Doubly nonlinear evolution inclusion of second order with non-variational perturbation

Aras Bacho*

October 23, 2019

Abstract

We investigate the existence of weak solutions to the abstract Cauchy problem for the following doubly nonlinear evolution inclusion of second order

\[
\begin{aligned}
&u''(t) + \partial \Psi(u'(t)) + \partial E_t(u(t)) + B(t, u(t), u'(t)) \ni f(t), \\
&u(0) = u_0, \quad u'(0) = v_0,
\end{aligned}
\]

on a separable and reflexive Banach space \( V \). Under suitable assumptions on the functionals \( E_t \) and \( \Psi \) as well as the perturbation \( B \) and the external force \( f \), the existence of solutions is established by showing the convergence of a variational approximation scheme using tools from the theory of convex analysis and the theory of gradient systems.

1 Introduction

In this article, results on the doubly non-linear evolution equation

\[
\begin{aligned}
&u''(t) + A(t)u(t) + B(t, u(t), u'(t)) = f(t), \\
&u(0) = u_0, \quad u'(0) = v_0
\end{aligned}
\]

are provided. General inclusions of type (1.1) can be seen as an abstract formulation of second order in time equations coming from physics as the Kevin–Voigt model or the peridynamic model in elasticity theory, the non-linear Klein–Gordon equation from quantum mechanics and other nonlinear wave equations arising in mechanics and quantum mechanics, a equation for describing a vibrating membrane, see, e.g., [EmT10, EmŠ11, EmŠ13, RoT17, BMR12, Rou05].

Abstract evolution equations of second order of the form

\[ u''(t) + A(t)u(t) + B(t)u(t) = f(t), \quad t \in (0, T), \]

have been studied by several authors under various conditions and assumptions on the acting operators. Emmrich and Šiška have shown in [EmŠ13] the existence of weak solutions considering the case when \( A : V_A \to V_A^* \) is a linear, bounded, strongly positive and symmetric, and \( B : V_B \to V_B^* \) is a demicontinuous and bounded potential operator.

*Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany.
operator such that a Andrews–Ball type condition holds, i.e., \((B + \lambda A) : V \to V^*\) is monotone with \(V := V_A \cap V_B\), \(V_A\) and \(V_B\) being suitable separable and reflexive Banach spaces, whereas in [EST15] the authors consider the reverse case, i.e., the operator \(A\) is supposed to be hemicontinuous, monotone, coercive and satisfy a suitable growth condition and the operator \(B\) is supposed to be linear, bounded, symmetric, and strongly positive. A doubly nonlinear case has been investigated in [EmT10, EmT11, EmŠ11] with similar assumptions on the operators with the crucial assumption on \(B(t) = B_0 + B_1(t)\) being the sum of a linear, bounded, symmetric and strongly positive operator \(B_0\) and a compact perturbation \(B_1(t)\). The case where at least one operator is assumed to be multivalued has also been studied by several authors. Just recently Rossi and Thomas studied in [RoT17] a coupled system of a rate-dependent and a rate-independent process. The rate-dependent process being of second order in time modelling visco-elastic solids with inertia while the rate-independent process is a generalized gradient flow modeling visco-elastic solids.

A special class of evolution equations of second order are the so-called Hamiltonian or energy preserving systems, i.e., when \(A \equiv 0\) and \(B\) is a potential operator. In this case no dissipative effects occur, e.g., in the Klein–Gordon equation without viscous regularization. This situation is in general more delicate and there are very few abstract results known. By means of the semigroup theory, Barbu and Brézis [Bar76, Bre72] obtained the existence of strong solutions in a Hilbert space setting. We also refer the reader to the monographs [Bar76, GGZ74, Rou09] for a detailed study of evolution equations.

## 2 Analytical framework

In what follows, let \((U, \| \cdot \|_U), (V, \| \cdot \|)\) and \((W, \| \cdot \|_W)\) be real, reflexive and separable Banach spaces and let \((H, \cdot , \cdot)\) be a Hilbert space with the norm \(| \cdot |\) induced by the inner product \((\cdot , \cdot)\), such that the dense and continuous embeddings

\[
U \hookrightarrow V \hookrightarrow W \hookrightarrow H \cong H^* \hookrightarrow W^* \hookrightarrow V^* \hookrightarrow U^* 
\]

hold. Furthermore, we assume the embeddings \(U \hookrightarrow W\) as well as \(V \hookrightarrow H\) to be compact.

First, in order to give the inclusion (1.1) a meaning, we define for a proper functional \(F : V \to (-\infty, +\infty]\) the (Fréchet) subdifferential of \(F\), which is given by the multivalued map \(\partial F : V \to 2^{V^*}\) with

\[
\partial F(u) := \left\{ \xi \in V^* : \liminf_{v \to u} \frac{F(v) - F(u) - \langle \xi, v - u \rangle}{\|v - u\|} \geq 0 \right\}.
\]

Its elements are called subgradients, where \((\cdot , \cdot)\) denotes the duality pairing between the Banach space \(V\) and its topological dual space \(V^*\). The effective domain of \(F\) and the domain of its subdifferential \(\partial F\), we denote by \(D(F) := \{ v \in V \mid F(v) < +\infty\}\) and \(D(\partial F) := \{ v \in V : \partial F(v) \neq \emptyset \}\), respectively. If the set of subgradients of \(F\) at a given point \(u\) is nonempty, we say that \(F\) is subdifferentiable at \(u\). The following lemma states that the subdifferential of the sum of a subdifferential function and a Fréchet differentiable function equals the sum of the subdifferentials of both functions.
Lemma 2.1. Let \( F : V \to (-\infty, +\infty] \) and \( G : V \to (-\infty, +\infty] \) be subdifferentiable and Fréchet differentiable at \( u \in D(\partial F) \cap D(\partial G) \neq \emptyset \), respectively. Then, there holds

\[
\partial(F + G)(u) = \partial F(u) + DFG(u).
\]

Proof. This follows immediately from the definition of a subdifferential. \( \square \)

Since we are dealing with convex and \( \lambda \)-convex functions, we would like to express the subdifferential in an equivalent way which is easier to handle. The function \( F \) is called \( \lambda \)-convex with parameter \( \lambda \in \mathbb{R} \) if

\[
F(tu + (1-t)v) \leq tF(u) + (1-t)F(v) + \frac{\lambda}{2}t(1-t)\|u - v\|^2
\]

for all \( u, v \in D(F) \) and \( t \in (0, 1) \). If \( \lambda < 0 \) and \( \lambda = 0 \), then the function is called strongly convex and convex, respectively. It is not difficult to show that for a \( \lambda \)-convex and proper function \( F \), the subdifferential of \( F \) is equivalently given by

\[
\partial F(u) = \left\{ \xi \in V^* : F(u) \leq F(v) + \langle \xi, u - v \rangle + \frac{\lambda}{2}\|u - v\|^2 \text{ for all } v \in V \right\}.
\]

It follows immediately that for \( \lambda < 0 \), the subdifferential of a \( \lambda \)-convex function \( F \) is strongly positive, i.e.,

\[
\langle \xi - \zeta, u - v \rangle \geq -\lambda\|u - v\|^2
\]

for all \( \xi \in \partial F(u) \) and \( \zeta \in \partial F(v) \).

Let us now introduce an important tool from the theory of convex analysis. For a proper, lower semicontinuous and convex function \( F : V \to (-\infty, +\infty] \), we define the so-called convex conjugate (or Legendre–Fenchel transform) \( F^* : V^* \to (-\infty, +\infty] \) by

\[
F^*(\xi) := \sup_{u \in V} \{ \langle \xi, u \rangle - F(u) \}, \quad \xi \in V^*.
\]

By definition, we directly obtain the Fenchel–Young inequality

\[
\langle \xi, u \rangle \leq F(u) + F^*(\xi), \quad v \in V, \xi \in V^*.
\]

It can be checked that the convex conjugate itself is proper, lower semicontinuous and convex, see, e.g., Ekeland and Témam [EkT74]. If, in addition, we assume \( F(0) = 0 \), then \( F^*(0) = 0 \) holds as well. The following lemma gives a characterisation of a subgradient of a function \( F \) in terms of its convex conjugate.

Lemma 2.2. Let \( F : V \to (-\infty, +\infty] \) be a proper, lower semicontinuous and convex functional and let \( F^* : V^* \to (-\infty, +\infty] \) be the convex conjugate of \( F \). Then for all \( (u, \xi) \in V \times V^* \), the following assertions are equivalent:

i) \( \xi \in \partial F(u) \) in \( V^* \);

ii) \( u \in \partial F^*(\xi) \) in \( V \);

iii) \( \langle \xi, u \rangle = F(u) + F^*(\xi) \) in \( \mathbb{R} \).

Proof. Ekeland and Témam [EkT74, Prop. 5.1 and Cor. 5.2 on pp. 21]. \( \square \)
2.1 Assumptions

In this section, we collect all assumptions concerning the potential \( \Psi \), the energy functional \( \mathcal{E} \), the perturbation \( B \) as well as the external force \( f \). Henceforth, we refer to the inclusion (1.1) in the given framework as the damped inertial system \((U, V, W, H, \mathcal{E}, \Psi, B, f)\). The assumptions we consider are essentially the same as given in [BEM19], where the same evolution inclusion has been investigated after neglecting the acceleration term \( u''(t) \). Involving the acceleration term makes the situation more delicate. The assumptions for the dissipation potential \( \Psi \) are the following:

\[ (2.\Psi) \text{ Quadratic dissipation potential.} \] There exist a strongly positive, symmetric and continuous bilinear form \( a : V \times V \to \mathbb{R} \) such that \( \Psi(v) = a(v, v) \).

Remark 2.3.

\( i ) \) Assumption \((2.\Psi)\) yields the convexity and continuity of the map \( v \mapsto \Psi(v) \). Furthermore, \( \Psi \) is Fréchet differentiable with the Fréchet derivative given by a strongly positive, linear bounded and symmetric operator \( A : V \to V^* \) such that \( \partial \Psi(v) = \{ Av \} \) and the potential is expressed by \( \Psi(v) = \frac{1}{2} \langle Av, v \rangle \). Assumption \((2.\Psi)\) implies that the Legendre–Fenchel transform \( \Psi^* \) is convex, continuous, finite everywhere, i.e., \( D(\Psi^*) = V^* \) and can explicitly expressed by \( \Psi^*(\xi) = \frac{1}{2} \langle \xi, A^{-1} \xi \rangle \), where \( A^{-1} : V^* \to V \) is also continuous, symmetric and strongly positive which follows from the Lax–Milgram theorem.

