Weighted energy-dissipation functionals for doubly nonlinear evolution

Akagi, Goro

Ulisse, Stefanelli

(Citation)
Journal of Functional Analysis,260(9):2541-2578

(Issue Date)
2011-05-01

(Resource Type)
journal article

(Version)
Accepted Manuscript

(URL)
https://hdl.handle.net/20.500.14094/90001967

Weighted energy-dissipation functionals for doubly nonlinear evolution

Goro Akagi* and Ulisse Stefanelli†

Abstract

This paper is concerned with the Weighted Energy-Dissipation (WED) functional approach to doubly nonlinear evolutionary problems. This approach consists in minimizing (WED) functionals defined over entire trajectories. We present the features of the WED variational formalism and analyze the related Euler-Lagrange problems. Moreover, we check that minimizers of the WED functionals converge to the corresponding limiting doubly nonlinear evolution. Finally, we present a discussion on the functional convergence of sequences of WED functionals and present some application of the abstract theory to nonlinear PDEs.

Keywords: Doubly nonlinear equations; Variational principle; Γ-convergence.

AMS Mathematics Subject Classification (2010): 35K55.

1 Introduction

This paper is concerned with the analysis of the *Weighted Energy-Dissipation* (WED) functional $I_\varepsilon : L^p(0,T;V) \to (-\infty,\infty]$ given by

$$I_\varepsilon(u) := \int_0^T e^{-t/\varepsilon}\left(\psi(u'(t)) + \frac{1}{\varepsilon}\phi(u(t))\right) \mathrm{d}t.$$

Here, $t \in [0,T] \mapsto u(t) \in V$ is a given trajectory in a uniformly convex Banach space V, u' is the time derivative, $p \in [2,\infty)$, $\psi, \phi : V \to (-\infty,\infty]$ are convex functionals, and ψ has p-growth.

The WED functional arises as a new tool in order to possibly reformulate dissipative evolution problems in a variational fashion. In particular, minimizers u_ε of the WED functional I_ε taking a given initial value $u_\varepsilon(0) = u_0$ are expected to converge as $\varepsilon \to 0$ to solutions of the *doubly nonlinear* Cauchy problem

$$\partial\psi(u'(t)) + \partial\phi(u(t)) \ni 0, \quad 0 < t < T, \qquad u(0) = u_0 \tag{1. 1}$$

*Shibaura Institute of Technology, 307 Fukasaku, Minuma-ku, Saitama-shi, Saitama 337-8570, Japan. E-mail: g-akagi@sic.shibaura-it.ac.jp

†IMATI - CNR, v. Ferrata 1, I-27100 Pavia, Italy. E-mail: ulisse.stefanelli@imati.cnr.it

(here ∂ is the subdifferential, see §2.1). The differential problem (1. 1) expresses a balance between the system of *conservative actions* modeled by the gradient $\partial\phi$ of the *energy* ϕ and that of *dissipative actions* described by the gradient $\partial\psi$ of the *dissipation* ψ. This in particular motivates the terminology *WED* as the energy ϕ and dissipation ψ appear in I_ε along with the parameter $1/\varepsilon$ and the exponentially decaying weight $t \mapsto \exp(-t/\varepsilon)$.

The doubly nonlinear dissipative relation (1. 1) is extremely general and stands as a paradigm for dissipative evolution. Indeed, let us remark that the formulation (1. 1) includes the case of gradient flows, which corresponds to the choice of a quadratic dissipation ψ. Consequently, the interest in providing a variational approach to (1. 1) is evident, for it would pave the way to the application of general methods of the Calculus of Variations to a variety of nonlinear dissipative evolution problems.

This perspective has recently attracted attention and, particularly, the WED formalism has already been matter of consideration. At first, the WED functional approach has been addressed by Mielke & Ortiz [19] in the rate-independent case, namely for a positively 1-homogeneous dissipation ψ ($p = 1$). By requiring the compactness of sublevels for ϕ, in [19] it is checked that the limit $\varepsilon \to 0$ can be rigorously performed and minimizers of the WED functionals converge to suitably weak solutions of the corresponding limiting problem. These results are then extended and combined with time-discretization in [21].

Out of the rate-independent realm, the only available results for the WED functionals are for the gradient flow case $p = 2$ (particularly $\psi(\cdot) = |\cdot|^2/2$). In [10] Conti & Ortiz provide two concrete examples of nontrivial relaxations of WED functionals connected with applications in Mechanics. In particular, they show the possibility of tackling via the WED functional approach some specific micro-structure evolution problem and the respective scaling analysis. The general gradient flow case is addressed in [22] where the limit $\varepsilon \to 0$ is checked and the analysis is combined with time-discretization. In this case, the convexity of ϕ plays a crucial role and no compactness is assumed. Finally, the relaxation of the WED functional related to the evolution by mean curvature of cartesian surfaces is addressed in [27].

Our focus here is on more general cases $p \in [2, \infty)$ instead. By assuming p-growth and differentiability for ψ and some growth restriction and the compactness of the sublevels for ϕ we are able to prove that minimizers u_ε of I_ε converge to a solution of (1. 1) (paper [2] contains another result in this direction under a different assumption frame).

The limit $\varepsilon \to 0$ is clearly the crucial issue for the WED theory and it is usually referred to as the *causal limit*. This name is suggested by the facts that the Euler-Lagrange equation for I_ε turns out to be *elliptic-in-time* (hence *non-causal*) and that the causality of the limiting problem (1. 1) is restored as $\varepsilon \to 0$. More precisely, let X be a second reflexive Banach space which is densely and compactly embedded in V. Then, we shall prove that the Euler-Lagrange equation for I_ε under the constraint $u_\varepsilon(0) = u_0$ reads

$$-\varepsilon \frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{d}_V\psi(u_\varepsilon'(t))) + \mathrm{d}_V\psi(u_\varepsilon'(t)) + \partial_X\phi_X(u_\varepsilon(t)) \ni 0 \quad \text{in } X^*, \quad 0 < t < T, \tag{1. 2}$$

$$u_\varepsilon(0) = u_0, \tag{1. 3}$$

$$\mathrm{d}_V\psi(u_\varepsilon'(T)) = 0, \tag{1. 4}$$

where $\phi_X : X \to [0, \infty]$ is the restriction of ϕ onto X, and $\mathrm{d}_V\psi$ and $\partial_X\phi_X$ are the Gâteaux differential of ψ and the subdifferential of ϕ_X, respectively (see §2.1 for definitions). Hence, the Euler-Lagrange equation (1. 2) for I_ε stands as an *elliptic regularization in time* of (1. 1) (note the final condition (1. 4)). In particular, by formally taking the limit in the Euler-Lagrange equation (1. 2)–(1. 4) as $\varepsilon \to 0$, the following *causal* problem is recovered

$$\mathrm{d}_V\psi(u'(t)) + \partial_V\phi(u(t)) \ni 0 \ \text{ in } V^*, \quad 0 < t < T, \tag{1. 5}$$

$$u(0) = u_0. \tag{1. 6}$$

Note that the existence of global solutions for (1. 5)–(1. 6) was proved by Colli [9] in our very functional setting and it is hence out of question here. Instead, we concentrate on the possibility of recovering solutions to (1. 5)–(1. 6) via the minimization of the WED functionals I_ε and the causal limit $\varepsilon \to 0$. To this aim, we shall start from establishing the existence of strong solutions to the Euler-Lagrange system (1. 2)–(1. 4) which, apparently, was never considered before.

A second issue of this paper is the discussion of the functional convergence as $h \to 0$ of a sequence of WED functionals $I_{\varepsilon,h}$ in the form

$$I_{\varepsilon,h}(u) = \int_0^T e^{-t/\varepsilon}\Big(\psi_h(u'(t)) + \frac{1}{\varepsilon}\phi_h(u(t))\Big)\mathrm{d}t$$

with initial constraints $u(0) = u_{0,h} \in D(\phi_h)$ and two sequences of convex functionals $\psi_h, \phi_h : V \to (-\infty, \infty]$ depending on the additional parameter $h > 0$. We shall provide sufficient conditions under which $I_{\varepsilon,h} \to I_\varepsilon$ in the so-called *Mosco* sense (see Definition 6.1). In particular, our sufficient conditions consist of *separate* Γ-liminf conditions for ψ_h and ϕ_h as well as a suitable *joint recovery sequence condition* in the same spirit of [20]. The present functional-convergence results are new even in the gradient flow case $p = 2$.

Before closing this section let us mention that elliptic-in-time regularizations of parabolic problems are classical in the linear case and some results can be found in the monograph by Lions & Magenes [18]. As for the nonlinear case, one has to recall the paper by Ilmanen [16] where the WED is used in order to prove the existence and partial regularity of the so-called Brakke mean curvature flow of varifolds.

Apart from the WED formalism, a number of alternative contributions to other variational formulations to nonlinear evolutionary problems, e.g., Brézis-Ekeland's principle [7, 8], have been considered in order to characterize entire trajectories as critical points of functionals (see also [5], [12, 13, 14], [15] for linear cases, and [24, 25], [32, 33], [11], [28, 29, 30] for nonlinear cases). The advantages of the WED formalism over former variational approaches are that it relies on a true minimization procedure (plus passage to the causal limit) and that it directly applies to doubly nonlinear evolution equations.

This is the plan of the paper. Section 2 is devoted to enlist and comment assumptions and present some preliminary facts to be used throughout. In Section 3, we prove the existence of strong solutions of the Euler-Lagrange equation (1. 2)–(1. 4), whereas Section 4 brings to a proof of the coincidence between global minimizers for I_ε and strong solutions of the Euler-Lagrange equation. In Section 5, we check for the causal limit $\varepsilon \to 0$ and Section 6 is concerned with the functional convergence of the sequence of WED functionals

$I_{\varepsilon,h}$ as $h \to 0$. A typical example of a doubly nonlinear PDE fitting the current analysis is

$$\alpha(u_t) - \nabla \cdot (|\nabla u|^{m-2}\nabla u) = 0$$

with α monotone, non-degenerate, and polynomially growing at ∞. Details on the respective WED functional approach as well as its approximation by functional convergence are presented in Section 7. Eventually, the appendix contains a proof of a technical lemma.

2 Assumptions and preliminary material

2.1 Notation, subdifferential, Gâteaux differential

Let us collect in the following some preliminary material along with relevant notation.

Let φ be a proper (i.e., $\varphi \not\equiv \infty$), lower semicontinuous and convex functional from a normed space E into $(-\infty, \infty]$. Then the subdifferential operator $\partial_E \varphi : E \to E^*$ of φ is defined by

$$\partial_E \varphi(u) := \{\xi \in E^*;\ \varphi(v) - \varphi(u) \geq \langle \xi, v - u \rangle_E \ \text{ for all } \ v \in E\}$$

with the domain $D(\partial_E \varphi) := \{u \in D(\varphi);\ \partial_E \varphi(u) \neq \emptyset\}$ and obvious notation for the duality pairing. It is known that $\partial_E \varphi$ is maximal monotone in $E \times E^*$ ([6], [4]).

The functional φ is said to be Gâteaux differentiable at u (resp., in E), if there exists $\xi \in E^*$ such that

$$\lim_{h \to 0} \frac{\varphi(u + he) - \varphi(u)}{h} = \langle \xi, e \rangle_E \quad \text{for any } e \in E$$

at u (resp., for all $u \in E$). In this case, ξ is called a Gâteaux derivative of φ at u and denoted by $\mathrm{d}_E \varphi(u)$. We can naturally define an operator $\mathrm{d}_E \varphi$ from E into E^*. If φ is Gâteaux differentiable at u, then the set $\partial_E \varphi(u)$ consists of the single element $\mathrm{d}_E \varphi(u)$.

Throughout this paper, we denote by A the graph of a possibly multivalued operator $A : E \to E^*$. Hence $[u, \xi] \in A$ means that $u \in D(A)$ and $\xi \in A(u)$.

2.2 Assumptions

Let V and V^* be a uniformly convex Banach space and its dual space with norms $|\cdot|_V$ and $|\cdot|_{V^*}$, respectively, and a duality pairing $\langle \cdot, \cdot \rangle_V$ and let X be a reflexive Banach space with a norm $|\cdot|_X$ and a duality pairing $\langle \cdot, \cdot \rangle_X$ such that

$$X \hookrightarrow V \quad \text{and} \quad V^* \hookrightarrow X^*$$

with densely defined compact canonical injections. Let $\psi : V \to [0, \infty)$ be a Gâteaux differentiable convex functional and let $\phi : V \to [0, \infty]$ be a proper lower semicontinuous convex functional.

Let $p \in [2, \infty)$ and $m \in (1, \infty)$ be fixed, and introduce our basic assumptions:

(A1) there exist constants $C_1, C_2 > 0$ such that $C_1|u|_V^p \le \psi(u) + C_2$ for all $u \in V$;

(A2) there exist constants $C_3, C_4 > 0$ such that $|\mathrm{d}_V\psi(u)|_{V^*}^{p'} \le C_3|u|_V^p + C_4$ for all $u \in V$;

(A3) there exists a non-decreasing function ℓ_1 in $\mathbb{R}$ such that

$$|u|_X^m \le \ell_1(|u|_V)\,(\phi(u)+1) \qquad \text{for all } u \in D(\phi);$$

(A4) there exists a non-decreasing function ℓ_2 in $\mathbb{R}$ such that

$$|\eta|_{X^*}^{m'} \le \ell_2(|u|_V)\,(|u|_X^m+1) \qquad \text{for all } [u,\eta] \in \partial_X\phi_X,$$

where $\phi_X : X \to [0,\infty]$ denotes the restriction of ϕ on X.

Note that, by (A2) and the definition of subdifferential, the continuity of ψ in V also follows. Furthermore, we can also verify by (A3) and (A4) that ϕ_X is continuous in X and $D(\partial_X\phi_X) = X$. Moreover, from the definition of subdifferential and (A1), it also holds that

$$\text{(A1)}' \qquad C_1|u|_V^p \le \langle \mathrm{d}_V\psi(u), u\rangle_V + C_2' \qquad \text{for all } u \in V$$

with $C_2' := C_2 + \psi(0) \ge 0$. Let us manipulate (A2) in order to get

$$\text{(A2)}' \qquad \psi(u) \le \psi(0) + \langle \mathrm{d}_V\psi(u), u\rangle_V \le \psi(0) + \frac{1}{p'}|\mathrm{d}_V\psi(u)|_{V^*}^{p'} + \frac{1}{p}|u|_V^p \le C_3'(|u|_V^p + 1)$$

for all $u \in V$ with $C_3' := \psi(0) + C_4 + C_3 + 1 \ge 0$. Similarly, by (A4), we can also obtain

$$\phi(u) \le \ell_3(|u|_V)\,(|u|_X^m+1) \qquad \text{for all } u \in X$$

with a non-decreasing function ℓ_3 in $\mathbb{R}$.

Finally, let us give a precise definition of WED functionals of our main interest.

$$I_\varepsilon(u) = \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon}\left(\psi(u'(t)) + \frac{1}{\varepsilon}\phi(u(t))\right)\mathrm{d}t & \text{if } u \in W^{1,p}(0,T;V),\ u(0)=u_0 \\ & \quad \text{and } \psi(u'(\cdot)), \phi(u(\cdot)) \in L^1(0,T), \\ \infty & \text{else} \end{cases}$$

with an initial data $u_0 \in V$ and a parameter $\varepsilon > 0$. Then we remark that

$$D(I_\varepsilon) = \{u \in L^m(0,T;X) \cap W^{1,p}(0,T;V);\ u(0) = u_0\},$$

as the above remarks imply

$$\begin{aligned} & e^{-T/\varepsilon}\left(C_1\int_0^T |u'(t)|_V^p \mathrm{d}t + \frac{1}{\varepsilon\ell_1(\|u\|_{L^\infty(0,T;V)})}\int_0^T |u(t)|_X^m \mathrm{d}t - C_2T - \frac{T}{\varepsilon}\right) \\ \le\ & \int_0^T e^{-t/\varepsilon}\left(\psi(u'(t)) + \frac{1}{\varepsilon}\phi(u(t))\right)\mathrm{d}t \\ \le\ & C_3'\left(\int_0^T |u'(t)|_V^p \mathrm{d}t + T\right) + \frac{\ell_3(\|u\|_{L^\infty(0,T;V)})}{\varepsilon}\left(\int_0^T |u(t)|_X^m \mathrm{d}t + T\right), \end{aligned}$$

where $\|u\|_{L^\infty(0,T;V)} := \operatorname{ess\,sup}_{t\in[0,T]} |u(t)|_V$, for all $u \in L^m(0,T;X) \cap W^{1,p}(0,T;V)$.

