall u,v \in H \text{ such that } \|u\|_H \leq R, \tag{2.2.12}$$

for $c_{11,R} = (c_{7,R}+1)/2$.

By virtue of (2.2.9) and (2.2.11), for each $R > 0$, we can find positive constants $c_{12,R}, c_{13,R} > 0$ such that

$$(\mathrm{d}_2\psi(u_1,v_1) - \mathrm{d}_2\psi(u_2,v_2), v_1 - v_2) \geq c_{12,R}\|v_1-v_2\|_H^2 - c_{13,R}\|u_1-u_2\|_H^2\|v_2\|_H^2 \tag{2.2.13}$$

for $u_1, u_2, v_1, v_2 \in H$ satisfying $\|u_1\|, \|u_2\| \leq R$. Indeed, setting $v_\theta := (1-\theta)v_2 + \theta v_1$ for $\theta \in [0,1]$, we observe that

$$\begin{aligned}
&(\mathrm{d}_2\psi(u_1,v_1) - \mathrm{d}_2\psi(u_2,v_2), v_1 - v_2)\\
&\quad\geq (\mathrm{d}_2\psi(u_1,v_1) - \mathrm{d}_2\psi(u_1,v_2), v_1-v_2) + (\mathrm{d}_2\psi(u_1,v_2) - \mathrm{d}_2\psi(u_2,v_2), v_1-v_2)\\
&\quad\geq \int_0^1 \left(\frac{\mathrm{d}}{\mathrm{d}\theta}\mathrm{d}_2\psi(u_1,v_\theta), v_1-v_2\right)\mathrm{d}\theta - \|\mathrm{d}_2\psi(u_1,v_2) - \mathrm{d}_2\psi(u_2,v_2)\|_H\|v_1-v_2\|_H\\
&\quad= \int_0^1 (\mathrm{d}_{22}\psi(u_1,v_\theta)(v_1-v_2), v_1-v_2)\,\mathrm{d}\theta - \|\mathrm{d}_2\psi(u_1,v_2) - \mathrm{d}_2\psi(u_2,v_2)\|_H\|v_1-v_2\|_H\\
&\quad\geq c_{10,R}\|v_1-v_2\|_H^2 - c_{8,R}\|u_1-u_2\|_H\|v_2\|_H\|v_1-v_2\|_H\\
&\quad\geq \frac{c_{10,R}}{2}\|v_1-v_2\|_H^2 - \frac{c_{8,R}^2}{2c_{10,R}}\|u_1-u_2\|_H^2\|v_2\|_H^2
\end{aligned}$$

so that (2.2.13) follows by choosing $c_{12,R} = c_{10,R}/2$ and $c_{13,R} = c_{8,R}^2/(2c_{10,R})$.

We now prove a Mosco convergence result, which in particular applies to the setting of Assumption (A2).

Lemma 2.2.4 (Mosco convergence)**.** *Let $\psi \in C^1(H\times H;\mathbb{R}_+)$ fulfill (2.2.6). Then, for all $u_\lambda \to u$ in H one has that $\psi(u_\lambda,\cdot) \to \psi(u,\cdot)$ in the sense of Mosco in H.*

Proof. Fix $u_1, u_2, v \in H$ and let $u_\theta := (1-\theta)u_1 + \theta u_2$ for $\theta \in (0,1)$. Owing to the fact that $\psi \in C^1(H\times H,\mathbb{R}_+)$, by using (2.2.6) one computes

$$\begin{aligned}
|\psi(u_1,v) - \psi(u_2,v)| &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta}\psi(u_\theta,v)\,\mathrm{d}\theta = \int_0^1 (\mathrm{d}_1\psi(u_\theta,v), u_1-u_2)\,\mathrm{d}\theta\\
&\leq \|u_1-u_2\|_H\int_0^1 \|\mathrm{d}_1\psi(u_\theta,v)\|_H\,\mathrm{d}\theta \leq c_{5,R}(1+\|v\|_H^2)\|u_1-u_2\|_H.
\end{aligned}$$

This entails the pointwise convergence $\psi(u_\lambda,v) \to \psi(u,v)$, so that, for all v one can choose the constant recovery sequence $v_\lambda = v$.

On the other hand, let $v_\lambda \rightharpoonup v$ in H. Then it follows that

$$\begin{aligned}
\psi(u_\lambda,v_\lambda) &\geq \psi(u,v_\lambda) - |\psi(u,v_\lambda) - \psi(u_\lambda,v_\lambda)|\\
&\geq \psi(u,v_\lambda) - c_{5,R}(1+\|v_\lambda\|_H^2)\|u-u_\lambda\|_H,
\end{aligned}$$

which implies that $\liminf_{\lambda\to 0}\psi(u_\lambda,v_\lambda) \geq \psi(u,v)$. $\square$

The argument of Lemma 2.2.4 can be easily generalized to integral functionals on $L^2(0,T;H)$, also in presence of an arbitrary weight. More precisely, for all nonnegative weights $\mu \in L^\infty(0,T)$, one also has that

$$\int_0^T \mu\psi(u_k,\cdot) \to \int_0^T \mu\psi(u,\cdot) \quad \text{in the sense of Mosco on } L^2(0,T;H) \tag{2.2.14}$$

whenever $u_k \to u$ in $C([0,T];H)$. Note that in the following we resort to the latter convergence by choosing $\mu(t) = t$.

Given Assumption (A1), we readily check that $\partial_X \phi^1|_X \equiv A$ and $\partial\phi^1(u) = Au \in H$ for all $u \in D(\partial\phi) = \{v \in X \ : \ Av \in H\}$. Moreover, on account of assumptions (2.2.2)–(2.2.3) and (2.2.12) we can conclude that $K(u_0)$ from (2.1.3) actually corresponds to

$$K(u_0) = \{u \in H^1(0,T;H) \cap L^2(0,T;X) \ : \ u(0) = u_0 \text{ and } \phi^2(u) \in L^1(0,T)\}.$$

2.2.4 Main result

Before stating our main result, i.e., Theorem 2.2.5, let us specify the Cauchy problem for the gradient flow (2.1.1) as

$$\xi + \eta^1 + \eta^2 = 0 \quad \text{in } H, \text{ a.e. in } (0,T), \tag{2.2.15}$$
$$\xi = \mathrm{d}_2\psi(u,\dot u) \quad \text{in } H, \text{ a.e. in } (0,T), \tag{2.2.16}$$
$$\eta^1 = Au \quad \text{in } H, \text{ a.e. in } (0,T), \tag{2.2.17}$$
$$\eta^2 \in \partial\phi^2(u) \quad \text{in } H, \text{ a.e. in } (0,T), \tag{2.2.18}$$
$$u(0) = u_0. \tag{2.2.19}$$

The *Euler–Lagrange* problem corresponding to the minimization of W^ε given by (2.1.2) on $K(u_0)$ reads

$$-\varepsilon\dot\xi_\varepsilon + \xi_\varepsilon(t) + \varepsilon\gamma_\varepsilon + \eta^1_\varepsilon + \eta^2_\varepsilon = 0 \text{ in } X^*, \text{ a.e. in } (0,T), \tag{2.2.20}$$
$$\xi_\varepsilon = \mathrm{d}_2\psi(u_\varepsilon,\dot u_\varepsilon) \quad \text{in } H, \text{ a.e. in } (0,T), \tag{2.2.21}$$
$$\gamma_\varepsilon = \mathrm{d}_1\psi(u_\varepsilon,\dot u_\varepsilon) \quad \text{in } H, \text{ a.e. in } (0,T), \tag{2.2.22}$$
$$\eta^1_\varepsilon = Au_\varepsilon \text{ in } X^*, \text{ a.e. in } (0,T), \tag{2.2.23}$$
$$\eta^2_\varepsilon \in \partial\phi^2(u_\varepsilon) \text{ in } H, \text{ a.e. in } (0,T), \tag{2.2.24}$$
$$u_\varepsilon(0) = u_0, \tag{2.2.25}$$
$$\varepsilon\xi_\varepsilon(T) = 0. \tag{2.2.26}$$

Note the occurrence of the natural Neumann condition $\varepsilon\xi_\varepsilon(T) = 0$, which turns out to be instrumental for proving *a priori* estimates and degenerates for $\varepsilon \to 0$.

Our main result reads as follows.

Theorem 2.2.5 (WED Approach). *Under Assumptions* (A1)–(A2) *we have the following*:

(i) *The* WED *functional* W^ε *admits a global minimizer* $u_\varepsilon \in K(u_0)$;

(ii) *Given a global minimizer* u_ε *of* W^ε, *the quintuple* $(u_\varepsilon, \xi_\varepsilon, \gamma_\varepsilon, \eta^1_\varepsilon, \eta^2_\varepsilon)$ *with* $\xi_\varepsilon = \mathrm{d}_2\psi(u_\varepsilon,\dot u_\varepsilon)$, $\gamma_\varepsilon = \mathrm{d}_1\psi(u_\varepsilon,\dot u_\varepsilon)$, $\eta^1_\varepsilon = Au_\varepsilon$, *and* $\eta^2_\varepsilon = \varepsilon\dot\xi_\varepsilon - \xi_\varepsilon - \varepsilon\gamma_\varepsilon - \eta^1_\varepsilon \in \partial\phi^2(u_\varepsilon)$ *belongs to*

$$[H^1(0,T;H) \cap L^2(0,T;X)] \times L^2(0,T;H) \times L^1(0,T;H)$$
$$\times L^2(0,T;X^*) \times L^r(0,T;H),$$

and is a strong solution of the Euler–Lagrange problem (2.2.20)–(2.2.26);

(iii) *For any sequence* $\varepsilon_k \to 0$*, there exists a (not relabeled) subsequence such that* $(u_{\varepsilon_k}, \xi_{\varepsilon_k}, \eta^1_{\varepsilon_k}, \eta^2_{\varepsilon_k}) \to (u, \xi, \eta^1, \eta^2)$ *weakly* in*

$$[H^1(0,T;H) \cap L^2(0,T;X)] \times L^2(0,T;H) \times L^2(0,T;X^*) \times L^r(0,T;H),$$

(u, ξ, η^1, η^2) *is a strong solution of the gradient flow* (2.2.15)–(2.2.19) *with* $\eta_1 = Au \in L^2(0,T;H)$.

The remainder of the paper is devoted to the proof of Theorem 2.2.5. More precisely, parts (i) and (ii) are proved in Section 2.3 by means of a regularization procedure, and part (iii) is proved in Section 2.4.

A *caveat* on notation: in the following we use the same symbol c to indicate a positive constant, possibly depending on $u_0, c_i, c_{i,R}, \ell$, but independent of ε and of the regularization parameter λ (introduced below). Note that the constant c may change, even within the same line.

2.3 Existence of solutions to the Euler–Lagrange problem

This section is devoted to the proof of Theorem 2.2.5.(i)-(ii). We show that W^ε admits global minimizers u_ε and that such a global minimizer is a strong solution of the Euler–Lagrange problem (2.2.20)–(2.2.26). To this aim, in Subsection 2.3.1 we define a regularized functional $W^{\varepsilon\lambda}$ for $\lambda > 0$ by replacing ϕ^1 and ϕ^2 by their *Moreau–Yosida regularizations*, and check that $W^{\varepsilon\lambda}$ admits global minimizers $u_{\varepsilon\lambda}$. These $u_{\varepsilon\lambda}$ are proved to solve a regularized Euler–Lagrange problem in Subsection 2.3.2. In Subsection 2.3.3, we discuss the integrability of $u_{\varepsilon\lambda}$. The chain rule for the potential ψ is addressed in Subsection 2.3.4. This allows to derive in Subsection 2.3.5 some *a priori* estimates on the approximation $u_{\varepsilon\lambda}$ independently of λ. In Subsection 2.3.6, we take the limit $\lambda \to 0$ (up to subsequences) and prove that $u_\varepsilon = \lim_{\lambda\to 0} u_{\varepsilon\lambda}$ solves the Euler–Lagrange problem (2.2.20)–(2.2.26). Then, in Subsection 2.3.7, we check that such u_ε minimizes W^ε, which proves Theorem 2.2.5.(i). Eventually, we show that all global minimizers of $W^{\varepsilon\lambda}$ solve the Euler–Lagrange problem (2.2.20)–(2.2.26), so that Theorem 2.2.5.(ii) holds, as well.

2.3.1 The regularized functional

For every $\lambda > 0$, let us introduce the regularized WED functional $W^{\varepsilon\lambda} : L^2(0,T;H) \to [0,\infty]$ as

$$W^{\varepsilon\lambda}(u) = \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon}\Big(\varepsilon\psi(J_\lambda u(t), \dot u(t)) + \phi^1_\lambda(u(t)) + \phi^2_\lambda(u(t))\Big)\,\mathrm{d}t & \text{if } u \in K_0(u_0), \\ \infty & \text{else,} \end{cases}$$

where

$$K_0(u_0) = \{u \in W^{1,2}(0,T;H) \ : \ u(0) = u_0\}.$$

Here, $\phi^1_\lambda \in C^{1,1}(H;\mathbb{R})$ is the *Moreau–Yosida regularization* of ϕ^1 in H, namely

$$\phi^1_\lambda(u) = \inf_{v\in H}\left(\frac{1}{2\lambda}\|u-v\|^2_H + \phi^1(v)\right) = \frac{1}{2\lambda}\|u - J_\lambda u\|^2_H + \phi^1(J_\lambda u). \tag{2.3.1}$$

The map $J_\lambda : H \to H$ is the *resolvent* of $\partial\phi^1$ at level λ, namely the linear solution operator $J_\lambda : u \in H \mapsto J_\lambda u$ of the problem

$$J_\lambda u - u + \lambda\partial\phi^1(J_\lambda u) = 0 \ \text{ for every } u \in H.$$

We will also use the standard notation $A_\lambda : H \to H$ for the Yosida approximation $A_\lambda u = (u - J_\lambda u)/\lambda = \mathrm{d}\phi^1_\lambda(u)$ of A. Analogously, we let $\phi^2_\lambda \in C^{1,1}(H;\mathbb{R})$ be the Moreau–Yosida regularization of ϕ^2 in H and I_λ be the corresponding resolvent of $\partial\phi^2$ at level λ.

Lemma 2.3.1 (Global Minimizers of the regularized functional)**.** *The regularized* WED *functional $W^{\varepsilon\lambda}$ admits global minimizers in $K_0(u_0)$.*

Proof. The assertion follows by the Direct Method. At first, note that the constant-in-time function $u(t) \equiv u_0$ belongs to $K_0(u_0)$, which is then not empty. Let $(u_n)_n$ be an *infimizing sequence*, namely, $u_n \in K_0(u)$ with $W^{\varepsilon\lambda}(u_n) \to \inf_{K_0(u_0)} W^{\varepsilon\lambda}$. As $W^{\varepsilon\lambda}$ is coercive on $H^1(0,T;H)$ by assumption (2.2.5), we get

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

Hence, up to not a relabeled subsequence, one has

$$u_n \rightharpoonup u \ \text{ in } H^1(0,T;H). \tag{2.3.2}$$

As ϕ^1_λ and ϕ^2_λ are weakly lower semicontinuous on H we get

$$\int_0^T e^{-t/\varepsilon}\phi^j_\lambda(u) \le \liminf_{n\to\infty}\int_0^T e^{-t/\varepsilon}\phi^j_\lambda(u_n) \quad \text{for } j = 1,2.$$

Note that $u_0 = u_n(0) \rightharpoonup u(0)$ in H, so that $u(0) = u_0$. This in particular implies that $u \in K_0(u_0)$. As J_λ is a contraction in H, we obtain

$$\left\|\frac{\mathrm{d}}{\mathrm{d}t}(J_\lambda u_n)\right\|_{L^2(0,T;H)} \le \|\dot{u}_n\|_{L^2(0,T;H)} \le c$$

and from the coercivity (2.2.2) of ϕ^1 we get

$$\|J_\lambda u_n\|_{L^2(0,T;X)} \le c.$$

Then, by applying the Aubin–Lions–Simon Lemma 2.2.1, again up to a not relabeled subsequence, we obtain the strong convergence

$$J_\lambda u_n \to J_\lambda u \ \text{ in } C([0,T];H). \tag{2.3.3}$$

Moreover, since convergences (2.3.2)–(2.3.3) hold, relying on (2.2.14) we have

$$\int_0^T e^{-t/\varepsilon}\psi(J_\lambda u, \dot{u}) \le \liminf_{n\to\infty}\int_0^T e^{-t/\varepsilon}\psi(J_\lambda u_n, \dot{u}_n).$$

This in turn guarantees that

$$W^{\varepsilon\lambda}(u) \le \liminf_{n\to\infty} W^{\varepsilon\lambda}(u_n) = \inf_{K_0(u_0)} W^{\varepsilon\lambda},$$

showing that u is a global minimizer of $W^{\varepsilon\lambda}$. ☐

2.3.2 Regularized Euler–Lagrange problem

We now compute the first variation of the functional $W^{\varepsilon\lambda}$ at a minimum point $u_{\varepsilon\lambda} \in K_0(u_0)$. Fix $v \in V := \{v \in H^1(0,T;H) \ : \ v(0) = 0\}$, so that $u_{\varepsilon\lambda} + sv \in K_0(u_0)$ for all $s \in \mathbb{R}$, and compute the derivative of the real differentiable function $s \mapsto W^{\varepsilon\lambda}(u_{\varepsilon\lambda} + sv)$ at $s = 0$. As $u_{\varepsilon\lambda}$ is a global minimizer of $W^{\varepsilon\lambda}$, one gets

$$
\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}s}\Big[W^{\varepsilon\lambda}(u_{\varepsilon\lambda} + sv)\Big]_{s=0} \\
&= \int_0^T e^{-t/\varepsilon}\frac{\mathrm{d}}{\mathrm{d}s}\Big[\varepsilon\psi(J_\lambda(u_{\varepsilon\lambda}+sv), \dot{u}_{\varepsilon\lambda}+s\dot{v}) + \phi^1_\lambda(u_{\varepsilon\lambda}+sv) + \phi^2_\lambda(u_{\varepsilon\lambda}+sv)\Big]_{s=0}.
\end{aligned}
$$

Since J_λ is linear we readily obtain

$$
0 = \int_0^T e^{-t/\varepsilon}\Big[\varepsilon(\xi_{\varepsilon\lambda}, \dot{v}) + \varepsilon(\gamma_{\varepsilon\lambda}, J_\lambda v) + (\eta^1_{\varepsilon\lambda}, v) + (\eta^2_{\varepsilon\lambda}, v)\Big], \tag{2.3.4}
$$

where we have used the notation $\xi_{\varepsilon\lambda} = \mathrm{d}_2\psi(J_\lambda u_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda})$, $\gamma_{\varepsilon\lambda} = \mathrm{d}_1\psi(J_\lambda u_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda})$, $\eta^1_{\varepsilon\lambda} = A_\lambda u_{\varepsilon\lambda}$, and $\eta^2_{\varepsilon\lambda} = \mathrm{d}\phi^2_\lambda(u_{\varepsilon\lambda})$. Integrating by parts the first term on the right-hand side of equation (2.3.4) and recalling that v vanishes at $t = 0$, we get

$$
0 = \int_0^T e^{-t/\varepsilon}\Big[-\varepsilon(\dot{\xi}_{\varepsilon\lambda}, v) + (\xi_{\varepsilon\lambda}, v) + \varepsilon(J^*_\lambda\gamma_{\varepsilon\lambda}, v) + (\eta^1_{\varepsilon\lambda}, v) + (\eta^2_{\varepsilon\lambda}, v)\Big] + \varepsilon e^{-T/\varepsilon}(\xi_{\varepsilon\lambda}(T), v(T)).
$$

Here, we have again used the fact that J_λ is linear and continuous and we have indicated by J^*_λ its adjoint. Since the last integral equation holds for every $v \in V$, by further imposing $v(T) = 0$ we conclude that $u_{\varepsilon\lambda}$ solves the regularized Euler–Lagrange problem

$$
\begin{aligned}
&-\varepsilon\dot{\xi}_{\varepsilon\lambda}(t) + \xi_{\varepsilon\lambda}(t) + \varepsilon J^*_\lambda\gamma_{\varepsilon\lambda}(t) + \eta^1_{\varepsilon\lambda}(t) + \eta^2_{\varepsilon\lambda}(t) = 0 \quad \text{in } H, \text{ a.e. in } (0,T), &(2.3.5)\\
&\xi_{\varepsilon\lambda}(t) = \mathrm{d}_2\psi(J_\lambda u_{\varepsilon\lambda}(t), \dot{u}_{\varepsilon\lambda}(t)) \quad \text{in } H, \text{ a.e. in } (0,T), &(2.3.6)\\
&\gamma_{\varepsilon\lambda}(t) = \mathrm{d}_1\psi(J_\lambda u_{\varepsilon\lambda}(t), \dot{u}_{\varepsilon\lambda}(t)) \quad \text{in } \quad H, \text{ a.e. in } (0,T), &(2.3.7)\\
&\eta^1_{\varepsilon\lambda}(t) = A_\lambda u_{\varepsilon\lambda}(t) \quad \text{in } H \text{ a.e. in } (0,T), &(2.3.8)\\
&\eta^2_{\varepsilon\lambda}(t) = \mathrm{d}\phi^2_\lambda(u_{\varepsilon\lambda}(t)) \quad \text{in } H \text{ a.e. in } (0,T), &(2.3.9)\\
&u_{\varepsilon\lambda}(0) = u_0. &(2.3.10)
\end{aligned}
$$

Eventually, by considering $v \in V$ with $v(T) \neq 0$, we additionally deduce the terminal condition

$$
\varepsilon\xi_{\varepsilon\lambda}(T) = 0. \tag{2.3.11}
$$

Note that the latter and (2.2.12) imply that $\dot{u}(T) = 0$, as well.

2.3.3 Improved integrability

We have proved that the regularized WED functional $W^{\varepsilon\lambda}$ admits a global minimizer $u_{\varepsilon\lambda} \in K_0(u_0)$ and that this solves the regularized Euler–Lagrange problem (2.3.5)–(2.3.11). As $u_{\varepsilon\lambda}$ belongs to $H^1(0,T;H)$, by recalling assumptions (2.2.6) and (2.2.8) we have that

$J^*_\lambda\gamma_{\varepsilon\lambda} \in L^1(0,T;H)$, $\xi_{\varepsilon\lambda} \in L^2(0,T;H)$, and, by comparison in equation (2.3.5), that $-\varepsilon\dot{\xi}_{\varepsilon\lambda} + \eta^1_{\varepsilon\lambda} + \eta^2_{\varepsilon\lambda} \in L^1(0,T;H)$. By using the Lipschitz continuity of A_λ and $\mathrm{d}\phi^2_\lambda$ for $\lambda > 0$ fixed we can sharpen the statement on the integrability as follows.

Lemma 2.3.2 (Improved integrability)**.** *Let* $u_{\varepsilon\lambda} \in K_0(u_0)$ *minimize* $W^{\varepsilon\lambda}$*. Then,*

$$\begin{aligned}
&u_{\varepsilon\lambda} \in W^{2,\infty}(0,T;H),\\
&\xi_{\varepsilon\lambda} = \mathrm{d}_2\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \in W^{1,\infty}(0,T;H),\\
&\gamma_{\varepsilon\lambda} = \mathrm{d}_1\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \in L^\infty(0,T;H),\\
&\eta^1_{\varepsilon\lambda} = A_\lambda u_{\varepsilon\lambda} \in L^\infty(0,T;H),\\
&\eta^2_{\varepsilon\lambda} = \mathrm{d}\phi^2_\lambda(u_{\varepsilon\lambda}) \in L^\infty(0,T;H).
\end{aligned}$$

Proof. As A_λ, $\mathrm{d}\phi^2_\lambda : H \to H$ are Lipschitz continuous with Lipschitz constant $1/\lambda$, we have

$$\|\eta^1_{\varepsilon\lambda}\|_{L^\infty(0,T;H)} \leq \frac{1}{\lambda}\|u_{\varepsilon\lambda}\|_{L^\infty(0,T;H)} \leq c_\lambda, \tag{2.3.12}$$

and

$$\|\eta^2_{\varepsilon\lambda}\|_{L^\infty(0,T;H)} \leq \frac{1}{\lambda}\|u_{\varepsilon\lambda}\|_{L^\infty(0,T;H)} + c \leq c_\lambda,$$

where, here and in the following, we explicitly record that the constant c_λ depends on λ. Given the fact that $-\varepsilon\dot\xi_{\varepsilon\lambda} + \eta^1_{\varepsilon\lambda} + \eta^2_{\varepsilon\lambda} = -\xi_{\varepsilon\lambda} - \varepsilon J^*_\lambda \gamma_{\varepsilon\lambda} \in L^1(0,T;H)$, we deduce that $\dot\xi_{\varepsilon\lambda} \in L^1(0,T;H)$. In particular,

$$\|\xi_{\varepsilon\lambda}\|_{W^{1,1}(0,T;H)} \leq c_\lambda.$$

This, along with (2.2.12), implies that $u_{\varepsilon\lambda} \in W^{1,\infty}(0,T;H)$ and ultimately

$$\|\gamma_{\varepsilon\lambda}\|_{L^\infty(0,T;H)} \leq c_\lambda. \tag{2.3.13}$$

Then, by using estimates (2.3.12)–(2.3.13), and by a comparison in equation (2.3.5) we get

$$\|\xi_{\varepsilon\lambda}\|_{W^{1,\infty}(0,T;H)} \leq \frac{c_\lambda}{\varepsilon}.$$

As $\xi_{\varepsilon\lambda} = \mathrm{d}_2\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda})$, by using the smoothness of $\mathrm{d}_2\psi$ and differentiating in time one formally gets

$$\dot\xi_{\varepsilon\lambda} = \mathrm{d}_{21}\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) J_\lambda \dot u_{\varepsilon\lambda} + \mathrm{d}_{22}\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda})\ddot u_{\varepsilon\lambda}.$$

The latter computation can indeed be made rigorous. In fact, one has that $J_\lambda \dot u_{\varepsilon\lambda} \in L^\infty(0,T;H)$ as $\dot u_{\varepsilon\lambda} \in L^\infty(0,T;H)$ and J_λ is Lipschitz continuous in H. Assumption (2.2.10) entails that

$$\|\mathrm{d}_{21}\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) J_\lambda \dot u_{\varepsilon\lambda}\|_H \leq c_{9,R}\|\dot u_{\varepsilon\lambda}\|_H \|J_\lambda \dot u_{\varepsilon\lambda}\|_H,$$

with $R = \|J_\lambda u_{\varepsilon\lambda}\|_{L^\infty(0,T;H)}$, hence $\mathrm{d}_{21}\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) J_\lambda \dot u_{\varepsilon\lambda} \in L^\infty(0,T;H)$. By comparison

$$\mathrm{d}_{22}\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda})\ddot u_{\varepsilon\lambda} = \dot\xi_{\varepsilon\lambda} - \mathrm{d}_{21}\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) J_\lambda \dot u_{\varepsilon\lambda} \in L^\infty(0,T;H)$$

and one can use the coercivity of $\mathrm{d}_{22}\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda})$, i.e., assumption (2.2.11), to get that $\ddot u_{\varepsilon\lambda} \in L^\infty(0,T;H)$, as well. This proves the assertion.

□

2.3.4 Chain rule for ψ

In view of deriving *a priori* estimates, it is paramount to prove a chain rule for ψ. Specifically, we have the following.

