On the Nonsmooth Analysis of Doubly Nonlinear Evolution Inclusions of First and Second Order with Applications

vorgelegt von
M. Sc.
Aras Bacho

an der Fakultät II – Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
– Dr. rer. nat. –

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr. Martin Skutella
Gutachter:	Prof. Dr. Etienne Emmrich
Gutachter:	Prof. Dr. Eduard Feireisl
Gutachter:	Prof. Dr. Alexander Mielke

Tag der wissenschaftlichen Aussprache: 17. Juni 2021

Berlin 2021

Le savant n'étudie pas la nature parce que cela est utile; il l'étudie parce qu'il y prend plaisir et il y prend plaisir parce qu'elle est belle. Si la nature n'était pas belle, elle ne vaudrait pas la peine d'être connue, la vie ne vaudrait pas la peine d'être vécue. Je ne parle pas ici, bien entendu, de cette beauté qui frappe les sens, de la beauté des qualités et des apparences; non que j'en fasse fi, loin de là, mais elle n'a rien à faire avec la science; je veux parler de cette beauté plus intime qui vient de l'ordre harmonieux des parties, et qu'une intelligence pure peut saisir.

— Jules Henri Poincaré (1854-1912)

Acknowledgements

I would like to thank my supervisor, Prof. Dr. Etienne Emmrich, for his continuous support since my bachelor studies and for fully believing in me so that I could follow my own ideas. I would also like to thank him for giving me the opportunity to attend numerous conferences all over the world, and for his many hints and comments on my dissertation. I thank Prof. Dr. Alexander Mielke for agreeing to act as a reviewer and for his collaboration within the SFB 910. I also thank Prof. Dr. Eduard Feireisl for agreeing to act as a reviewer of this thesis and for enriching our group during his Einstein professorship. I also thank his research team Anna Abbatiello, Danica Basaric, and Nilasis Chaudhuri for the many fun moments in the group. In particular, I thank André Eikmeier, Lukas Geuter, and Dr. Monika Eisenmann for the many corrections that made this thesis readable. I would also like to thank Richard Duong, Dr. Christian Kreusler, Dr. Robert Lasarzik, and Mathieu Rosière for interesting discussions, and our secretary Alexandra Schulte who took a lot of bureaucracy off our hands and provided us with enough pen and paper.

A special thanks goes to my wonderful family who has always supported me in many ways during all phases of my live and always encouraged me to become more.

To my parents.

Abstract

Many phenomena in nature are characterized by discontinuous processes. To describe them by suitable mathematical models is often a challenge. Evolution inclusions can be a suitable means to model such discontinuous processes mathematically.
The present work is devoted to the nonsmooth analysis of doubly nonlinear evolution inclusions of first and second order with leading subdifferential operators and non-monotone and non-variational perturbations using methods from the theory of convex analysis. The thesis is divided into two parts.
In the first part, we prove the existence of strong solutions to abstract Cauchy problems for perturbed generalized gradient flows for a certain class of nonlinear and monotone subdifferential operators acting on the time derivative of the solution, and nonlinear and non-monotone subdifferential operators acting on the solution as well as a certain class of perturbations. As an application of the abstract existence result, we show the existence of weak solutions of an initial-boundary value problem.
In the second part, we prove the existence of strong solutions to abstract Cauchy problems for doubly nonlinear evolution inclusions of second order. In doing so, we treat the equations with linear and nonlinear damping separately. In the case of linear damping, we consider a special class of leading linear potential operators acting on the time derivative of the solution, and nonlinear subdifferential operators acting on the solution. In the case of nonlinear damping, we consider the reverse case. In both cases, we allow a perturbation which depends nonlinearly on the solution as well as its time derivative. As an application of the abstract existence results, we prove the existence of weak solutions to certain initial-boundary value problems.

Zusammenfassung

Viele Phänomene in der Natur sind durch unstetige Prozesse charakterisiert. Diese durch ein geeignetes mathematisches Modell zu beschreiben, stellt oftmals eine Herausforderung dar. Evolutionsinklusionen können ein geeignetes Mittel sein, solche unstetigen Prozesse mathematisch zu modellieren.
Die vorliegende Arbeit widmet sich der nichtglatten Analyse von doppelt nichtlinearen Evolutionsinklusionen erster und zweiter Ordnung mit führenden Subdifferentialoperatoren und nicht-monotonen und nicht-variationellen Störungen mit Methoden aus der Theorie der konvexen Analysis. Die Arbeit ist in zwei Teile gegliedert.
Im ersten Teil weisen wir die Existenz von starken Lösungen zu abstrakten Cauchy Problemen für gestörte verallgemeinerte Gradientenflüsse für eine bestimmte Klasse von nichtlinearen und monotonen Subdifferentialoperatoren, welche auf die Zeit-Ableitung der Lösung wirken, und nichtlinearen und nicht-monotonen Subdifferentialoperatoren, welche auf die Lösung wirken, sowie einer bestimmten Klasse von Störungen nach. Als Anwendung des abstrakten Existenzresultats, zeigen wir die Existenz von schwachen Lösungen eines Anfangs-Randwertproblems.
Im zweiten Teil weisen wir die Existenz von starken Lösungen zu abstrakten Cauchy Problemen für doppelt nichtlineare Evolutionsinklusionen zweiter Ordnung nach. Dabei behandeln wir hyperbolische Gleichungen mit linearer und nichtlinearer Dämpfung gesondert. Im Fall der linearen Dämpfung betrachten wir eine spezielle Klasse von führenden linearen Potentialoperatoren, welche auf die Zeit-Ableitung der Lösung wirken, und nichtlinearen Potentialoperatoren, welche auf die Lösung wirken. Im Fall der nichtlinearen Dämpfung betrachten wir den umgekehrten Fall. In beiden Fällen erlauben wir eine Störung, welche nichtlinear von der Lösung sowie ihrer Zeit-Ableitung abhängt. Als Anwendung der abstrakten Existenzresultate weisen wir die Existenz von schwachen Lösungen zu gewissen Anfangs-Randwertproblemen nach.

Contents

Chapter 1

Introduction

People have always wrestled with understanding the essence of all things, whether it's been philosophers, mathematicians, artists, poets, composers, theologians, or natural scientists. As the German poet and natural scientist Johann Wolfgang von Goethe once formulated in his work Faust I, people desire to understand "what holds the world together at its innermost core"[1]. Attempts at answering this question are provided in a variety of methods, ranging from using music and painting, to language and mathematics. And even if these attempts may all seem different and illuminate only partial aspects of the truth (whatever that truth may be), they all agree on one thing: there lies inherent within objects a certain harmony, which is usually identified and described by its simplicity and beauty. The idea that objects carry such intrinsic qualities was first developed by Greek philosopher of antiquity, Plato, who assigned to every sensually perceptible object an abstract metaphysical form. Regardless of whether these abstractions are real, they undoubtedly contribute towards establishing and recognizing deeper connections between objects. This is especially true within the discipline of mathematics, which thrives on abstracting concrete objects. Here, abstraction serves as an indispensable means to find common structures and gain heuristic insights of apparently different objects for further examination. If a class of objects is successfully described on an abstract level, suitable methods can be developed to investigate these objects on an individual and class level, and in abstract and concrete forms. The present thesis is devoted to the study of abstract Cauchy problems for doubly nonlinear evolution inclusions of first and second order.

The history of Cauchy problems have their roots in several places. In 1926, the Austrian physicist Erwin Schrödinger postulated in his seminal work [149] the linear partial differential equation

$$i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{x},t) = \left(-\frac{\hbar}{2m}\Delta + V(\boldsymbol{x},t)\right)\Psi(\boldsymbol{x},t)$$

to describe a quantum mechanical state of a non-relativistic system and thus laid the foundation of quantum mechanics. Here, Ψ denotes the so-called wave function with the probability density $|\Psi(\boldsymbol{x},t)|^2$ which can be interpreted as the probability of a particle to stay in the point $\boldsymbol{x}$ in space and at the time t. In 1930, the English

[1] This a part of the English translation of line 382 in Goethe's original work [159]: "Dass ich erkenne, was die Welt im Innersten zusammenhält"

physicist Paul Dirac [60] generalized the Schrödinger equation to describe more general situations, which include relativistic effects, therefore improving upon the equation postulated by Schrödinger. He introduced the generalized equation in the so-called bra-ket notation[2] by

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle, \tag{1.0.1}$$

where $\hat{H}$ is the so-called Hamilton operator which acts on and takes values in an abstract Hilbert space and generates the time evolution of the quantum state described by $|\Psi\rangle$. The function $|\Psi\rangle$ can be seen as the abstract function associated to Ψ which can be related to each other via $|\Psi(t)\rangle(\boldsymbol{x}) = \Psi(\boldsymbol{x},t)$[3]. In 1933, Paul Dirac and Erwin Schrödinger received the Nobel prize "for the discovery of new productive forms of atomic theory".

Although the equation (1.0.1) can be seen as the first abstract evolution equation, the notion of an abstract Cauchy problem would not be formalized as such until 1952[4] by the American mathematician Einar Hille [92] following the concept of an Cauchy problem first coined in 1923 by the French mathematician Jacques Hadamard [90] for concrete problems. Hille investigated the abstract Cauchy problem

$$\begin{cases} u'(t) = Au(t), \quad t > 0, \\ u(0) = u_0 \end{cases} \tag{1.0.2}$$

with A being a linear, unbounded, and self-adjoint operator acting on a Banach space X. He gave necessary and sufficient conditions for the operator A to generate a strongly continuous semigroup (C_0-semigroup) of contractions which is directly related to the existence and uniqueness of *mild*[5] and *classical solutions*[6]. He rediscovered this result, known as the Hille–Yosida theorem, independently of the Japanese mathematician Kôsaku Yosida who already gave a proof in 1948 [162]. The abstract Cauchy problem (1.0.2) has been subsequently studied by many authors and also extended to the non-autonomous, i.e., the time-dependent case [93, 97, 98, 110, 132]. Nevertheless, it took more than two decades to prove the nonlinear counterpart of the Hille–Yosida theorem, which has been provided in 1971 by the American mathematicians Thomas M. Liggett and Michael G. Crandall [49]. They proved

[2] Although the bra-ket is attributed creatively to Paul Dirac, it was already introduced in the form $[\cdot|\cdot]$ in 1862 by the German mathematician Hermann Grassmann [88] to describe an inner product.

[3] Nowadays, most mathematicians use for the notational convenience the same notation for the concrete function and the corresponding abstract function.

[4] However, the term "evolution equation" goes back to the French mathematicians Laurent Schwartz [150] which introduced it in 1950. We refer the reader to Hazewinkel [91] and Fattorini [82] for more historical remarks on Cauchy problems.

[5] A function $u : [0,+\infty) \to X$ is called mild solution if $u(t) \in \mathrm{dom}(A)$ for almost every $t > 0$, $Au(t)$ is locally Bochner integrable, and there holds $u(t) = u_0 + \int_0^t Au(s)\,\mathrm{d}s$ for all $t > 0$, see Pazy [132].

[6] A function $u : [0,+\infty) \to X$ is called classical solution to (1.0.2) if $u(t) \in \mathrm{dom}(A)$ for all $t > 0$, it is continuous on $[0,+\infty)$ and continuously differentiable on $(0,+\infty)$, and it satisfies (1.0.2) pointwise everywhere.

in particular the existence of *strong solutions*, i.e., an absolutely continuous function with a certain regularity, which fulfills the initial condition and satisfies the equation pointwise almost everywhere, to the abstract CAUCHY problem for the more general nonlinear evolution inclusion

$$u'(t) + Bu(t) \ni 0, \quad t > 0,$$

for accretive operators B. This has first been extended in 1973 by the French mathematician Haïm BRÉZIS [32] to the case

$$Au'(t) + Bu(t) \ni f(t), \quad t \in (0, T), \tag{1.0.3}$$

where A is a linear, unbounded, and self-adjoint operator and B is a maximal monotone operator on a HILBERT space. This has been extended further in 1975 by the Romanian mathematician Viorel BARBU to the fully nonlinear case on a HILBERT spacem, where he assumed that both operators are subdifferential or subgradient operators, i.e., $A = \partial\psi$ and $B = \partial\phi$ for proper, lower semicontinuous, and convex functionals ψ and ϕ, see Section 2.2. This leads to the so-called *generalized gradient system*

$$\partial\psi(u'(t)) + \partial\phi(u(t)) \ni f(t), \quad t \in (0, T),$$

The equation[7] (1.0.3) is also referred to as the doubly nonlinear evolution equation of first type, whereas the equation

$$(Au(t))' + Bu(t) \ni f(t), \quad t \in (0, T), \tag{1.0.4}$$

is referred to as the doubly nonlinear evolution equation of second type [84]. Since the equation of second type was more interesting among mathematicians and physicists from an application point of view, it has been studied more extensively in the early 70s, see Section 1.2.

Regarding evolution equations of second order of the type

$$u''(t) + Au'(t) + Bu(t) = f(t), \quad t \in (0, T), \tag{1.0.5}$$

one can formally reduce it to a system of equations of first order by introducing the unknown variable $v = u'$ obtaining

$$\begin{pmatrix} v(t) \\ u(t) \end{pmatrix}' + \begin{pmatrix} A & B \\ -I & 0 \end{pmatrix} \begin{pmatrix} v(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} f(t) \\ 0 \end{pmatrix}, \quad t \in (0, T),$$

where I denotes the identity. However, this reduction can lead to the *well-posedness*[8] of the problem under relatively strong assumptions on the operators A und B, e.g., the linearity or LIPSCHITZ continuity, which reduces the number of application enormously. First results to fully nonlinear evolution equations have been obtained

[7] Although it is strictly speaking an inclusion we will sometimes refer to an inclusion as equation.

[8] The notion of the well-posedness of a CAUCHY problem has been introduced by Jacques HADAMARD [90] and describes CAUCHY problems where existence and uniqueness of solutions that continuously depend on the given data can be shown.

in 1965 by the French mathematician Jacques-Louis LIONS and the American mathematician Walter A. STRAUSS in their seminal work [108] where they showed 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'. Since then, many contributions have been made to nonlinear evolution equations; we will include the most recent ones in Section 1.2.

This is the point of departure for the present work, which addresses the existence of strong solutions to the abstract CAUCHY problem for nonlinear evolution inclusions of first and second order of the type

$$\partial \Psi_{u(t)}(u'(t)) + \partial \mathcal{E}_t(u(t)) \ni B(t, u(t)), \quad t \in (0,T), \tag{1.0.6}$$

and

$$u''(t) + \partial \Psi_{u(t)}(u'(t)) + \partial \mathcal{E}_t(u(t)) + B(t, u(t), u'(t)) \ni f(t), \quad t \in (0,T), \tag{1.0.7}$$

which has not been studied before in the generality presented here, where Ψ, $\mathcal{E}$, and B are called the dissipation potential, the energy functional, and the perturbation, respectively, which satisfy certain conditions.

1.1 Structure of the thesis

This thesis is organized as follows. It consists of two parts: the first part deals with evolution inclusions of first order, and the second part with evolution inclusions of second order.