2.3 Coincidence between $\partial_X \phi_X$ and $\partial_V \phi$

The following proposition shows some relationship between $\partial_V \phi$ and $\partial_X \phi_X$.

Proposition 2.1. *Let V and X be normed spaces such that $X \subset V$ with a continuous canonical injection. Let ϕ be a proper, lower semicontinuous and convex functional from V into $(-\infty, \infty]$. Moreover, let ϕ_X be the restriction of ϕ onto X. If $D(\phi) \subset X$, then*

$$D(\partial_V \phi) = \{w \in D(\partial_X \phi_X);\ \partial_X \phi_X(w) \cap V^* \neq \emptyset\}, \tag{2. 1}$$

and moreover,

$$\partial_V \phi(u) = \partial_X \phi_X(u) \cap V^* \quad \textit{for} \ \ u \in D(\partial_V \phi). \tag{2. 2}$$

Proof. We first note that $V^* \subset X^*$. Let $u \in D(\partial_V \phi)$ and $f \in \partial_V \phi(u)$ be fixed. For any $v \in D(\phi_X) \subset X$, noting that $D(\partial_V \phi) \subset D(\phi) \subset X$ by assumption, we find that

$$\begin{aligned} \phi_X(v) - \phi_X(u) &= \phi(v) - \phi(u) \\ &\geq \langle f, v - u \rangle_V = \langle f, v - u \rangle_X, \end{aligned}$$

which implies $u \in D(\partial_X \phi_X)$ and $f \in \partial_X \phi_X(u) \cap V^*$.

Conversely, let $u \in \{w \in D(\partial_X \phi_X);\ \partial_X \phi_X(w) \cap V^* \neq \emptyset\}$ and $f \in \partial_X \phi_X(u) \cap V^*$ be fixed. For $v \in D(\phi) \subset X$, it follows that

$$\begin{aligned} \phi(v) - \phi(u) &= \phi_X(v) - \phi_X(u) \\ &\geq \langle f, v - u \rangle_X = \langle f, v - u \rangle_V, \end{aligned}$$

which gives $u \in D(\partial_V \phi)$ and $f \in \partial_V \phi(u)$. Thus (2. 1) and (2. 2) hold. □

2.4 Representation of subdifferentials in $L^p(0, T; V)$

We provide here a result on the possible representation of the subdifferential of an integral functional. This representation turns out to be useful later on. We set $\mathcal{V} := L^p(0, T; V)$ and define the two functionals I^1_ε, $I^2_\varepsilon : \mathcal{V} \to [0, \infty]$ by

$$I^1_\varepsilon(u) := \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon} \psi(u'(t)) \mathrm{d}t & \text{if} \ \ u \in W^{1,p}(0, T; V), \ u(0) = u_0, \\ \infty & \text{else} \end{cases}$$

(note that $\psi(u'(\cdot)) \in L^1(0, T)$ for all $u \in W^{1,p}(0, T; V)$ by (A2)$'$) and

$$I^2_\varepsilon(u) := \begin{cases} \displaystyle\int_0^T \frac{1}{\varepsilon} e^{-t/\varepsilon} \phi(u(t)) \mathrm{d}t & \text{if} \ \ u \in L^m(0, T; X), \\ \infty & \text{else.} \end{cases}$$

Then it is obvious that

$$I_\varepsilon(u) = I^1_\varepsilon(u) + I^2_\varepsilon(u) \quad \text{for} \ \ u \in D(I_\varepsilon) = D(I^1_\varepsilon) \cap D(I^2_\varepsilon).$$

Moreover, I_ε^1 and I_ε^2 are proper, (weakly) lower semicontinuous and convex in $\mathcal{V}$.

Let us discuss the representation of the subdifferential operator $\partial_{\mathcal{V}} I_\varepsilon^1$. Define the operator $\mathcal{A} : \mathcal{V} \to \mathcal{V}^*$ by

$$\mathcal{A}(u)(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-t/\varepsilon}\mathrm{d}_V\psi(u'(t))\right) \quad \text{for} \ \ u \in D(\mathcal{A})$$

with the domain

$$D(\mathcal{A}) = \left\{ u \in D(I_\varepsilon^1);\ \mathrm{d}_V\psi(u'(\cdot)) \in W^{1,p'}(0,T;V^*),\ \mathrm{d}_V\psi(u'(T)) = 0 \right\}.$$

We have the following result.

Proposition 2.2 (Identification of $\mathcal{A}$)**.** *It holds that* $\mathcal{A} = \partial_{\mathcal{V}} I_\varepsilon^1$.

Proof. It can be easily seen that $\mathcal{A} \subset \partial_{\mathcal{V}} I_\varepsilon^1$. Hence it remains to prove the inverse inclusion. Set $\mathcal{W} := W^{1,p}(0,T;V)$. Define two functionals $J, K : \mathcal{W} \to [0,\infty]$ by

$$\begin{aligned} J(u) \ &:= \ \int_0^T e^{-t/\varepsilon}\psi(u'(t))\mathrm{d}t, \\ K(u) \ &:= \ \begin{cases} 0 & \text{if} \ \ u(0) = u_0, \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

and denote by $I_{\varepsilon,\mathcal{W}}^1$ the restriction of I_ε^1 on $\mathcal{W}$ (hence $I_{\varepsilon,\mathcal{W}}^1 = J + K$). Then, J is Gâteaux differentiable in $\mathcal{W}$. Indeed, let $u, e \in \mathcal{W}$ and let $h \in \mathbb{R}$. Then,

$$\frac{J(u+he) - J(u)}{h} = \int_0^T e^{-t/\varepsilon}\frac{\psi(u'(t)+he'(t)) - \psi(u'(t))}{h}\mathrm{d}t.$$

Since ψ is Gâteaux differentiable in V, the integrand of the right-hand side converges to $e^{-t/\varepsilon}\langle \mathrm{d}_V\psi(u'(t)), e'(t)\rangle_V$ for almost every $t \in (0,T)$ as $h \to 0$. Now, we easily compute that

$$\langle \mathrm{d}_V\psi(u'(t)), e'(t)\rangle_V \le \frac{\psi(u'(t)+he'(t)) - \psi(u'(t))}{h} \le \langle \mathrm{d}_V\psi(u'(t)+he'(t)), e'(t)\rangle_V$$

and, by means of (A2), we obtain

$$\begin{aligned} &\left|\frac{\psi(u'(t)+he'(t)) - \psi(u'(t))}{h}\right| \le \left(|\mathrm{d}_V\psi(u'(t))|_{V^*} + |\mathrm{d}_V\psi(u'(t)+he'(t))|_{V^*}\right)|e'(t)|_V \\ &\le \frac{C_3}{p'}|u'(t)|_V^p + \frac{2C_4}{p'} + \frac{C_3}{p'}|u'(t)+he'(t)|_V^p + \frac{2}{p}|e'(t)|_V^p \in L^1(0,T). \end{aligned}$$

Hence, by dominated convergence we deduce that J is Gâteaux differentiable in $\mathcal{W}$ and

$$\begin{aligned} \frac{J(u+he) - J(u)}{h} \ &\to \ \int_0^T e^{-t/\varepsilon}\langle \mathrm{d}_V\psi(u'(t)), e'(t)\rangle_V \mathrm{d}t \\ &=: \ \langle\langle \mathrm{d}_{\mathcal{W}} J(u), e\rangle\rangle_{\mathcal{W}}, \quad \text{for all } e \in \mathcal{W}, \end{aligned}$$

where $\langle\langle \cdot,\cdot\rangle\rangle_{\mathcal{W}}$ stands for the duality pairing between $\mathcal{W}$ and $\mathcal{W}^*$.

Moreover, K is proper, lower semicontinuous, and convex in $\mathcal{W}$, and we have

$$\langle\langle f, e\rangle\rangle_{\mathcal{W}} = 0 \quad \text{for all } [u,f] \in \partial_{\mathcal{W}} K \text{ and } e \in \mathcal{W} \text{ with } e(0) = 0.$$

Therefore, since $D(J) = \mathcal{W}$, we find that

$$\partial_{\mathcal{W}} I^1_{\varepsilon,\mathcal{W}} = \mathrm{d}_{\mathcal{W}} J + \partial_{\mathcal{W}} K$$

with the domain

$$D(\partial_{\mathcal{W}} I^1_{\varepsilon,\mathcal{W}}) = \{u \in \mathcal{W};\ u(0) = u_0\}.$$

Now, from the fact that $\mathcal{W} \subset \mathcal{V}$ and $D(I^1_\varepsilon) \subset \mathcal{W}$, it follows that

$$\partial_{\mathcal{V}} I^1_\varepsilon \subset \partial_{\mathcal{W}} I^1_{\varepsilon,\mathcal{W}}.$$

Let $[u,f] \in \partial_{\mathcal{V}} I^1_\varepsilon$ (hence $u(0) = u_0$). Then,

$$\int_0^T e^{-t/\varepsilon} \langle \mathrm{d}_V \psi(u'(t)), e'(t)\rangle_V \mathrm{d}t = \int_0^T \langle f(t), e(t)\rangle_V \mathrm{d}t$$

for any $e \in \mathcal{W}$ with $e(0) = 0$. Hence the function $t \mapsto e^{-t/\varepsilon} \mathrm{d}_V \psi(u'(t))$ belongs to $W^{1,p'}(0,T;V^*)$ and is such that

$$f(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-t/\varepsilon} \mathrm{d}_V \psi(u'(t))\right) \quad \text{for a.a. } t \in (0,T).$$

Moreover, we can also observe that $\mathrm{d}_V \psi(u'(T)) = 0$ from the arbitrariness of $e(T) \in V$. Thus $u \in D(\mathcal{A})$ and $f = \mathcal{A}(u)$. Consequently, $\mathcal{A}$ coincides with $\partial_{\mathcal{V}} I^1_\varepsilon$. □

2.5 Integration by parts at Lebesgue points in vector spaces

For $m, p \in (1,\infty)$, let the space $L^m(t_1,t_2;X) \cap W^{1,p}(t_1,t_2;V)$ ($t_1, t_2 \in [0,T]$ with $t_1 < t_2$) be endowed with the norm $\|\cdot\|_{L^m_X \cap W^{1,p}_V(t_1,t_2)}$ given by

$$\|u\|_{L^m_X \cap W^{1,p}_V(t_1,t_2)} := \|u\|_{L^m(t_1,t_2;X)} + \|u\|_{W^{1,p}(t_1,t_2;V)}$$

and, classically,

$$\|u\|_{L^m(t_1,t_2;X)} = \left(\int_{t_1}^{t_2} |u(t)|^m_X \mathrm{d}t\right)^{1/m}, \quad \|u\|_{W^{1,p}(t_1,t_2;V)} = \|u\|_{L^p(t_1,t_2;V)} + \|u'\|_{L^p(t_1,t_2;V)}.$$

Moreover, the space $L^m(t_1,t_2;X) \cap L^p(t_1,t_2;V)$ is analogously defined, and its dual space can be identified with

$$\begin{aligned} & L^{m'}(t_1,t_2;X^*) + L^{p'}(t_1,t_2;V^*) \\ = \ & \{f_1 + f_2;\ f_1 \in L^{m'}(t_1,t_2;X^*) \text{ and } f_2 \in L^{p'}(t_1,t_2;V^*)\}. \end{aligned}$$

Furthermore, the duality pairing between $L^m(t_1,t_2;X) \cap L^p(t_1,t_2;V)$ and its dual space can be written by

$$\langle\langle f, u\rangle\rangle_{L^m_X \cap L^p_V(t_1,t_2)} = \int_{t_2}^{t_1} \langle f_1(t), u(t)\rangle_X \mathrm{d}t + \int_{t_2}^{t_1} \langle f_2(t), u(t)\rangle_V \mathrm{d}t \quad \text{for } f = f_1 + f_2.$$

Then it follows immediately that

$$\|f\|_{L^{m'}_{X^*}+L^{p'}_{V^*}(t_1,t_2)} \leq \|f_1\|_{L^{m'}(t_1,t_2;X^*)} + \|f_2\|_{L^{p'}(t_1,t_2;V^*)} \quad \text{for } f = f_1 + f_2. \tag{2. 3}$$

In case $(t_1, t_2) = (0, T)$, we omit $(0, T)$ in the notation of the norms and the duality pairing. We shall be needing an integration by parts formula in this functional space setting.

Proposition 2.3 (Integration by parts)**.** *Let $m, p \in (1, \infty)$ and let $u \in L^m(0, T; X) \cap W^{1,p}(0, T; V)$ and $\xi \in L^{p'}(0, T; V^*)$ be such that $\xi' \in L^{m'}(0, T; X^*) + L^{p'}(0, T; V^*)$. Let $t_1, t_2 \in (0, T)$ be Lebesgue points of the function $t \mapsto \langle \xi(t), u(t) \rangle_V$. Then it holds that*

$$\langle\langle \xi', u \rangle\rangle_{L^m_X \cap L^p_V(t_1,t_2)} = \langle \xi(t_2), u(t_2) \rangle_V - \langle \xi(t_1), u(t_1) \rangle_V - \int_{t_1}^{t_2} \langle \xi(t), u'(t) \rangle_V \mathrm{d}t. \tag{2. 4}$$

Proof. There exist $f_1 \in L^{m'}(0, T; X^*)$ and $f_2 \in L^{p'}(0, T; V^*)$ such that $\xi' = f_1 + f_2$. Put

$$F_i(t) := \int_0^t f_i(s) \mathrm{d}s \quad \text{for } i = 1, 2.$$

Then, $F_1 \in W^{1,m'}(0, T; X^*)$, $F_2 \in W^{1,p'}(0, T; V^*)$ and $\xi(t) = F_1(t) + F_2(t) + \xi(0)$. We observe that

$$\lim_{\tau \to 0} \int_0^{T-\tau} \left| \frac{F_1(t+\tau) - F_1(t)}{\tau} - f_1(t) \right|_{X^*}^{m'} \mathrm{d}t = 0$$

and the analogue holds for F_2. Hence, we have

$$\begin{aligned} &\langle\langle \xi', u \rangle\rangle_{L^m_X \cap L^p_V(t_1,t_2)} \\ &= \lim_{\tau \to 0} \int_{t_1}^{t_2} \left(\left\langle \frac{F_1(t+\tau) - F_1(t)}{\tau}, u(t) \right\rangle_X + \left\langle \frac{F_2(t+\tau) - F_2(t)}{\tau}, u(t) \right\rangle_V \right) \mathrm{d}t \\ &= \lim_{\tau \to 0} \int_{t_1}^{t_2} \left\langle \frac{\xi(t+\tau) - \xi(t)}{\tau}, u(t) \right\rangle_V \mathrm{d}t. \end{aligned}$$

Moreover, since t_1, t_2 are Lebesgue points of $\langle \xi(\cdot), u(\cdot) \rangle_V$, it follows that

$$\begin{aligned} &\int_{t_1}^{t_2} \left\langle \frac{\xi(t+\tau) - \xi(t)}{\tau}, u(t) \right\rangle_V \mathrm{d}t \\ &= \frac{1}{\tau} \int_{t_2}^{t_2+\tau} \langle \xi(t), u(t) \rangle_V \mathrm{d}t - \frac{1}{\tau} \int_{t_1}^{t_1+\tau} \langle \xi(t), u(t) \rangle_V \mathrm{d}t \\ &\quad - \int_{t_1}^{t_2} \left\langle \xi(t+\tau), \frac{u(t+\tau) - u(t)}{\tau} \right\rangle_V \mathrm{d}t \\ &\to \langle \xi(t_2), u(t_2) \rangle_V - \langle \xi(t_1), u(t_1) \rangle_V - \int_{t_1}^{t_2} \langle \xi(t), u'(t) \rangle_V \mathrm{d}t \end{aligned}$$

as $\tau \to 0$. Thus we obtain (2. 4). □

3 The Euler-Lagrange equation

This section brings to a proof of the existence of strong solutions for the Euler-Lagrange equation (1. 2)–(1. 4) related to the WED functional I_ε. Hence, the value of the parameter ε is kept fixed throughout this section. We shall be concerned with the following precise notion of solution.