Lemma 2.3.3 (Chain rule)**.** *Let $u_{\varepsilon\lambda}$ minimize $W^{\varepsilon\lambda}$ in $K_0(u_0)$. Then, for all $t \in [0,T]$ the following holds*

$$\int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\psi(J_\lambda u_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}(\xi_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) + \int_0^t \Big(-(\dot{\xi}_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) + (\gamma_{\varepsilon\lambda}, J_\lambda \dot{u}_{\varepsilon\lambda})\Big). \tag{2.3.14}$$

Proof. As $J_\lambda u_{\varepsilon\lambda}$, $\dot{u}_{\varepsilon\lambda} \in W^{1,\infty}(0,T;H)$ from Lemma 2.3.2 and $\psi \in C^2(H\times H;\mathbb{R}_+)$, one has that $t \in (0,T) \mapsto \psi(J_\lambda u_{\varepsilon\lambda}(t), \dot{u}_{\varepsilon\lambda}(t))$ is differentiable and

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi(J_\lambda u_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) = (\mathrm{d}_1\psi(J_\lambda u_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}), J_\lambda \dot{u}_{\varepsilon\lambda}) + (\mathrm{d}_2\psi(J_\lambda u_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}), \ddot{u}_{\varepsilon\lambda}).$$

More precisely, from (2.2.6) and (2.2.8) we deduce that

$$\begin{aligned}\left|\frac{\mathrm{d}}{\mathrm{d}t}\psi(J_\lambda u_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda})\right| &\le \|\mathrm{d}_1\psi(J_\lambda u_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda})\|_H \|J_\lambda \dot{u}_{\varepsilon\lambda}\|_H + \|\mathrm{d}_2\psi(J_\lambda u_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda})\|_H \|\ddot{u}_{\varepsilon\lambda}\|_H \\ &\le (c_{5,R} + c_{7,R})(1 + \|\dot{u}_{\varepsilon\lambda}\|^2)(\|J_\lambda \dot{u}_{\varepsilon\lambda}\|_H + \|\ddot{u}_{\varepsilon\lambda}\|_H)\end{aligned}$$

with $R = \|J_\lambda u_{\varepsilon\lambda}\|_{L^\infty(0,T;H)}$, proving that $t \mapsto \psi(J_\lambda u_{\varepsilon\lambda}(t), \dot{u}_{\varepsilon\lambda}(t)) \in W^{1,\infty}(0,T)$. In particular, the classical chain rule holds

$$\int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\psi(J_\lambda u_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) = \int_0^t (\xi_{\varepsilon\lambda}, \ddot{u}_{\varepsilon\lambda}) + \int_0^t (\gamma_{\varepsilon\lambda}, J_\lambda \dot{u}_{\varepsilon\lambda})$$

and equation (2.3.14) readily follows. □

2.3.5 A priori estimates

In this section, we prove *a priori* estimates on the global minimizers $u_{\varepsilon\lambda}$ independently of λ. To start with, test the regularized Euler–Lagrange equation (2.3.5) by $\dot{u}_{\varepsilon\lambda}$ and integrate it over $(0,T)$ to get

$$\int_0^T \Big(-\varepsilon(\dot{\xi}_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) + (\xi_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) + \varepsilon(\gamma_{\varepsilon\lambda}, J_\lambda \dot{u}_{\varepsilon\lambda}) + (\eta^1_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) + (\eta^2_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda})\Big) = 0.$$

Owing to the definitions (2.3.8) and (2.3.9) of $\eta^1_{\varepsilon\lambda}$ and $\eta^2_{\varepsilon\lambda}$, using the chain rule (2.3.14) from Lemma 2.3.3 with $t = T$ and recalling that $\varepsilon\xi_{\varepsilon\lambda}(T) = 0$, we obtain

$$\begin{aligned}0 = \varepsilon\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\psi(J_\lambda u_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) + \varepsilon(\xi_{\varepsilon\lambda}(0), \dot{u}_{\varepsilon\lambda}(0)) + \int_0^T (\xi_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) \\ + \int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\Big(\phi^1_\lambda(u_{\varepsilon\lambda}) + \phi^2_\lambda(u_{\varepsilon\lambda})\Big).\end{aligned}$$

By integrating the first and last terms and using $\dot{u}_{\varepsilon\lambda}(T) = 0$ and $\psi(\cdot, 0) = 0$, we infer that

$$\begin{aligned}0 = \varepsilon\Big(-\psi(J_\lambda u_0, \dot{u}_{\varepsilon\lambda}(0)) + (\xi_{\varepsilon\lambda}(0), \dot{u}_{\varepsilon\lambda}(0))\Big) + \phi^1_\lambda(u_{\varepsilon\lambda}(T)) - \phi^1_\lambda(u_0) \\ + \phi^2_\lambda(u_{\varepsilon\lambda}(T)) - \phi^2_\lambda(u_0) + \int_0^T (\xi_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}).\end{aligned}$$

This ensures that

$$\int_0^T (\xi_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) + \phi^1_\lambda(u_{\varepsilon\lambda}(T)) + \phi^2_\lambda(u_{\varepsilon\lambda}(T)) \le \phi^1_\lambda(u_0) + \phi^2_\lambda(u_0). \tag{2.3.15}$$

Here, we also used the fact that $(\xi_{\varepsilon\lambda}(0), \dot{u}_{\varepsilon\lambda}(0)) - \psi(J_\lambda u_0, \dot{u}_{\varepsilon\lambda}(0)) = \psi^*(J_\lambda u_0, \xi_{\varepsilon\lambda}(0)) \geq 0$ due to the identity (2.2.1) applied to $v \mapsto \psi(J_\lambda u_0, v)$, as well as $\psi(J_\lambda u_0, 0) = 0$. By means inequality (2.2.12) and the nonnegativity of ϕ^1_λ and ϕ^2_λ we deduce that

$$\int_0^T \|\dot{u}_{\varepsilon\lambda}\|^2_H \leq c. \tag{2.3.16}$$

Via (2.2.6) and (2.2.8) this implies that

$$\int_0^T \|\xi_{\varepsilon\lambda}\|^2_H \leq c, \tag{2.3.17}$$

and

$$\int_0^T \|\gamma_{\varepsilon\lambda}\|_H \leq c. \tag{2.3.18}$$

We test again (2.3.5) by $\dot{u}_{\varepsilon\lambda}$ and integrate it on $(0,t)$ with $t \in (0,T)$. Again from the chain rule (2.3.14) we obtain

$$\begin{aligned}
&\phi^1_\lambda(u_{\varepsilon\lambda}(t)) + \phi^2_\lambda(u_{\varepsilon\lambda}(t)) + \int_0^t (\xi_{\varepsilon\lambda}, \dot{u}_{\varepsilon\lambda}) + \varepsilon\Big((\xi_{\varepsilon\lambda}(0), \dot{u}_{\varepsilon\lambda}(0)) - \psi(J_\lambda u_0, \dot{u}_{\varepsilon\lambda}(0))\Big) = \\
&= \varepsilon\Big((\xi_{\varepsilon\lambda}(t), \dot{u}_{\varepsilon\lambda}(t)) - \psi(J_\lambda u_{\varepsilon\lambda}(t), \dot{u}_{\varepsilon\lambda}(t))\Big) + \phi^1_\lambda(u_0) + \phi^2_\lambda(u_0).
\end{aligned}$$

Thanks to (2.2.12) and the fact that $\varepsilon((\xi_{\varepsilon\lambda}(0), \dot{u}_{\varepsilon\lambda}(0)) - \psi(J_\lambda u_0, \dot{u}_{\varepsilon\lambda}(0))) \geq 0$ we deduce that

$$\begin{aligned}
\phi^1_\lambda(u_{\varepsilon\lambda}(t)) + \phi^2_\lambda(u_{\varepsilon\lambda}(t)) \leq{} & \varepsilon\Big((\xi_{\varepsilon\lambda}(t), \dot{u}_{\varepsilon\lambda}(t)) - \psi(J_\lambda u_{\varepsilon\lambda}(t), \dot{u}_{\varepsilon\lambda}(t))\Big) \\
& + \phi^1_\lambda(u_0) + \phi^2_\lambda(u_0).
\end{aligned}$$

Integrating it once more on $(0,T)$ we get

$$\int_0^T \phi^1_\lambda(u_{\varepsilon\lambda}) + \int_0^T \phi^2_\lambda(u_{\varepsilon\lambda}) \leq c\varepsilon + T\phi^1_\lambda(u_0) + T\phi^2_\lambda(u_0) \leq c. \tag{2.3.19}$$

From the bound on $\phi^1_\lambda(u_{\varepsilon\lambda})$, by using assumption (2.2.2), we obtain

$$\int_0^T \|J_\lambda u_{\varepsilon\lambda}\|^2_X \leq c. \tag{2.3.20}$$

Hence, thanks to assumption (2.2.3), we also get

$$\int_0^T \|A_\lambda u_{\varepsilon\lambda}\|^2_{X^*} \leq c.$$

From the bound on $\phi^2_\lambda(u_{\varepsilon\lambda})$, by using $\eta^2_{\varepsilon\lambda} = \partial\phi^2_\lambda(u_{\varepsilon\lambda}) \in \partial\phi^2(I_\lambda u_{\varepsilon\lambda})$, $\phi^2(I_\lambda u_{\varepsilon\lambda}) \leq \phi^2_\lambda(u_{\varepsilon\lambda})$, and assumption (2.2.4), we have

$$\int_0^T \|\eta^2_{\varepsilon\lambda}\|^r_H \leq c. \tag{2.3.21}$$

Eventually, by comparison in the regularized Euler–Lagrange equation (2.3.5), we get

$$\varepsilon\|\dot{\xi}_{\varepsilon\lambda}\|_{L^1(0,T;H)+L^2(0,T;X^*)} \leq c. \tag{2.3.22}$$

2.3.6 Passage to the limit as $\lambda \to 0$

We aim at checking that by taking the limit of $(u_{\varepsilon\lambda}, \xi_{\varepsilon\lambda}, \gamma_{\varepsilon\lambda}, \eta^1_{\varepsilon\lambda}, \eta^2_{\varepsilon\lambda})$ as $\lambda \to 0$ we obtain a strong solution of the Euler–Lagrange problem (2.2.20)–(2.2.26). Thanks to estimates (2.3.16)–(2.3.18) and (2.3.20)–(2.3.22) we find $(u_\varepsilon, \xi_\varepsilon, \gamma_\varepsilon, \eta^1_\varepsilon, \eta^2_\varepsilon)$ such that, up to not relabeled subsequences, the following convergences hold:

$$u_{\varepsilon\lambda} \rightharpoonup u_\varepsilon \quad \text{in} \quad H^1(0,T;H), \tag{2.3.23}$$
$$J_\lambda u_{\varepsilon\lambda} \rightharpoonup \beta_\varepsilon \quad \text{in} \quad H^1(0,T;H) \cap L^2(0,T;X), \tag{2.3.24}$$
$$J^*_\lambda \gamma_{\varepsilon\lambda} \rightharpoonup \gamma_\varepsilon \quad \text{in} \quad (H^1(0,T;H))^*, \tag{2.3.25}$$
$$\gamma_{\varepsilon\lambda} \rightharpoonup \tilde{\gamma}_\varepsilon \quad \text{in} \quad (H^1(0,T;H))^*, \tag{2.3.26}$$
$$A_\lambda u_{\varepsilon\lambda} \rightharpoonup \eta^1_\varepsilon \quad \text{in} \quad L^2(0,T;X^*), \tag{2.3.27}$$
$$\eta^2_{\varepsilon\lambda} \rightharpoonup \eta^2_\varepsilon \quad \text{in} \quad L^r(0,T;H), \tag{2.3.28}$$
$$\xi_{\varepsilon\lambda} \overset{*}{\rightharpoonup} \xi_\varepsilon \quad \text{in} \quad L^2(0,T;H) \cap BV([0,T];X^*), \tag{2.3.29}$$
$$\dot{\xi}_{\varepsilon\lambda} \rightharpoonup \dot{\xi}_\varepsilon \quad \text{in} \quad (H^1(0,T;H))^* + L^2(0,T;X^*). \tag{2.3.30}$$

Convergences (2.3.25), (2.3.26), and (2.3.30) are obtained from the estimate (2.3.18) by considering the continuous embedding of $L^1(0,T;H)$ into the Hilbert space $(H^1(0,T;H))^*$ and by noting that $\|J^*_\lambda\|_{\mathcal{L}(H)} \leq 1$, where $\|\cdot\|_{\mathcal{L}(H)}$ denotes the operator norm for bounded linear operators in H. Thanks to the Aubin–Lions–Simon Lemma 2.2.1, from convergence (2.3.24) we deduce the strong convergence

$$J_\lambda u_{\varepsilon\lambda} \to \beta_\varepsilon \ \text{ in } C([0,T];H). \tag{2.3.31}$$

We can pass to the limit into the regularized Euler–Lagrange equation (2.3.5) and obtain

$$-\varepsilon\dot{\xi}_\varepsilon + \xi_\varepsilon + \varepsilon\gamma_\varepsilon + \eta^1_\varepsilon + \eta^2_\varepsilon = 0 \ \text{ in } \ (H^1(0,T;H))^* + L^2(0,T;X^*). \tag{2.3.32}$$

Let us now identify the limit functions. We proceed in subsequent steps.

• *Step 1: Identification* $\tilde{\gamma}_\varepsilon = \gamma_\varepsilon$. By means of convergence (2.3.25) we have, for all $v \in H^1(0,T;H)$,

$$\langle J^*_\lambda \gamma_{\varepsilon\lambda}, v\rangle_{H^1(0,T;H)} \to \langle \gamma_\varepsilon, v\rangle_{H^1(0,T;H)}.$$

On the other hand, as J_λ is linear, we get

$$\langle J^*_\lambda \gamma_{\varepsilon\lambda}, v\rangle_{H^1(0,T;H)} = \langle \gamma_{\varepsilon\lambda}, J_\lambda v\rangle_{H^1(0,T;H)} \to \langle \tilde{\gamma}_\varepsilon, v\rangle_{H^1(0,T;H)},$$

where we have used the strong convergence of $J_\lambda v$ to v in $H^1(0,T;H)$ and the weak convergence (2.3.26) of $\gamma_{\varepsilon\lambda}$ to $\tilde{\gamma}_\varepsilon$ in $(H^1(0,T;H))^*$. Hence, we have $\tilde{\gamma}_\varepsilon = \gamma_\varepsilon$.

• *Step 2: Strong convergence of $u_{\varepsilon\lambda}$ to u_ε.* Recalling the definition of the Moreau–Yosida regularization (2.3.1) and bound (2.3.19) on $\phi^1_\lambda(u_{\varepsilon\lambda})$ in $L^1(0,T;\mathbb{R}_+)$, we get the following strong convergence:

$$\|u_{\varepsilon\lambda} - J_\lambda u_{\varepsilon\lambda}\|^2_{L^2(0,T;H)} \leq 2\lambda \|\phi^1_\lambda(u_{\varepsilon\lambda})\|_{L^1(0,T)} \to 0. \tag{2.3.33}$$

Let $\varphi \in L^2(0,T;H)$ be given. Using the weak convergence (2.3.23) of $u_{\varepsilon\lambda}$ in $H^1(0,T;H)$ along with the strong convergence (2.3.33), we infer that $J_\lambda u_{\varepsilon\lambda}$ weakly converges to u_ε in $L^2(0,T;H)$, namely,

$$\left|\int_0^T (\varphi, J_\lambda u_{\varepsilon\lambda} - u_\varepsilon)\right| \leq \int_0^T |(\varphi, J_\lambda u_{\varepsilon\lambda} - u_{\varepsilon\lambda})| + \left|\int_0^T (\varphi, u_{\varepsilon\lambda} - u_\varepsilon)\right| \to 0.$$

Hence, thanks to the uniqueness of the limit and to strong convergence (2.3.31) of $J_\lambda u_{\varepsilon\lambda}$ to β_ε in $C([0,T];H)$, we deduce that

$$J_\lambda u_{\varepsilon\lambda} \to u_\varepsilon \ \text{ in } C([0,T];H). \tag{2.3.34}$$

From the strong convergences (2.3.33) and (2.3.34) we deduce that

$$\|u_{\varepsilon\lambda} - u_\varepsilon\|_{L^2(0,T;H)} \le \|u_{\varepsilon\lambda} - J_\lambda u_{\varepsilon\lambda}\|_{L^2(0,T;H)} + \|J_\lambda u_{\varepsilon\lambda} - u_\varepsilon\|_{L^2(0,T;H)} \to 0,$$

which implies the pointwise convergence, up to a not relabeled subsequence,

$$u_{\varepsilon\lambda}(t) \to u_\varepsilon(t) \ \text{ in } H, \ \text{for a.e. } t \in (0,T). \tag{2.3.35}$$

This pointwise convergence and bound (2.3.16) on $u_{\varepsilon\lambda}$ allow us to apply the Vitali–Lebesgue Theorem (or Hölder's inequality) and obtain the strong convergence

$$u_{\varepsilon\lambda} \to u_\varepsilon \ \text{ in } L^q(0,T;H) \ \text{ for any } q < \infty. \tag{2.3.36}$$

• *Step 3: Identification of η^1_ε.* As $A_\lambda u_{\varepsilon\lambda} = A(J_\lambda u_{\varepsilon\lambda})$, A is linear and continuous from X to X^*, and $J_\lambda u_{\varepsilon\lambda}$ weakly converges to u_ε in $L^2(0,T;X)$ from (2.3.24) and (2.3.34), it follows that

$$A(J_\lambda u_{\varepsilon\lambda}) \rightharpoonup Au_\varepsilon \ \text{ in } L^2(0,T;X^*). \tag{2.3.37}$$

In particular,

$$\eta^1_\varepsilon = Au_\varepsilon \ \text{ in } X^*, \ \text{a.e. in } (0,T).$$

• *Step 4: Identification of η^2_ε.* From the definition (2.3.1) of the Moreau–Yosida regularization and bound (2.3.19) on $\phi^2_\lambda(u_{\varepsilon\lambda})$ in $L^1(0,T)$ we get

$$\|u_{\varepsilon\lambda} - I_\lambda u_{\varepsilon\lambda}\|^2_{L^2(0,T;H)} \le 2\lambda \|\phi^2_\lambda(u_{\varepsilon\lambda})\|_{L^1(0,T)} \to 0,$$

where we recall that $I_\lambda = (I + \lambda\partial\phi^2)^{-1}$ denotes the resolvent for $\partial\phi^2_\lambda$. This, along with the strong convergence (2.3.36), implies

$$I_\lambda u_{\varepsilon\lambda} \to u_\varepsilon \ \text{ in } L^2(0,T;H). \tag{2.3.38}$$

From the convexity of ϕ^2, thanks to the strong convergence (2.3.38) and to the weak convergence (2.3.28), implies that

$$\int_0^T (\eta^2_{\varepsilon\lambda}, I_\lambda u_{\varepsilon\lambda}) \to \int_0^T (\eta^2_\varepsilon, u_\varepsilon),$$

so that the identification

$$\eta^2_\varepsilon \in \partial\phi^2(u_\varepsilon) \ \text{ in } H, \ \text{a.e. in } (0,T)$$

follows by [15, Prop. 2.5, p. 27] .

• *Step 5: Identification of ξ_ε.*

As $\xi_{\varepsilon\lambda} \in W^{1,\infty}(0,T;H)$ and $u_{\varepsilon\lambda} \in H^1(0,T;H)$, we have that $t \mapsto (\xi_{\varepsilon\lambda}(t), u_{\varepsilon\lambda}(t)) \in H^1(0,T)$. Also using (2.3.11) we can integrate by parts and get

$$\int_t^T (\varepsilon\xi_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) = -\varepsilon(\xi_{\varepsilon\lambda}(t), u_{\varepsilon\lambda}(t)) + \int_t^T (-\varepsilon\dot\xi_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \quad \forall t \in [0,T].$$

We now rewrite the term in $\varepsilon\dot{\xi}_{\varepsilon\lambda}$ by using the regularized Euler–Lagrange equation (2.3.5) and integrate again over $(0,T)$ getting

$$\begin{aligned}\int_0^T\int_t^T(\varepsilon\xi_{\varepsilon\lambda},\dot{u}_{\varepsilon\lambda}) &= -\varepsilon\int_0^T(\xi_{\varepsilon\lambda},u_{\varepsilon\lambda})-\varepsilon\int_0^T\int_t^T(J_\lambda^*\gamma_{\varepsilon\lambda},u_{\varepsilon\lambda})\\ &-\int_0^T\int_t^T(\xi_{\varepsilon\lambda},u_{\varepsilon\lambda})-\int_0^T\int_t^T(\eta^2_{\varepsilon\lambda},u_{\varepsilon\lambda})-\int_0^T\int_t^T(\eta^1_{\varepsilon\lambda},u_{\varepsilon\lambda}).\end{aligned}\tag{2.3.39}$$

Our aim is to pass to the lim sup as $\lambda\to0$ in (2.3.39). Let us consider each term in the right-hand side separately. By means of the strong convergence (2.3.36) of $u_{\varepsilon\lambda}$ to u_ε in $L^q(0,T;H)$ for any $q<\infty$ and of the weak convergence (2.3.29) of $\xi_{\varepsilon\lambda}$ to ξ_ε in $L^2(0,T;H)$ we obtain

$$\int_t^T(\xi_{\varepsilon\lambda},u_{\varepsilon\lambda})\to\int_t^T(\xi_\varepsilon,u_\varepsilon)\quad\forall t\in[0,T],\tag{2.3.40}$$

as well as

$$\int_0^T\int_t^T(\xi_{\varepsilon\lambda},u_{\varepsilon\lambda})\to\int_0^T\int_t^T(\xi_\varepsilon,u_\varepsilon).\tag{2.3.41}$$

The $J_\lambda^*\gamma_{\varepsilon\lambda}$-term on the right-hand side of (2.3.39) can be treated as follows

$$\begin{aligned}\int_0^T\int_t^T(J_\lambda^*\gamma_{\varepsilon\lambda},u_{\varepsilon\lambda}) &= \int_0^T t\,(J_\lambda^*\gamma_{\varepsilon\lambda},u_{\varepsilon\lambda})=\int_0^T(\gamma_{\varepsilon\lambda},tJ_\lambda u_{\varepsilon\lambda})\\ &=\int_0^T(\gamma_{\varepsilon\lambda},tJ_\lambda u_{\varepsilon\lambda}-tu_\varepsilon)+\int_0^T(\gamma_{\varepsilon\lambda},tu_\varepsilon).\end{aligned}$$

The strong convergence (2.3.34) of $J_\lambda u_{\varepsilon\lambda}$ to u_ε in $C([0,T];H)$ and bound (2.3.18) of $\gamma_{\varepsilon\lambda}$ in $L^1(0,T;H)$ ensure that the first integral in the above right-hand side converges to 0 as $\lambda\to0$. Moreover, the weak convergence (2.3.26) of $\gamma_{\varepsilon\lambda}$ to γ_ε in $(H^1(0,T;H))^*$ ensures that

$$\int_0^T(\gamma_{\varepsilon\lambda},tu_\varepsilon)\to\langle\gamma_\varepsilon,tu_\varepsilon\rangle_{H^1(0,T;H)}.$$

All in all, we have proved that

$$\int_0^T\int_t^T(J_\lambda^*\gamma_{\varepsilon\lambda},u_{\varepsilon\lambda})\to\langle\gamma_\varepsilon,tu_\varepsilon\rangle_{H^1(0,T;H)}.\tag{2.3.42}$$

From convergences (2.3.36) and (2.3.28) we get

$$\int_t^T(\eta^2_{\varepsilon\lambda},u_{\varepsilon\lambda})\to\int_t^T(\eta^2_\varepsilon,u_\varepsilon)\quad\forall t\in[0,T]$$

as well as

$$\int_0^T\int_t^T(\eta^2_{\varepsilon\lambda},u_{\varepsilon\lambda})\to\int_0^T\int_t^T(\eta^2_\varepsilon,u_\varepsilon).\tag{2.3.43}$$

As for the last term in the right-hand side of (2.3.39), for all $t\in[0,T]$ we use the fact that

$$\int_t^T\langle A_\lambda u_{\varepsilon\lambda},J_\lambda u_{\varepsilon\lambda}\rangle_X=\int_t^T(A_\lambda u_{\varepsilon\lambda},u_{\varepsilon\lambda})-\lambda\int_t^T\|A_\lambda u_{\varepsilon\lambda}\|^2_H\leq\int_t^T(A_\lambda u_{\varepsilon\lambda},u_{\varepsilon\lambda}),$$

and rewrite it as

$$-\int_t^T (A_\lambda u_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \le -\int_t^T \langle A_\lambda u_{\varepsilon\lambda}, J_\lambda u_{\varepsilon\lambda}\rangle_X. \tag{2.3.44}$$

Since A is maximal monotone and $A_\lambda u_{\varepsilon\lambda} = A(J_\lambda u_{\varepsilon\lambda})$, we have

$$0 \le \int_t^T \langle A_\lambda u_{\varepsilon\lambda} - Au_\varepsilon, J_\lambda u_{\varepsilon\lambda} - u_\varepsilon\rangle_X,$$

from which, thanks to the weak convergences (2.3.37) and (2.3.24) of $A_\lambda u_{\varepsilon\lambda}$ to Au_ε in $L^2(0,T;X^*)$ and of $J_\lambda u_{\varepsilon\lambda}$ to u_ε in $L^2(0,T;X)$, respectively, it follows that

$$\int_0^T\int_t^T \langle Au_\varepsilon, u_\varepsilon\rangle_X \le \liminf_{\lambda\to 0} \int_0^T\int_t^T \langle A_\lambda u_{\varepsilon\lambda}, J_\lambda u_{\varepsilon\lambda}\rangle_X,$$

that is

$$\limsup_{\lambda\to 0}\Big(-\int_0^T\int_t^T \langle A_\lambda u_{\varepsilon\lambda}, J_\lambda u_{\varepsilon\lambda}\rangle_X\Big) \le -\int_0^T\int_t^T \langle Au_\varepsilon, u_\varepsilon\rangle_X. \tag{2.3.45}$$

Eventually, by combining inequalities (2.3.44) and (2.3.45) we get

$$\limsup_{\lambda\to 0}\Big(-\int_0^T\int_t^T (A_\lambda u_{\varepsilon\lambda}, u_{\varepsilon\lambda})\Big) \le -\int_0^T\int_t^T \langle Au_\varepsilon, u_\varepsilon\rangle_X. \tag{2.3.46}$$

We now pass to the $\limsup$ as $\lambda \to 0$, in equation (2.3.39). Using convergences (2.3.40)–(2.3.43) and inequality (2.3.46) we obtain

$$\begin{aligned}\limsup_{\lambda\to 0}\int_0^T\int_t^T (\varepsilon\xi_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \le\ & -\varepsilon\int_0^T (\xi_\varepsilon, u_\varepsilon) - \langle \gamma_\varepsilon, tu_\varepsilon\rangle_{H^1(0,T;H)} \\ & -\int_0^T\int_t^T \langle Au_\varepsilon, u_\varepsilon\rangle_X - \int_0^T\int_t^T (\eta_\varepsilon^2, u_\varepsilon) - \int_0^T\int_t^T (\xi_\varepsilon, u_\varepsilon).\end{aligned}$$

By using the limit equation (2.3.32) and the fact that $\eta_\varepsilon^1 = Au_\varepsilon$, we rewrite the above right-hand side as

$$\limsup_{\lambda\to 0}\int_0^T\int_t^T (\varepsilon\xi_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \le\ -\varepsilon\int_0^T (\xi_\varepsilon, u_\varepsilon) + \int_0^T \langle -\varepsilon\dot\xi_\varepsilon, u_\varepsilon\rangle_{H^1(0,T;H)\cap L^2(0,T;X)}.$$

Finally, via the integration by parts formula of Lemma 2.2.2, it follows that

$$\limsup_{\lambda\to 0}\int_0^T\int_t^T (\varepsilon\xi_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \le \int_0^T\int_t^T (\varepsilon\xi_\varepsilon, \dot u_\varepsilon)$$

or, equivalently,

$$\limsup_{\lambda\to 0}\int_0^T t(\varepsilon\xi_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \le \int_0^T t(\varepsilon\xi_\varepsilon, \dot u_\varepsilon). \tag{2.3.47}$$

Moving from this, we can now identify $\xi_\varepsilon = \mathrm{d}_2\psi(u_\varepsilon, \dot u_\varepsilon)$ by following the classical argument from [9, Lemma 3.57, p. 356 and Prop. 3.59, p. 361]. Owing to (2.2.14) for $\mu(t) = t$, by the strong convergence (2.3.34) we can get that $\int_0^T t\psi(J_\lambda u_{\varepsilon\lambda}, \cdot) \to \int_0^T t\psi(u_\varepsilon, \cdot)$ on $L^2(0,T;H)$ in the sense of Mosco. In particular, for any $v \in L^2(0,T;H)$ there exists a recovery sequence $v_\lambda \to v$ in $L^2(0,T;H)$ such that

$$\limsup_{\lambda\to 0}\int_0^T t\psi(J_\lambda u_{\varepsilon\lambda}, v_\lambda) = \int_0^T t\psi(u_\varepsilon, v).$$

On the other hand, since $\dot u_{\varepsilon\lambda} \rightharpoonup \dot u_\varepsilon$ in $L^2(0,T;H)$, we get

$$\liminf_{\lambda\to 0}\int_0^T t\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \geq \int_0^T t\psi(u_\varepsilon, \dot u_\varepsilon).$$

As $\xi_{\varepsilon\lambda} = \mathrm{d}_2\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda})$ one has that

$$\int_0^T t\psi(J_\lambda u_{\varepsilon\lambda}, v_\lambda) - \int_0^T t\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \geq \int_0^T t(\xi_{\varepsilon\lambda}, v_\lambda - \dot u_{\varepsilon\lambda}).$$

By passing to the lim sup as $\lambda \to 0$ and by using (2.3.47) we obtain

$$\int_0^T t\psi(u_\varepsilon, v) - \int_0^T t\psi(u_\varepsilon, \dot u_\varepsilon) \geq \int_0^T t(\xi_\varepsilon, v - \dot u_\varepsilon).$$

As $v \in L^2(0,T;H)$ is arbitrary, this corresponds to

$$\xi_\varepsilon = \mathrm{d}_2\psi(u_\varepsilon, \dot u_\varepsilon) \ \text{ in } H, \ \text{ a.e. in } (0,T).$$

• *Step 6: Strong convergence of* $\sqrt{t}\dot u_{\varepsilon\lambda}$. As $\xi_{\varepsilon\lambda} = \mathrm{d}_2\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda})$ and $\xi_\varepsilon = \mathrm{d}_2\psi(u_\varepsilon, \dot u_\varepsilon)$, from (2.2.13), (2.3.34), and (2.3.47) we deduce that

$$\begin{aligned} c_{12,R}\int_0^T t\|\dot u_{\varepsilon\lambda} - \dot u_\varepsilon\|_H^2 \leq & \int_0^T t\,(\xi_{\varepsilon\lambda} - \xi_\varepsilon, \dot u_{\varepsilon\lambda} - \dot u_\varepsilon) \\ & + c_{13,R}\sup_{t\in[0,T]}\|J_\lambda u_{\varepsilon\lambda} - u_\varepsilon\|_H^2 \int_0^T t\|\dot u_\varepsilon\|_H^2 \to 0, \end{aligned}$$

which implies

$$\sqrt{t}\,\dot u_{\varepsilon\lambda} \to \sqrt{t}\,\dot u_\varepsilon \quad \text{strongly in } L^2(0,T;H). \tag{2.3.48}$$

In particular, we have that

$$\dot u_{\varepsilon\lambda}(t) \to \dot u_\varepsilon(t) \quad \text{in} \ \ H, \ \text{ for a.e. } t \in (0,T). \tag{2.3.49}$$

• *Step 7: Identification of* γ_ε. As $\dot u_{\varepsilon\lambda}(t) \to \dot u_\varepsilon(t)$ in H for almost every $t \in (0,T)$ from (2.3.49), $J_\lambda u_{\varepsilon\lambda}(t) \to u_\varepsilon(t)$ in H for every $t \in [0,T]$ from (2.3.34), and $\mathrm{d}_1\psi \in C^1(H\times H;H)$ we deduce that

$$\mathrm{d}_1\psi(J_\lambda u_{\varepsilon\lambda}(t), \dot u_{\varepsilon\lambda}(t)) \to \mathrm{d}_1\psi(u_\varepsilon(t), \dot u_\varepsilon(t)) \ \text{ in } H, \ \text{ for a.e. } t \text{ in } (0,T).$$

Eventually, since by (2.2.6)

$$\|\mathrm{d}_1\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda})\|_H \leq c_{5,R}(1 + \|\dot u_{\varepsilon\lambda}\|_H^2),$$

we apply a generalization of the Dominated Convergence Theorem (see Lemma 2.2.3) and we infer that $\mathrm{d}_1\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \to \mathrm{d}_1\psi(u_\varepsilon, \dot u_\varepsilon)$ in $L^1(0,T;H)$, namely

$$\gamma_\varepsilon = \mathrm{d}_1\psi(u_\varepsilon, \dot u_\varepsilon) \ \text{ in } H, \ \text{ a.e. in } (0,T).$$

As $\gamma_\varepsilon \in L^1(0,T;H)$, we find that $\dot\xi_\varepsilon \in L^1(0,T;H) + L^2(0,T;X^*)$, so that equation (2.3.32) holds in $L^1(0,T;H) + L^2(0,T;X^*)$. In conclusion, we have

$$-\varepsilon\dot\xi_\varepsilon + \xi_\varepsilon + \varepsilon\gamma_\varepsilon + \eta_\varepsilon^1 + \eta_\varepsilon^2 = 0 \ \text{ in } X^*, \ \text{ a.e. in } (0,T).$$

2.3.7 Minimization of the WED Functional W^ε

In this section, we show that u_ε minimizes the WED functional W^ε.

