Abstract

The existence of strong solutions to the abstract Cauchy problem for the non-linearly damped inertial system

\[
\begin{aligned}
\left\{
\begin{array}{l}
\ddot{u}(t) + \partial \Psi_u(t)(u'(t)) + \partial \mathcal{E}_t(u(t)) + B(t, u(t), u'(t)) \ni f(t), \\
u(0) = u_0, \quad u'(0) = v_0,
\end{array}
\right.
\end{aligned}
\tag{1.1}
\]

on a Gelfand triplet type of framework is proven. Under mild assumptions on the dissipation potential \(\Psi_u\) and differentiability assumption on the energy functional, the existence of strong solutions is shown by establishing the convergence of a semi-implicit time discretization scheme.

1 Introduction

In this article, we investigate the abstract Cauchy problem

\[
\begin{aligned}
\left\{
\begin{array}{l}
\ddot{u}(t) + \partial \Psi_u(t)(u'(t)) + \partial \mathcal{E}_t(u(t)) + B(t, u(t), u'(t)) \ni f(t), \\
u(0) = u_0, \quad u'(0) = v_0,
\end{array}
\right.
\end{aligned}
\tag{1.1}
\]

where again \(\Psi_u\) denotes the dissipation potential, \(\mathcal{E}_t\) the energy functional, \(B\) the perturbation, and \(f\) the external force. Here, the dissipation potential \(\Psi_u\) is, in general, nonlinear, non-quadratic, nonsmooth, and depends nonlinearly on the state \(u\). The energy functional \(\mathcal{E}_t = \mathcal{E}_1^1 + \mathcal{E}_2^2\) is the sum of a functional \(\mathcal{E}_1^1\) that is defined by a strongly positive, symmetric, and bounded bilinear form and a strongly continuous \(\lambda\)-convex functional \(\mathcal{E}_2^2\). The perturbation \(B\) is a strongly continuous perturbation of \(\partial \Psi_u\) and \(\partial \mathcal{E}_t\). An illustrative example in this framework is, in the smooth setting, given by

\[
\partial_t u - \nabla \cdot \left( g(u) \left| \nabla \partial_t u \right|^{q-2} \nabla \partial_t u \right) - \Delta u + W'(u) + b(u, \partial_t u) = f,
\]

where \(q > 1\), \(W : \mathbb{R} \to \mathbb{R}\) is a \(\lambda\)-convex and continuously differentiable function with Lipschitz continuous derivative, \(b : \mathbb{R} \to \mathbb{R}\) is a lower order perturbation, and \(f : \mathbb{R} \to \mathbb{R}\) an external force. The energy functional and the dissipation potential are given by

\[
\mathcal{E}(u) = \int_\Omega \left( \frac{1}{2} \left| \nabla u(x) \right|^2 + W(u(x)) \right) \, dx \quad \text{and} \quad \Psi_u(v) = \frac{1}{q} \int_\Omega g(u(x)) \left| \nabla v(x) \right|^q \, dx,
\]

\footnote{Ludwig-Maximilians-Universität München, Mathematisches Institut, Theresienstr. 39, 80333 München, Germany.}
and the perturbation is given by
\[ \langle B(u,v), w \rangle_{L^2} = \int_{\Omega} b(u(x), v(x)) w(x) \, dx. \]

In the section 3, we consider examples where we actually employ the full generality of our theory developed in this paper.

Evolution equations of second order where the operator acting on the time derivative of the solution is nonlinear has been studied by very few authors. Lions and Strauss [LiS65] showed in their seminal work the well-posedness of the Cauchy problem for the doubly nonlinear evolution equation
\[ u''(t) + A(t, u(t), u'(t)) + Bu(t) = f(t), \quad t \in (0, T), \]
where \( B \) is an unbounded, self-adjoint, and linear operator and \( A \) is a fully nonlinear operator which satisfies a monotonicity type condition. The peculiarity in this work is the assumption that the operators \( A \) and \( B \) are defined on different spaces, whose intersection is densely and continuously embedded in both spaces. This implies that the solution \( u \) takes values in a different space than its time derivative \( u' \). Emmrich and Thalhammer [EŠT15] showed, based on the techniques used by Lions and Strauss, the existence of solutions to the Cauchy problem for
\[ u''(t) + A(t)u'(t) + B(t)u(t) = f(t), \quad t \in (0, T), \]
where for each \( t \in [0, T] \), \( A(t) : V_A \to V_A^* \) is a hemicontinuous operator that satisfy a suitable growth condition such that \( A + \kappa I \) is monotone and coercive, and the operator \( B(t) = B_0 + C(t) : V_B \to V_B^* \) is the sum of a linear, bounded, symmetric, and strongly positive operator and a strongly continuous perturbation \( C(t) \). As in [LiS65], the authors assume neither that \( V_A \) is continuously embedded in \( V_B \) nor the reverse case. The assumptions on \( A \) imply that \( A + \kappa I \) is maximal monotone and therefore not necessarily a potential operator. Therefore, the result obtained here only partially generalizes the above mentioned results. However, to the best of the authors’ knowledge, results on the existence of strong solutions for multivalued operators \( A \) do not exist in the literature.

1.1 Preliminaries

For a proper functional \( F : X \to (-\infty, +\infty] \) on a BANACH space \((X, \| \cdot \|_X)\), we denote with the multivalued map \( \partial F : X \rightrightarrows X^* \), the (FRÉCHET) subdifferential of \( F \) defined by
\[ \partial F(u) := \left\{ \xi \in X^* : \lim\inf_{v \to u} \frac{F(v) - F(u) - \langle \xi, v - u \rangle_{X^* \times X}}{\|v - u\|_X} \geq 0 \right\}, \]
where \( \langle \cdot, \cdot \rangle \) denotes the duality pairing between the BANACH space \( X \) and its topological dual space \( X^* \). The elements of the subdifferential are also called subgradients. The effective domain of \( F \) and the domain of its subdifferential \( \partial F \), we denote by \( D(F) := \{ v \in X \mid F(v) < +\infty \} \) and \( D(\partial F) := \{ v \in X : \partial F(v) \neq \emptyset \} \), 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 GÂTEAUX differentiable function equals the sum of the subdifferentials of both functions. More precisely, there holds
Lemma 1.1. Let $F_1 : X \to (-\infty, +\infty]$ and $F_2 : X \to (-\infty, +\infty]$ be subdifferentiable and FRÉCHET differentiable at $u \in D(\partial G) \cap D(\partial G) \neq \emptyset$, respectively. Then, there holds

$$\partial(F_1 + F_2)(u) = \partial F_1(u) + DF_2(u),$$

where $DF_2$ denotes the FRÉCHET derivative of $F_2$.

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

Unlike the differential operator, the subdifferential operator is, in general, not linear. However, the following lemma shows under which assumptions this is the case.

Lemma 1.2. Let $F_1 : X \to (-\infty, +\infty]$ and $F_2 : X \to (-\infty, +\infty]$ be proper, lower semicontinuous and convex, and if there is a point $\tilde{u} \in \text{dom}(F_1) \cap \text{dom}(F_2)$ where $F_2$ is continuous, we have

$$\partial(F_1 + F_2)(v) = \partial F_1(v) + \partial F_2(v) \quad \text{for all } v \in X.$$  \hfill (1.2)

If $F_2$ is, in addition GÂTEAUX differentiable on $V$, there holds $\partial F_2(v) = \{D_G F_2(v)\}$ and we have

$$\partial(F_1 + F_2)(v) = \partial F_1(v) + D_G F_2(v) \quad \text{for all } v \in X,$$

where $D_G F_2$ denotes the GÂTEAUX derivative of $F_2$.

Proof. Proposition 5.3. on p. 23 and Proposition 5.6 on p. 26 in [EkT99]. \hfill \square

Furthermore, we recall an important tool from the theory of convex analysis. For a proper, lower semicontinuous and convex function $F : X \to (-\infty, +\infty]$, we define the so-called LEGENDRE–FENCHEL transform (or convex conjugate) $F^* : X^* \to (-\infty, +\infty]$ by

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

By definition, we directly obtain the Fenchel–Young inequality

$$\langle \xi, u \rangle \leq F(u) + F^*(\xi), \quad v \in X, \xi \in X^*.$$

It is easily checked that the transform itself is proper, lower semicontinuous and convex, see, e.g., EKELAND and TÉMAM [EkT99]. If, in addition, we assume $F(0) = 0$, then $F^*(0) = 0$ holds as well. We may ask how the convex conjugate of the sum of two functions can be explicitly expressed in terms of the two functions.

Lemma 1.3. Let $F_1 : X \to (-\infty, +\infty]$ and $F_2 : X \to (-\infty, +\infty]$ be a proper, lower semicontinuous and convex functional such that

$$\bigcap_{\lambda \geq 0} \lambda (\text{dom}(F_1) - \text{dom}(F_2))$$

is a closed vector space.

Moreover, let $F_1^* : X^* \to (-\infty, +\infty]$ and $F_2^* : X^* \to (-\infty, +\infty]$ be the associated convex conjugate of $F_1$ and $F_2$, respectively. Then, there holds

$$(F_1 + F_2)^*(\xi) = \min_{\eta \in X^*} \left( F_1^*(\xi - \eta) + F_2^*(\eta) \right) \quad \text{for all } \xi \in X^*.$$  \hfill (1.3)
Proof. Theorem 1.1, pp. 126, in BREZIS u. ATTOUCH [AtB86, Theorem 1.1, pp. 126]. □

For an illustration of the lemma, we consider

**Example 1.4.** Let \((X, \| \cdot \|_X)\) and \((Y, \| \cdot \|_Y)\) a BANACH spaces such that both \(X\) and \(Y\) are continuously embedded into another BANACH space \(Z\), and such that \(X \cap Y\), equipped with the norm \(\| \cdot \|_{X \cap Y} = \| \cdot \|_X + \| \cdot \|_Y\), is dense in both \(X\) and \(Y\). Then, the space \(X \cap Y\) becomes a BANACH space itself and the dual space can be identified as \(X^* + Y^*\) with the dual norm \(\| \xi \|_{X^* + Y^*} = \inf_{\xi_1 \in X^*, \xi_2 \in Y^*} \max \{ \| \xi_1 \|_{X^*}, \| \xi_2 \|_{Y^*} \}\). Now, let for \(p, q \in (1, +\infty)\) the functions \(F_1, F_2 : X \cap Y \rightarrow \mathbb{R}\) be given by

\[
F_1(u) = \frac{1}{p} \| u \|_{X^*}^p, \quad F_2(u) = \frac{1}{q} \| u \|_{Y^*}^q, \quad u \in X \cap Y.
\]

Then, according to Lemma 1.3, the convex conjugate \((F_1 + F_2)^* : X^* + Y^* \rightarrow (-\infty, +\infty]\) given by

\[
(F_1 + F_2)^*(\xi) = \min_{\xi_1 \in X^*, \xi_2 \in Y^*} \left( \frac{1}{p} \| \xi_1 \|_{X^*}^p + \frac{1}{q} \| \xi_2 \|_{Y^*}^q \right) \quad \text{for all } \xi \in X^* + Y^*, \tag{1.4}
\]

where \(p' > 1\) and \(q' > 1\) denotes the conjugate exponent of \(p\) and \(r\), respectively, i.e., fulfilling \(1/p + 1/p' = 1\) and \(1/q + 1/q' = 1\).

Further, we recall that if \(X\) and \(Y\) are separable and reflexive Banach spaces, then the space \(X \cap Y\) equipped with the norm \(\| \cdot \|_{X \cap Y}\) is a separable and reflexive BANACH space. Furthermore, the duality pairing between \(X \cap Y\) and \(X^* + Y^*\) is given by

\[
\langle f, v \rangle_{X^* \times Y^*} = \langle f_1, u \rangle_{X^* \times X} + \langle f_2, u \rangle_{Y^* \times Y}, \quad u \in X \cap Y,
\]

for all \(v \in X \cap Y\) and any partition \(f = f_1 + f_2\) with \(f_1 \in X\) and \(f_2 \in Y,\) for any \(p \in [1, +\infty],\) there holds \(L^p(0, T; X) \cap L^p(0, T; Y) = L^p(0, T; X \cap Y),\) where the measurability immediately follows from the PETTIS theorem, see, e.g., DIESTEL & UHL [DiU77, Theorem 2, p. 42]. If the continuous embedding \(X \hookrightarrow Y\) holds, then

\[
\langle f, v \rangle_{X^* \times X} = \langle f, v \rangle_{Y^* \times Y} \quad \text{if } v \in X \text{ and } f \in Y^*.
\]

see, e.g., BRÉZIS [BrÄ11, Remark 3, pp. 136] and GAJEWSKI et al. [GGZ74, Kapitel 1, §5].