Definition 3.1 (Strong solution)**.** *A function* $u : [0,T] \to V$ *is said to be a* strong solution *of* (1. 2)–(1. 4) *if the following conditions are satisfied*:

$$u \in L^m(0,T;X) \cap W^{1,p}(0,T;V), \tag{3. 1}$$

$$\xi(\cdot) := \mathrm{d}_V\psi(u'(\cdot)) \in L^{p'}(0,T;V^*) \quad \text{and} \quad \xi' \in L^{m'}(0,T;X^*) + L^{p'}(0,T;V^*), \tag{3. 2}$$

there exists $\eta \in L^{m'}(0,T;X^*)$ *such that*

$$\eta(t) \in \partial_X\phi_X(u(t)), \quad -\varepsilon\xi'(t) + \xi(t) + \eta(t) = 0 \quad \text{in } X^* \quad \text{for a.a. } t \in (0,T), \tag{3. 3}$$

$$u(0) = u_0 \quad \text{and} \quad \xi(T) = 0. \tag{3. 4}$$

Remark 3.2. By definition, ξ belongs to $W^{1,\sigma}(0,T;X^*)$ with $\sigma := \min\{m',p'\} > 1$. We particularly deduce that $\xi \in C([0,T];X^*)$, and hence $\xi(t)$ lies in X^* at every $t \in [0,T]$.

The main result of this section is the following.

Theorem 3.3 (Existence of strong solutions)**.** *Assume that* (A1)–(A4) *are all satisfied. For every* $u_0 \in D(\phi)$*, the Euler equation* (1. 2)–(1. 4) *admits a strong solution satisfying*

$$\int_0^T |u_\varepsilon'(t)|_V^p \mathrm{d}t \leq \frac{1}{C_1}\left(\phi(u_0) + C_2'T + \varepsilon\psi(0)\right), \tag{3. 5}$$

$$\int_0^T \phi(u_\varepsilon(t))\mathrm{d}t \leq \left(\phi(u_0) + C_2'T + \varepsilon\psi(0)\right)T + \varepsilon\int_0^T \langle\xi_\varepsilon(t), u_\varepsilon'(t)\rangle_V \mathrm{d}t, \tag{3. 6}$$

$$\begin{aligned}\int_0^T \langle\eta_\varepsilon(t), u_\varepsilon(t)\rangle_X \mathrm{d}t \leq{}& -\langle\varepsilon\xi_\varepsilon(0), u_0\rangle_X - \int_0^T \langle\varepsilon\xi_\varepsilon(t), u_\varepsilon'(t)\rangle_V \mathrm{d}t \\ & - \int_0^T \langle\xi_\varepsilon(t), u_\varepsilon(t)\rangle_V \mathrm{d}t,\end{aligned} \tag{3. 7}$$

$$\int_0^T \langle\xi_\varepsilon(t), u_\varepsilon'(t)\rangle_V \mathrm{d}t \leq -\phi(u_\varepsilon(T)) + \phi(u_0) + \varepsilon\psi(0). \tag{3. 8}$$

The rest of this section is devoted to a proof of Theorem 3.3. The strategy of the proof is quite classical: we introduce suitable approximating problems by replacing ϕ with its Yosida approximation ϕ_λ, establish a-priori estimates independently of λ, and finally pass to the limit as $\lambda \to 0$. For the sake of clarity, we split this proof in subsequent subsections.

3.1 Approximating problem

Let us start by introducing the following approximate problems for $\lambda > 0$:

$$-\varepsilon\xi'_{\varepsilon,\lambda}(t) + \xi_{\varepsilon,\lambda}(t) + \eta_{\varepsilon,\lambda}(t) = 0 \ \text{ in } V^*, \qquad 0 < t < T, \tag{3. 9}$$

$$\xi_{\varepsilon,\lambda}(t) = \mathrm{d}_V\psi(u'_{\varepsilon,\lambda}(t)), \quad \eta_{\varepsilon,\lambda}(t) = \partial_V\phi_\lambda(u_{\varepsilon,\lambda}(t)) \ \text{ in } V^*, \qquad 0 < t < T, \tag{3. 10}$$

$$u_{\varepsilon,\lambda}(0) = u_0, \tag{3. 11}$$

$$\xi_{\varepsilon,\lambda}(T) = 0, \tag{3. 12}$$

where $\partial_V\phi_\lambda$ is the Yosida approximation of $\partial_V\phi$. Here we recall that $\partial_V\phi_\lambda$ coincides with the subdifferential operator of the Moreau-Yosida regularization ϕ_λ of ϕ given by

$$\phi_\lambda(u) := \inf_{v\in V}\left(\frac{1}{2\lambda}|u - v|_V^2 + \phi(v)\right) = \frac{1}{2\lambda}|u - J_\lambda u|_V^2 + \phi(J_\lambda u),$$

where J_λ is the resolvent for $\partial_V\phi$ (see [4] for more details). We also recall by definition that

$$F_V(J_\lambda u - u) + \lambda\partial_V\phi_\lambda(u) = 0 \quad \text{ for all } \ u \in V, \tag{3. 13}$$

where $F_V : V \to V^*$ denotes the duality mapping between V and V^*. Then, a strong solution $u_{\varepsilon,\lambda}$ of (3. 9)–(3. 12) on $[0, T]$ will be obtained as a global minimizer for the functional $I_{\varepsilon,\lambda} : \mathcal{V} \to [0, \infty]$ given by

$$I_{\varepsilon,\lambda}(u) = \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon}\left(\psi(u'(t)) + \frac{1}{\varepsilon}\phi_\lambda(u(t))\right)\mathrm{d}t & \text{if } u \in W^{1,p}(0,T;V),\ u(0) = u_0 \\ & \quad \text{and } \phi_\lambda(u(\cdot)) \in L^1(0,T), \\ \infty & \text{otherwise.} \end{cases}$$

More precisely, we have the following.

Lemma 3.4 (Solvability of the approximating problem)**.** *For each* $\varepsilon, \lambda > 0$*, the functional* $I_{\varepsilon,\lambda}$ *admits a global minimizer* $u_{\varepsilon,\lambda}$ *on* $\mathcal{V}$*. Moreover,* $u_{\varepsilon,\lambda}$ *is a strong solution of* (3. 9)–(3. 12) *and*

$$\xi_{\varepsilon,\lambda} \in W^{1,p'}(0,T;V^*) \quad \textit{and} \quad \eta_{\varepsilon,\lambda} \in L^{p'}(0,T;V^*). \tag{3. 14}$$

Proof. We observe that $I_{\varepsilon,\lambda}$ is proper, lower semicontinuous and convex in $\mathcal{V}$. Moreover, $I_{\varepsilon,\lambda}$ is coercive in $\mathcal{V}$ by (A1), i.e.,

$$I_{\varepsilon,\lambda}(u) \to \infty \quad \text{ if } \ \|u\|_{\mathcal{V}} \to \infty.$$

Hence, $I_{\varepsilon,\lambda}$ admits a (global) minimizer $u_{\varepsilon,\lambda}$ for each $\lambda > 0$.

We now define the functional $I^2_{\varepsilon,\lambda} : \mathcal{V} \to [0, \infty]$ as

$$I^2_{\varepsilon,\lambda}(u) := \begin{cases} \displaystyle\int_0^T \frac{1}{\varepsilon}e^{-t/\varepsilon}\phi_\lambda(u(t))\mathrm{d}t & \text{if } \ \phi_\lambda(u(\cdot)) \in L^1(0,T), \\ \infty & \text{otherwise.} \end{cases}$$

Note that, as $p \geq 2$, we have

$$\begin{aligned}\int_0^T \phi_\lambda(u(t))\mathrm{d}t &\leq \frac{1}{2\lambda}\int_0^T |u(t)-v|_V^2\mathrm{d}t + \phi(v)T \\ &< \infty \quad \text{for all } \ u \in \mathcal{V}\end{aligned}$$

with any $v \in D(\phi)$. In particular, $D(I^2_{\varepsilon,\lambda}) = \mathcal{V}$. Thus, we can deduce that $\partial_{\mathcal{V}} I^1_\varepsilon + \partial_{\mathcal{V}} I^2_{\varepsilon,\lambda}$ is maximal monotone in $\mathcal{V} \times \mathcal{V}^*$, and therefore

$$\partial_{\mathcal{V}} I_{\varepsilon,\lambda} = \partial_{\mathcal{V}} I^1_\varepsilon + \partial_{\mathcal{V}} I^2_{\varepsilon,\lambda}.$$

By Proposition 2.2, $\partial_{\mathcal{V}} I^1_\varepsilon$ coincides with the operator $\mathcal{A}$. Moreover, we can verify in a standard way (see, e.g., [17]) that, for $[u,f] \in \mathcal{V} \times \mathcal{V}^*$,

$$[u,f] \in \partial_{\mathcal{V}} I^2_{\varepsilon,\lambda} \quad \text{if and only if} \quad [u(t),f(t)] \in \frac{1}{\varepsilon}e^{-t/\varepsilon}\partial_V \phi_\lambda \ \text{ for a.a. } t \in (0,T).$$

Therefore, the assertion follows from global minimality, namely $\partial_{\mathcal{V}} I_{\varepsilon,\lambda}(u_{\varepsilon,\lambda}) \ni 0$. □

Remark 3.5 (Need for the approximation)**.** As for the Euler equation of I_ε, the argument of the proof of Lemma 3.4 does not apply, for it would not be clear how to check for the coincidence between $\partial_{\mathcal{V}} I_\varepsilon$ and the sum $\partial_{\mathcal{V}} I^1_\varepsilon + \partial_{\mathcal{V}} I^2_\varepsilon$; indeed $D(\partial_{\mathcal{V}} I^2_\varepsilon)$ might have no interior point in $\mathcal{V}$. Such a difficulty is one of reasons why we handle a weak formulation (1. 2)–(1. 4) of the Euler equation $\partial_{\mathcal{V}} I_\varepsilon(u) \ni 0$ in this paper.

3.2 A priori estimates

From here on we simply write $u_\lambda, \xi_\lambda, \eta_\lambda$ instead of $u_{\varepsilon,\lambda}, \xi_{\varepsilon,\lambda}, \eta_{\varepsilon,\lambda}$, respectively. Testing relation (3. 9) on $u'_\lambda(t)$ and integrating over $(0,T)$, we have

$$-\varepsilon \int_0^T \langle \xi'_\lambda(t), u'_\lambda(t)\rangle_V \,\mathrm{d}t + \int_0^T \langle \xi_\lambda(t), u'_\lambda(t)\rangle_V \,\mathrm{d}t + \phi_\lambda(u_\lambda(T)) - \phi_\lambda(u_0) = 0.$$

We now use the Neumann boundary condition (3. 12) in order to get

$$\begin{aligned}\int_0^T \langle \xi'_\lambda(t), u'_\lambda(t)\rangle_V \,\mathrm{d}t &= -\langle \xi_\lambda(0), u'_\lambda(0)\rangle_V - \int_0^T \langle \xi_\lambda(t), u''_\lambda(t)\rangle_V \,\mathrm{d}t \\ &\leq \psi(0) - \psi(u'_\lambda(0)) - \psi(u'_\lambda(T)) + \psi(u'_\lambda(0)) \\ &= \psi(0) - \psi(u'_\lambda(T)) \leq \psi(0). \end{aligned} \tag{3. 15}$$

Note that this calculation is presently just formal, for u need not to belong to $W^{2,p}(0,T;V)$. This procedure can however be rigorously justified and we have collected some detail in Lemma 7.3. Moreover, the following holds as well

$$\int_0^t \langle \xi'_\lambda(s), u'_\lambda(s)\rangle_V \mathrm{d}s \leq \langle \xi_\lambda(t), u'_\lambda(t)\rangle_V + \psi(0) \quad \text{for all } \ t \in \mathcal{L}_\lambda, \tag{3. 16}$$

where the set $\mathcal{L}_\lambda$ is defined by

$$\begin{aligned}\mathcal{L}_\lambda := \Big\{ & t \in (0,T); \ u_\lambda \text{ is differentiable in } V \text{ at } t \\ & \text{and } \ t \text{ is a Lebesgue point of } t \mapsto \langle \xi_\lambda(t), u'_\lambda(t)\rangle_V \Big\}.\end{aligned}$$

By using relation (3. 15) we have

$$\int_0^T \langle \xi_\lambda(t), u_\lambda'(t)\rangle_V \mathrm{d}t + \phi_\lambda(u_\lambda(T)) \le \phi_\lambda(u_0) + \varepsilon\psi(0). \tag{3. 17}$$

Hence, by relation (A1)$'$, one obtains

$$C_1 \int_0^T |u_\lambda'(t)|_V^p \mathrm{d}t + \phi_\lambda(u_\lambda(T)) \le \phi_\lambda(u_0) + C_2'T + \varepsilon\psi(0) \tag{3. 18}$$

and

$$\int_0^T |u_\lambda'(t)|_V^p \mathrm{d}t \le \frac{1}{C_1}\left(\phi(u_0) + C_2'T + \varepsilon\psi(0)\right), \tag{3. 19}$$

$$\int_0^T |\xi_\lambda(t)|_{V^*}^{p'} \mathrm{d}t + \sup_{t\in[0,T]} |u_\lambda(t)|_V + \sup_{t\in[0,T]} |J_\lambda u_\lambda(t)|_V \le C. \tag{3. 20}$$

Here, we have used assumption (A2) and the fact that $|J_\lambda u|_V \le C(|u|_V + 1)$ for all $u \in V$ and $\lambda > 0$ (see [6], [4]). Hence, by testing equation (3. 9) on $u_\lambda'(t)$ and integrating over $(0,t)$ we obtain by (3. 16)

$$\begin{aligned} &C_1 \int_0^t |u_\lambda'(\tau)|_V^p d\tau + \phi_\lambda(u_\lambda(t)) \\ &\le \ \phi(u_0) + C_2'T + \varepsilon\langle \xi_\lambda(t), u_\lambda'(t)\rangle_V + \varepsilon\psi(0) \quad \text{for all } \ t \in \mathcal{L}_\lambda. \end{aligned}$$

As the set $\mathcal{L}_\lambda$ has full Lebesgue measure, i.e., the measure of $(0,T) \setminus \mathcal{L}_\lambda$ is zero, by integrating both sides over $(0,T)$ again, we deduce that

$$\int_0^T \phi_\lambda(u_\lambda(t))\mathrm{d}t \le \left(\phi(u_0) + C_2'T + \varepsilon\psi(0)\right)T + \varepsilon \int_0^T \langle \xi_\lambda(t), u_\lambda'(t)\rangle_V \mathrm{d}t. \tag{3. 21}$$

Finally, from the above estimates we obtain

$$\int_0^T \phi_\lambda(u_\lambda(t))\mathrm{d}t \le C. \tag{3. 22}$$

Since $\phi(J_\lambda u_\lambda(t)) \le \phi_\lambda(u_\lambda(t))$ and $\eta_\lambda(t) \in \partial_V\phi(J_\lambda u_\lambda(t)) \subset \partial_X\phi_X(J_\lambda u_\lambda(t))$, it follows from assumptions (A3) and (A4) that

$$\int_0^T |J_\lambda u_\lambda(t)|_X^m \mathrm{d}t + \int_0^T |\eta_\lambda(t)|_{X^*}^{m'} \mathrm{d}t \le C. \tag{3. 23}$$

Eventually, a comparison in equation (3. 9) yields

$$\|\varepsilon\xi_\lambda'\|_{L^{m'}_{X^*} + L^{p'}_{V^*}} \le C, \tag{3. 24}$$

which, in particular, implies

$$\sup_{t\in[0,T]} |\varepsilon\xi_\lambda(t)|_{X^*} \le C. \tag{3. 25}$$

3.3 Passage to the limit

From the a priori estimates of Subsection 3.2, we have, for some not relabelled subsequences,

$$
\begin{aligned}
u_\lambda &\to u && \text{weakly in } W^{1,p}(0,T;V), && (3.\ 26)\\
J_\lambda u_\lambda &\to v && \text{weakly in } L^m(0,T;X), && (3.\ 27)\\
\xi_\lambda &\to \xi && \text{weakly in } L^{p'}(0,T;V^*), && (3.\ 28)\\
\eta_\lambda &\to \eta && \text{weakly in } L^{m'}(0,T;X^*), && (3.\ 29)\\
\xi'_\lambda &\to \xi' && \text{weakly in } L^{m'}(0,T;X^*) + L^{p'}(0,T;V^*). && (3.\ 30)
\end{aligned}
$$

That is

$$-\varepsilon\xi' + \xi + \eta = 0 \tag{3. 31}$$

and the estimate (3. 5) follows directly from the bound (3. 19).