For any given $v \in K(u_0)$, consider the regularized WED functional

$$W^{\varepsilon\lambda}(v) = \int_0^T e^{-t/\varepsilon}\Big(\varepsilon\psi(J_\lambda v, \dot v) + \phi^1_\lambda(v) + \phi^2_\lambda(v)\Big).$$

From the basic properties of the resolvent J_λ and of the Moreau–Yosida regularizations ϕ^1_λ and ϕ^2_λ, since $J_\lambda v \to v$ in H almost everywhere in $(0,T)$, by the continuity of ψ we deduce that $\psi(J_\lambda v, \dot v) \to \psi(v, \dot v)$ almost everywhere in $(0,T)$, that $\phi^2_\lambda(v) \to \phi^2(v)$ a.e. in $t \in (0,T)$, and that $\phi^1_\lambda(v) \to \phi^1(v)$ a.e. in $t \in (0,T)$. Moreover, by (2.2.12) it holds that

$$0 \le \varepsilon\psi(J_\lambda v, \dot v) + \phi^1_\lambda(v) + \phi^2_\lambda(v) \le \varepsilon c\left(1 + \|\dot v\|_H^2\right) + \phi^1(v) + \phi^2(v) \in L^1(0,T).$$

Thus, we may apply the Dominated Convergence Theorem and obtain

$$W^{\varepsilon\lambda}(v) \to W^\varepsilon(v) \ \text{ as } \lambda \to 0. \tag{2.3.50}$$

As $u_{\varepsilon\lambda} \in K_0(u_0)$ is a global minimizer of $W^{\varepsilon\lambda}$, we have

$$W^{\varepsilon\lambda}(v) \ge W^{\varepsilon\lambda}(u_{\varepsilon\lambda}) \ \text{ for any } v \in K(u_0). \tag{2.3.51}$$

Convergences (2.3.23), (2.3.38), and (2.3.34), the lower semicontinuity of ϕ^1 and ϕ^2, and the convergence (2.2.14) imply that

$$\begin{aligned}\liminf_{\lambda\to 0} W^{\varepsilon\lambda}(u_{\varepsilon\lambda}) &= \liminf_{\lambda\to 0} \int_0^T e^{-t/\varepsilon}\Big(\varepsilon\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) + \phi^1_\lambda(u_{\varepsilon\lambda}) + \phi^2_\lambda(u_{\varepsilon\lambda})\Big)\\ &\ge \liminf_{\lambda\to 0} \int_0^T e^{-t/\varepsilon}\Big(\varepsilon\psi(J_\lambda u_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) + \phi^1(J_\lambda u_{\varepsilon\lambda}) + \phi^2(I_\lambda u_{\varepsilon\lambda})\Big) \ge W^\varepsilon(u_\varepsilon).\end{aligned} \tag{2.3.52}$$

By combining convergence (2.3.50) with inequalities (2.3.51) and (2.3.52) we obtain

$$W^\varepsilon(v) = \lim_{\lambda\to 0} W^{\varepsilon\lambda}(v) \ge \liminf_{\lambda\to 0} W^{\varepsilon\lambda}(u_{\varepsilon\lambda}) \ge W^\varepsilon(u_\varepsilon) \quad \text{for any } v \in K(u_0),$$

which proves that $u_\varepsilon \in K(u_0)$ minimizes W^ε. In case u_ε is the unique global minimizer of W^ε, we have proved that u_ε solves the Euler–Lagrange problem (2.2.20)–(2.2.26) and Theorem 2.2.5.(ii) follows.

In case W^ε has more global minimizers, the convergence analysis from Section 2.3.6 applies just to those resulting as limits $\lim_{\lambda\to 0} u_{\varepsilon\lambda}$ of (subsequence of) global minimizers $u_{\varepsilon\lambda}$ of the regularized WED functional $W^{\varepsilon\lambda}$. In this case, we are forced to refine the convergence argument by arguing by penalization. Let $\hat u_\varepsilon \in K(u_0)$ be a fixed global minimizer of W^ε and define the penalized functionals $\hat W^\varepsilon$ and $\hat W^{\varepsilon\lambda}$ by

$$\hat W^\varepsilon(u) = \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon}(\varepsilon\psi(u,\dot u) + \hat\phi(u))\,\mathrm{d}t & \text{if } u \in K(u_0),\\ \infty & \text{else}\end{cases} \tag{2.3.53}$$

with

$$\hat\phi(u) = \phi^1(u) + \phi^2(u) + \frac{1}{2}\|u - \hat u_\varepsilon\|_H^2,$$

and

$$\hat{W}^{\varepsilon\lambda}(u) = \int_0^T e^{-t/\varepsilon}(\varepsilon\psi(J_\lambda u, \dot{u}) + \hat{\phi}_\lambda(u)),$$

with

$$\hat{\phi}_\lambda(u) = \phi^1_\lambda(u) + \phi^2_\lambda(u) + \frac{1}{2}\|u - \hat{u}_\varepsilon\|^2_H.$$

From the definition (2.3.53) of $\hat{W}^\varepsilon$ we readily check that $\hat{u}_\varepsilon$ is the unique global minimizer of $\hat{W}^\varepsilon$ in $K(u_0)$, and that $\hat{W}^\varepsilon(\hat{u}_\varepsilon) = W^\varepsilon(\hat{u}_\varepsilon) = \min_{v\in K(u_0)} W^\varepsilon(v)$. Arguing as in Section 2.3.1, we can show that $\hat{W}^{\varepsilon\lambda}$ admits a global minimizer $\tilde{u}_{\varepsilon\lambda}$. Moreover, $\tilde{u}_{\varepsilon\lambda}$ satisfies a penalized version of the regularized Euler–Lagrange problem (2.3.5)–(2.3.11), namely,

$$\begin{aligned}
&-\varepsilon\dot{\tilde{\xi}}_{\varepsilon\lambda} + \tilde{\xi}_{\varepsilon\lambda} + \varepsilon J^*_\lambda\tilde{\gamma}_{\varepsilon\lambda} + \tilde{\eta}^1_{\varepsilon\lambda} + \tilde{\eta}^2_{\varepsilon\lambda} = -\tilde{u}_{\varepsilon\lambda} + \hat{u}_\varepsilon \quad \text{in } H, \text{ a.e. in } (0,T), && (2.3.54)\\
&\tilde{\xi}_{\varepsilon\lambda} = \mathrm{d}_2\psi(J_\lambda\tilde{u}_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) \quad \text{in } H, \text{ a.e. in } (0,T), && (2.3.55)\\
&\tilde{\gamma}_{\varepsilon\lambda} = \mathrm{d}_1\psi(J_\lambda\tilde{u}_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) \quad \text{in } H, \text{ a.e. in } (0,T), && (2.3.56)\\
&\tilde{\eta}^1_{\varepsilon\lambda} = A_\lambda\tilde{u}_{\varepsilon\lambda} \quad \text{in } H, \text{ a.e. in } (0,T), && (2.3.57)\\
&\tilde{\eta}^2_{\varepsilon\lambda} = \partial\phi^2_\lambda(\tilde{u}_{\varepsilon\lambda}) \quad \text{in } H, \text{ a.e. in } (0,T), && (2.3.58)\\
&\tilde{u}_{\varepsilon\lambda}(0) = u_0, && (2.3.59)\\
&\varepsilon\tilde{\xi}_{\varepsilon\lambda}(T) = 0. && (2.3.60)
\end{aligned}$$

We now aim at recovering estimates (2.3.16)–(2.3.22) in this penalized setting. We consider the time integral over $(0,T)$ of the test of equation (2.3.54) on $\dot{\tilde{u}}_{\varepsilon\lambda}$ getting

$$\int_0^T \Big(-\varepsilon(\dot{\tilde{\xi}}_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) + (\tilde{\xi}_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) + \varepsilon(\tilde{\gamma}_{\varepsilon\lambda}, J_\lambda\dot{\tilde{u}}_{\varepsilon\lambda}) + (\tilde{\eta}^1_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) + (\tilde{\eta}^2_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda})\Big) = -\int_0^T (\tilde{u}_{\varepsilon\lambda} - \hat{u}_\varepsilon, \dot{\tilde{u}}_{\varepsilon\lambda}). \tag{2.3.61}$$

The right-hand side term can be treated as follows:

$$-\int_0^T (\tilde{u}_{\varepsilon\lambda} - \hat{u}_\varepsilon, \dot{\tilde{u}}_{\varepsilon\lambda}) = -\frac{1}{2}\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\|\tilde{u}_{\varepsilon\lambda}\|^2_H + \int_0^T (\hat{u}_\varepsilon, \dot{\tilde{u}}_{\varepsilon\lambda}).$$

Inserting this equality into (2.3.61) we get

$$\int_0^T \Big(-\varepsilon(\dot{\tilde{\xi}}_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) + (\tilde{\xi}_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) + \varepsilon(\tilde{\gamma}_{\varepsilon\lambda}, J_\lambda\dot{\tilde{u}}_{\varepsilon\lambda}) + (\tilde{\eta}^1_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) + (\tilde{\eta}^2_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda})\Big) = -\frac{1}{2}\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\|\tilde{u}_{\varepsilon\lambda}\|^2_H + \int_0^T (\hat{u}_\varepsilon, \dot{\tilde{u}}_{\varepsilon\lambda}),$$

which by Lemma 2.3.3 and $\varepsilon\tilde{\xi}_{\varepsilon\lambda}(T) = 0$ can be rearranged as follows:

$$\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\Big(\varepsilon\psi(J_\lambda\tilde{u}_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) + \phi^1_\lambda(\tilde{u}_{\varepsilon\lambda}) + \phi^2_\lambda(\tilde{u}_{\varepsilon\lambda}) + \frac{1}{2}\|\tilde{u}_{\varepsilon\lambda}\|^2_H\Big) + \int_0^T (\tilde{\xi}_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) + \varepsilon(\tilde{\xi}_{\varepsilon\lambda}(0), \dot{\tilde{u}}_{\varepsilon\lambda}(0)) = \int_0^T (\hat{u}_\varepsilon, \dot{\tilde{u}}_{\varepsilon\lambda}).$$

The latter entails that

$$\begin{aligned}
&\int_0^T (\tilde{\xi}_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) + \varepsilon\Big((\tilde{\xi}_{\varepsilon\lambda}(0), \dot{\tilde{u}}_{\varepsilon\lambda}(0)) - \psi(J_\lambda \tilde{u}_{\varepsilon\lambda}(0), \dot{\tilde{u}}_{\varepsilon\lambda}(0))\Big) \\
&\quad + \phi^1_\lambda(\tilde{u}_{\varepsilon\lambda}(T)) + \phi^2_\lambda(\tilde{u}_{\varepsilon\lambda}(T)) + \frac{1}{2}\|\tilde{u}_{\varepsilon\lambda}(T)\|^2_H \\
&\le \phi^1_\lambda(\tilde{u}_{\varepsilon\lambda}(0)) + \phi^2_\lambda(\tilde{u}_{\varepsilon\lambda}(0)) + \frac{1}{2}\|\tilde{u}_{\varepsilon\lambda}(0)\|^2_H + \delta\int_0^T \|\dot{\tilde{u}}_{\varepsilon\lambda}\|^2_H + \frac{1}{4\delta}\int_0^T \|\hat{u}_\varepsilon\|^2_H, \qquad (2.3.62)
\end{aligned}$$

where $\delta > 0$ will be chosen small, see below. From the boundedness of the terms $\phi^1_\lambda(\tilde{u}_{\varepsilon\lambda}(0)) = \phi^1_\lambda(u_0)$, $\phi^2(\tilde{u}_{\varepsilon\lambda}(0)) = \phi^2_\lambda(u_0)$, and $\|\hat{u}_\varepsilon\|_{L^2(0,T;H)}$ and from the nonnegativity of all pointwise terms in the left-hand side of (2.3.62) we infer that

$$\int_0^T (\tilde{\xi}_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) \le c + \delta \int_0^T \|\dot{\tilde{u}}_{\varepsilon\lambda}\|^2_H.$$

Hence, by choosing δ small enough, property (2.2.12) implies that

$$\int_0^T \|\dot{\tilde{u}}_{\varepsilon\lambda}\|^2_H \le c. \qquad (2.3.63)$$

By means of the bound (2.3.63), inequality (2.3.62) leads to estimates analogous to (2.3.17)–(2.3.22), where $(u_{\varepsilon\lambda}, \xi_{\varepsilon\lambda}, \gamma_{\varepsilon\lambda}, \eta^1_{\varepsilon\lambda}, \eta^2_{\varepsilon\lambda})$ are replaced by their penalized counterparts $(\tilde{u}_{\varepsilon\lambda}, \tilde{\xi}_{\varepsilon\lambda}, \tilde{\gamma}_{\varepsilon\lambda}, \tilde{\eta}^1_{\varepsilon\lambda}, \tilde{\eta}^2_{\varepsilon\lambda})$. Hence, we may extract convergent subsequences and proceed precisely as in Section 2.3.6. Eventually, we obtain, for some not relabeled subsequences, the following convergences:

$$\begin{aligned}
\tilde{u}_{\varepsilon\lambda} &\rightharpoonup \tilde{u}_\varepsilon && \text{in} \quad H^1(0,T;H), && (2.3.64)\\
\tilde{u}_{\varepsilon\lambda} &\to \tilde{u}_\varepsilon && \text{in} \quad L^q(0,T;H) \quad \forall q < \infty, && (2.3.65)\\
J^*_\lambda \tilde{\gamma}_{\varepsilon\lambda} &\rightharpoonup \tilde{\gamma}_\varepsilon && \text{in} \quad (H^1(0,T;H))^*, && (2.3.66)\\
\tilde{\eta}^1_{\varepsilon\lambda} &\rightharpoonup \tilde{\eta}^1_\varepsilon && \text{in} \quad L^2(0,T;X^*), && (2.3.67)\\
\tilde{\eta}^2_{\varepsilon\lambda} &\rightharpoonup \tilde{\eta}^2_\varepsilon && \text{in} \quad L^r(0,T;H), && (2.3.68)\\
\tilde{\xi}_{\varepsilon\lambda} &\overset{*}{\rightharpoonup} \tilde{\xi}_\varepsilon && \text{in} \quad L^2(0,T;H) \cap BV([0,T];X^*), && (2.3.69)\\
\dot{\tilde{\xi}}_{\varepsilon\lambda} &\rightharpoonup \dot{\tilde{\xi}}_\varepsilon && \text{in} \quad (H^1(0,T;H))^* + L^2(0,T;X^*), && (2.3.70)
\end{aligned}$$

where $(\tilde{u}_\varepsilon, \tilde{\xi}_\varepsilon, \tilde{\gamma}_\varepsilon, \tilde{\eta}^1_\varepsilon, \tilde{\eta}^2_\varepsilon)$ solve

$$-\varepsilon\dot{\tilde{\xi}}_\varepsilon + \tilde{\xi}_\varepsilon + \varepsilon\tilde{\gamma}_\varepsilon + \tilde{\eta}^1_\varepsilon + \tilde{\eta}^2_\varepsilon = -\tilde{u}_\varepsilon + \hat{u}_\varepsilon \quad \text{in} \quad (H^1(0,T;H))^* + L^2(0,T;X^*). \qquad (2.3.71)$$

To identify the limit functions we proceed as in Section 2.3.6, the only difference being in equation (2.3.39), namely in the identification of $\tilde{\xi}_\varepsilon = \lim_{\lambda\to 0}\tilde{\xi}_{\varepsilon\lambda}$. In this case, an additional term arises from the penalization. Indeed, we have

$$\begin{aligned}
\int_{t_1}^{t_2} (\varepsilon\tilde{\xi}_{\varepsilon\lambda}, \dot{\tilde{u}}_{\varepsilon\lambda}) &= \varepsilon(\tilde{\xi}_{\varepsilon\lambda}(t_2), \tilde{u}_{\varepsilon\lambda}(t_2)) - \varepsilon(\tilde{\xi}_{\varepsilon\lambda}(t_1), \tilde{u}_{\varepsilon\lambda}(t_1)) + \int_{t_1}^{t_2} (-\varepsilon\dot{\tilde{\xi}}_{\varepsilon\lambda}, \tilde{u}_{\varepsilon\lambda}) \\
&= \varepsilon(\tilde{\xi}_{\varepsilon\lambda}(t_2), \tilde{u}_{\varepsilon\lambda}) - \varepsilon(\tilde{\xi}_{\varepsilon\lambda}(t_1), \tilde{u}_{\varepsilon\lambda}(t_1)) + \int_{t_1}^{t_2} (-\tilde{\xi}_{\varepsilon\lambda} - \varepsilon J^*_\lambda\tilde{\gamma}_{\varepsilon\lambda} - \tilde{\eta}^1_{\varepsilon\lambda} - \tilde{\eta}^2_{\varepsilon\lambda} - \tilde{u}_{\varepsilon\lambda} + \hat{u}_\varepsilon, \tilde{u}_{\varepsilon\lambda}),
\end{aligned}$$

where in the last equality we have used the penalized version of the regularized Euler–Lagrange equation (2.3.54). To treat the additional term we note that the strong convergence (2.3.65) of

$\tilde{u}_{\varepsilon\lambda}$ to $\tilde{u}_\varepsilon$ in $L^q(0,T;H)$ for all $q<\infty$ implies the strong convergence of $\tilde{u}_{\varepsilon\lambda}-\hat{u}_\varepsilon$ to $\tilde{u}_\varepsilon-\hat{u}_\varepsilon$ in $L^2(0,T;H)$. Hence, we have that

$$\int_{t_1}^{t_2}(\tilde{u}_{\varepsilon\lambda}-\hat{u}_\varepsilon,\tilde{u}_{\varepsilon\lambda})\to\int_{t_1}^{t_2}(\tilde{u}_\varepsilon-\hat{u}_\varepsilon,\tilde{u}_\varepsilon).$$

We can hence proceed as in Section 2.3.6 and obtain,

$$\tilde{\xi}_\varepsilon=\mathrm{d}_2\psi(\tilde{u}_\varepsilon,\dot{\tilde{u}}_\varepsilon)\quad\text{in } H,\ \text{a.e. in } (0,T), \tag{2.3.72}$$
$$\tilde{\gamma}_\varepsilon=\mathrm{d}_1\psi(\tilde{u}_\varepsilon,\dot{\tilde{u}}_\varepsilon)\quad\text{in } H,\ \text{a.e. in } (0,T), \tag{2.3.73}$$
$$\tilde{\eta}^1_\varepsilon=A\tilde{u}_\varepsilon\quad\text{in } X^*,\ \text{a.e. in } (0,T), \tag{2.3.74}$$
$$\tilde{\eta}^2_\varepsilon\in\partial\phi^2(\tilde{u}_\varepsilon)\quad\text{in } H,\ \text{a.e. in } (0,T), \tag{2.3.75}$$
$$\tilde{u}_\varepsilon(0)=u_0, \tag{2.3.76}$$
$$\varepsilon\tilde{\xi}_\varepsilon(T)=0, \tag{2.3.77}$$

as well as

$$-\varepsilon\dot{\tilde{\xi}}_\varepsilon+\tilde{\xi}_\varepsilon+\varepsilon\tilde{\gamma}_\varepsilon+\tilde{\eta}^1_\varepsilon+\tilde{\eta}^2_\varepsilon=-\tilde{u}_\varepsilon+\hat{u}_\varepsilon\quad\text{in } X^* \text{ a.e. in } (0,T).$$

Eventually, $\tilde{u}_\varepsilon$ minimizes $\tilde{W}^\varepsilon$. Since $\tilde{W}^\varepsilon$ admits the unique global minimizer $\hat{u}_\varepsilon$ we have that $\tilde{u}_\varepsilon=\hat{u}_\varepsilon$. By substituting $\tilde{u}_\varepsilon=\hat{u}_\varepsilon$ in (2.3.71)–(2.3.77) we find that the global minimizer $\hat{u}_\varepsilon$ solves the Euler–Lagrange problem (2.2.20)–(2.2.26). The proof of Theorem 2.2.5.(ii) is hence complete.

2.4 The causal limit $\varepsilon\to 0$

This section is devoted to a proof of Theorem 2.2.5.(iii). First, we provide *a priori* estimates on global minimizers $(u_\varepsilon)_\varepsilon$ which are independent of ε. Then, we consider the limit $\varepsilon\to 0$ and show that $u=\lim_{\varepsilon\to 0}u_\varepsilon$ solves the gradient flow (2.2.15)–(2.2.19).

As the estimates in Subsection 2.3.5 do not depend on ε, we may pass them to the lim inf for $\varepsilon\to 0$. Hence, by using the lower semicontinuity of the norm and of ϕ^1 and ϕ^2, it is straightforward to deduce that

$$\begin{aligned}&\|u_\varepsilon\|_{H^1(0,T;H)\cap L^2(0,T;X)}+\|\xi_\varepsilon\|_{L^2(0,T;H)}+\|\gamma_\varepsilon\|_{L^1(0,T;H)}\\&\quad+\|Au_\varepsilon\|_{L^2(0,T;X^*)}+\|\eta^2_\varepsilon\|_{L^r(0,T;H)}+\|\varepsilon\dot{\xi}_\varepsilon\|_{L^1(0,T;H)+L^2(0,T;X^*)}\le c.\end{aligned} \tag{2.4.1}$$

We aim at proving that the solutions $(u_\varepsilon,\xi_\varepsilon,\eta^1_\varepsilon,\eta^2_\varepsilon)$ of the Euler–Lagrange problem (2.2.20)–(2.2.26) converge to a solution (u,ξ,Au,η^2) of the gradient flow (2.2.15)–(2.2.19). Estimate (2.4.1) implies that, up to not relabeled subsequences, the following convergences hold:

$$u_\varepsilon\rightharpoonup u\quad\text{in}\quad H^1(0,T;H)\cap L^2(0,T;X), \tag{2.4.2}$$
$$\xi_\varepsilon\rightharpoonup\xi\quad\text{in}\quad L^2(0,T;H), \tag{2.4.3}$$
$$\varepsilon\gamma_\varepsilon\to 0\quad\text{in}\quad L^1(0,T;H), \tag{2.4.4}$$
$$Au_\varepsilon\rightharpoonup Au\quad\text{in}\quad L^2(0,T;X^*), \tag{2.4.5}$$
$$\eta^2_\varepsilon\rightharpoonup\eta^2\quad\text{in}\quad L^r(0,T;H), \tag{2.4.6}$$
$$\varepsilon\dot{\xi}_\varepsilon\rightharpoonup 0\quad\text{in}\quad (H^1(0,T;H))^*+L^2(0,T;X^*). \tag{2.4.7}$$

As the sequence u_ε is bounded in $H^1(0,T;H) \cap L^2(0,T;X)$, the Aubin–Lions–Simon Lemma 2.2.1 implies that

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

The above convergences allow us to pass to the limit in the Euler–Lagrange equation (2.2.20) and to obtain

$$\xi + Au + \eta^2 = 0 \quad \text{in } X^*, \ \text{a.e. in } (0,T). \tag{2.4.8}$$

Let us now proceed with the identification of the limits η and ξ. Note preliminarily that a comparison in equation (2.4.8) guarantees that $Au \in H$ almost everywhere in $(0,T)$. In particular, we have that

$$Au = \partial\phi^1(u) \quad \text{in } H, \ \text{a.e. } (0,T),$$

so that equation (2.4.8) actually holds in H.

As u_ε strongly converges to u in $L^2(0,T;H)$ and η_ε^2 weakly converges to η^2 in $L^r(0,T;H)$, by applying [15, Prop. 2.5, p. 27] we find that

$$\eta \in \partial\phi(u) \ \text{ in } H, \ \text{a.e. in } (0,T).$$

We now pass to the limit as $\lambda \to 0$ in inequality (2.3.15). Recalling the convergences (2.3.29), (2.3.34), and (2.3.48), as well as the fact that $(\xi_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \geq 0$ a.e., for all $\delta > 0$ one has that

$$\begin{aligned}
\int_\delta^T (\xi_\varepsilon, \dot u_\varepsilon) &= \lim_{\lambda\to 0}\int_\delta^T (\xi_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \leq \lim_{\lambda\to 0}\int_0^T (\xi_{\varepsilon\lambda}, \dot u_{\varepsilon\lambda}) \\
&\leq \limsup_{\lambda\to 0}\Big(-\phi_\lambda^1(u_{\varepsilon\lambda}(T)) - \phi_\lambda^2(u_{\varepsilon\lambda}(T)) + \phi_\lambda^1(u_0) + \phi_\lambda^2(u_0)\Big) \leq -\phi(u_\varepsilon(T)) + \phi(u_0).
\end{aligned}$$

By taking $\delta \to 0$ we hence deduce that

$$\int_0^T (\xi_\varepsilon, \dot u_\varepsilon) \leq -\phi(u_\varepsilon(T)) + \phi(u_0).$$

Passing to the lim sup in the above inequality, we obtain

$$\limsup_{\varepsilon\to 0}\int_0^T (\xi_\varepsilon, \dot u_\varepsilon) \leq -\phi(u(T)) + \phi(u_0) = \int_0^T (-Au - \eta^2, \dot u) = \int_0^T (\xi, \dot u),$$

where the last equality follows from the already established limit equation (2.4.8). As $\xi_\varepsilon \rightharpoonup \xi$ in $L^2(0,T;H)$, $\dot u_\varepsilon \rightharpoonup \dot u$ in $L^2(0,T;H)$ and $u_\varepsilon \to u$ in $C([0,T];H)$, [15, Prop 2.5, p. 27] along with (2.2.14) entails that

$$\xi = \mathrm{d}_2\psi(u,\dot u) \ \text{ in } H, \ \text{a.e. in } (0,T),$$

and the assertion of Theorem 2.2.5.(iii) follows.

2.5 Application

In this section, we present an application of our abstract theory to a nonlocal problem. Let $T > 0$, Ω be an open $C^{1,1}$ bounded domain in $\mathbb{R}^d$ for $d \in \mathbb{N}$, and consider

$$g(k * u)u_t - \Delta u + \beta(u) \ni 0 \qquad \text{in } (0,T)\times\Omega, \tag{2.5.1}$$

$$u = 0 \qquad \text{on } (0,T)\times\partial\Omega, \tag{2.5.2}$$

$$u(0,\cdot) = u_0 \qquad \text{in } \Omega. \tag{2.5.3}$$

Here, the function $g \in C^{2,1}_{\rm loc}(\mathbb{R})$ is assumed to be uniformly positive, namely, $0 < \alpha \leq g(\cdot)$ for some positive constant α. Moreover, $\beta = \partial\hat\beta$ is the subdifferential of a proper, convex lower semicontinuous function $\hat\beta : \mathbb{R} \to [0,\infty]$ such that $\hat\beta(0) = 0$ with

$$|\beta(r)| \leq c(1+|r|) \ \text{ and } \ r\beta(r) \geq \frac{1}{c}|r|^2 - c \qquad \forall r \in \mathbb{R}$$

for some positive constant c. We indicate by $u_0 \in H^1_0(\Omega)$ the given initial datum. For all $u \in L^2(\mathbb{R}^d)$ we indicate by $\bar u \in L^2(\mathbb{R}^d)$ the corresponding trivial extension, namely, $\bar u = u$ on Ω and $\bar u = 0$ elsewhere. We assume to be given a kernel $k \in L^2(\mathbb{R}^d)$, so that the symbol $k * u$ denotes the convolution of k with the trivial extension $\bar u$, namely,

$$(k*u)(x) = \int_{\mathbb{R}^d} k(x-y)\bar u(y)\,\mathrm{d}y \quad \forall x \in \Omega.$$

Equation (2.5.1) may arise, in some specific setting, within the frame of the so-called Kirchhoff-like nonlocal models, see [17, 19, 20, 28] and references therein.