The next lemma establishes a deep connection between the subgradient of a function and its convex conjugate.

**Lemma 1.5.** Let \(V\) be a BANACH space and let \(F : V \rightarrow (-\infty, +\infty]\) be a proper, lower semicontinuous and convex functional and let \(F^* : V^* \rightarrow (-\infty, +\infty]\) 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 \(\mathbb{R}\).

\[1\]See, e.g., Chapter I Section 5 in [GGZ74] for more details.
Proof. Proposition 5.1 and Corollary 5.2 on pp. 21 in [EkT99].

In the next result, we show

**Lemma 1.6.** Let the functionals \( f, f_n : [0, T] \times X \to (-\infty, +\infty) \) be given and fulfill the assumptions of Lemma 1.5, and let \( p \in (1, +\infty) \). Furthermore, let \((v_n)_{n \in \mathbb{N}} \subset L^p(0, T; X)\) and \((\xi_n)_{n \in \mathbb{N}} \subset L^{p'}(0, T; X^*)\) with \( v_n \to v \) in \( L^p(0, T; X) \) and \( \xi_n \to \xi \) in \( L^{p'}(0, T; X^*) \) as \( n \to \infty \) where \( F_n \) is the integral functional associated to \( f_n \). If

\[
\int_0^T (f(t, v(t)) + f^*(t, \xi(t))) \, dt \leq \liminf_{n \to \infty} \int_0^T (f_n(t, v_n(t)) + f^*_n(t, \xi_n(t))) \, dt
\]

and there holds

\[
\limsup_{n \to \infty} \int_0^T \langle \xi_n(t) - \xi(t), v_n(t) - v(t) \rangle \, dt \leq 0,
\]

then \( \xi(t) \in \partial f(t, v(t)) \) a.e. in \((0, T)\) and

\[
\int_0^T (f(t, v(t)) + f^*(t, \xi(t))) \, dt = \lim_{n \to \infty} \int_0^T (f_n(t, v_n(t)) + f^*_n(t, \xi_n(t))) \, dt
\]

**Proof.** Lemma 2.4.4 in BACHO [Bac21].

## 2 Topological assumptions and main result

We assume that \((U, \| \cdot \|_U), (V, \| \cdot \|_V), (W, \| \cdot \|_W)\) and \((\bar{W}, \| \cdot \|_{\bar{W}})\) are real, reflexive, and separable BANACH spaces such that \( U \cap V \) is separable and reflexive and that \((H, |\cdot|, \langle \cdot, \cdot \rangle)\) is a HILBERT space with norm \(|\cdot|\) induced by the inner product \(\langle \cdot, \cdot \rangle\).

Similarly, we assume again the following dense, continuous and compact embeddings

\[
\begin{align*}
U \cap V &\overset{d}{\hookrightarrow} U \overset{c}{\hookrightarrow} \bar{W} \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} \bar{W}^* \overset{d}{\hookrightarrow} U^* \overset{d}{\hookrightarrow} V^* + U^*, \\
U \cap V &\overset{d}{\hookrightarrow} V \overset{c}{\hookrightarrow} W \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} W^* \overset{d}{\hookrightarrow} V^* + U^*,
\end{align*}
\]

and if the perturbation does not explicitly depend on \( u \) or \( u' \), then we do not assume \( U \overset{c}{\hookrightarrow} \bar{W} \) or \( V \overset{c}{\hookrightarrow} W \), respectively. We further assume \( V \hookrightarrow W \) if \( E^*_t \neq 0 \), see Condition (2. Ea). We note that we neither assume \( U \hookrightarrow V \) nor \( V \hookrightarrow U \). Since in this case the subdifferential of \( \Psi_u \) is nonlinear, we refer to the inclusion (1.1) in the given framework as **nonlinearly damped inertial system** \((U, V, W, \bar{W}, H, \mathcal{E}, \Psi, B, f)\).

We first collect all the assumptions for the energy functional \( \mathcal{E}_t \), the dissipation potential \( \Psi_u \), the perturbation \( B \) as well as the external force \( f \), and discuss them subsequently. We start with the assumptions for the dissipation potential \( \Psi \).

(2.\( \Psi_a \)) **Dissipation potential.** For every \( u \in U \), let \( \Psi_u : V \to [0, +\infty) \) be a lower semicontinuous and convex functional with \( \Psi(0) = 0 \) such that the mapping \((u, v) \mapsto \Psi_u(v)\) is \( \mathcal{B}(U) \otimes \mathcal{B}(V) \)-measurable.

(2.\( \Psi_b \)) **Superlinearity.** The functional \( \Psi \) satisfies the following growth condition, i.e., there exists a positive real number \( q > 1 \) such that for all \( R > 0 \) there exist positive constants \( c_R, C_R > 0 \) such that for all \( u \in U \) with \( \sup_{t \in [0, T]} |\mathcal{E}_t(u)| \leq R \), there holds

\[
c_R(\|v\|_V^q - 1) \leq \Psi_u(v) \leq C_R(\|v\|_V^q + 1)
\]

for all \( v \in V, t \in [0, T] \).
(2.1) \textbf{Lower semicontinuity of } \Psi_u + \Psi_u^*: \text{ For all sequences } v_n \to v \text{ in } L^q(0, T; V), 
\eta_n \rightharpoonup \eta \text{ in } L^q(0, T; V^*), \text{ and } u_n(t) \to u(t) \text{ in } U \text{ for all } t \in [0, T] \text{ as } n \to \infty \text{ such that supp}_t \in [0, T], n \in \mathbb{N} \text{ and } \eta_n(t) \in \partial \Psi_{u_n(t)}(v_n(t)) \text{ a.e. in } t \in (0, T) \text{ for all } n \in \mathbb{N}, \text{ there holds}
\int_0^T (\Psi_{u(t)}(v(t)) + \Psi_{u(t)}^*(\eta(t))) \, dt \leq \liminf_{n \to \infty} \int_0^T (\Psi_{u_n(t)}(v_n(t)) + \Psi_{u_n(t)}^*(\eta_n(t))) \, dt.

For the solvability of problem (1.1), only the previous assumptions are required. If we additionally assume the uniform monotonicity of \partial \Psi_u, we obtain stronger convergence of the discrete time-derivatives \nabla r_n in the space L^q(0, T; V), see Lemma 2.7.

(2.2) \textbf{Uniform monotonicity of } \partial \Psi. \text{ For all } R > 0, \text{ there exists a constant } \mu_R > 0 \text{ such that}
\langle \xi - \eta, v - w \rangle_{V^* \times V} \geq \mu_R \|v - w\|_{V^{\max\{1,q\}}}

for all \( \xi \in \partial \Psi_u(v), \eta \in \partial \Psi_u(w) \) and \( u, v, w \in \{ \tilde{v} \in V : \mathcal{E}_t(\tilde{v}) \leq R \} \).

\textbf{Remark 2.1.}

\begin{enumerate}
\item[i)] We recall that the conjugate \( \Psi_u^* : V^* \to \mathbb{R} \) is lower semicontinuous and convex itself and that the growth condition (2.1) implies the following growth condition for the conjugate \( \Psi^* \): for all \( R > 0 \), there exist positive numbers \( \bar{c}_R, \bar{C}_R > 0 \) such that for all \( u \in U \) with \( \sup_{t \in [0, T]} \mathcal{E}_t(u) \leq R \), there holds
\( \bar{c}_R(\|\xi\|_{V^*}^q - 1) \leq \Psi_u^*(\xi) \leq \bar{C}_R(\|\xi\|_{V^*}^q + 1) \) \text{ for all } \( \xi \in V^* \),
where \( q^* = q/(q - 1) \).

\item[ii)] We can allow more general time-dependent dissipation potentials \( \Psi_u : [0, T] \times V \to [0, +\infty) \) by assuming all the previous assumptions uniformly in \( t \in [0, T] \).
\end{enumerate}

Now, we proceed with the assumptions for the energy functional.

(2.3) \textbf{Basic properties}. \text{ For all } t \in [0, T], \text{ the functional } \mathcal{E}_t : U \to \mathbb{R} \text{ is the sum of functionals } \mathcal{E}^1 : U \to \mathbb{R} \text{ and } \mathcal{E}^2_t : \tilde{W} \to \mathbb{R}. \text{ The functional } \mathcal{E}^1(\cdot) = \frac{1}{2} \overline{b}(\cdot, \cdot) \text{ is induced by a bounded, symmetric, and strongly positive bilinear form } b : U \times U \to \mathbb{R}, \text{ i.e., there exist constants } \mu, \alpha > 0 \text{ such that}
\begin{align*}
b(u, v) & \leq \alpha \|u\|_U \|v\|_U \quad \text{for all } u, v \in U, \\
\mu \|u\|_U^2 & \leq b(u, u) \quad \text{for all } u \in U.
\end{align*}

(2.4) \textbf{Bounded from below}. \mathcal{E}_t \text{ is bounded from below uniformly in time, i.e., there exists a constant } C_0 \in \mathbb{R} \text{ such that}
\mathcal{E}_t(u) \geq C_0 \quad \text{for all } u \in U \text{ and } t \in [0, T].

Since a potential is uniquely determined up to a constant, we assume without loss of generality \( C_0 = 0 \).

(2.5) \textbf{Coercivity}. \text{ For every } t \in [0, T], \mathcal{E}_t \text{ has bounded sublevel sets in } U.
(2. Ed) Control of the time derivative. For all \( u \in U \), the mapping \( t \mapsto \mathcal{E}_t^2(u) \) is in \( C([0, T]) \cap C^1(0, T) \) and its derivative \( \partial_t \mathcal{E}_t^2 \) is controlled by the function \( \mathcal{E}_t^2 \), i.e., there exists \( C_1 > 0 \) such that

\[
|\partial_t \mathcal{E}_t^2(u)| \leq C_1 \mathcal{E}_t^2(u) \quad \text{for all } t \in (0, T) \text{ and } u \in V.
\]

Furthermore, for all sequences \( (u_n)_{n \in \mathbb{N}}, u \subset D \) with \( u_n \to u \) as \( n \to \infty \) and \( \sup_{n \in \mathbb{N}, t \in [0, T]} \mathcal{E}_t(u_n) < +\infty \), there holds

\[
\limsup_{n \to \infty} \partial_t \mathcal{E}_t^2(u_n) \leq \partial_t \mathcal{E}_t^2(u) \quad \text{for a.e. } t \in (0, T).
\]

(2. Ee) Fréchet differentiability. For all \( t \in [0, T] \), the mapping \( u \mapsto \mathcal{E}_t^2(u) \) is Fréchet differentiable on \( \widetilde{W} \) with derivative \( D\mathcal{E}_t^2 \) such that the mapping \( (t, u) \mapsto D\mathcal{E}_t^2(u) \) is continuous as a mapping from \( [0, T] \times \widetilde{W} \) to \( U^* \) on sublevel sets of the energy, i.e., for all \( R > 0 \) and sequences \( (u_n)_{n \in \mathbb{N}}, u \subset \widetilde{W} \) and \( (t_n)_{n \in \mathbb{N}}, t \subset [0, T] \) with \( \sup_{t \in [0, T], n \in \mathbb{N}} \mathcal{E}_t(u_n) < +\infty \), \( u_n \to u \) in \( \widetilde{W} \), and \( t_n \to t \) as \( n \to \infty \), there holds

\[
\lim_{n \to \infty} \|D\mathcal{E}_{t_n}^2(u_n) - D\mathcal{E}_t^2(u)\|_{U^*} = 0.
\]

(2. Ef) \( \lambda \)-convexity. There exists a non-negative real number \( \lambda \geq 0 \) such that

\[
\mathcal{E}_t^2(\vartheta u + (1 - \vartheta)v) \leq \vartheta \mathcal{E}_t^2(u) + (1 - \vartheta)\mathcal{E}_t^2(v) + \vartheta(1 - \vartheta)\lambda |u - v|^2
\]

for all \( t \in [0, T], \vartheta \in [0, 1] \) and \( u, v \in U \).