In **Chapter 2**, we give an introduction to the theory of convex analysis for the analysis of nonsmooth functionals. We introduce the associated terminology and notation, and present the results required for proofs in later chapters. These results are sourced from existing literature, which are specified in the chapter, or proven by the author, if they were not found in the literature. Section 2.1 is devoted to the basic notions, such as the lower semicontinuity and the λ-convexity of a function. In Section 2.2, we introduce the notion of a subdifferential and characterize the subdifferential for λ-convex and for differentiable functionals. Then, in Section 2.3, we define the LEGENDRE–FENCHEL transformation of a function and represent a connection to its subdifferential that is essential for this work. Since the functionals we are working with do not exhibit any kind of differentiability, it is necessary in our existence proof to smooth, in an intermediate step, the functional acting on the first derivative of the sought solution. This is done by the p-MOREAU–YOSIDA regularization, which will be defined in Section 2.5 and for which we prove important properties in Theorem 2.5.2. Theorem 2.5.2 is independent of our main results presented from Chapter 3, 5, and 6, and has not been published. In Section 2.4, we define the Γ- and MOSCO-convergence for s sequence of functionals, and in Section 2.6, we introduce parameterized YOUNG measures.

Part I

Chapter 3 is devoted to the strong solvability of the CAUCHY problem for the doubly nonlinear evolution inclusion (1.0.6). In order to show the existence of solutions, we use a semi-implicit EULER-method and establish convergence of the approximation scheme which is formulated in Section 3.1. In Section 3.2, we collect the assumptions concerning the functionals $\mathcal{E}_t$ and Ψ_u and the operator B. After a discussion of the assumptions in Section 3.2.1, we present the main result in Theorem 3.2.3. In Sections 3.3 and 3.4, we show a discrete energy-dissipation inequality and derive from it a priori estimates for the interpolations. In Section 3.5, we show the compactness of the interpolations in suitable spaces and characterize the limit points by using parameterized YOUNG measures. We then prove the main result in Section 3.6, first by proving the existence of strong solutions to the regularized problem, and then by concluding the proof by passing to the limit as $\varepsilon \searrow 0$ and showing the existence strong solutions of the CAUCHY problem for (1.0.6), which fulfills an energy dissipation balance. The results of this chapter have been published in BACHO, EMMRICH & MIELKE [21] with stronger assumptions, which will be discussed more in detail in Section 3.2.

In **Chapter 4**, we show the existence of an initial-boundary value problem with nonlinear constraints as an application of the theorem provided in Chapter 3.

Part II

In **Chapter 5**, we consider the CAUCHY problem for the second-order evolution inclusions (1.0.7) which we refer to as *linearly damped inertial system*. Here, we discuss two cases. In the first case, we assume that $\partial\Psi$ is a linear, bounded, strongly positive, and self-adjoint operator, and in the second case, we allow a strongly continuous nonlinear perturbation of the linear part. In both cases, the operator $\partial\mathcal{E}_t$ is nonlinear and the subdifferential of a proper, sequentially weakly lower semicontinuous, and λ-convex functional $\mathcal{E}_t$, and the perturbation B is a fully nonlinear and strongly continuous operator acting on u and u'. Here, the functionals Ψ and $\mathcal{E}_t$ are defined on different spaces for which we assume not that either of the two spaces is continuously embedded in the other one. For both cases, we show the existence of strong solutions of the CAUCHY problem for (1.0.7), which fulfills an energy-dissipation inequality. The results of this chapter have been prepublished under stronger assumptions in BACHO [20]. The precise assumptions are presented in Section 5.1. After a discussion of the assumptions in Section 5.1.1, we present the main result in Theorem 5.1.4. The steps of the proof of the main result has the same structure as the proof of the main result in Chapter 3, and is based on showing the convergence of a semi-implicit time discretization of the inclusion (1.0.7). This is accomplished by first showing the solvability of the variational approximation scheme based on the time discretization in Section 5.2. In Section 5.3, a discrete energy dissipation inequality is shown and a priori estimates are derived. In Section 5.4, we show the compactness of the interpolations and pass then to the limit as the step size vanishes in Section 5.5.

In **Chapter 6**, we switch the properties of the dissipation potential and the energy functional, and allow the dissipation potential further dependence on the state u. More precisely, we show the existence of strong solutions through the main assumption that the leading or dominating part of $\partial\mathcal{E}_t$ is a linear, bounded, strongly positive, and self-adjoint operator and $\partial\Psi_u$ is the subdifferential of a proper, lower semicontinuous and convex operator Ψ_u of p-growth. The functionals Ψ_u and $\mathcal{E}_t$ again act on different spaces for which we again assume that neither of the two spaces is continuously embedded in the other one. The perturbation is a fully nonlinear and strongly continuous operator acting on u and u'. Under these assumptions (which will be made more precise in Section 6.1), we show the existence of a strong solution of the Cauchy problem for (1.0.7), which fulfills an energy-dissipation balance and will be presented in Theorem 6.1.4 in the same section. The proof of the aforementioned theorem is divided into the same steps as in Chapter 5. The results presented in this chapter are novel and have not been published before.

In **Chapter 7**, we then apply the theorems proved in Chapters 5 and 6 to some concrete examples to demonstrate the range of possible applications. In Sections 7.1 and 7.2, we consider differential inclusions, which fits into the framework of Chapter 5. In Section 7.3 and Section 7.4, we consider the equations of the martensitic transformation in shape-memory alloys and a viscous regularization of the Klein–Gordon equation. Finally, in Section 7.5, we consider a differential inclusion with nonlinear damping.

1.2 Literature review

Results on abstract evolution inclusion or equations of type (1.0.3) and (1.0.5) have been provided by several authors under various conditions and assumptions on the operators A and B, as well as the underlying spaces the operators are acting on. Here we give an overview of the most recent literature to nonlinear evolution equations of first and second order.

Evolution equations of first order

In the above mentioned work of Colli & Visintin [45], the authors work in their analytical framework with a Gelfand triple[9] where V is compactly embedded in H. Under the assumption that $\mathrm{dom}(A) = H$ and $\mathrm{dom}(B) \subset V$, and that either

[9]A triple (V, H, V^*) of vector spaces is called Gelfand triple if H is a Hilbert space and V is a reflexive Banach space, which has the dual space V^*, such that the following dense and continuous embeddings hold: $V \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} V^*$, where H has been identified with its dual space H^* via the Riesz isomorphism.

B is a subdifferential operator such that the potential and the operator A satisfy certain coercivity conditions or that A is a subdifferential operator and B is a Lipschitz continuous and strongly monotone operator, they showed the existence of a strong solution with $u \in \mathrm{H}^1(0,T;H) \cap \mathrm{L}^\infty(0,T;V)$ if $f \in \mathrm{L}^2(0,T;H)$ in the first case and a strong solution $u \in \mathrm{H}^1(0,T;V)$ if $f \in \mathrm{H}^1(0,T;V^*)$ in the second case to (1.0.3) in H and V^*, respectively. Similar results for operators of p-growth with $1 < p < +\infty$ have been obtained by Arai [14], Barbu[24], Senba [151], and Colli [44]. In order to obtain solutions, the aforementioned authors use regularization and approximation techniques to construct functions that approximate a solution. Some of the techniques include the Moreau–Yosida regularization and the Yosida approximation for the operators, see Section 2.5.

A more elegant approach has been made by Stefanelli by using the celebrated Brézis–Ekeland variational principle [26, 34, 36] in order to characterize and show the existence of strong solutions to the Cauchy problem

$$\begin{cases} \partial\psi(u'(t)) + \partial\phi(u(t)) \ni f(t) & \text{for a.e. } t \in (0,T), \\ u(0) = u_0, \end{cases} \tag{1.2.1}$$

for proper, lower semicontinuous and convex functionals $\phi, \psi : V \to (-\infty, +\infty]$ defined on a reflexive Banach space V with norm $\|\cdot\|_V$. The Brézis–Ekeland variational principle states that a function is a solution to the parabolic equation (1.2.1) if and only if it solves a minimization problem despite the equation in (1.2.1) not having a variational structure. Rewriting (1.2.1) into the form

$$\begin{cases} \xi(t) + \partial\phi(u(t)) \ni f(t) & \text{for a.e. } t \in (0,T), \\ \xi(t) \in \partial\psi(u'(t)) & \text{for a.e. } t \in (0,T), \\ u(0) = u_0, \end{cases} \tag{1.2.2}$$

Stefanelli [155] showed that the couple $(u,\xi) \in \mathrm{W}^{1,p}(0,T;V) \times \mathrm{L}^p(0,T;V)$ with $1 \le p < +\infty$ solves (1.2.2) if and only if (u,ξ) minimizes $I : \mathrm{W}^{1,p}(0,T;V) \times \mathrm{L}^p(0,T;V) \to [0,+\infty]$ with

$$\begin{aligned} I(v,\eta) = {} & \left(\int_0^T \left(\psi(v'(t)) + \psi^*(\eta(t)) - \langle \eta(t), v'(t)\rangle_{V^*\times V}\right) \mathrm{d}t + \phi(v(t)) - \phi(v(0)) \right)^+ \\ & + \left(\int_0^T \left(\phi(v(t)) + \phi^*(f(t)-\eta(t)) - \langle f(t)-\eta(t), v(t)\rangle_{V^*\times V}\right) \mathrm{d}t \right) \\ & + \|v(0) - u_0\|_V^2, \end{aligned}$$

where $\langle\cdot,\cdot\rangle_{V^*\times V}$ denotes the duality pairing between V and its dual space V^*, ϕ^* and ψ^* again the conjugate functionals of ϕ and ψ (see Section 2.3), respectively, and $x^+ := \max\{x, 0\}, x \in \mathbb{R}$. Furthermore, he showed that the assumption that ψ has p-growth with $p > 1$ and ϕ has compact sublevel sets in V is sufficient to obtain coercivity of the functional I with respect to a certain topology. He also established a result on Γ-convergence by giving sufficient conditions for a sequence of solutions of (1.2.2) to converge to another solution of (1.2.2). All the results presented previously rely heavily on the convexity of the functional ϕ or the maximal monotonicity of the operator B. Mielke, Rossi & Savaré overcame this problem by using the De Giorgi's energy-dissipation principle, which states that under suitable conditions solutions to the system (1.2.2) with $f = 0$ can be characterized as absolutely continuous functions satisfying the so-called energy-dissipation balance

$$\phi(u(t)) + \int_0^T (\psi(u'(t)) + \psi^*(-\xi(t)))\,\mathrm{d}t = \phi(u_0).$$

The idea of the existence result is based on a metric space formulation of gradient flows introduced by De Giorgi, Marino & Tosques in their pioneering work [54] where one replaces the Fréchet derivatives by suitable metric derivatives[10]. Based on the metric formulation, they showed first in [142] the existence of absolutely continuous curves with values in a separable metric space. The main assumptions are that ψ is proper, lower semicontinuous, convex, and has superlinear growth, and that $\mathcal{E}_t$ is lower semicontinuous, satisfies a chain rule, has compact sublevel sets, and its subdifferential satisfies a certain closedness condition. These results have been further generalized by the same authors in [122] when the metric space is a reflexive Banach space where they considered the Cauchy problem for the generalized gradient system of the form

$$\begin{cases} \partial\Psi_{u(t)}(u'(t)) + \mathrm{F}_t(u(t)) \ni 0 & \text{for a.e. } t \in (0,T), \\ \mathcal{E}_t(u(t)) \subset \mathrm{F}_t(u(t)) & \text{for a.e. } t \in (0,T), \end{cases}$$

by allowing a time-dependence of $\mathcal{E}_t$ and a state-dependence of Ψ_u. We extended this result to perturbed gradient systems by incorporating a non-variational and non-monotone perturbation B in form of (1.0.6) into the equation and by avoiding further a certain regularity assumption for Ψ_u which has been accomplished by regularization arguments, see Chapter 3 and Section 3.2 for the precise assumptions. However, the results do not include the case where Ψ has at most linear growth, which is strictly related to rate independent systems where Ψ is positively homogeneous of degree one, i.e. $\Psi(\alpha v) = |\alpha|\Psi(v)$ for all $\alpha \in \mathbb{R}$. For results on rate-independent systems, see Section 1.3.

Perturbed gradient systems have already been investigated by Brézis [32] and Ôtani [126, 127], Akagi [4] and Akagi & Melchionna [5]. In [32, 126, 127], the authors investigate the case when $A = I$ and $B = \partial\phi$ for a proper, lower semicontinuous, and convex functional ϕ on a separable Hilbert space where the operator B is perturbed by a Lipschitz or multivalued operator that satisfies certain growth and continuity conditions. The doubly nonlinear case has been studied by Akagi and Akagi & Melchionna. In [4], the author assumed that $\Psi = \Psi_t$ and $\mathcal{E}$ are both proper, lower semicontinuous and convex, and B is a multi-valued operator satisfying certain growth and continuity conditions in a Gelfand triple framework. For different growth conditions for B, he showed local and global existence results. In [5], the authors assumed in addition the Gâteaux differentiability of Ψ but allowing more non-convex functionals of the form $\mathcal{E} = \mathcal{E}_1 - \mathcal{E}_2$, where $\mathcal{E}_i, i = 1, 2$, are

[10] We refer the reader to Ambrosio Gigli & Savaré [10] for a detailed treatise of gradient flows in metric spaces. See also Ambrosio [8, 9] and the introduction of Chapter 3, where we elaborate more on the metric formulation.

again supposed to be convex. We remark that these kinds of non-convex functionals can also be treated in our framework, see [122, 142, 143]. The perturbation is supposed to be time-independent and continuous, which is also a special case in our setting. Furthermore, the author works on a GELFAND triple (V, H, V^*) which excludes the case where ϕ has p-growth with $p \in (1, 2)$. In addition, it has been assumed that V is an uniformly convex BANACH space, which in particular is reflexive. Thus, while the latter work is completely covered by our work, we do not include multi-valued perturbations and do not consider time-dependent functionals Ψ_t. Also, our work only focuses on the second type of abstract doubly nonlinear equations and only covers abstract doubly nonlinear equations of the first type (1.0.4) if A is FRÉCHET differentiable with an invertible derivative $\mathrm{D}A$, so that formally $(Au(t))' = [\mathrm{D}Au(t)]u'(t)$ and therefore $\Psi_u(v) = \frac{1}{2}\langle [\mathrm{D}Au]v, v\rangle$. For results on this case, we refer to GRANGE & MIGNOT [87], BAMBERGER [23], BARBU [25], DIBENEDETTO & SHOWALTER [56], MAITRE & WITOMSKI [111], AIZICOVICI & HOKKANEN [2, 3], MATAS & MERKER [113] and the references therein. For nonlinear equations with VOLTERRA operators, we refer to, e.g., GAJEWSKI, GRÖGER & ZACHARIAS [84, Kapitel V], GILARDI & STEFANELLI [85, 86], EIKMEIER & EMMRICH [66], EIKMEIER, EMMRICH & KREUSLER [67] and the references therein.

In the case of single-valued operators, EMMRICH & VALLET [78] investigate the CAUCHY problem for the equation of BARENBLATT-type

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

where A is a hemicontinuous, monotone and coercive operator and B a strongly continuous operator. The operator A is, in particular, maximal monotone (see, e.g., BARBU [26, Theorem 2.4, p. 36]) but not necessarily cyclical monotone and therefore not necessarily a subdifferential operator, or in this case, potential operator, see BRÉZIS [32, Chapter II, Section 7, p. 38]. The operator B is not supposed to satisfy any monotonicity assumption. For these types of equations, see also BAUZET & VALLET [29] and the references therein. For abstract evolution equations, we also refer to the monographs ROUBÍČEK [145, Part II], WLOKA [160, Chapter IV], and ZEIDLER [164, Chapter 30].