\( ii ) \) A state dependence of the dissipation potential \( \Psi_u \) is admissible under the assumptions that the state dependence is continuous in the sense of Mosco-convergence, the potentials \( \Psi_u \) and \( \Psi^*_u \) are coercive uniformly in \( u \in V \) in sublevels of \( \mathcal{E} \) and the Nemitskij operator associated to the Fréchet derivative \( A(t) : V \to V^* \) of \( \Psi_u(v) \) with respect to \( v \in V \) is weak to weak continuous in suitable Bochner spaces. See, e.g., [Att84, BEM19, MRS13] for more details.

\( iii ) \) We remark, that we could also allow a time dependent dissipation potential when we assume a differentiable time dependence and a uniform strongly positivity and boundedness of \( A(t) : V \to V^* \) and a slight modification of Assumption \((2.Ec)\) and \((2.Bb)\).

Let us also collect the assumptions for the energy functional \( \mathcal{E} \).

\[ (2.Ea) \text{ Constant domain.} \] For all \( t \in [0, T] \), the functional \( \mathcal{E}_t : V \to (-\infty, +\infty] \) is proper and weakly lower semicontinuous with the time-independent effective domain \( \text{dom}(\mathcal{E}_t) \equiv D \subset U \) for all \( t \in [0, T] \).

\[ (2.Eb) \text{ Bounded from below.} \] \( \mathcal{E}_t \) is bounded from below uniformly in time, i.e., there exists a constant \( C_0 \in \mathbb{R} \) such that
\[
\mathcal{E}_t(u) \geq C_0 \quad \text{for all } u \in V \text{ and } t \in [0, T].
\]
Since a potential is uniquely determined up to a constant, we assume without loss of generality \( C_0 = 0 \).

\[ (2.Ec) \text{ Coercivity.} \] There exists \( \tau_0 > 0 \) such that for every \( t \in [0, T] \) and \( u_0 \in V \) the map
\[
u \mapsto \mathcal{E}_t(u) + \tau_0 \Psi((u - u_0)/\tau_0)
\]
has bounded sublevels in \( U \).
(2.Ed) **Control of the time derivative.** For all \( u \in D \), the map \( t \mapsto \mathcal{E}_t(u) \) is continuous in \([0, T]\) and differentiable in \((0, T)\) and its derivative \( \partial_t \mathcal{E}_t \) is controlled by the function \( \mathcal{E}_t \), i.e., there exists \( C_1 > 0 \) such that
\[
|\partial_t \mathcal{E}_t(u)| \leq C_1 \mathcal{E}_t(u) \quad \text{for all } t \in (0, T) \text{ and } u \in D.
\]

(2.Ee) **Closedness of \( \partial \mathcal{E} \).** For all sequences \( t_n : [0, T] \to [0, T], \; n \in \mathbb{N}, \)
\( (u_n)_{n \in \mathbb{N}} \subset L^\infty(0, T; U) \) and \( (\xi_n)_{n \in \mathbb{N}} \subset L^2(0, T; U^*) \) such that
a) \( t_n(t) \to t \) for a.a. \( t \in [0, T], \)

b) \( \exists C_2 > 0 : \sup_{\{n \in \mathbb{N} \mid t \in [0, T]\}} \mathcal{G}(u_n(t)) \leq C_2, \)

c) \( \xi_n(t) \in \partial \mathcal{E}_{t_n(t)}(u_n(t)) \text{ a.e. in } (0, T), \)

d) \( u_n \rightharpoonup^* u \text{ in } L^\infty(0, T; U), \; \|\sigma_h u_n - u_n\|_{L^2(0, T; V)} \to 0 \) as \( h \to 0 \) uniformly in \( n \in \mathbb{N} \) and \( \xi_n \rightharpoonup \xi \text{ in } L^2(0, T; U^*), \)

e) \( \limsup_{n \to \infty} \int_0^T (\xi_n(t), u_n(t))_{U^* \times U} dt \leq \int_0^T (\xi(t), u(t))_{U^* \times U} dt, \)

we have the relations
\[
\xi(t) \in \partial \mathcal{E}_t(u(t)) \text{ in } U^*, \quad \mathcal{E}_{t_n(t)}(u_n(t)) \to \mathcal{E}_t(u(t)) \quad \text{as } n \to \infty
\]
and
\[
\limsup_{n \to \infty} \partial_t \mathcal{E}_{t_n(t)}(u_n(t)) \leq \partial_t \mathcal{E}_t(u(t)) \quad \text{for a.a. } t \in (0, T).
\]

(2.Ef) **\( \lambda \)-convexity.** There exists \( \lambda \leq 0 \) such that for every \( t \in [0, T] \), the energy functional \( \mathcal{E}_t \) is \( \lambda \)-convex.

(2.Eg) **Control of the subgradient.** Let \( u \in D(\partial_t \mathcal{E}_t) \) and \( \xi \in \partial_t \mathcal{E}_t(u) \). Then there exist constants \( \hat{C} > 0 \) and \( \sigma > 0 \) such that
\[
\|\xi\|_{U^*} \leq \hat{C}(1 + \mathcal{E}(u) + \|u\|_V),
\]
where \( \partial_t \mathcal{E} \) denotes the subdifferential of \( \mathcal{E} \) on the space \( U \).

We first give a few relevant comments on these assumptions that will be important later on.

**Remark 2.4.**

i) From Assumption (2.Ed), we deduce with **GRONWALL’s lemma** the chain of inequalities
\[
e^{-\hat{C}|t-s|} \mathcal{E}_s(u) \leq \mathcal{E}_t(u) \leq e^{\hat{C}|t-s|} \mathcal{E}_s(u) \quad \text{for all } s, t \in [0, T], \; u \in D. \tag{2.1}
\]

In particular, there exists a constant \( C_1 > 0 \) such that
\[
\mathcal{G}(u) = \sup_{t \in [0, T]} \mathcal{E}_t(u) \leq C_1 \inf_{t \in [0, T]} \mathcal{E}_t(u) \quad \text{for all } u \in D. \tag{2.2}
\]

ii) If the sequence \( (u_n)_{n \in \mathbb{N}} \) in (2.Ec) is bounded in \( H^1(0, T; V) \), the condition \( \|\sigma_h u_n - u_n\|_{L^2(0, T; h, V)} \to 0 \) as \( h \to 0 \) uniformly in \( n \in \mathbb{N} \) holds obviously. Though, this can not stated in this way since we want to verify the conditions for a sequence of piecewise constant functions.
Finally, we make the following assumptions on the non-variational non-monotone perturbation $B$ and the force $f$.

(2.Ba) **Continuity.** The map $(t, u, v) \mapsto B(t, u, v) : [0, T] \times W \times H \to V^*$ is continuous on sublevels of $\mathcal{G}$, i.e., for every sequence $(t_n, u_n, v_n) \to (t, u, v)$ in $[0, T] \times W \times H$ with $\sup_{n \in \mathbb{N}} \mathcal{G}(u_n) \leq R$, there holds $B(t_n, u_n, v_n) \to B(t, u, v)$ in $V^*$.

(2.Bb) **Control of $B$.** There exist constants $\beta > 0$ as well as $c \in (0, 1)$ and $\tilde{c} \in [0, 1)$ with $c + \tilde{c} < 1$ such that
\[
c \Psi^* \left( \frac{-B(t, u, v)}{c} \right) \leq \beta (1 + \mathcal{E}_t(u) + |v|^2) + \tilde{c} \Psi(v)
\]
for all $u \in D$, $v \in V$, $t \in [0, T]$.

(2.f) **External force.** The external force is square integrable as function with values in $H$, i.e., $f \in L^2(0; T; H)$.

**Remark 2.5.** We note that Assumption (2.Ba) ensures that the Nemytskij operator associated to $B$ maps strongly measurable functions contained in sublevels of $\mathcal{G}$ into strongly measurable functions, i.e., for all strongly measurable functions $u$ and $v$ with $\sup_{t \in [0, T]} \mathcal{G}(u(t)) \leq R$, the map $t \mapsto B(t, u(t), v(t))$ is strongly measurable.

Having all assumptions given, we are in the position to state the main result which includes the definition of a notion of a solution to (1.1).

**Theorem 2.6** (Existence result). Let the damped inertial system $(U, V, W, H, \mathcal{E}, \Psi, B, f)$ be given and fulfil Assumptions (2.E), (2.Ψ) and (2.B) as well as Assumption (2.f). Then for all initial conditions $u_0 \in D$, $v_0 \in H$, there exists a solution to (1.1), i.e., there exist functions
\[
u \in L^\infty(0; T; U) \cap AC([0, T]; V) \cap W^{1, \infty}(0, T; H) \cap H^2(0, T; U^*) \text{ and } \xi \in L^\infty(0, T; U^*)
\]
with $u(0) = u_0$ and $u'(0) = v_0$ such that
\[
\xi(t) \in \partial \mathcal{E}_t(u(t)) \text{ and } f(t) \in u''(t) + \partial \Psi(u'(t)) + \xi(t) + B(t, u(t), u'(t)) \text{ in } U^*
\]
for almost all $t \in (0, T)$ and such that the energy dissipation inequality
\[
\frac{1}{2} |u'(t)|^2 + \mathcal{E}_t(u(t)) + \int_s^t (\Psi(u'(r)) + \Psi^*(S(r) - \xi(r) - u''(r)) \ dr
\]
\[
\leq \frac{1}{2} |u'(s)|^2 + \mathcal{E}_s(u(s)) + \int_s^t \partial \mathcal{E}_t(u(r)) \ dr + \int_s^t (S(r), u'(r)) \ dr,
\]
holds for all $0 < t \leq T$ for $s = 0$, and a.a. $s \in (0, t)$, where $S(r) := f(r) - B(r, u(r), u'(r))$, $r \in [0, T]$.