(2. Eg) Control of \( D\mathcal{E}_t^2 \). There exist positive constants \( C_2 > 0 \) and \( \sigma > 0 \) such that

\[
\|D\mathcal{E}_t^2(u)\|_{W^*}^{\sigma} \leq C_3(1 + \mathcal{E}_t^2(u) + \|u\|_{W^*}) \quad \text{for all } t \in [0, T], u \in \widetilde{W}.
\]

Again, several remarks are in order.

Remark 2.2.

i) The assumptions on the quadratic form \( \mathcal{E}^1 \) imply that the Fréchet derivative \( D\mathcal{E}^1 \) is given by a linear, bounded, symmetric and strongly positive operator \( E \in \mathcal{L}(V, V^*) \) such that \( \mathcal{E}^1(u) = \frac{1}{2}\langle Eu, u \rangle \) is strongly convex and therefore sequentially weakly lower semicontinuous. Furthermore, the corresponding Nemitskiĭ operator is a linear and bounded map from \( L^2(0, T; V) \) to \( L^2(0, T; V^*) \) and hence weak-to-weak continuous from \( L^2(0, T; V) \) to \( L^2(0, T; V^*) \).

ii) From Assumption (2. Ed), it follows after integration

\[
\sup_{t \in [0, T]} \mathcal{E}_t^2(u) \leq e^{C_i T} \inf_{t \in [0, T]} \mathcal{E}_t^2(u),
\]

\[
|\mathcal{E}_t^2(u) - \mathcal{E}_s^2(u)| \leq e^{C_i T} \sup_{r \in [0, T]} \mathcal{E}_r^2(u)|s - t| \quad \text{for all } u \in U, s, t \in [0, T].
\]

iii) The derivative of the \( \lambda \)-convex energy functional is characterized by the inequality

\[
\mathcal{E}_t^2(u) - \mathcal{E}_t^2(v) \leq \langle D\mathcal{E}_t^2(u), u - v \rangle_{U^*, U} + \lambda |u - v|^2
\]

for all \( t \in [0, T], u, v \in U \). In fact, the \( \lambda \)-convexity can be replaced by the latter inequality, since we only make use of (2.2) in order to obtain a priori estimates, see Lemma 2.6.
We recall that the Fréchet differentiability of $\mathcal{E}_t$ implies the subdifferentiability of $\mathcal{E}_t$ and the subdifferential is a singleton with $\partial \mathcal{E}(u) = \{D\mathcal{E}(u)\}$.

Finally, we collect the assumptions concerning the perturbation $B$ and the external force $f$.

(2.Ba) **Continuity.** The mapping $B: [0, T] \times \bar{W} \times W \to V^*$ is continuous on sublevel sets of $\mathcal{E}_t$, i.e., for every converging sequence $(t_n, u_n, v_n) \to (t, u, v)$ in $[0, T] \times \bar{W} \times W$ with $\sup_{n \in \mathbb{N}} G(u_n) < +\infty$, there holds $B(t_n, u_n, v_n) \to B(t, u, v)$ in $V^*$ as $n \to \infty$.

(2.Bb) **Control of the growth.** There exist positive constants $\beta > 0$ and $c, \nu \in (0, 1)$ such that

$$c \Psi_u^* \left(\frac{-B(t, u, v)}{c}\right) \leq \beta(1 + \mathcal{E}_t(u) + |v|^2 + \Psi(v)^\nu)$$

for all $u \in U, v \in V, t \in [0, T]$.

(8.f) **External force.** There holds $f \in L^2(0, T; H)$.

**Remark 2.3.** If the growth condition (2.$\Psi_b$) for $\Psi_u$ holds uniformly in $u \in U$, then more general external forces $f \in L^1(0, T; H) + L^q(0, T; V^*)$ can be considered.

### 2.1 Discussion of the assumptions

Again, we want to discuss the assumptions more in detail.

As the name suggests, we consider in this case evolution equations of second order with nonlinear damping, i.e., where $\partial \Psi_u(t)$ is nonlinear and in general multi-valued. This restricts us to the case where the principle part of the operator $\partial \mathcal{E}$ is linear. The principle parts of $\partial \Psi_u(t)$ and $\partial \mathcal{E}$ are defined on spaces for which we assume not that either of the two spaces is continuously embedded in the other one. As mentioned in the introduction, this has not been studied before. However, for single valued operators, a similar case has been investigated by Lions & Strauss [LiS65] and Emmrich & Thalhammer [EmT11].

Ad (2.$\Psi$). The conditions for the dissipation potentials are similar to those in [BEM19] for perturbed gradient systems. In contrast to the superlinearity condition in [BEM19], we assume here that $\Psi_u$ has $p$-growth on sublevel sets of $\mathcal{E}_t$, which allows us to employ an integration by parts formula for the second derivative $u''$ proven in Emmrich & Thalhammer [EmT11], see Lemma 2.7 below.

We remark that the liminf estimate in Condition (2.$\Psi_c$) is already implied by the Mosco-convergence $\Psi_{u_n} \overset{M}{\rightharpoonup} \Psi_u$ for all sequences $u_n \rightharpoonup u$, see [Ste08]. The prototypical examples for state-independent dissipation potential which fulfill Condition (2.$\Psi_a$)-(2.$\Psi_d$) are

$$\Psi(v) = \int_{\Omega} \left(\frac{1}{p} |\nabla v(x)|^p + |\nabla v(x)| \right) \, dx \quad \text{or} \quad \Psi(v) = \int_{\Omega} \left(\frac{1}{p} |v(x)|^p + |v(x)| \right) \, dx$$
on $V = W_0^{1,p}(\Omega)$ or $V = L^p(\Omega)$ with $p \in (1, +\infty)$, respectively. For state-independent dissipation potentials more general integral functionals of the form

$$\Psi_u(v) = \int_\Omega \psi(x, u(x), v(x), \nabla u(x), \nabla v(x)) \, dx$$

can be considered on appropriate Sobolev spaces, where $\psi$ is a proper, lower semicontinuous and convex function satisfying certain growth and continuity conditions.

Ad (2.E). The crucial assumption we make for the energy functional $E_t = E^1 + E^2_t$ is that the leading part $E^1$ is defined by a bounded, symmetric, and strongly positive bilinear form $b : U \times U \to \mathbb{R}$. All other assumptions concern the strongly continuous perturbation $E^2_t$ which are very similar to those made for the energy functional for linearly damped inertial systems. The main difference is that we assume a Fréchet differentiability of $E^2_t$. Having discussed all assumptions, we are in a position to state the main result which again includes the notion of solution to (1.1).

**Theorem 2.4 (Existence result).** Let the nonlinearly damped inertial system 

$$(U, V, W, \dot{W}, H, E, \Psi, B, f)$$

be given and fulfill Assumptions (2.E), (2.A)-(2.C) as well as (2.B) and Assumption (2.F). Then, for every $u_0 \in U$ and $v_0 \in H$, there exists a solution to (1.1), i.e., there exist functions $u \in C([0, T]; U) \cap W^{1,\infty}(0, T; H) \cap W^{2,p}(0, T; U^* + V^*)$ with $u-u_0 \in W^{1,p}(0, T; V)$ and $\eta \in L^p(0, T; V^*)$ satisfying the initial conditions $u(0) = u_0$ in $U$ and $u'(0) = v_0$ in $H$ such that

$$\eta(t) \in \partial \Psi_{u(t)}(u'(t)) \quad \text{and} \quad u''(t) + \eta(t) + D\mathcal{E}_t(u) + B(t, u(t), u'(t)) = f(t) \quad \text{in } U^* + V^* \quad (2.3)$$

for almost every $t \in (0, T)$. Furthermore, the energy-dissipation balance

$$\frac{1}{2} |u'(t)|^2 + \mathcal{E}_t(u(t)) + \int_0^t \left( \Psi_{u(t)}(u'(r)) + \Psi_{u(t)}(S(r) - D\mathcal{E}_r(r) - u''(r)) \right) \, dr$$

$$= \frac{1}{2} |v_0|^2 + \mathcal{E}_0(u_0) + \int_0^t \partial \mathcal{E}_r(u(r)) \, dr + \int_0^t \langle S(r), u'(r) \rangle_{V^* \times V} \, dr \quad (2.4)$$

holds for almost every $t \in (0, T)$, where $S(r) := f(r) - B(r, u(r), u'(r))$, $r \in [0, T]$, and if $V \hookrightarrow U$, then (2.4) holds for all $t \in [0, T]$.

### 2.2 Variational approximation scheme

The proof of Theorem 2.4 again relies on a semi-implicit time discretization scheme. Therefore, we will proceed in a similar way to the case in the previous section. The main difference and difficulty arises in identifying the (a priori) weak limits associated with the nonlinear terms $D\mathcal{E}$ and $\partial \Psi$. Again, for $N \in \mathbb{N}\setminus\{0\}$, let

$$I_\tau = \{0 = t_0 < t_1 < \cdots < t_n = n\tau < \cdots < t_N = T\}$$

be an equidistant partition of the time interval $[0, T]$ with step size $\tau := T/N$, where we again omit the dependence of the nodes from the partition on the step size. Discretizing inclusion (1.1) in a semi-implicit manner yields

$$\frac{V^n_{\tau} - V^{n-1}_{\tau}}{\tau} + \partial \Psi_{U^{n-1}_{\tau}}(V^n_{\tau}) + D\mathcal{E}_{t^n_{\tau}}(U^n_{\tau}) + B \left(t_n, U^{n-1}_{\tau}, V^{n-1}_{\tau} \right) \ni f^n_{\tau} \quad \text{in } U^* + V^* \quad (2.5)$$
for \( n = 1, \ldots , N \) with \( V^n_\tau = \frac{U^n_\tau - U^{n-1}_\tau}{\tau} \). The value \( U^n_\tau \) is to be determined recursively from the variational approximation scheme

\[
\begin{aligned}
U^0_\tau \in U \cap V \text{ and } V^0_\tau \in V \text{ are given; whenever } U^1_\tau, \ldots , U^{n-1}_\tau \in D \cap V \text{ are known, find } U^n_\tau \in J_{\tau,t_{n-1}} \left( U^{n-1}_\tau, U^{n-2}_\tau, B(t_n, U^{n-1}_\tau, V^{n-1}_\tau) - f^n_\tau \right)
\end{aligned}
\]

(2.6)

for \( n = 1, \ldots , N \), where \( J_{\tau,t}(v,w;\eta) := \arg\min_{u \in \mathcal{U} \cap \mathcal{V}} \Phi (r,t,v,w;\eta;u) \) and \( U^{-1}_\tau = U^0_\tau - V^0_\tau \).

Thus, Lemma 2.5 ensures that minimizer of the mapping

\[
\Phi (r,t,v,w;\eta;u) = \frac{1}{2r^2}u^2 - 2\nu + w^2 + r\Psi_\nu \left( \frac{u-v}{r} \right) + \mathcal{E}_t(u) - \langle \zeta, u \rangle_{V^* \times V}
\]

for \( r \in \mathbb{R}^{>0}, t \in [0,T] \) with \( r + t \leq T \) as well as \( v \in V, w \in H \) and \( \zeta \in V^* \). The solvability of the discrete problem (2.6) and that every solution fulfills the EULER–LAGRANGE equation (2.5) is ensured by the following lemma.

**Lemma 2.5.** Let the nonlinearly damped inertial system \((U,V,W,H,\mathcal{E},\Psi)\) be given and fulfill the Conditions (2. Ea)-(2. Ec), (2. Ec), (2. Ef) and (2. Ea)-(2. Pb). Furthermore, let \( r \in (0,T) \) and \( t \in [0,T] \) with \( r + t \leq T \) as well as \( v \in V, w \in H \) and \( \zeta \in V^* \). Then, the set \( J_{r,t}(v,w;\eta) \) is non-empty and single valued if \( r \leq \frac{1}{2\lambda} \), where \( \lambda \) is from (2. Ef).

Furthermore, to every \( u \in J_{r,t}(v,w;\eta) \), there exists \( \eta \in \partial V\Psi_\nu \left( \frac{u-r}{\lambda} \right) \subset V^* \) such that

\[
\frac{u - 2\nu - w}{r^2} + \eta + \mathcal{D}\mathcal{E}_t(u) + \zeta = 0 \quad \text{in } U^* + V^*.
\]

**Proof.** The proof follows immediately from the direct methods of the calculus of variations as well as Lemma 1.2. \( \square \)

Thus, Lemma 2.5 ensures that minimizer of the mapping

\[
u \mapsto \Phi (r,t_{n-1},U^{n-1}_\tau, U^{n-2}_\tau, B(t_n, U^{n-1}_\tau, V^{n-1}_\tau) - f^n_\tau;u),
\]

fulfill the EULER–LAGRANGE equation (2.5).