Evolution equations of second order

Results on abstract evolution equations of second order are in general much more delicate and difficult. The reason is that, roughly speaking, the equations possess the additional term $\partial_{tt}u$ describing the propagation of waves that, as opposed to parabolic equations, has a nonsmoothing effect in the time evolution for the solution u. As a consequence, much less existence results are available.

Evolution inclusions of second order of the form

$$u''(t) + A(t)u'(t) + B(t)u(t) \ni f(t), \quad t \in (0, T),$$

i.e., in the multivalued case, have been studied by ROSSI & THOMAS in [144] where $A(t) = A : V \to V^*$ is a linear, bounded, strongly positive and symmetric defined on the reflexive and separable BANACH space, and $B(t) = \partial\mathcal{E}_t$ is the subdifferential of a λ-convex functional with effective domain in a reflexive and separable Banach space U. In the framework of the Gelfand quintuplet

$$U \overset{d}{\hookrightarrow} V \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} V^* \overset{d}{\hookrightarrow} U^*$$

and under the assumption that $\mathcal{E}_t$ satisfies a chain rule and $\partial\mathcal{E}_t$ satisfies a closedness condition, they showed the existence of a strong solution. While this work is completely covered with the result of Chapter 5, we allow further a strongly continuous non-monotone and non-variational perturbation that depends on u and u' as well as a nonlinear monotone perturbation of A of variational type. We furthermore do not assume the rather restrictive assumption that $U \overset{d}{\hookrightarrow} V$. Furthermore, the strong closedness condition of $\partial\mathcal{E}_t$ assumed in [144] excludes the application to nonlinear elastodynamics where the operator satisfies a so-called Andrews–Ball type condition, see Section 7.3. In Emmrich & Šiška [74], the authors develop an abstract theory in the smooth setting with the application to nonlinear elastodynamics. They prove the existence of strong solutions for the case where $A : V_A \to V_A^*$ is linear, bounded, strongly positive and symmetric, and $B : V_B \to V_B^*$ is supposed to be demicontinuous and a bounded potential operator. In addition, B satisfies an Andrews–Ball-type condition, meaning that $(B + \lambda A) : V \to V^*$ is monotone where $V := V_A \cap V_B$ is densely and continuously embedded into the separable and reflexive Banach spaces, V_A and V_B, for which we assume not that either of the two spaces is continuously embedded in the other one. Since we allow a more general nonsmooth functional $\mathcal{E}$, this result is also covered by the main result presented in Chapter 5. The case where the operator A is nonlinear has also been discussed by several authors. Apart from the well-posedness result of Lions & Strauss [108], Emmrich & Thalhammer [75] showed the existence of solutions 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 $C(t)$ with the same assumptions on V_A and V_B as above. As mentioned before, the assumptions on A imply that $A + \kappa I$ is maximal monotone and therefore not necessarily a potential operator. Therefore, the result obtained in Chapter 6 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.

Doubly nonlinear evolution inclusions where the leading parts of A and B are both nonlinear and contain in the applications the same order of spatial derivatives, is unfortunately not treatable in our framework. The main difficulty that arises in showing the existence of global (weak, strong or classical) solutions is the identification of the weak limits, which arise after applying a discretization method, in both nonlinearities, A and B. If one of the operators is linear, then the identification is usually accomplished by using monotonicity, compactness or fixed-point arguments. However, in some concrete problems, this can been shown by exploiting the special structure of the operators. For example, Puhst [134] showed the existence of weak solutions under the assumption that the operators A and B are nonlocal operators. Friedman & Nečas [83] showed the existence of weak solutions under the assumptions that the operators are potential operators that are twice differentiable such that the Hessian matrices are uniformly positive definite and bounded. Bulíček, Málek & Rajagopal [38] and Bulíček, Kaplický & Steinhauer [37] showed the existence of weak solutions under the assumptions that the operators satisfy strong monotonicity, Lipschitz, and growth conditions, which has been shown to be classical solutions under stronger regularity conditions on the operators.

For further results on nonlinear evolution equations, we refer to Leray [105], Dionne [59], Emmrich & Thalhammer [76, 77], Emmrich, & Šiška [73] including stochastic perturbations, Emmrich,Šiška & Thalhammer [75] for a numerical analysis, Emmrich, Šiška & Wróblewska-Kamińska [79] and Ruf [146] for results on Orlicz spaces, and the monographs Lions [106], Lions & Magenes [107, Chapitre 3.8], Barbu [24, Chapter V], Wloka [160, Chapter V], Zeidler [164, Chapter 33], Roubíček [145, Chapter 11] and the references therein.

The list of literature presented in this section is not intended to be exhaustive.

1.3 Outlook

There are still many open questions concerning doubly nonlinear abstract evolution inclusions of first and second order with respect to their generalizations, and a corresponding solution concept to them. Some of these questions are directly related to our work and will be discussed here. The following list of problems is, of course, not intended to be exhaustive.

Non-reflexive Banach spaces

The assumption that the underlying spaces are reflexive Banach spaces excludes many spaces, including the function spaces $C(\Omega), \mathrm{L}^1(\Omega)$ and $\mathrm{L}^\infty(\Omega)$, in general Orlicz spaces, the space of functions with bounded variation, the space of Radon measures, etc., and therefore excludes many important applications. Therefore, it is interesting to consider problems on Banach spaces that are not reflexive. By employing the theory of semigroups, this has been accomplished by Hille [92] and Crandall & Liggett [49] where they show the existence and uniqueness of mild solutions to the parabolic equation (1.0.2) for unbounded, linear, and self-adjoint operators or nonlinear accretive operators. The result for nonlinear accretive operators can be extended to the case where A is perturbed by a locally Lipschitz continuous operator, see, e.g., Barbu [26, Theorem 4.8, p. 150,]. An important factor in the existence result of mild solutions is the fact that in the definition of a mild solution, the solution is not required to possess any vector differentiability. This is a problem if one asks for strong solutions, i.e., absolutely continuous functions with a certain regularity that fulfill the differential inclusion pointwise almost everywhere. The problem is based on the fact that absolutely continuous functions $u : [0, T] \to X$ with values in a Banach space do not possess the so-called Radon–Nikodým property[11]; these functions are in general not differentiable almost everywhere[12]. This problem has been overcome for gradient flows of type (1.2.1) by introducing a metric formulation of the gradient flow equation, where one replaces the derivative u' of an absolutely continuous function $u : [0,T] \to X$ by its metric derivative $|u'|$ defined by

$$|u'|(t) := \lim_{s\to t} \frac{d(u(t), u(s))}{|t-s|},$$

which always exists for almost every $t \in (0,T)$, see AMBROSIO et al. [10, Theorem 1.1.2, p. 24]. However, due to the lack of a linear structure of the underlying space, there is a need for an appropriate definition of a perturbed gradient flow in metric spaces.

Rate independent systems.

An essential condition to obtain the existence of strong solutions is the superlinearity of the dissipation potential Ψ_u and its convex conjugate Ψ_u^*. The superlinearity guarantees that the derivatives of the approximate solutions are equi-integrable, so that we obtain a solution which is absolutely continuous. This is no longer given if Ψ_u has at most linear growth. Nevertheless, this is an interesting case from a mathematical and physical point of view and leads to the notion of so-called rate-independent systems, which refers to systems where the dissipation potential is homogeneously positive of degree one, i.e., $\Psi_u(\lambda v) = |\lambda|\Psi(v)$ for all $\lambda > 0, v \in V$ which implies that $\partial\Psi(\lambda v) = \partial\Psi(v)$ for all $v \in V$. Therefore, the class of solutions to rate-independent systems is time scale invariant. Due to the lack of superlinearity, the analysis of rate-independent systems are completely different from the case studied here and therefore necessitates a different solution concept. Relying on the so-called energetic formulation, rate-independent systems have been extensively studied for the unperturbed case by MIELKE and coauthors, see, e.g., [115, 116, 120, 121, 123] and the references therein. In the energetic formulation, a curve $u : [0,T] \to V$ is called an energetic solution to a rate-independent system if it fulfills the global stability condition

$$\mathcal{E}_t(u(t)) \leq \Psi(u(t) - v) + \mathcal{E}_t(v) \quad \text{for all } v \in V,$$

and the energy balance

$$\mathrm{Var}_\Psi(u;[0,t]) + \mathcal{E}_t(u(t)) = \mathcal{E}_0(u(t)) + \int_0^t \partial_r \mathcal{E}_r(u(r))\,\mathrm{d}r \quad \text{for all } t \in [0,T],$$

[11] A BANACH space X does possess the RADON–NIKODÝM property if and only if every absolutely continuous function $u : [0,T] \to X$ is differentiable amost everywhere in which case there holds $u(t) - u(s) = \int_s^t u'(r)\,\mathrm{d}r$ for all $s,t \in [0,T]$. A sufficient condition for a BANACH space X to have the RADON–NIKODÝM property is the reflexivity of X or if X is separable and the dual space of another BANACH space, see p. 217 and pp. 61 in DIESTEL & UHL [58] for more sufficient and necessary conditions and for the definition of the RADON–NIKODÝM property, respectively.

[12] See, e.g., EMMRICH [71, Beispiel 7.1.21, p. 162] for an example of an abstract function with values in a non-reflexive BANACH space which is nowhere differentable.

where

$$\mathrm{Var}_\Psi(v;[a,b]) = \sup\{\sum_{k=1}^{N} \Psi(v(t_k) - v(t_{k-1})) : a = t_0 < \cdots < t_M = b\}$$

is the total variation of a function $v : [0,T] \to V$ on $[a,b] \subset [0,T]$ induced by Ψ. The question is whether the solution concepts can be modified in a mathematically and physically reasonable way to a perturbed problem. We refer the interested reader to Mielke & Roubíček [124] for a detailed treatise of rate-independent systems.

Periodicity of solutions.

An important question to address is the periodicity of solutions to abstract evolution inclusions of first and second order. In the generality of our setting, there are no such results known. However, under stronger assumptions, there are results available for doubly nonlinear evolution inclusions of first order. Akagi & Stefanelli [6] have shown the existence of periodic solutions to the doubly nonlinear case where A is a maximal monotone operator of at most linear growth and B is the subdifferential of a proper, lower semicontinuous and convex functional. Within the class of subdifferential operators, this has very recently been extended by Koike, Ôtani & Uchida [103] to the case where A is the Gâteaux derivative and B is the subdifferential of proper, lower semicontinuous and convex functionals of polynomial growth defined on an uniformly convex Banach space. To the authors' best knowledge, there are no more results available for doubly nonlinear equations of the first type (1.0.3). In contrast, the existence of periodic solutions to the second type of equation (1.0.4) has been vigorously studied by many authors, see, e.g., [3, 94, 100, 101] and the references therein. For evolution equations of second order, the existence of solutions has been shown in Gajewski et al. [84] when the operator A is radially continuous, monotone and coercive and B is a linear, bounded, positive, and self-adjoint operator, and both operators are defined on a Hilbert space.

Chapter 2

An Introduction to Convex Analysis

In this preliminary chapter, we will introduce some useful tools from the theory of convex analysis, and try to highlight their general importance and their relevance in regard to the present work.

The theory of convex analysis deals in essence with the study of convex functions and convex sets, and has numerous applications in various areas, e.g., convex optimization, economics, mechanics and numerical analysis. The application in convex optimization was motivated by the seminal results obtained in linear programming, where minimization problems of linear functionals over polytopes, which are expressed by linear constraints, are studied, in the hope of obtaining similar results for nonlinear functionals subject to nonlinear constraints. As a result, the duality principle from linear programming was extended to nonlinear problems and led to the notion of the LEGENDRE–FENCHEL transformation and the subdifferentiability, which will be defined in Section 2.3 and 2.2, respectively. Besides, we will introduce the MOREAU–YOSIDA regularization in Section 2.5, the MOSCO-convergence in Section 2.4, and parameterized YOUNG measures in Section 2.6.

This chapter is mainly based on the excellent and self-contained monographs of EKELAND & TEMAM [69], ROCKAFELLAR [139], BARBU & PRECUPANU [27], BARBU [26]. More specific literature will be mentioned in the related sections.

2.1 Preliminaries and notation

The main objects of our study are defined on BANACH spaces. However, many properties and tools we present in this chapter are also available on more general spaces. Therefore, if not otherwise specified, we consider a real BANACH space X equipped with the norm $\|\cdot\|$ and we denote with X^* its topological dual space equipped with the norm $\|\cdot\|_*$. The duality pairing between X and X^* is denoted by $\langle\cdot,\cdot\rangle$. Furthermore, we denote with $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ the *extended real line*. On $\overline{\mathbb{R}}$, we define a total order by setting $-\infty \leq a \leq +\infty$ for all $a \in \overline{\mathbb{R}}$ so that the set can be equipped with the order topology. We extend the arithmetic operations of $\mathbb{R}$

to $\overline{\mathbb{R}}$ by setting

$$\begin{aligned}
&a + \infty = +\infty + a = +\infty, \quad a \neq -\infty,\\
&a - \infty = -\infty + a = -\infty, \quad a \neq +\infty,\\
&a(\pm\infty) = (\pm\infty)a = \pm\infty, \quad a \in (0, \infty],\\
&a(\pm\infty) = (\pm\infty)a = \mp\infty, \quad a \in [-\infty, 0),\\
&\frac{a}{\pm\infty} = 0, \qquad a \in \mathbb{R},\\
&\frac{\pm\infty}{a} = \pm\infty, \quad a \in (0, +\infty),\\
&\frac{\pm\infty}{a} = \mp\infty, \quad a \in (-\infty, 0).
\end{aligned}$$

Many real world problems are optimization problems of the form

$$\inf_{v \in C} \tilde{f}(v)$$

where the objective is to find a value $v \in C$ in an arbitrary set $C \subset X$ which minimizes the real-valued functional $\tilde{f} : C \to \mathbb{R}$. The functional $\tilde{f}$ is often not defined outside of the set C. Nevertheless, one can introduce the *extended functional* $f : X \to \overline{\mathbb{R}}$ which takes values in the extended real line $\overline{\mathbb{R}}$ by setting

$$f(v) = \begin{cases} \tilde{f}(v) & \text{if } v \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

Introducing extended functionals, not only simplifies the analysis from a notational point of view, but has also the advantage that properties of sets can be translated to properties of the functionals. If we consider, e.g., the *indicator function* $\imath_C : C \to \overline{\mathbb{R}}$ of a convex and closed set $C \in X$, defined by

$$\imath_C(v) = \begin{cases} 0 & \text{if } v \in C, \\ +\infty & \text{otherweise,} \end{cases}$$

then, there holds $v \in C$ if and only if $\imath_C(v) < +\infty$. Furthermore, it can be shown that C is closed or convex if and only if $\imath_C$ is a lower semicontinuous or convex function, respectively. As a consequence, one can focus on optimization problems with functionals that are defined on the whole space such as

$$\inf_{v \in X} f(v). \tag{2.1.1}$$

A necessary condition for the solvability[1] of such a minimization problem is indeed that the set C is non-empty, or, in other words, the extended functional f is not identically $+\infty$, i.e., if the *effective domain*

$$\mathrm{dom}(f) := \{v \in X : f(v) < +\infty\}$$

[1]The minimization problem (2.1.1) is called solvable if there exists at least one element in X that minimizes f and where f is finite.

of f is non-empty. We call the functional f *proper* if it has a non-empty effective domain and if it takes nowhere the value $-\infty$. In what follows, we constantly assume that $f : X \to (-\infty, +\infty]$ is an extended and proper functional.