### 2.2 Semi-implicit variational approximation scheme

The proof of theorem (2.6) is based on the construction of strong solutions to (1.1) via a semi-implicit time discretization scheme. More specifically, we will employ a semi-implicit Euler method where all terms will be discretized implicitly except for the non-variational...
perturbation term \( B \) in order to obtain a variational approximation scheme to inclusion (1.1). To illustrate the idea, let for \( N \in \mathbb{N} \setminus \{0\} \)
\[
I_\tau = \{ 0 = t_0 < t_1 < \cdots < t_n = n\tau < \cdots < t_N = T \}
\]
be an equidistant partition of the time interval \([0, T]\) with step size \( \tau := T/N \), where we omit the dependence of supporting points on the step size for simplicity. The discretisation of (1.1) is then given by
\[
\frac{V^n_r - V^{n-1}_r}{\tau} + \partial \Psi(V^n_r) + \partial \mathcal{E}_{t_n}(U^n_r) + B(t_{n-1}, U^{n-1}_r, V^{n-1}_r) \ni f^n_r \quad \text{in} \ V^* \tag{2.4}
\]
for \( n = 1, \ldots, N \), where \( V^n_r := \frac{u^n_r - u^{n-1}_r}{\tau} \) and \( f^n_r := \int_{t_{n-1}}^{t_n} f(\sigma) \, d\sigma \). The values \( U^n_r \approx u(t_n) \) and \( V^n_r \approx u(t_n) \) shall approximate the exact solution and its time derivative, and are to determine recursively from (2.4). By Lemma 2.1, it can be seen that the approximate value \( U^n_r \) is characterized as the solution to the Euler–Lagrange equation associated to the map
\[
u \mapsto \Phi(\tau, t_n-1, U^{n-1}_r, U^{n-2}_r, f^n_r - B(t_{n-1}, U^{n-1}_r, V^{n-1}_r); u),
\]
where
\[
\Phi(r, t, v, w, \eta; u) = \frac{1}{2r^2}|u - 2v - w|^2 + r\Psi\left(\frac{u-v}{r}\right) + \mathcal{E}_{t+r}(u) - \langle \eta, u \rangle
\]
for \( r \in \mathbb{R}^+, t \in [0, T] \) with \( r + t \in [0, T] \), \( u \in D, v, w \in V \), and \( \eta \in V^* \). We end up with the recursive scheme
\[
\begin{aligned}
U^0_r \text{ and } U^{-1}_r \text{ are given; whenever } U^{j-1}_r, \ldots, U^{n-1}_r \in V \text{ are known,}
\end{aligned}
\]
\[
\begin{aligned}
&\text{find } U^n_r \in J_{r,t_n-1}(U^{n-1}_r, U^{n-2}_r, f^n_r - B(t_{n-1}, U^{n-1}_r, V^{n-1}_r))
\end{aligned}
\]
for \( n = 1, \ldots, N \), where \( J_{r,t}(v, w; \eta) := \arg \min_{u \in U} \Phi(r, t, v, w, \eta; u) \). For notational convenience, we define the so-called Moreau–Yosida regularisation
\[
\Phi_{r,t}(v, w; \eta) = \inf_{u \in U} \Phi(r, t, v, w, \eta; u).
\]
(2.7)

The following lemma assures the solvability of the variational scheme (2.6).

**Lemma 2.7.** Let the system \((U, V, W, H, \mathcal{E}, \Psi)\) be given and let the Assumptions (2.Ea)-(2.Ec) and (2.Ψ) be fulfilled. Then for all \( r \in (0, r_0), t \in [0, T] \) with \( r + t \leq T, v, w \in V \) and \( \eta \in V^* \), the set \( J_{r,t}(v, w; \eta) \) is nonempty, where \( r_0 \) is from (2.Ec).

**Proof.** Let \( u \in D, v, w \in V, \eta \in V^* \) and \( r \in (0, r_0), t \in [0, T] \) with \( r + t \leq T \) be given. First of all, the Fenchel–Young inequality and the boundedness of the energy from below yield
\[
\Phi(r, t, v, w, \eta; u) = \frac{1}{2r^2}|u - 2v + w|^2 + r\Psi\left(\frac{u-v}{r}\right) + \mathcal{E}_{t+r}(u) - \langle \eta, u \rangle
\]
\[
= \frac{1}{2r^2}|u - 2v + w|^2 + \frac{1}{r}\Psi\left(\frac{u-v}{r}\right) + \mathcal{E}_{t+r}(u) - \langle \eta, u - v \rangle - \langle \eta, v \rangle
\]
\[
\geq \frac{1}{2r^2}|u - 2v + w|^2 + \left(\frac{1}{r} - \varepsilon\right)\Psi\left(\frac{u-v}{r}\right) + \mathcal{E}_{t+r}(u)
\]
\[
- \varepsilon\Psi^*\left(\frac{\eta}{\varepsilon}\right) - \langle \eta, v \rangle
\]
(2.8)
where $\epsilon = 1/r + 1/\tau_0$. This implies $\Phi_{r,t}(v, w; \eta) > -\infty$, i.e., there exists a global minimum of $\Phi(r, t, v, w; \eta; \cdot)$. On the other hand, we observe that

$$
\Phi_{r,t}(v, w; \eta) \leq \frac{1}{2r^2} \left| u_0 - 2v + w \right|^2 + r\Phi \left( \frac{u_0 - v}{r} \right) + \varepsilon_{t+r}(u_0) - \langle \eta, u_0 \rangle
$$

for any $u_0 \in D$, so that $\Phi_{r,t}(v, w; \eta) < +\infty$ holds as well. It remains to show that the global minimum is achieved by an element of $D$. In order to show that, let $(u_n)_{n \in \mathbb{N}} \subset V$ be a minimizing sequence for $\Phi(r, t, v, w; \eta; \cdot)$. From (2.8), we deduce that $(u_n)_{n \in \mathbb{N}} \subset V$ is contained in a sublevel set of $\tau_0 \Phi ((\cdot - v)/\tau_0) + \varepsilon_{t+r}((\cdot)$ and thus by Assumption (2.Ec) bounded in $U$. Hence, there exists a subsequence (not relabeled) which converges weakly in $V$ towards a limit $\tilde{u} \in U$. By the weak lower semicontinuity of the map $u \mapsto \Phi(r, t, v, w; \eta; u)$, we have

$$
\Phi(r, t, v, w; \eta; \tilde{u}) \leq \liminf_{n \to \infty} \Phi(r, t, v, w; \eta; u_n) = \inf_{\eta \in \Phi} \Phi(r, t, v, w; \eta; \tilde{v}),
$$

and therefore $u \in J_{r,t}(v, w; \eta) \neq \emptyset$ and $u \in D$. \hfill \Box

### 2.3 A priori estimates

Since the preceding lemma assures the solvability of the approximation scheme (2.6), we are now able to define linear and constant interpolation functions which will interpolate the values $(U^n_\tau)_{n=0}^N$ and $(V^n_\tau)_{n=0}^N$ for every $\tau > 0$, respectively, and we will derive a priori estimates for them. The interpolation functions shall approximate the wanted solution to (1.1) and its derivative, and are therefore also referred to as approximate solution to (1.1). In order to define the approximate solutions, let the initial values $u_0 \in D$ and $v_0 \in V$ and the time step $\tau > 0$ be fixed. Further, let $(U^n_\tau)_{n=1}^N \subset D$ be the sequence of approximate values obtained from the variational approximation scheme (2.6) for $U^0_\tau := U_0$ and $U^{-1}_\tau := u_0 - \tau v_0$. Moreover, let $(\xi^n_\tau)_{n=1}^N \subset V^*$ be a sequence of subgradients of the energy determined by (2.4), i.e. $\xi^n_\tau \in \partial E_{t_\tau}(U^n_\tau)$, $i = 1, \ldots, N$. The piecewise constant and linear interpolation functions are defined by

$$
\overline{U}_\tau(0) = \underline{U}_\tau(0) = \tilde{U}_\tau(0) := U^0_\tau = u_0 \quad \text{and} \quad \overline{U}_\tau(t) := U^{n-1}_\tau, \quad \tilde{U}_\tau(t) := \frac{t}{\tau}U^{n-1}_\tau + \frac{t-n-1}{\tau}U^n_\tau \quad \text{for } t \in [t_{n-1}, t_n),
$$

and

$$
\underline{U}_\tau(t) := U^n_\tau \quad \text{for } t \in (t_{n-1}, t_n] \quad \text{and} \quad \underline{U}_\tau(T) = U^N_\tau, \quad n = 1, \ldots, N.
$$

as well as

$$
\overline{V}_\tau(0) = \underline{V}_\tau(0) = \tilde{V}_\tau(0) := V^0_\tau = v_0 \quad \text{and} \quad \overline{V}_\tau(t) := V^{n-1}_\tau, \quad \tilde{V}_\tau(t) := \frac{t}{\tau}V^{n-1}_\tau + \frac{t-n-1}{\tau}V^n_\tau \quad \text{for } t \in [t_{n-1}, t_n),
$$

and

$$
\underline{V}_\tau(t) := V^n_\tau \quad \text{for } t \in (t_{n-1}, t_n] \quad \text{and} \quad \underline{V}_\tau(T) = V^N_\tau, \quad n = 1, \ldots, N,
$$

where we recall $V^n_\tau = \frac{U^n_\tau - U^{n-1}_\tau}{\tau}$ for $n = 1, \ldots, N$. We notice that $\tilde{U}_\tau = \nabla \tilde{v}_\tau$ in the weak sense. Furthermore, we define the functions $\xi_\tau : [0, T] \to V^*$ and $f_\tau : [0, T] \to H$ by

$$
\xi_\tau(t) = \xi^n_\tau, \quad f_\tau(t) = f^n_\tau \quad \text{for } t \in [t_{n-1}, t_n), \quad n = 1, \ldots, N,
$$

and

$$
\xi_\tau(T) = \xi^N_\tau \quad \text{and} \quad f_\tau(T) = f^N_\tau.
$$
2.3 A priori estimates

For notational convenience, we also introduce the piecewise constant functions \( \bar{t}_\tau \) : \([0, T] \to [0, T] \) and \( \underline{t}_\tau \) : \([0, T] \to [0, T] \) given by

\[
\begin{align*}
\bar{t}_\tau(0) &:= 0 \quad \text{and} \quad \bar{t}_\tau(t) := t_n \quad \text{for} \ t \in (t_{n-1}, t_n], \\
\underline{t}_\tau(T) &:= T \quad \text{and} \quad \underline{t}_\tau(t) := t_{n-1} \quad \text{for} \ t \in [t_{n-1}, t_n), \ n = 1, \ldots, N.
\end{align*}
\]

Obviously, there holds \( \bar{t}_\tau(t) \to t \) and \( \underline{t}_\tau(t) \to t \) as \( \tau \to 0 \). Henceforth, we use the abbreviated notation \( B_\tau(t) := B(\underline{t}_\tau(t), \bar{t}_\tau(t), \underline{v}_\tau(t)) \), \( t \in [0, T] \).