The WED functional corresponding to (2.5.1)–(2.5.3) is

$$W^\varepsilon(u) = \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon} \int_\Omega \left(\varepsilon g(k*u)\frac{|u_t|^2}{2} + \frac{|\nabla u|^2}{2} + \hat\beta(u)\right) \mathrm{d}x\,\mathrm{d}t & \text{if } u \in K(u_0), \\ \infty & \text{else,} \end{cases}$$

with $K(u_0)$ given by

$$K(u_0) = \{u \in H^1(0,T;L^2(\Omega)) \cap L^2(0,T;H^1_0(\Omega)) \ : \ u(0) = u_0\}.$$

Theorem 2.5.1. *Under the above assumptions, it holds that*

(i) *W^ε admits a global minimizer $u_\varepsilon \in K(u_0)$*

(ii) *$u_\varepsilon \to u$ (up to subsequences) weakly in $H^1(0,T;L^2(\Omega)) \cap L^2(0,T;H^1_0(\Omega))$ where u solves (2.5.1)–(2.5.3) in a strong sense, that is, $u \in L^2(0,T;H^2(\Omega))$ and there exists $\xi \in L^2(0,T;L^2(\Omega))$ such that $\xi \in \beta(u)$ a.e. in $(0,T)\times\Omega$ and*

$$g(k*u)u_t - \Delta u + \xi = 0 \qquad \text{a.e. in } (0,T)\times\Omega.$$

Proof. A variational formulation of (2.5.1)–(2.5.3) is obtained by letting $H = L^2(\Omega)$, $X = H^1_0(\Omega)$, and by defining

$$\psi(u,v) = \int_\Omega g(k*u)\frac{|v|^2}{2}\,\mathrm{d}x, \quad \phi^1(u) = \frac{1}{2}\int_\Omega |\nabla u|^2\,\mathrm{d}x, \quad \phi^2(u) = \int_\Omega \hat\beta(u)\,\mathrm{d}x.$$

In order to prove the statement, it suffices to show that Theorem 2.2.5 is applicable. By letting $A := \partial\phi_1$ we readily check that $A(u) = -\Delta u$ and $\beta(u) = \partial\phi^2(u)$ and $\eta^2 \in \partial\phi^2(u)$ if and only if $\eta^2 \in \beta(u)$ almost everywhere for u, $\eta^2 \in H$. Assumption (A1) is clearly satisfied.

Let us hence concentrate on Assumption (A2). To begin with, we check that $\psi \in C^0(H\times H;\mathbb{R}_+)$. Take a sequence $(u_n,v_n)_n \subset H\times H$ with $(u_n,v_n) \to (u,v)$ in $H\times H$ and use the Young inequality to bound

$$
\begin{aligned}
|\psi(u_n,v_n)-\psi(u,v)| &\le |\psi(u_n,v_n)-\psi(u_n,v)|+|\psi(u_n,v)-\psi(u,v)| \\
&\le \sup_{|\rho|\le\|k\|_{L^2(\mathbb{R}^d)}\|u_n\|_H} g(\rho)\left|\int_\Omega\left(\frac{|v_n|^2}{2}-\frac{|v|^2}{2}\right)\mathrm{d}x\right| \\
&\quad+\sup_{|\rho|\le\|k\|_{L^2(\mathbb{R}^d)}\|u_n\|_H} |g'(\rho)|\|k*(u_n-u)\|_H\left(\int_\Omega\frac{|v|^2}{2}\,\mathrm{d}x\right) \\
&\le \sup_{|\rho|\le\|k\|_{L^2(\mathbb{R}^d)}\|u_n\|_H} g(\rho)\,\frac{1}{2}\left|\|v_n\|_H^2-\|v\|_H^2\right| \\
&\quad+\sup_{|\rho|\le\|k\|_{L^2(\mathbb{R}^d)}\|u_n\|_H} |g'(\rho)|\,\frac{1}{2}\|k\|_{L^2(\mathbb{R}^d)}\|u_n-u\|_H\|v\|_H^2\to 0.
\end{aligned}
\tag{2.5.4}
$$

Let us now turn to the proof that $\psi\in C^1(H\times H;\mathbb{R}_+)$. We start by computing the Gâteaux derivative $\mathrm{d}_{G,1}\psi(u,v)$ in (u,v) of ψ with respect to the first variable along the direction $w\in H$ getting

$$(\mathrm{d}_{G,1}\psi(u,v),w)=\int_\Omega g'(k*u)\frac{|v|^2}{2}k*w\,\mathrm{d}x.$$

Arguing as above one readily proves that the function $u\mapsto(\mathrm{d}_{G,1}\psi(u,v),w)$ is continuous in H. On the other hand, as the map $w\mapsto(\mathrm{d}_{G,1}\psi(u,v),w)$ is linear and continuous, one concludes that the Gâteaux derivative $\mathrm{d}_{G,1}\psi$ and the partial Fréchet differential $\mathrm{d}_1\psi$ coincide (and are continuous). Analogously, we compute the Gâteaux derivative of $\psi(u,v)$ with respect to the second variable along the direction $w\in H$, namely,

$$(\mathrm{d}_{G,2}\psi(u,v),w)=\int_\Omega g(k*u)vw\,\mathrm{d}x.$$

Both functions $v\mapsto(\mathrm{d}_{G,2}\psi(u,v),w)$ and $w\mapsto(\mathrm{d}_{G,2}\psi(u,v),w)$ are linear and continuous in H. This proves that ψ is Fréchet differentiable with respect to the second variable with a continuous partial Fréchet differential, namely, $\mathrm{d}_2\psi=\mathrm{d}_{G,2}\psi$.

In order to prove that $\psi\in C^2(H\times H;\mathbb{R}_+)$, we compute the second Gâteaux derivatives of ψ. Owing to the fact that g is twice differentiable, by fixing two directions w, $z\in H$ we have

$$
\begin{aligned}
(\mathrm{d}_{G,11}\psi(u,v)w,z)&=\int_\Omega g''(k*u)\frac{|v|^2}{2}(k*z)(k*w)\,\mathrm{d}x,\\
(\mathrm{d}_{G,21}\psi(u,v)w,z)&=\int_\Omega g'(k*u)(k*z)vw\,\mathrm{d}x,\\
(\mathrm{d}_{G,12}\psi(u,v)w,z)&=\int_\Omega g'(k*u)(k*w)vz\,\mathrm{d}x,\\
(\mathrm{d}_{G,22}\psi(u,v)w,z)&=\int_\Omega g(k*u)wz\,\mathrm{d}x.
\end{aligned}
$$

Again, by suitably modifying the argument of (2.5.4), taking into account that $g''\in C^{0,1}(\mathbb{R})$ and $k\in L^2(\mathbb{R}^d)$ one can prove that $(u,v)\mapsto(\mathrm{d}_{G,ij}\psi(u,v)w,z)$ is continuous in $H\times H$ for all $(w,z)\in H\times H$ and all i, $j=1,2$. At the same time, $(w,z)\mapsto(\mathrm{d}_{G,ij}\psi(u,v)w,z)$ is linear and continuous for all $(u,v)\in H\times H$ and i, $j=1,2$. Hence, ψ is twice Fréchet differentiable, with continuous second Fréchet differential $\mathrm{d}^2\psi$.

For all $u\in H$, the map $v\in H\mapsto\int_\Omega g(k*u)|v|^2/2\,\mathrm{d}x$ is convex and we have $\psi(u,0)=0$.

We now proceed in proving conditions (2.2.5)–(2.2.11). Condition (2.2.5) follows straightforwardly from the fact that $g(\cdot) \geq \alpha > 0$ and

$$\psi(u,v) \geq \frac{\alpha}{2}\|v\|_H^2$$

so that one can choose $c_4 = 2/\alpha$. To check inequality (2.2.6) we compute

$$\|\mathrm{d}_1\psi(u,v)\|_H = \sup_{\|w\|_H\neq 0} \frac{(\mathrm{d}_{G,1}\psi(u,v),w)}{\|w\|_H} = \sup_{\|w\|_H\neq 0} \frac{1}{\|w\|_H}\int_\Omega g'(k*u)\frac{|v|^2}{2}k*w\,\mathrm{d}x.$$

Assuming $\|u\|_H \leq R$ and using the Young inequality we get

$$\|\mathrm{d}_1\psi(u,v)\|_H \leq \sup_{\|w\|_H\neq 0} \frac{1}{\|w\|_H}\left(\sup_{|\rho|\leq\|k\|_{L^2(\mathbb{R}^d)}R}|g'(\rho)|\right)\frac{1}{2}\|v\|_H^2\|k\|_{L^2(\mathbb{R}^d)}\|w\|_H.$$

Condition (2.2.6) follows then by choosing $c_{5,R} := \frac{1}{2}\|k\|_{L^2(\mathbb{R}^d)}\sup_{|\rho|\leq\|k\|_{L^2(\mathbb{R}^d)}R}|g'(\rho)|$. Similarly we have that

$$\begin{aligned}\|\mathrm{d}_1\psi(u_1,v)-\mathrm{d}_1\psi(u_2,v)\|_H &= \sup_{\|w\|_H\neq 0}\frac{(\mathrm{d}_{G,1}\psi(u_1,v)-\mathrm{d}_{G,1}\psi(u_2,v),w)}{\|w\|_H}\\ &= \sup_{\|w\|_H\neq 0}\frac{1}{\|w\|_H}\int_\Omega (g'(k*u_1)-g'(k*u_2))\frac{|v|^2}{2}k*w\,\mathrm{d}x.\end{aligned}$$

As g' is locally Lipschitz continuous, by using again the Young inequality we get

$$\begin{aligned}&\|\mathrm{d}_1\psi(u_1,v)-\mathrm{d}_1\psi(u_2,v)\|_H\\ &\leq \sup_{\|w\|_H\neq 0}\frac{1}{\|w\|_H}\int_\Omega \sup_{|\rho|\leq 2\|k\|_{L^2(\mathbb{R}^d)}R}|g''(\rho)||k*(u_1-u_2)|\frac{|v|^2}{2}|k*w|\,\mathrm{d}x\\ &\leq \sup_{\|w\|_H\neq 0}\frac{\sup_{|\rho|\leq 2\|k\|_{L^2(\mathbb{R}^d)}R}|g''(\rho)|}{\|w\|_H}\|k*(u_1-u_2)\|_{L^\infty(\Omega)}\frac{\|v\|_H^2}{2}\|k*w\|_{L^\infty(\Omega)}\\ &\leq \sup_{\|w\|_H\neq 0}\frac{\sup_{|\rho|\leq 2\|k\|_{L^2(\mathbb{R}^d)}R}|g''(\rho)|}{\|w\|_H}\|k\|_{L^2(\mathbb{R}^d)}\|u_1-u_2\|_H\frac{\|v\|_H^2}{2}\|k\|_{L^2(\mathbb{R}^d)}\|w\|_H.\end{aligned}$$

Then, by choosing $c_{6,R} := \frac{1}{2}\sup_{|\rho|\leq 2\|k\|_{L^2(\mathbb{R}^d)}R}|g''(\rho)|\|k\|_{L^2(\mathbb{R}^d)}^2$, we get (2.2.7).

Let us now consider

$$\|\mathrm{d}_2\psi(u,v)\|_H = \sup_{\|w\|_H\neq 0}\frac{(\mathrm{d}_{G,2}\psi(u,v),w)}{\|w\|_H} = \sup_{\|w\|_H\neq 0}\frac{1}{\|w\|_H}\int_\Omega g(k*u)vw\,\mathrm{d}x.$$

The Young inequality and $\|u\|_H \leq R$ yield

$$\|\mathrm{d}_2\psi(u,v)\|_H \leq \sup_{\|w\|_H\neq 0}\frac{1}{\|w\|_H}\|g(k*u)\|_{L^\infty(\Omega)}\|v\|_H\|w\|_H \leq \sup_{\rho\leq\|k\|_{L^2(\mathbb{R}^d)}R} g(\rho)\,\|v\|_H,$$

so that (2.2.8) follows by choosing $c_{7,R} := \sup_{\rho\leq\|k\|_{L^2(\mathbb{R}^d)}R}|g(\rho)|^2$.

We now proceed toward (2.2.9) by computing

$$\begin{aligned}\|\mathrm{d}_2\psi(u_1,v)-\mathrm{d}_2\psi(u_2,v)\|_H &= \sup_{\|w\|_H\neq 0}\frac{(\mathrm{d}_{G,2}\psi(u_1,v)-\mathrm{d}_{G,2}\psi(u_2,v),w)}{\|w\|_H}\\ &= \sup_{\|w\|_H\neq 0}\frac{1}{\|w\|_H}\int_\Omega (g(k*u_1)-g(k*u_2))vw\,\mathrm{d}x.\end{aligned}$$

As g is locally Lipschitz continuous, by applying the Young inequality we get

$$\begin{aligned}
&\|\mathrm{d}_2\psi(u_1,v)-\mathrm{d}_2\psi(u_2,v)\|_H\\
&\quad\le \sup_{\|w\|_H\neq 0}\frac{1}{\|w\|_H}\int_\Omega \sup_{|\rho|\le 2\|k\|_{L^2(\mathbb{R}^d)}R}|g'(\rho)||k*(u_1-u_2)||v||w|\,\mathrm{d}x\\
&\quad\le \sup_{\|w\|_H\neq 0}\frac{\sup_{|\rho|\le 2\|k\|_{L^2(\mathbb{R}^d)}R}|g'(\rho)|}{\|w\|_H}\|k*(u_1-u_2)\|_{L^\infty(\Omega)}\|v\|_H\|w\|_H\\
&\quad\le \sup_{|\rho|\le 2\|k\|_{L^2(\mathbb{R}^d)}R}|g'(\rho)|\|k\|_{L^2(\mathbb{R}^d)}\|u_1-u_2\|_H\|v\|_H,
\end{aligned}$$

so that (2.2.9) follows by choosing $c_{8,R}:=\sup_{|\rho|\le 2\|k\|_{L^2(\mathbb{R}^d)}R}|g'(\rho)|\|k\|_{L^2(\mathbb{R}^d)}$.

We now consider

$$\begin{aligned}
\|\mathrm{d}_{21}\psi(u,v)w\|_H &= \sup_{\|z\|_H\neq 0}\frac{(\mathrm{d}_{G,21}\psi(u,v)w,z)}{\|z\|_H}\\
&= \sup_{\|z\|_H\neq 0}\frac{1}{\|z\|_H}\int_\Omega g'(k*u)(k*z)vw\,\mathrm{d}x.
\end{aligned}$$

Thanks to the Young inequality and to $\|u\|\le R$ we obtain

$$\begin{aligned}
\|\mathrm{d}_{21}\psi(u,v)w\|_H &\le \sup_{\|z\|_H\neq 0}\frac{1}{\|z\|_H}\|g'(k*u)\|_{L^\infty(\Omega)}\|k*z\|_{L^\infty(\Omega)}\|v\|_H\|w\|_H\\
&\le \sup_{\rho\le\|k\|_{L^2(\mathbb{R}^d)}R}|g'(\rho)|\|k\|_{L^2(\mathbb{R}^d)}\|v\|_H\|w\|_H,
\end{aligned}$$

so that (2.2.10) follows for $c_{9,R}:=\sup_{\rho\le\|k\|_{L^2(\mathbb{R}^d)}R}|g'(\rho)|\|k\|_{L^2(\mathbb{R}^d)}$. By computing

$$(\mathrm{d}_{22}\psi(u,v)w,w)=\int_\Omega g(k*u)|w|^2\,\mathrm{d}x$$

and recalling that $g(\cdot)\ge\alpha>0$, one obtains (2.2.11) with $c_{10,R}:=\alpha$.

Having checked Assumption (A2), we are in the position of applying Theorem 2.2.5 which concludes the proof. $\square$

Before closing this section, let us remark that the abstract theory applies to other classes of PDEs, as well. For instance one could consider the PDE problem

$$\begin{aligned}
g(\|u\|^2_{L^2(\Omega)})u_t-\Delta u+\beta(u)\ni 0 &\qquad \text{a.e. in } (0,T)\times\Omega,\\
u=0 &\qquad \text{a.e. on } (0,T)\times\partial\Omega,\\
u(0,x)=u_0(x) &\qquad \text{a.e. in } \Omega.
\end{aligned}$$

Under the same assumptions above we can prove an analogous version of Theorem 2.5.1.

Acknowledgement

This research was funded in whole or in part by the Austrian Science Fund (FWF) grants 10.55776/F65, 10.55776/I4354, 10.55776/I5149, 10.55776/P32788, and 10.55776/W1245. For open access purposes, the authors have applied a CC BY public copyright license to any author accepted manuscript version arising from this submission. US acknowledges financial support from the OeAD-WTZ grant CZ 05/2024. GA is also supported by JSPS KAKENHI Grant Numbers JP21KK0044, JP21K18581, JP20H01812 and JP20H00117 as well as the Research Institute for Mathematical Sciences, an International Joint Usage/Research Center located in Kyoto University.

Chapter 3

Optimal control of gradient flows via the Weighted Energy-Dissipation method

This chapter consists of my submitted manuscript with Takeshi Fukao and Ulisse Stefanelli [29].

Abstract

We consider a general optimal control problem in the setting of gradient flows. Two approximations of the problem are presented, both relying on the variational reformulation of gradient-flow dynamics via the Weighted-Energy-Dissipation variational approach. This consists in the minimization of global-in-time functionals over trajectories, combined with a limit passage. We show that the original nonpenalized problem and the two successive approximations admits solutions. Moreover, resorting to a Γ-convergence analysis we show that penalised optimal controls converge to nonpenalized one as the approximation is removed.

3.1 Introduction

This paper is concerned with an optimal control problem for abstract gradient flows in Hilbert spaces. We are interested in finding a solution to the following problem

$$\min_{(y,u)\in H^1(0,T;H)\times U} P(y,u). \tag{$\mathcal{P}$}$$

The *control* $u : [0,T] \to H$ and the gradient flow $y : [0,T] \to H$ are trajectories in the Hilbert space H and $T > 0$ is some final reference time. The set U of *admissible controls* is assumed to be compact in $L^2(0,T;H)$ and the functional P is prescribed as by

$$P(y,u) = \begin{cases} J(y,u) & \text{if } y = S(u), \\ \infty & \text{else,} \end{cases}$$

where J is a given *target functional*, which is assumed to be lower semicontinuous with respect to the weak $\times$ strong topology in $H^1(0,T;H) \times L^2(0,T;H)$. Moreover, $S(u)$ indicates the unique solution $y \in H^1(0,T;H)$, given $u \in L^2(0,T;H)$ to the *gradient flow problem*

$$\dot{y}(t) + \partial\phi(y(t)) \ni u(t) \quad \text{for a.e. } t \in (0,T), \tag{3.1.1}$$

$$y(0) = y^0. \tag{3.1.2}$$

The dot in (3.1.1) indicates the time derivative of y. The functional $\phi : H \to (-\infty, \infty]$ is assumed to be proper, lower semicontinuous, and κ-convex for some $\kappa \in \mathbb{R}$, i.e., $y \mapsto \psi(u) = \phi(y) - \kappa\|y\|^2/2$ is convex. In particular, $\partial\phi(u) = \partial\psi(u) + \kappa u$, where $\partial\psi(u)$ is the subdifferential in the sense of convex analysis. Finally, $y^0 \in D(\partial\phi)$.

The optimal control Problem $(\mathcal{P})$ can be readily proved to admit a solution, namely an optimal pair (y,u). The focus of this note is on the possible approximation of Problem $(\mathcal{P})$ by means of two penalized problems. The departing point for such approximation is the so called *Weighted Energy-Dissipation* (WED) approach to the gradient flow problem. This consists in the minimization of a family of ε-dependent global-in-time *WED* functionals $W_\varepsilon : H^1(0,T;H) \times L^2(0,T;H) \to (-\infty, \infty]$ given by

$$W_\varepsilon(y,u) = \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon}\left(\frac{\varepsilon}{2}\|\dot{y}(t)\|^2 + \phi(y(t)) - (u,y)\right) dt & \text{if } (y,u) \in K(y^0) \times U, \\ \infty & \text{else,} \end{cases} \tag{3.1.3}$$

where $\varepsilon > 0$ and the convex closed set $K(y^0)$ is given by

$$K(y^0) = \{y \in H^1(0,T;H) : y(0) = y^0 \in H,\, \phi \circ y \in L^1(0,T)\}.$$

Indeed, given $u \in U$, the link between the minimization of $W_\varepsilon(\cdot, u)$ and the gradient flow problem is revealed by computing the corresponding Euler-Lagrange problem

$$-\varepsilon\ddot{y}_\varepsilon(t) + \dot{y}_\varepsilon(t) + \partial\phi(y_\varepsilon(t)) \ni u(t) \quad \text{for a.e. } t \in (0,T), \tag{3.1.4}$$

$$y_\varepsilon(0) = y^0, \tag{3.1.5}$$

$$\varepsilon\dot{y}_\varepsilon(T) = 0. \tag{3.1.6}$$

The latter is nothing but an elliptic-in-time regularization of the gradient flow problem, which is recovered by taking $\varepsilon \to 0$. More precisely, owing to [45], for ε small enough one can uniquely define

$$y_\varepsilon^u = \arg\min W_\varepsilon(\cdot, u)$$

and prove that $y_\varepsilon^u \rightharpoonup y = S(u)$ in $H^1(0,T;H)$, see Proposition 3.2.1 below. Note the occurrence of the final condition (3.1.6), which eventually is dropped in the $\varepsilon \to 0$ limit.

In the setting of gradient flows, the WED approach has been applied to mean curvature [31, 55], to periodic solvability [32], to micro-structure evolution [21], to the incompressible Navier-Stokes system [10, 47], and to stochastic PDEs [50]. The linear case is mentioned in the classical PDE textbook by Evans [27, Problem 3, p. 487].

The general theory for κ-convex energies ϕ can be found in [45], while the nonconvex setting is discussed in [7] and [40] deals with nonpotential perturbations. A stability result via Γ-convergence is in [35] and [39] uses the WED approach for studying symmetries of solutions. In the more general setting of metric spaces, curves of maximal slope have also been studied [48, 49].

Besides the gradient flow case, the WED approach has also been considered in the doubly nonlinear setting, see [42, 44] for rate-independent and [2, 3, 4, 5, 6] for rate-dependent theories.

In addition, the hyperbolic case of semilinear waves has been considered in [53, 56], also including forcings [38, 57, 58, 59] or dissipative terms [1, 37, 52]. Applications to dynamic fracture [23, 33], dynamic plasticity [24], and Lagrangian mechanics [36] are also available.

The aim of this paper is that of investigating two approximations of Problem $(\mathcal{P})$, based on the WED functionals W_ε. At first, we approximate the constraint $y = S(u)$ by considering the optimal control problem

$$\min_{(y,u)\in H^1(0,T;H)\times U} P_\varepsilon(y,u), \tag{$\mathcal{P}_\varepsilon$}$$

where the approximating functional P_ε is defined for all $\varepsilon > 0$ as

$$P_\varepsilon(y,u) = \begin{cases} J(y,u) & \text{if } y \in \arg\min W_\varepsilon(\cdot,u), \\ \infty & \text{else.} \end{cases}$$

This essentially amounts to replacing the gradient-flow constraint (3.1.1)-(3.1.2) by its elliptic regularization (3.1.4)-(3.1.6). Problem $(\mathcal{P}_\varepsilon)$ is still a constrained optimization. Nevertheless the constraint is itself expressed by a minimization, turning $(\mathcal{P}_\varepsilon)$ into a *bilevel* optimization problem. The internal minimization layer is actually a convex problem, which is therefore accessible to efficient optimization techniques. Another distintive feature of this approach is that WED minimizers y_ε^u turn out to be more regular than the corresponding gradient flows $y \in S(u)$, see Proposition 3.2.1.

As a second option, we consider a penalization of the constraint $y \in \arg\min W_\varepsilon(\cdot,u)$ in $(\mathcal{P}_\varepsilon)$, tuning to an unconstrained optimal control problem, namely,

$$\min P_{\varepsilon\lambda}(y,u), \tag{$\mathcal{P}_{\varepsilon\lambda}$}$$

with $P_{\varepsilon\lambda}$ given for all ε, $\lambda > 0$ as

$$P_{\varepsilon\lambda}(y,u) = J(y,u) + \frac{1}{\lambda}(W_\varepsilon(y,u) - M_\varepsilon^u).$$

Here, $\lambda > 0$ serves as penalization parameter and M_ε^u denotes the minimum value of $W_\varepsilon(\cdot,u)$, given u, i.e., $M_\varepsilon^u = W_\varepsilon(y_\varepsilon^u,u)$. In particular, the functional $y \mapsto W_\varepsilon(y,u) - M_\varepsilon^u$ is nonnegative and vanishes iff $y = y_\varepsilon^u$. Although the unconstrained formulation of $(\mathcal{P}_{\varepsilon\lambda})$ is appealing, one shall stress that the computation of y_ε^u, i.e., the minimization of $W_\varepsilon(\cdot,u)$ is still needed for evaluating the minimum value M_ε^u, see Lemma 3.3.1 below. A remarkable feature, however, is that the functional $y \mapsto W_\varepsilon(y,u) - M_\varepsilon^u$ controls the distance $\|y - y_\varepsilon^u\|^2_{H^1(0,T;H)}$, which may be of some applicative interest.

Our main result is Theorem 3.2.2 below. We first check that Problems $(\mathcal{P})$, $(\mathcal{P}_\varepsilon)$, and $(\mathcal{P}_{\varepsilon\lambda})$ are solvable, namely, there exist optimal pairs (y,u), $(y_\varepsilon,u_\varepsilon)$, and $(y_{\varepsilon\lambda},u_{\varepsilon\lambda})$, respectively (Theorem 3.2.2.i). Then, we prove that

$$P_{\varepsilon\lambda} \xrightarrow{\Gamma} P_\varepsilon \quad \text{and} \quad P_\varepsilon \xrightarrow{\Gamma} P$$

in the sense of Γ-convergence with respect to the weak×strong topology of $H^1(0,T;H) \times L^2(0,T;H)$, which we indicate as τ. Upon checking the coercivity of P_ε and $P_{\varepsilon\lambda}$ with respect to topology τ, this allows us to prove that, for some not relabeled subsequences,

$$(y_\varepsilon,u_\varepsilon) \xrightarrow{\tau} (y,u) \quad \text{and} \quad (y_{\varepsilon\lambda},u_{\varepsilon\lambda}) \xrightarrow{\tau} (y_\varepsilon,u_\varepsilon)$$

as $\varepsilon \to 0$ and $\lambda \to 0$, respectively, (Theorem 3.2.2.ii-iii).

Eventually, we tackle the joint $(\varepsilon,\lambda) \to (0,0)$ in the specific case $\lambda = \lambda_\varepsilon$ with $\limsup_{\varepsilon\to 0} \lambda_\varepsilon \varepsilon^{-3} e^{T/\varepsilon} = 0$ proving that

$$P_{\varepsilon\lambda} \xrightarrow{\Gamma} P.$$

Again due to coerciveness, this entails that $(y_{\varepsilon,\lambda}) \xrightarrow{\mathcal{T}} (y,u)$ along some not relabeled subsequence (Theorem 3.2.2.iv).

The plan of the paper is the following. In Section 3.2, we introduce the assumptions and state our main result, namely, Theorem 3.2.2. Section 3.3 collects a series of lemmas, which will be used throughout. The proof of Theorem 3.2.2 is in Sections 3.4-3.5. Specifically, existence for the Problems $(\mathcal{P})$, $(\mathcal{P}_\varepsilon)$, and $(\mathcal{P}_{\varepsilon\lambda})$ is checked in Section 3.4 and convergences as $\varepsilon \to 0$ and $\lambda \to 0$ are discussed in Section 3.5.

3.2 Statement of main results

In this section, we introduce assumptions, recall some result from [45], and state our main result, i.e., Theorem 3.2.2. Let us start by assuming that

(A1) H is a real Hilbert space, $T > 0$, and $U \subset L^2(0,T;H)$ is nonempty and compact.

We indicate by $\|\cdot\|$ and by $(\cdot,\cdot)$ the norm and the inner product in H, respectively, and we indicate by $\|\cdot\|_E$ the norm in the generic Banach space E.

Concerning the functional ϕ we ask that

(A2) $\phi : H \to (-\infty,\infty]$ is proper, κ-convex, and lower semicontinuous, and $y^0 \in D(\partial\phi)$.