Further properties of f that are indispensable for guaranteeing the solvability of (2.1.1), are indeed the sequential (weak) lower semicontinuity, convexity and coercivity, which are defined for extended functionals in the same manner as for real valued functionals by respecting the arithmetic operations on $\overline{\mathbb{R}}$.

Definition 2.1.1 *Let (X, τ) be a topological space. The functional $f : X \to (-\infty, +\infty]$ is called lower semicontinuous in $u \in X$ if there holds*

$$f(u) \leq \liminf_{v \to u} f(v).$$

The functional f is called sequentially lower semicontinuous in $u \in X$ if for all sequences $(u_n)_{n\in\mathbb{N}} \subset X$ with $u_n \to u$ as $n \to \infty$ there holds

$$f(u) \leq \liminf_{n \to \infty} f(u_n).$$

Finally, the functional f is called lower semicontinuous or sequentially lower semicontinuous if it is lower semicontinuous or sequentially semicontinuous at every point. If $\tau = \sigma(X, X^*)$ is the weak topology, we say f is sequentially weakly lower semicontinuous or weakly lower semicontinuous.

Since we work with functionals that can take the value $+\infty$, it is beneficial to give equivalent characterizations of the lower semicontinuity in terms of the *epigraph* and the *sublevel sets* of f, which is very useful in practice.

Lemma 2.1.2 *Let (X, τ) be a topological space and $f : X \to (-\infty, +\infty]$ be a proper function. Then, the following assertions are equivalent:*

i) *The functional f is (sequentially) lower semicontinuous.*

ii) *For all $\gamma \in \mathbb{R}$, the sublevel set $\{v \in V : f(v) \leq \gamma\}$ is (sequentially) closed in V.*

iii) *The epigraph of f, defined by*

$$epi\ f := \{(v, \gamma) \in V \times \mathbb{R} : f(v) \leq \gamma\},$$

is (sequentially) closed in $V \times \mathbb{R}$.

Proof. This is proven in Dixmier [61, Theorem 7.4.11, p. 79]. □

Since on a metric space, the sequential lower semicontinuity and lower semicontinuity coincide, we will not distinguish between both terms. Furthermore, it is readily seen that weak lower semicontinuity implies lower semicontinuity. The converse holds true for convex functions, see, e.g., Brézis [35, Corollary 3.8, p. 61]. However, this, in general, does not hold true for λ-convex functionals.

Definition 2.1.3 *Let $\lambda \in \mathbb{R}$. Then, the functional $f : X \to (-\infty, +\infty]$ is called λ-convex if for all $u, v \in X$ and $t \in (0, 1)$ there holds*

$$f(tv + (1-t)u) \leq tf(v) + (1-t)f(u) + \lambda t(1-t)\|v - u\|^2. \tag{2.1.2}$$

The functional f is called convex if it is 0-convex and strictly convex if it is 0-convex and the inequality (2.1.2) is strict for all $u \neq v$.

Remark 2.1.4 If X is a HILBERT space, the λ-convexity of f is equivalent to the convexity of the functional $f + \lambda\| \cdot \|^2$ when the norm $\| \cdot \|$ is induced by the inner product on X.

2.2 Subdifferential calculus

From the theory of calculus of variations, it is well-known that solutions (in particular stationary solutions) to a large class of partial differential equations correspond, by the variation principle, to critical or stationary points of functionals, which are also called *energy functionals*. Critical points of a functional are those points where the (GÂTEAUX) derivative of the functional is zero. For example, if $\bar{v} \in X$ solves the minimization problem (2.1.1), and the functional f is GÂTEAUX differentiable in $\bar{v} \in X$, then by FERMAT's theorem, the point $\bar{v} \in X$ is a critical point of f, i.e.,

$$\mathrm{D}_G f(\bar{v}) = 0, \tag{2.2.1}$$

where D_G denotes the GÂTEAUX derivative of f. The equation (2.2.1) is also called EULER–LAGRANGE equation associated to f.

Even if we deal with instationary (time-dependent) problems, we will encounter minimization problems of the form (2.1.1) after discretizing the evolution inclusions (1.0.6) and (1.0.7) in time and solving the discretized inclusions, see Sections 3.1, 5.2, and 6.2. However, the functionals we deal with are, in general, not GÂTEAUX differentiable. Therefore, we need a generalization of FERMAT's theorem for a non-differentiable functional f, which in fact is given by the (FRÉCHET) subdifferential of f. The (FRÉCHET) subdifferential or subderivative of f is a generalized notion of derivative, and is, unlike the weak derivative, a locally defined object, and, in general, a multi-valued map from X to X^*.

Definition 2.2.1 (**Fréchet subdifferential**) *Let $f : X \to \overline{\mathbb{R}}$ be proper and $u \in \mathrm{dom}(f)$. Then, the FRÉCHET subdifferential $\partial f : X \rightrightarrows X^*$ of f at the point u is defined by the set*

$$\partial f(u) := \left\{ w \in X^* : \liminf_{v \to u} \frac{f(v) - f(u) - \langle w, v - u \rangle}{\|v - u\|} \geq 0 \right\} \tag{2.2.2}$$

and the elements of ∂f are called subgradients. Furthermore, the domain of ∂f is defined by

$$\mathrm{dom}(\partial f) := \{ u \in \mathrm{dom}(f) : \partial f(u) \neq \emptyset \}.$$

Finally, f is called subdifferentiable at the point $u \in \mathrm{dom}(f)$ if $u \in \mathrm{dom}(\partial f)$.

We refer to the FRÉCHET subdifferential simply as subdifferential. If we want to highlight that the subdifferential of f has been taken on the space X, we write $\partial_X f$. The reason for that is that the subdifferential always depends on the topology of the underlying space. However, we can always extend the functional f to a larger space which contains the space X by setting the value to $+\infty$ outside its domain so that specifying the subdifferential is in certain cases useful. We note that endowing the space X with an equivalent norm does not change the set (2.2.2). One can also easily check that the subdifferential $\partial f(u)$ is a closed and convex set for all $u \in \operatorname{dom}(\partial f)$. Moreover, from the definition of subdifferentiability, it is readily seen that Fréchet differentiable functionals are in particular subdifferentiable, and that the subdifferential becomes a singleton with the Fréchet derivative as a single value. Therefore, the subdifferential is indeed a generalized notion of differentiability. Similarly, for a Gâteaux-differentiable and convex function, the subdifferential contains only the Gâteaux derivative, which is stated in Lemma 2.2.7 below.

The following lemma gives a characterization of the subdifferential of a λ-convex functional, which is very useful in practice. The same characterization for convex functionals is often used as a definition of the subdifferential. The lemma gives also a sufficient condition for the graph $\operatorname{Gr}(\partial f) := \{(u, \partial f(u)) \subset X \times X^* : u \in X\}$ of ∂f to be strongly-weakly closed.

Lemma 2.2.2 *Let $f : X \to (-\infty, +\infty]$ be subdifferentiable in $u \in \operatorname{dom}(\partial f)$. Then, the following assertions hold:*

i) Let f be λ-convex with $\lambda \in \mathbb{R}$. Then, $\xi \in \partial f(u)$ if and only if

$$f(u) - f(v) \leq \langle \xi, u - v\rangle + \lambda\|u - v\|^2 \quad \text{for all } v \in X. \tag{2.2.3}$$

If f is lower semicontinuous, then $\operatorname{Gr}(\partial f)$ is strongly-weakly closed.

ii) Let f be Gâteaux differentiable on a convex set $\mathcal{A} \subset X$. Then, f is convex over $\mathcal{A}$ if and only if

$$f(u) - f(v) \leq \langle f'(u), u - v\rangle \quad \text{for all } u, v \in \mathcal{A}.$$

Proof. Ad *i*). Let $\xi \in \partial f(u)$. Since the inequality (2.2.3) is trivially fulfilled for all $v \in X \setminus \operatorname{dom}(f)$, it is sufficient to show (2.2.3) for all $v \in \operatorname{dom}(f)$. Therefore, let $v \in \operatorname{dom}(f)$. Since the inequality (2.2.3) for $v = u$ is obviously fulfilled, we assume $v \neq u$. Then, by definition

$$\begin{aligned}
0 &\leq \liminf_{\tilde{v} \to u} \frac{f(\tilde{v}) - f(u) - \langle \xi, \tilde{v} - u\rangle}{\|\tilde{v} - u\|} \\
&\leq \liminf_{t \to 0^+} \frac{f(u + t(v-u)) - f(u) - \langle \xi, t(v-u)\rangle}{\|t(v-u)\|} \\
&\leq \liminf_{t \to 0^+} \frac{tf(v) + (1-t)f(u) + t(1-t)\lambda\|u-v\|^2 - f(u) - \langle \xi, t(v-u)\rangle}{\|t(v-u)\|} \\
&= \frac{f(v) - f(u) + \lambda\|u-v\|^2 - \langle \xi, v - u\rangle}{\|v-u\|} \quad \text{for all } v \in V,
\end{aligned}$$

where we have used the λ-convexity of f. The converse is clearly fulfilled. Now, let $u_n \to u$ in V and $\xi_n \rightharpoonup \xi$ in V^* as $n \to \infty$ be convergent sequences with $\xi_n \in \partial f(u_n)$ for all $n \in \mathbb{N}$. Then, the lower semicontinuity of f and the characterization (2.2.3) yields

$$\begin{aligned}
f(u) - f(v) &\leq \liminf_{n \to \infty} (f(u_n) - f(v)) \\
&\leq \liminf_{n \to \infty} \left(\langle \xi_n, u_n - v\rangle + \lambda\|u_n - v\|^2\right) \\
&= \langle \xi, u - v\rangle + \lambda\|u - v\|^2 \quad \text{for all } v \in X,
\end{aligned}$$

whence $\xi \in \partial f(u)$.
Ad *ii*). This follows from Ekeland & Temam [69, Proposition 5.3 & 5.4]. $\square$

Remark 2.2.3 Let $u, v \in \operatorname{dom}(\partial f)$. Then, the characterization (2.2.3) immediately implies

$$-2\lambda\|u-v\|^2 \leq \langle \xi - \eta, v - u\rangle \quad \text{for all } \xi \in \partial f(u), \eta \in \partial f(v).$$

The operator ∂f is called *strongly monotone* if $\lambda < 0$ and *monotone* if $\lambda = 0$. This definition indeed coincides with the definition of strong monotonicity and monotonicity for single valued operators.

Hence, using the subdifferential, we see that $\bar{v}$ being a global or local minimizer of f implies

$$0 \in \partial f(\bar{v}),$$

and the reverse holds true when f is convex. In the next example, we see a subdifferentiable function that admits a global minimizer where the function is not Fréchet differentiable. We also see an example of a function that is not subdifferentiable.

Example 2.2.4 Let $h : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ with

$$h(x) = \begin{cases} x & \text{if } x < 0, \\ \frac{1}{2}x^2 & \text{otherwise,} \end{cases} \quad \text{and} \quad g(x) = -|x|, \quad x \in \mathbb{R},$$

be given. Since the functions h and g are differentiable on $\mathbb{R}\backslash\{0\}$, there holds $\partial h(x) = \{h'(x)\}$ and $\partial g(x) = \{g'(x)\}$ for all $\mathbb{R}\backslash\{0\}$. A simple calculation shows that while the subdifferential of h at $x = 0$ is given by the closed interval $[-1, 0]$, the subdifferential of g at $x = 0$ is empty. Hence, the subdifferential of h and g are given by

$$\partial h(x) = \begin{cases} -1, & \text{if } x \in (-\infty, 0), \\ [-1, 0], & \text{if } x = 0, \\ x, & \text{if } x \in (0, +\infty) \end{cases}, \qquad \partial g(x) = \begin{cases} 1, & \text{if } x \in (-\infty, 0), \\ \emptyset, & \text{if } x = 0, \\ -1, & \text{if } x \in (0, +\infty), \end{cases}$$

which is illustrated in figure 2.1.

As Figure 2.1 nicely illustrates, the subdifferential of the convex functional f in the point $x = 0$ contains all subgradients which are tangential to the graph at the point $x = 0$. In particular, there holds $\xi = 0 \in \partial f(0)$ and therefore $x = 0$ is a global minimizer of f.

An important question is whether the subdifferential operator is additive, i.e., the subdifferential of the sum of two functionals equals the sum of the subdifferential of the individual functionals. In general, this is not true. Choose, e.g., $f_1 = f$ and $f_2 = -f$ with f being any functional which is not subdifferentiable. Clearly, the sum of the functionals is subdifferentiable, but does not equal the sum of their subdifferentials.

The following lemmas give sufficient conditions so that subdifferential operator is indeed additive.

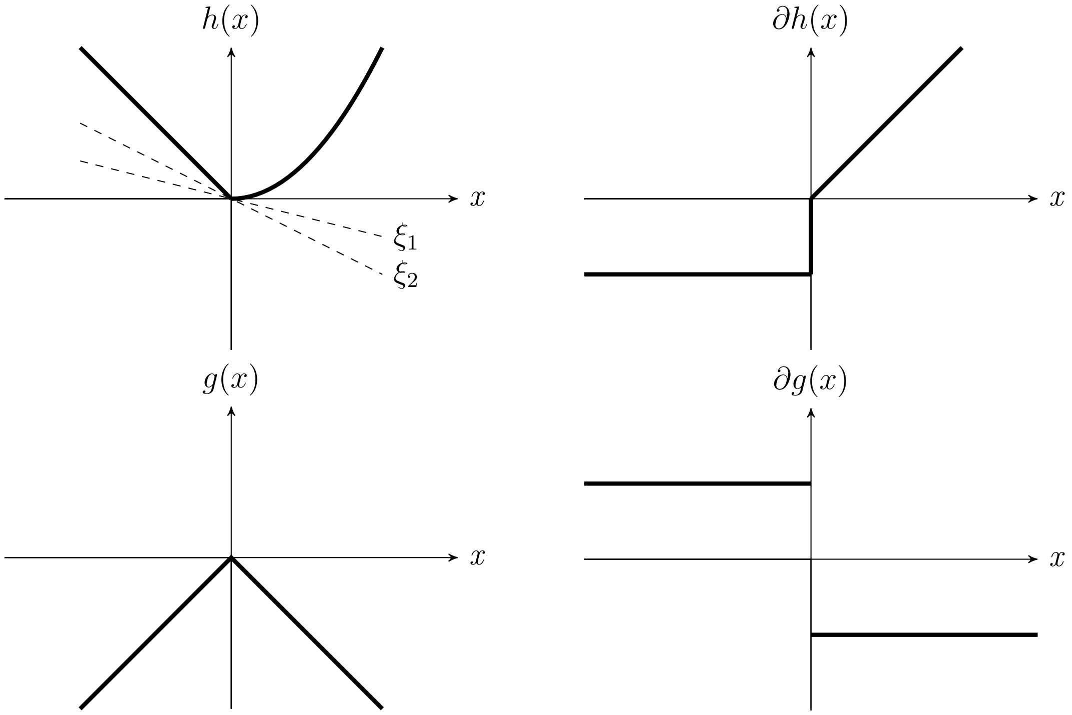

Figure 2.1: The figure shows the graph of the functions h and g, and their subdifferential ∂h and ∂g, respectively. It also shows the subgradients $\xi_1 = -\frac{1}{4}$ and $\xi_2 = -\frac{1}{2}$ of h at the point $x = 0$.