Note that u_λ is equicontinuous in $C([0,T];V)$ with respect to λ from the bound (3. 19) and put $v_\lambda(t) := J_\lambda u_\lambda(t) - u_\lambda(t)$. By (3. 13) and the monotonicity of $\partial_V\phi$, we have

$$\langle F_V(v_\lambda(t+h)) - F_V(v_\lambda(t)), J_\lambda u_\lambda(t+h) - J_\lambda u_\lambda(t)\rangle_V \le 0,$$

which, together with estimate (3. 20), implies

$$\langle F_V(v_\lambda(t+h)) - F_V(v_\lambda(t)), v_\lambda(t+h) - v_\lambda(t)\rangle_V \le C|u_\lambda(t+h) - u_\lambda(t)|_V.$$

Hence, the right-hand side goes to zero as $h \to 0$ uniformly for $\lambda > 0$. Since V is uniformly convex, thanks to [31], for each $R > 0$ there exists a strictly increasing function m_R on $[0,\infty)$ such that $m_R(0) = 0$ and

$$m_R(|u-v|_V) \le \langle F_V(u) - F_V(v), u-v\rangle_V \quad \text{for } u,v \in B_R := \{u \in V;\ |u|_V \le R\}.$$

Namely, $v_\lambda(\cdot)$ is equicontinuous in $C([0,T];V)$ for $\lambda > 0$, and so is $J_\lambda u_\lambda(\cdot)$. Recalling that X is compactly embedded in V, by Theorem 3 of [26], we deduce from estimate (3. 23) that

$$J_\lambda u_\lambda \to v \quad \text{strongly in } C([0,T];V). \tag{3. 32}$$

By the integral estimate (3. 22), we have

$$\int_0^T |u_\lambda(t) - J_\lambda u_\lambda(t)|_V^2 \mathrm{d}t \le 2\lambda \int_0^T \phi_\lambda(u_\lambda(t))\mathrm{d}t \le 2\lambda C \to 0.$$

Therefore, we have $u = v$ and the bound in (3. 20) entails that

$$u_\lambda \to u \quad \text{strongly in } L^q(0,T;V) \tag{3. 33}$$

with an arbitrary $q \in [1,\infty)$. Hence, the strong convergences (3. 32) and (3. 33) yield

$$
\begin{aligned}
J_\lambda u_\lambda(t) &\to u(t) && \text{strongly in } V \ \text{ for all } \ t \in [0,T], && (3.\ 34)\\
u_\lambda(t) &\to u(t) && \text{strongly in } V \ \text{ for a.a. } \ t \in (0,T). && (3.\ 35)
\end{aligned}
$$

Since V^* is compactly embedded in X^*, estimates (3. 20) and (3. 24) entail

$$\xi_\lambda \to \xi \quad \text{strongly in } C([0,T];X^*), \tag{3. 36}$$

which also implies $\xi(T) = 0$.

Put now $p(t) := \liminf_{\lambda\to 0} |\xi_\lambda(t)|_{V^*}^{p'}$, and note that $p \in L^1(0,T)$ by Fatou's Lemma and (3. 20). Then $p(t) < \infty$ for a.a. $t \in (0,T)$, and for such $t \in (0,T)$ we can take a subsequence $\lambda_n^t \to 0$ (possibly depending on t) such that

$$\xi_{\lambda_n^t}(t) \to \xi(t) \quad \text{weakly in } V^*. \tag{3. 37}$$

We shall now check for the almost everywhere relations

$$\eta(t) \in \partial_X \phi_X(u(t)), \quad \xi(t) = \mathrm{d}_V \psi(u'(t)).$$

Let us start from the former. Define the subset $\mathcal{L} \subset (0,T)$ by

$$\mathcal{L} := \Big\{ t \in (0,T); \quad \begin{array}{l} t \text{ is a Lebesgue point of the function } t \mapsto \langle \xi(t), u(t)\rangle_V, \text{ and} \\ \text{for any sequence } \lambda_n \to 0, \text{ there exists a subsequence} \\ \lambda_{n'} \to 0 \text{ such that } \langle \xi_{\lambda_{n'}}(t), u_{\lambda_{n'}}(t)\rangle_V \to \langle \xi(t), u(t)\rangle_V \Big\}. \end{array}$$

Note that the convergences (3. 35) and (3. 37) entail that $\mathcal{L}$ has full Lebesgue measure. For arbitrary $t_1, t_2 \in \mathcal{L}$ with $t_1 \le t_2$, we have

$$\begin{aligned} \int_{t_1}^{t_2} \langle \eta_\lambda(t), J_\lambda u_\lambda(t)\rangle_X \mathrm{d}t &= \int_{t_1}^{t_2} \langle \eta_\lambda(t), u_\lambda(t)\rangle_V \mathrm{d}t - \lambda \int_{t_1}^{t_2} |\eta_\lambda(t)|_{V^*}^2 \mathrm{d}t \\ &\le \int_{t_1}^{t_2} \langle \eta_\lambda(t), u_\lambda(t)\rangle_V \mathrm{d}t. \end{aligned} \tag{3. 38}$$

On the other hand, from equation (3. 9) it follows that

$$\int_{t_1}^{t_2} \langle \eta_\lambda(t), u_\lambda(t)\rangle_V \mathrm{d}t = \int_{t_1}^{t_2} \langle \varepsilon\xi_\lambda'(t), u_\lambda(t)\rangle_V \mathrm{d}t - \int_{t_1}^{t_2} \langle \xi_\lambda(t), u_\lambda(t)\rangle_V \mathrm{d}t. \tag{3. 39}$$

Moreover, since $u_\lambda \in W^{1,p}(0,T;V)$ and $\xi_\lambda \in W^{1,p'}(0,T;V^*)$ (hence the function $t \mapsto \langle \xi_\lambda(t), u_\lambda(t)\rangle_V$ is differentiable-in-time for almost all $t \in (0,T)$), we note that

$$\begin{aligned} &\int_{t_1}^{t_2} \langle \varepsilon\xi_\lambda'(t), u_\lambda(t)\rangle_V \mathrm{d}t \\ &\quad = \varepsilon\langle \xi_\lambda(t_2), u_\lambda(t_2)\rangle_V - \varepsilon\langle \xi_\lambda(t_1), u_\lambda(t_1)\rangle_V - \int_{t_1}^{t_2} \langle \varepsilon\xi_\lambda(t), u_\lambda'(t)\rangle_V \mathrm{d}t \end{aligned} \tag{3. 40}$$

(note that (3. 38)–(3. 40) also hold for any $t_1, t_2 \in [0,T]$ with $t_1 \le t_2$).

By the definition of $\mathcal{L}$,

$$\langle \xi_{\lambda_n}(t_i), u_{\lambda_n}(t_i)\rangle_V \to \langle \xi(t_i), u(t_i)\rangle_V \quad \text{for } i = 1, 2 \tag{3. 41}$$

with a subsequence $\lambda_n \to 0$ (possibly depending on t_1, t_2). By a standard argument for monotone operators, it follows from the weak convergences (3. 26) and (3. 28) that

$$\liminf_{\lambda \to 0} \int_{t_1}^{t_2} \langle \xi_\lambda(t), u'_\lambda(t) \rangle_V \mathrm{d}t \geq \int_{t_1}^{t_2} \langle \xi(t), u'(t) \rangle_V \mathrm{d}t \tag{3. 42}$$

(it also follows for any $t_1, t_2 \in [0, T]$ with $t_1 \leq t_2$).

Combining these facts and using Proposition 2.3, we deduce that

$$\begin{aligned}
&\limsup_{\lambda_n \to 0} \int_{t_1}^{t_2} \langle \eta_{\lambda_n}(t), J_{\lambda_n} u_{\lambda_n}(t) \rangle_X \mathrm{d}t \\
&\leq \ \varepsilon \langle \xi(t_2), u(t_2) \rangle_V - \varepsilon \langle \xi(t_1), u(t_1) \rangle_V \\
&\qquad - \int_{t_1}^{t_2} \langle \varepsilon \xi(t), u'(t) \rangle_V \mathrm{d}t - \int_{t_1}^{t_2} \langle \xi(t), u(t) \rangle_V \mathrm{d}t \\
&= \ \langle\langle \varepsilon \xi', u \rangle\rangle_{L^m_X \cap L^p_V(t_1, t_2)} - \int_{t_1}^{t_2} \langle \xi(t), u(t) \rangle_V \mathrm{d}t \\
&= \ \int_{t_1}^{t_2} \langle \eta(t), u(t) \rangle_X \mathrm{d}t.
\end{aligned}$$

By exploiting the maximal monotonicity of $\partial_X \phi_X$ in $X \times X^*$ (see also Lemma 1.2 of [6] and Proposition 1.1 of [17]), we conclude that $\eta(t)$ belongs to $\partial_X \phi_X(u(t))$ for a.a. $t \in (t_1, t_2)$. It also follows that

$$\lim_{\lambda_n \to 0} \int_{t_1}^{t_2} \langle \eta_{\lambda_n}(t), J_{\lambda_n} u_{\lambda_n}(t) \rangle_X \mathrm{d}t = \int_{t_1}^{t_2} \langle \eta(t), u(t) \rangle_X \mathrm{d}t. \tag{3. 43}$$

Moreover, from the arbitrariness of $t_1, t_2 \in \mathcal{L}$ and the fact that $(0, T) \setminus \mathcal{L}$ is negligible, we also conclude that $\eta(t) \in \partial_X \phi_X(u(t))$ for a.a. $t \in (0, T)$.

Here let us prove the energy inequality (3. 7). Test $\eta_\lambda(t)$ on $u_\lambda(t)$ and integrate over $(0, T)$. We have

$$\begin{aligned}
\int_0^T \langle \eta_\lambda(t), u_\lambda(t) \rangle_V \mathrm{d}t &= \int_0^T \langle \varepsilon \xi'_\lambda(t), u_\lambda(t) \rangle_V \mathrm{d}t - \int_0^T \langle \xi_\lambda(t), u_\lambda(t) \rangle_V \mathrm{d}t \\
&= -\langle \varepsilon \xi_\lambda(0), u_0 \rangle_V - \int_0^T \langle \varepsilon \xi_\lambda(t), u'_\lambda(t) \rangle_V \mathrm{d}t \\
&\quad - \int_0^T \langle \xi_\lambda(t), u_\lambda(t) \rangle_V \mathrm{d}t.
\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}\int_0^T \langle \eta(t), u(t)\rangle_X \mathrm{d}t &\leq \liminf_{\lambda_n\to 0}\int_0^T \langle \eta_{\lambda_n}(t), J_{\lambda_n}u_{\lambda_n}(t)\rangle_X \mathrm{d}t\\ &\overset{(3.\ 38)}{\leq} \limsup_{\lambda_n\to 0}\int_0^T \langle \eta_{\lambda_n}(t), u_{\lambda_n}(t)\rangle_V \mathrm{d}t\\ &= -\lim_{\lambda_n\to 0}\langle \varepsilon\xi_{\lambda_n}(0), u_0\rangle_V - \liminf_{\lambda_n\to 0}\int_0^T \langle \varepsilon\xi_{\lambda_n}(t), u'_{\lambda_n}(t)\rangle_V \mathrm{d}t\\ &\quad -\lim_{\lambda_n\to 0}\int_0^T \langle \xi_{\lambda_n}(t), u_{\lambda_n}(t)\rangle_V \mathrm{d}t\\ &\overset{(3.\ 42)}{\leq} -\langle \varepsilon\xi(0), u_0\rangle_X - \int_0^T \langle \varepsilon\xi(t), u'(t)\rangle_V \mathrm{d}t\\ &\quad -\int_0^T \langle \xi(t), u(t)\rangle_V \mathrm{d}t,\end{aligned}$$

which leads us to estimate (3. 7).

Let us next check that $\xi(\cdot) = \mathrm{d}_V\psi(u'(\cdot))$ almost everywhere. Let $t_1, t_2 \in \mathcal{L}$ be again fixed and consider the same sequence λ_n as before. For notational simplicity, λ_n will be denoted by λ. We have

$$\begin{aligned}&\int_{t_1}^{t_2} \langle \varepsilon\xi_\lambda(t), u'_\lambda(t)\rangle_V \mathrm{d}t\\ &\overset{(3.\ 40)}{=} \varepsilon\langle \xi_\lambda(t_2), u_\lambda(t_2)\rangle_V - \varepsilon\langle \xi_\lambda(t_1), u_\lambda(t_1)\rangle_V - \int_{t_1}^{t_2} \langle \varepsilon\xi'_\lambda(t), u_\lambda(t)\rangle_V \mathrm{d}t\\ &\overset{(3.\ 9)}{=} \varepsilon\langle \xi_\lambda(t_2), u_\lambda(t_2)\rangle_V - \varepsilon\langle \xi_\lambda(t_1), u_\lambda(t_1)\rangle_V - \int_{t_1}^{t_2} \langle \xi_\lambda(t), u_\lambda(t)\rangle_V \mathrm{d}t\\ &\quad -\int_{t_1}^{t_2} \langle \eta_\lambda(t), u_\lambda(t)\rangle_V \mathrm{d}t\\ &\overset{(3.\ 38)}{\leq} \varepsilon\langle \xi_\lambda(t_2), u_\lambda(t_2)\rangle_V - \varepsilon\langle \xi_\lambda(t_1), u_\lambda(t_1)\rangle_V - \int_{t_1}^{t_2} \langle \xi_\lambda(t), u_\lambda(t)\rangle_V \mathrm{d}t\\ &\quad -\int_{t_1}^{t_2} \langle \eta_\lambda(t), J_\lambda u_\lambda(t)\rangle_X \mathrm{d}t =: \mathrm{RHS}.\end{aligned}$$

Hence by convergences (3. 41), (3. 43), and Proposition 2.3, for $\lambda_n \to 0$ we have

$$\begin{aligned}\mathrm{RHS} &\to \varepsilon\langle \xi(t_2), u(t_2)\rangle_V - \varepsilon\langle \xi(t_1), u(t_1)\rangle_V\\ &\quad -\int_{t_1}^{t_2} \langle \xi(t), u(t)\rangle_V \mathrm{d}t - \int_{t_1}^{t_2} \langle \eta(t), u(t)\rangle_X \mathrm{d}t\\ &= \langle\langle \varepsilon\xi', u\rangle\rangle_{L^m_X\cap L^p_V(t_1,t_2)} + \int_{t_1}^{t_2} \langle \varepsilon\xi(t), u'(t)\rangle_V \mathrm{d}t\\ &\quad -\int_{t_1}^{t_2} \langle \xi(t), u(t)\rangle_V \mathrm{d}t - \int_{t_1}^{t_2} \langle \eta(t), u(t)\rangle_X \mathrm{d}t\\ &\overset{(3.\ 9)}{=} \int_{t_1}^{t_2} \langle \varepsilon\xi(t), u'(t)\rangle_V \mathrm{d}t.\end{aligned}$$

Therefore

$$\limsup_{\lambda \to 0} \int_{t_1}^{t_2} \langle \xi_\lambda(t), u'_\lambda(t) \rangle_V \, \mathrm{d}t \le \int_{t_1}^{t_2} \langle \xi(t), u'(t) \rangle_V \, \mathrm{d}t. \tag{3. 44}$$

Thanks to the demiclosedness of the maximal monotone operator $u \mapsto \mathrm{d}_V\psi(u(\cdot))$ in $L^p(0,T;V) \times L^{p'}(0,T;V^*)$ and Proposition 1.1 of [17], $\xi(t)$ coincides with $\mathrm{d}_V\psi(u'(t))$ for a.a. $t \in (0,T)$, and moreover,

$$\lim_{\lambda \to 0} \int_{t_1}^{t_2} \langle \xi_\lambda(t), u'_\lambda(t) \rangle_V \mathrm{d}t = \int_{t_1}^{t_2} \langle \xi(t), u'(t) \rangle_V \mathrm{d}t \quad \text{for all } t_1, t_2 \in \mathcal{L}. \tag{3. 45}$$

As in (3. 21), we can derive by (3. 32) and (3. 45) that

$$\int_{t_1}^{t_2} \phi(u(t)) \mathrm{d}t \le (\phi(u_0) + C'_2 T + \varepsilon\psi(0))\,(t_2 - t_1) + \varepsilon \int_{t_1}^{t_2} \langle \xi(t), u'(t) \rangle_V \mathrm{d}t$$

for all $t_1, t_2 \in \mathcal{L}$ with $t_1 \le t_2$ Letting $(t_1, t_2) \to (0, T)$, we obtain (3. 6).