With the above defined functions, we are in the position to show useful a priori estimates.

**Lemma 2.8** (A priori estimates). Let the system \((U, V, W, H, E, \Psi, B, f)\) be given and satisfy the Assumptions (2.E), (2.Ψ), (2.B) as well as Assumption (2.f). Furthermore, let \( \hat{U}_\tau, \hat{V}_\tau, \hat{W}_\tau, \hat{H}_\tau, \hat{E}_\tau, \hat{\Psi}_\tau \) and \( f_\tau \) be the interpolation functions defined in (2.10)-(2.12) associated to the given values \( u_0 \in D, v_0 \in V \) and the step size \( \tau > 0 \). Then, the discrete energy inequality

\[
\begin{align*}
\frac{1}{2} \left| \nabla \tau(t) \right|^2 + \mathcal{E}_\tau(t) + \int_{t(t)}^{t(T)} \left( \Psi(\nabla \tau(r)) + \Psi^*(S_\tau(r) - \hat{V}_\tau'(r) - \xi_\tau(r)) \right) dr \\
\leq \frac{1}{2} \left| \nabla \tau(s) \right|^2 + \mathcal{E}_\tau(s) + \int_{t(t)}^{t(T)} \partial_t \mathcal{E}_\tau(\hat{U}_\tau(r)) dr + \int_{t(t)}^{t(T)} \langle S_\tau(r), \nabla \tau(r) \rangle dr \\
+ \tau \frac{\lambda}{2} \int_{t(t)}^{t(T)} \left| \nabla \tau(r) \right|^2 dr,
\end{align*}
\]

holds for all \( 0 \leq s < t \leq T \), where \( S_\tau(r) := f_\tau(r) - B_\tau(r), r \in [0, T] \). Moreover, there exist positive constants \( M, \tau^* > 0 \) such that the estimates

\[
\begin{align*}
\sup_{t \in [0, T]} \left| \nabla \tau(t) \right| &\leq M, \\
\sup_{t \in [0, T]} \mathcal{E}_\tau(\hat{U}_\tau(t)) &\leq M, \\
\sup_{t \in [0, T]} |\partial_t \mathcal{E}_\tau(\hat{U}_\tau(t))| &\leq M, \\
\int_0^T \left( \Psi \left( \nabla \tau(r) \right) + \Psi^* \left( S_\tau(r) - \hat{V}_\tau'(r) - \xi_\tau(r) \right) \right) dr &\leq M
\end{align*}
\]

hold for all \( 0 < \tau \leq \tau^* \). Besides, the families

\[
\begin{align*}
(\nabla \tau)_{0 < \tau \leq \tau^*} &\subset L^2(0, T; V), \\
(\hat{V}_\tau')_{0 < \tau \leq \tau^*} &\subset L^2(0, T; U^*), \\
(B_\tau)_{0 < \tau \leq \tau^*} &\subset L^2(0, T; V^*), \\
(\xi_\tau)_{0 < \tau \leq \tau^*} &\subset L^\infty(0, T; U^*)
\end{align*}
\]

are uniformly bounded with respect to \( \tau \) in the respective spaces. Finally, there holds

\[
\sup_{t \in [0, T]} \left( \left\| \bar{U}_\tau(t) - \hat{U}_\tau(t) \right\| + \left\| \bar{V}_\tau(t) - \hat{V}_\tau(t) \right\| + \left\| \nabla \tau(t) - \hat{V}_\tau'(t) \right\|_{U^*} + \left\| \nabla \tau(t) - \nabla \tau(t) \right\|_{U^*} \right) \to 0
\]

as \( \tau \to 0 \).

**Proof.** Let \( (\bar{U}_\tau^n)_{n=1}^N \subset D \) be the approximate values obtained from the variational approximation scheme (2.6) and let \( (\xi_\tau^n)_{n=1}^N \subset V^* \) be the associated subgradients. Then, as already mentioned, by Lemma 2.1 the approximate value \( \bar{U}_\tau^n \) solves the EULER–LAGRANGE equation (2.4), viz.

\[
S_\tau^n = \frac{V^n - V^{n-1}}{\tau} - \xi_\tau^n \in \partial \Psi(\bar{V}_\tau^n) \quad \text{and} \quad \xi_\tau^n \in \partial \mathcal{E}_\tau^n(\bar{U}_\tau^n),
\]

(2.18)
where \( S^n_\tau := f^n_\tau - B(t_{n-1}, U^n_{\tau-1}, V^n_{\tau-1}) \). Due to Lemma 2.2, the first inclusion is equivalent to
\[
\Psi(V^n_\tau) + \Psi^\ast\left( S^n_\tau - \frac{V^n_\tau - V^n_{\tau-1}}{\tau} - \xi^n_\tau \right) = \left\langle S^n_\tau - \frac{V^n_\tau - V^n_{\tau-1}}{\tau} - \xi^n_\tau, V^n_\tau \right\rangle
\] (2.19)
and the second one implies
\[
- \left\langle \xi^n_\tau, U^n_\tau - U^n_{\tau-1} \right\rangle \leq \mathcal{E}_{t_{n-1}}(U^n_{\tau-1}) - \mathcal{E}_{t_n}(U^n_\tau) + \frac{\lambda}{2} \| U^n_\tau - U^n_{\tau-1} \|^2
\]
\[
= \mathcal{E}_{t_{n-1}}(U^n_{\tau-1}) - \mathcal{E}_{t_n}(U^n_\tau) + \int_{t_{n-1}}^{t_n} \partial_r \mathcal{E}_r(U^n_{\tau}) \, dr + \frac{\lambda}{2} \| U^n_\tau - U^n_{\tau-1} \|^2
\] (2.20)
for all \( n = 1, \ldots, N \). Using the identity
\[
\left( V^n_\tau - V^n_{\tau-1}, V^n_\tau \right) = \frac{1}{2} \left( |V^n_\tau|^2 - |V^n_{\tau-1}|^2 + |V^n_\tau - V^n_{\tau-1}|^2 \right),
\]
and the fact that \( \langle w, v \rangle = \langle w, v \rangle \) whenever \( w \in H \), we end up with
\[
\frac{1}{2} |V^n_\tau|^2 + \mathcal{E}_{t_{n-1}}(U^n_{\tau-1}) + \tau \Psi(V^n_\tau) + \tau \Psi^\ast\left( S^n_\tau - \frac{V^n_\tau - V^n_{\tau-1}}{\tau} - \xi^n_\tau \right) - \tau \left\langle S^n_\tau, V^n_\tau \right\rangle
\]
\[
\leq \frac{1}{2} |V^n_\tau|^2 + \mathcal{E}_{t_{n-1}}(U^n_{\tau-1}) + \int_{t_{n-1}}^{t_n} \partial_r \mathcal{E}_r(U^n_{\tau}) \, dr + \frac{\lambda}{2} \| U^n_\tau - U^n_{\tau-1} \|^2
\] (2.21)
for all \( n = 1, \ldots, N \), which, by summing up the inequalities, implies (2.13). In order to show the bounds (2.14) and (2.15), we make use of the following estimates: First,
\[
\tau \left\langle S^n_\tau, V^n_\tau \right\rangle \leq c \tau \Psi(V^n_\tau) + c \Psi^\ast\left( \frac{-B(t_{n-1}, U^n_{\tau-1}, V^n_{\tau-1})}{c} \right) + \frac{\tau}{2} |f^n_\tau|^2 (1 + |V^n_\tau|^2)
\]
\[
\leq c \tau \Psi(V^n_\tau) + \tau \beta (1 + \mathcal{E}_{t_{n-1}}(U^n_{\tau-1}) + |V^n_{\tau-1}|^2) + \tau \tilde{c} \Psi(V^n_{\tau-1})
\]
\[
+ \frac{\tau}{2} |f^n_\tau|^2 (1 + |V^n_\tau|^2),
\] (2.22)
where we used Assumption (2.Bb) and \( U^n_{\tau-1} - \tau V^n_{\tau-1} = U^n_{\tau-2} \) in the last inequality. Second, by the uniform strong positivity of the dissipation potential
\[
\mu \| U^n_\tau - U^n_{\tau-1} \|^2 \leq \Psi(U^n_\tau - U^n_{\tau-1}) \leq \tau \Psi\left( \frac{U^n_\tau - U^n_{\tau-1}}{\tau} \right)
\]
for \( \tau \leq 1 \), where we used the fact that \( \tau \mapsto \tau \Psi(\frac{\tau}{\tau}) \) is monotonically decreasing. Finally, we use the estimate following from the energetic control of power (2.Ed)
\[
\int_{t_{n-1}}^{t_n} \partial_r \mathcal{E}_r(U^n_{\tau}) \, dr \leq \int_{t_{n-1}}^{t_n} C \mathcal{E}_r(U^n_{\tau}) \, dr \leq C \int_{t_{n-1}}^{t_n} \mathcal{G}(U^n_{\tau}) \, dr.
\]
Inserting all preceding inequalities in (2.21) and summing up all inequalities from 1 to \( n \), there exist constants \( C_2, C_3 > 0 \) such that
\[
\frac{1}{2} |V^n_\tau|^2 + \frac{1}{C_1} \mathcal{G}(U^n_\tau) + \int_0^{t_n} \left( (1 - \alpha) \Psi(\nabla_\tau(r)) + \Psi^\ast\left( S_\tau(r) - \nabla_\tau(r) - \xi_\tau(r) \right) \right) \, dr
\]
\[
\leq C_3 \left( |v_0|^2 + \mathcal{E}_0(u_0) + T + \| f \|_{L^1(0,T;H)} + \Psi(v_0) \right)
\]
\[
+ C_2 \int_0^{t_n} \left( (1 + |f_\tau(r)|) |\nabla_\tau(r)|^2 + \mathcal{G}(\nabla_\tau(r)) \right) \, dr,
\]
where $\alpha := c + \tilde{c} + \tau \frac{1}{2p} < 1$ for $\tau < \tau^* := \min\{\frac{2M}{N}(1 - c - \tilde{c}), r^*, 1\}$ with $r^* > 0$ from Assumption (2.Bb). Then by the GRONWALL lemma there exists a constant $M > 0$ such that (2.14) and (2.15) are satisfied. Further, due to the coercivity of $\Psi$ and $\Psi^*$, the uniform boundedness of $(\overline{V}_{\tau})_{0 < \tau \leq \tau^*} \subset L^2(0, T; V)$ and $(S_{\tau} - \tilde{V}_{\tau} - \xi_{\tau})_{0 < \tau \leq \tau^*} \subset L^2(0, T; V^*)$ follows. The boundedness of $(B_{\tau})_{0 < \tau \leq \tau^*} \subset L^2(0, T; V^*)$ uniformly in $\tau$ is a consequence of Assumption (2.Bb): for $M > 0$, there exists $\mu_M > 0$ such that