Lemma 2.2.5 *Let $f : X \to (-\infty, +\infty]$ be given by $f = f_1 + f_2$, where $f_1 : X \to (-\infty, +\infty]$ is subdifferentiable and $f_2 : X \to \mathbb{R}$ is* FRÉCHET *differentiable in $u \in \mathrm{dom}(f_1)$. Then, f is subdifferentiable in u and the subdifferential is given by*

$$\begin{aligned} \partial f(u) &= \partial f_1(u) + Df_2(u) \\ &= \{\xi + Df_2(u) : \xi \in \partial f_1(u)\}, \end{aligned}$$

where $Df_2(u)$ is the FRÉCHET *derivative of f_2 in u.*

Proof. This immediately follows from the definition of the subdifferential. □

Combining Lemma 2.2.2 and Lemma 2.2.5, we obtain

Corollary 2.2.6 *Under the assumptions of Lemma 2.2.5, let $f_1 : X \to (-\infty, +\infty]$ be convex. Furthermore, let $\mathrm{dom}(\partial f_1) \neq \emptyset$ and f_2 be* FRÉCHET *differentiable at the point $u \in D(\partial f_1)$. Then, $\xi \in \partial f(u)$ if and only if*

$$f_1(u) + \langle \xi - Df_2(u), v - u \rangle \leq f_1(v) \quad \textit{for all } v \in X.$$

If f_2 is convex, then u is a global minimizer of f if and only if $-Df_2(u) \in \partial f_1(u)$, i.e.,

$$f_1(u) + \langle -Df_2(u), v - u \rangle \leq f_1(v) \quad \textit{for all } v \in X.$$

The previous results deal with the case where at least one functional is differentiable and do therefore not answer the question of the additivity of the subdifferential operator when both functionals are non-differentiable. For general finite valued functionals $f_1, f_2 : X \to \mathbb{R}$, there holds

$$\partial\left(f_1 + f_2\right)(v) \supset \partial f_1(v) + \partial f_2(v),$$

which immediately follows from the definition. However, this inclusion is not useful since the right-hand side might be empty, while the left-hand side is non-empty. The following lemma gives a satisfying answer to that question.

Lemma 2.2.7 (Variational sum rule) *Let $f_1 : X \to (-\infty, +\infty]$ and $f_2 : X \to (-\infty, +\infty]$ be proper, lower semicontinuous, and convex. Furthermore, assume that there exists a point $\tilde{u} \in \operatorname{dom}(f_1) \cap \operatorname{dom}(f_2)$ where f_2 is continuous. Then, there holds*

$$\partial(f_1 + f_2)(v) = \partial f_1(v) + \partial f_2(v) \quad \textit{for all } v \in X.$$

If, in addition, f_2 is GÂTEAUX differentiable on V with GÂTEAUX derivative $D_G f_2$, then there holds $\partial f_2(v) = \{D_G f_2(v)\}$ and we obtain

$$\partial(f_1 + f_2)(v) = \partial f_1(v) + D_G f_2(v) \quad \textit{for all } v \in X.$$

Proof. This has been proven in EKELAND & TEMAM [69, Proposition 5.3. & 5.6]. □

With the variational sum rule, we are able to decompose subgradients of $f_1 + f_2$ in terms of the subgradients of f_1 and f_2. Apart from that, we are also interested in a special chain rule for the subdifferential of two composite functions $\Lambda : X \to Y$ and $f : Y \to (-\infty, +\infty]$, which are defined on BANACH spaces X and Y.

Lemma 2.2.8 *Let $\Lambda : X \to Y$ be a linear, bounded operator and $f : Y \to (-\infty, +\infty]$ be a proper, lower semicontinuous, and convex functional. If there exists a point $\Lambda\tilde{u} \in Y$ with $\tilde{u} \in X$, where f is finite and continuous, then for all $u \in X$, there holds*

$$\partial(f \circ \Lambda)(u) = \Lambda^* \partial f(\Lambda u) \quad \textit{for all } u \in X,$$

where $\Lambda^* : Y^* \to X^*$ denotes the adjoint operator of Λ.

Proof. This has been proven in EKELAND & TEMAM [69, Proposition 5.7]. □

For the operator Λ, we have in particular in mind the gradient operator ∇ which has as adjoint the divergence operator div, see Section 7.1.

2.3 Legendre–Fenchel transformation

If we take a closer look into the characterization (2.2.3) of the subdifferential of a proper, lower semicontinuous, and convex functional, we find in particular

$$\langle \xi, u \rangle \geq \sup_{v \in X} \{\langle \xi, v \rangle - f(v)\} + f(u). \tag{2.3.1}$$

We define the LEGENDRE–FENCHEL transformation $f^* : X^* \to \overline{\mathbb{R}}$ of f by

$$f^*(\eta) := \sup_{v \in X} \{\langle \eta, v\rangle - f(v)\}, \quad \eta \in X^*,$$

which is also called the *convex conjugate*, FENCHEL *conjugate*, or simply, *conjugate* of f. We can then formulate 2.3.1 in terms of f and its convex conjugate by the inequality

$$\langle \xi, u\rangle \geq f(u) + f^*(\xi). \tag{2.3.2}$$

Hence by Lemma 2.2.2, the latter inequality holds for $u \in V$ and $\xi \in V^*$ if and only if ξ is the subgradient of f in u, i.e., $\xi \in \partial f(u)$. We note that by the FENCHEL–YOUNG inequality

$$\langle \eta, v\rangle \leq f(v) + f^*(\eta) \quad \text{for all } u \in X, \xi \in X^*,$$

which, by definition, is always fulfilled, we can replace the inequality (2.3.2) with an equality.

We found another characterization of the subdifferential in terms of the conjugate function. This, among others, is stated in the following lemma.

Lemma 2.3.1 *Let $f : X \to (-\infty, +\infty]$ be a proper, lower semicontinuous, and convex functional and let $f^* : X^* \to (-\infty, +\infty]$ be the conjugate of f. Then, for all $(u, \xi) \in X \times X^*$, the following assertions are equivalent:*

i) $\xi \in \partial f(u)$ *in* X^*;

ii) $u \in \partial f^*(\xi)$ *in* X;

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

Proof. This has been proven in EKELAND & TEMAM in [69, Proposition 5.1 & Corollary 5.2]. □

The preceding lemma reveals a deep relationship between f and f^*, and is crucial in the existence result presented in the following chapters. Hence, it is useful to study the properties of the LEGENDRE–FENCHEL transformation.

Lemma 2.3.2 *Let $f : X \to (-\infty, +\infty]$ be a proper, lower semicontinuous, and convex function. Then, f is continuous over the interior of its effective domain and the LEGENDRE–FENCHEL transformation f^* is proper, lower semicontinuous, and convex. If $f \geq 0$ on X and $f(0) = 0$, then $f^* \geq 0$ on X^* and $f^*(0) = 0$. Furthermore, if X is reflexive, then there holds $f^{**} = (f^*)^* = f$. Finally, for general functions $h, g : X \to \overline{\mathbb{R}}$ with $h \leq g$ on X, there holds $g^* \leq h^*$ on X^*.*

Proof. The fact that f is continuous over the interior of its effective domain, and f^* is proper, lower semicontinuous, and convex, follows from Corollary 2.5, pp. 13 and the discussion in Section 4 in EKELAND & TEMAM [69]. If $f \geq 0$ on X and $f(0) = 0$, then

$$f^*(\xi) = \sup_{v \in V}\{\langle \xi, v\rangle - f(v)\} \geq \langle \xi, 0\rangle - f(0) = 0$$

and

$$f^*(0) = \sup_{v \in V}\{-f(v)\} = 0.$$

If X is reflexive then $f^{**} = (f^*)^* = f$ follows from EKELAND & TEMAM [69, Proposition 4.1, p. 18]. The last assertion immediately follows from the definition of the LEGENDRE–FENCHEL transformation. □

In order to provide a better understanding of Lemma 2.3.1, we consider the following examples.

Example 2.3.3 Let $p \in [1, +\infty)$ and $f : \mathbb{R} \to [0, +\infty]$ be defined by $f(x) = \frac{1}{p}|x|^p, x \in \mathbb{R}$. We want to calculate the conjugate $f^* : \mathbb{R} \to [0, +\infty]$. To do so, we distinguish the cases $p = 1$ and $p > 1$. First, let $p = 1$. We note that by Lemma 2.3.2, there holds $f^* \geq 0$ on $\mathbb{R}$. Then, for $y \in [-1, 1]$, we obtain

$$f^*(y) = \sup_{x \in \mathbb{R}}\{xy - |x|\} \leq 0$$

whence $f^*(y) = 0$. For $y \in \mathbb{R}$ with $|y| > 1$, there holds

$$\begin{aligned} f^*(y) &= \sup_{x \in \mathbb{R}}\{xy - |x|\} \\ &\geq \sup_{x \in \mathbb{R}}\{(|y| - 1)|x|\} = +\infty \end{aligned}$$

Therefore, the conjugate is given by the indicator function, i.e., $f^* = \imath_{[-1,1]}$. Now, let $p > 1$. Then, by the differential calculus, $x \in \mathbb{R}$ maximizes the function $\mathbb{R} \ni \tilde{x} \mapsto \tilde{x}y - \frac{1}{p}|\tilde{x}|^p$ for $y \in \mathbb{R}$ if and only if $y = x^{p-2}x$. Hence, the conjugate is given by $f^*(y) = \frac{1}{p^*}|y|^{p^*}, y \in \mathbb{R}$, where $p^* = p/(p-1)$ is the conjugate exponent. For $p \in (1, \infty)$, the FENCHEL–YOUNG inequality reads

$$xy \leq \frac{1}{p}|x|^p + \frac{1}{p^*}|y|^{p^*} \quad \text{for all } x, y \in \mathbb{R}. \tag{2.3.3}$$

which, in fact, is YOUNG's inequality. Lemma 2.3.1 gives now an optimality criteria for YOUNG's inequality: equality holds in 2.3.3 if and only if $y = x^{p-2}x$.

The conjugate functional on an infinite-dimensional space is, in general, difficult to calculate explicitly. However, if the functional is radial, i.e., if it depends only on the length of a vector, then we can reduce the calculation of the conjugate functional to the one-dimensional case, as it is shown in the following example.

Example 2.3.4 Let $f : \mathbb{R} \to [0, +\infty]$ be a proper and even function. We define $F : X \to [0, +\infty]$ by

$$F(u) := f(\|u\|), \quad u \in X.$$

Then, the conjugate $F^* : X^* \to (-\infty, +\infty]$ of F is given by

$$\begin{aligned} F^*(\xi) &= \sup_{v \in X}\{\langle \xi, v\rangle - F(v)\} \\ &= \sup_{r \geq 0} \sup_{v \in X,\, \|v\| = r}\{\langle \xi, v\rangle - f(\|v\|)\} \\ &= \sup_{r \geq 0}\{r\|\xi\|_* - f(r)\} \\ &= f^*(\|\xi\|_*) \quad \text{for all } \xi \in X. \end{aligned}$$

We obtain for, e.g., $f(x) = \frac{1}{p}|x|^p$ with $p \in [1, +\infty)$ the functional $F(v) = \frac{1}{p}\|v\|^p$ whose conjugate is given by $F^*(\xi) = \frac{1}{p^*}\|\xi\|_*^{p^*}$ if $p > 1$ and $F^*(\xi) = \imath_{\overline{B}_{X^*}(0,1)}(\xi)$ if $p = 1$, where $\overline{B}_{X^*}(0,1)$ denotes the closed unit ball in X^*. Here, Lemma 2.3.1 gives a characterization of the subdifferential of F, the so-called p-duality map denoted by F_X^p, see Section 2.5 for more details. Hence, $\xi \in \partial F(u) = F_X^p(u), u \in \mathrm{dom}(\partial F)$ if and only if

$$\langle \xi, u \rangle = \frac{1}{p}\|u\|^p + \frac{1}{p^*}\|\xi\|_*^{p^*}$$

which by YOUNG's inequality holds true if and only if $\langle \xi, u \rangle = \|u\|\|\xi\|_*$ and $\|u\|^p = \|\xi\|_*^{p^*}$.

Thus, we obtain a real-valued formula that entirely describes the relation between ξ and $\partial f(u)$ in an infinite-dimensional vector space. In fact, this will allow us to reformulate the generalized gradient flow equation by a real-valued equation, see the introduction Chapter 3, where we elaborate more on this and stress why this is crucial for our approach.

We may also ask, how the conjugate of the sum of two functionals can explicitly be expressed in terms of conjugate of the individual functionals.

Lemma 2.3.5 *Let $(X, \|\cdot\|)$ be a BANACH space, $f_1 : X \to (-\infty, +\infty]$ and $f_2 : X \to (-\infty, +\infty]$ be proper, lower semicontinuous, and convex functionals such that*

$$\bigcap_{\lambda \geq 0} \lambda\left(\mathrm{dom}(f_1) - \mathrm{dom}(f_2)\right) \qquad \textit{is a closed vector space,}$$

where $A - B := \{a - b : a \in A, b \in B\}$ for two sets $A, B \subset X$. Moreover, let $f_1^*, f_2^* : X^* \to (-\infty, +\infty]$ be the associated conjugate functional of f_1 and f_2. Then, there holds

$$(f_1 + f_2)^*(\xi) = \min_{\eta \in X^*}\left(f_1^*(\xi - \eta) + f_2^*(\eta)\right) \quad \textit{for all } \xi \in X^*. \tag{2.3.4}$$

Proof. This is proven in ATTOUCH & BRÉZIS [18, Theorem 1.1, pp. 126]. □

For an illustration of the preceding lemma, we consider

Example 2.3.6 Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two BANACH spaces such that $X \cap Y$, equipped with the norm $\|\cdot\|_{X\cap Y} = \|\cdot\|_X + \|\cdot\|_Y$, is dense in both X and Y. Furthermore, we assume that X and Y are each continuously embedded into another BANACH space Z. 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 $\|\xi\|_{X^*+Y^*} = \inf_{\substack{\xi_1 \in X^*, \xi_2 \in Y^* \\ \xi = \xi_1 + \xi_2}} \max\{\|\xi_1\|_{X^*}, \|\xi_2\|_{Y^*}\}$, see, e.g., GAJEWSKI et al. [84, Chapter I, Section 5]. Let $p, q \in (1, +\infty)$ and the functionals $f_1, f_2 : X \cap Y \to \mathbb{R}$ be defined 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 2.3.5, the conjugate $(f_1+f_2)^*: X^*+Y^* \to (-\infty,+\infty]$ is given by

$$(f_1+f_2)^*(\xi) = \min_{\substack{\xi_1\in X^*,\xi_2\in Y^*\\ \xi=\xi_1+\xi_2}} \left(\frac{1}{p^*}\|\xi_1\|_{X^*}^{p^*} + \frac{1}{q^*}\|\xi_2\|_{Y^*}^{q^*}\right) \quad \text{for all } \xi \in X^*+Y^*, \tag{2.3.5}$$

where $p^*>1$ and $q^*>1$ again denotes the conjugate exponent of p and q, respectively. Applying YOUNG's inequality to (2.3.5), we obtain the estimates

$$\begin{aligned} (f_1+f_2)(u) &\geq C\|u\|_{X\cap Y} - C && \text{for all } u \in X\cap Y \\ (f_1+f_2)^*(\xi) &\geq C\|\xi\|_{X^*+Y^*} - C && \text{for all } \xi \in X^*+Y^* \end{aligned}$$

for some constant $C>0$. We will make use of the latter estimates in Chapter 5 by choosing $X=\mathrm{L}^2(0,T;V)$ and $Y=\mathrm{L}^r(0,T;W)$.