By the weak lower semicontinuity of ϕ in V, it follows from convergence (3. 34) that

$$\liminf_{\lambda \to 0} \phi_\lambda(u_\lambda(T)) \ge \liminf_{\lambda \to 0} \phi(J_\lambda u_\lambda(T)) \ge \phi(u(T)).$$

Hence combining (3. 17) with (3. 42), we can derive (3. 8). This completes a proof of Theorem 3.3.

4 Minimizers of WED functionals

In this short section, we are concerned with the existence and characterization of minimizers of the WED functional I_ε in $\mathcal{V} := L^p(0,T;V)$. Our aim is to prove that every minimizer u_ε of I_ε coincides with a strong solution of (1. 2)–(1. 4) which is a limit of global minimizers $u_{\varepsilon,\lambda}$ for $I_{\varepsilon,\lambda}$ as $\lambda \to 0$, provided that either ψ or ϕ is strictly convex.

Let us start by defining the minimizers of I_ε as follows.

Definition 4.1 (Minimizer)**.** *A function $u \in \mathcal{V}$ is said to be a minimizer of I_ε in $\mathcal{V}$ if $\partial_{\mathcal{V}} I_\varepsilon(u) \ni 0$.*

The main result of this section is the following.

Theorem 4.2 (Existence and characterization of minimizers)**.** *Assume* (A1)–(A4)*. For each $u_0 \in D(\phi)$, the strong solution of* (1. 2)–(1. 4) *obtained in Theorem* 3.3 *is a minimizer of I_ε in $\mathcal{V}$. Moreover, if either ψ or ϕ is strictly convex, then the minimizer is unique.*

Our proof of this theorem is divided into the following two lemmas.

Lemma 4.3 (Strong solutions are minimizers)**.** *Let u_ε be a strong solution of* (1. 2)–(1. 4) *obtained in Theorem* 3.3*. Then, u_ε is a minimizer of I_ε in $\mathcal{V}$.*

Proof. By Lemma 3.4, we have obtained a global minimizer $u_{\varepsilon,\lambda} \in \mathcal{V}$ of $I_{\varepsilon,\lambda}$, namely,

$$I_{\varepsilon,\lambda}(v) \geq I_{\varepsilon,\lambda}(u_{\varepsilon,\lambda}) \quad \text{for all} \quad v \in D(I_\varepsilon).$$

By passing to the limit as $\lambda \to 0$ and using dominated convergence, we get

$$I_{\varepsilon,\lambda}(v) \to I_\varepsilon(v).$$

Moreover, by the weak lower semicontinuity of $I_\varepsilon^1, I_\varepsilon^2$ in $\mathcal{V}$, we also deduce from the convergences (3. 26) and (3. 32) that

$$\begin{aligned}
\liminf_{\lambda\to 0} I_{\varepsilon,\lambda}(u_{\varepsilon,\lambda}) &= \liminf_{\lambda\to 0} \int_0^T e^{-t/\varepsilon}\left(\psi(u'_{\varepsilon,\lambda}(t)) + \frac{1}{\varepsilon}\phi_\lambda(u_{\varepsilon,\lambda}(t))\right) \mathrm{d}t \\
&\geq \liminf_{\lambda\to 0}\left(I_\varepsilon^1(u_{\varepsilon,\lambda}) + I_\varepsilon^2(J_\lambda u_{\varepsilon,\lambda})\right) \\
&\geq I_\varepsilon^1(u_\varepsilon) + I_\varepsilon^2(u_\varepsilon) = I_\varepsilon(u_\varepsilon).
\end{aligned}$$

Therefore $I_\varepsilon(v) \geq I_\varepsilon(u_\varepsilon)$ for all $v \in D(I_\varepsilon)$, namely $0 \in \partial_{\mathcal{V}} I_\varepsilon(u_\varepsilon)$. □

Lemma 4.4 (Minimizers are unique)**.** *Suppose that either ψ or ϕ is strictly convex in V. Then, for each $\varepsilon > 0$, I_ε admits a unique minimizer.*

Proof. In both cases the functional I_ε turns out to be strictly convex in $\mathcal{V}$ and the assertion follows. □

5 The causal limit

In this section we ascertain the fundamental issue of the WED approach. Namely, we prove that the minimizers u_ε of the WED functionals I_ε converge as $\varepsilon \to 0$. Our main result is the following.

Theorem 5.1 (Causal limit)**.** *Assume* (A1)–(A4) *and that either ψ or ϕ is strictly convex. Let $u_0 \in D(\phi)$ and let u_ε be a minimizer of I_ε on $\mathcal{V} := L^p(0,T;V)$. Then, there exist a sequence $\varepsilon_n \to 0$ and a limit u such that*

$$\begin{aligned}
u_{\varepsilon_n} \to u \qquad & \textit{strongly in } C([0,T];V), \\
& \textit{weakly in } W^{1,p}(0,T;V) \cap L^m(0,T;X),
\end{aligned}$$

and u is a strong solution of (1. 5)–(1. 6)*.*

Proof. For each $\varepsilon > 0$, let u_ε be the unique minimizer of I_ε on $\mathcal{V}$. By Theorem 4.2, u_ε is a strong solution of (1. 2)–(1. 4) satisfying estimates (3. 5)–(3. 8).

Since $u_\varepsilon(0) = u_0$, it follows from estimate (3. 5) that

$$\sup_{t\in[0,T]} |u_\varepsilon(t)|_V \leq C. \tag{5. 1}$$

Furthermore, by assumption (A2),

$$\int_0^T |\xi_\varepsilon(t)|_{V^*}^{p'} \mathrm{d}t \le C. \tag{5. 2}$$

Hence, by taking a suitable (non relabelled) sequence $\varepsilon \to 0$,

$$u_\varepsilon \to u \qquad \text{weakly in } W^{1,p}(0,T;V), \tag{5. 3}$$

$$\xi_\varepsilon \to \xi \qquad \text{weakly in } L^{p'}(0,T;V^*). \tag{5. 4}$$

Combining the bounds (3. 6), (3. 5), and (5. 2), we deduce from assumption (A3) that

$$\int_0^T |u_\varepsilon(t)|_X^m \mathrm{d}t \le C. \tag{5. 5}$$

Hence, one has

$$u_\varepsilon \to u \quad \text{weakly in } L^m(0,T;X).$$

Moreover, it classically follows from estimates (3. 5) and (5. 5) that

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

which also implies

$$u_\varepsilon(t) \to u(t) \quad \text{strongly in } V \text{ for all } \ t \in [0,T] \tag{5. 6}$$

and $u(0) = u_0$.

By assumption (A4) together with the bounds in (5. 1) and (5. 5), we have

$$\int_0^T |\eta_\varepsilon(t)|_{X^*}^{m'} \mathrm{d}t \le C, \tag{5. 7}$$

which implies

$$\eta_\varepsilon \to \eta \quad \text{weakly in } L^{m'}(0,T;X^*). \tag{5. 8}$$

By using the equation (1. 2) along with the estimates (5. 2) and (5. 7), we find that

$$\|\varepsilon\xi_\varepsilon'\|_{L^{m'}_{X^*}+L^{p'}_{V^*}} \le C. \tag{5. 9}$$

Thus

$$\varepsilon\xi_\varepsilon' \to 0 \quad \text{weakly in } L^{m'}(0,T;X^*) + L^{p'}(0,T;V^*). \tag{5. 10}$$

Hence

$$\xi + \eta = 0. \tag{5. 11}$$

Moreover, for each $v \in X$, it follows from the final condition (1. 4) and the latter convergence that

$$\langle \varepsilon\xi_\varepsilon(t), v\rangle_X = \left\langle \int_T^t \varepsilon\xi_\varepsilon'(t)\mathrm{d}t, v \right\rangle_X = \int_T^t \langle \varepsilon\xi_\varepsilon'(t), v\rangle_X \mathrm{d}t \to 0,$$

which leads us to

$$\varepsilon\xi_\varepsilon(t) \to 0 \quad \text{weakly in } X^* \text{ for each } t \in [0,T]. \tag{5. 12}$$

We next claim that $\eta(t) \in \partial_X \phi_X(u(t))$ for almost all $t \in (0,T)$. Indeed, by estimate (3. 7),

$$\begin{aligned}\int_0^T \langle \eta_\varepsilon(t), u_\varepsilon(t)\rangle_X \mathrm{d}t &\le -\langle \varepsilon\xi_\varepsilon(0), u_0\rangle_X - \int_0^T \langle \varepsilon\xi_\varepsilon(t), u_\varepsilon'(t)\rangle_V \mathrm{d}t \\ &\quad - \int_0^T \langle \xi_\varepsilon(t), u_\varepsilon(t)\rangle_V \mathrm{d}t \\ &\to -\int_0^T \langle \xi(t), u(t)\rangle_V \mathrm{d}t.\end{aligned}$$

Hence, we have

$$\limsup_{\varepsilon\to 0} \int_0^T \langle \eta_\varepsilon(t), u_\varepsilon(t)\rangle_X \mathrm{d}t \le -\int_0^T \langle \xi(t), u(t)\rangle_X \mathrm{d}t = \int_0^T \langle \eta(t), u(t)\rangle_X \mathrm{d}t.$$

Therefore, by using the demiclosedness of the maximal monotone operator $\partial_X \phi_X$ in $L^m(0,T;X) \times L^{m'}(0,T;X^*)$ and applying Proposition 1.1 of [17], we conclude that $\eta(t) \in \partial_X \phi_X(u(t))$ for almost all $t \in (0,T)$. Furthermore, since $\xi \in L^{p'}(0,T;V^*)$, by Proposition 2.1, we have $\eta(t) \in \partial_V \phi(u(t))$ for almost every $t \in (0,T)$.

Let us now check that $\xi(t) = \mathrm{d}_V\psi(u'(t))$ for almost every $t \in (0,T)$. By passing to the lim sup as $\varepsilon \to 0$ into estimate (3. 8) with the aid of the strong convergence (5. 6) and the lower semicontinuity of ϕ in the weak topology of V, we obtain

$$\begin{aligned}\limsup_{\varepsilon\to 0} \int_0^T \langle \xi_\varepsilon(t), u_\varepsilon'(t)\rangle_V \mathrm{d}t &\le -\liminf_{\varepsilon\to 0} \phi(u_\varepsilon(T)) + \phi(u_0) \le -\phi(u(T)) + \phi(u_0) \\ &= \int_0^T \langle -\eta(t), u'(t)\rangle_V \mathrm{d}t \overset{(5.\ 11)}{=} \int_0^T \langle \xi(t), u'(t)\rangle_V \mathrm{d}t.\end{aligned}$$

Thus, we have $\xi(t) = \mathrm{d}_V\psi(u'(t))$ for almost every $t \in (0,T)$. Consequently, u solves the limiting problem (1. 5)–(1. 6) on $[0,T]$. □

Remark 5.2. If one is interested in proving the convergence of strong solutions u_ε of (1. 2)–(1. 4) satisfying energy inequalities (3. 5)–(3. 8) as $\varepsilon \to 0$ to a strong solution of (1. 5)–(1. 6), the strict convexity of ϕ and ψ need not be assumed

6 Mosco-convergence of WED functionals

We shall prepare here some convergence result for sequences of WED functionals at fixed level ε. In particular, we present sufficient conditions for the convergence as $h \to \infty$ of the sequence of WED functionals $I_{\varepsilon,h}$ given by

$$I_{\varepsilon,h}(u) = \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon}\Big(\psi_h(u'(t)) + \frac{1}{\varepsilon}\phi_h(u(t))\Big)\mathrm{d}t & \text{if } u \in W^{1,p}(0,T;V) \cap L^m(0,T;X), \\ & \quad \text{and } u(0) = u_{0,h}, \\ \infty & \text{otherwise} \end{cases}$$

with initial data $u_{0,h} \in X$, a Gâteaux differentiable convex functional $\psi_h : V \to [0,\infty)$ and a proper, lower semicontinuous convex functional $\phi_h : V \to [0,\infty]$ for $h \in \mathbb{N}$.

Throughout this section, we assume with no further specific mention that the functionals ψ_h and ϕ_h fulfill the general assumptions (A1)–(A4) with constants independent of h. In particular, by letting

$$\mathcal{Z} := L^m(0,T;X) \cap W^{1,p}(0,T;V),$$

we easily check that $I_{\varepsilon,h}$ are bounded from below in $\mathcal{Z}$ uniformly for $h \in \mathbb{N}$. Hence, the global minimizers u_h of $I_{\varepsilon,h}$ are bounded in $\mathcal{Z}$ for all $h \in \mathbb{N}$.

Let us now make precise our notion of functional convergence in the following.

Definition 6.1 (Mosco-convergence in $\mathcal{Z}$)**.** *The functional $I_{\varepsilon,h}$ is said to* Mosco-converge *to I_ε in $\mathcal{Z}$ as $h \to \infty$ if the following two conditions hold*:

(i) (Liminf condition) *Let $u_h \to u$ weakly in $\mathcal{Z}$ as $h \to \infty$. Then*

$$\liminf_{h\to\infty} I_{\varepsilon,h}(u_h) \geq I_\varepsilon(u).$$

(ii) (Existence of recovery sequences) *For every $u \in \mathcal{Z}$ and sequence $k_h \to \infty$ in $\mathbb{N}$, there exist a subsequence (k'_h) of (k_h) and a recovery sequence (u_h) in $\mathcal{Z}$ such that*

$$u_h \to u \quad \text{strongly in } \mathcal{Z} \quad \text{and} \quad I_{\varepsilon,k'_h}(u_h) \to I_\varepsilon(u) \quad \text{as} \;\; h \to \infty.$$

Note that Mosco-convergence is classical (see [3], [23]), and corresponds to the usual notion of Γ-convergence with respect to both the strong and the weak topology in $\mathcal{Z}$. Mosco-convergence of the driving functionals arises as the natural requirement in order to deduce the convergence of the related differential problems (see [3, Thm. 3.74(2), p. 388] for gradient flows and [28, Lemma 7.1] for doubly nonlinear evolutions).

Our sufficient conditions for Mosco-convergence are stated in the following.

(H0) (Separability of spaces) V and X are separable.