**2.3 Discrete Energy-Dissipation inequality and a priori estimates**

In this section, we derive a priori estimates to the approximate solutions. Thus, let the initial values \( u_0 \in U \cap V \) and \( v_0 \in V \) as well as the time step \( \tau > 0 \) be given and fixed. As before, we will assume more general initial values in the main existence result and approximate by suitable sequences of values. Then, for given approximate values \( (U^n_\tau)_{n=1}^N \) with \( U^0_\tau := u_0 \) and \( V^0_\tau = v_0 \) obtained from the variational approximation scheme (2.6), we define the piecewise constant and linear interpolations \( \overline{U}_\tau, \underline{U}_\tau, \overline{V}_\tau, \underline{V}_\tau, \overline{V}_\tau, \xi_\tau, f_\tau \) as well as \( \overline{U}_\tau \) and \( \underline{U}_\tau \) as in [Bac20]. Furthermore, by Lemma 2.5, there exists a sequence \( (\eta^n_\tau)_{n=1}^N \subset V^* \) of subgradients fulfilling \( \eta^n_\tau \in \partial V\Psi_\nu (U^n_\tau) \), \( n = 1, \ldots , N \), such that

\[
\frac{V^n_\tau - V^{n-1}_\tau}{\tau} + \eta^n_\tau + \mathcal{D}\mathcal{E}_t(U^n_\tau) + B(t_n, U^{n-1}_\tau, V^{n-1}_\tau) = f^n_\tau \quad \text{in } U^* + V^*, \quad n = 1, \ldots , N.
\]

Then, we define the measurable function \( \eta_\tau : [0,T] \rightarrow V^* \) by

\[
\eta_\tau(t) = \eta^n_\tau \quad \text{for } t \in (t_{n-1}, t_n], \quad n = 1, \ldots , N, \quad \text{and } \quad \eta_\tau(T) = \eta^N_\tau.
\]

Having defined the interpolations, we are in the position to show the a priori estimates in the following lemma.
2.3 DISCRETE ENERGY-DISSIPATION INEQUALITY AND A PRIORI ESTIMATES

Lemma 2.6 (A priori estimates). Let the system \((U, V, W, H, \mathcal{E}, \Psi, B, f)\) be given and satisfy the Assumptions (2.E), (2.Ψ), (2.B) as well as Assumption (2.f). Furthermore, let \(U_\tau, U_{\tau-}, \tilde{U}_\tau, V_\tau, \tilde{V}_\tau, \eta_\tau \) and \(f_\tau\) be the interpolations associated with the given values \(u_0 \in U \cap V, v_0 \in V\) and the step size \(\tau > 0\). Then, the discrete energy-dissipation inequality

\[
\int_{\tilde{T}_\tau(s)}^{\tilde{T}_\tau(\tau)} \left( \psi_{U_\tau}\left(\nabla_\tau(r)\right) + \psi_{U_{\tau-}}^\ast \left( S_\tau(r) - \tilde{\nabla}_\tau(r) - D\mathcal{E}_{U_\tau}(\nabla_\tau(r)) \right) \right) dr \\
+ \frac{1}{2} \left| \nabla_\tau(t) \right|^2 + \mathcal{E}_{\tilde{U}_\tau}(\nabla_\tau(t)) \\
\leq \frac{1}{2} \left| \nabla_\tau(s) \right|^2 + \mathcal{E}_{\tilde{U}_\tau}(\nabla_\tau(s)) + \int_{\tilde{T}_\tau(s)}^{\tilde{T}_\tau(\tau)} \partial_t \mathcal{E}_\tau(U_\tau(r)) dr + \int_{\tilde{T}_\tau(s)}^{\tilde{T}_\tau(\tau)} \left( S_\tau(r) - \tilde{\nabla}_\tau(r) - D\mathcal{E}_{U_\tau}(\nabla_\tau(r)) \right) U_{\tau-} U \, dr \\
+ \tau \lambda \int_{\tilde{T}_\tau(s)}^{\tilde{T}_\tau(\tau)} \left| \nabla_\tau(r) \right|^2 dr
\]

(2.8)

holds for all \(0 \leq s < t \leq T\), where we have introduced the short-hand notation \(S_\tau(r) := f_\tau(r) - B(\tilde{U}_\tau(r), U_{\tau-}(r), \tilde{V}_\tau(r))\), \(r \in [0, T]\). Moreover, there exist positive constants \(M, \tau^* > 0\) such that the estimates

\[
\sup_{t \in [0, T]} \left| \nabla_\tau(t) \right| \leq M, \quad \sup_{t \in [0, T]} \mathcal{E}_\tau(U_\tau(t)) \leq M, \quad \sup_{t \in [0, T]} \left| \partial_t \mathcal{E}_\tau(U_\tau(t)) \right| \leq M, \quad (2.9)
\]

\[
\int_0^T \left( \psi_{U_\tau}\left(\nabla_\tau(r)\right) + \psi_{U_{\tau-}}^\ast \left( S_\tau(r) - \tilde{\nabla}_\tau(r) - D\mathcal{E}_{U_\tau}(\nabla_\tau(r)) \right) \right) dr \leq M
\]

(2.10)

hold for all \(0 < \tau \leq \tau^*\). In particular, the families of functions

\[
(U_\tau)_{0 < \tau \leq \tau^*} \subset L^\infty(0, T; U),
\]

\[
(V_\tau)_{0 < \tau \leq \tau^*} \subset L^q(0, T; V),
\]

\[
(\eta_\tau)_{0 < \tau \leq \tau^*} \subset L^q(0, T; V^*),
\]

\[
(\tilde{V}_\tau^*)_{0 < \tau \leq \tau^*} \subset L^q(0, T; U^* + V^*),
\]

\[
(B_\tau)_{0 < \tau \leq \tau^*} \subset L^2(0, T; V^*),
\]

\[
(D\mathcal{E}_\tau^2(U_\tau))_{0 < \tau \leq \tau^*} \subset L^\infty(0, T; \tilde{W}^*),
\]

(2.11)

are uniformly bounded with respect to \(\tau\) in the respective spaces, where \(q^* > 0\) is the conjugate exponent to \(q > 1\) and \(\nu \in (0, 1)\) being from Assumption (2.Bb). Finally, there holds

\[
\sup_{t \in [0, T]} \left( \| U_\tau(t) - U_\tau(t) \|_V + \| \tilde{U}_\tau(t) - U_\tau(t) \|_V \right) \to 0
\]

\[
\sup_{t \in [0, T]} \left( \| \nabla_\tau(t) - \tilde{\nabla}_\tau(t) \|_{U^* + V^*} + \| V_\tau(t) - \nabla_\tau(t) \|_{U^* + V^*} \right) \to 0
\]

(2.12)

as \(\tau \to 0\).

Proof. Let \((U^n)_{n=1}^{N} \subset U \cap V\) be the approximative values obtained from the variational approximation scheme (2.6) which satisfy by Lemma 1.2 the EULER–LAGRANGE equation

\[
f^n - B(t_n, U_{\tau-}^{n-1}, V_{\tau-}^{n-1}) - \frac{V^n - V^{n-1}_\tau}{\tau} - D\mathcal{E}_{\tau_n}(U^n_\tau) = \eta^n_\tau \in \partial_{V} \Psi_{U_{\tau-}^{n-1}}(V^n_\tau)
\]

(2.13)
for all \( n = 1, \ldots, N \). According to Lemma 1.5, inclusion (2.13) is equivalent to

\[
\Psi_{U^n_{\tau}}(V^n_{\tau}) + \Psi^*_{U^n_{\tau}} \left( f^n_{\tau} - B(t_n, U^{n-1}_{\tau}, V^{n-1}_{\tau}) - \frac{V^n_{\tau} - V^{n-1}_{\tau}}{\tau} - D\varepsilon_n(U^n_{\tau}) \right)
\]

\[
= \left( f^n_{\tau} - B(t_n, U^{n-1}_{\tau}, V^{n-1}_{\tau}) - \frac{V^n_{\tau} - V^{n-1}_{\tau}}{\tau} - D\varepsilon_n(U^n_{\tau}), V^n_{\tau} \right)_{V^* \times V}, \quad n = 1, \ldots, N.
\]

Furthermore, the enhanced FRÉCHET subdifferentiability (2. Ef) yields

\[
- \left( D\varepsilon_n(U^n_{\tau}), U^n_{\tau} - U^{n-1}_{\tau} \right)_{(U^* + V^*) \times (U \cup V)} \leq E_n(U^n_{\tau} - U^{n-1}_{\tau})^2
\]

\[
- \varepsilon_n(U^n_{\tau})
\]

\[
= E_{n-1}(U^{n-1}_{\tau}) - E_n(U^n_{\tau}) + \lambda |U^n_{\tau} - U^{n-1}_{\tau}|^2 + \int_{t_{n-1}}^{t_n} \partial_r \varepsilon_r(U^{n-1}_{\tau}) \, dr
\]

for all \( n = 1, \ldots, N \). Employing the identity

\[
(V^n_{\tau} - V^{n-1}_{\tau}, V^n_{\tau}) = \frac{1}{2} |V^n_{\tau}|^2 - \frac{1}{2} |V^{n-1}_{\tau}|^2 + \frac{1}{2} |V^n_{\tau} - V^{n-1}_{\tau}|^2,
\]

we obtain

\[
\frac{1}{2} |V^n_{\tau}|^2 + E_n(U^n_{\tau}) + \tau \Psi_{U^n_{\tau}}(V^n_{\tau}) + \tau \Psi^*_{U^n_{\tau}} \left( S^n_{\tau} - \frac{V^n_{\tau} - V^{n-1}_{\tau}}{\tau} - D\varepsilon_n(U^n_{\tau}) \right)
\]

\[
\leq \frac{1}{2} |V^{n-1}_{\tau}|^2 + E_{n-1}(U^{n-1}_{\tau}) + \int_{t_{n-1}}^{t_n} \partial_r \varepsilon_r(U^{n-1}_{\tau}) \, dr + \lambda \int_{t_{n-1}}^{t_n} |V^r_{\tau}|^2 \, dr + \tau \left( S^r_{\tau}, V^r_{\tau} \right)_{V^* \times V}
\]

for all \( n = 1, \ldots, N \), where \( S^n_{\tau} := f^n_{\tau} - B(t_n, U^{n-1}_{\tau}, V^{n-1}_{\tau}) \), \( n = 1, \ldots, N \). Summing up the inequalities over \( n \) yields (2.8). The estimates (2.9) and (2.10) are obtained by employing the discrete version of GRONWALL’S lemma (Lemma A.2) taking further into account that by Condition (2. Ed), there holds

\[
\partial_r \varepsilon^2_r(U^{n-1}_{\tau}) \leq C_1 E^2_r(U^{n-1}_{\tau})
\]

for all \( r \in (0, T) \) and \( n = 1, \ldots, N \). The estimates (2.9) and (2.10) in turn imply in view of Assumption (2. Ea),(2. Ec),(2. Eg) as well as Assumption (2.Ψb) and Remark 2.1 the uniform bounds (2.11a)-(2.11f) and the convergences (2.12).