We continue with addressing the following problem: let $f(t,\cdot)\to(-\infty,+\infty]$ be a proper, lower semicontinuous, and convex functional for each $t\in[0,T]$ and define the *integral functional*

$$F(v) = \begin{cases} \int_0^T f(t,v(t))\,\mathrm{d}t & \text{if } f(\cdot,v(\cdot))\in \mathrm{L}^1(0,T), \\ +\infty & \text{otherwise.} \end{cases} \tag{2.3.6}$$

We want to know whether the properties of f are inherited by F. The following lemma provides sufficient condition to give a positive answer to this question. Furthermore, it gives a relation between the subdifferential of F on BOCHNER-LEBESGUE spaces and f. To ensure that the mapping $t\mapsto f(t,v(t))$ is LEBESGUE measurable for any BOCHNER measurable (strongly measurable) functional $v:[0,T]\to X$, we introduce the notion of a *normal integrand*, which was introduced by the American mathematician Ralph T. ROCKAFELLAR [138, 140, 141] in order to study integrals of the form (2.3.6) for a wider class of integrands f than the classical CARATHÉODORY function[2].

We denote with $\mathscr{L}_{(0,T)}$ the LEBESGUE σ-algebra of the interval $[0,T]$ and with $\mathscr{B}(X)$ the BOREL σ-algebra of X. A functional $f:[0,T]\times X\to(-\infty,+\infty]$ is called *normal integrand* if it is $\mathscr{L}_{(0,T)}\otimes\mathscr{B}(X)$-measurable on $[0,T]\times X$ and for a.e. $t\in(0,T)$ the mapping $v\mapsto f(t,v)$ is lower semicontinuous on X. Note that if f is a normal integrand and X is a separable BANACH space, then by the PETTIS theorem, see, e.g., DIESTEL & UHL [58, Theorem 2, p. 42], the mapping $t\mapsto f(t,v(t))$ is LEBESGUE measurable for any BOCHNER measurable functional $v:[0,T]\to X$.

The BOCHNER–LEBESGUE spaces[3] are as usual denoted by $\mathrm{L}^p(0,T;X)$ for $p\in[1,+\infty]$.

Theorem 2.3.7 *Let X be a separable and reflexive BANACH space and $f:[0,T]\times X\to(-\infty,+\infty]$ be a normal integrand such that $f(t,\cdot):X\to(-\infty,+\infty]$ is for a.e. $t\in(0,T)$ a proper, lower semicontinuous and convex functional. Denote*

[2] The functional $f(t,\cdot)\to\mathbb{R}$ is called CARATHÉODORY functional if the mapping $t\mapsto f(t,v)$ is LEBESGUE measurable for all $v\in X$ and the mapping $v\mapsto f(t,v)$ is continuous for almost every $t\in(0,T)$.

[3] See ROUBÍČEK [145, Section 1.5] for a definition of the BOCHNER–LEBESGUE spaces.

with $f^* : [0,T] \times X^* \to (-\infty, +\infty]$ *the conjugate functional given by* $f^*(t,\cdot) = (f(t,\cdot))^*, t \in [0,T]$*, and assume that there exists constants* $\alpha, \alpha^*, \beta, \beta^* > 0$ *such that*

$$f(t,v) + \alpha\|v\| + \beta \geq 0 \quad \textit{for a.e. } t \in [0,T] \textit{ and all } v \in X,$$

and

$$f^*(t,\xi) + \alpha^*\|\xi\|_* + \beta^* \geq 0 \quad \textit{for a.e. } t \in [0,T] \textit{ and all } \xi \in X^*.$$

Then following assertions hold

i) *The functional* $f^* : [0,T] \times X^* \to (-\infty, +\infty]$ *is a normal integrand, and if* F *is proper, then the conjugate functional* $F^* : \mathrm{L}^{p^*}(0,T;X^*) \to \overline{\mathbb{R}}$ *is proper, lower semicontinuous and convex, and is given by the integral functional*

$$F^*(\xi) = \begin{cases} \int_0^T f^*(t,\xi(t))\,\mathrm{d}t & \textit{if } f^*(\cdot,\xi(\cdot)) \in \mathrm{L}^1(0,T), \\ +\infty & \textit{otherwise.} \end{cases}$$

ii) *The functional* F *is lower semicontinuous and convex on* $\mathrm{L}^p(0,T;X)$*, and there holds* $F(v) > -\infty$ *for all* $v \in \mathrm{L}^p(0,T;X)$*.*

iii) *Let* F *be proper, and let* $v \in \mathrm{dom}(F)$ *and* $\xi \in \mathrm{L}^{p^*}(0,T;X^*)$*. Then,* $\xi \in \partial F(v) \subset \mathrm{L}^{p^*}(0,T;X^*)$ *if and only if* $\xi(t) \in \partial f(t,v(t)) \subset X^*$ *for a.e.* $t \in (0,T)$*.*

Proof. Assertions *i)* and *ii)* follow from Kenmochi [99] and Rockafellar [140, Proposition 2 & Theorem 2] as well as Lemma 2.3.2, respectively. Assertion *iii)* follows from *i), ii)*, Lemma 2.3.1, and the fact that

$$\int_0^T \left(f(t,v(t)) + f^*(t,\xi(t)) - \langle \xi(t), v(t)\rangle\right) = 0 \tag{2.3.7}$$

if and only if

$$f(t,v(t)) + f^*(t,\xi(t)) - \langle \xi(t), v(t)\rangle = 0 \quad \text{a.e. in } (0,T),$$

which in turn follows from the fact that the integrand in (2.3.7) is by the Fenchel–Young inequality always non-negative. □

2.4 Mosco-convergence

In this section, we introduce the notion of the Mosco-*convergence*, which was originally introduced by the Italian mathematician Umberto Mosco [125] in order to study variational inequalities. Before we motivate the Mosco-convergence, we provide a definition.

Definition 2.4.1 *A sequence of functionals* $f_n : X \to (-\infty, +\infty]$ *converges to* $f : X \to (-\infty, +\infty]$ *in the sense of Mosco (we write* $f_n \xrightarrow{\mathrm{M}} f$*) if and only if for all* $u \in X$

$$\begin{cases} a) & f(u) \leq \liminf_{n\to\infty} f_n(u_n) \quad \textit{for all } u_n \rightharpoonup u \textit{ in } X, \\ b) & \exists \hat{u}_n \to u \textit{ in } V \textit{ such that } f(u) \geq \limsup_{n\to\infty} f_n(\hat{u}_n). \end{cases}$$

We note that the implication in b) can be replaced by $f(u) = \lim_{n\to\infty} f_n(\hat{u}_n)$ since the other direction of the inequality already follows from a). The existence of a strongly convergent sequence in b) is often referred to as the *recovery sequence*. We note further that constant sequences of functions do not, in general, converge in the sense of Mosco since the functional is by a) assumed to be weakly lower semicontinuous. However, if we deal with functionals that are lower semicontinuous and convex, and thus weakly lower semicontinuous, then constant sequences converge in the sense of Mosco. The Mosco-convergence is related to the notion of *Γ-convergence*[4] where the convergences in a) and b) in Definition 2.4.1 are assumed to hold with respect to the same topology, which usually is either the strong topology or the weak topology. The Γ-convergence gives a sufficient condition to conclude that a sequence of solutions u_n to the minimization problems

$$\inf_{v\in X} f_n(v)$$

converge in a certain topology to a solution to a limiting minimization problem as $n \to \infty$. The Mosco-convergence, which is a stronger notion of convergence, provides a sufficient condition to conclude that a sequence of subgradients converge to a subgradient of a limiting functional as we will see. Therefore, the Mosco-convergence and the Γ-convergence are very useful tools in, e.g., phase transitions, homogenization theory, dimension reduction, the formalization of the passage of a discrete model to a continuous model, etc., see [30, 114, 117–119, 147, 152]. We refer the interested reader to the monographs Braides [31] and Dal Maso [50] for an introduction to Γ-convergence.

In Lemma 2.2.2, we have seen that the lower semicontinuity and λ-convexity of a functional yields the strong-weak closedness of the graph of its subdifferential. However, sometimes we do only have weakly convergent sequences $u_n \rightharpoonup u$ and $\xi_n \rightharpoonup \xi$ with $\xi_n \in \partial f(u_n)$, $n \in \mathbb{N}$, at our disposal which is, in general, not enough to conclude $\xi \in \partial f(u)$. However, for a proper, lower semicontinuous, and convex functional, a sufficient condition to make this conclusion is in fact given by the limsup estimate

$$\limsup_{n\to\infty} \langle \xi_n - \xi, u_n - u\rangle \le 0.$$

This holds even true for maximal monotone operators, see, e.g., Brézis, Crandall & Pazy [33, Lemma 1.2], which in particular contain the set of subdifferential operators of proper, lower semicontinuous, and convex functionals, see Rockafellar [137, Theorem 4]. If we consider a sequence of functionals $(f_n)_{n\in\mathbb{N}}$ so that $\xi_n \in \partial f(u_n)$ is replaced by $\xi_n \in \partial f_n(u_n)$, our next question is: what type of convergence for the sequence $(f_n)_{n\in\mathbb{N}}$ to a functional f is sufficient to conclude $\xi \in \partial f(u)$. In fact, as we mentioned before, such a convergence is given by the Mosco-convergence.

Lemma 2.4.2 *Let $f, f_n : X \to (-\infty, +\infty]$ be proper, lower semicontinuous, and convex functionals for all $n \in \mathbb{N}$, and denote with $f_n^*, f^* : X^* \to (-\infty, +\infty]$ the*

[4]The Γ-convergence has originally been introduced by the Italian mathematician Ennio De Giorgi [51–53, 55] in a series of articles, a couple of years after the introduction of the Mosco-convergence, where he studied Green functions.

associated conjugate functionals. Moreover, let $v_n \rightharpoonup v$ in X and $\xi_n \rightharpoonup \xi$ in X^* as $n \to \infty$ with $\xi_n \in \partial f_n(u_n), n \in \mathbb{N}$ such that

$$\limsup_{n\to\infty} \langle \xi_n - \xi, u_n - u \rangle \leq 0.$$

If

$$f_n \xrightarrow{\mathrm{M}} f \quad or \quad f(u) + f^*(\xi) \leq \liminf_{n\to\infty} \left(f_n(u_n) + f_n^*(\xi_n) \right),$$

then $u \in \partial f(u)$ and

$$\lim_{n\to\infty} f_n(u_n) = f(u), \ \lim_{n\to\infty} f_n^*(\xi_n) = f^*(\xi) \quad or \quad f(u) + f^*(\xi) = \lim_{n\to\infty} \left(f(u_n) + f^*(\xi_n) \right),$$

respectively.

Proof. We assume first that $f_n \xrightarrow{\mathrm{M}} f$. Let $v \in X$, then by the Mosco-convergence of $(f_n)_{n\in\mathbb{N}}$ there exists a strongly convergent sequence $\hat{v}_n \to v$ in X as $n \to \infty$ such that $f(v) \geq \limsup_{n\to\infty} f_n(\hat{v}_n)$. With the liminf estimate $a)$ for the Mosco-convergence, we obtain

$$\begin{aligned} f(u) - f(v) &\leq \liminf_{n\to\infty} f_n(u_n) - \limsup_{n\to\infty} f_n(\hat{v}_n) \\ &\leq \liminf_{n\to\infty} \left(f_n(u_n) - f_n(\hat{v}_n) \right) \\ &= \liminf_{n\to\infty} \langle \xi_n, u_n - \hat{v}_n \rangle \\ &\leq \limsup_{n\to\infty} \langle \xi_n, u_n - \hat{v}_n \rangle \\ &= \langle \xi, u - v \rangle \quad \text{for all } v \in X, \end{aligned}$$

whence $\xi \in f(u)$. Now, let $\hat{u}_n \to u$ in X as $n \to \infty$ such that $f(u) \geq \limsup_{n\to\infty} f_n(\hat{u}_n)$. Then, we obtain

$$\begin{aligned} f(u) &\leq \liminf_{n\to\infty} f_n(u_n) \\ &\leq \limsup_{n\to\infty} f_n(u_n) \\ &\leq \limsup_{n\to\infty} \left(\langle \xi_n, u_n - \hat{u}_n \rangle + f_n(\hat{u}_n) \right) \\ &\leq \limsup_{n\to\infty} \left(\langle \xi_n - \xi, u_n - \hat{u}_n \rangle + \langle \xi, u_n - \hat{u}_n \rangle + f_n(\hat{u}_n) \right) \\ &\leq \limsup_{n\to\infty} \langle \xi_n - \xi, u_n - \hat{u}_n \rangle + \limsup_{n\to\infty} \langle \xi, u_n - \hat{u}_n \rangle + \limsup_{n\to\infty} f_n(\hat{u}_n) \\ &\leq f(u), \end{aligned}$$

and hence $\lim_{n\to\infty} f_n(u_n) = f(u)$. Exploiting Lemma 2.3.1, we also obtain

$$\begin{aligned}
f^*(\xi) &= \langle \xi, u\rangle - f(u) \\
&= \lim_{n\to\infty} \left(\langle \xi_n, \hat{u}_n\rangle - f_n(\hat{u}_n)\right) \\
&\le \liminf_{n\to\infty} \sup_{v\in X} \left(\langle \xi_n, v\rangle - f_n(v)\right) \\
&= \liminf_{n\to\infty} \left(\langle \xi_n, u_n\rangle - f_n(u_n)\right) \\
&= \liminf_{n\to\infty} f_n^*(\xi_n) \\
&\le \limsup_{n\to\infty} f_n^*(\xi_n) \\
&= \limsup_{n\to\infty} \left(\langle \xi_n, u_n\rangle - f_n(u_n)\right) \\
&= \limsup_{n\to\infty} \left(\langle \xi_n - \xi, u_n - u\rangle + \langle \xi, u\rangle - f_n(u_n)\right) \\
&\le \limsup_{n\to\infty} \langle \xi_n - \xi, u_n - u\rangle + \langle \xi, u\rangle - \lim_{n\to\infty} f_n(u_n) \\
&\le \langle \xi, u\rangle - f(u) \\
&= f^*(\xi)
\end{aligned}$$

from which $\lim_{n\to\infty} f_n^*(\xi_n) = f^*(\xi)$ follows. Now, we assume that $f(u) + f^*(\xi) \le \liminf_{n\to\infty} (f_n(u_n) + f_n^*(\xi_n))$. Then, with Lemma 2.3.1 and the Fenchel–Young inequality, we find

$$\begin{aligned}
\langle \xi, u\rangle &\le f(u) + f^*(\xi) \\
&\le \liminf_{n\to\infty} \left(f_n(u_n) + f_n^*(\xi_n)\right) \\
&\le \limsup_{n\to\infty} \left(f_n(u_n) + f_n^*(\xi_n)\right) \\
&\le \limsup_{n\to\infty} \langle \xi_n, u_n\rangle \\
&\le \langle \xi, u\rangle
\end{aligned}$$

whence $\xi \in \partial f(u)$ and $f(u) + f^*(\xi) = \lim_{n\to\infty} (f_n(u_n) + f_n^*(\xi_n))$. $\square$

In view of Lemma 2.3.1, we obtain the same implication in the previous result by replacing $f_n \xrightarrow{\mathrm{M}} f$ with $f_n^* \xrightarrow{\mathrm{M}} f^*$. So it seems natural to assume that there is a relation between these two convergences. In fact, Attouch [17, Theorem 3.18, p. 295] has shown that they are equivalent if the underlying Banach space X is reflexive. Based on that, Stefanelli showed the following equivalence.