$$
\mu_M \int_0^T \|B_{\tau}(r)\|_*^2 \, dr \leq \int_0^T \Psi^* (B(U_c(r), U_{\tau}(r), \overline{V}_c(r))) \, dr \\
\leq \int_0^T \left( \beta(1 + \xi_{\tau}(r))(\overline{U}_c(r)) + \|\nabla_{\tau}(r)\|^2 + \tilde{c} \Psi(\overline{U}_c(r)) \right) \, dr \\
\leq \text{const.}
$$

uniformly in $\tau$. Since $(f_{\tau})_{0 < \tau \leq \tau^*}$ is uniformly integrable in $L^2(0, T; H)$, we deduce that $(\tilde{V}_{\tau} + \chi_{\tau})_{0 < \tau \leq \tau^*}$ is uniformly bounded in $L^2(0, T; V^*)$ with respect to $\tau$ as well. Finally, Assumption (2.Eg) implies uniform bounds for $(\xi_{\tau})_{0 < \tau \leq \tau^*}$ and $(\tilde{V}_{\tau})_{0 < \tau \leq \tau^*}$ in $L^\infty(0, T; U^*)$ and $L^2(0, T; U^*)$, respectively. It remains to show the uniform convergences (2.17). But this follows immediately from the uniform integrability of $(\tilde{V}_{\tau})_{0 < \tau \leq \tau^*} \subset L^2(0, T; U^*)$ and $(\overline{V}_{\tau})_{0 < \tau \leq \tau^*} \subset L^2(0, T; V)$ since there holds

$$
\|\tilde{U}_{\tau}(t) - \overline{U}_{\tau}(t)\| \leq \|\overline{U}_{\tau}(t) - \overline{\nabla}_{\tau}(t)\| = \int_{\tau(t)}^{\overline{\tau}(t)} \|\tilde{U}_{\tau}(r)\| \, dr \quad \text{and}
$$

$$
\|\tilde{V}_{\tau}(t) - \overline{V}_{\tau}(t)\|_{U^*} \leq \|\overline{V}_{\tau}(t) - \overline{V}_{\tau}(t)\|_{U^*} = \int_{\tau(t)}^{\overline{\tau}(t)} \|\tilde{V}_{\tau}(r)\|_{U^*} \, dr
$$

for all $t \in [0, T]$. \hfill \Box

### 2.4 Limit passage and completion of the proof

This section is devoted to the existence of convergent subsequences of the approximate solutions in some proper BOCHNER spaces in order to pass to the limit in the discrete inclusion (2.4) as the step size vanishes. As we will see, we will indeed obtain a solution to the CAUCHY problem (1.1). For this purpose, we will make use of compactness properties of bounded sets in reflexive and separable spaces with respect to the weak topology. We elaborate on this in the next result.

**Lemma 2.9 (Compactness).** Under the same assumptions of Lemma 2.8, let $(\tau_n)_{n \in \mathbb{N}}$ be a vanishing sequence of positive real numbers and let $u_0 \in D$ and $v_0 \in V$. Then there exists a subsequence, still denoted by $(\tau_n)_{n \in \mathbb{N}}$, a pair of functions $(u, \xi)$ fulfilling $u(0) = u_0$ in $U$, $u'(0) = v_0$ in $V$ and $u \in C_{w}([0, T]; U) \cap AC([0, T]; V) \cap C^1([0, T]; H) \cap H^2(0, T; U^*)$, \hfill \Box
\[ \xi \in L^{\infty}(0, T; U^*) \text{ such that the following convergences hold} \]

\[
\begin{align*}
U_{r_n}, \bar{U}_{r_n}, \hat{U}_{r_n} & \rightharpoonup u \quad \text{in } L^{\infty}(0, T; U), \\
\bar{U}_{r_n} & \to u \quad \text{in } C_w([0, T]; U), \\
U_{r_n}(t), \bar{U}_{r_n}(t) & \to u(t) \quad \text{in } U \text{ for all } t \in [0, T], \\
\hat{U}_{r_n} & \to u \quad \text{in } C([0, T]; W), \\
\hat{U}_{r_n} & \rightharpoonup^* u \quad \text{in } W^{1,\infty}(0, T; H), \\
\nabla_{r_n}, \bar{V}_{r_n} & \to u' \quad \text{in } L^2(0, T; V), \\
\bar{V}_{r_n}, \hat{V}_{r_n} & \to u' \quad \text{in } L^p(0, T; H) \text{ for all } p \geq 1 \text{ with } u' \in C([0, T]; H), \\
\hat{V}_{r_n} & \rightharpoonup^* u' \quad \text{in } L^\infty(0, T; H) \cap H^1(0, T; U^*), \\
\nabla_{r_n}(t), \bar{V}_{r_n}(t) & \to u'(t) \quad \text{in } H \text{ for a.a. } t \in (0, T), \\
\hat{V}_{r_n} & \to u' \quad \text{in } C_w([0, T]; H), \\
\nabla_{r_n}(t), \nabla_{r_n}(t) & \to u'(t) \quad \text{in } H \text{ for all } t \in [0, T], \\
\xi_{r_n} & \rightharpoonup u \quad \text{in } L^\infty(0, T; U^*), \\
B_{r_n} & \to B(\cdot, u(\cdot), u'(\cdot)) \quad \text{in } L^2(0, T; V^*). 
\end{align*}
\]

**Proof.** Let \( \bar{U}_{r_n}, \nabla_{r_n}, \bar{V}_{r_n}, \nabla_{r_n}, \xi_{r_n} \) as well as \( f_{r_n} \) be the interpolation functions with the initial values \( u_0 \in D \) and \( v_0 \in V \) as defined in (2.10)-(2.12). Since all spaces were supposed to be separable and reflexive, we note that if a \textsc{Banach} space \( X \) is separable and reflexive, the spaces \( L^p(0, T; X) \) become for \( 1 < p < \infty \) also separable and reflexive, whereas \( L^{\infty}(0, T; X) \) becomes the dual space of the separable space \( L^1(0, T; X^*) \). So, as a consequence of the \textsc{Banach–Alaoglu} theorem, bounded sets in \( L^p(0, T; X) \), \( 1 < p < \infty \), and \( L^{\infty}(0, T; X) \) are relatively compact with respect to the weak and weak* topology, respectively. In view of the a priori estimates (2.14) and (2.15) and Assumption (2.Ec), this implies that the sequence \( (\bar{U}_{r_n}) \) is bounded in \( L^{\infty}(0, T; U) \). Together with the bounds (2.16), this already yields the existence of convergence subsequences (denoted as before) such that (2.23a), (2.23e),(2.23k),(2.23h) and (2.23l) hold. We remark, that the limit functions can be identified with \( u \) and \( u' \) by standard arguments. The convergence of \( \hat{U}_{r_n} \to u \) in \( C_w([0, T]; V) \) follows directly from Lemma A.2.4 in [Puh16] stating that for a bounded sequence in \( L^{\infty}(0, T; V) \) having weak derivatives which are bounded in \( L^p(0, T; V) \), \( p \in (1, \infty) \), there exists a subsequence converging against an absolutely continuous function with respect to the topology of \( C_w([0, T]; V) \). Then, using the uniform boundedness of the interpolation functions in \( U \), i.e., (2.14) and the convergence (2.23a), we obtain (2.23b) and (2.23c). Along the same lines, there holds (2.23j) and (2.23k). In order to show (2.23d), we make use of the \textsc{Arzelà–Ascoli} theorem, i.e., we show that the sequence of piecewise linear interpolations are pointwise contained in a relatively compact subset of the space \( W \) and are pointwise equi-continuous. The former follows from the boundedness of \( \hat{U}_{r_n}(t) \) in \( U \) uniformly in \( t \in [0, T] \) and for all \( n \in \mathbb{N} \) following from the a priori estimates and the compact embedding of \( U \) in \( W \). The latter, on the other hand, is a consequence of the estimate

\[
\begin{align*}
\|\hat{U}_{r_n}(t) - \hat{U}_{r_n}(s)\|_W & \leq \int_s^t \|\hat{U}_{r_n}'(r)\|_W \, dr \leq |t - s|^{1/2} \left( \int_s^t \|\hat{U}_{r_n}'(r)\|_W^2 \, dr \right)^{1/2} \\
& \leq |t - s|^{1/2} \|\hat{U}_{r_n}'\|_{L^2(0, T; W)} \leq C|t - s|^{1/2} \|\hat{U}_{r_n}\|_{L^2(0, T; V)} \\
& \leq |t - s|^{1/2} C M,
\end{align*}
\]
which in turn again follows from the a priori estimates. Now, we seek to apply theorem (A.2) to the sequence $\nabla_{\tau_n}$ with $M_+ = V, B = H$ and $Y = U^*$. Assumption i) and ii) theorem (A.2) are obviously fulfilled. Assumption iii) follows for $p = \infty$ directly from the a priori estimate (2.14), whereas the compliance of assumption iv) for $r = 2$ can be seen by

$$\|\sigma_{\tau_n} \nabla_{\tau_n} - \nabla_{\tau_n}\|_{L^2(0,T-\tau_n;U^*)} = \tau_n \|\nabla_{\tau_n}'\|_{L^2(0,T;U^*)} \leq \tau_n M,$$

and hence (2.23g). This, in turn, implies pointwise convergence of the very sequence almost everywhere in $[0,T]$, i.e., (2.23i). The assertion for $\tilde{V}_{\tau_n}$ can be shown analogously. Further, from the convergences (2.23d), (2.23i), (2.17) and Assumptions (2.B), we deduce with LEBESGUES dominated convergence theorem the strong convergence of the perturbation (2.23m). Finally, thanks to (2.23b) and (2.23i), the initial conditions are also fulfilled by $u$ and $u'$.