In particular, the *effective domain* $D(\phi) = \{v \in H \ : \ \phi(v) < \infty\}$ is not empty. Moreover,

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

and $D(\psi) = D(\phi)$. Equivalently, one can state κ-convexity of ϕ as

$$\phi(rw + (1-r)v) \le r\phi(w) + (1-r)\phi(v) - \frac{\kappa}{2} r(1-r)\|w-v\|^2 \quad \forall w,v \in H,\ 0 \le r \le 1.$$

Correspondingly, we define the (Fréchet) subdifferential $\partial\phi : H \to 2^H$ as $\partial\phi(y) = \partial\psi(y) + \kappa y$, where $\partial\psi$ is the subdifferential of ψ in the sense of convex analysis [16]. This implies that one has $\eta \in \partial\phi(y)$ iff $y \in D(\phi)$ and

$$(\eta, w-y) \le \phi(w) - \phi(y) + \frac{\kappa}{2}\|w-y\|^2 \quad \forall w \in D(\phi),$$

$D(\partial\phi) = \{y \in D(\phi) \ : \ \partial\phi(y) \neq \emptyset\} = D(\partial\psi)$, and $\partial\phi(y)$ is convex and closed, for all $y \in D(\partial\phi)$. In particular, for all $y \in D(\partial\phi)$ one has that $\partial\phi(y)$ has a unique element of minimal norm, which we indicate by $(\partial\phi(y))^\circ$.

The next proposition summarizes results from [45].

Proposition 3.2.1 (WED approach to gradient flows)**.** *Under assumptions* (A1)-(A2)*, there exists $\varepsilon_0 > 0$ so that for all $\varepsilon \in (0,\varepsilon_0)$ and all $u \in U$ the functional $W_\varepsilon(\cdot,u)$ is κ_ε-convex in $H^1(0,T;H)$ for some $\kappa_\varepsilon > 0$. In particular, there exists a unique minimizer $y^u_\varepsilon = \arg\min W_\varepsilon(\cdot,u)$. One has that $y^u_\varepsilon \in H^2(0,T;H)$ is the unique solution of the Euler-Lagrange problem* (3.1.4)-(3.1.6) *and that $\eta^u_\varepsilon = u + \varepsilon\ddot{y}^u_\varepsilon - \dot{y}^u_\varepsilon \in L^2(0,T;H)$ and fulfills $\eta^u_\varepsilon \in \partial\phi(y^u_\varepsilon)$ a.e. Moreover, there exists a nondecreasing function $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ independent of ε such that*

$$\begin{aligned} &\varepsilon\|\ddot{y}^u_\varepsilon\|_{L^2(0,T;H)} + \varepsilon^{1/2}\|\dot{y}^u_\varepsilon\|_{L^\infty(0,T;H)} + \|\dot{y}^u_\varepsilon\|_{L^2(0,T;H)} + \|\eta^u_\varepsilon\|_{L^2(0,T;H)} \\ &\quad \le \ell(\|u\|_{L^2(0,T;H)} + \|y^0\| + \|(\partial\phi(y^0))^\circ\|), \end{aligned} \tag{3.2.1}$$

where $(\partial\phi(y^0))^\circ$ is the element of minimal norm in $\partial\phi(y^0)$. As $\varepsilon \to 0$, one has that y^u_ε converges to $y \in S(u)$ weakly in $H^1(0,T;H)$ and strongly in $H^\sigma(0,T;H)$ for all $\sigma \in (0,1)$.

In all of the following, we will tacitly assume that $\varepsilon \in (0, \varepsilon_0)$, so that Proposition 3.2.1 holds. In particular, $y^u_\varepsilon = \arg\min W_\varepsilon(\cdot, u)$ is well-defined.

Concerning the target functional J we assume that

(A3) $J : H^1(0,T;H) \times L^2(0,T;H) \to \mathbb{R}_+$, $\mathbb{R}_+ := [0,\infty)$, is lower semicontinuous with respect to τ. Moreover, for all $u \in L^2(0,T;H)$ given, $y \mapsto J(y,u)$ is upper semicontinuous with respect to the strong $H^\sigma(0,T;H)$ topology, for some $\sigma \in (0,1)$.

A possible example of functional J fulfilling (A3) is

$$J(y,u) = f(y(T)) + \int_0^T g(y,u)\,\mathrm{d}t$$

where $f : H \to \mathbb{R}$ is continuous and $g : H \times H \to \mathbb{R}$ is continuous and bounded.

Let us recall that, given functionals F_ρ, $F : H^1(0,T;H) \times L^2(0,T;H) \to \mathbb{R} \cup \{\infty\}$ for $\rho > 0$, we say that the sequence $(F_\rho)_\rho$ *Γ-converges* to F with respect to the topology τ and we write $F = \Gamma_\tau \lim_{\rho\to 0} F_\rho$ if the following conditions hold

(i) (Γ-lim inf inequality) For every $(y_\rho, u_\rho) \xrightarrow{\tau} (y,u)$ we have

$$F(y,u) \le \liminf_{\rho\to 0} F_\rho(y_\rho, u_\rho); \tag{3.2.2}$$

(ii) (Recovery sequence) For every $(\hat{y}, \hat{u}) \in H^1(0,T;H) \times L^2(0,T;H)$ there exists $(\hat{y}_\rho, \hat{u}_\rho) \xrightarrow{\tau} (\hat{y}, \hat{u})$ such that

$$\limsup_{\rho\to 0} F_\rho(\hat{y}_\rho, \hat{u}_\rho) \le F(\hat{y}, \hat{u}). \tag{3.2.3}$$

We now state our main result.

Theorem 3.2.2 (WED approach to optimal control)**.** *Assume* (A1)-(A3)*. Then:*

i) *For all $\varepsilon, \lambda > 0$ Problems $(\mathcal{P})$, $(\mathcal{P}_\varepsilon)$, and $(\mathcal{P}_{\varepsilon\lambda})$ admit solutions.*

ii) *$P_\varepsilon \xrightarrow{\Gamma} P$ as $\varepsilon \to 0$. Any sequence $(y_\varepsilon, u_\varepsilon)_\varepsilon$ of solutions to Problem $(\mathcal{P}_\varepsilon)$ admits a not relabeled subsequence such that $(y_\varepsilon, u_\varepsilon) \xrightarrow{\tau} (y,u)$ where (y,u) solves Problem $(\mathcal{P})$.*

iii) *$P_{\varepsilon\lambda} \xrightarrow{\Gamma} P_\varepsilon$ as $\lambda \to 0$, for all $\varepsilon > 0$ fixed. Any sequence $(y_{\varepsilon\lambda}, u_{\varepsilon\lambda})_\lambda$ of solutions to the Problem $(\mathcal{P}_{\varepsilon\lambda})$ admits a not relabeled subsequence such that $(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \xrightarrow{\tau} (y_\varepsilon, u_\varepsilon)$ where $(y_\varepsilon, u_\varepsilon)$ solves Problem $(\mathcal{P}_\varepsilon)$.*

iv) *Let $\lambda = \lambda_\varepsilon$ with $\limsup_{\varepsilon\to 0} \lambda_\varepsilon \varepsilon^{-3} e^{-T/\varepsilon} = 0$. Then $P_{\varepsilon\lambda_\varepsilon} \xrightarrow{\Gamma} P$ as $\varepsilon \to 0$. Any sequence of solutions $(y_{\varepsilon\lambda_\varepsilon}, u_{\varepsilon\lambda_\varepsilon})$ to Problem $(\mathcal{P}_{\varepsilon\lambda})$ admits a not relabeled subsequence such that $(y_{\varepsilon\lambda_\varepsilon}, u_{\varepsilon\lambda_\varepsilon}) \xrightarrow{\tau} (y,u)$ where (y,u) solves Problem $(\mathcal{P})$.*

The proof of Theorem 3.2.2 is given in Section 3.4-3.5. More precisely, we give a proof of Theorem 3.2.2.i in Section 3.4 whereas Theorem 3.2.2.ii-iv is proved in Section 3.5.

3.2.1 An example

Before closing this section, we give an illustration of the results by resorting to simple ODE example. In particular, we consider the ODE

$$\dot{y} + y = u, \quad y(0) = 1 \tag{3.2.4}$$

with $U = \{u \,:\, u(t) = u_0 e^{-t} \text{ for some } u_0 \in [0,1]\}$. We are interested in minimizing

$$J(y,u) = \frac{1}{2}\int_0^1 (y(t) - e^{-t})^2 \,\mathrm{d}t + \frac{1}{2}\int_0^1 t^2 (u(t) - e^{-t})^2 \,\mathrm{d}t.$$

Note that this fits in the theory by letting $H = \mathbb{R}$, $T = 1$, and $\phi(y) = y^2/2$. In particular, U is clearly compact into $L^2(0,1)$.

Problem $\mathcal{P}$ reads

$$\min_{u_0\in[0,1]} \left\{ J(y,u) \ :\ \dot{y}(t) + y(t) = u_0 e^{-t},\ y(0) = 1 \right\},$$

and can be directly solved. For all u_0 the solution of (3.2.4) is $y(t) = (1 + tu_0)e^{-t}$ and we have

$$\begin{aligned} J(y,u) &= \frac{1}{2}\int_0^1 u_0^2 t^2 e^{-2t}\,\mathrm{d}t + \frac{1}{2}\int_0^1 (u_0 - 1)^2 t^2 e^{-2t}\,\mathrm{d}t \\ &= \frac{1}{8}(1 - 5e^{-2})(u_0^2 + (u_0 - 1)^2) \end{aligned}$$

which is minimized at $u_0 = 1/2$ with value $(1 - 5e^{-2})/16$ and a corresponding optimal trajectory $y(t) = (1 + t/2)e^{-t}$.

Let us now turn to Problem $\mathcal{P}_\varepsilon$ which can be written as

$$\min_{u_0\in[0,1]} \left\{ J(y,u) \ :\ -\varepsilon\ddot{y}(t) + \dot{y}(t) + y(t) = u_0 e^{-t},\ y(0) = 1,\ y'(1) = 0 \right\}.$$

Given $u \in U$, the only solution y_ε^u to

$$-\varepsilon\ddot{y}(t) + \dot{y}(t) + y(t) = u_0 e^{-t}, \quad y(0) = 1, \quad y'(1) = 0$$

is given by

$$y_\varepsilon^u(t) = c_\varepsilon^- e^{r_\varepsilon^- t} + c_\varepsilon^+ e^{r_\varepsilon^+ t} - \frac{u_0}{\varepsilon} e^{-t},$$

where

$$\begin{aligned} r_\varepsilon^- &= \frac{1 - \sqrt{4\varepsilon + 1}}{2\varepsilon}, \qquad r_\varepsilon^+ = \frac{1 + \sqrt{4\varepsilon + 1}}{2\varepsilon}, \\ c_\varepsilon^- &= \frac{\frac{u_0}{\varepsilon e} + r_\varepsilon^+ (1 + \frac{u_0}{\varepsilon}) e^{r_\varepsilon^+}}{r_\varepsilon^+ e^{r_\varepsilon^+} - r_\varepsilon^- e^{r_\varepsilon^-}}, \qquad c_\varepsilon^+ = 1 + \frac{u_0}{\varepsilon} - c_\varepsilon^-. \end{aligned}$$

The value of J at (y_ε^u, u) can be explicitly computed as a function of u_0 as

$$\begin{aligned} u_0 \mapsto j_\varepsilon(u_0) &:= J(y_\varepsilon^u, u_0 e^{-t}) \\ &= \frac{1}{2}\Bigg(\frac{(c_\varepsilon^-)^2}{2r_\varepsilon^-}(e^{2r_\varepsilon^-} - 1) + \frac{(c_\varepsilon^+)^2}{2r_\varepsilon^+}(e^{2r_\varepsilon^+} - 1) - \frac{1}{2}\left(\frac{u_0}{\varepsilon} + 1\right)^2 (e^{-2} - 1) \\ &\quad + \frac{2c_\varepsilon^- c_\varepsilon^+}{r_\varepsilon^- + r_\varepsilon^+}(e^{r_\varepsilon^- + r_\varepsilon^+} - 1) - \frac{2c_\varepsilon^-}{r_\varepsilon^- - 1}\left(\frac{u_0}{\varepsilon} + 1\right)(e^{r_\varepsilon^- - 1} - 1) \\ &\quad - \frac{2c_\varepsilon^+}{r_\varepsilon^+ - 1}\left(\frac{u_0}{\varepsilon} + 1\right)(e^{r_\varepsilon^+ - 1} - 1) \Bigg) + \frac{1}{8}(1 - 5e^{-2})(u_0 - 1)^2 \end{aligned}$$

A tedious but elementary computation ensures that $u_0 \mapsto J(y^u_\varepsilon, u_0 e^{-t})$ converges to $u_0 \mapsto (1/8)(1-5e^{-2})(u_0^2 + (u_0-1)^2)$ uniformly on $[0,1]$. In particular, the minimum value of j_ε on $[0,1]$ converges to that of j_0, namely $(1-5e^{-2})/16$, and the minimizers of j_ε converge to the minimizer $1/2$, as well.

We conclude by considering Problem $\mathcal{P}_{\varepsilon\lambda}$. This amount to the following

$$\min_{\substack{u_0\in[0,1],\\ y\in H^1(0,1),\\ y(0)=1}} \left\{ J(y,u) + \frac{1}{\lambda}\left(\int_0^1 e^{-t/\varepsilon}\left(\frac{\varepsilon}{2}|\dot{y}(t)|^2 + \frac{1}{2}|y(t)|^2 - u_0 e^{-t} y(t)\right) \mathrm{d}t - M^u_\varepsilon \right)\right\}.$$

In order to tackle this minimization problem, one needs to evaluate M^u_ε, which generally calls for another minimization. In this example however one can use the above expression for y^u_ε and explicitly compute M^u_ε

$$\begin{aligned} M^u_\varepsilon = & \frac{\varepsilon^2(r^-_\varepsilon)^2(c^-_\varepsilon)^2}{4\varepsilon r^-_\varepsilon - 2}(e^{2r^-_\varepsilon - 1/\varepsilon} - 1) + \frac{\varepsilon^2(r^+_\varepsilon)^2(c^+_\varepsilon)^2}{4\varepsilon r^+_\varepsilon - 2}(e^{2r^+_\varepsilon - 1/\varepsilon} - 1)\frac{u_0^2}{4\varepsilon + 2}(e^{-2-1/\varepsilon} - 1) \\ & + \frac{c^-_\varepsilon c^+_\varepsilon r^-_\varepsilon r^+_\varepsilon}{\varepsilon(r^-_\varepsilon + r^+_\varepsilon) - 1}(e^{r^-_\varepsilon + r^+_\varepsilon - 1/\varepsilon} - 1) + \frac{\varepsilon c^-_\varepsilon}{\varepsilon r^-_\varepsilon - 1}(e^{r^-_\varepsilon - 1/\varepsilon} - 1) \\ & + \frac{\varepsilon c^+_\varepsilon}{\varepsilon r^+_\varepsilon - 1}(e^{r^+_\varepsilon - 1/\varepsilon} - 1) + \frac{u_0}{1+\varepsilon}(e^{-1-1/\varepsilon} - 1). \end{aligned}$$

3.3 Preliminary lemmas

Before moving to the proof of Theorem 3.2.2, we present in this section some preliminary lemmas, which will be used throughout and which complement the analysis in [45].

To start with, let us explicitly remark that, although ϕ is just κ-convex, the classical tools from convex analysis apply to $\partial\phi$, as well. In particular, we have that

$$y_n \rightharpoonup y, \ \eta_n \rightharpoonup \eta, \ \eta_n \in \partial\phi(y_n), \ \limsup_n (\eta_n, y_n) \le (\eta, y) \ \Rightarrow \ \eta \in \partial\phi(y). \tag{3.3.1}$$

Indeed, one has that $\eta_n - \kappa y_n \in \partial\psi(y_n)$, $\eta_n - \kappa y_n \rightharpoonup \eta - \kappa y$, and

$$\limsup_n (\eta_n - \kappa y_n, y_n) \le \limsup_n (\eta_n, y_n) - \liminf_n \kappa \|y_n\|^2 \le (\eta - \kappa y, y)$$

so that using [15, Prop 2.5, p. 27] one finds $\eta - \kappa y \in \partial\psi(y)$, which entails $\eta \in \partial\phi(y)$. The identification (3.3.1) equivalently holds in its integrated form for sequences y_n and η_n weakly converging in $L^2(0,T;H)$, namely.

$$\begin{aligned} & y_n \rightharpoonup y, \ \eta_n \rightharpoonup \eta \text{ in } L^2(0,T;H), \ \eta_n \in \partial\phi(y_n) \text{ a.e.}, \\ & \limsup_n \int_0^T (\eta_n, y_n)\,\mathrm{d}t \le \int_0^T (\eta, y)\,\mathrm{d}t & (3.3.2) \\ & \Rightarrow \ \eta \in \partial\phi(y) \text{ a.e.} & (3.3.3) \end{aligned}$$

One can also readily prove the following generalizaton to the κ-convex case of the classical chain rule [16, Lemma 3.3, p. 73]

$$\begin{aligned} & y \in H^1(0,T;H), \ \eta \in L^2(0,T;H), \ \eta \in \partial\phi(u) \text{ a.e. in } (0,T) & (3.3.4) \\ & \Rightarrow \ \phi \circ y \text{ is absolutely continuous on } [0,T] \text{ and} & (3.3.5) \\ & \frac{\mathrm{d}}{\mathrm{d}t}\phi \circ y = (\eta, y) \text{ a.e. in } (0,T). & (3.3.6) \end{aligned}$$

In the following, we use the symbol c to indicate a generic positive constant, possibly depending on T, U, y^0 but independent on ε and λ and possibly varying from line to line.

We are now ready to present the lemmas.

Lemma 3.3.1 (Value of M_ε^u)**.** *For all $u \in U$, recalling that $y_\varepsilon^u = \arg\min W_\varepsilon(\cdot, u)$ and $M_\varepsilon^u = W_\varepsilon(y_\varepsilon^u, u)$ we have*

$$\begin{aligned} M_\varepsilon^u = &-\frac{\varepsilon^2}{2}\|\dot{y}_\varepsilon^u(0)\|^2 - \varepsilon e^{-T/\varepsilon}\phi(y_\varepsilon^u(T)) + \varepsilon\phi(y^0) \\ &+ \varepsilon\int_0^T e^{-t/\varepsilon}(u, \dot{y}_\varepsilon^u)\,\mathrm{d}t - \int_0^T e^{-t/\varepsilon}(u, y_\varepsilon^u)\,\mathrm{d}t. \end{aligned}$$

Proof. The trajectory y_ε^u solves the Euler-Lagrange problem (3.1.4)-(3.1.6). By taking the scalar product with $e^{-t/\varepsilon}\dot{y}_\varepsilon^u \in H^1(0,T;H)$ in equation (3.1.4), integrating in time, and using the chain rule (3.3.4), we get

$$\int_0^T e^{-t/\varepsilon}\left(-\frac{\varepsilon}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\dot{y}_\varepsilon^u\|^2 + \|\dot{y}_\varepsilon^u\|^2 + \frac{\mathrm{d}}{\mathrm{d}t}\phi(y_\varepsilon^u) - (u, \dot{y}_\varepsilon^u)\right)\mathrm{d}t = 0.$$

By integrating by parts the first and the third term, one obtains

$$\begin{aligned} &\left[e^{-t/\varepsilon}\left(-\frac{\varepsilon}{2}\|\dot{y}_\varepsilon^u\|^2 + \phi(y_\varepsilon^u)\right)\right]_0^T - \int_0^T e^{-t/\varepsilon}(u, \dot{y}_\varepsilon^u)\,\mathrm{d}t \\ &\quad + \frac{1}{\varepsilon}\int_0^T e^{-t/\varepsilon}(u, y_\varepsilon^u)\,\mathrm{d}t + \frac{1}{\varepsilon}M_\varepsilon^u = 0, \end{aligned}$$

where we used the definition (3.1.3) of the WED functional. The thesis follows from conditions (3.1.5)-(3.1.6). □

Lemma 3.3.2 (Continuity of the map $u \mapsto y_\varepsilon^u$)**.** *Let $(u_k)_k \subset U$ be such that $u_k \to u$ in $L^2(0,T;H)$ and let $\eta_\varepsilon^{u_k} = \varepsilon\ddot{y}_\varepsilon^{u_k} - \dot{y}_\varepsilon^{u_k} + u_k$. Up to not relabeled subsequences, one has*

$$\begin{aligned} y_\varepsilon^{u_k} &\rightharpoonup y_\varepsilon^u \quad \text{in} \quad H^2(0,T;H), \\ \eta_\varepsilon^{u_k} &\rightharpoonup \eta_\varepsilon^u \quad \text{in} \quad L^2(0,T;H) \end{aligned}$$

where $\eta_\varepsilon^u = \varepsilon\ddot{y}_\varepsilon^u - \dot{y}_\varepsilon^u + u$.

Proof. From the uniform estimate (3.2.1) we may extract not relabeled subsequences such that $y_\varepsilon^{u_k} \rightharpoonup y$ in $H^2(0,T;H)$ and $\eta_\varepsilon^{u_k} \rightharpoonup \eta$ in $L^2(0,T;H)$ and get

$$-\varepsilon\ddot{y}(t) + \dot{y}(t) + \eta(t) = u(t) \quad \text{in } H \text{ a.e. } t \in (0,T).$$

As $y_\varepsilon^{u_k}(t) \rightharpoonup y(t)$ and $\dot{y}_\varepsilon^{u_k}(t) \rightharpoonup \dot{y}(t)$ for all $t \in [0,T]$ we have that $y(0) = y^0$ and $\varepsilon\dot{y}(T) = 0$. In order to conclude the proof it hence suffices to check that $\eta \in \partial\phi(y)$ a.e. Take the scalar lim sup of the integral over $(0,T)$ of the scalar product between $\eta_\varepsilon^{u_k}$ and $y_\varepsilon^{u_k}$. Using equation (3.1.4) at level k, we obtain

$$\begin{aligned}
&\limsup_{k\to\infty}\int_0^T (\eta_\varepsilon^{u_k}, y_\varepsilon^{u_k})\,\mathrm{d}t\\
&\quad = \limsup_{k\to\infty}\left(\varepsilon\int_0^T (\ddot{y}_\varepsilon^{u_k}, y_\varepsilon^{u_k})\,\mathrm{d}t - \int_0^T (\dot{y}_\varepsilon^{u_k}, y_\varepsilon^{u_k})\,\mathrm{d}t + \int_0^T (u_k, y_\varepsilon^{u_k})\,\mathrm{d}t\right)\\
&\quad = \limsup_{k\to\infty}\left(-\varepsilon(\dot{y}_\varepsilon^{u_k}(0), y^0) - \varepsilon\int_0^T \|\dot{y}_\varepsilon^{u_k}\|^2\mathrm{d}t\right.\\
&\qquad \left. - \frac{1}{2}\|y_\varepsilon^{u_k}(T)\|^2 + \frac{1}{2}\|y^0\|^2 + \int_0^T (u_k, y_\varepsilon^{u_k})\,\mathrm{d}t\right)
\end{aligned}$$

where we also used the conditions (3.1.5)-(3.1.6). Owing to the above convergences we infer

$$\begin{aligned}
&\limsup_{k\to\infty}\int_0^T (\eta_\varepsilon^{u_k}, y_\varepsilon^{u_k})\,\mathrm{d}t\\
&\quad \le -\varepsilon(\dot{y}(0), y^0) - \varepsilon\int_0^T \|\dot{y}\|^2\mathrm{d}t - \frac{1}{2}\|y(T)\|^2 + \frac{1}{2}\|y^0\|^2 + \int_0^T (u, y)\,\mathrm{d}t\\
&\quad = \int_0^T (\eta, y)\,\mathrm{d}t.
\end{aligned}$$

This implies that $\eta \in \partial\phi(y)$ via (3.3.3), see [15, Prop. 2.5, p. 27]. Hence, y in the unique solution to the Euler-Lagrange problem (3.1.4)-(3.1.6) and $y = y_\varepsilon^u$, $\eta = \eta_\varepsilon^u$ by Proposition 3.2.1. $\square$

Lemma 3.3.3 (Coercivity of $P_{\varepsilon\lambda}$). *We have*

$$\varepsilon^3 e^{-T/\varepsilon}\|y - y_\varepsilon^u\|^2_{H^1(0,T;H)} \le W_\varepsilon(y,u) - M_\varepsilon^u \quad \forall (y,u) \in K(y^0)\times U. \tag{3.3.7}$$

In particular, the sublevels of $\lambda\varepsilon^{-3}e^{T/\varepsilon}P_{\varepsilon\lambda}(\cdot,u)$ are bounded in $H^1(0,T;H)$ independently of $u \in U$.

Proof. Let us start by rewriting $W_\varepsilon(y,u) - M_\varepsilon^u$ as

$$\begin{aligned}
&W_\varepsilon(y,u) - M_\varepsilon^u\\
&\quad = \int_0^T e^{-t/\varepsilon}\left(\frac{\varepsilon}{2}\|\dot{y}\|^2 + \frac{\kappa}{2}\|y\|^2\right)\mathrm{d}t + \left(\int_0^T e^{-t/\varepsilon}\Big(\psi(y) - (u,y)\Big)\,\mathrm{d}t - M_\varepsilon^u\right)\\
&\quad =: Q_\varepsilon(y) + R_\varepsilon(y,u),
\end{aligned}$$

where the functional $R_\varepsilon(\cdot,u)$ is convex and the quadratic functional Q_ε can be written in terms of $v = e^{-t/(2\varepsilon)}y$ as

$$\begin{aligned}
Q_\varepsilon(y) &= \int_0^T \left(\frac{\varepsilon}{2}\|\dot{v}\|^2 + \frac{1+4\varepsilon\kappa}{8\varepsilon}\|v\|^2\right)\mathrm{d}t + \frac{1}{4}\|v(T)\|^2 - \frac{1}{4}\|v(0)\|^2\\
&=: V_\varepsilon(v) + \frac{1}{4}\|v(T)\|^2 - \frac{1}{4}\|v(0)\|^2.
\end{aligned}$$

We now use the fact that V_ε is quadratic. For all v_1, $v_2 \in H^1(0,T;H)$ and all $r \in (0,1)$ one computes

$$\begin{aligned}
&V_\varepsilon(rv_1+(1-r)v_2)\\
&=\frac{\varepsilon}{2}\int_0^T \|r\dot v_1+(1-r)\dot v_2\|^2\,\mathrm{d}t+\frac{1+4\varepsilon\kappa}{8\varepsilon}\int_0^T\|rv_1+(1-r)v_2\|^2\,\mathrm{d}t\\
&=\frac{\varepsilon}{2}\int_0^T\big(r^2\ \|\dot v_1\|^2+(1-r)^2\|\dot v_2\|^2+2r(1-r)(\dot v_1,\dot v_2)\big)\ \mathrm{d}t\\
&\quad+\frac{1+4\varepsilon\kappa}{8\varepsilon}\int_0^T\big(r^2\ \|v_1\|^2+(1-r)^2\|v_2\|^2+2r(1-r)(v_1,v_2)\big)\ \mathrm{d}t\\
&=\frac{\varepsilon}{2}\int_0^T\big(r\|\dot v_1\|^2+(1-r)\|\dot v_2\|^2-r(1-r)\|\dot v_1-\dot v_2\|^2\big)\ \mathrm{d}t\\
&\quad+\frac{1+4\varepsilon\kappa}{8\varepsilon}\int_0^T\big(r\|v_1\|^2+(1-r)\|v_2\|^2-r(1-r)\|v_1-v_2\|^2\big)\ \mathrm{d}t\\
&=rV_\varepsilon(v_1)+(1-r)V_\varepsilon(v_2)-r(1-r)\int_0^T\left(\frac{\varepsilon}{2}\|\dot v_1-\dot v_2\|^2+\frac{1+4\varepsilon\kappa}{8\varepsilon}\|v_1-v_2\|^2\right)\mathrm{d}t,
\end{aligned}$$

which, for $\varepsilon<\kappa$ small enough, implies the following inequality

$$V_\varepsilon(rv_1+(1-r)v_2)\le rV_\varepsilon(v_1)+(1-r)V_\varepsilon(v_2)-\varepsilon\frac{r(1-r)}{2}\|v_1-v_2\|^2_{H^1(0,T;H)}. \tag{3.3.8}$$

This in particular entails that

$$\begin{aligned}
&W_\varepsilon(ry+(1-r)y^u_\varepsilon,u)-M^u_\varepsilon\\
&\quad=V_\varepsilon(rv+(1-r)v^u_\varepsilon)+\frac14\|rv(T)+(1-r)v^u_\varepsilon(T)\|^2\\
&\qquad-\frac14\|rv(0)+(1-r)v^u_\varepsilon(0)\|^2+R_\varepsilon(ry+(1-r)y^u_\varepsilon,u)\\
&\quad\le rV_\varepsilon(v)+(1-r)V_\varepsilon(v^u_\varepsilon)-\varepsilon\frac{r(1-r)}{2}\|v-v^u_\varepsilon\|^2_{H^1(0,T;H)}\\
&\qquad+\frac{r}{4}\|v(T)\|^2+\frac{1-r}{4}\|v^u_\varepsilon(T)\|^2-\frac14\|y^0\|^2+rR_\varepsilon(y,u)+(1-r)R_\varepsilon(y^u_\varepsilon,u)\\
&\quad=r(W_\varepsilon(y,u)-M^u_\varepsilon)+(1-r)(W_\varepsilon(y^u_\varepsilon,u)-M^u_\varepsilon)-\varepsilon\frac{r(1-r)}{2}\|v-v^u_\varepsilon\|^2_{H^1(0,T;H)}\\
&\quad=r(W_\varepsilon(y,u)-M^u_\varepsilon)-\varepsilon\frac{r(1-r)}{2}\|v-v^u_\varepsilon\|^2_{H^1(0,T;H)},
\end{aligned}\tag{3.3.9}$$

where we used the convexity of the maps $v\mapsto R_\varepsilon(v,u)$ and $v\mapsto\|v\|^2$, inequality (3.3.8), the fact that $v(0)=v^u_\varepsilon(0)=y^0$, and $W_\varepsilon(y^u_\varepsilon,u)=M^u_\varepsilon$. By observing that

$$W_\varepsilon(ry+(1-r)y^u_\varepsilon,u)-M^u_\varepsilon\ge\min_y W_\varepsilon(y,u)-M^u_\varepsilon=W_\varepsilon(y^u_\varepsilon,u)-M^u_\varepsilon=0,$$

inequality (3.3.9) implies that

$$r(W_\varepsilon(y,u)-M^u_\varepsilon)\ge\varepsilon\frac{r(1-r)}{2}\|v-v^u_\varepsilon\|^2_{H^1(0,T;H)}.$$