Lemma 2.4.3 *Let X be a reflexive Banach space and let $f, f_n : X \to (-\infty, +\infty]$ be proper, lower semicontinuous, and convex functionals for all $n \in \mathbb{N}$, and denote with $f^*, f_n^* : X \to (-\infty, +\infty]$ the associated conjugate functionals. Then, $f_n \xrightarrow{\mathrm{M}} f$ if and only if*

$$\begin{cases}
a) & f(u) \le \inf\{\liminf_{n\to\infty} f_n(u_n) : u_n \rightharpoonup u \text{ in } X\}, \\
b) & f^*(\xi) \le \inf\{\liminf_{n\to\infty}, f_n^*(\xi_n) : \xi_n \rightharpoonup \xi \text{ in } X^*\}, \\
c) & (f_n^*)_{n\in\mathbb{N}} \text{ is uniformly proper,}
\end{cases}$$

where point iii) means that there exists a bounded sequence $(\xi_n)_{n\in\mathbb{N}} \subset X^*$ such that $\xi_n \in \mathrm{dom}(f_n^*)$ for all $n \in \mathbb{N}$.

Proof. This has been proven in Stefanelli [155, Lemma 4.1]. □

Lemma 2.4.3 gives a characterization of the Mosco-convergence in terms of a functional and its conjugate without assuming the existence of a recovery sequence, which makes it easier to verify in practice. The lemma also shows that the Mosco-convergence $f_n \xrightarrow{\mathrm{M}} f$ actually implies the liminf estimate for the sum $f_n + f_n^*$ in Lemma 2.4.2.

We want to employ the previous results in Chapter 3 where we study perturbed gradient systems and in Chapter 6 where we study nonlinearly damped inertial systems by choosing $f_n = \Psi_{u_n}$. More precisely, we will choose $f_n = \Psi_{\underline{U}_{\tau_n}(t)}$ where $\underline{U}_{\tau_n}$ are the piecewise constant interpolations, see Section 3.4. In Chapter 3, we will obtain a strong convergence of the sequence $(\underline{U}_{\tau_n})_{n\in\mathbb{N}}$ uniformly on $[0,T]$, which makes it reasonable to assume the Mosco-convergence of the sequence $(\Psi_{u_n})_{n\in\mathbb{N}}$ for strongly convergent sequences $u_n \to u$. However, for nonlinearly damped inertial systems, we only obtain a pointwise weak convergence of $(\underline{U}_{\tau_n})_{n\in\mathbb{N}}$ so that assuming the Mosco-convergence of the sequence $(\Psi_{u_n})_{n\in\mathbb{N}}$ is too restrictive and not necessary as we will see. Therefore, we will assume in Chapter 6 a liminf estimate for the sum $\Psi_{\underline{U}_{\tau_n}(t)} + \Psi^*_{\underline{U}_{\tau_n}(t)}$ on suitable Bochner–Lebesgue spaces, which is already implied by the Mosco-convergence. The following lemma shows that this will be sufficient in order to obtain the weak-weak closedness of the graph of the subdifferential.

Lemma 2.4.4 *Let the functionals $f, f_n : [0,T] \times X \to (-\infty,+\infty]$ be given and fulfill the assumptions of Theorem 2.3.7, and let $p \in (1,+\infty)$. Furthermore, let $(v_n)_{n\in\mathbb{N}} \subset \mathrm{L}^p(0,T;X)$ and $(\xi_n)_{n\in\mathbb{N}} \subset \mathrm{L}^{p^*}(0,T;X^*)$ with $\xi_n \in \partial F_n(v_n)$ such that $v_n \rightharpoonup v$ in $\mathrm{L}^p(0,T;X)$ and $\xi_n \rightharpoonup \xi$ in $\mathrm{L}^{p^*}(0,T;X^*)$ as $n \to \infty$ where F_n is the integral functional associated to f_n. If*

$$\int_0^T (f(t,v(t)) + f^*(t,\xi(t)))\,\mathrm{d}t \le \liminf_{n\to\infty} \int_0^T (f_n(t,v_n(t)) + f_n^*(t,\xi_n(t)))\,\mathrm{d}t \tag{2.4.1}$$

and there holds

$$\limsup_{n\to\infty} \int_0^T \langle \xi_n(t) - \xi(t), v_n(t) - v(t)\rangle\,\mathrm{d}t \le 0,$$

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

$$\int_0^T (f(t,v(t)) + f^*(t,\xi(t)))\,\mathrm{d}t = \lim_{n\to\infty} \int_0^T (f_n(t,v_n(t)) + f_n^*(t,\xi_n(t)))\,\mathrm{d}t$$

Proof. This immediately follows from Lemma 2.4.2 and Theorem 2.3.7. □

Remark 2.4.5 It has been shown in Stefanelli [155, Lemma 4.1] that under the assumptions of Lemma 2.4.4, the convergence $f_n \xrightarrow{\mathrm{M}} f$ implies $F_n \xrightarrow{\mathrm{M}} F$, which in turn implies the liminf estimate (2.4.1).

2.5 The Moreau–Yosida regularization

In this section, we study for a general proper, lower semicontinuous, and convex functional $f : X \to (-\infty,+\infty]$ on a normed space $(X, \|\cdot\|)$ the properties of the so-called *Moreau–Yosida regularization*

$$f_\varepsilon(u) = \inf_{v\in X}\left\{\frac{1}{2\varepsilon}\|u-v\|^2 + f(v)\right\} \quad , u \in X,$$

of f, where $\varepsilon > 0$ is called the *regularization parameter*. In this section, we are primarily guided by BARBU [26] and BARBU & PRECUPANU [27]. We show in this section to what extent the geometrical properties of the dual space X^* are translated to the regularity properties of the regularization f_ε. Roughly speaking, the better the geometrical properties of the dual space X^* are, the better the regularization becomes.

We recall the definition of the duality map $F_X : X \rightrightarrows X^*$, which is given by $F_X(v) := \{\xi \in X^* : \langle \xi, v\rangle = \|v\|^2 = \|\xi\|_*^2\}$. As we have shown in Example 2.3.4, the duality map is given by the subdifferential of $\frac{1}{2}\|\cdot\|^2$, i.e., $F_X(u) = \partial(\frac{1}{2}\|u\|^2)$ for all $u \in X$. Furthermore, it is well-known that the set $F_X(u)$ is in every point $u \in X$ non-empty, convex, bounded, and weak∗-closed [5], see, e.g., BARBU & PRECUPANU [27, Section 1.2.4]. The duality map has also a geometrical interpretation: by the HAHN–BANACH theorem[6], there holds

$$\|u\| = \max_{\substack{\zeta\in X^*\\ \|\zeta\|_*=1}} \langle \zeta, u\rangle = \max_{\substack{\zeta\in X^*\\ \|\zeta\|_*=\|u\|}} \frac{\langle \zeta, u\rangle}{\|u\|} \geq \frac{\langle \xi, u\rangle}{\|u\|} \quad \text{for all } \xi \in X \text{ with } \|\xi\|_* = \|u\|.$$

Thus, an element of the dual space belongs to the duality map $\xi^* \in F_X(u)$ if and only if it solves the maximization problem

$$\max_{\substack{\zeta\in X^*\\ \|\zeta\|_*=\|u\|}} \frac{\langle \zeta, u\rangle}{\|u\|}, \tag{2.5.1}$$

for which the set of maximizers is non-empty. In other words, ξ^* generates a closed supporting hyperplane to the closed ball $\overline{B}(0, \|u\|)$. We call a norm *smooth* if and only if the duality map is single-valued, or geometrically speaking, each supporting hyperplane which passes through a boundary point of the sphere $S(0, \|u\|)$ with radius $\|u\|$ is also a tangential hyperplane. We call a normed space smooth if there is an equivalent smooth norm. From (2.5.1), it is readily seen that if the dual space X^* is strictly convex, i.e., the dual norm $\|\cdot\|_*$ is strictly convex, the element which generates the supporting hyperplane is unique, meaning that the duality map $F_V(u)$ is single-valued. In this case, the duality map is also demicontinuous [7], which implies that the norm on X is GÂTEAUX differentiable. If the dual space X^* is uniformly convex[8], then the duality map is uniformly continuous on every bounded subset of X and the norm on X is uniformly FRÉCHET differentiable in the sense that the limit

$$\lim_{\lambda\to 0} \frac{\|u+\lambda v\| - 1}{\lambda}$$

[5]Therefore, $F_X(u)$ is weak∗-compact.

[6]See, e.g., BRÉZIS [35, Theorem 1.1, p. 1].

[7]A map $f : X \to Y$ between two normed spaces X and Y is called demicontinuous if it is strong-to-weak* continuous.

[8]The normed space X is called uniformly convex if for every $0 < \varepsilon \leq 2$ there exists $\delta > 0$ such that for any two vectors $x, y \in X$ with $\|x\|_X = \|y\|_X = 1$ the condition $\|x-y\|_X \geq \varepsilon$ implies that $\left\|\frac{x+y}{2}\right\|_X \leq 1-\delta$. An uniformly convex space is in particular strict convex.

exists uniformly in $x, y \in S(0,1)$, see [26, 102]. Obviously, the regularity of the norm of a BANACH space is deeply related to the geometrical properties of its dual space. If X is a reflexive BANACH space, then by the renorming theorem due to ASPLUND [16], there exist always equivalent norms $\|\cdot\|$ of X and $\|\cdot\|_*$ of the dual space X^* such that both X and X^* equipped with these norms are strictly convex and smooth, see BARBU & PRECUPANU [27, Theorem 1.105, p. 36]. Consequently, a reflexive BANACH space can be equipped with an equivalent GÂTEAUX differentiable norm such that the duality map is demicontinuous. It is well-known that a HILBERT space, in particular, is reflexive and that the duality map is identical with the RIESZ isomorphism between the HILBERT space and its dual. For a more detailed discussion of the geometry of BANACH spaces, and in particular, concerning the duality maps, we refer the interested reader to [24, 26, 27, 39, 40, 57, 102, 165].

The question arises, if and to what extent the properties of the duality map are related to the regularization properties of the MOREAU-YOSIDA regularization. We will see that the properties of the duality map are inherited by the subdifferential of the MOREAU-YOSIDA regularization. In fact, we will answer the question for the more general so-called p-MOREAU–YOSIDA regularization, which is for $p > 1$ given by

$$f_\varepsilon(u) = \inf_{v \in X} \left\{ \frac{\varepsilon}{p} \left\| \frac{u-v}{\varepsilon} \right\|^p + f(v) \right\} \quad , u \in X. \tag{2.5.2}$$

The reason why we want to study p-MOREAU–YOSIDA regularization is simply because it maintains the growth of the functional f if it has p-growth, see Section 3.2.

The following lemma shows some basic properties of the p-MOREAU–YOSIDA regularization on general normed spaces.

Lemma 2.5.1 *Let $f : X \to (-\infty, +\infty]$ be a proper and convex functional, and, for $\varepsilon > 0$ and $p > 1$, let f_ε be the p-MOREAU–YOSIDA regularization defined by* (2.5.2)*. Then, f_ε is finite, convex, and locally LIPSCHITZ continuous on X. If f is in addition lower semicontinuous, and X is a reflexive BANACH space, then the infimum in $f_\varepsilon(u) = \inf_{v \in X} \left\{ \frac{\varepsilon}{p} \left\| \frac{u-v}{\varepsilon} \right\|^p + f(v) \right\}$ is attained at every point $u \in X$.*

Proof. Let $\tilde{u} \in \operatorname{dom}(f) \neq \emptyset$. Then, on the one hand, there holds

$$f_\varepsilon(u) \leq \frac{1}{p\varepsilon^{p-1}} \|u - \tilde{u}\|^p + f(\tilde{u}) < \infty \quad \text{for every } u \in X. \tag{2.5.3}$$

On the other hand, by EKELAND & TEMAM [69, Proposition 3.1, p. 14], there exists an affine linear minorant to f, i.e., there exist $\xi \in X^*$ and $\alpha \in \mathbb{R}$ such that

$$f(v) \geq \alpha + \langle \xi, v \rangle \quad \text{for all } v \in X,$$

so that $f_\varepsilon(u) > -\infty$ for every $u \in X$, whence $\operatorname{dom}(f_\varepsilon) = X$. Now, let for $\lambda \in (0,1)$ and $u_1, u_2 \in V$, $(v_n^i)_{n \in \mathbb{N}} \subset X$ be a minimizing sequence for $f_\varepsilon(u_i)$ for $i = 1, 2$. We set $w_n := \lambda v_n^1 + (1-\lambda)v_n^2$, $n \in \mathbb{N}$. Then, by the convexity of f, there holds

$$\begin{aligned} f_\varepsilon(\lambda u_1 + (1-\lambda)u_2) &= \inf_{v\in V}\left\{\frac{1}{p\varepsilon^{p-1}}\|\lambda u_1 + (1-\lambda)u_2 - v\|^p + f(v)\right\} \\ &\le \frac{1}{p\varepsilon^{p-1}}\|\lambda u_1 + (1-\lambda)u_2 - w_n\|^p + f(w_n) \\ &\le \lambda\left(\frac{1}{p\varepsilon^{p-1}}\|u_1 - v_n^1\|^p + f(v_n^1)\right) \\ &\quad + (1-\lambda)\left(\frac{1}{p\varepsilon^{p-1}}\|u_2 - v_n^2\|^p + f(v_n^2)\right) \\ &\to \lambda f_\varepsilon(u_1) + (1-\lambda)f_\varepsilon(u_2) \quad \text{as } n\to\infty, \end{aligned}$$

which means that f_ε is convex. We note that by (2.5.3), f_ε is bounded on every open bounded set of X. Hence, EKELAND & TEMAM [69, by Corollary 2.4, p. 12], f_ε is locally LIPSCHITZ continuous on X. Finally, if X is a reflexive BANACH space, then the infimum in $f_\varepsilon(u) = \inf_{v\in X}\left\{\frac{\varepsilon}{p}\left\|\frac{u-v}{\varepsilon}\right\|^p + f(v)\right\}$ is attained at every point $u \in X$ by the direct method of calculus of variations. □

In the next theorem, we will show properties of the p-MOREAU–YOSIDA regularization under the assumption that X is reflexive such that, by the renorming theorem, X and X^* are simultaneously strictly convex and smooth. Before we progress to the next theorem, we recall that the p-duality map F_X^p is given by $F_X^p := \partial\frac{1}{p}\|\cdot\|^p$ for $p > 1$. Then, since the mapping $v \mapsto \frac{1}{p}\|v\|^p$ is continuous and convex on X, EKELAND & TEMAM [69, Proposition 5.1 & 5.2, Corollary 5.1, pp. 21] ensure that F_X^p is a bounded and set-valued map such that $F_X^p(u)$ is non-empty, convex, and weak*-closed for all $u \in X$, and by Example 2.3.4, characterized by

$$F_X^p(u) = \{\xi \in X^* : \langle \xi, u\rangle = \|u\|^p = \|\xi\|_*^{p^*}\}. \tag{2.5.4}$$

As for $p = 2$, if the dual space is strictly convex, then by KIEN [102, Proposition 2.3] and AKAGI & MELCHIONNA [5, Lemma 19], the p-duality map is demicontinuous, single-valued, and monotone in the sense that

$$\langle F_X^p(u) - F_X^p(v), u - v\rangle \ge \left(\|u\|^{p-1} - \|v\|^{p-1}\right)(\|u\| - \|v\|) \quad \text{for all } u, v \in X.$$

With the above-mentioned properties of the p-duality map, we are able to proof in the following theorem that the p-MOREAU–YOSIDA regularization is under suitable conditions GÂTEAUX differentiable with a demicontinuous GÂTEAUX derivative. This result generalizes and follows the proof of Theorem 2.58, p. 98, in BARBU [26] where the case $p = 2$ has been studied.