### 2.4 Compactness

In this section, we prove the (weak) compactness of the approximate solutions in suitable BOCHNER spaces in order to pass to the limit in the weak formulation of the discrete inclusion (2.5) as the step size vanishes. After identifying all the weak limits, we will indeed obtain a solution to the CAUCHY problem (1.1). The compactness result is given in the following lemma whose proof follow along the same line as the corresponding Lemma in [Bac20] for linearly damped inertial systems. Therefore, we will prove the assertions which differ from the previously mentioned lemma.
Lemma 2.7 (Compactness). Under the same assumptions of Lemma 2.6, let \((\tau_n)_{n \in \mathbb{N}}\) be a vanishing sequence of step sizes and let \(u_0 \in U \cap V\) and \(v_0 \in V\). Then, there exists a subsequence, still denoted by \((\tau_n)_{n \in \mathbb{N}}\), a pair of functions \((u, \eta)\) with

\[
  u \in C_w([0, T]; U) \cap W^{1, q}(0, T; V) \cap W^{1, \infty}(0, T; H) \cap W^{2,q^*}(0, T; U^* + V^*) \\
  \eta \in L^{q^*}(0, T; V^*),
\]

and fulfilling the initial values \(u(0) = u_0\) in \(U\) and \(u'(0) = v_0\) in \(H\) such that the following convergences hold

\[
\begin{align*}
  \hat{U}_{\tau_n}, \hat{U}_{\tau_n}, \hat{U}_{\tau_n} & \rightharpoonup u & \text{in } L^\infty(0, T; U \cap V), \quad (2.14a) \\
  \hat{U}_{\tau_n}(t), \hat{U}_{\tau_n}(t), \hat{U}_{\tau_n}(t) & \to u(t) & \text{in } U \text{ for all } t \in [0, T], \quad (2.14b) \\
  \hat{U}_{\tau_n}(t) & \to u(t) & \text{in } V \text{ for all } t \in [0, T], \quad (2.14c) \\
  \hat{U}_{\tau_n} & \to u & \text{in } L^r(0, T; \bar{W}) \text{ for any } r \geq 1, \quad (2.14d) \\
  \hat{U}_{\tau_n}(t), \hat{U}_{\tau_n}(t), \hat{U}_{\tau_n}(t) & \to u(t) & \text{in } \bar{W} \text{ for all } t \in [0, T], \quad (2.14e) \\
  \nabla_{\tau_n}, \nabla_{\tau_n}, \nabla_{\tau_n} & \rightharpoonup u' & \text{in } L^q(0, T; V) \cap L^\infty(0, T; H), \quad (2.14f) \\
  \nabla_{\tau_n} & \to u' & \text{in } L^p(0, T; H) \text{ for any } p \geq 1, \quad (2.14g) \\
  \hat{V}_{\tau_n}(t), \hat{V}_{\tau_n}(t), \hat{V}_{\tau_n}(t) & \to u'(t) & \text{in } H \text{ for all } t \in [0, T], \quad (2.14h) \\
  \eta_{\tau_n} & \to \eta & \text{in } L^{q^*}(0, T; V^*), \quad (2.14i) \\
  E\hat{U}_{\tau_n} & \to Eu & \text{in } L^2(0, T; U^*), \quad (2.14j) \\
  D\mathcal{E}_t^2(\hat{U}_{\tau_n}) & \to D\mathcal{E}_t^2(u) & \text{in } L^r(0, T; \bar{W}^*) \text{ for any } r \geq 1, \quad (2.14k) \\
  \hat{V}_{\tau_n}'' & \to u'' & \text{in } L^{\max(2,q^*)}(0, T; U^* + V^*), \quad (2.14l) \\
  f_{\tau_n} & \to f & \text{in } L^2(0, T; H), \quad (2.14m) \\
  B_{\tau_n} & \to B(\cdot, u(\cdot), u'(\cdot)) & \text{in } L^r(0, T; V^*), \quad (2.14n)
\end{align*}
\]

Furthermore, if the dissipation potential satisfies in addition Assumption (2.Ψd), then, there holds

\[
\begin{align*}
  \nabla_{\tau_n} & \to u' & \text{in } L^{\max(2,q)}(0, T; U), \quad (2.15a) \\
  \hat{U}_{\tau_n} & \to u & \text{in } C([0, T]; U). \quad (2.15b)
\end{align*}
\]

Finally, the function \(u\) satisfies the inequality

\[
\begin{align*}
  \frac{1}{2}|v_0|^2 - \frac{1}{2}|u'(t)|^2 + \mathcal{E}_0(u_0) - \mathcal{E}_1(u(t)) +\int_0^t \partial_t \mathcal{E}_r(u(r)) \, dr \\
  \leq -\int_0^t \langle u''(r) + D \mathcal{E}_r(u(r)), u'(r) \rangle_{V^* \times V} \quad (2.16)
\end{align*}
\]

for almost every \(t \in (0, T)\).

Proof. We restrict the proof by only showing the convergence (2.14j),(2.14k), (2.15a), and (2.15b) and note that the remainder of the proof can be proved in the same manner as in [Bac20]. First, convergence (2.14j) follows from Remark 2.2 i) and the weak convergence (2.14a). Further, from the growth condition (2. Eg), we obtain

\[
\|D\mathcal{E}_t^2(\hat{U}_{\tau_n}(t))\|_{\bar{W}^*} \leq C_3(1 + \mathcal{E}^2(\hat{U}_{\tau_n}(t)) + \|\hat{U}_{\tau_n}(t)\|_{\bar{W}})
\]
and in view of the a priori estimates (2.9),
\[ \|D \mathcal{E}_r^2(t)(\mathcal{U}_r(t))\|_{W^*} \leq C \quad \text{for all } t \in [0, T]. \]

Together with the convergence (2.14e) and the continuity condition (2. Ee), this leads to (2.14k). The last assertions (2.15a) and (2.15b) follow immediately from Assumption (2.Ψd) and
\[
\limsup_{n \to \infty} \int_0^T \|V_{\tau_n}(r) - u'(r)\|_{V^*}^{\max(p, 2)} \, dr \\
\leq \limsup_{n \to \infty} \int_0^T (\eta_h(r) - \eta(r), V_{\tau_n}(r) - u'(r))_{V^* \times V} \, dr \leq 0
\]
and \( \eta(t) \in \partial V \Phi_u(t)(u'(t)) \) a.e. in \((0, T)\), which we will show in the proof of the main result. It remains to show the inequality (2.16). The difficulty in proving the aforementioned inequality is that we are not allowed to split the duality pairing in the integral on the right-hand side and consider each integral separately. However, since (2.16) is a slight modification of Lemma 6 in EMMRICH & THALHAMMER [EmT11], we follow their proof and regularize the function \( u' \) by its so-called STEKLOV average. For a function \( v \in L^p(0, T; X), p \geq 1 \), defined on a BANACH space \( X \) and being extended by zero outside \([0, T]\), the STEKLOV average is, for sufficiently small \( h > 0 \), given by
\[
S_h v(t) := \frac{1}{2h} \int_{t-h}^{t+h} v(r) \, dr.
\]

It is readily seen that \( S_h v \in L^p(0, T; X) \) and \( \|S_h v\|_{L^p(0, T; X)} \leq \|v\|_{L^p(0, T; X)} \). Furthermore, it can be shown by a regularization argument that \( S_h v \to v \) in \( L^p(0, T; X) \) as \( h \to 0 \), see, e.g., DIESTEL & UHL [DiU77, Theorem 9, p. 49].

Defining \( K v(t) = \int_0^t v(r) \, dr \), we commence with calculating
\[
\begin{align*}
- \int_s^t \langle (S_h u')'(r) + D \mathcal{E}_r(u_0 + (K S_h u')(r)), (S_h u')'(r) \rangle_{V^* \times V} \, dr \\
- \int_s^t \langle (S_h u')'(r) + E(u_0 + (K S_h u')(r)) + D \mathcal{E}_r^2(u_0 + (K S_h u')(r)), (S_h u')'(r) \rangle_{V^* \times V} \, dr \\
= \frac{1}{2} |(S_h u')(s)|^2 - \frac{1}{2} |(S_h u')(t)|^2 + D \mathcal{E}_r^1(u_0 + (K S_h u')(s)) - D \mathcal{E}_r^1(u_0 + (K S_h u')(t)) \\
+ \mathcal{E}_r^2(u_0 + (K S_h u')(s)) - \mathcal{E}_r^2(u_0 + (K S_h u')(t))
\end{align*}
\]
for all \( s, t \in [0, T] \) where we have applied the integration by parts formula, since the duality pairing can be split now due to the fact that \((S_h u')(t) = \frac{1}{2h}(\dot{u}(t+h) - \dot{u}(t-h))\), where \( \dot{u} \) is a continuous extension of \( u \) outside \([0, T]\) which makes sense, since \( u \in L^\infty(0, T; U) \cap W^{1,1}(0, T; H) \subset C_w([0, T]; U) \) and therefore \( S_h u' \in L^2(0, T; U) \). However, we are not allowed to perform the limit passage after splitting up all the integrals, since the duality pairing in the limit would not be well defined because we only know that \( u' + D \mathcal{E}_r(u) \in L^p(0, T; V^*) \). Nevertheless, since we have assumed \( V \hookrightarrow \overline{W} \), we can treat the term involving \( D \mathcal{E}_r^2 : \overline{W} \to \overline{W}^* \hookrightarrow V^* \) separately. First, taking into account
\[
u_0 + (K S_h u')(t) = u_0 + \frac{1}{2h} \int_{t-h}^{t+h} \dot{u}(r) \, dr - \frac{1}{2h} \int_{-h}^{h} \dot{u}(r) \, dr
\]
and that \( u \in C_w([0, T]; U) \subset C([0, T]; \overline{W}) \) since \( U \subset \hookrightarrow \overline{W} \), there holds
\[
\lim_{h \to 0} (u_0 + (K S_h u')) = u \quad \text{in } C([0, T]; \overline{W}). \tag{2.17}
\]
Finally, by the continuity of $\mathcal{E}_t^2$ and $D\mathcal{E}_t^2$, the convergences (2.17) and $S_h u' \to u'$ in $L^q(0, T; V)$ as $h \to 0$, there holds
\[
\begin{align*}
&= -\int_s^t \langle D\mathcal{E}_t^2(u(r)), u'(r) \rangle_{V^* \times V} \, dr \\
&= \lim_{h \to 0} \int_s^t \langle D\mathcal{E}_t^2(u_0 + (KS_h u')(r)), (S_h u')(r) \rangle_{V^* \times V} \, dr \\
&= \lim_{h \to 0} \left( \mathcal{E}_s^2(u_0 + (KS_h u')(s)) - \mathcal{E}_0^2(u_0 + (KS_h u')(t)) \right) \\
&= \mathcal{E}_s^2(u(s)) - \mathcal{E}_0^2(u(t)) 
\end{align*}
\] (2.18)
for all $s, t \in [0, T]$. Second, it has been shown in Emmrich & Thalhammer [EmT11, Lemma 6] that
\[
-\int_0^t \langle u''(r) + E(u(r)), u'(r) \rangle_{V^* \times V} \, dr \\
\leq \frac{1}{2} |v_0|^2 - \frac{1}{2} |u'(t)|^2 + \mathcal{E}^{1}(u_0) - \mathcal{E}^{1}(u(t))
\]
for almost every $t \in (0, T)$. The latter inequality together with (2.18) implies (2.16), which completes the proof.

\[\square\]

### 2.5 Proof of Theorem 2.4

Let $u_0 \in U, v_0 \in H$ and $(\tau_n)_{n \in N}$ be a vanishing sequence of positive step sizes. Let $(u_0^k)_{k \in \mathbb{N}} \subset U \cap V$ and $(v_0^k)_{k \in \mathbb{N}} \subset V$ be such that $u_0^k \to u_0$ in $U$ and $v_0^k \to v_0$ in $H$ as $k \to \infty$. We let $k \in \mathbb{N}$ be fixed and we denote the interpolations associated with the initial data $u_0^k$ and $v_0^k$ as before. Henceforth, we suppress the dependence of the interpolations on $k$ for simplicity. By the previous lemma, there exists a subsequence (relabeled as before) of the interpolations and limit functions $u \in C_u([0, T]; U) \cap W^{1, \infty}(0, T; H) \cap W^{1,q}(0, T; V^*) \cap W^{2,r^*}(0, T; U^* + V^*)$ (notice that $u_0^k \in U \cap V$) and $u(0) = u_0^k$ in $U$ and $u'(0) = v_0^k$ in $H$ such that the convergences (2.14) hold, where we again suppress the dependence of the limit functions on $k$. First, we prove that the inclusion (2.3) holds. To do so, we note that the Euler–Lagrange equation (2.13) reads
\[
\begin{align*}
\tilde{V}_{\tau_n}(t) + &\eta_{\tau_n}(t) + D\mathcal{E}_{\tau_n}(t)(\overline{U}_{\tau_n}(t)) + S_{\tau_n}(t) = 0 \quad \text{in } U^* + V^*, \\
&\eta_n(t) \in \partial_v\Psi_{U_{\tau_n}}(\nabla_{\tau_n}(t)) 
\end{align*}
\] (2.19)
for all $t \in (0, T)$, where $S_{\tau_n}(t) = B(\overline{U}_{\tau_n}(t), \overline{V}_{\tau_n}(t), \overline{U}_{\tau_n}(t)) - f_{\tau_n}(t), t \in [0, T]$. Testing equation (2.19) with $w \in L^{\max(2,q)}(0, T; U \cap V)$, we obtain
\[
\int_0^T \langle \tilde{V}_{\tau_n}(t) + \eta_{\tau_n}(t) + D\mathcal{E}_{\tau_n}(s)(\nabla_{\tau_n}(s)), S_{\tau_n}(r), w(r) \rangle_{(U^* + V^*) \times (U \cap V)} \, dt = 0.
\]
Then, with the aid of the convergences (2.14), we are allowed to pass to the limit in the weak formulation obtaining
\[
\int_0^T \langle u''(r) + \eta(r) + D\mathcal{E}_a(u(s)) + B(t, u(r), u'(r)) - f(r), w(r) \rangle_{(U^* + V^*) \times (U \cap V)} \, dr = 0
\]
for all \( w \in L^{\max\{2,q\}}(0, T; U \cap V) \). Then, by a density argument and the fundamental lemma of calculus of variations, we deduce

\[
u''(t) + \eta(t) + D\mathcal{E}_t(u(t)) + B(t, u(t), u'(t)) = f(t) \quad \text{in } U^* + V^*
\]

for a.e. \( t \in (0, T) \). We shall identify the weak limit \( \eta \) as subgradient of the dissipation potential almost everywhere, i.e., \( \eta(t) \in \partial_V \Psi_{u(t)}(u'(t)) \) for almost every \( t \in (0, T) \). For that purpose, we will employ Lemma 1.6 with \( f_n(t, v) = \Psi_{\mathcal{L}u(t)}(v) \) and \( f(t, v) = \Psi_{u(t)}(v) \) for all \( v \in X = V \) and \( n \in \mathbb{N} \). Assumption (1.5) is already fulfilled by Condition (2.\( \Psi_c \)).