We conclude this section with the proof of Theorem 2.6. We first show that the limit function obtained from the previous lemma is indeed a solution to the CAUCHY problem. Let $u_0 \in D, v_0 \in H$ and the vanishing sequence of step sizes $(\tau_n)_{n \in \mathbb{N}}$ be given. We remark that for the a priori estimates, we necessarily needed the initial datum $v_0$ to be in $V$. Therefore, let $(v^k_n)_{k \in \mathbb{N}} \subset V$ be a sequence such that $v^k_n \to v_0$ in $H$. Henceforth, we assume $k \in \mathbb{N}$ to be fixed and we define the interpolation functions associated to $u_0$ and $v^k_k$ as in the previous lemma omitting for notational convenience the dependence on $k$. Then, we obtain again by the previous lemma the existence of convergent subsequences (labeled as before) of the interpolation functions and a limit function $u \in L^\infty(0,T;U) \cap H^1(0,T;V) \cap C^1([0,T];H) \cap H^2(0,T;U^*)$ with $u(0) = u_0 \in U$ and $u'(0) = v^k_0 \in V$, where again we omit the dependence of the limit function on $k$. Now, the discrete inclusion (2.18) reads in the weak formulation

$$\int_0^T \langle S_{\tau_n}(r) - \tilde{V}_{\tau_n}'(r) - \xi_{\tau_n}(r) - A\nabla_{\tau_n}(r), w(r) \rangle \, dr = 0 \quad \text{for all } w \in L^2(0,T;V),$$

where $S_{\tau_n} = f_{\tau_n} - B_{\tau_n}$. Since $\Psi$ is defined by a uniformly positive quadratic form, the FRÉCHET derivative is a linear bounded and strongly positive operator $A : V \to V^*$ which implies that the associated NEMITSKIJ operator $A : L^2(0,T;V) \to L^2(0,T;V^*)$ is well defined, linear and bounded. Therefore the NEMITSKIJ operator is weak-to-weak continuous and we can pass to the limit. Also the convergence of $S_{\tau_n}$ to $f - B(\cdot,u(\cdot),u'(\cdot))$ strongly $L^2(0,T;V^*)$ as $n \to \infty$ is justified by the previous lemma. Since $\tilde{V}_{\tau_n}' \to u''$ in $L^2(0,T;U^*)$ and $\xi_{\tau_n} \to \xi$ in $L^\infty(0,T;U^*)$, we are able to identify the weak limit of $(\tilde{V}_{\tau_n}' + \xi_{\tau_n})$ in $L^2(0,T;V^*)$ as $(u'' + \xi) \in L^2(0,T;V^*)$ although each term lies in the weaker space $L^2(0,T;U^*)$. Therefore, we are allowed to pass to the limit in the weak formulation there as well. Then, by a density argument and the fundamental lemma of calculus of variations, we deduce

$$u''(t) + Au'(t) + \xi(t) + B(t,u(t),u'(t)) = f(t) \quad \text{in } V^* \text{ a.a. in } (0,T). \quad (2.24)$$

It remains to show that $\xi(t) \in \partial E(u(t))$ in $U^*$ a.a. in $(0,T)$. In order to show this, we employ the closedness condition (2.Ee) which mimics the result for maximal monotone operators in Appendix A.1. Since we have already shown all assumptions to hold true except for Assumption d) and e). Assumption d) follows immediately from

$$\|\sigma_{\tau_n} \nabla_{\tau_n} - \nabla_{\tau_n}\|_{L^2(0,T-\tau_n;V)} = \tau_n \|\nabla_{\tau_n}'\|_{L^2(0,T;V)} \leq \tau_n M,$$
whereas Assumption e) is verified by the following calculations: Let $t \in [0, T]$. Then, we have
\[
\int_0^t \langle \xi_{\tau_n}(r), \overline{U}_{\tau_n}(r) \rangle_{U^* \times U} \, dr = \int_0^t \langle S_{\tau_n}(r) - \tilde{V}_{\tau_n}'(r) - A\overline{V}_{\tau_n}(r), \overline{U}_{\tau_n}(r) \rangle_{U^* \times U} \, dr \\
= \int_0^t \langle S_{\tau_n}(r), \overline{U}_{\tau_n}(r) \rangle_{U^* \times U} \, dr \\
- \int_0^t \langle \tilde{V}_{\tau_n}'(r), \overline{U}_{\tau_n}(r) \rangle_{U^* \times U} \, dr \\
- \int_0^t \langle A\overline{V}_{\tau_n}(r), \overline{U}_{\tau_n}(r) \rangle_{U^* \times U} \, dr \\
=: I_1 + I_2 + I_3.
\]

The convergence of the first integral is due to the strong convergence of $S_{\tau_n}$ against $f - B(\cdot, u(\cdot), u'(\cdot))$ in $L^2(0, T; V^*)$. For the second integral, we observe, that by the discrete integration by parts rule, there holds
\[
- \int_0^t \langle \tilde{V}_{\tau_n}'(r), \overline{U}_{\tau_n}(r) \rangle_{U^* \times U} \, dr \\
= \int_{\tau_n}^t \langle \tilde{U}_{\tau_n}'(r), \tilde{V}_{\tau_n}'(r - \tau_n) \rangle \, dr - \langle \overline{U}_{\tau_n}(t), \tilde{V}_{\tau_n}'(t) \rangle - \langle \overline{U}_{\tau_n}(\tau_n), v_0 \rangle.
\]
Thus, by (2.23d), (2.23g) and (2.23k)
\[
\lim_{n \to \infty} I_2 = \int_0^t (u'(r), u'(r)) \, dr - (u(t), u'(t)) + (u_0, v_0) = \int_0^t (u''(r), u(r))_{U^* \times U} \, dr.
\]

Finally, observing that
\[
\int_0^t \langle A\overline{V}_{\tau_n}(r), \overline{U}_{\tau_n}(r) \rangle_{U^* \times U} \, dr = \int_0^t \langle A\overline{U}_{\tau_n}(r), \overline{V}_{\tau_n}(r) \rangle_{V^* \times V} \, dr \\
= \sum_{k=1}^N \langle A\overline{U}_{\tau_n}^k, \overline{U}_{\tau_n}^k - \overline{U}_{\tau_n}^{k-1} \rangle \\
\leq \sum_{k=1}^N \left( \psi(\overline{U}_{\tau_n}^k) - \psi(\overline{U}_{\tau_n}^{k-1}) \right) \\
= \psi(\overline{U}_{\tau_n}^N) - \psi(u_0) \\
= \psi(\overline{U}_{\tau_n}(\bar{t}_{\tau_n}(t))) - \psi(u_0),
\]
we obtain by the weakly lower semicontinuity of the dissipation potential and its differentiability
\[
\limsup_{n \to \infty} I_3 \leq \limsup_{n \to \infty} \left( \psi(u_0) - \psi(\overline{U}_{\tau_n}(\bar{t}_{\tau_n}(t))) \right) \\
= - \liminf_{n \to \infty} \left( \psi(\overline{U}_{\tau_n}(\bar{t}_{\tau_n}(t))) - \psi(u_0) \right) \\
\leq \psi(u_0) - \psi(u(t)) \\
\leq \int_0^t \langle Au(r), u'(r) \rangle \, dr \\
= \int_0^t \langle Au'(r), u(r) \rangle \, dr.
\]
Hence, by the closedness condition (2Ec), there holds \( \xi(t) \in \partial E(u(t)) \) in \( U^* \) and \( \mathcal{E}_{\tau_n(t)}(\overline{U}_{\tau_n(t)}) \to \mathcal{E}_{t}(u(t)) \) as well as
\[
\limsup_{n \to \infty} \partial \mathcal{E}_{\tau_n(t)}(u_n(t)) \leq \partial \mathcal{E}_{t}(u(t))
\] (2.25)
for a.a. \( t \in (0, T) \). It remains to show that the energy dissipation inequality holds. Let \( t \in [0, T] \) and \( \mathcal{N} \subset (0, T) \) a set of measure zero such that \( \mathcal{E}_{\tau_n(s)}(\overline{U}_{\tau_n(s)}) \to \mathcal{E}_{t}(u(s)) \) and \( V_{\tau_n(s)} \to u'(s) \) for each \( s \in [0, T] \setminus \mathcal{N} \). Then, exploiting the convergences (2.23) and (2.25) as well as the condition (2Ed), we obtain from the discrete energy dissipation inequality
\[
\frac{1}{2}|u'(t)|^2 + \mathcal{E}_{t}(u(t)) + \int_{s}^{t} (\Psi(u'(r)) + \Psi^*(S(r) - \xi(r) - u''(r))) \, dr
\]
\[
\leq \liminf_{n \to \infty} \left( \frac{1}{2} |V_{\tau_n(t)}|^2 + \mathcal{E}_{\tau_n(t)}(\overline{U}_{\tau_n(t)})
\right.
\]
\[
+ \int_{\tau_n(s)}^{\tau_n(t)} (\Psi(V_{\tau_n(r)}) + \Psi^*(S_{\tau_n(r)} - \hat{V}_{\tau_n(r)}') - \xi_{\tau_n(r)}) \, dr
\]
\[
\leq \limsup_{n \to \infty} \left( \frac{1}{2} |V_{\tau_n(s)}|^2 + \mathcal{E}_{\tau_n(s)}(\overline{U}_{\tau_n(s)})
\right)
\]
\[
+ \int_{\tau_n(s)}^{\tau_n(t)} \partial \mathcal{E}_{r}(\overline{U}_{\tau_n(r)}) \, dr + \int_{\tau_n(s)}^{\tau_n(t)} (\mathcal{E}_{r}(u(r)) \, dr
\]
\[
= \frac{1}{2}|u'(s)|^2 + \mathcal{E}_{s}(u(s)) + \int_{s}^{t} \partial \mathcal{E}_{r}(u(r)) \, dr + \int_{s}^{t} (\mathcal{E}_{r}(u'(r)) \, dr,
\]
where \( S(r) = f(r) - B(r, u(r), u'(r)) \). The last step in the proof relies in showing, that there exists also a solution to (1.1) with the initial conditions \( u(0) = u_0 \) and \( u'(0) = \nu_0 \). We remember, that for each \( k \in \mathbb{N} \), there exists a solution \( u_k \) with \( u_k(0) = u_0 \) and \( u'_k(0) = \nu_0 \), where \( \nu_k \to \nu_0 \) in \( H \) as \( k \to \infty \). Let \( (u_k)_{k \in \mathbb{N}} \) be a sequence of such solutions. From the discrete energy dissipation balance, we find in an analogous way as before the same bounds for \( u_k \) in suitable Bochner spaces so that the rest of the proof works in exactly the same manner which show the assertion.