(H1) (Liminf condition for ϕ_h in X) Let (u_h) be a sequence in X such that $u_h \to u$ weakly in X. Then
$$\liminf_{h\to\infty} \phi_h(u_h) \geq \phi(u).$$

(H2) (Liminf condition for ψ_h in V) Let (u_h) be a sequence in V such that $u_h \to u$ weakly in V. Then
$$\liminf_{h\to\infty} \psi_h(u_h) \geq \psi(u).$$

(H3) (Existence of joint recovery sequences for ϕ_h and ψ_h in X) Let (k_h) be a sequence in $\mathbb{N}$ such that $k_h \to \infty$. Let (u_h) be a sequence in X such that

$$u_h \to u \quad \text{strongly in } X \quad \text{and} \quad \phi_{k_h}(u_h) \to \phi(u) \quad \text{as} \;\; h \to \infty.$$

Then, for every $v \in X$, $\tau > 0$, there exists a sequence $(v_{\tau,h})$ in X such that

$$
\begin{aligned}
&v_{\tau,h} \to v \quad \text{strongly in } X,\\
&\psi_{k_h}\left(\frac{v_{\tau,h} - u_h}{\tau}\right) \to \psi\left(\frac{v-u}{\tau}\right), \quad \phi_{k_h}(v_{\tau,h}) \to \phi(v) \quad \text{as } h \to \infty.
\end{aligned}
$$

(H4) (Convergence of initial data) $u_{0,h} \in X$, $u_{0,h} \to u_0$ strongly in X and $\phi_h(u_{0,h}) \to \phi(u_0)$ as $h \to \infty$.

The reader should notice that we are not requiring for the *separate* functional convergence $\phi_h \to \psi$ and $\phi_h \to \phi$ here (Γ- or Mosco-). In particular, our proof makes a crucial use of the possibility of finding a *joint recovery sequence* as of assumption (H3). Let us comment that the occurrence of such joint condition is not at all unexpected. Indeed, a similar joint recovery condition has been proved to be necessary and sufficient for passing to the limit in sequences of rate-independent evolution problems in an *energetic* form in [20], namely for $p = 1$. Moreover, let us note that in case $p = 2$, the concrete construction of an analogous joint recovery sequence is at the basis of the relaxation proof in [10].

The main result of this section is stated as follows.

Theorem 6.2 (Mosco-convergence of $I_{\varepsilon,h}$)**.** *Assume* (H0)–(H4)*. Then, the functionals $I_{\varepsilon,h}$ Mosco-converge in $\mathcal{Z}$ to I_ε as $h \to \infty$.*

We shall provide a proof of this theorem in the next subsection. Still, let us first point out a corollary, whose immediate proof is omitted.

Corollary 6.3 (Minimizers converge to a minimizer)**.** *Under the assumptions of Theorem* 6.2*, let u_h be a global minimizer of $I_{\varepsilon,h}$ for $h \in \mathbb{N}$ such that $u_{k_h} \to u$ weakly in $\mathcal{Z}$ as $h \to \infty$ along with some sequence $k_h \to \infty$ in $\mathbb{N}$. Then u minimizes I_ε.*

6.1 Proof of Theorem 6.2

We provide here a proof of Theorem 6.2 by establishing conditions (i) and (ii) of Definition 6.1. Condition (ii) of Definition 6.1 is proved in a smooth case first and then generalized.

6.1.1 Liminf inequality

By using Corollary 4.4 of [28], we can derive from (H0)–(H2) that

$$\liminf_{h\to\infty} \int_0^T e^{-t/\varepsilon}\phi_h(u_h(t))\mathrm{d}t \geq \int_0^T e^{-t/\varepsilon}\phi(u(t))\mathrm{d}t$$

if $u_h \to u$ weakly in $L^m(0,T;X)$;

$$\liminf_{h\to\infty} \int_0^T e^{-t/\varepsilon}\psi_h(u_h'(t))\mathrm{d}t \geq \int_0^T e^{-t/\varepsilon}\psi(u'(t))\mathrm{d}t$$

if $u_h' \to u'$ weakly in $L^p(0,T;V)$. Let (u_h) be a sequence in $D(I_{\varepsilon,h})$ such that $u_h \to u$ weakly in $\mathcal{Z}$. Then we can take a subsequence (k_h) of (h) such that $u_{k_h} \to u$ strongly in $C([0,T];V)$ by the compact embedding $X \hookrightarrow V$, and therefore, $u(0) = u_0$ by (H4). It follows that

$$\liminf_{h\to\infty} I_{\varepsilon,h}(u_h) \geq I_\varepsilon(u).$$

Thus the Liminf condition (i) follows.

6.1.2 Recovery sequence for $u \in C^1([0,T];X)$

Let us next prove the existence of recovery sequences of $u \in D(I_\varepsilon)$ for $I_{\varepsilon,h}$. We first treat the case that $u \in C^1([0,T];X)$ and $u(0) = u_0$, which also leads us to $u \in D(I_\varepsilon)$. Our recovery sequence will be constructed from an approximation of u. Let $N \in \mathbb{N}$ be fixed and set $\tau := T/N$, $u_\tau^i := u(i\tau)$ and $t^i := \tau i$ for $i = 0, 1, \ldots, N$ (i.e., $t^0 = 0$ and $t^N = T$). Define the piecewise linear interpolant $\hat{u}_\tau \in D(I_\varepsilon)$ by

$$\hat{u}_\tau(t) = \alpha_\tau^i(t) u_\tau^i + (1 - \alpha_\tau^i(t)) u_\tau^{i+1} \quad \text{for } \ t \in [t^i, t^{i+1}),$$

where $\alpha_\tau^i(t) := (t^{i+1} - t)/\tau$, and a piecewise forward constant interpolant $\bar{u}_\tau \in L^\infty(0,T;X)$ by

$$\bar{u}_\tau(t) = u_\tau^{i+1} \quad \text{for } \ t \in [t^i, t^{i+1}).$$

As $u \in C^1([0,T];X)$, it follows that

$$\hat{u}_\tau \to u \qquad \text{strongly in } W^{1,\infty}(0,T;X), \tag{6. 1}$$

$$\bar{u}_\tau \to u \qquad \text{strongly in } L^\infty(0,T;X). \tag{6. 2}$$

Now, we find that

$$\begin{aligned} I_\varepsilon(u) &= \int_0^T e^{-t/\varepsilon} \left(\psi(u'(t)) + \frac{1}{\varepsilon}\phi(u(t)) \right) \mathrm{d}t \\ &= \int_0^T e^{-t/\varepsilon} \left(\psi(\hat{u}_\tau'(t)) + \frac{1}{\varepsilon}\phi(\bar{u}_\tau(t)) \right) \mathrm{d}t \\ &\quad + \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon} \Big(\phi(u(t)) - \phi(\bar{u}_\tau(t)) \Big) \mathrm{d}t + \int_0^T e^{-t/\varepsilon} \Big(\psi(u'(t)) - \psi(\hat{u}_\tau'(t)) \Big) \mathrm{d}t \\ &=: \mathcal{I}_{1,\tau} + \mathcal{I}_{2,\tau} + \mathcal{I}_{3,\tau}. \end{aligned}$$

Since ϕ is convex, letting $\eta(t) \in \partial_X \phi_X(u(t))$ and $\bar{\eta}_\tau(t) \in \partial_X \phi_X(\bar{u}_\tau(t))$, we can exploit (A4) and convergence (6. 2) in order to check that

$$\begin{aligned} \mathcal{I}_{2,\tau} &\leq \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon} \langle \eta(t), u(t) - \bar{u}_\tau(t) \rangle_X \, \mathrm{d}t \\ &\leq \frac{C}{\varepsilon} \left(\int_0^T |u(t)|_X^m \mathrm{d}t + T \right)^{1/m'} \|u - \bar{u}_\tau\|_{L^m(0,T;X)} \to 0 \quad \text{ as } \tau \to 0 \end{aligned}$$

with $C = \sup_{t\in[0,T]} \ell_2(|u(t)|_V)^{1/m'}$ and

$$\begin{aligned}\mathcal{I}_{2,\tau} &\geq \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon} \langle \bar{\eta}_\tau(t), u(t) - \bar{u}_\tau(t)\rangle_X \, \mathrm{d}t \\ &\geq -\frac{C_\tau}{\varepsilon}\left(\int_0^T |\bar{u}_\tau(t)|_X^m \mathrm{d}t + T\right)^{1/m'} \|u - \bar{u}_\tau\|_{L^m(0,T;X)} \to 0 \quad \text{as } \tau \to 0\end{aligned}$$

with $C_\tau = \sup_{t\in[0,T]} \ell_2(|\bar{u}_\tau(t)|_V)^{1/m'}$, which is bounded as $\tau \to 0$ by (6. 2). Hence $\mathcal{I}_{2,\tau} = o(1; \tau \to 0)$, where we wrote $o(1;\tau \to 0)$ instead of $o(1)$ in order to enforce which parameter is supposed to be infinitesimal.

Moreover, by letting $\xi(t) = \mathrm{d}_V\psi(u'(t))$ and $\bar{\xi}_\tau(t) = \mathrm{d}_V\psi(\hat{u}_\tau'(t))$, we use (A2) and convergence (6. 1) in such a way that

$$\begin{aligned}\mathcal{I}_{3,\tau} &\leq \int_0^T e^{-t/\varepsilon} \langle \xi(t), u'(t) - \hat{u}_\tau'(t)\rangle_V \, \mathrm{d}t \\ &\leq \left(C_3 \int_0^T |u'(t)|_V^p \mathrm{d}t + C_4 T\right)^{1/p'} \|u' - \hat{u}_\tau'\|_{L^p(0,T;V)} \to 0 \quad \text{as } \tau \to 0\end{aligned}$$

and

$$\begin{aligned}\mathcal{I}_{3,\tau} &\geq \int_0^T e^{-t/\varepsilon} \left\langle \bar{\xi}_\tau(t), u'(t) - \hat{u}_\tau'(t)\right\rangle_V \, \mathrm{d}t \\ &\geq -\left(C_3 \int_0^T |\hat{u}_\tau'(t)|_V^p \mathrm{d}t + C_4 T\right)^{1/p'} \|u' - \hat{u}_\tau'\|_{L^p(0,T;V)} \to 0 \quad \text{as } \tau \to 0.\end{aligned}$$

Thus, we observe that $\mathcal{I}_{3,\tau} = o(1; \tau \to 0)$, and therefore

$$I_\varepsilon(u) = \mathcal{I}_{1,\tau} + o(1; \tau \to 0). \tag{6. 3}$$

For $\tau > 0$, let us define a difference operator δ_τ by

$$\delta_\tau \chi^{i+1} := \frac{\chi^{i+1} - \chi^i}{\tau} \qquad \text{for a vector } \{\chi^i\}_{i=0,1,\ldots,N}.$$

Then $\mathcal{I}_{1,\tau}$ can be written as follows:

$$\begin{aligned}\mathcal{I}_{1,\tau} &= \sum_{i=0}^{N-1} \int_{t^i}^{t^{i+1}} e^{-t/\varepsilon} \left(\psi(\hat{u}_\tau'(t)) + \frac{1}{\varepsilon}\phi(\bar{u}_\tau(t))\right) \mathrm{d}t \\ &= \sum_{i=0}^{N-1} \left(\int_{t^i}^{t^{i+1}} e^{-t/\varepsilon} \mathrm{d}t\right) \left(\psi(\delta_\tau u_\tau^{i+1}) + \frac{1}{\varepsilon}\phi(u_\tau^{i+1})\right). \end{aligned} \tag{6. 4}$$

Let (k_h) be a sequence in $\mathbb{N}$ such that $k_h \to \infty$. Set $u_{\tau,h}^0 := u_{0,k_h}$. Then by (H3) and (H4), we can take a sequence $(u_{\tau,h}^1)$ in X such that

$$\begin{aligned}&u_{\tau,h}^1 \to u_\tau^1 \quad \text{strongly in } X, \\ &\psi_{k_h}\left(\frac{u_{\tau,h}^1 - u_{\tau,h}^0}{\tau}\right) \to \psi\left(\frac{u_\tau^1 - u_\tau^0}{\tau}\right), \quad \phi_{k_h}(u_{\tau,h}^1) \to \phi(u_\tau^1).\end{aligned}$$

Hence iterating this process $(N-1)$ times, we can further obtain $(u^i_{\tau,h})$ in X for $i = 2, 3, \ldots, N$ such that

$$u^i_{\tau,h} \to u^i_\tau \quad \text{strongly in } X, \tag{6. 5}$$

$$\psi_{k_h}\left(\frac{u^i_{\tau,h} - u^{i-1}_{\tau,h}}{\tau}\right) \to \psi\left(\frac{u^i_\tau - u^{i-1}_\tau}{\tau}\right), \quad \phi_{k_h}(u^i_{\tau,h}) \to \phi(u^i_\tau). \tag{6. 6}$$

Define the piecewise linear interpolant $\hat{u}_{\tau,h} \in D(I_{\varepsilon,h})$ and the piecewise forward constant interpolant $\bar{u}_{\tau,h} \in L^\infty(0,T;X)$ as above and, by convergence (6. 5), we get

$$\hat{u}_{\tau,h} \to \hat{u}_\tau \quad \text{strongly in } W^{1,\infty}(0,T;X) \quad \text{as } h \to \infty. \tag{6. 7}$$

Therefore, for each $\tau > 0$ we can choose $h_\tau \in \mathbb{N}$ such that

$$\|\hat{u}_{\tau,h_\tau} - \hat{u}_\tau\|_{L^m_X \cap W^{1,p}_V} < \tau \quad \text{and} \quad h_\tau > \tau^{-1}.$$

Combining this fact with convergence (6. 1), we also deduce that

$$\begin{aligned} \|\hat{u}_{\tau,h_\tau} - u\|_{L^m_X \cap W^{1,p}_V} &\leq \|\hat{u}_{\tau,h_\tau} - \hat{u}_\tau\|_{L^m_X \cap W^{1,p}_V} + o(1; \tau \to 0) \\ &\leq \tau + o(1; \tau \to 0), \end{aligned}$$

which implies

$$\hat{u}_{\tau,h_\tau} \to u \quad \text{strongly in } \mathcal{Z} \quad \text{as } \tau \to 0.$$

As for the convergence of $I_{\varepsilon,k_h}(\hat{u}_{\tau,h})$, we calculate

$$\begin{aligned} I_{\varepsilon,k_h}(\hat{u}_{\tau,h}) &= \int_0^T e^{-t/\varepsilon}\left(\psi_{k_h}(\hat{u}'_{\tau,h}(t)) + \frac{1}{\varepsilon}\phi_{k_h}(\bar{u}_{\tau,h}(t))\right)\mathrm{d}t \\ &\quad + \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}\Big(\phi_{k_h}(\hat{u}_{\tau,h}(t)) - \phi_{k_h}(\bar{u}_{\tau,h}(t))\Big)\mathrm{d}t \\ &= \sum_{i=0}^{N-1}\left(\int_{t^i}^{t^{i+1}} e^{-t/\varepsilon}\mathrm{d}t\right)\left(\psi_{k_h}(\delta_\tau u^{i+1}_{\tau,h}) + \frac{1}{\varepsilon}\phi_{k_h}(u^{i+1}_{\tau,h})\right) \\ &\quad + \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}\Big(\phi_{k_h}(\hat{u}_{\tau,h}(t)) - \phi_{k_h}(\bar{u}_{\tau,h}(t))\Big)\mathrm{d}t \\ &= \mathcal{I}_{1,\tau,h} + \mathcal{I}_{2,\tau,h}. \end{aligned} \tag{6. 8}$$

Then, by the above-stated convergences (6. 6) and (6. 4),

$$\mathcal{I}_{1,\tau,h} \to \sum_{i=0}^{N-1}\left(\int_{t^i}^{t^{i+1}} e^{-t/\varepsilon}\mathrm{d}t\right)\left(\psi(\delta_\tau u^{i+1}_\tau) + \frac{1}{\varepsilon}\phi(u^{i+1}_\tau)\right) = \mathcal{I}_{1,\tau} \quad \text{as } h \to \infty. \tag{6. 9}$$

Hence, it remains to handle $\mathcal{I}_{2,\tau,h}$.