Hence, it remains to show that

\[
\limsup_{n \to \infty} \int_0^T \langle \eta_n(t), \nabla_{\tau_n}(t) \rangle_{V^* \times V} dt \leq \int_0^T \langle \eta(t), u'(t) \rangle_{V^* \times V} dt.
\]

In order to show the latter limes superior estimate, we use the fact that \( \eta_{\tau_n} \) can be expressed through the remaining terms of the Euler–Lagrange equation (2.19). Therefore, we will split the integral on the left-hand side of (2.20) and note first that

\[
- \int_0^t \langle \hat{V}'_{\tau_n}(r), \nabla_{\tau_n}(r) \rangle_{V^* \times V} dr
= - \int_0^t \langle \hat{V}'_{\tau_n}(r), \hat{V}_{\tau_n}(r) \rangle_{V^* \times V} dr + \int_0^t \langle \hat{V}'_{\tau_n}(r), \hat{V}_{\tau_n}(r) - \nabla_{\tau_n}(r) \rangle_{V^* \times V} dr
= \frac{1}{2} |v_0|^2 - \frac{1}{2} |\hat{V}_{\tau_n}(t)|^2 + \int_0^t \langle \hat{V}'_{\tau_n}(r), \hat{V}_{\tau_n}(r) - \nabla_{\tau_n}(r) \rangle_{V^* \times V} dr
\leq \frac{1}{2} |v_0|^2 - \frac{1}{2} |\hat{V}_{\tau_n}(t)|^2,
\]

where we used the fundamental theorem of calculus for the absolutely continuous function \( t \mapsto \frac{1}{2} |\hat{V}_{\tau_n}(t)|^2 \) and that the estimate

\[
\int_0^t \langle \hat{V}'_{\tau_n}(r), \hat{V}_{\tau_n}(r) - \nabla_{\tau_n}(r) \rangle_{V^* \times V} dr
= \sum_{i=1}^{m-1} \int_{t_{i-1}}^{t_i} \left( \frac{V^i_{\tau_n} - V^{i-1}_{\tau_n}}{\tau_n}, \frac{r - t_{i-1} + V^{i-1}_{\tau_n} t_i - r}{\tau_n} - V^i_{\tau_n} \right) dr
+ \int_{t_{m-1}}^{t} \left( \frac{V^m_{\tau_n} - V^{m-1}_{\tau_n}}{\tau_n}, \frac{r - t_{m-1} + V^{m-1}_{\tau_n} t_m - r}{\tau_n} - V^m_{\tau_n} \right) dr
= - \sum_{i=1}^{m-1} \int_{t_{i-1}}^{t_i} \left( \frac{V^i_{\tau_n} - V^{i-1}_{\tau_n}}{\tau_n}, \left( \frac{V^i_{\tau_n} - V^{i-1}_{\tau_n}}{\tau_n} \right) \frac{t_i - r}{\tau_n} \right) dr
- \int_{t_{m-1}}^{t} \left( \frac{V^m_{\tau_n} - V^{m-1}_{\tau_n}}{\tau_n}, \left( \frac{V^m_{\tau_n} - V^{m-1}_{\tau_n}}{\tau_n} \right) \frac{t_m - r}{\tau_n} \right) dr
= - \sum_{i=1}^{m-1} \int_{t_{i-1}}^{t_i} \frac{t_i - r}{\tau_n} |V^i_{\tau_n} - V^{i-1}_{\tau_n}|^2 dr
- \int_{t_{m-1}}^{t} \frac{t_m - r}{\tau_n} |V^m_{\tau_n} - V^{m-1}_{\tau_n}|^2 dr \leq 0
\]

with \( t \in (t_{m-1}, t_m] \) for some \( m \in \{1, \ldots, N\} \).

We continue with the term involving the derivative of the energy functional and start with the linear part:
where we used
\[ \int_0^t \langle E(\tilde{U}_{\tau_n}(r) - \overline{U}_{\tau_n}(r)), \nabla_{\tau_n}(r) \rangle_{V^* \times U} \, dr \leq 0, \]
which can be shown in the same way as above by using the strong positivity of \( E \). As for the nonlinear part, we obtain by employing the \( \lambda \)-convexity of \( \mathcal{E}_t^2 \) that
\[
- \int_0^t \langle D\mathcal{E}_t^2(\overline{U}_{\tau_n}(r)), \nabla_{\tau_n}(r) \rangle_{V^* \times U} \, dr \\
= - \sum_{i=1}^{m-1} \langle D\mathcal{E}_t^2(U_{\tau_n}^i), U_{\tau_n}^i - U_{\tau_n}^{i-1} \rangle_{V^* \times U} - \frac{t - t_m - 1}{\tau_n} \langle D\mathcal{E}_t^2(U_{\tau_n}^m), U_{\tau_n}^m - U_{\tau_n}^{m-1} \rangle_{V^* \times U} \\
\leq - \sum_{i=1}^{m-1} \left( \mathcal{E}_t^2(U_{\tau_n}^{i-1}) - \mathcal{E}_t^2(U_{\tau_n}^i) - \lambda \left| U_{\tau_n}^i - U_{\tau_n}^{i-1} \right|^2 \right) \\
- \frac{t - t_m - 1}{\tau_n} \left( \mathcal{E}_t^2(U_{\tau_n}^m) - \mathcal{E}_t^2(U_{\tau_n}^{m-1}) - \lambda \left| U_{\tau_n}^m - U_{\tau_n}^{m-1} \right|^2 \right) \\
= - \sum_{i=1}^{m-1} \left( \mathcal{E}_t^2(U_{\tau_n}^{i-1}) - \mathcal{E}_t^2(U_{\tau_n}^i) + \int_{t_{i-1}}^{t_i} \partial_r \mathcal{E}_r^2(U_{\tau_n}^i) \, dr + \lambda \tau_n \left| V_{\tau_n}^i \right|^2 \right) \\
+ \frac{t_m - t}{\tau_n} \left( \mathcal{E}_t^2(U_{\tau_n}^m) - \mathcal{E}_t^2(U_{\tau_n}^{m-1}) - \lambda \left| U_{\tau_n}^m - U_{\tau_n}^{m-1} \right|^2 \right) \\
= \mathcal{E}_0^2(u_0) - \mathcal{E}_t^2(\overline{U}_{\tau_n}(t)) + \int_0^{\overline{T}_{\tau_n}(t)} \partial_r \mathcal{E}_r^2(\overline{U}_{\tau_n}(r)) \, dr + I_n(t),
\]
where
\[
I_n(t) = \frac{t_m - t}{\tau_n} \left( \mathcal{E}_t^2(U_{\tau_n}^m) - \mathcal{E}_t^2(U_{\tau_n}^{m-1}) - \lambda \left| U_{\tau_n}^m - U_{\tau_n}^{m-1} \right|^2 \right) + \lambda \tau_n \int_0^{T_{\tau_n}(t)} \left| \nabla_{\tau_n}(r) \right|^2 \, dr.
\]
Now, we want to make use of the inequality (2.16). However, the aforementioned inequality only holds true for almost every \( t \in (0, T) \). Therefore, we take a sequence of increasing values \( (\beta_i)_{i \in \mathbb{N}} \), \( \beta_i \in (0, T) \) for all \( i \in \mathbb{N} \), converging to \( T \) for which (2.16) holds true. Then, choosing \( t = \beta_i \), we obtain with the convergences (2.14b), (2.14h), (2.14e), and (2.14g), the sequential weak lower semicontinuity of \( \mathcal{E}_t^1 \) and \( \left| \cdot \right| \) and the continuity of \( \mathcal{E}_t^1 \), the limes superior condition and growth condition (2. Ed) on \( \partial_r \mathcal{E}_r^2 \) and FÄTUO’S Lemma that
\[
\limsup_{n \to \infty} - \int_0^{\beta_i} \langle \hat{\nabla}_{\tau_n}^t(r), D\mathcal{E}_{\tau_n}(r) \rangle_{V^* \times V} \, dr \\
\leq \limsup_{n \to \infty} \left( \frac{1}{2} |v_0|^2 - \frac{1}{2} \left| \hat{V}_{\tau_n}(\beta_i) \right|^2 + \mathcal{E}_0(u_0) - \mathcal{E}_t^1(\hat{U}_{\tau_n}(\beta_i)) - \mathcal{E}_t^2(\overline{U}_{\tau_n}(\beta_i)) \\
+ \int_0^{\overline{T}_{\tau_n}(\beta_i)} \partial_r \mathcal{E}_r(\overline{U}_{\tau_n}(r)) \, dr + I_n(t) \right) \\
\leq \frac{1}{2} |v_0|^2 - \frac{1}{2} \left| u'(\beta_i) \right|^2 + \mathcal{E}_0(u_0) - \mathcal{E}_t^1(u(\beta_i)) + \int_0^{\beta_i} \partial_r \mathcal{E}_r(u(r)) \, dr.
\]
Since \( u \in C_w([0,T];U) \) and \( u' \in L^\infty(0,T;H) \cap W^{1,1}(0,T;U^* + V^*) \subset C_w([0,T];H) \), Lemma 2.7 then yields
\[
\frac{1}{2}|v_0|^2 - \frac{1}{2}|u'(|\beta_1)|^2 + \mathcal{E}_0(u_0) - \mathcal{E}_\beta(u(\beta_1)) + \int_0^{\beta_1} \partial_t \mathcal{E}_r(u(r)) \, dr
\leq - \int_0^{\beta_1} (u''(r) + D\mathcal{E}_r(u(r)), u'(r))_{V^* \times V}.
\]

Then, in view of the convergences (2.14m) and (2.14n), the Euler–Lagrange equation (2.19), we obtain
\[
\limsup_{n \to \infty} \int_0^{\beta_1} \langle \eta_n(t), \nabla \tau_n(t) \rangle_{V^* \times V} \, dt
= \limsup_{n \to \infty} \int_0^{\beta_1} \langle S_{\tau_n}(t) - \tilde{V}_{\tau_n}(t) - D\mathcal{E}_{\tau_n(t)}(\nabla \tau_n(t)), \nabla \tau_n(t) \rangle_{V^* \times V} \, dt
\leq \int_0^{\beta_1} \langle f(t) - B(t, u(t), u'(t)) - u''(t) - D\mathcal{E}_r(u(t)), u'(t) \rangle_{V^* \times V} \, dt
= \int_0^{\beta_1} \langle \eta(t), u'(t) \rangle_{V^* \times V} \, dt.
\]

Together with Condition (2.\Psi c) and Lemma 1.6, this implies \( \eta(t) \in \partial \Psi_{\bar{u}(t)}(u'(t)) \) for almost every \( t \in (0, \beta_1) \) for all \( l \in \mathbb{N} \). Letting \( l \to \infty \) leads to \( \eta(t) \in \partial \Psi_{\bar{u}(t)}(u'(t)) \) for almost every \( t \in (0, T) \). This shows for each \( k \in \mathbb{N} \) the existence of a function \( u \) satisfying the inclusion (2.3), and the initial values \( u(0) = u_0^k \in U \cap V \) and \( u'(0) = v_0^k \in V \). Denote with \( (u_k)_{k \in \mathbb{N}} \) the sequence of solutions to the associated sequence of initial values, and with \( (\eta_k)_{k \in \mathbb{N}} \) the subgradients of \( \Psi_{\bar{u}(t)}(u'_k(t)) \). In the last step, we want to show that there exists a limit function \( u \) which satisfies (2.3) and (2.4) as well as the initial values \( u(0) = u_0 \) in \( U \) and \( u'(0) = v_0 \) in \( H \). We recall that \( u_0^k \to u_0 \) in \( U \) and \( v_0^k \to v_0 \) in \( H \) as \( k \to \infty \). The next steps are the following:

1. We derive a priori estimates based on the energy-dissipation inequality (2.4),
2. We show compactness of the sequences \( (u_k)_{k \in \mathbb{N}} \) and \( (\eta_k)_{k \in \mathbb{N}} \) in appropriate spaces,
3. We pass to the limit in the inclusion 2.3 as \( k \to \infty \).