Assume now that $r>0$, divide by r, and take $r\to0$ to get

$$W_\varepsilon(y,u)-M^u_\varepsilon\ge\frac{\varepsilon}{2}\|v-v^u_\varepsilon\|^2_{H^1(0,T;H)}.$$

Eventually, putting $v = e^{-t/(2\varepsilon)}y$ and $v^u_\varepsilon = e^{-t/(2\varepsilon)}y^u_\varepsilon$ we obtain

$$
\begin{aligned}
&W_\varepsilon(y,u) - M^u_\varepsilon \geq \frac{\varepsilon}{2}\|v - v^u_\varepsilon\|^2_{H^1(0,T;H)} \\
&\geq \frac{\varepsilon}{2}\int_0^T e^{-t/\varepsilon}\bigg(\|y-y^u_\varepsilon\|^2 + \|\dot y - \dot y^u_\varepsilon\|^2 + \frac{\|y-y^u_\varepsilon\|^2}{4\varepsilon^2} - \frac{1}{\varepsilon}(y-y^u_\varepsilon, \dot y - \dot y^u_\varepsilon)\bigg)\,\mathrm{d}t \\
&\geq \frac{\varepsilon}{2}\int_0^T e^{-t/\varepsilon}\bigg(\|y-y^u_\varepsilon\|^2 + \|\dot y - \dot y^u_\varepsilon\|^2 + \frac{\|y-y^u_\varepsilon\|^2}{4\varepsilon^2} - \frac{1}{1+2\varepsilon^2}\|\dot y - \dot y^u_\varepsilon\|^2 \\
&\quad - \frac{1+2\varepsilon^2}{4\varepsilon^2}\|y-y^u_\varepsilon\|^2\bigg)\,\mathrm{d}t \\
&\geq \varepsilon^3 e^{-T/\varepsilon}\|y-y^u_\varepsilon\|^2_{H^1(0,T;H)}
\end{aligned}
$$

for $\varepsilon \in (0,1]$, which proves (3.3.7). In particular, we have

$$
\begin{aligned}
\|y\|^2_{H^1(0,T;H)} &\leq 2\|y-y^u_\varepsilon\|^2_{H^1(0,T;H)} + 2\|y^u_\varepsilon\|^2_{H^1(0,T;H)} \\
&\leq c + 2\lambda\varepsilon^{-3}e^{T/\varepsilon}P_{\varepsilon\lambda}(y,u),
\end{aligned}
$$

which concludes the proof. □

Lemma 3.3.4. *Let $u_\varepsilon \in U$ with $u_\varepsilon \to u$ in $L^2(0,T;H)$. Then, up to subsequences $y^{u_\varepsilon}_\varepsilon \rightharpoonup y$ in $H^1(0,T;H)$ and $y^{u_\varepsilon}_\varepsilon \to y$ in $H^\sigma(0,T;H)$ for any $\sigma \in (0,1)$ with $y = S(u)$.*

Proof. Letting $\eta_\varepsilon = u_\varepsilon + \varepsilon\ddot y^{u_\varepsilon}_\varepsilon - \dot y^{u_\varepsilon}_\varepsilon$, the uniform bound (3.2.1) and up to not relabeled subsequences we have that

$$
\begin{aligned}
\varepsilon\ddot y^{u_\varepsilon}_\varepsilon &\rightharpoonup 0 \quad \text{in } L^2(0,T;H), && (3.3.10)\\
y^{u_\varepsilon}_\varepsilon &\rightharpoonup y \quad \text{in } H^1(0,T;H), && (3.3.11)\\
\eta^{u_\varepsilon}_\varepsilon &\rightharpoonup \tilde\eta \quad \text{in } L^2(0,T;H), && (3.3.12)
\end{aligned}
$$

for some functions y and $\tilde\eta$ with $\dot y + \tilde\eta = u$.

We shall now check that $y^{u_\varepsilon}_\varepsilon$ is indeed a Cauchy sequence in $C^0([0,T];H)$. Let $y^{u_\varepsilon}_\varepsilon$ and $y^{u_\mu}_\mu$ be solutions of the problem (3.1.4)-(3.1.6) at level ε and μ, respectively. Consider the difference of the Euler-Lagrange equation (3.1.4) at level ε and the one at level μ and take its scalar product with the function $w := y^{u_\varepsilon}_\varepsilon - y^{u_\mu}_\mu$. By letting $\eta_\mu = u_\mu + \mu\ddot y^{u_\mu}_\mu - \dot y^{u_\mu}_\mu$ and integrating in time over $(0,t)$ we get

$$
\begin{aligned}
&-\int_0^t (\varepsilon\ddot y^{u_\varepsilon}_\varepsilon - \mu\ddot y^{u_\mu}_\mu, w)\,\mathrm{d}t + \int_0^t \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\|w\|^2\,\mathrm{d}t + \int_0^t (\eta_\varepsilon - \eta_\mu, y^{u_\varepsilon}_\varepsilon - y^{u_\mu}_\mu)\,\mathrm{d}t \\
&\quad = \int_0^t (u_\varepsilon - u_\mu, w)\,\mathrm{d}t.
\end{aligned}
$$

The κ-convexity of ϕ, the fact that $w(0) = 0$, and an integration by parts give

$$
\begin{aligned}
&\varepsilon\int_0^t \|\dot w\|^2\,\mathrm{d}t + \kappa\int_0^t \|w\|^2\,\mathrm{d}t + \frac{1}{2}\|w(t)\|^2 \\
&\quad \leq (\varepsilon\dot y^{u_\varepsilon}_\varepsilon(t) - \mu\dot y^{u_\mu}_\mu(t), w(t)) - (\varepsilon-\mu)\int_0^t (\dot y^{u_\mu}_\mu, \dot w)\,\mathrm{d}t + \int_0^t (u_\varepsilon - u_\mu, w)\,\mathrm{d}t.
\end{aligned}
$$

Young's inequality allows then to deduce

$$
\begin{aligned}
&\varepsilon \int_0^t \|\dot{w}\|^2 \,\mathrm{d}t + \kappa \int_0^t \|w\|^2 \,\mathrm{d}t + \frac{1}{4}\|w(t)\|^2 \\
&\quad \le 2\varepsilon^2 \|\dot{y}_\varepsilon^{u_\varepsilon}\|^2_{C^0([0,T];H)} + 2\mu^2 \|\dot{y}_\mu^{u_\mu}\|^2_{C^0([0,T];H)} \\
&\quad + (\varepsilon+\mu)\|\dot{y}_\mu^{u_\mu}\|_{L^2(0,T;H)}\|\dot{w}\|_{L^2(0,T;H)} + \|u_\varepsilon - u_\mu\|_{L^2(0,T;H)}\|w\|_{L^2(0,T;H)}. \qquad (3.3.13)
\end{aligned}
$$

The Gagliardo-Nirenberg inequality [16, Comments (*iii*), p. 233], [46, Theorem 1, p. 734], and bound (3.2.1) give

$$
\begin{aligned}
&\|\dot{y}_\varepsilon^{u_\varepsilon}\|_{C^0([0,T];H)} \le c\|\ddot{y}_\varepsilon^{u_\varepsilon}\|^{1/2}_{L^2(0,T;H)}\|\dot{y}_\varepsilon^{u_\varepsilon}\|^{1/2}_{L^2(0,T;H)} + c\|\dot{y}_\varepsilon^{u_\varepsilon}\|_{L^2(0,T;H)} \le c\varepsilon^{-1/2}, \\
&\|\dot{y}_\mu^{u_\mu}\|_{C^0([0,T];H)} \le c\|\ddot{y}_\mu^{u_\mu}\|^{1/2}_{L^2(0,T;H)}\|\dot{y}_\mu^{u_\mu}\|^{1/2}_{L^2(0,T;H)} + c\|\dot{y}_\mu^{u_\mu}\|_{L^2(0,T;H)} \le c\mu^{-1/2}.
\end{aligned}
$$

We can hence use (3.3.13) and the fact that $\|\dot{y}_\mu^{u_\mu}\|_{L^2(0,T;H)} \le c$ and $\|\dot{w}\|_{L^2(0,T;H)} \le c$ to obtain

$$
\varepsilon \int_0^t \|\dot{w}\|^2 \,\mathrm{d}t + \frac{1}{4}\|w(t)\|^2 \le c(\varepsilon+\mu) + c\|u_\varepsilon - u_\mu\|_{L^2(0,T;H)} - \kappa \int_0^t \|w\|^2 \,\mathrm{d}t.
$$

This entails that

$$
\|y_\varepsilon - y_\mu\|_{C^0([0,T];H)} \le c\Big(\varepsilon + \mu + \|u_\varepsilon - u_\mu\|_{L^2(0,T;H)}\Big)^{1/2} \qquad (3.3.14)
$$

where we also applied the Gronwall Lemma in case $\kappa < 0$. The strong convergence

$$
y_\varepsilon \to y \ \text{ in } C^0([0,T];H), \qquad (3.3.15)
$$

follows.

To identify the limit function $\tilde{\eta}$, we use convergences (3.3.12) and (3.3.15) to get

$$
\limsup_{\varepsilon\to 0} \int_0^T (\eta_\varepsilon, y_\varepsilon)\,\mathrm{d}t = \int_0^T (\tilde{\eta}, y)\,\mathrm{d}t.
$$

This implies that $\tilde{\eta} \in \partial\phi(u)$ a.e. by (3.3.3). The initial condition (3.1.2) follows from (3.3.15), so that $y = S(u)$. As this limit is unique, the whole sequence $(y_\varepsilon^{u_\varepsilon})_\varepsilon$ converges and no extraction of a subsequence is actually necessary.

We now make use of the interpolation space $(C^0([0,T];H), H^1(0,T;H))_{\sigma,1}$ for $\sigma \in (0,1)$, whose elements are those functions $w \in C^0([0,T];H)$ such that

$$
\|w\|_{(C^0([0,T];H),H^1(0,T;H))_{\sigma,1}} := \int_0^\infty r^{-\sigma-1}K(r,w)\,\mathrm{d}r < \infty,
$$

where $K : (0,\infty) \times C^0([0,T];H) \to [0,\infty)$ is defined as

$$
\begin{aligned}
K(r,w) := \inf\Big\{&\|w_0\|_{C^0([0,T];H)} + r\|w_1\|_{H^1(0,T;H)} : \\
&w_0 \in C^0([0,T];H),\ w_1 \in H^1(0,T;H),\ w = w_0 + w_1\Big\}.
\end{aligned}
$$

see the classical reference [12]. We will also use that there exists $c > 0$ such that

$$
\|w\|_{(C^0([0,T];H),H^1(0,T;H))_{\sigma,1}} \le c\|w\|^{1-\sigma}_{C^0([0,T];H)}\|w\|^{\sigma}_{H^1(0,T;H)} \quad \forall w \in H^1(0,T;H),
$$

see [18, Lemma 2.1.i]. As $(C^0([0,T];H), H^1(0,T;H))_{\sigma,1} \subset H^\sigma(0,T;H)$ [12, Theorem 6.2.4, p. 142], the uniform bound of $y_\varepsilon^{u_\varepsilon}$ in $H^1(0,T;H)$ and estimate (3.3.14) imply that $y_\varepsilon^{u_\varepsilon} \to y$ in $H^\sigma(0,T;H)$. Indeed, one has

$$\begin{aligned}\|y_\varepsilon^{u_\varepsilon} - y_\mu^{u_\mu}\|_{H^\sigma(0,T;H)} &\le c\|y_\varepsilon^{u_\varepsilon} - y_\mu^{u_\mu}\|_{(C^0([0,T];H),H^1(0,T;H))_{\sigma,1}} \\ &\le c\|y_\varepsilon^{u_\varepsilon} - y_\mu^{u_\mu}\|^{1-\sigma}_{C^0([0,T];H)}\|y_\varepsilon^{u_\varepsilon} - y_\mu^{u_\mu}\|^\sigma_{H^1(0,T;H)} \\ &\le c\left(\varepsilon + \mu + \|u_\varepsilon - u_\mu\|_{L^2(0,T;H)}\right)^{(1-\sigma)/2} \to 0.\end{aligned}$$

□

Lemma 3.3.5. *Let $(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \in H^1(0,T;H) \times U$ be such that $P_{\varepsilon\lambda}(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \le c$, with c independent of λ. As $\lambda \to 0$, up to not relabeled subsequences we have that $(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \xrightarrow{\tau} (y_\varepsilon^{u_\varepsilon}, u_\varepsilon)$.*

Proof. From the compact injection $U \subset\subset L^2(0,T;H)$ we have $u_{\varepsilon\lambda} \to u_\varepsilon$ in $L^2(0,T;H)$ along some not relabeled subsequence. Lemma 3.3.3 implies that $(y_{\varepsilon\lambda})_\lambda$ is bounded in $H^1(0,T;H)$. Hence, up to not relabeled subsequences, $y_{\varepsilon\lambda} \rightharpoonup y_\varepsilon$ in $H^1(0,T;H)$.

As $P_{\varepsilon\lambda}(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) < c$ and J is nonnegative we have

$$0 \le W_\varepsilon(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) - m_\varepsilon(u_{\varepsilon\lambda}) \le c\lambda.$$

Recall from Section 3.4.3 that $(y,u) \mapsto W_\varepsilon(y,u) - M_\varepsilon^u$ is lower semicontinuous with respect to the topology τ. Hence, by passing to the lim inf as $\lambda \to 0$, we get

$$W_\varepsilon(y_\varepsilon, u_\varepsilon) = M_\varepsilon^{u_\varepsilon},$$

proving indeed that $y_\varepsilon = y_\varepsilon^{u_\varepsilon}$, which concludes the argument. □

Lemma 3.3.6. *Let $\lambda = \lambda_\varepsilon$, with $\limsup_{\varepsilon\to 0} \lambda_\varepsilon \varepsilon^{-3} e^{T/\varepsilon} = 0$ and let $(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \in H^1(0,T;H) \times U$ be such that $P_{\varepsilon\lambda}(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \le c$, with c independent of ε. Then, as $\varepsilon \to 0$, up to not relabeled subsequences $(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \xrightarrow{\tau} (y,u)$ with $y = S(u)$.*

Proof. Let $(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \in H^1(0,T;H) \times U$ be such that $P_{\varepsilon\lambda}(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \le c$. As the injection $U \subset\subset L^2(0,T;H)$ is compact, we get $u_{\varepsilon\lambda} \to u$ in $L^2(0,T;H)$ along some not relabeled subsequence. Using Lemma 3.3.3, we have

$$\limsup_{\varepsilon\to 0} \|y_{\varepsilon\lambda} - y_\varepsilon^{u_{\varepsilon\lambda}}\|^2_{H^1(0,T;H)} \le \limsup_{\varepsilon\to 0} \lambda_\varepsilon \varepsilon^{-3} e^{T/\varepsilon} P_{\varepsilon\lambda}(y_{\varepsilon\lambda}, u_{\varepsilon\lambda}) = 0. \tag{3.3.16}$$

On the other hand, Lemma 3.3.4 implies that $y_\varepsilon^{u_{\varepsilon\lambda}} \rightharpoonup y$ in $H^1(0,T;H)$ and $y_\varepsilon^{u_{\varepsilon\lambda}} \to y$ in $H^\sigma(0,T;H)$ for all $\sigma \in (0,1)$, with $y = S(u)$. This implies that $y_{\varepsilon\lambda} \rightharpoonup y$ in $H^1(0,T;H)$.

The thesis follows then from (3.3.16) and Lemma (3.3.4) by simply observing that

$$\|y_{\varepsilon\lambda} - y\|_{H^\sigma(0,T;H)} \le c\|y_{\varepsilon\lambda} - y_\varepsilon^{u_{\varepsilon\lambda}}\|_{H^1(0,T;H)} + \|y_\varepsilon^{u_{\varepsilon\lambda}} - y\|_{H^\sigma(0,T;H)} \to 0.$$ □

3.4 Existence: proof of Theorem 3.2.2.i

In this section, we existence for Problems $(\mathcal{P})$, $(\mathcal{P}_\varepsilon)$, and $(\mathcal{P}_{\varepsilon\lambda})$, namely Theorem 3.2.2.i.

3.4.1 Well-posedness for Problem $\mathcal{P}$

Taking any $\tilde{u} \in U$ letting $\tilde{y} = S(\tilde{u})$ one has that $0 \leq P(\tilde{y}, \tilde{u}) = J(\tilde{y}, \tilde{u}) < \infty$. In particular $\inf_{H^1(0,T;H)\times U} P \in [0, \infty)$.

Let $(y_k, u_k)_k \subset H^1(0,T;H)\times U$ be an infimizing sequence for P, that is $P(y_k, u_k) \to \inf_{H^1(0,T;H)\times U} P$. The strong convergence

$$u_k \to u \qquad \text{in } L^2(0,T;H) \tag{3.4.1}$$

follows from the compact injection $U \subset\subset L^2(0,T;H)$, up to some not relabeled subsequence. For all $k > 0$, $y_k \in S(u_k)$ solves the gradient flow

$$\dot{y}_k(t) + \eta_k(t) = u_k(t) \quad \text{in } H, \quad \text{a.e. } t \in (0,T), \tag{3.4.2}$$

$$\eta_k(t) \in \partial\phi(y_k(t)) \quad \text{in } H, \quad \text{a.e. } t \in (0,T), \tag{3.4.3}$$

$$y_k(0) = y^0. \tag{3.4.4}$$

As each term in equation (3.4.2) is in $L^2(0,T;H)$, we use the chain rule (3.3.4) in order to compute

$$\begin{aligned}
&\int_0^T \|\dot{y}_k\|^2 \,\mathrm{d}t + \int_0^T \|\eta_k\|^2 \,\mathrm{d}t = \int_0^T \|\dot{y}_k + \eta_k\|^2 \,\mathrm{d}t - 2\int_0^T (\dot{y}_k, \eta_k) \,\mathrm{d}t \\
&= \int_0^T \|u_k\|^2 \,\mathrm{d}t - 2\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\phi(y_k) \,\mathrm{d}t = \int_0^T \|u_k\|^2 \,\mathrm{d}t - 2\phi(y_k(T)) + 2\phi(y^0) \\
&\leq \int_0^T \|u_k\|^2 \,\mathrm{d}t - 2((\partial\phi(y^0))^\circ, y_k(T) - y^0) \\
&\leq \int_0^T \|u_k\|^2 \,\mathrm{d}t + \frac{1}{2}\int_0^T \|\dot{y}_k\|^2 \mathrm{d}t + 2T\|(\partial\phi(y^0))^\circ\|^2
\end{aligned}$$

where we used the equation (3.4.2) as well as the fact that $y^0 \in D(\partial\phi)$. We have obtained $\|\dot{y}_k\|_{L^2(0,T;H)} + \|\eta_k\|_{L^2(0,T;H)} \leq c$, which yields, up to not relabeled subsequences, that

$$y_k \rightharpoonup y \qquad \text{in } H^1(0,T;H), \tag{3.4.5}$$

$$\eta_k \rightharpoonup \eta \qquad \text{in } L^2(0,T;H), \tag{3.4.6}$$

for some limit functions y, η with $\dot{y} + \eta = u$ a.e. As $y_k(0) \rightharpoonup y(0)$, we have that $y(0) = y^0$. Moreover,

$$\begin{aligned}
&\limsup_{k\to\infty} \int_0^T (\eta_k, y_k) \,\mathrm{d}t = \limsup_{k\to\infty} \int_0^T (u_k - \dot{y}_k, y_k) \,\mathrm{d}t \\
&= \int_0^T (u, y) \,\mathrm{d}t - \liminf_{k\to\infty} \frac{1}{2}\|y_k(T)\|^2 + \frac{1}{2}\|y^0\|^2 \\
&\leq \int_0^T (u - \dot{y}, y) \,\mathrm{d}t = \int_0^T (\eta, y) \,\mathrm{d}t.
\end{aligned}$$

Again (3.3.3) entails that $\eta \in \partial\phi(y)$ in $L^2(0,T;H)$. The limit function y hence solves then the gradient flow problem, namely, $y \in S(u)$. Eventually, convergences (3.4.1) and (3.4.5) together with Assumption (A3) imply

$$\inf_{H^1(0,T;H)\times U} P \leq J(y,u) \leq \liminf_{k\to\infty} J(y_k, u_k) = \inf_{H^1(0,T;H)\times U} P,$$

so that (y, u) actually solves $(\mathcal{P})$.

3.4.2 Well-posedness for Problem $\mathcal{P}_\varepsilon$

Choosing an arbitrary $\tilde{u} \in U$ and letting $\tilde{y} \in S(\tilde{u})$ one has that $0 \leq P_\varepsilon(\tilde{y}\,\tilde{u}) = J(\tilde{y}, \tilde{u}) < \infty$. In particular, $\inf_{H^1(0,T;H)\times U} P_\varepsilon \in [0, \infty)$.

Let $(y_k, u_k)_k \subset H^1(0,T;H)\times U$ be an infimizing sequence for P_ε, namely, $P_\varepsilon(y_k, u_k) \to \inf_{H^1(0,T;H)\times U} P_\varepsilon$. We have, $P_\varepsilon(y_k, u_k) = J(y_k, u_k)$ and $y_k \in \arg\min W_\varepsilon(\cdot, u_k)$, namely, $y_k = y_\varepsilon^{u_k}$. The strong convergence $u_k \to u$ in $L^2(0,T;H)$ follows by the compact injection $U \subset\subset L^2(0,T;H)$, up to a not relabeled subsequence. For each $k > 0$, there exists $\eta_k \in L^2(0,T;H)$ such that

$$\begin{aligned} -\varepsilon\ddot{y}_k(t) + \dot{y}_k(t) + \eta_k(t) &= u_k(t) \quad \text{for a.e. } t \in (0,T), \\ \eta_k(t) &\in \partial\phi(y_k(t)) \quad \text{for a.e. } t \in (0,T), \\ y_k(0) &= y^0, \\ \varepsilon\dot{y}_k(T) &= 0. \end{aligned}$$

Lemma 3.3.2 ensures that, up to not relabeled subsequences, $y_k \rightharpoonup y$ in $H^2(0,T;H)$, $\eta_k \rightharpoonup \eta$ in $L^2(0,T;H)$ where $\eta \in \partial\phi(y)$ a.e., and y solves the Euler-Lagrange problem (3.1.4)-(3.1.6) corresponding to u. In particular, $y = y_\varepsilon^u$. Together with Assumption (A3), these convergences imply

$$\inf_{H^1(0,T;H)\times U} P_\varepsilon \leq J(y,u) \leq \liminf_{k\to\infty} J(y_k, u_k) = \inf_{H^1(0,T;H)\times U} P_\varepsilon$$

which proves that (y, u) actually minimizes P_ε.

3.4.3 Well-posedness for Problem $\mathcal{P}_{\varepsilon\lambda}$

Given any $\tilde{u} \in U$ we have that $0 \leq P_{\varepsilon\lambda}(y_\varepsilon^{\tilde{u}}, \tilde{u}) = J(y_\varepsilon^{\tilde{u}}, \tilde{u}) < \infty$. This proves that $\inf_{H^1(0,T;H)\times U} P_{\varepsilon\lambda} \in [0, \infty)$.

Let $(y_k, u_k)_k \subset H^1(0,T;H)\times U$ be an infimizing sequence for $P_{\varepsilon\lambda}$, namely, such that $P_{\varepsilon\lambda}(y_k, u_k) \to \inf_{H^1(0,T;H)\times U} P_{\varepsilon\lambda}$. The strong convergence $u_k \to u$ in $L^2(0,T;H)$ follows from the compact injection $U \subset\subset L^2(0,T;H)$, up to a not relabeled subsequence. Using Lemma 3.3.3, we have that $\|y_k\|_{H^1(0,T;H)} \leq c$. Then, up to not relabeled subsequences, the following convergence hold

$$y_k \rightharpoonup y \qquad \text{in } H^1(0,T;H). \tag{3.4.7}$$

Since $H^1(0,T;H) \subset C^0([0,T];H)$, we have that $y_k(t) \rightharpoonup y(t)$ for all $t \in [0,T]$. In particular, the initial condition $y(0) = y^0$ is satisfied.

Lemma 3.3.2 implies that

$$y_\varepsilon^{u_k} \rightharpoonup y_\varepsilon^u \qquad \text{in } H^2(0,T;H), \tag{3.4.8}$$

and y_ε^u satisfies the Euler-Lagrange problem (3.1.4)-(3.1.6) corresponding to u. By Assumption (A3), we have that $J(y,u) \leq \liminf_n J(y_n, u_n)$.

We hence reduce ourselves to check that $(y,u) \mapsto W_\varepsilon(y,u) - M_\varepsilon^u$ is lower semicontinuous with respect to the topology τ. By using Lemma 3.3.1, we have

$$\begin{aligned} &W_\varepsilon(y_k, u_k) - M_\varepsilon^{u_k} \\ &\quad = \int_0^T e^{-t/\varepsilon}\left(\frac{\varepsilon}{2}\|\dot{y}_k\|^2 + \phi(y_k) - (y_k, u_k)\right) \mathrm{d}t + \frac{\varepsilon^2}{2}\|\dot{y}_\varepsilon^{u_k}(0)\|^2 \\ &\qquad + \varepsilon e^{-T/\varepsilon}\phi(y_\varepsilon^{u_k}(T)) - \varepsilon\phi(y^0) - \varepsilon\int_0^T e^{-t/\varepsilon}(u_k, \dot{y}_\varepsilon^{u_k})\,\mathrm{d}t + \int_0^T e^{-t/\varepsilon}(u_k, y_\varepsilon^{u_k})\,\mathrm{d}t. \end{aligned}$$

Taking the lim inf as $k \to \infty$ and using convergences (3.4.1) and (3.4.8), the lower semicontinuity of $\|\cdot\|$ and of ϕ, and the fact that we have $\dot{y}_\varepsilon^{u_k}(t) \rightharpoonup \dot{y}_\varepsilon^u(t)$ in H for every $t \in [0,T]$ from convergence (3.4.8), we obtain

$$\begin{aligned}
&\liminf_{k\to\infty}\left(W_\varepsilon(y_k,u_k) - M_\varepsilon^{u_k}\right) \\
&\geq \int_0^T e^{-t/\varepsilon}\left(\frac{\varepsilon}{2}\|\dot{y}\|^2 + \phi(y) - (y,u)\right)\mathrm{d}t + \frac{\varepsilon^2}{2}\|\dot{y}_\varepsilon^u(0)\|^2 \\
&\quad + \varepsilon e^{-T/\varepsilon}\phi(y_\varepsilon^u(T)) - \varepsilon\phi(y^0) - \varepsilon\int_0^T e^{-t/\varepsilon}(u,\dot{y}_\varepsilon^u)\,\mathrm{d}t + \int_0^T e^{-t/\varepsilon}(u,y_\varepsilon^u)\,\mathrm{d}t \\
&= W_\varepsilon(y,u) - M_\varepsilon^u.
\end{aligned}$$

We have checked that $P_{\varepsilon\lambda}(y,u) \leq \liminf_k P_{\varepsilon\lambda}(y_k,u_k) = \inf_{H^1(0,T;H)\times U} P_{\varepsilon\lambda}$, proving that (y,u) minimizes $P_{\varepsilon\lambda}$.

3.5 Convergence: proof of Theorem 3.2.2.ii-iv

3.5.1 Proof of Theorem 3.2.2.ii.

Let us start by checking the Γ-convergence $P_\varepsilon \xrightarrow{\Gamma} P$. We focus first on the Γ-lim inf inequality (3.2.2). Assume $(y_\varepsilon,u_\varepsilon) \xrightarrow{\tau} (y,u)$. Without loss of generality, $\sup_\varepsilon P_\varepsilon(y_\varepsilon,u_\varepsilon) < \infty$. Then, $P_\varepsilon(y_\varepsilon,u_\varepsilon) = J(y_\varepsilon,u_\varepsilon)$, $y_\varepsilon = y_\varepsilon^{u_\varepsilon}$, and we have

$$\liminf_{\varepsilon\to0} P_\varepsilon(y_\varepsilon,u_\varepsilon) = \liminf_{\varepsilon\to0} J(y_\varepsilon,u_\varepsilon) \geq J(y,u),$$

where we used the lower semicontinuity of J. The identification $J(y,u) = P(y,u)$, and hence the inequality (3.2.2), follows then from Lemma 3.3.4.