\[ \square \]

2.5 Application

In this section we want to apply the abstract result to a concrete non-trivial example which does not fit into the framework of the cited authors. Let \( \Omega \subset \mathbb{R}^d \) be a bounded domain with a Lipschitz boundary. We consider the Cauchy problem to the doubly nonlinear evolution equation of second order
\[
(P) \begin{cases}
-\Delta u_t - \Delta P u + (u^2 - 1) u \pm |u|^{q-1} \pm |u_t|^{r-1} = 0 \quad \text{in } \Omega \times (0, T), \\
u(0, x) = u_0(x) \quad \text{on } \Omega, \\
u'(0, x) = \nu_0(x) \quad \text{on } \Omega, \\
u(t, x) = 0 \quad \text{on } \partial \Omega \times [0, T].
\end{cases}
\]
Choosing \( U = \mathcal{W}^1_0(\Omega), V = \mathcal{H}^1_0(\Omega), W = \mathcal{L}^q(\Omega) \) and \( H = \mathcal{L}^2(\Omega) \), the dissipation potential \( \Psi : V \to \mathbb{R} \) and the energy functional \( \mathcal{E} : V \to (-\infty, +\infty] \) are given by
\[
\Psi(v) = \frac{1}{2} \int_{\Omega} |\nabla v(x)|^2 \, dx \quad \text{and} \quad \mathcal{E}(u) = \int_{\Omega} \left( \frac{1}{p} |\nabla u(x)|^p + \frac{1}{4} (u^2 - 1)^2 \right) \, dx,
\]
respectively, whereas the perturbation \( B : W \times H \rightarrow V^* \) is defined by
\[
\langle B(u, v), w \rangle_{V^* \times V} = \langle B(u, v), w \rangle_{W^* \times W} = \int_{\Omega} (\pm |u(x)|^{q-1} \pm |v(x)|^{r-1}) w(x) \, dx.
\]
Choosing, e.g.,
\[
d \geq 3, \, p \in (d, +\infty), \, q \in (1, p/2 + 1) \quad \text{and} \quad r \in (1, 2],
\]
by the RELLICh–KONDRAChOV theorem, \( U \) and \( V \) are compactly embedded in \( W \) and \( H \), respectively. Since the dissipation potential is state independent, it is induced by the bilinear form \( a : V \times V \rightarrow \mathbb{R} \)
\[
a(v, w) := \frac{1}{2} \int_{\Omega} \nabla v \cdot \nabla w \, dx
\]
and therefore satisfies all conditions. The conditions (2.Eb)-(2.Ed) are obviously fulfilled by the energy functional. In order to verify (2.Ea), we note that every convex and lower semicontinuous functional on a BANACH space is weakly lower semicontinuous, and that a functional is lower semicontinuous with respect to a given topology if and only if its sublevel sets are closed with respect to that topology. Knowing that, we observe that the energy
\[
\mathcal{E}(u) = \int_{\Omega} \left( \frac{1}{p} |\nabla u(x)|^p + \frac{1}{4} (u^2 - 1)^2 \right) \, dx
\]
\[
= \int_{\Omega} \left( \frac{1}{p} |\nabla u(x)|^p + \frac{1}{4} (u^4 - 2u^2 + 1) \right) \, dx = W(u) - \frac{1}{2} \int_{\Omega} u^2 \, dx \tag{2.26}
\]
is the sum of a convex and a concave function on \( V \) where the concave function is due to the compact embedding of \( H^1_0 \) in \( L^2 \) continuous with respect to the weak topology. This implies \( \mathcal{E} \) to be weakly lower semicontinuous on \( V \). In fact, the convex part of the energy is perturbed by the negative HILBERT space norm of \( L^2 \) squared which by the parallelogram law and the embedding \( H^1_0 \rightarrow L^2 \) leads to the \( \Lambda \)-convexity of \( \mathcal{E} \) with \( \Lambda := C \) being the constant of the very same embedding. Now, we show the closedness property (2.Ec). First, we note that for each \( u \in D(\partial \mathcal{E}) \), there holds \( \xi \in \partial \mathcal{E}(u) = \partial W(u) - u \) if and only if \( \xi = -\text{div}(|\nabla u|^{p-2} \nabla u) + (u^2 - 1)u = \zeta - u \in V^* \), where \( \zeta \in \partial W(u) \). Now, let \( u_n \rightarrow^* u \) in \( L^\infty(0, T; U) \cap H^1(0, T; V) \) and \( \xi_n \rightarrow^* \xi \in L^2(0, T; U^*) \) such that \( \xi_n(t) \in \partial \mathcal{E}(u_n(t)) \) for almost all \( t \in (0, T) \) and
\[
\limsup_{n \rightarrow \infty} \int_0^T \langle \xi_n(t), u_n(t) \rangle_{U^* \times U} \, dt \leq \int_0^T \langle \xi(t), u(t) \rangle_{U^* \times U} \, dt. \tag{2.27}
\]
Defining \( \zeta_n := \xi_n + u_n \), there holds \( \zeta_n \in \partial W(u_n) \) and \( \zeta_n \rightarrow \zeta := \xi + u \in L^2(0, T; U^*) \). By the LIONS-AUBIN lemma, we obtain strong convergence of \( u_n \rightarrow u \) in \( C([0, T]; L^2) \). Thus, in view of (2.28), we deduce
\[
\limsup_{n \rightarrow \infty} \int_0^T \langle \zeta_n(t), u_n(t) \rangle_{U^* \times U} \, dt \leq \int_0^T \langle \zeta(t), u(t) \rangle_{U^* \times U} \, dt. \tag{2.28}
\]
Since \( W \) is convex, by Lemma A.1., Lemma A.2. and Theorem A.4., there holds \( \zeta(t) \in \partial W(u(t)) \) in \( W^{-1,p'} \) and \( W(u_n(t)) \rightarrow W(u(t)) \) as \( n \rightarrow \infty \) for a.a. \( t \in (0, T) \). Therefore
The control of the subgradient of $\mathcal{E}$ follows from the following calculations: Let $u \in D(\partial \mathcal{E})$ and $\xi \in \partial \mathcal{E}$. Then,

$$
\langle \xi, v \rangle_{U^* \times U^*} = \int_{\Omega} \left( |\nabla u|^p - 2 \nabla u \cdot \nabla v + (u^2 - 1)uv \right) dx 
\leq C \left( \left\| u \right\|_{W_0^{1,p}}^{p-1} + \left\| (u^2 - 1)u \right\|_{L^{(p+1)/p}} \right) \left\| v \right\|_{W_0^{1,p}} 
\leq C \left( 1 + \frac{1}{p} \left\| u \right\|_{W_0^{1,p}}^p + \left\| u \right\|_{W_0^{1,p}} \right) \left\| v \right\|_{W_0^{1,p}} 
= C \left( 1 + \mathcal{E}(u) + \left\| u \right\|_{u} \right) \left\| v \right\|_{W_0^{1,p}} \quad \text{for all } v \in W_0^{1,p},
$$

where $C > 0$ denotes a generic constant, whence (2.Eg). Finally, we verify the assumptions on the perturbation $B$. The continuity condition (2.Ba) can easily be checked. Ad (2.Bb): Let $u \in D(\mathcal{E})$ and $v, w \in V$. Then, by the Hölder and Young inequality as well as the Sobolev embedding theorem, there holds

$$
\langle B(u, v), w \rangle_{V^* \times V} = \int_{\Omega} (\pm |u(x)|^{q-1} \pm |v(x)|^{r-1})w(x) dx 
\leq C \left( \left\| u \right\|_{L^{q-1}(\Omega)}^{q-1} + \left\| v \right\|_{L^{r-1}(\Omega)}^{r-1} \right) \left\| w \right\|_{L^{2d/(d-2)}} 
\leq C \left( \left\| u \right\|_{W_0^{1,p}}^{q-1} + \left\| v \right\|_{W_0^{1,p}}^{r-1} \right) \left\| w \right\|_{H_0^1} 
\leq c (C(1 + \mathcal{E}(u)) + \tilde{c} \mathcal{E}(v))^{1/2} \left\| w \right\|_{H_0^1},
$$

where again $C > 0$ denotes a generic constant and $c, \tilde{c} \in (0, 1)$ with $c + \tilde{c} < 1$. Noticing that the convex conjugate of $\Psi$ is given by $\Psi^*(\xi) = \frac{1}{2} \left\| \xi \right\|_s^2$ for all $\xi \in V^*$, we conclude (2.Bb). Therefore, we obtain for any initial data $u_0 \in D(\mathcal{E}) = W_0^{1,p} \cap L^4$ and $v_0 \in L^2$ the existence of a solution

$$
u \in L^\infty(0, T; W_0^{1,p}) \cap H^1(0, T; H^1) \cap C^1([0, T]; L^2) \cap H^2(0, T; W^{-1,p'}),$$

to (P) a.e. in time in $H^{-1}$ such that the energy dissipation inequality

$$
\frac{1}{2} \left\| u'(t) \right\|_{0,2}^2 + \mathcal{E}(u(t)) + \int_s^t (\Psi(u'(r)) + \Psi^*(B(u(r), u'(r)) - u''(r) - \xi(r)) dr 
\leq \frac{1}{2} \left\| u'(s) \right\|_{0,2}^2 + \mathcal{E}(u(s)) + \int_s^t \langle B(u(r), u'(r)), u'(r) \rangle_{H^{-1} \times H_0^1} dr
$$

holds for all $t \in [0, T]$ for $s = 0$ and a.a. $s \in (0, t)$, where $\xi(t) = -\Delta_p u(t)$ for a.a. $t \in (0, T)$.

### A Appendix

#### A.1 Maximal monotone operators

In this section, we collect some important results from the theory of maximal monotone operators. We shall henceforth assume that $B$ is a real Banach space with its dual space denoted by $B^*$. Furthermore, we denote by $\langle \cdot, \cdot \rangle : B^* \times B \to \mathbb{R}$ the dual pairing between $B^*$ and $B$. We recall that a multivalued operator $A : B \to 2^{B^*}$ is said to be monotone if for all $v_1, v_2 \in D(A) := \{ v \in B \mid A(v) \neq \emptyset \}$ and all $x_i \in A(v_i), i = 1, 2$, there holds

$$
\langle x_1 - x_2, v_1 - v_2 \rangle \geq 0.
$$
A monotone operator is said to be maximal monotone if its graph is not contained properly in the graph of any other monotone operator.