Ad 1. Let \( t \in [0,T] \) and \( \mathcal{N} \subset (0,T) \) a set of measure zero such that \( \mathcal{E}_{\tau_n(s)}(\nabla \tau_n(s)) \to \mathcal{E}_r(u(s)) \) and \( \nabla \tau_n(s) \to u'(s) \) for each \( s \in [0,T] \setminus \mathcal{N} \). Then, employing the convergences (2.14), we obtain
\[
\frac{1}{2}|u'_k(t)|^2 + \mathcal{E}_r(u_k(t)) + \int_0^t \left( \Psi_{u_k(r)}(u'_k(r)) + \Psi_{u_k(r)}'(S_k(r) - D\mathcal{E}_r(u_k(r)) - u''_k(r)) \right) \, dr
\leq \liminf_{n \to \infty} \left( \frac{1}{2} |\nabla \tau_n(t)|^2 - \mathcal{E}_\tau(t) \langle \nabla \tau_n(t), \nabla \tau_n(t) \rangle \right)
+ \int_0^{\tau_{\tau_n}(t)} \left( \Psi_{\bar{u}_\tau(r)}(\nabla \tau_n(r)) + \Psi_{\bar{u}_\tau(r)}'(S_{\tau_n}(r) - D\mathcal{E}_{\tau_n(r)}(\nabla \tau_n(r)) - \tilde{V}_{\tau_n}(r)) \right) \, dr
\leq \limsup_{n \to \infty} \left( \frac{1}{2} |v_0^k|^2 + \mathcal{E}_0(u_0^k) + \int_0^{\tau_{\tau_n}(t)} \partial_t \mathcal{E}_r(\nabla \tau_n(r)) \, dr \right)
+ \int_0^{\tau_{\tau_n}(t)} \langle S_{\tau_n}(r), \nabla \tau_n(r) \rangle_{V^* \times V} \, dr + \tau \lambda \int_0^{\tau_{\tau_n}(t)} \|V_r(r)\|^2 \, dr
= \frac{1}{2} |v_0^k|^2 + \mathcal{E}_0(u_0^k) + \int_0^t \partial_t \mathcal{E}_r(u_k(r)) \, dr + \int_0^t \langle S_{\tau_n}(r), u'_k(r) \rangle_{V^* \times V} \, dr.
\]
for all $t \in [0, T]$, where $S_k(r) = f(r) - B(r, u_k(r), u'_k(r))$. Again, taking into account Condition (2. Ed), (2.Bb), and (2.Bb), we obtain with the lemma of Gronwall (Lemma A.1)

$$\frac{1}{2}|u'_k(t)|^2 + \mathcal{E}_k(u_k(t)) + \int_0^t (\Psi(u'_k(r)) + \Psi'(S(r) - D\mathcal{E}_r(u_k(r)) - u''_k(r))) \, dr \leq C_B.$$ 

for all $t \in [0, T]$ for a constant $C_B > 0$.

Ad 2. With the same reasoning as for the interpolations, we obtain the convergences

\begin{align*}
&u_k \overset{\star}{\rightharpoonup} u \quad \text{in} \quad L^\infty(0, T; U), \quad (2.21a) \\
&u - u_0 \overset{\star}{\rightharpoonup} u - u_0 \quad \text{in} \quad L^\infty(0, T; V), \quad (2.21b) \\
&u_k(t) \to u(t) \quad \text{in} \quad U \quad \text{for all} \quad t \in [0, T], \quad (2.21c) \\
&u_k(t) - u_0^k \to u(t) - u_0 \quad \text{in} \quad V \quad \text{for all} \quad t \in [0, T], \quad (2.21d) \\
&u_k \to u \quad \text{in} \quad L^r(0, T; \widetilde{W}) \quad \text{for any} \quad r \geq 1, \quad (2.21e) \\
&u_k(t) \to u(t) \quad \text{in} \quad \tilde{W} \quad \text{for all} \quad t \in [0, T], \quad (2.21f) \\
&u'_k(t) \oversett{\star}{\to} u' \quad \text{in} \quad L^q(0, T; V) \cap L^\infty(0, T; H), \quad (2.21g) \\
&u'_k(t) \to u' \quad \text{in} \quad L^p(0, T; H) \quad \text{for all} \quad p \geq 1, \quad (2.21h) \\
&u''_k(t) \to u'' \quad \text{in} \quad H \quad \text{for all} \quad t \in [0, T], \quad (2.21i) \\
&\eta_{\tau_n} \overset{\star}{\to} \eta \quad \text{in} \quad L^q(0, T; V^*), \quad (2.21j) \\
&Eu_k \rightharpoonup Eu \quad \text{in} \quad L^2(0, T; U^*), \quad (2.21k) \\
&D\mathcal{E}_t^2(u_k) \to D\mathcal{E}_t^2(u) \quad \text{in} \quad L^r(0, T; U^*) \quad \text{for any} \quad r \geq 1, \quad (2.21l) \\
&u''_k \oversett{\star}{\to} u'' \quad \text{in} \quad L^\min(2, q^*)\,(0, T; U^* + V^*), \quad (2.21m) \\
&B(\cdot, u_k, u'_k) \to B(\cdot, u, u') \quad \text{in} \quad L^r(0, T; V^*), \quad (2.21n)
\end{align*}

and if $\Psi_u$ satisfies (2.Ψd), then

\begin{align*}
&u'_k \to u' \quad \text{in} \quad L^{\max(2, q)}(0, T; U), \\
&u_k \to u \quad \text{in} \quad C([0, T]; U).
\end{align*}

Ad 3. Therefore, $u \in C_w([0, T]; U) \cap W^{1,\infty}([0, T]; H) \cap W^{2, q^*}(0, T; U^* + V^*)$ with $u - u_0 \in W^{1,q}(0, T; V)$ and $\eta \in L^{q^*}(0, T; V^*)$ satisfies the initial conditions $u(0) = u_0$ in $U$ and $u'(0) = v_0$ in $H$. Along the same lines as for the interpolations, we obtain with Condition (2.Ψc) and Lemma 1.6 that $u$ and $\eta$ satisfy the inclusions (2.3). It remains to show the energy-dissipation balance (2.4). The inequality

$$\frac{1}{2}|u'(t)|^2 + \mathcal{E}_t(u(t)) + \int_0^t \left(\Psi_{u(r)}(u'(r)) + \Psi'_{u(r)}(S(r) - D\mathcal{E}_r(u(r)) - u''(r))\right) \, dr \leq \frac{1}{2}|v_0|^2 + \mathcal{E}_0(u_0) + \int_0^t \partial_r \mathcal{E}_r(u(r)) \, dr + \int_0^t (S(r), u'(r))_{V^* \times V} \, dr,$$

for all $t \in [0, T]$ with $S(r) = f(r) - B(r, u(r), u'(r))$ is obtained by passing to the limit as $k \to \infty$ while taking into account the convergences (2.14). Then, employing again (2.16)
and the Fenchel–Young inequality, we obtain
\[
\int_0^t \left( \Psi_{u(t)}(u'(r)) + \Psi_{u(r)}^*(S(r) - D\mathcal{E}_r(u(r)) - u''(r)) \right) \, dr \\
\leq \frac{1}{2} |v_0|^2 - \frac{1}{2} |u'(t)|^2 + \mathcal{E}_0(u_0) - \mathcal{E}_T(u(t)) + \int_0^t \partial_r \mathcal{E}_r(u(r)) \, dr \\
+ \int_0^t \langle S(r), u'(r) \rangle_{V^* \times V} \, dr \\
\leq \int_0^t \langle D\mathcal{E}_r(u(r)) - u''(r), u'(r) \rangle_{V^* \times V} \, dr + \int_0^t \langle S(r), u'(r) \rangle_{V^* \times V} \, dr \\
= \int_0^t (S(r) - D\mathcal{E}_r(u(r)) - u''(r), u'(r))_{V^* \times V} \, dr \\
\leq \int_0^t \left( \Psi_{u(t)}(u'(r)) + \Psi_{u(r)}^*(S(r) - D\mathcal{E}_r(u(r)) - u''(r)) \right) \, dr
\]
for almost every \( t \in (0, T) \). Now, if \( V \to U \), then the inequality (2.16) indeed holds as equality for all \( t \in [0, T] \) by the classical integration by parts formula. This shows (2.4), and hence the completion of the proof. \( \square \)

Remark 2.8. The proof of Theorem 2.4 reveals that one can consider dissipation potentials that depend on a parameter \( \varepsilon \). In this case, the Condition (2.Ψa) is assumed to hold for every \( \varepsilon \geq 0 \) while Condition (2.Ψb) holds uniformly in \( \varepsilon \geq 0 \). Condition (2.Ψc) can either be replaced with the Mosco-convergence \( \Psi_{u_n}^\varepsilon \to M, \Psi_u^0 \) for every sequence \( u_n \to u \) as \( \varepsilon \to 0 \), or with a more general liminf estimate (1.5).

3 Applications

In this section, we want to apply the abstract result developed in the previous sections to a concrete examples. First, we discuss in detail a mathematical example to illustrate the strength of the theory and highlight the case (a) and (b). Then, we also discuss some physically meaningful example.

3.1 A viscous regularization of the Klein–Gordon equation

The following example is a nonlinearly damped inertial system and can be interpreted as a viscous regularization KLEIN–GORDON equation. The equations supplemented with initial and boundary conditions are given by

\[
(P4) \quad \begin{cases}
\partial_t u - \nabla \cdot p - \Delta u + b(u) = f & \text{in } \Omega_T, \\
p(x, t) \in \partial_v \psi(x, u(x, t), \nabla \partial_t u(x, t)) & \text{a.e. in } \Omega_T, \\
u(x, 0) = u_0(x) & \text{on } \Omega, \\
u'(x, 0) = v_0(x) & \text{on } \Omega, \\
u(x, t) = 0 & \text{on } \partial \Omega \times [0, T], \\
\frac{\partial v}{\partial v}(x, t) = 0 & \text{on } \partial \Omega \times [0, T].
\end{cases}
\]

If \( \psi = 0 \) and \( b(u) = \gamma u \) for a constant \( \gamma > 0 \), then the equation in (P4) reduces to the classical KLEIN–GORDON equation, which is a relativistic wave equation with applications in relativistic quantum mechanics that is related to the SCHRÖDINGER equation.
3.1 A VISCOUS REGULARIZATION OF THE KLEIN–GORDON EQUATION

We make the following assumptions on the functions $\psi$ and $b$. For simplicity, we choose $d = 1$ and note that the case $d \geq 2$ can be (under stronger assumptions) be treated in a similar way.

(3.a) The function $\psi : \Omega \times \mathbb{R} \times \mathbb{R} \to [0, +\infty)$ is a CARATHÉODORY function such that $\psi(x, y, \cdot)$ is a proper, lower semicontinuous, and convex, and $\psi(y, y, 0) = 0$ for almost every $x \in \Omega$ and all $y \in \mathbb{R}$.

(3.b) There exists a real number $q > 1$ and positive constants $c_\psi, C_\psi > 0$ such that

$$c_\psi^R (|z|^q - 1) \leq \psi(x, y, z) \leq C_\psi^R (1 + |z|^q)$$

for a.e. $x \in \Omega$ and all $z \in \mathbb{R}^m$, $y \in \mathbb{R}$, $|y| \leq R$.