To prove the recovery-sequence condition (3.2.3), we can assume, without loss of generality, that $P(\hat{y},\hat{u}) < \infty$. Then, we have $\hat{y} = S(\hat{u})$. As a recovery sequence we choose $\hat{u}_\varepsilon = \hat{u}$ and $\hat{y}_\varepsilon = y_\varepsilon^{\hat{u}}$. Then, we have the identification $P_\varepsilon(\hat{y}_\varepsilon,\hat{u}_\varepsilon) = J(\hat{y}_\varepsilon,\hat{u})$. Moreover, Lemma 3.3.4 ensures convergence $\hat{y}_\varepsilon \to \hat{y}$ in $H^\sigma(0,T;H)$ for all $\sigma \in (0,1)$. Assumption (A3) entails that

$$\limsup_{\varepsilon\to0} P_\varepsilon(\hat{y}_\varepsilon,\hat{u}_\varepsilon) = \limsup_{\varepsilon\to0} J(\hat{y}_\varepsilon,\hat{u}) \leq J(\hat{y},\hat{u}),$$

where we used the upper semicontinuity of $J(\cdot,\hat{u})$ in the strong topology of $H^\sigma(0,T;H)$. The identification $J(\hat{y},\hat{u}) = P(\hat{y},\hat{u})$ follows again from Lemma 3.3.4, which in particular ensures that $\hat{y} = S(\hat{u})$.

As the functionals P_ε are equicoercive in $H^1(0,T;H) \times L^2(0,T;H)$ from Proposition 3.2.1, Theorem 3.2.2.ii follows from the Fundamental Theorem of Γ-convergence [22, Thm. 7.4, p. 69].

3.5.2 Proof of Theorem 3.2.2.iii.

In order to prove the Γ-convergence $P_{\varepsilon\lambda} \xrightarrow{\Gamma} P_\varepsilon$, let us first check the Γ-lim inf inequality (3.2.2). Without loss of generality, let $(y_{\varepsilon\lambda},u_{\varepsilon\lambda}) \xrightarrow{\tau} (y_\varepsilon,u_\varepsilon)$ as $\lambda \to 0$ be such that $\sup_\lambda P_{\varepsilon\lambda}(y_{\varepsilon\lambda},u_{\varepsilon\lambda}) < \infty$. Then, we get

$$\liminf_{\lambda\to0} P_{\varepsilon\lambda}(y_{\varepsilon\lambda},u_{\varepsilon\lambda}) \geq \liminf_{\lambda\to0} J(y_{\varepsilon\lambda},u_{\varepsilon\lambda}) \geq J(y_\varepsilon,u_\varepsilon),$$

where the last inequality follows from the lower semicontinuity of J. Eventually, the identification $J(y_\varepsilon, u_\varepsilon) = P_\varepsilon(y_\varepsilon, u_\varepsilon)$ directly follows from Lemma 3.3.5.

A recovery sequence is given by $(\hat{y}_{\varepsilon\lambda}, \hat{u}_{\varepsilon\lambda}) = (y_\varepsilon^{\hat{u}_\varepsilon}, \hat{u}_\varepsilon)$. In fact, we readily obtain

$$\limsup_{\lambda\to 0} P_{\varepsilon\lambda}(y_\varepsilon^{\hat{u}_\varepsilon}, \hat{u}_{\varepsilon\lambda}) = \limsup_{\lambda\to 0} J(y_\varepsilon^{\hat{u}_\varepsilon}, \hat{u}_\varepsilon) = P_\varepsilon(\hat{y}_\varepsilon, \hat{u}_\varepsilon).$$

The equicoerciveness of the functionals $P_{\varepsilon\lambda}$ in $H^1(0,T;H) \times L^2(0,T;H)$ from Lemma 3.3.3 and an application of the Fundamental Theorem of Γ-convergence [22, Thm. 7.4, p. 69] conclude the proof.

3.5.3 Proof of Theorem 3.2.2.iv.

Recall that $\lambda = \lambda_\varepsilon$ is such that $\limsup_{\varepsilon\to 0} \lambda_\varepsilon \varepsilon^{-3} e^{T/\varepsilon} = 0$. In order to check that $P_{\varepsilon\lambda_\varepsilon} \xrightarrow{\Gamma} P$ we start from the Γ-lim inf inequality (3.2.2). Without loss of generality, we can consider $\sup_\varepsilon P_{\varepsilon\lambda_\varepsilon}(y_{\varepsilon\lambda_\varepsilon}, u_{\varepsilon\lambda_\varepsilon}) < \infty$. Hence, from the definition of $P_{\varepsilon\lambda_\varepsilon}$, we deduce

$$\liminf_{\varepsilon\to 0} P_{\varepsilon\lambda_\varepsilon}(y_{\varepsilon\lambda_\varepsilon}, u_{\varepsilon\lambda_\varepsilon}) \geq \liminf_{\varepsilon\to 0} J(y_{\varepsilon\lambda_\varepsilon}, u_{\varepsilon\lambda_\varepsilon}) \geq J(y, u),$$

where we used the fact that $W_\varepsilon(y_{\varepsilon\lambda_\varepsilon}, u_{\varepsilon\lambda_\varepsilon}) - M_\varepsilon^{u_{\varepsilon\lambda_\varepsilon}} \geq 0$ and the lower semicontinuity of J. The identification $J(y,u) = P(y,u)$ follows then from Lemma 3.3.6.

As regards the recovery-sequence condition (3.2.3), we first note that, without loss of generality, one can assume $P(\hat{y}, \hat{u}) < \infty$, which implies $\hat{y} = S(\hat{u})$. We choose the recovery sequence $(y_\varepsilon^{\hat{u}}, \hat{u})$, so that $P_{\varepsilon\lambda_\varepsilon}((y_\varepsilon^{\hat{u}}, \hat{u}) = J(y_\varepsilon^{\hat{u}}, \hat{u})$. Lemma 3.3.4 implies the convergence $y_\varepsilon^{\hat{u}} \to \hat{y}$ in $H^\sigma(0,T;H)$ for all $\sigma \in (0,1)$. By exploiting the upper semicontinuity of $J(\cdot, \hat{u})$ in the strong $H^\sigma(0,T;H)$ topology for some $\sigma \in (0,1)$ from assumption (A3), this entails that

$$\limsup_{\varepsilon\to 0} P_{\varepsilon\lambda_\varepsilon}(y_\varepsilon^{\hat{u}}, \hat{u}) = \limsup_{\varepsilon\to 0} J(y_\varepsilon^{\hat{u}}, \hat{u}) \leq J(\hat{y}, \hat{u}).$$

As we have that $\hat{y} = S(\hat{u})$, the equality $J(\hat{y}, \hat{u}) = P(\hat{y}, \hat{u})$ follows.

Having proved the Γ convergence, the statement follows from the equicoerciveness of the functionals $P_{\varepsilon\lambda_\varepsilon}$ in $H^1(0,T;H) \times L^2(0,T;H)$ from Lemma 3.3.3 by applying again [22, Thm. 7.4, p. 69].

Acknowledgements

T.F. acknowledges the support from the JSPS KAKENHI Grant-in-Aid for Scientific Research(C), Japan, Grant Number 21K03309. U.S. is supported by the Austrian Science Fund (FWF) through projects 10.55776/F65, 10.55776/I4354, 10.55776/I5149, and 10.55776/P32788.

Chapter 4

A global variational approach to the dynamics of elastic bodies

Abstract

In this chapter, we investigate a global-in-time variational approach to the dynamics of hyperelastic materials based on the Weighted Energy-Dissipation (WED) variational theory. A family of ε-dependent functionals defined over entire trajectories is introduced. We prove that these functionals admit minimizers and that such minimizers converge to a solution of hyperelastic dynamics as $\varepsilon \to 0$.

4.1 Introduction

In this chapter, we introduce and investigate a variational approach of the dynamics of hyperelastic materials. We address the situation without external volume forces on the time half line $t > 0$ as modeled by the following equation,

$$\rho y_{tt} - \nabla \cdot \mathrm{D}W(\nabla y) + \delta \nabla^4 : \nabla^4 y = 0 \quad \text{in } (0, \infty) \times \Omega. \tag{4.1.1}$$

Here, the *reference configuration* Ω is a bounded domain in $\mathbb{R}^3$ of class C^4, $y : (0, \infty) \times \Omega \to \mathbb{R}^3$ is the *deformation*, $\rho > 0$ is the referential material *density* (assumed to be space homogeneous), and δ is a *length-scale* parameter. The symbol D denotes the Gâteaux derivative, W is the *stored elastic energy density*, and the fourth-order operator ∇^4 is given by $\nabla \otimes \nabla \otimes \nabla \otimes \nabla$. The nonlinear map $W : \mathbb{R}^{3\times 3} \to [0, \infty]$ is taken to be the sum of a convex $C^{1,1}_{\mathrm{loc}}$ functional and a term inversely proportional to the determinant of its argument, so that $W(F) \to \infty$ as $\det F \to 0^+$. The existence of solutions of equation (4.1.1), subject to the initial conditions

$$y(0, x) = y^0, \quad y_t(0, x) = y^1, \tag{4.1.2}$$

is proven via the Weighted Energy-Dissipation (WED) variational approach. Indeed, we define an ε-dependent family of convex global-in-time functionals with an exponentially decaying weight. The minimizers of these functionals correspond to elliptic-in-time regularizations of equation (4.1.1). The core of this approach is proving that those minimizers converge to the solution of the equation (4.1.1) as ε goes to 0. In particular, for all $\varepsilon > 0$, we define the WED functional $G_\varepsilon : L^2(\mathbb{R}_+ \times \Omega) \to [0,\infty]$ via

$$G_\varepsilon(y) = \begin{cases} \int_0^\infty \int_\Omega e^{-t/\varepsilon} \left(\frac{\varepsilon^2\rho}{2}|y_{tt}|^2 + W(\nabla y) + \frac{\delta}{2}|\nabla^4 y|^2 \right) \mathrm{d}x\,\mathrm{d}t & \text{if } y \in K(y^0,y^1), \\ \infty & \text{else,} \end{cases}$$

where the domain $K(y^0, y^1)$ depends on the initial conditions and is given by

$$K(y^0,y^1) = \{y \in \mathcal{H} \ : \det \nabla y(t,x) > 0 \text{ a.e.} \in \mathbb{R}_+ \times \Omega\,,\ y(0,\cdot) = y^0\,,\ y_t(0,\cdot) = y^1\},$$

and $\mathcal{H}$ stands for (see Subsection 4.2.1 for details)

$$\mathcal{H} = H^2(\mathbb{R}_+, e^{-t/\varepsilon}; L^2(\Omega)) \cap L^2(\mathbb{R}_+, e^{-t/\varepsilon}; H^4(\Omega)).$$

Here, the initial condition $y^0 \in H^4(\Omega)$ is such that $\min_{\overline{\Omega}} \det \nabla y^0 =: \alpha > 0$ and the initial deformation velocity y^1 is taken to be in $H^4(\Omega)$, as well. In our main result, Theorem 4.3.1, we prove that

(i) for all $\varepsilon > 0$, there exists a global minimizer $y^\varepsilon \in K(y^0, y^1)$ of G_ε and

(ii) by taking the limit $\varepsilon \to 0$, y^ε converge to a solution of (4.1.1) up to subsequences.

Remark 4.1.1. Condition $\det \nabla y(t,x) > 0$ for almost every $(t,x) \in (0,T) \times \Omega$ in the definition of $K(y^0,y^1)$ implies that the displacement $y(t,\cdot)$ is locally orientation-preserving and invertible.

Remark 4.1.2. From the embedding $H^3(\Omega) \hookrightarrow C^{1,1/2}(\overline{\Omega})$, we get that $\nabla y^0 \in C^{1,\gamma}(\overline{\Omega})$ for some $\gamma \in (0,1)$.

The chapter is organized as follows. Section 4.2 collects some preliminary material. In Section 4.3, we state our main result, namely Theorem 4.3.1, which is then proved in Section 4.4.

4.2 Preliminaries

This section introduces the functional setting and presents the proof of some preliminary results. In particular, in Subsection 4.2.1 we introduce the weighted Lebesgue-Bochner spaces and a corresponding Poincaré-type inequality, see Lemma 4.2.1. Subsection 4.2.2 deals with some interpolation results, namely Proposition 4.2.3 and Theorem 4.2.4. The latter, ensures that the spatial gradient of the minimizers of G_ε has a uniformly positive determinant everywhere in space and time (see Corollary 4.4.2). The positivity of the determinant of the deformation gradient is a crucial issue in Elasticity theory. In particular, Theorem 4.2.4 allows us to compute the Euler-Lagrange equation for G_ε.

4.2.1 Function spaces

Let $q \in [1,\infty)$, $\varepsilon > 0$, and B be a Banach space. The *weighted Lebesgue-Bochner* space (see also [3, Section 3]) is defined as

$$L^q(\mathbb{R}_+, e^{-t/\varepsilon}\mathrm{d}t; B) = \{y \in \mathcal{M}(\mathbb{R}_+; B) \ : \ t \mapsto |y(t)|_B^q e^{-t/\varepsilon} \in L^1(\mathbb{R}_+)\},$$

where $\mathcal{M}(\mathbb{R}_+; B)$ denotes the space of strongly measurable functions with values in B and $|\cdot|_B$ is the norm in B. The weighted Sobolev-Bochner spaces $W^{1,p}(\mathbb{R}_+, e^{-t/\varepsilon}\mathrm{d}t; B)$ are defined analogously. Eventually we recall a Poincaré-type inequality [3, Lemma 2].

Lemma 4.2.1. *Let $\varepsilon > 0$, $1 \le p < \infty$, and B be a reflexive Banach space. There exists a constant $C = C(\varepsilon, p) > 0$ such that*

$$\|y\|^p_{L^p(\mathbb{R}_+, e^{-t/\varepsilon}\mathrm{d}t;B)} \le C\Big(|y(0)|^p_B + \|y_t\|^p_{L^p(\mathbb{R}_+, e^{-t/\varepsilon}\mathrm{d}t;B)}\Big) \quad \forall y \in W^{1,p}(\mathbb{R}_+, e^{-t/\varepsilon}\mathrm{d}t; B).$$

4.2.2 Interpolation results

Let us record the following embedding result from [41, Prop. 3.2, p. 1217].

Lemma 4.2.2. *Let $p \in (1, \infty)$ and $\Omega \subset \mathbb{R}^d$ be a bounded open domain with C^1 boundary. Let $s, r \ge 0$, $\alpha \in (0, 2)$, and $\beta > 0$ be given. If $v \in W^{s+\alpha,p}(0, \hat{T}; W^{r,p}(\Omega)) \cap W^{s,p}(0, \hat{T}; W^{r+\beta,p}(\Omega))$ then, for all $\theta \in [0, 1]$ one has that $v \in W^{s+\theta\alpha,p}(0, \hat{T}; W^{r+(1-\theta)\beta,p}(\Omega))$ and*

$$\|v\|_{W^{s+\theta\alpha,p}(0,\hat{T};W^{r+(1-\theta)\beta,p}(\Omega))} \le \|v\|^{\theta}_{W^{s+\alpha,p}(0,\hat{T};W^{r,p}(\Omega))} \|v\|^{1-\theta}_{W^{s,p}(0,\hat{T};W^{r+\beta,p}(\Omega))}.$$

We are now in the position to prove the following embedding result.

Proposition 4.2.3. *Let Ω be a bounded open subset of $\mathbb{R}^3$ with C^1 boundary. If $v \in H^2(0, \hat{T}; L^2(\Omega)) \cap L^2(0, \hat{T}; H^4(\Omega))$ then, for all $\theta \in (1/4, 3/8)$ there exists a constant $\gamma^* = \gamma^*(\Omega, \theta) > 0$ such that $\nabla v \in C^{0,\gamma^*}([0, \hat{T}] \times \overline{\Omega})$ and we have*

$$\|\nabla v\|_{C^{0,\gamma^*}([0,\hat{T}]\times\overline{\Omega})} \le \|v\|^{\theta}_{H^2(0,\hat{T};L^2(\Omega))} \|v\|^{1-\theta}_{L^2(0,\hat{T};H^4(\Omega))}.$$

Proof. Lemma 4.2.2 with $s = r = 0, \alpha = 2 - \nu, p = 2, \beta = 4$, for $\nu > 0$ small, implies

$$H^{2-\nu}(0, \hat{T}; L^2(\Omega)) \cap L^2(0, \hat{T}; H^4(\Omega)) \hookrightarrow H^{(2-\nu)\theta}(0, \hat{T}; H^{4(1-\theta)}(\Omega)).$$

Hence, in particular, $\nabla v \in H^{(2-\nu)\theta}(0, \hat{T}; H^{4(1-\theta)-1}(\Omega))$. By first choosing $\theta > \frac{1}{2(2-\nu)}$, by means of the Sobolev Embedding Theorem for fractional Sobolev spaces [26, Theorem 8.2, p. 562] we get the embedding

$$H^{(2-\nu)\theta}(0, \hat{T}; H^{4(1-\theta)-1}(\Omega)) \hookrightarrow C^{0,\gamma}([0, \hat{T}]; H^{4(1-\theta)-1}(\Omega)),$$

with $\gamma = (2 - \nu)\theta - (1/2) > 0$. By restricting our analysis to the case $\theta < 1/2$ we can write $H^{4(1-\theta)-1}(\Omega) = H^{1+\eta}(\Omega)$, where $\eta = 2(1 - 2\theta) \in (0, 1)$. By applying again the Sobolev Embedding Theorem for fractional Sobolev spaces [26, Theorem 6.7, p. 557], we get $H^{1+\eta}(\Omega) \hookrightarrow W^{1,p^*}(\Omega)$ where $p^* = 6/(3 - 2\eta)$. Eventually, by considering $\theta < 3/8$ we get $\eta \in (1/2, 1)$ and then $p^* > 3$. Hence, we first infer $H^{1+\eta}(\Omega) \hookrightarrow C^{0,\hat{\gamma}}(\overline{\Omega})$ where $\hat{\gamma} = \eta - 1/2 \in (0, 1/2)$ and then we conclude that $\nabla v \in C^{0,\gamma}([0, \hat{T}]; C^{0,\hat{\gamma}}(\overline{\Omega}))$. By defining $\gamma^* := \min\{\gamma, \hat{\gamma}\}$ we obtain the assertion. □

We conclude this section by recording the following result [30, Lemma 4.1, p. 868].

Theorem 4.2.4. *Let $\Omega_T := \Omega \times (0, T) \subset \mathbb{R}^{d+1}$ be a bounded Lipschitz domain, $J \in C^{0,\gamma}(\overline{\Omega_T})$, and $\gamma \in (0, 1)$. Suppose that $J(t, x) > 0$ almost everywhere in Ω_T and that there exist two positive constants L, M such that*

$$\int_{\Omega_T} (J(\tau, z))^{-s} \le L, \qquad \|J\|_{C^{0,\gamma}(\overline{\Omega_T})} \le M,$$

for some $s \ge (d + 1)/\gamma$. Then, there exists a constant $\widetilde{c} = \widetilde{c}(\Omega, M, L, s, d)$ such that

$$J(t, x) \ge \widetilde{c} \qquad \forall (t, x) \in \overline{\Omega_T}.$$

4.3 Assumptions and main results

This section is devoted to stating of our main result. We start by recalling some classical definitions and by introducing our assumption on data. Let SL(3) stand for the *special linear group*, namely the group of matrices in $\mathbb{R}^{3\times 3}$ with determinant equal to 1, let SO(3) stand for the *special orthogonal group*, namely the group of rotations with determinant 1, and let $\mathrm{GL}_+(3)$ stand for the subset of the *general linear group*, namely the group matrices in $\mathbb{R}^{3\times 3}$ with positive determinant, namely,

$$\begin{aligned}
\mathrm{SL}(3) &:= \{Q \in \mathbb{R}^{3\times 3} \mid \det Q = 1\},\\
\mathrm{SO}(3) &:= \{Q \in \mathrm{SL}(3) \mid Q^T Q = QQ^T = I\},\\
\mathrm{GL}_+(3) &:= \{Q \in \mathbb{R}^{3\times 3} \mid \det Q > 0\},
\end{aligned}$$

where we indicate by I the identity matrix in $\mathbb{R}^{3\times 3}$. From now on, we assume that Ω is a bounded domain in $\mathbb{R}^3$ with a C^1 boundary. We assume the following.

(A0) $y^0, y^1 \in H^4(\Omega)$ with $\min_{\overline{\Omega}} \nabla y^0 =: \alpha > 0$.

(A1) The stored elastic energy $W : \mathbb{R}^{3\times 3} \to [0,\infty]$ is given by

$$W(F) = \begin{cases} H(F) + \mu(\det F)^{-s} & \text{if } F \in \mathrm{GL}_+(3),\\ \infty & \text{else,} \end{cases}$$

with $\mu, s > 0$ and it fulfills

$$W(RF) = W(F) \qquad \forall F \in \mathrm{GL}_+(d), \quad \forall R \in \mathrm{SO}(3).$$

Here, $H \in C^3(\mathrm{GL}_+(3))$ is a convex functional so that there exist $q \in (1,\infty)$ and positive constants c_1, c_2 such that

$$\begin{aligned}
&\frac{1}{c_1}|F|^q - c_1 \le\ H(F) \le c_1\Big(1 + |F|^q\Big) \qquad \forall F \in \mathrm{GL}_+(3),\\
&|\mathrm{D}H(F)|^{q'} \le c_2\Big(1 + |F|^q\Big) \qquad \forall F \in \mathrm{GL}_+(d), \quad q' = q/(q-1).
\end{aligned}$$

A consequence of Assumption (A1) is that

$$\begin{aligned}
&|F|^q \le c_3(DH(F) : F + 1) \qquad \forall F \in \mathrm{GL}_+(3),\\
&H(F) \le c_4(DH(F) : F + 1) \le c_5(|F|^q + 1) \qquad \forall F \in \mathrm{GL}_+(3),
\end{aligned}$$

for some positive constants c_3, c_4, c_5.

Theorem 4.3.1 (WED Approach). *Under Assumptions* (A0)-(A1)*:*

(i) *The* WED *functional G_ε admits a global minimizer $y^\varepsilon \in K(y^0, y^1)$;*

(ii) *Letting y^ε minimize G_ε, for any sequence ε_k there exists a (not relabeled) subsequence $(y^{\varepsilon_k}, \mathrm{D}W(\nabla y^{\varepsilon_k}))$, such that $(y^{\varepsilon_k}, \mathrm{D}W(\nabla y^{\varepsilon_k})) \to (y, \mathrm{D}W(\nabla y))$ weakly* in*

$$[W^{1,\infty}(0,T; L^2(\Omega)) \cap L^\infty(0,T; H^4(\Omega))] \times L^\infty(0,T; L^m(\Omega)),$$

for any $m < 6$ where, y solves (4.1.1) *and fulfills the initial conditions* (4.1.2)*. In addition,*

$$y_t \in L^\infty(\mathbb{R}_+; L^2(\Omega)) \quad \textit{and} \quad \nabla^4 y \in L^\infty(\mathbb{R}_+; L^2(\Omega)).$$

4.4 Proof of Theorem 4.3.1

This section is devoted to the proof of Theorem 4.3.1. In Subsection 4.4.1, we introduce the time rescaling $t/\varepsilon = s$. Subsection 4.4.2 is then devoted to the proof of the existence of minimizers of the time-rescaled functional J_ε, which is the WED functional G_ε under the change of variable $t/\varepsilon = s$. In Subsection 4.4.3, we adapt a variational technique by SERRA & TILLI [53] to our case in order to get a priori estimates. In Subsection 4.4.4 and Subsection 4.4.5, we pass to the limit and we identify the limit functions. Eventually, we provide sharper regularity statements on the solutions in Subsection 4.4.6.

A caveat on notation: in the following we denote by c a generic positive constant depending on data but independent of ε, δ, and λ, possibly changing from line to line.

4.4.1 Time rescaling

Given the WED functional G_ε, the change of variable $t/\varepsilon = s$ brings to the time-rescaled functional

$$J_\varepsilon(v) = \begin{cases} \displaystyle\int_0^\infty \int_\Omega e^{-t}\left(\frac{\rho}{2}|v_{tt}|^2 + \varepsilon^2 W(\nabla v) + \frac{\varepsilon^2\delta}{2}|\nabla^4 v|^2\right) \mathrm{d}x\,\mathrm{d}t & \text{for } v \in K(y^0, \varepsilon y^1), \\ \infty & \text{else.} \end{cases}$$

In particular, letting $y \in K(y^0, y^1)$ we have that $v(t,x) = y(\varepsilon t, x)$ belongs to $K(y^0, \varepsilon y^1)$, and viceversa, and

$$J_\varepsilon(v) = \varepsilon G_\varepsilon(y).$$

4.4.2 Existence of minimizers

Lemma 4.4.1. *J_ε admits a minimizer $v \in K(y^0, \varepsilon y^1)$.*

Proof. We first check that $K(y^0, \varepsilon y^1)$ is not empty. Consider $\overline{v} = y^0 + \varepsilon\psi(t)y^1$, where $\psi \in C_c^\infty[0,\infty)$ with $\psi(0) = 0$ and $\psi'(0) = 1$, and $\varepsilon > 0$ is such that $|\nabla\overline{v}(x,t) - \nabla y^0(x)| \leq R$ for all $(t,x) \in [0,\infty) \times \overline{\Omega}$ and $R > 0$ given. Moreover, we assume

$$\varepsilon L_p|\nabla y^1(x)||\psi(t)| \leq \frac{\alpha}{2}, \qquad \forall (t,x) \in [0,\infty) \times \overline{\Omega},$$

where L_p is the Lipschitz constant of the map $F \mapsto \det F$ on $\{F \in \mathbb{R}^{3\times 3} \,|\, \sup_{x\in\overline{\Omega}}|F - \nabla y^0(x)| \leq R\}$. Then, one has

$$\det(\nabla\overline{v}) = \det(\nabla y^0(x) + \varepsilon\psi(t)\nabla y^1(x)) \geq \det(\nabla y^0(x)) - L_p|\varepsilon\psi(t)\nabla y^1(x)|,$$

from which we readily get $\det(\nabla y^0(x) + \varepsilon\psi(t)\nabla y^1(x)) \geq \alpha/2 > 0$. This implies that $J_\varepsilon(\overline{v})$ is bounded, namely,

$$\begin{aligned} J_\varepsilon(\overline{v}) &= \int_0^\infty \int_\Omega e^{-t}\left(\frac{\rho}{2}|\overline{v}_{tt}|^2 + \varepsilon^2 H(\nabla\overline{v}) + \varepsilon^2\mu(\det(\nabla\overline{v}))^{-s} + \frac{\varepsilon^2\delta}{2}|\nabla^4\overline{v}|^2\right) \\ &\leq \varepsilon^2 \int_\Omega \left(\frac{c\rho}{2}|y^1|^2 + c + \frac{\delta}{2}(|\nabla^4 y^0|^2 + \varepsilon^2\|\psi\|_\infty^2|\nabla^4 y^1|^2)\right) \\ &\leq c\varepsilon^2. \end{aligned} \tag{4.4.1}$$

Hence, $\overline{v} \in K(y^0, \varepsilon y^1)$ with $J_\varepsilon(\overline{v}) < \infty$.

We prove the existence of minimizers by the Direct Method. Let $\{v^k\}_k$ be an *minimizing sequence* for J_ε in $K(y^0, \varepsilon y^1)$, namely $J_\varepsilon(v^k) \to \inf_{K(y^0,\varepsilon y^1)} J_\varepsilon \geq 0$. Since J_ε is coercive on $\mathcal{K} := H^2(\mathbb{R}_+, e^{-t}; L^2(\Omega)) \cap L^2(\mathbb{R}, e^{-t}; H^4(\Omega))$, we have

$$\|v^k\|_{\mathcal{K}} \leq c, \tag{4.4.2}$$

from which we get, up to a not relabeled subsequence, that

$$v^k \rightharpoonup v \quad \text{in} \quad \mathcal{K}. \tag{4.4.3}$$

Note that $y^0 = v^k(0) \rightharpoonup v(0)$ in $L^2(\Omega)$, so that $v(0) = y^0$ and $\varepsilon y^1 = v^k_t(0) \rightharpoonup v_t(0)$ in $L^2(\Omega)$, and $v_t(0) = v^1$. To prove that v is actually the minimum of J_ε we have to show that

$$J_\varepsilon(v) \leq \liminf_{k\to\infty} J_\varepsilon(v^k).$$

By using the lower semicontinuity of the norms and the convergence (4.4.3) it is straightforward to obtain

$$\int_0^\infty \int_\Omega e^{-t}\left(\frac{\rho}{2}|v_{tt}|^2 + \frac{\varepsilon^2\delta}{2}|\nabla^4 v|^2\right) \leq \liminf_{k\to\infty} \int_0^\infty \int_\Omega e^{-t}\left(\frac{\rho}{2}|v^k_{tt}|^2 + \frac{\varepsilon^2\delta}{2}|\nabla^4 v^k|^2\right).$$

For the stored elastic energy W we proceed as follows. Let $T > 0$ be fixed. As the bound (4.4.2) entails

$$\|v^k\|_{H^2(0,T;L^2(\Omega))\cap L^2(0,T;H^4(\Omega))} \leq ce^{T/\varepsilon},$$

by applying Proposition 4.2.3 we get

$$\|\nabla v^k\|_{C^{0,\gamma^*}([0,T]\times\overline{\Omega})} \leq ce^{T/\varepsilon}, \tag{4.4.4}$$

and by means of the Aubin-Lions Lemma we get the strong convergence

$$v^k \to v \quad \text{in } L^2(0,T; H^3(\Omega)). \tag{4.4.5}$$

The bound (4.4.4) and the fact that

$$\int_0^T e^{-t} \int_\Omega (\det \nabla v^k)^{-s} < c$$

allow us to apply Theorem 4.2.4. Hence, we obtain that $\det \nabla v^k \geq C(T, \varepsilon) > 0$ everywhere and thus, by exploiting the strong convergence (4.4.5) we get the pointwise convergence $W(\nabla v^k) \to W(\nabla v)$ a.e. in $(0,T) \times \Omega$. We can now use a standard diagonal argument and extract a subsequence such that $W(\nabla v^k) \to W(\nabla v)$ a.e. in $(0, \infty) \times \Omega$. Thus, by applying Fatou's Lemma, we get

$$\int_0^\infty \int_\Omega e^{-t} W(\nabla v) \leq \liminf_{k\to\infty} \int_0^\infty \int_\Omega e^{-t} W(\nabla v^k),$$

which concludes the proof. □

Corollary 4.4.2. *There exists a constant $K = K(\alpha, \mu, s, d, T, |\Omega|) > 0$ such that every minimizer $v^* \in K(y^0, \varepsilon y^1)$ of J_ε satisfies*

$$\det \nabla v^*(t, x) > K \quad \forall (x,t) \in [0,T] \times \overline{\Omega}.$$

In particular, K is independent of the minimizer v^*.