From the convexity of ϕ_h,

$$\begin{aligned}
\mathcal{I}_{2,\tau,h} &= \sum_{i=0}^{N-1}\int_{t^i}^{t^{i+1}}\frac{e^{-t/\varepsilon}}{\varepsilon}\Big(\phi_{k_h}\big(\alpha_\tau^i(t)u_{\tau,h}^i+(1-\alpha_\tau^i(t))u_{\tau,h}^{i+1}\big)-\phi_{k_h}(u_{\tau,h}^{i+1})\Big)\mathrm{d}t\\
&\le \sum_{i=0}^{N-1}\int_{t^i}^{t^{i+1}}\frac{e^{-t/\varepsilon}}{\varepsilon}\Big(\alpha_\tau^i(t)\phi_{k_h}(u_{\tau,h}^i)+(1-\alpha_\tau^i(t))\phi_{k_h}(u_{\tau,h}^{i+1})-\phi_{k_h}(u_{\tau,h}^{i+1})\Big)\mathrm{d}t\\
&= \sum_{i=0}^{N-1}\left(\int_{t^i}^{t^{i+1}}\frac{e^{-t/\varepsilon}}{\varepsilon}\alpha_\tau^i(t)\mathrm{d}t\right)\Big(\phi_{k_h}(u_{\tau,h}^i)-\phi_{k_h}(u_{\tau,h}^{i+1})\Big).
\end{aligned}$$

Here, again by (6. 6), we get

$$\begin{aligned}
&\sum_{i=0}^{N-1}\left(\int_{t^i}^{t^{i+1}}\frac{e^{-t/\varepsilon}}{\varepsilon}\alpha_\tau^i(t)\mathrm{d}t\right)\Big(\phi_{k_h}(u_{\tau,h}^i)-\phi_{k_h}(u_{\tau,h}^{i+1})\Big)\\
&\quad= o_\tau(1;h\to\infty)+\sum_{i=0}^{N-1}\left(\int_{t^i}^{t^{i+1}}\frac{e^{-t/\varepsilon}}{\varepsilon}\alpha_\tau^i(t)\mathrm{d}t\right)\Big(\phi(u_\tau^i)-\phi(u_\tau^{i+1})\Big).
\end{aligned}$$

Set $\eta_\tau^i\in\partial_X\phi_X(u_\tau^i)$ for $i=0,1,\dots,N$. Then, noticing that

$$\phi(u_\tau^i)-\phi(u_\tau^{i+1})\le\left\langle\eta_\tau^i,u_\tau^i-u_\tau^{i+1}\right\rangle_X\le|\eta_\tau^i|_{X^*}|u_\tau^i-u_\tau^{i+1}|_X,$$

by assumption (A4) and the strong convergence (6. 2), we obtain

$$\begin{aligned}
&\sum_{i=0}^{N-1}\left(\int_{t^i}^{t^{i+1}}\frac{e^{-t/\varepsilon}}{\varepsilon}\alpha_\tau^i(t)\mathrm{d}t\right)\Big(\phi(u_\tau^i)-\phi(u_\tau^{i+1})\Big)\\
&\quad\le \frac{1}{\varepsilon}\sum_{i=0}^{N-1}\frac{\tau}{2}|\eta_\tau^i|_{X^*}|u_\tau^i-u_\tau^{i+1}|_X\\
&\quad\le C\left(\int_\tau^T|\bar{u}_\tau(t)-\bar{u}_\tau(t-\tau)|_X\mathrm{d}t+\tau|u_\tau^1-u_\tau^0|_X\right)\to 0\quad\text{as }\ \tau\to 0.
\end{aligned}$$

Along these very same lines, it is possible to deduce an analogous estimate from below and we conclude that

$$\mathcal{I}_{2,\tau,h}\le o_\tau(1;h\to\infty)+o(1;\tau\to 0).$$

Combining now the latter with the convergence (6. 9) and the decomposition (6. 8), we deduce that

$$I_{\varepsilon,k_h}(\hat{u}_{\tau,h})=\mathcal{I}_{1,\tau,h}+\mathcal{I}_{2,\tau,h}\le I_\varepsilon(u)+o_\tau(1;h\to\infty)+o(1;\tau\to 0).$$

Hence, for each $\tau>0$ we can extract a sequence $h_\tau\in\mathbb{N}$ such that

$$I_{\varepsilon,k_{h_\tau}}(\hat{u}_{\tau,h_\tau})\le I_\varepsilon(u)+\tau+o(1;\tau\to 0)\quad\text{and}\quad h_\tau>\tau^{-1}.$$

Thus, one has

$$\limsup_{\tau\to 0}I_{\varepsilon,k_{h_\tau}}(\hat{u}_{\tau,h_\tau})\le I_\varepsilon(u).$$

6.1.3 Recovery sequence for general u

Let us now discuss the general case $u \in D(I_\varepsilon)$, i.e., $u \in \mathcal{Z}$ and $u(0) = u_0$. Let $v(t) := u(t) - u_0$ for $t \in [0,T]$. By using a standard mollification argument, we can construct $v_n \in C^1([0,T];X)$ for all $n \in \mathbb{N}$ such that

$$v_n \to v \quad \text{strongly in } \mathcal{Z} \text{ as } n \to \infty \quad \text{and} \quad v_n(0) = 0.$$

Now let $w_n(t) := v_n(t) + u_0$ for $t \in [0,T]$. Then $w_n \in C^1([0,T];X)$ satisfies

$$w_n \to u \quad \text{strongly in } \mathcal{Z} \quad \text{and} \quad w_n(0) = u_0.$$

By virtue of assumptions (A2) and (A4), the functions $u \mapsto J(u)$ and $u \mapsto I^2_\varepsilon(u)$ are continuous in $W^{1,p}(0,T;V)$ and $L^m(0,T;X)$, respectively (see Subsection 2.4). Hence, by relabeling the sequence (w_n) and using the continuity of ϕ_X and ψ in X and V, respectively, we can say

$$\|w_n - u\|_{L^m_X \cap W^{1,p}_V} < \frac{1}{n} \quad \text{and} \quad |I_\varepsilon(w_n) - I_\varepsilon(u)| < \frac{1}{n}.$$

Now, by the above-proved existence of a recovery sequence in the smooth case, for each $n \in \mathbb{N}$, we can take a subsequence (k^n_h) of (k_h) and a sequence (u^n_h) in $\mathcal{Z}$ such that

$$u^n_h \to w_n \text{ strongly in } \mathcal{Z} \quad \text{and} \quad I_{\varepsilon,k^n_h}(u^n_h) \to I_\varepsilon(w_n) \quad \text{as } h \to \infty$$

at each $n \in \mathbb{N}$. Finally, by using a diagonal argument, we can choose a sequence (u_n) in $\mathcal{Z}$ and subsequence (k'_n) of (k_h) such that

$$u_n \to u \quad \text{strongly in } \mathcal{Z} \quad \text{and} \quad I_{\varepsilon,k'_n}(u_n) \to I_\varepsilon(u) \quad \text{as } n \to \infty.$$

Consequently, the recovery sequence condition (ii) of Definition 6.1 holds.

7 Applications

In this section, we present doubly nonlinear problems and apply the above-detailed abstract theory to them. Let Ω be a bounded domain of $\mathbb{R}^N$ with smooth boundary $\partial\Omega$. We start with the following doubly nonlinear parabolic equation (DNP):

$$\alpha(u_t) - \Delta^a_m u = 0 \quad \text{in} \quad \Omega \times (0,T), \tag{7. 1}$$

$$u = 0 \quad \text{on} \quad \partial\Omega \times (0,T), \tag{7. 2}$$

$$u(\cdot,0) = u_0 \quad \text{in} \quad \Omega, \tag{7. 3}$$

where $\alpha : \mathbb{R} \to \mathbb{R}$ and Δ^a_m is the so-called m-Laplace operator with a coefficient function $a : \Omega \to \mathbb{R}$ given by

$$\Delta^a_m u = \nabla \cdot \left(a(x)|\nabla u|^{m-2}\nabla u\right), \quad 1 < m < \infty.$$

Here we assume that

(a1) $u_0 \in W_0^{1,m}(\Omega)$, $a \in L^\infty(\Omega)$ and $a_1 \le a(x) \le a_2$ for a.e. $x \in \Omega$ with some $a_1, a_2 > 0$.

(a2) α is maximal monotone in $\mathbb{R}$. Moreover, there exist $p \in [2,\infty)$ and constants $C_5, C_6 > 0$ such that

$$C_5|s|^p - \frac{1}{C_5} \le A(s) \quad \text{and} \quad |\alpha(s)|^{p'} \le C_6(|s|^p + 1) \quad \text{for all } s \in \mathbb{R},$$

where $A(s) := \int_0^s \alpha(\sigma)d\sigma$ for $s \in \mathbb{R}$.

Note that α is continuous in $\mathbb{R}$ by (a2).

In order to recast (DNP) into an abstract Cauchy problem, we set

$$V = L^p(\Omega) \quad \text{and} \quad X = W_0^{1,m}(\Omega)$$

and define two functionals $\psi, \phi : V \to [0,\infty]$ by

$$\psi(u) = \int_\Omega A(u(x))\mathrm{d}x,$$

$$\phi(u) = \begin{cases} \dfrac{1}{m}\displaystyle\int_\Omega a(x)|\nabla u(x)|^m \mathrm{d}x & \text{if } u \in W_0^{1,m}(\Omega), \\ \infty & \text{otherwise.} \end{cases}$$

Assume

(a3) $p < m^* := Nm/(N-m)_+$.

Then, by the Rellich-Kondrachov compact embedding theorem, we observe that $X \hookrightarrow V$ compactly. We find that ψ is of class C^1 in V and $\mathrm{d}_V\psi(u) = \alpha(u)$. In particular, the bounds in (A1) and (A2) immediately follow from (a2). Furthermore, ϕ_X is of class C^1 in X, and $\partial_X\phi_X(u) = -\Delta_m^a u$ equipped with the boundary condition $u|_{\partial\Omega} = 0$, and conditions (A3) and (A4) hold. Thus, (DNP) is reduced into the abstract doubly nonlinear problem (1. 5)–(1. 6). The existence of strong solutions of such a problem has been already discussed in [9].

Our current interest lies in the elliptic regularizations $(\mathrm{ER})_\varepsilon$ of (DNP) of the form:

$$-\varepsilon\alpha(u_t)_t + \alpha(u_t) - \Delta_m^a u = 0 \quad \text{in} \quad \Omega \times (0,T), \tag{7. 4}$$

$$u = 0 \quad \text{on} \quad \partial\Omega \times (0,T), \tag{7. 5}$$

$$u(\cdot,0) = u_0 \quad \text{in} \quad \Omega, \tag{7. 6}$$

$$\alpha(u(\cdot,T)) = 0 \quad \text{in} \quad \Omega \tag{7. 7}$$

for $\varepsilon > 0$. By applying our abstract theory, in particular, Theorems 3.3, 4.2 and 5.1, we have

Theorem 7.1 (WED approach to (DNP))**.** *Under* (a1)–(a3)*,* $(\mathrm{ER})_\varepsilon$ *admits a strong solution* $u_\varepsilon \in L^m(0,T;W_0^{1,m}(\Omega)) \cap W^{1,p}(0,T;L^p(\Omega))$*. Moreover,* u_ε *is the unique minimizer* *of the* WED *functional* $I_\varepsilon : L^p(0,T;L^p(\Omega)) \to [0,\infty]$ *given by*

$$I_\varepsilon(u) = \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon} \left(\int_\Omega A(u_t(x,t))\mathrm{d}x + \frac{1}{\varepsilon m} \int_\Omega a(x)|\nabla u(x,t)|^m \mathrm{d}x \right) \mathrm{d}t \\ \quad \textit{if} \ \ u(\cdot,0) = u_0 \ \ \textit{in } \Omega \ \ \textit{and} \ \ u \in L^m(0,T;W_0^{1,m}(\Omega)) \cap W^{1,p}(0,T;L^p(\Omega)), \\ \infty \quad \textit{otherwise.} \end{cases}$$

Furthermore, u_{ε_n} *converges to a strong solution* u *of* (DNP) *in the following sense*:

$$u_{\varepsilon_n} \to u \qquad \begin{aligned} &\textit{strongly in } C([0,T];L^p(\Omega)), \\ &\textit{weakly in } L^m(0,T;W_0^{1,m}(\Omega)) \cap W^{1,p}(0,T;L^p(\Omega)) \end{aligned}$$

along with some sequence $\varepsilon_n \to 0$.

We next consider the following sequence of doubly nonlinear problems $(\mathrm{DNP})_h$ for $h \in \mathbb{N}$:

$$\alpha_h(u_t) - \Delta_m^{a_h} u = 0 \quad \text{in} \quad \Omega \times (0,T), \tag{7. 8}$$

$$u = 0 \quad \text{on} \quad \partial\Omega \times (0,T), \tag{7. 9}$$

$$u(\cdot,0) = u_{0,h} \quad \text{in} \quad \Omega \tag{7. 10}$$

with functions $u_{0,h} : \Omega \to \mathbb{R}$, $a_h : \Omega \to \mathbb{R}$ and $\alpha_h : \mathbb{R} \to \mathbb{R}$. Aizicovici & Yan [1] proved the convergence theorem for $(\mathrm{DNP})_h$ under appropriate conditions on the convergences of $u_{0,h}$, a_h and α_h as $h \to \infty$. As in (DNP), let us introduce elliptic regularizations $(\mathrm{ER})_{\varepsilon,h}$ of $(\mathrm{DNP})_h$ given by

$$-\varepsilon\alpha_h(u_t)_t + \alpha_h(u_t) - \Delta_m^{a_h} u = 0 \quad \text{in} \quad \Omega \times (0,T), \tag{7. 11}$$

$$u = 0 \quad \text{on} \quad \partial\Omega \times (0,T), \tag{7. 12}$$

$$u(\cdot,0) = u_{0,h} \quad \text{in} \quad \Omega, \tag{7. 13}$$

$$\alpha_h(u(\cdot,T)) = 0 \quad \text{in} \quad \Omega. \tag{7. 14}$$

Then, the same conclusions as in Theorem 7.1 hold also for $(\mathrm{DNP})_h$ and $(\mathrm{ER})_{\varepsilon,h}$ and the corresponding WED functionals given by

$$I_{\varepsilon,h}(u) = \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon} \left(\int_\Omega A_h(u_t(x,t))\mathrm{d}x + \frac{1}{\varepsilon m} \int_\Omega a_h(x)|\nabla u(x,t)|^m \mathrm{d}x \right) \mathrm{d}t \\ \quad \text{if} \ \ u(\cdot,0) = u_{0,h} \ \ \text{in } \Omega \ \ \text{and} \ \ u \in L^m(0,T;W_0^{1,m}(\Omega)) \cap W^{1,p}(0,T;L^p(\Omega)), \\ \infty \quad \text{otherwise} \end{cases}$$

with $A_h(s) = \int_0^s \alpha_h(\sigma)d\sigma$ for $s \in \mathbb{R}$.

Finally, let us discuss the Mosco-convergence of $I_{\varepsilon,h}$ under the following assumptions.

(h1) Condition (a1) holds with functions a and u_0 replaced by a_h and $u_{0,h}$, respectively, and the respective constants independent of h. Moreover, $a_h(x) \to a(x)$ for a.e. $x \in \Omega$ and $u_{0,h} \to u_0$ strongly in $W_0^{1,m}(\Omega)$ as $h \to \infty$.

(h2) Condition (a2) holds with α replaced by α_h and respective constants independent of h. Moreover, $A_h \xrightarrow{\Gamma} A$ as $h \to \infty$, i.e., the following (i) and (ii) hold:

(i) for every sequence $s_h \to s$ as $h \to \infty$, $A(s) \le \liminf_{h\to\infty} A_h(s_h)$,

(ii) for every $s \in \mathbb{R}$, there exists a sequence $s_h \to s$ such that $A_h(s_h) \to A(s)$.

More precisely, we can prove

Theorem 7.2 (Mosco-convergence of $I_{\varepsilon,h}$). *Assume* (a1)–(a2), (h1)–(h2). *Then,* $I_{\varepsilon,h}$ *Mosco-converges to* I_ε *on* $\mathcal{Z} := L^m(0,T;X) \cap W^{1,p}(0,T;V)$ *as* $h \to \infty$.