Often in nonlinear problems, a closedness property of nonlinear operators is needed in order to identify weak limits. The following result, given by H. Brézis, M.G. Crandall and A. Pazy, states under which conditions maximal monotone operator is closed with respect to an appropriate topology.

**Lemma A.1.** Let $B$ be a real reflexive Banach space. Let $A : B \to 2^{B^*}$ be maximal monotone, and let $\xi_n \in A(v_n)$ and $\xi \in A(v)$ be such that $\xi_n \rightharpoonup \xi$ in $B$ and $v_n \to v$ in $B^*$ as $n \to \infty$. Moreover, let either

$$\limsup_{n,m \to \infty} \langle \xi_n - \xi_m, v_n - v_m \rangle \leq 0,$$

or

$$\limsup_{n \to \infty} \langle \xi_n - \xi, v_n - v \rangle \leq 0,$$

then $\xi \in A(v)$ and $\langle \xi_n, v_n \rangle \to \langle \xi, v \rangle$.

**Proof.** Lemma 1.2 in [BCP70].

The following theorem which was proven by R.T. Rockafellar in 1966, reveals an important relation between the subdifferential of a functional and a maximal monotone operator.

**Lemma A.2.** Let $B$ be a real Banach space, and let $f : B \to (-\infty, +\infty]$ be a proper, lower semicontinuous and convex functional. Then, the subdifferential $\partial F$ is maximal monotone.

**Proof.** Theorem 4 in [Roc66].

**Remark A.3.** In fact, the converse holds true if the multivalued operator is maximal cyclically monotone. In this case, there exists a proper convex functional $f$ such that $\partial f = A$, we refer to [Roc66] for more details.

Finally, the following result, given by N. Kenmochi, states to what extent the properties of a functional on a Banach space and its subdifferential is inherited by the Nemitskii operator on a appropriate Bochner space and vice versa.

**Theorem A.4.** Let $B$ be a real reflexive Banach space, and let $f(t, \cdot) : B \to (-\infty, +\infty]$ be for each $t \in [0, T]$ a proper, lower semicontinuous and convex functional such that for each (Bochner) measurable function $v : [0, T] \to B$, the map $t \mapsto f(t, v(t))$ is (Lebesgue) measurable and there exists constants $\alpha, \beta \in \mathbb{R}$ such that

$$f(t, v) + \alpha\|v\|_B + \beta \geq 0 \quad \text{for all } t \in [0, T] \text{ and } v \in V.$$

For $p \in (1, +\infty)$, we define the Functional $F : L^p(0, T; B) \to (-\infty, +\infty]$ by

$$F(u) = \begin{cases} 
\int_0^T f(t, u(t)) \, dt & \text{if } f(\cdot, u(\cdot)) \in L^1(0, T), \\
+\infty & \text{otherwise}.
\end{cases}$$
A.2 Compactness result

Then, $F$ lower semicontinuous and convex with $F(u) > -\infty$ on $L^p(0,T;B)$. Assume further that for each $t \in [0,T]$ and each $z \in B$ with $f(t,z) < +\infty$, there exists a function $v \in L^p(0,T;B)$ such that $v(t) = z$, $f(\cdot,v(\cdot)) \in L^1(0,T)$, $v$ is right-continuous at $t$ and
\[
\limsup_{s \to t} f(s,v(s)) \leq f(t,z).
\]

Let $u \in D(\partial F)$, then
\[
\xi \in \partial_{L^p(0,T;B)}F(u) \subset L^p(0,T;B^*) \text{ if and only if } \xi(t) \in \partial_B f(t,v(t)) \text{ for a.a. } t \in (0,T).
\]

Proof. Proposition 1.1 in [Ken75].

A.2 Compactness result

In this section, we provide a version of Lions–Aubin or Lions–Aubin–Simon lemma, a well-established strong compactness result for Bochner spaces. This version is also known as the Lions–Aubin–Dubinskii lemma and deals with the case of piecewise constant functions in time which avoids the construction of weakly time differentiable functions. Although, we need a special case of this lemma, we want to state it in its full generality.

We call a set $M_+ \subset B$ a seminormed nonnegative cone in a Banach space $B$ if the following conditions hold: for all $u \in M_+$ and $c \geq 0$, there holds $cu \in M_+$, and if there exists a function $[\cdot] : M_+ \to [0,\infty)$ such that $|u| = 0$ iff $u = 0$ and $|cu| = c|u|$ for all $c \geq 0$. Furthermore, we say $M_+$ is continuously embedded in $B$, noting $M_+ \hookrightarrow B$, if there exists $C > 0$ such that
\[
\|u\|_B \leq C[u] \text{ for all } u \in M_+.
\]

**Theorem A.3** [Lions–Aubin–Dubinskii] Let $B$ and $Y$ be Banach spaces and $M_+$ be a seminormed nonnegative cone in $B$. Let either $1 \leq p < \infty$ and $r = 1$ or $p = \infty$ and $r > 1$. Let $(u)_{\tau_n} \subset L^p(0,T;M_+ \cap Y)$ be a sequence of functions being constant on each subinterval $((k-1)\tau_n,k\tau_n]$, $1 \leq k \leq n$, $T = n\tau_n$ such that the following assumptions hold

i) $M_+ \hookrightarrow B$ compactly,

ii) for all sequences $(w_n)_n \subset B \cap Y$ with $w_n \to w$ in $B$ and $w_n \to 0$ in $Y$ imply that $w = 0$,

iii) $(u)_{\tau_n}$ is bounded in $L^p(0,T;M_+)$,

iv) there exists $C > 0$ such that for all $n \in \mathbb{N}$, $\|\sigma_{\tau_n} u_{\tau_n} - u_{\tau_n}\|_{L^p(0,T-\tau_n;Y)} \leq C\tau_n$, where $\sigma_{\tau_n} u := u(\cdot + h)$.

Then, if $p < \infty$, $(u)_{\tau_n}$ is relatively compact in $L^p(0,T;B)$ and if $p = \infty$, there exists a subsequence of $(u)_{\tau_n}$ converging in $L^q(0,T;B)$ for all $1 \leq q < \infty$ to a limit function belonging to $C([0,T];B)$.

**Proof.** Theorem 2.

**Remark A.5.** We note, that the same conclusion can easily shown for piecewise constant functions which are defined to be constant on left closed intervals.
References

[Att84] H. Attouch. *Variational Convergence of Functions and Operators.* Pitman Advanced Publishing Program. Pitman, 1984.

[Bar76] V. Barbu. *Nonlinear Semigroups and Differential Equations in Banach spaces.* Noordhoff, Leyden, 1976.

[BCP70] H. Brézis, M. G. Crandall, and A. Pazy. Perturbations of nonlinear maximal monotone sets in Banach space. *Comm. Pure Appl. Math.*, 23, 123–144, 1970.

[BEM19] A. Bacho, E. Emmrich, and A. Mielke. An existence result and evolutionary $\Gamma$-convergence for perturbed gradient systems. *J. Evol. Equ.*, 19(2), 479–522, 2019.

[BMR12] M. Bulíček, J. Málek, and K. R. Rajagopal. On Kelvin-Voigt model and its generalizations. *Evol. Equ. Control Theory*, 1(1), 17–42, 2012.

[Bre72] H. Brezis. Équations d’évolution du second ordre associées à des opérateurs monotones. *Israel J. Math.*, 12, 51–60, 1972.

[Ekt74] I. Ekeland and R. Temam. *Analyse Convexe et Problèmes Variationnels.* Dunod, 1974.

[EmŠ11] E. Emmrich and D. Šiška. Full discretisation of second-order nonlinear evolution equations: strong convergence and applications. *Comput. Methods Appl. Math.*, 11(4), 441–459, 2011.

[EmŠ13] E. Emmrich and D. Šiška. Evolution equations of second order with nonconvex potential and linear damping: existence via convergence of a full discretization. *J. Differential Equations*, 255(10), 3719–3746, 2013.

[EmT10] E. Emmrich and M. Thalhammer. Convergence of a time discretisation for doubly nonlinear evolution equations of second order. *Found. Comput. Math.*, 10(2), 171–190, 2010.

[EmT11] E. Emmrich and M. Thalhammer. Doubly nonlinear evolution equations of second order: existence and fully discrete approximation. *J. Differential Equations*, 251(1), 82–118, 2011.

[EŠT15] E. Emmrich, D. Šiška, and M. Thalhammer. On a full discretisation for nonlinear second-order evolution equations with monotone damping: construction, convergence, and error estimates. *Found. Comput. Math.*, 15(6), 1653–1701, 2015.

[GGZ74] H. Gajewski, K. Gröger, and K. Zacharias. *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen.* Akademie-Verlag, Berlin, 1974. Mathematische Lehrbücher und Monographien, II. Abteilung, Mathematische Monographien, Band 38.

[Ken75] N. Kenmochi. Some nonlinear parabolic variational inequalities. *Israel J. Math.*, 22(3–4), 304–331, 1975.

[MRS13] A. Mielke, R. Rossi, and G. Savaré. Nonsmooth analysis of doubly nonlinear evolution equations. *Calc. Var. Part. Diff. Equns.*, 46(1-2), 253–310, 2013.
[Puh16] D. Puhst. Zur Existenztheorie nichtlokaler nichtlinearer Evolutionsgleichungen mit Anwendung in der Peridynamik. PhD thesis, Institut für Mathematik, Technische Universität Berlin, 2016.

[Roc66] R. T. Rockafellar. Characterization of the subdifferentials of convex functions. Pacific J. Math., 17, 497–510, 1966.

[RoT17] R. Rossi and M. Thomas. Coupling rate-independent and rate-dependent processes: existence results. SIAM J. Math. Anal., 49(2), 1419–1494, 2017.

[Rou05] T. Roubíček. Nonlinear Partial Differential Equations with Applications. Birkhäuser Verlag, Basel, 2005.

[Rou09] T. Roubíček. Rate independent processes in viscous solids at small strains. Math. Methods Appl. Sci. (MMAS), 32(7), 825–862, 2009. Erratum Vol. 32(16) p. 2176.