Proof. Arguing along the lines of the proof of Lemma 4.4.1 we deduce that we have $J_\varepsilon \leq c$ where $c > 0$ is given and v is any minimizer of J_ε. This in particular entails that $\|\nabla v\|_{C^{0,\gamma^*}([0,T]\times\overline{\Omega})} \leq c$ and

$$\int_0^T e^{-t} \int_\Omega (\det \nabla v)^{-s} < c$$

independently of the minimizer. Then, by applying Theorem 4.2.4 we get the thesis. $\square$

4.4.3 A priori estimates

We now adapt to our case the variational argument originally proposed by Serra and Tilli in [53].

Proposition 4.4.3. *Let $v^\varepsilon \in K(y^0, \varepsilon y^1)$ be a minimizer of J_ε. Then there exists a constant $c = c(y^0, y^1) > 0$ such that*

$$J_\varepsilon(v^\varepsilon) \leq c\varepsilon^2.$$

Proof. As v^ε is a minimizer for J_ε, we have $J_\varepsilon(v^\varepsilon) \leq J_\varepsilon(\overline{v})$. The thesis follows then from inequality (4.4.1). $\square$

Let us define the function $L_\varepsilon \in L^1(\mathbb{R}_+, e^{-t}\mathrm{d}t)$ as

$$L_\varepsilon(t) := \int_\Omega \left(\frac{\rho}{2}|v^\varepsilon_{tt}|^2 + \varepsilon^2 W(\nabla v^\varepsilon) + \frac{\varepsilon^2\delta}{2}|\nabla^4 v^\varepsilon|^2 \right) \mathrm{d}x \quad \text{for a.e. } t \geq 0.$$

We introduce, for every $t \geq 0$, the nonnegative and nonincreasing function $H_\varepsilon \in W^{1,1}(0,\infty)$ given by

$$H_\varepsilon(t) := \int_t^\infty e^{-s} L_\varepsilon(s)\, \mathrm{d}s.$$

Note that $H_\varepsilon(0) = J_\varepsilon(v)$ and $\lim_{t\to\infty} H_\varepsilon(t) = 0$. We also define

$$D_\varepsilon(t) := \frac{\rho}{2}\int_\Omega |v^\varepsilon_{tt}|^2\, \mathrm{d}x \qquad \text{for a.e. } t > 0,$$

and

$$I_\varepsilon(t) := \frac{\rho}{4}\int_\Omega |v^\varepsilon_t|^2\, \mathrm{d}x \qquad \forall t \geq 0.$$

It is easy to check that $I_\varepsilon \in W^{1,1}(0,\infty)$ with

$$\frac{\mathrm{d}I_\varepsilon(t)}{\mathrm{d}t} = \frac{\rho}{2}\int_\Omega v^\varepsilon_t \cdot v^\varepsilon_{tt}\, \mathrm{d}x \qquad \text{for a.e. } t > 0.$$

Eventually, for almost every $t > 0$, we define the function

$$E_\varepsilon(t) := I_\varepsilon(t) - \frac{\mathrm{d}I_\varepsilon(t)}{\mathrm{d}t} + \frac{e^t}{2}H_\varepsilon(t).$$

Let us now state some results from [53], proved for semilinear waves, which still hold in our case.

Lemma 4.4.4 ([53])**.** *Let $v^\varepsilon \in K(y^0, \varepsilon y^1)$ be a minimizer of J_ε. Then, there exists a constant $c = c(y^0, y^1) > 0$ such that*

$$\|v^\varepsilon_{tt}\|^2_{L^2(\mathbb{R}_+, e^{-t}\mathrm{d}t; L^2(\Omega))} + \|v^\varepsilon_t\|^2_{L^2(\mathbb{R}_+, e^{-t}\mathrm{d}t; L^2(\Omega))} \leq c\varepsilon^2, \quad \|v^\varepsilon\|^2_{L^2(\mathbb{R}_+, e^{-t}\mathrm{d}t; L^2(\Omega))} \leq c.$$

Moreover, the function $e^{-t}I_\varepsilon(t) \in W^{1,1}(\mathbb{R}_+)$ and

$$\lim_{T\to\infty} e^{-T} I_\varepsilon(T) = 0.$$

The following Lemma is a straightforward extension to the elastic setting of the analogous result from [53].

Lemma 4.4.5. *Let $v^\varepsilon \in K(y^0, \varepsilon y^1)$ be a minimizer of J_ε. Then, there exists a constant $c = c(y^0, y^1) > 0$ such that*

$$\int_t^{t+1}\int_\Omega \left(W(\nabla v^\varepsilon) + \frac{\delta}{2}|\nabla^4 v^\varepsilon|^2 \right) \mathrm{d}x\,\mathrm{d}s \le c, \quad \int_\Omega |v_t^\varepsilon(x,t)|^2\,\mathrm{d}x \le c\varepsilon^2 \quad \forall t \ge 0,$$

and

$$\left| \int_\Omega v_{tt}^\varepsilon(t,x)h(x)\,dx \right| \le c\varepsilon^2 \|\nabla^4 h\|_{L^2(\Omega)}, \quad \forall h \in H^4(\Omega) \text{ a.e. } t > 0.$$

Moreover, v satisfies the Euler-Lagrange equation

$$\int_0^\infty \int_\Omega e^{-t}\left(\rho v_{tt}^\varepsilon \cdot \eta_{tt} + \varepsilon^2 \mathrm{D}W(\nabla v^\varepsilon) : \nabla\eta + \delta\varepsilon^2 \nabla^4 v^\varepsilon : \nabla^4\eta \right) \mathrm{d}x\,\mathrm{d}t = 0,$$

for every test function $\eta \in K(0,0)$.

We rewrite the estimates in Lemma 4.4.5 in terms of y^ε given by letting $v^\varepsilon(t,x) = y^\varepsilon(\varepsilon t, x)$, $t \ge 0$, with $y^\varepsilon(0) = y^0$, and $y_t^\varepsilon(0) = y^1$, which is a minimizer of G_ε. We get

$$\int_t^{t+T}\int_\Omega \left(W(\nabla y^\varepsilon) + \frac{\delta}{2}|\nabla^4 y^\varepsilon|^2 \right) \mathrm{d}x\,\mathrm{d}s \le cT, \tag{4.4.6}$$

and

$$\int_\Omega |y_t^\varepsilon(t,x)|^2\,\mathrm{d}x \le c \quad \forall t \ge 0, \tag{4.4.7}$$

$$\int_\Omega |y^\varepsilon(t,x)|^2\,\mathrm{d}x \le c(1+t^2) \quad \forall t \ge 0, \tag{4.4.8}$$

$$\left| \int_\Omega y_{tt}(t,x)h(x)\,\mathrm{d}x \right| \le c\|\nabla^4 h\|_{L^2(\Omega)} \quad \forall h \in H^4(\Omega), \text{ for a.e. } t > 0. \tag{4.4.9}$$

Moreover, y^ε satisfies the Euler-Lagrange equation

$$\int_0^\infty \int_\Omega e^{-t/\varepsilon}\left(\varepsilon^2 \rho y_{tt}^\varepsilon \cdot \eta_{tt} + \mathrm{D}W(\nabla y^\varepsilon) : \nabla\eta + \delta \nabla^4 y^\varepsilon : \nabla^4\eta \right) \mathrm{d}x\,\mathrm{d}t = 0, \tag{4.4.10}$$

for every test function $\eta \in K(0,0)$. Note that one has

$$\mathrm{D}W(\nabla y^\varepsilon) = \mathrm{D}H(\nabla y^\varepsilon) + \mu\frac{\mathrm{Cof}(\nabla y^\varepsilon)}{[\det \nabla y^\varepsilon]^{s+1}}.$$

Via Assumption (A1), bound (4.4.6) implies that

$$\int_t^{t+T}\int_\Omega |\mathrm{D}H(\nabla y^\varepsilon)|^{q'}\,\mathrm{d}x\,\mathrm{d}s \le cT. \tag{4.4.11}$$

Moreover, by making use of the Hadamard inequality, bound (4.4.6) also implies (see [11, Proposition 5.1] for instance)

$$\int_t^{t+T}\int_\Omega |\mathrm{Cof}(\nabla y^\varepsilon)|^{q/2}\,\mathrm{d}x\,\mathrm{d}s \le cT,$$

which, along with Corollary 4.4.2, gives

$$\int_t^{t+T}\int_\Omega \left| \frac{\mathrm{Cof}\,\nabla y^\varepsilon}{[\det \nabla y^\varepsilon]^{s+1}} \right|^{q/2} \mathrm{d}x\,\mathrm{d}s \le cT. \tag{4.4.12}$$

Eventually, from bounds (4.4.6)–(4.4.7) we get (cf. [14, Theorem II.5.14, p. 101])

$$\|\nabla y^\varepsilon\|_{C^0([t,t+T];H^1(\Omega))} \le cT. \tag{4.4.13}$$

4.4.4 Passage to the limit

From (4.4.6)–(4.4.7) and (4.4.11)–(4.4.12), by a standard diagonal argument and by passing to a not relabeled subsequence, we can find y, ζ, and ξ such that for every $T > 0$ the following convergences hold

$$y^\varepsilon \rightharpoonup y \quad \text{in } W^{1,\infty}(0,T;L^2(\Omega)), \tag{4.4.14}$$

$$y^\varepsilon \rightharpoonup y \quad \text{in } L^2(0,T;H^4(\Omega)), \tag{4.4.15}$$

$$\mathrm{D}H(\nabla y^\varepsilon) \rightharpoonup \zeta \quad \text{in } L^{q'}(0,T;L^{q'}(\Omega)), \tag{4.4.16}$$

$$\frac{\mathrm{Cof}\,\nabla y^\varepsilon}{[\det \nabla y^\varepsilon]^{s+1}} \rightharpoonup \xi \quad \text{in } L^{q/2}(0,T;L^{q/2}(\Omega)). \tag{4.4.17}$$

By applying the Aubin-Lions Lemma, we also get

$$y^\varepsilon \to y \quad \text{in } L^2(0,T;H^3(\Omega)). \tag{4.4.18}$$

Let us now consider $\eta(t,x) = e^{t/\varepsilon}\varphi(t,x)$ with $\varphi \in C_0^\infty(\mathbb{R}_+ \times \mathbb{R}^3)$ as a test function in the Euler-Lagrange equation (4.4.10). By means of the convergences (4.4.14)–(4.4.17) we pass to the limit $\varepsilon \to 0$ in equation (4.4.10) and obtain

$$-\int_0^\infty \int_\Omega \rho y_t \cdot \varphi_t + \int_0^\infty \int_\Omega \Big((\zeta + \mu\xi) : \nabla\varphi + \delta \nabla^4 y : \nabla^4 \varphi \Big) = 0,$$

which proves that y is the distributional solution of $\rho y_{tt} - \nabla \cdot (\zeta + \mu\xi) + \delta\nabla^4 : \nabla^4 y = 0$.

4.4.5 Identification of the limit functions

From the continuity of the term $\mathrm{D}H$ and thanks to the strong convergence (4.4.18) we get the pointwise convergence

$$\mathrm{D}H(\nabla y^\varepsilon(t,x)) \to \mathrm{D}H(\nabla y(t,x)) \quad \text{for a.e. } (t,x) \in (0,T) \times \Omega.$$

Moreover, since by assumption $\mathrm{D}H$ is a locally Lipschitz continuous function, thanks to the bound (4.4.13) and by means of the Sobolev Embedding Theorem, we get

$$\|\mathrm{D}H(\nabla y^\varepsilon)\|_{C^0([t,t+T];H^1(\Omega))} \leq cT, \tag{4.4.19}$$

from which, by applying the Vitali-Lebesgue Convergence Theorem, we obtain

$$\mathrm{D}H(\nabla y^\varepsilon) \to \mathrm{D}H(\nabla y) \quad \text{in } L^q(0,T;L^m(\Omega)),$$

for all $q < \infty$ and for any $m < 6$.

As the map $F \mapsto \mathrm{Cof}\,F/[\det F]^{s+1}$ is locally Lipschitz continuous for $\det F$ uniformly well-separated from 0, the strong convergence (4.4.18) and Corollary 4.4.2 imply the pointwise convergence

$$\frac{\mathrm{Cof}\,\nabla y^\varepsilon}{[\det \nabla y^\varepsilon]^{s+1}} \to \frac{\mathrm{Cof}\,\nabla y}{[\det \nabla y]^{s+1}} \quad \text{a.e. } (0,T) \times \Omega.$$

Similarly to the bound (4.4.19), from the bound (4.4.13) we get

$$\left\| \frac{\mathrm{Cof}\,\nabla y^\varepsilon}{[\det \nabla y^\varepsilon]^{s+1}} \right\|_{C^0([t,t+T];L^6(\Omega))} \leq cT,$$

from which, by applying the Vitali-Lebesgue convergence Theorem, we get

$$\frac{\operatorname{Cof}\nabla y^\varepsilon}{[\det \nabla y^\varepsilon]^{s+1}} \to \frac{\operatorname{Cof}\nabla y}{[\det \nabla y]^{s+1}} \quad \text{in } L^q(0,T;L^m(\Omega)),$$

for all $q < \infty$ and for any $m < 6$. This proves that y is a distributional solution of equation (4.1.1).

4.4.6 Sharper regularity

Considering $\varepsilon \to 0$ in (4.4.6), by lower semicontinuity we get

$$\lim_{T\to 0} \frac{1}{T} \int_t^{t+T} \int_\Omega |\nabla^4 y(x,s)|^2 \,\mathrm{d}x\,\mathrm{d}s \le c \quad \forall t \ge 0.$$

As $\int_\Omega |\nabla^4 y(x,s)|^2$ is integrable in $[t, t+T]$ for all $t \ge 0$, the Lebesgue Differentiation Theorem implies

$$\lim_{T\to 0} \frac{1}{T} \int_t^{t+T} \int_\Omega |\nabla^4 y|^2 \,\mathrm{d}x\,\mathrm{d}s = \int_\Omega |\nabla^4 y(x,t)|^2 \,\mathrm{d}x \quad \text{a.e.}\, t \ge 0.$$

Hence, we get $\nabla^4 y \in L^\infty(\mathbb{R}_+; L^2(\Omega))$.

Eventually, the first bound in (4.4.7) implies $y_t \in L^\infty(\mathbb{R}_+; L^2(\Omega))$.

4.4.7 Convergence of the initial data

The condition $y(0) = y^0$ follows from (4.4.14), since $y^0 = y^\varepsilon(0) \rightharpoonup y(0)$ in $L^2(\Omega)$. To prove that $y_t(0) = y^1$ we proceed as follows: Bounds (4.4.7) and (4.4.9) provide uniform estimates of y_t^ε and y_{tt}^ε in $L^\infty(\mathbb{R}_+; L^2(\Omega))$ and $L^\infty(\mathbb{R}_+; H^{-4}(\Omega))$, respectively. By passing to a not relabeled subsequence as $\varepsilon \to 0$, we get that $y_t^\varepsilon \overset{*}{\rightharpoonup} y_t$ in $W^{1,\infty}(\mathbb{R}_+; H^{-4}(\Omega))$. Thus, by considering any $\varphi \in C_0^\infty[0,\infty)$ and any $h \in H^4(\Omega)$,

$$\begin{aligned}&\int_0^\infty \varphi(t)\left(\int_\Omega y_{tt}^\varepsilon(t,x)h(x)\,\mathrm{d}x\right)\mathrm{d}t \\ &\qquad = -\int_0^\infty \varphi_t(t)\left(\int_\Omega y_t^\varepsilon(t,x)h(x)\,\mathrm{d}x\right)\mathrm{d}t - \varphi(0)\int_\Omega y^1(x)h(x)\end{aligned}$$

and the identification $y_t(0) = y^1 \in H^{-4}(\Omega)$ follows by passing to the limit as ε goes to 0.

References

[1] G. Akagi, V. Bögelein, A. Marveggio, U. Stefanelli. Weighted Inertia-Dissipation-Energy approach to doubly nonlinear wave equations. Submitted, 2024. arXiv:2401.08856.

[2] G. Akagi, S. Melchionna. Elliptic-regularization of nonpotential perturbations of doubly-nonlinear gradient flows of nonconvex energies: a variational approach. *J. Convex Anal.* 25 (2018), no. 3, 861–898.

[3] G. Akagi, S. Melchionna, U. Stefanelli. Weighted energy-dissipation approach for doubly-nonlinear problems on the half line. *J. Evol. Equ.* 18 (2017), 49–74.

[4] G. Akagi, U. Stefanelli. A variational principle for doubly nonlinear evolution. *Appl. Math. Lett.* 23 (2010), 1120–1124.

[5] G. Akagi, U. Stefanelli. Weighted energy-dissipation functionals for doubly nonlinear evolution. *J. Funct. Anal.* 260 (2011), 2541–2578.

[6] G. Akagi, U. Stefanelli. Doubly nonlinear evolution equations as convex minimization. *SIAM J. Math. Anal.* 46 (2014), 1922–1945.

[7] G. Akagi, U. Stefanelli. A variational principle for gradient flows of nonconvex energies. *J. Convex Anal.* 23 (2016), 53–75.

[8] G. Akagi, U. Stefanelli, R. Voso. Weighted Energy-Dissipation approach to semilinear gradient flows with state-dependent dissipation. Submitted, 2024. arXiv:2404.03370.

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

[10] M. Bathory, U. Stefanelli. Variational resolution of outflow boundary conditions for incompressible Navier-Stokes. *Nonlinearity*, 35 (2022), no. 11, 5553–5592.

[11] B. Benešová, M. Kružík, A. Schlömerkemper. A note on locking materials and gradient polyconvexity. *Math. Models Methods Appl. Sci.* (2018), no.11, 1–35.

[12] J. Bergh, J. Löfström. *Interpolation spaces. An introduction.* Grundlehren der Mathematischen Wissenschaften 223. Springer-Verlag, Berlin, Germany (1976).

[13] V. Bögelein, F. Duzaar, P. Marcellini. Existence of evolutionary variational solutions via the calculus of variations. *J. Differential Equations.* 256 (2014), 3912–3942.

[14] F. Boyer, P. Fabrie. *Mathematical tools for the study of the incompressible Navier-Stokes equations and related models.* Springer, Applied Mathematical Sciences, Vol. 183, New York, 2013.

[15] H. Brezis. *Opérateur maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* (French) North-Holland Mathematics Studies, no. 5. Notas de Matemática (50). North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973.

[16] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Springer, New York, 2010.

[17] T. Caraballo, M. Herrera-Cobos, P. Marín-Rubio. Asymptotic behaviour of nonlocal p-Laplacian reaction-diffusion problems. *J. Math. Anal. Appl.* 459 (2018), 997–1015.

[18] S. N. Chandler-Wilde, D. P. Hewett, A. Moiola. Interpolation of Hilbert and Sobolev spaces: quantitative estimates and counterexamples. *Mathematika*, 61 (2015), no. 2, 414–443.

[19] M. Chipot, B. Lovat. On the asymptotic behaviour of some nonlocal problems. *Positivity*, 3 (1999), 65–81.

[20] M. Chipot, L. Molinet. Asymptotic behaviour of some nonlocal diffusion problems. *Appl. Anal.* 80 (2001), 279–315.

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

[22] G. Dal Maso. *An introduction to* Γ*-convergence.* Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston Inc., Boston, MA, 1993.

[23] G. Dal Maso, L. De Luca. A minimization approach to the wave equation on time-dependent domains. *Adv. Calc. Var.* 13 (2018), 425–436.

[24] E. Davoli, U. Stefanelli. Dynamic perfect plasticity as convex minimization. *SIAM J. Math. Anal.* 51 (2019), no. 2, 672–730.

[25] E. De Giorgi. Conjectures concerning some evolution problems. *Duke Math. J.* 81 (1996), 255–268.

[26] E. Di Nezza, G. Palatucci, E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.* 136 (2012), 521–573.

[27] L. C. Evans. *Partial differential equations.* Graduate Studies in Mathematics, volume 19. American Mathematical Society, Providence, 1998.

[28] J. Ferreira, H. B. De Oliveira. Parabolic reaction-diffusion systems with nonlocal coupled diffusivity terms. *Discrete Contin. Dyn. Syst.* 37 (2017), no. 5, 2431–2453.

[29] T. Fukao, U. Stefanelli, R. Voso. Optimal control of gradient flows via the Weighted Energy-Dissipation method. Submitted, 2024. arXiv:2403.15055.

[30] T. J. Healey, S. Krömer. Injective weak solutions in second-gradient nonlinear elasticity. *ESAIM Control Optim. Calc. Var.* 15 (2009), 863–871.

[31] T. Ilmanen. *Elliptic regularization and partial regularity for motion by mean curvature.* Memoirs of the American Mathematical Society, number 520. American Mathematical Society, Providence, 1994.

[32] N. Hirano. Existence of periodic solutions for nonlinear evolution equations in Hilbert spaces. *Proc. Amer. Math. Soc.* 120 (1994), 185–192.

[33] C. J. Larsen, M. Ortiz, C. L. Richardson. Fracture paths from front kinetics: relaxation and rate independence. *Arch. Ration. Mech. Anal.* 193 (2009), 539–583.

[34] C. Lienstromberg, S. Schiffer, R. Schubert. A variational approach to the non-newtonian Navier-Stokes equations. Submitted, 2023. arXiv:2312.03546.

[35] M. Liero, S. Melchionna. The weighted energy-dissipation principle and evolutionary Γ-convergence for doubly nonlinear problems. *ESAIM Control Optim. Calc. Var.* 25 (2019), no. 36, 1–38.

[36] M. Liero, U. Stefanelli. A new minimum principle for Lagrangian mechanics, *J. Nonlinear Sci.* 23 (2013), 179–204.

[37] M. Liero, U. Stefanelli. Weighted inertia-dissipation-energy functionals for semilinear equations. *Boll. Unione Mat. Ital. (9)*, 6 (2013), 1–27.

[38] E. Mainini, D. Percivale. On the weighted inertia-energy approach to forced wave equations. *J. Differential Equations*, 385 (2024), 121–154.

[39] S. Melchionna. A variational approach to symmetry, monotonicity, and comparison for doubly-nonlinear equations. *J. Math. Anal. Appl.* 456 (2017), 1303–1328.

[40] S. Melchionna. A variational principle for nonpotential perturbations of gradient flows of nonconvex energies. *J. Differential Equations*, 262 (2017), 3737–3758.

[41] M. Meyries, R. Schnaubelt. Interpolation, embeddings and traces of anisotropic fractional Sobolev spaces with temporal weights. *J. Funct. Anal.* 262 (2012), 1200–1229.

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

[43] A. Mielke, T. Roubíček. *Rate-independent systems. Theory and application.* Applied Mathematical Sciences, 193. Springer, New York, 2015.

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

[45] A. Mielke, U. Stefanelli. Weighted energy-dissipation functionals for gradient flows. *ESAIM Control Optim. Calc. Var.* 17 (2011), 52–85.

[46] L. Nirenberg. An extended interpolation inequality. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (3)*, 20 (1966), 733–737.

[47] M. Ortiz, B. Schmidt, U. Stefanelli. A variational approach to Navier-Stokes. *Nonlinearity*, 31 (2018), 5664–5682.

[48] R. Rossi, G. Savaré, A. Segatti, U. Stefanelli. A variational principle for gradient flows in metric spaces. *C. R. Math. Acad. Sci. Paris*, 349 (2011), 1224–1574.

[49] R. Rossi, G. Savaré, A. Segatti, U. Stefanelli. Weighted energy-dissipation functionals for gradient flows in metric spaces. *J. Math. Pures Appl. (9)*, 129 (2019), 1–66.

[50] L. Scarpa, U. Stefanelli. Stochastic PDEs via convex minimization. *Comm. Partial Differential Equations*, 46 (2021), no. 1, 66–97

[51] L. Scarpa, U. Stefanelli. The energy-dissipation principle for stochastic parabolic equations. *Adv. Math. Sci. Appl.* 30 (2021), no. 2, 429-–452.

[52] E. Serra, P. Tilli. A minimization approach to hyperbolic Cauchy problems. *J. Eur. Math. Soc.* (JEMS), 18 (2016), 2019–2044.

[53] E. Serra, P. Tilli. Nonlinear wave equations as limits of convex minimization problems: proof of a conjecture by De Giorgi. *Ann. of Math. (2)*, 175 (2012), 1551–1574.

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

[55] E. N. Spadaro, U. Stefanelli. A variational view at mean curvature evolution for linear growth functionals. *J. Evol. Equ.* 11 (2011), 793–809.

[56] U. Stefanelli. The De Giorgi conjecture on elliptic regularization. *Math. Models Methods Appl. Sci.* 21 (2011), 1377–1394.

[57] L. Tentarelli. On the extensions of the De Giorgi approach to nonlinear hyperbolic equations. *Rend. Semin. Mat. Univ. Politec. Torino*, 74 (2016), no. 3-4, 151–160.

[58] L. Tentarelli, P. Tilli. An existence result for dissipative nonhomogeneous hyperbolic equations via a minimization approach. *J. Differential Equations*, 266 (2019), no. 8, 5185–5208.

[59] L. Tentarelli, P. Tilli. De Giorgi's approach to hyperbolic Cauchy problems: the case of nonhomogeneous equations. *Comm. Partial Differential Equations*, 43 (2018), no. 4, 677–698.

[60] E. Zeidler. *Nonlinear functional analysis and its applications II/B.* Springer-Verlag, New York, 1990.

Zusammenfassung

Variationsprinzipien sind wichtige Instrumente zur Beschreibung und Analyse physikalischer Phänomene. Eine Methode, die Existenz von Lösungen für partielle Differentialgleichungen zu beweisen, besteht darin, ein Funktional zu minimieren, dessen Euler-Lagrange-Gleichung der gegebenen partiellen Differentialgleichung entspricht. Da es in manchen Fällen einfacher ist, stationäre Punkte für das entsprechende Funktional zu finden, als die Existenz von Lösungen auf der Ebene der Gleichungen zu beweisen, liefert diese Methode eine Möglichkeit des direkten Nachweises der Existenz einer Lösung. Die Herausforderung bei dem Ansatz ist es ein geeignetes Funktional zu finden.

Das sogenannte WED-Variationsprinzip (Weighted Energy-Dissipation) stellt einen alternativen Ansatz im Vergleich zu den klassischen Variationsprinzipien dar und bietet interessante mathematische Perspektiven zur Behandlung von partiellen Differentialgleichungen. Diese Methode besteht darin, ein Funktional zu finden, dessen Euler-Lagrange-Gleichung eine regularisierte Version der betrachteten partiellen Differentialgleichung ist. Diese Regularisierung führt zu einer höheren Flexibilität in Anwendungen und ermöglicht es, sowohl konservative als auch dissipative Systeme auf variationelle Weise zu beschreiben. Der Nachteil dieser Herangehensweise ist jedoch der Nachweis des sogenannten kausalen Limits (causal limit), also die abschließende Aufhebung der Regularisierung.

In dieser Dissertation behandeln wir drei Probleme mit Hilfe des WED-Ansatzes. Zunächst stellen wir einen neuartigen Ansatz zur Untersuchung semilinearer Gradientenflüsse mit zustandsabhängiger Dissipation vor. Dies führt dazu, dass die Euler-Lagrange-Gleichung spezifische zustandsabhängige dissipative Terme aufweist, welche minimale Integrierbarkeit in der Zeit zeigen, was das Problem auf technischer Ebene erheblich erschwert. Des weiteren untersuchen wir ein allgemeines Problem der optimalen Steuerung im Rahmen von Gradientenflüssen. Wir schlagen zwei Approximationen des Problems vor, die beide auf der Variationsformulierung der Gradientenflussdynamik über den WED-Ansatz beruhen. Schlussendlich stellen wir einen Variationsansatz für die Dynamik hyperelastischer Materialien mit Hilfe der WED-Variantionstheorie vor, welcher global in der Zeit ist.