Let (u_h) *be the sequence of unique global minimizers for* $I_{\varepsilon,h}$. *Then, there exists a sequence* $k_h \to \infty$ *in* $\mathbb{N}$ *such that*

$$u_{k_h} \to u \quad \text{weakly in } \mathcal{Z} \quad \text{as} \quad h \to \infty,$$

where u *minimizes* I_ε, *i.e.,* u *solves* $(\mathrm{ER})_\varepsilon$.

Proof. Let us check (H0)–(H4) for $I_{\varepsilon,h}$ as $h \to \infty$. As in [1], by using standard facts in [3], one can check (H1), (H2) and (H4) from (h1) and (h2). So it remains to prove (H3). Let (k_h) be a sequence in $\mathbb{N}$ such that $k_h \to \infty$. Let $u \in X$ and $u_h \in X$ be such that $u_h \to u$ strongly in X and $\phi_{k_h}(u_h) \to \phi(u)$ as $h \to \infty$. Let $v \in X$ and $\tau > 0$ be fixed. Set

$$v_h := u_h + v - u \in X.$$

Here we claim that $A_{k_h}(s) \to A(s)$ as $h \to \infty$ for all $s \in \mathbb{R}$. Indeed, by (ii) of (h2), for any $s \in \mathbb{R}$ we can take a sequence $s_h \to s$ such that $A_{k_h}(s_h) \to A(s)$ as $h \to \infty$. Hence by (a2) for α_h with C_6 independent of h, we have

$$\begin{aligned} A_{k_h}(s) - A(s) &= A_{k_h}(s) - A_{k_h}(s_h) + A_{k_h}(s_h) - A(s) \\ &\le C|s - s_h| + A_{k_h}(s_h) - A(s) \to 0 \quad \text{as} \quad h \to \infty, \end{aligned}$$

and obtain a similar estimate from below as well. Thus $A_{k_h}(s) \to A(s)$ as $h \to \infty$. Furthermore, we can derive

$$\psi_{k_h}(w) \to \psi(w) \quad \text{as} \quad h \to \infty \quad \text{for any } w \in V.$$

Indeed, for any $w \in V$, it follows that $A_{k_h}(w(x)) \to A(w(x))$ for a.e. $x \in \Omega$, and moreover, by (a2) for α_h with C_6 independent of h, dominated convergence yields $\psi_{k_h}(w) \to \psi(w)$ as $h \to \infty$. Thus

$$\psi_{k_h}\left(\frac{v_h - u_h}{\tau}\right) = \psi_{k_h}\left(\frac{v-u}{\tau}\right) \to \psi\left(\frac{v-u}{\tau}\right) \quad \text{as} \quad h \to \infty.$$

Moreover, since $v_h \to v$ strongly in X, it follows from (h1) that

$$\phi_{k_h}(v_h) = \frac{1}{m}\int_\Omega a_{k_h}(x)|\nabla v_h(x)|^m \mathrm{d}x \to \frac{1}{m}\int_\Omega a(x)|\nabla v(x)|^m \mathrm{d}x = \phi(v) \quad \text{as} \quad h \to \infty.$$

Thus, (H3) holds. Consequently, by Theorem 6.2 and Corollary 6.3, we obtain the desired. Note that the precompactness of global minimizers (u_h) is immediate. □

Appendix

We report here the details of a result used in Subsection 3.2. At first, let us recall some useful properties of the Legendre-Fenchel transform φ^* of a proper, lower semicontinuous and convex functional φ from a normed space E into $\to (-\infty,\infty]$ given by $\varphi^*(f) := \sup_{v\in E}\left\{\langle f,v\rangle_E - \varphi(v)\right\}$, for $f\in E^*$ (see, e.g., [4]):

(i) φ^* is proper, lower semicontinuous and convex in E^*;

(ii) $\varphi^*(f) = \langle f,u\rangle_E - \varphi(u)$ for all $[u,f]\in\partial_E\varphi$;

(iii) $u\in\partial_{E^*}\varphi^*(f)$ if and only if $f\in\partial_E\varphi(u)$.

Moreover, we observe that, whenever $\varphi : E\to[0,\infty]$, one has $\varphi^*(0) = -\inf_{v\in E}\varphi(v)\le 0$ and $\varphi^*(f)\ge -\varphi(0)$ for all $f\in E^*$.

Now, our claim reads,

Lemma 7.3. *The inequalities* (3. 15) *and* (3. 16) *can be rigorously justified within the frame of Subsection* 3.2.

Proof. Fix an arbitrary constant $\tau>0$ and define a backward difference operator δ_τ^- by

$$\delta_\tau^-\chi(t) = \frac{\chi(t)-\chi(t-\tau)}{\tau}$$

for functions χ defined on $[0,T]$ with values in a vector space and $t\ge\tau$. Test $\xi_\lambda'(t)$ by $u_\lambda'(t)$ and integrate over (t_0,T) with an arbitrary $t_0\in\mathcal{L}_\lambda$. Since $u_\lambda\in W^{1,p}(0,T;V)$ and $\xi_\lambda\in W^{1,p'}(0,T;V^*)$, we have

$$\int_{t_0}^T \langle\xi_\lambda'(s),u_\lambda'(s)\rangle_V\,\mathrm{d}s = \lim_{\tau\to0}\int_{t_0+\tau}^T\left\langle\xi_\lambda'(s),\delta_\tau^-u_\lambda(s)\right\rangle_V\mathrm{d}s.$$

Moreover, it follows from (3. 12) that

$$\begin{aligned}
&\int_{t_0+\tau}^T\left\langle\xi_\lambda'(s),\delta_\tau^-u_\lambda(s)\right\rangle_V\mathrm{d}s\\
&= \langle\xi_\lambda(T),\delta_\tau^-u_\lambda(T)\rangle_V - \langle\xi_\lambda(t_0+\tau),\delta_\tau^-u_\lambda(t_0+\tau)\rangle_V - \int_{t_0+\tau}^T\left\langle\xi_\lambda(s),\delta_\tau^-u_\lambda'(s)\right\rangle_V\mathrm{d}s\\
&= -\langle\xi_\lambda(t_0+\tau),\delta_\tau^-u_\lambda(t_0+\tau)\rangle_V - \frac{1}{\tau}\int_{t_0+\tau}^T\langle\xi_\lambda(s),u_\lambda'(s)-u_\lambda'(s-\tau)\rangle_V\,\mathrm{d}s. \qquad (7.\ 15)
\end{aligned}$$

Next, we observe that

$$\begin{aligned}
&\frac{1}{\tau}\int_{t_0+\tau}^T\langle\xi_\lambda(s),u_\lambda'(s)-u_\lambda'(s-\tau)\rangle_V\,\mathrm{d}s\\
&= \frac{1}{\tau}\int_{t_0+\tau}^T\langle\xi_\lambda(s),u_\lambda'(s)\rangle_V\mathrm{d}s - \frac{1}{\tau}\int_{t_0}^{T-\tau}\langle\xi_\lambda(s+\tau),u_\lambda'(s)\rangle_V\mathrm{d}s\\
&= \frac{1}{\tau}\int_{t_0}^{T-\tau}\langle\xi_\lambda(s)-\xi_\lambda(s+\tau),u_\lambda'(s)\rangle_V\mathrm{d}s - \frac{1}{\tau}\int_{t_0}^{t_0+\tau}\langle\xi_\lambda(s),u_\lambda'(s)\rangle_V\mathrm{d}s\\
&\quad + \frac{1}{\tau}\int_{T-\tau}^T\langle\xi_\lambda(s),u_\lambda'(s)\rangle_V\mathrm{d}s. \qquad (7.\ 16)
\end{aligned}$$

Using the fact that $u'_\lambda(s) \in \partial_{V^*}\psi^*(\xi_\lambda(s))$, where ψ^* denotes the Legendre-Fenchel transform of ψ, and the definition of subdifferentials, we obtain

$$\frac{1}{\tau}\int_{t_0}^{T-\tau} \langle \xi_\lambda(s) - \xi_\lambda(s+\tau), u'_\lambda(s)\rangle_V \mathrm{d}s \geq \frac{1}{\tau}\int_{t_0}^{T-\tau} \Big(\psi^*(\xi_\lambda(s)) - \psi^*(\xi_\lambda(s+\tau))\Big)\mathrm{d}s,$$

and, moreover,

$$\frac{1}{\tau}\int_{T-\tau}^{T} \langle \xi_\lambda(s), u'_\lambda(s)\rangle_V \mathrm{d}s \geq \frac{1}{\tau}\int_{T-\tau}^{T} \Big(\psi^*(\xi_\lambda(s)) - \psi^*(0)\Big)\mathrm{d}s.$$

Then, going back to equation (7. 16), one has

$$\begin{aligned}
&\frac{1}{\tau}\int_{t_0+\tau}^{T} \langle \xi_\lambda(s), u'_\lambda(s) - u'_\lambda(s-\tau)\rangle_V \,\mathrm{d}s \\
&\geq \frac{1}{\tau}\int_{t_0}^{t_0+\tau} \psi^*(\xi_\lambda(s))\mathrm{d}s - \frac{1}{\tau}\int_{t_0}^{t_0+\tau} \langle \xi_\lambda(s), u'_\lambda(s)\rangle_V \mathrm{d}s - \psi^*(0).
\end{aligned}$$

Hence, from equation (7. 15) we can compute that

$$\begin{aligned}
&\int_{t_0+\tau}^{T} \left\langle \xi'_\lambda(s), \delta_\tau^- u_\lambda(s)\right\rangle_V \mathrm{d}s \\
\leq\ & -\langle \xi_\lambda(t_0+\tau), \delta_\tau^- u_\lambda(t_0+\tau)\rangle_V - \frac{1}{\tau}\int_{t_0}^{t_0+\tau} \psi^*(\xi_\lambda(s))\mathrm{d}s \\
& + \frac{1}{\tau}\int_{t_0}^{t_0+\tau} \langle \xi_\lambda(s), u'_\lambda(s)\rangle_V \mathrm{d}s + \psi^*(0) \\
\to\ & -\langle \xi_\lambda(t_0), u'_\lambda(t_0)\rangle_V - \psi^*(\xi_\lambda(t_0)) + \langle \xi_\lambda(t_0), u'_\lambda(t_0)\rangle_V + \psi^*(0) \\
=\ & -\psi^*(\xi_\lambda(t_0)) + \psi^*(0) \leq \psi(0).
\end{aligned}$$

Here we used the facts that $t_0 \in \mathcal{L}_\lambda$, $\psi^* \geq -\psi(0)$, $\psi^*(0) \leq 0$ and the function $t \mapsto \psi^*(\xi_\lambda(t))$ is (absolutely) continuous on $[0,T]$ since $u'_\lambda(t) \in \partial_{V^*}\psi^*(\xi_\lambda(t))$, $u'_\lambda \in L^p(0,T;V)$ and $\xi_\lambda \in W^{1,p'}(0,T;V^*)$. Consequently,

$$\int_{t_0}^{T} \langle \xi'_\lambda(s), u'_\lambda(s)\rangle_V \,\mathrm{d}s \ \leq\ \psi(0) \quad \text{for all } \ t_0 \in \mathcal{L}_\lambda$$

and (3. 15) follows from the density of $\mathcal{L}_\lambda$.

Let now $t \in \mathcal{L}_\lambda$ be fixed. Arguing as above starting again from equation (7. 15) with t instead of T, we can also verify that

$$\begin{aligned}
&\int_{t_0+\tau}^{t} \left\langle \xi'_\lambda(s), \delta_\tau^- u_\lambda(s)\right\rangle_V \mathrm{d}s \\
\leq\ & \langle \xi_\lambda(t), \delta_\tau^- u_\lambda(t)\rangle_V - \langle \xi_\lambda(t_0+\tau), \delta_\tau^- u_\lambda(t_0+\tau)\rangle_V - \frac{1}{\tau}\int_{t_0}^{t_0+\tau} \psi^*(\xi_\lambda(s))\mathrm{d}s \\
& + \frac{1}{\tau}\int_{t_0}^{t_0+\tau} \langle \xi_\lambda(s), u'_\lambda(s)\rangle_V \mathrm{d}s + \psi^*(0) \\
\to\ & \langle \xi_\lambda(t), u'_\lambda(t)\rangle_V - \psi^*(\xi_\lambda(t_0)) + \psi^*(0) \\
\leq\ & \langle \xi_\lambda(t), u'_\lambda(t)\rangle_V + \psi(0) \quad \text{as } \ \tau \to 0,
\end{aligned}$$

and the inequality (3. 16) follows. □

Acknowledgements

The first author gratefully acknowledges the kind hospitality and support of the IMATI-CNR in Pavia, Italy, where this research was initiated, and the partial support of the SIT grant and the Grant-in-Aid for Young Scientists (B) # 19740073, MEXT. The second author is partly supported by FP7-IDEAS-ERC-StG Grant # 200947 (BioSMA).

References

[1] Aizicovici, S. and Yan, Q., Convergence theorems for abstract doubly nonlinear differential equations, Panamer. Math. J., **7** (1997), 1–17.

[2] Akagi, G. and Stefanelli, U., A variational principle for doubly nonlinear evolution, Preprint IMATI-CNR 11PV09/9/0, 2009.

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

[4] Barbu, V., *Nonlinear Semigroups and Differential Equations in Banach spaces*, Noordhoff, 1976.

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

[6] Brézis, H., Crandall, M.G. and Pazy, A., Perturbations of nonlinear maximal monotone sets in Banach space, *Comm. Pure. Appl. Math.* **23** (1970), 123–144.

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

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

[9] Colli, P., On some doubly nonlinear evolution equations in Banach spaces, Japan J. Indust. Appl. Math. **9** (1992), 181–203.

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

[11] Ghoussoub, N., *Self-dual partial differential systems and their variational principles*, Springer Monographs in Mathematics, Springer, New York, 2009.

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

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

[14] Gurtin, M.E., Variational principle for the time-dependent Schrödinger equation, J. Mathematical Phys., **6** (1965), 1506–1507.

[15] Hlaváček, I., Variational principles for parabolic equations, Apl. Mat., **14** (1969), 278–297.

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

[17] Kenmochi, N., Some nonlinear parabolic variational inequalities, Israel J. Math. **22** (1975), 304–331.

[18] Lions, J.-L. and Magenes, E., *Problèmes aux limites non homogènes et applications, Vol. 1*, Dunod, Paris, 1968.

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

[20] Mielke, A., Roubíček, T. and Stefanelli, U., Gamma-limits and relaxations for rate-independent evolutionary problems, Calc. Var. Partial Differential Equations, **31** (2008), 3:387–416.

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

[22] Mielke, A. and Stefanelli, U., Weighted energy-dissipation functionals for gradient flows, ESAIM Control Optim. Calc. Var., (2009), to appear.

[23] Mosco, U., Convergence of convex sets and of solutions of variational inequalities, Advances in Math., **3** (1969), 510–585.

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

[25] Nayroles, B., Un théorème de minimum pour certains systèmes dissipatifs, Variante hilbertienne, Travaux Sém. Anal. Convexe, **6** (1976), Exp. No. 2.

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

[27] Spadaro, E. and Stefanelli, U., A variational view at mean curvature evolution for cartesian surfaces, In preparation, 2009

[28] Stefanelli, U., The Brezis-Ekeland principle for doubly nonlinear equations, SIAM J. Control Optim., **47** (2008), 1615–1642.

[29] Stefanelli, U., A variational principle for hardening elastoplasticity, SIAM J. Math. Anal., **40** (2008), 623–652.

[30] Stefanelli, U., The discrete Brezis-Ekeland principle, J. Convex Anal., **16** (2009), 71–87.

[31] Prüß, J., A characterization of uniform convexity and applications to accretive operators, Hiroshima Math. J., **11** (1981), 229–234.

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

[33] Visintin, A., Extension of the Brezis-Ekeland-Nayroles principle to monotone operators, Adv. Math. Sci. Appl., **18** (2008), 633–650.