(3.c) The function $b : \Omega \to \mathbb{R}$ is a continuous function and there exist a real number $p > 1$ and a constant $C_b > 0$ such that

$$|b(u)| \leq C_b(|u|^{p-1} + 1) \quad \text{for all} \quad u \in \mathbb{R}.$$

Accordingly, the function spaces are given by $V = W_0^{1,q}(\Omega)$, $U = H_0^1(\Omega)$, $\bar{W} = L^{\max(p,2)}(\Omega)$ and $H = L^2(\Omega)$. Then, we identify the dissipation potential $\Psi : V \to \mathbb{R}$ and the energy functional $\mathcal{E} : U \to [0, +\infty)$ as

$$\Psi_u(v) = \int_\Omega \psi(x, u(x), \nabla v(x)) \, dx \quad \text{and} \quad \mathcal{E}(u) = \frac{1}{2} \int_\Omega |\nabla u(x)|^2 \, dx,$$

respectively. The perturbation $B : \bar{W} \to V^*$ is given by

$$\langle B(u), w \rangle_{\bar{W}^*, \bar{W}} = \int_\Omega b(u(x)) w(x) \, dx.$$

We note that the conjugate functional $\Psi_u^*$ can, in general, not be expressed as an integral functional over $\Omega$, since it is defined on $W^{-1,q^*}$. Obviously, $\mathcal{E}$ satisfies all Conditions 2. In view of the compact embedding $H_0^1(\Omega) \hookrightarrow C(\overline{\Omega})$ and FATOU’s lemma, it is readily that $\Psi_u$ satisfies Conditions (2.Ψa) and (2.Ψb). In order to verify Condition (2.Ψc), we show that for every sequence $u_n \to u$ in $U$ with $\sup_{n \in \mathbb{N}} \mathcal{E}(u_n) < +\infty$, there holds $\Psi_{u_n} \rightharpoonup \Psi_u$ as $n \to \infty$. As we mentioned in before, the MOSCO-convergence $\Psi_{u_n} \rightharpoonup \Psi_u$ implies the MOSCO-convergence of the related integral functionals that in turn implies Condition (2.Ψc). The liminf estimate in the MOSCO-convergence follows from IOFFE [Iof77, Theorem 3]. The limsup estimate is trivially fulfilled by choosing, for each $v \in V$, the constant sequence $v_n = v, n \in \mathbb{N}$, and the dominated convergence theorem.

If we assume $p \in (1, 2]$, and $f \in L^2(0, T; H)$, it is easy to check in the same way as in the previous examples that Conditions (2.Ba), (2.Bb), and (2.Bb) are also fulfilled. Therefore, Theorem 2.4 ensures that for every initial values $v_0 \in H$ and $u_0 \in U$, the existence of a solution $u \in C_w([0, T]; U) \cap W^{1,\infty}(0, T; H) \cap W^{2,q^*}(0, T; U^* + V^*)$ with $u - u_0 \in W^{1,q}(0, T; V)$ to (P4) satisfying the integral equation

$$\int_0^T \left( \langle u''v \rangle_{(U^* + V^*) \times (U \cap V)} + \int_\Omega p \cdot \nabla v + b(u)v \, dx \right) \, dt = \int_0^T \int_\Omega f(v) \, dx \, dt.$$
for all $v \in L^{\min(2,q^*)}(0,T;U^* + V^*)$ with $p(x,t) \in \partial_v \psi(x,u(x,t),\nabla \partial_t u(x,t))$ a.e. in $\Omega_T$, and the energy-dissipation balance

$$
\frac{1}{2} \|u'(t)\|_{L^2(\Omega)}^2 + \frac{1}{2} \|u(t)\|_{H^1_0(\Omega)}^2 + \int_0^t \left( \Psi_{u(t)}(u'(r)) + \Psi^*_u(f(r) - u''(r) - \Delta u(r)) \right) \, dr
$$

$$
= \frac{1}{2} \|v_0\|_{L^2(\Omega)}^2 + \frac{1}{2} \|u_0\|_{H^1_0(\Omega)}^2 + \int_0^t \langle f(r), u'(r) \rangle_{L^2(\Omega) \times L^2(\Omega)} \, dr
$$

holds for almost every $t \in (0,T)$ if $q \in (1,2)$ and for all $t \in (0,T)$ if $q \geq 2$.

### 3.2 Differential inclusion

In the final example, we consider a nonlinearly damped inertial system which can not be treated with the known abstract results. The differential inclusion supplemented with initial and boundary conditions is given by

$$
\begin{align*}
\partial_t u + |\partial_t u|^{q-2} \partial_t u + p - \Delta u &= f & \text{in } \Omega_T, \\
p(x,t) &\in \text{Sgn}(\partial_t u(x,t)) & \text{a.e. in } \Omega_T, \\
u(x,0) &= u_0(x) & \text{on } \Omega, \\
u(x,0) &= 0 & \text{on } \partial \Omega \times [0,T], \\
u(x,t) &= 0 & \text{on } \partial \Omega \times [0,T],
\end{align*}
$$

where $q \geq 2$ and $f \in L^2(0,T;H)$. We set $U = H^1_0(\Omega)$, $V = L^q(\Omega)$, and $H = L^2(\Omega)$. The dissipation potential $\Psi : V \to \mathbb{R}$ and the energy functional $E : U \to [0, +\infty]$ are given by

$$
\Psi(u) = \int_\Omega \left( \frac{1}{q} |v(x)|^q + |v(x)| \right) \, dx \quad \text{and} \quad E(u) = \frac{1}{2} \int_\Omega |\nabla u(x)|^2 \, dx,
$$

respectively. Consequently, $B = 0$ and $E^2_t = 0$. Again, all the assumptions are easily verified, so that Theorem 2.4 ensures for any initial values $v_0 \in H$ and $u_0 \in U$ the existence of a solution $u \in C_w([0,T];U) \cap W^{1,\infty}(0,T;H) \cap W^{2,q^*}(0,T;U^* + V^*)$ with $u - u_0 \in W^{1,q}(0,T;V)$ to (P5) fulfilling the integral equation

$$
\int_0^T \left( \langle u''(v), v' + V^* \rangle \right) \, dt + \int_\Omega |\partial_t u|^{q-2} \partial_t u v + \nabla u \cdot \nabla v \, dx \, dt
$$

for all $v \in L^{2,q^*}(0,T;U^* + V^*)$ with $p(t,x) \in \text{Sgn}(u(x,t))$ a.e. in $\Omega_T$, and the energy-dissipation balance

$$
\frac{1}{2} \|u'(t)\|_{L^2(\Omega)}^2 + \frac{1}{2} \|u(t)\|_{H^1_0(\Omega)}^2 + \int_0^t \left( \Psi(u'(r)) + \Psi^*(f(r) - u''(r) - \Delta u(r)) \right) \, dr
$$

$$
= \frac{1}{2} \|v_0\|_{L^2(\Omega)}^2 + \frac{1}{2} \|u_0\|_{H^1_0(\Omega)}^2 + \int_0^t \langle f(r), u'(r) \rangle_{L^2(\Omega) \times L^2(\Omega)} \, dr
$$

holds for almost every $t \in (0,T)$.

### A Appendix

#### A.1 The Gronwall lemma

In this section, we provide two versions of the Gronwall lemma, the discrete and the classical version. The GROWWALL lemma is indispensable for obtaining a priori estimates or to show stability or uniqueness results.
Lemma A.1 (Gronwall). Let $T \in (0, +\infty]$, $s \in [0, T)$, $a, b \in L^\infty(s, T)$, $\lambda \in L^1(s, T)$ with $\lambda(t) \geq 0$ almost everywhere in $(s, T)$ such that
\[
a(t) \leq b(t) + \int_s^t \lambda(r)a(r) \, dr \quad \text{a.e. in } (s, T),
\]
then, there holds
\[
a(t) \leq b(t) + \int_s^t e^{A(t)-A(s)} \lambda(r)b(r) \, dr \quad \text{a.e. in } (s, T),
\]
where $A(t) = \int_s^t \lambda(r) \, dr$, $t \in [s, T]$.

Proof. A proof can be found in Emmrich [Emm04, Lemma 7.3.1, pp.180]. □

Lemma A.2 (Discrete Gronwall). Let $A, \alpha \in [0, +\infty)$ and $\alpha_n, \tau_n \in [0, +\infty)$ for all $n \in \mathbb{N}$ be satisfying
\[
a_n \leq A + \alpha \sum_{k=1}^n \tau_k a_k \quad \text{for all } n \in \mathbb{N}, m := \sup_{n \in \mathbb{N}} \alpha \tau_n < 1.
\]
Then, setting $\beta = \alpha/(1-m)$, $B = A/(1-m)$ and $\tau_0 = 0$, there holds
\[
a_n \leq Be^{\beta \sum_{k=1}^{n-1} \tau_k} \quad \text{for all } n \in \mathbb{N}.
\]

Proof. A proof can be found in Ambrosio et al. [AGS08, Lemma 3.2.4, p. 68]. □

A.2 A compactness result

In this section, we provide a version of the 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.

Lemma A.3 (Lions–Aubin–Dubinskii). Let $X$, $B$ and $Y$ be Banach spaces such that the embedding $X \hookrightarrow B$ is compact and the embedding $B \hookrightarrow Y$ is continuous. Furthermore, let either $1 \leq p < \infty$ and $r = 1$ or $p = \infty$ and $r > 1$, and let $(u_n)_{n \in \mathbb{N}}$ be a sequence of functions that are constant on each subinterval $((k-1)\tau_n, k\tau_n)$, $1 \leq k \leq n$, $T = n\tau_n$ satisfying
\[
\tau_n^{-1} \|\sigma_{\tau_n} u_{\tau_n} - u_{\tau_n}\|_{L^r(0,T;\tau_n;Y)} + \|u_{\tau_n}\|_{L^p(0,T;X)} \leq C \quad \text{for all } n \in \mathbb{N}, \tag{A.1}
\]
where $\sigma_{\tau_n} u := u(\cdot + \tau_n)$ and $C > 0$ is a constant which is independent of $\tau$. If $p < \infty$, then $(u_{\tau_n})_{n \in \mathbb{N}}$ is relatively compact in $L^p(0,T;B)$ and if $p = \infty$, there exists a subsequence of $(u_{\tau_n})_{n \in \mathbb{N}}$ converging in $L^q(0,T;B)$ for all $1 \leq q < \infty$ to a limit function belonging to $C([0,T];B)$.

Proof. A proof can be found in Dreher & Jüngel [DrJ12, Theorem 1]. □
References

[AGS08] L. Ambrosio, N. Gigli, and G. Savaré. Gradient Flows in Metric Spaces and in the Space of Probability Measures. Lectures in Mathematics ETH Zürich. Birkhäuser, Basel, 2nd edition, 2008.

[AtB86] H. Attouch and H. Brézis. Duality for the sum of convex functions in general Banach spaces. In Aspects of mathematics and its applications, volume 34 of North-Holland Math. Library, pages 125–133. North-Holland, Amsterdam, 1986.

[Bac20] A. Bacho. Doubly nonlinear evolution inclusion of second order with non-variational perturbation. arXiv:1910.10128, 2020.

[Bac21] A. Bacho. On the nonsmooth analysis of doubly nonlinear evolution inclusions of first and second order with applications. Doctoral thesis, Technische Universität Berlin, Berlin, 2021.

[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.

[BrÄ11] H. Brézis. Functional Analysis, Sobolev Spaces and Partial Differential Equations. Springer-Verlag, New York, 2011.

[DiU77] J. Diestel and J. J. Uhl, Jr. Vector Measures. American Mathematical Society, Providence, R.I., 1977. Mathematical Surveys, No. 15.

[DrJ12] M. Dreher and A. Jüngel. Compact families of piecewise constant functions in $L^p(0,T;B)$. Nonlinear Anal., 75(6), 3072–3077, 2012.

[EkT99] I. Ekeland and R. Témam. Convex analysis and variational problems, volume 28 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.

[Emm04] E. Emmrich. Gewöhnliche und Operator-Differentialgleichungen. Vieweg, Wiesbaden, 2004.

[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.

[Iof77] A. D. Ioffe. On lower semicontinuity of integral functionals. I. SIAM J. Control Optimization, 15(4), 521–538, 1977.

[LiS65] J.-L. Lions and W. A. Strauss. Some non-linear evolution equations. Bull. Soc. Math. France, 93, 43–96, 1965.

[Ste08] U. Stefanelli. The Brezis-Ekeland principle for doubly nonlinear equations. SIAM J. Control Optim., 47(3), 1615–1642, 2008.