Theorem 2.5.2 *Let X be reflexive such that X and its dual X^* are strictly convex and smooth, and let $p > 1$ and $\varepsilon > 0$. Furthermore, let $f : X \to (-\infty, +\infty]$ be proper, lower semicontinuous, and convex. Then, the p-MOREAU–YOSIDA regularization is convex and locally LIPSCHITZ continuous and if f is strictly convex, so is f_ε. Moreover, $f_\varepsilon(u) = \inf_{v\in X}\left\{\frac{1}{p\varepsilon^{p-1}}\|u - v\|^p + f(v)\right\}$ attains at every point $u \in X$*

its unique minimizer denoted by $u_\varepsilon := \operatorname{argmin}_{v\in X}\left\{\frac{1}{p\varepsilon^{p-1}}\|u-v\|^p + f(v)\right\}$, *and* u_ε *satisfies the* EULER-LAGRANGE *equation*

$$0 \in F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right) + \partial f(u_\varepsilon). \tag{2.5.5}$$

Furthermore, f_ε *is* GÂTEAUX*-differentiable at every point* $u \in X$ *with the* GÂTEAUX*-derivative* $A_\varepsilon : X \to X^*$ *being demicontinuous on* X *and satisfying* $A_\varepsilon(u) = -F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right)$. *Moreover, the following assertions hold*

i) $f_\varepsilon(u) = \frac{\varepsilon}{p}\|A_\varepsilon(u)\|_*^{p^*} + f(u_\varepsilon)$ *for every* $u \in X$,

ii) $f(u_{\varepsilon_1}) \leq f_{\varepsilon_1}(u) \leq f_{\varepsilon_2}(u) \leq f(u)$ *for all* $u \in X$ *and all* $\varepsilon_1 \geq \varepsilon_2 > 0$,

iii) $\lim_{\varepsilon\to 0}\|u_\varepsilon - u\| = 0$ *for all* $u \in \operatorname{dom}(f)$,

iv) $\lim_{\varepsilon\to 0} f_\varepsilon(u) = f(u)$ *for every* $u \in X$.

Finally, the mapping $\varepsilon \mapsto f_\varepsilon(u)$ *is differentiable on* $(0,+\infty)$ *with*

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f_\varepsilon(u) = -\frac{1}{p^*\varepsilon^p}\|u_\varepsilon - u\|^p \quad \text{for all } \varepsilon > 0. \tag{2.5.6}$$

Proof. By Lemma 2.5.1, the p-MOREAU–YOSIDA regularization is convex and locally LIPSCHITZ continuous on X. Now, let f be strictly convex and let $u^0, u^1 \in X$ and $t \in (0,1)$. Then, we define $u^t = tu^0 + (1-t)u^1$ and assume

$$f_\varepsilon(u^t) = tf_\varepsilon(u^0) + (1-t)f_\varepsilon(u^1).$$

Then, using the convexity of $\|\cdot\|^p$ and f, we obtain

$$\begin{aligned}
tf_\varepsilon(u^0) + (1-t)f_\varepsilon(u^1) &= f_\varepsilon(u^t)\\
&= \inf_{v\in X}\left\{\frac{1}{p\varepsilon^{p-1}}\|u^t - v\|^p + f(v)\right\}\\
&\leq \frac{1}{p\varepsilon^{p-1}}\|u^t - (tu_\varepsilon^0 + (1-t)u_\varepsilon^1)\|^p + f(tu_\varepsilon^0 + (1-t)u_\varepsilon^1)\\
&\leq \frac{t}{p\varepsilon^{p-1}}\|u^0 - u_\varepsilon^0\|^p + \frac{(1-t)}{p\varepsilon^{p-1}}\|u^1 - u_\varepsilon^1\|^p \qquad (2.5.7)\\
&\quad + tf(u_\varepsilon^0) + (1-t)f(u_\varepsilon^1)\\
&= tf_\varepsilon(u^0) + (1-t)f_\varepsilon(u^1),
\end{aligned}$$

where $u_\varepsilon^i := \operatorname{argmin}_{v\in X}\left\{\frac{1}{p\varepsilon^{p-1}}\|u^i - v\|^p + f(v)\right\}, i = 0,1$. Therefore, the inequality (2.5.7) becomes an equality that implies the two equalities

$$\begin{aligned}
\frac{1}{p\varepsilon^{p-1}}\|t(u^0 - u_\varepsilon^0) + (1-t)(u^1 - u_\varepsilon^1)\|^p &= \frac{1}{p\varepsilon^{p-1}}\|u^t - (tu_\varepsilon^0 + (1-t)u_\varepsilon^1)\|^p\\
&= \frac{t}{p\varepsilon^{p-1}}\|u^0 - u_\varepsilon^0\|^p + \frac{(1-t)}{p\varepsilon^{p-1}}\|u^1 - u_\varepsilon^1\|^p
\end{aligned}$$

and

$$f(tu_\varepsilon^0 + (1-t)u_\varepsilon^1) = tf(u_\varepsilon^0) + (1-t)f(u_\varepsilon^1).$$

Then the strict convexity of the norm $\|\cdot\|$ implies $u^0 - u_\varepsilon^0 = u^1 - u_\varepsilon^1$ and the strict convexity of f implies $u_\varepsilon^0 = u_\varepsilon^1$ whence $u^0 = u^1$ and the strict convexity of f_ε.

The strict convexity of the norm also implies that the resolvent operator $J_\varepsilon(u) := \operatorname{argmin}_{v\in X}\left\{\frac{1}{p\varepsilon^{p-1}}\|u-v\|^p + f(v)\right\}$ is single-valued for every $u \in X$ and satisfies by Lemma 2.2.7 the inclusion (2.5.5). We define $A_\varepsilon(u) := -F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right)$, and note that from the characterization (2.5.4) of the p-duality map, there holds

$$\begin{aligned} f_\varepsilon(u) &= \frac{\varepsilon}{p}\left\|\frac{u_\varepsilon - u}{\varepsilon}\right\|^p + f(u_\varepsilon) \\ &= \frac{\varepsilon}{p}\left\|F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right)\right\|_*^{p^*} + f(u_\varepsilon) \\ &= \frac{\varepsilon}{p}\left\|A_\varepsilon(u)\right\|_*^{p^*} + f(u_\varepsilon). \end{aligned}$$

If we show that the operator A_ε is the GÂTEAUX derivative of f_ε, $i)$ follows. First, AKAGI & MELCHIONNA [5, Lemma 19] have shown that the operator $A_\varepsilon : X \to X^*$ is demicontinuous, i.e., for all sequences $u_n \to u$ in X as $n \to \infty$, there holds $A_\varepsilon(u_n) \rightharpoonup A_\varepsilon(u)$ in X^* as $n \to \infty$. Second, we show that $A_\varepsilon(u)$ belongs to the subdifferential $\partial f_\varepsilon(u)$ for every $u \in V$. Let $u, v \in X$ and $u_\varepsilon = J_\varepsilon(u), v_\varepsilon = J_\varepsilon(v)$. Then, in view of (2.5.4) and the fact that $A_\varepsilon(u) = -F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right) \in \partial f(u)$, we find

$$\begin{aligned} f_\varepsilon(u) - f_\varepsilon(v) &= \frac{\varepsilon}{p}\left\|\frac{u_\varepsilon - u}{\varepsilon}\right\|^p + f(u_\varepsilon) - \frac{\varepsilon}{p}\left\|\frac{v_\varepsilon - v}{\varepsilon}\right\|^p - f(v_\varepsilon) \\ &\le \frac{\varepsilon}{p}\left\|\frac{u_\varepsilon - u}{\varepsilon}\right\|^p - \frac{\varepsilon}{p}\left\|\frac{v_\varepsilon - v}{\varepsilon}\right\|^p - \left\langle F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right), u_\varepsilon - v_\varepsilon\right\rangle \\ &= \frac{\varepsilon}{p}\left\|\frac{u_\varepsilon - u}{\varepsilon}\right\|^p - \frac{\varepsilon}{p}\left\|\frac{v_\varepsilon - v}{\varepsilon}\right\|^p - \left\langle F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right), u_\varepsilon - u\right\rangle \\ &\quad - \left\langle F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right), u - v\right\rangle - \left\langle F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right), v - v_\varepsilon\right\rangle \\ &\le \frac{\varepsilon}{p}\left\|\frac{u_\varepsilon - u}{\varepsilon}\right\|^p - \frac{\varepsilon}{p}\left\|\frac{v_\varepsilon - v}{\varepsilon}\right\|^p - \varepsilon\left\|\frac{u_\varepsilon - u}{\varepsilon}\right\|^p \\ &\quad - \left\langle F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right), u - v\right\rangle + \frac{\varepsilon}{p^*}\left\|F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right)\right\|_*^{p^*} + \frac{\varepsilon}{p}\left\|\frac{v - v_\varepsilon}{\varepsilon}\right\|^p \\ &= -\left\langle F_X^p\left(\frac{u_\varepsilon - u}{\varepsilon}\right), u - v\right\rangle \\ &= \langle A_\varepsilon(u), u - v\rangle \quad \text{for all } v \in X, \end{aligned} \tag{2.5.8}$$

whence $A_\varepsilon(u) \in \partial f_\varepsilon(u)$. Subtracting each side of (2.5.8) by $\langle A_\varepsilon(v), u - v\rangle$, we obtain

$$0 \le f_\varepsilon(u) - f_\varepsilon(v) - \langle A_\varepsilon(v), u - v\rangle \le \langle A_\varepsilon(u) - A_\varepsilon(v), u - v\rangle \tag{2.5.9}$$

for all ε and $u, v \in X$. Choosing $u = v + tw$, where $t > 0$ and $w \in X$, and dividing (2.5.9) by t, we obtain

$$\lim_{t\searrow 0}\frac{f_\varepsilon(v + tw) - f_\varepsilon(v)}{t} = \langle A_\varepsilon(v), w\rangle \qquad \text{for all } w \in V,$$

where we used the demicontinuity of A_ε. Hence, the functional f_ε is GÂTEAUX differentiable with derivative A_ε. We prove now the assertion *ii*). The chain of inequalities $f(u_\varepsilon) \leq f_\varepsilon(u) \leq f(u)$ follows immediately from the definition of the p-MOREAU–YOSIDA regularization. To conclude *ii*), it remains to show that the mapping $\varepsilon \mapsto f_\varepsilon(u)$ is monotonically decreasing on $(0,\infty)$ for every fixed $u \in X$. Let $u \in X$ and $0 < \varepsilon_2 < \varepsilon_1$. Then, by the definition of a minimizer

$$\begin{aligned}
f_{\varepsilon_2}(u) &= \frac{\varepsilon_2}{p}\left\|\frac{u_{\varepsilon_2}-u}{\varepsilon_2}\right\|^p + f(u_{\varepsilon_2})\\
&\leq \frac{\varepsilon_2}{p}\left\|\frac{u_{\varepsilon_1}-u}{\varepsilon_2}\right\|^p + f(u_{\varepsilon_1})\\
&= \left(\frac{1}{p\varepsilon_2^{p-1}} - \frac{1}{p\varepsilon_1^{p-1}}\right)\|u_{\varepsilon_1}-u\|^p + \frac{\varepsilon_1}{p}\left\|\frac{u_{\varepsilon_1}-u}{\varepsilon_1}\right\|^p + f(u_{\varepsilon_1})\\
&= \left(\frac{1}{p\varepsilon_2^{p-1}} - \frac{1}{p\varepsilon_1^{p-1}}\right)\|u_{\varepsilon_1}-u\|^p + f_{\varepsilon_1}(u) \qquad (2.5.10)\\
&\leq f_{\varepsilon_1}(u),
\end{aligned}$$

where we used $0 < \varepsilon_2 < \varepsilon_1$ in the last inequality. Now, we aim to show (2.5.6). First, switching the roles of ε_1 and ε_2 in the inequality (2.5.10), and dividing both sides by $\varepsilon_1 - \varepsilon_2 > 0$, we obtain the chain of inequalities

$$\begin{aligned}
&\frac{1}{p(\varepsilon_2\varepsilon_1)^{p-1}}\left(\frac{\varepsilon_1^{p-1}-\varepsilon_2^{p-1}}{\varepsilon_1-\varepsilon_2}\right)\|u_{\varepsilon_2}-u\|^p\\
&\leq -\frac{f_{\varepsilon_1}(u)-f_{\varepsilon_2}(u)}{\varepsilon_1-\varepsilon_2} \qquad (2.5.11)\\
&\leq \frac{1}{p(\varepsilon_2\varepsilon_1)^{p-1}}\left(\frac{\varepsilon_1^{p-1}-\varepsilon_2^{p-1}}{\varepsilon_1-\varepsilon_2}\right)\|u_{\varepsilon_1}-u\|^p
\end{aligned}$$

for all $0 < \varepsilon_2 < \varepsilon_1$, which also implies

$$\|u_{\varepsilon_2}-u\| \leq \|u_{\varepsilon_1}-u\| \quad \text{for all } 0 < \varepsilon_2 < \varepsilon_1. \qquad (2.5.12)$$

Second, since the real-valued mapping $\varepsilon \mapsto f_\varepsilon(u)$ is, by LEBESGUE's differentiation theorem for monotone functions[9],for every fixed $u \in X$ monotone, it is differentiable almost everywhere and there holds

$$\frac{\mathrm{d}f_\varepsilon(u)}{\mathrm{d}\varepsilon^+} \leq \frac{\mathrm{d}f_\varepsilon(u)}{\mathrm{d}\varepsilon^-} \quad \text{for all } \varepsilon > 0, u \in X,$$

where $\frac{\mathrm{d}f_\varepsilon(u)}{\mathrm{d}\varepsilon^+}$ and $\frac{\mathrm{d}f_\varepsilon(u)}{\mathrm{d}\varepsilon^-}$ denote the right and left derivative of $\tilde{\varepsilon} \mapsto f_{\tilde{\varepsilon}}(u)$ in $\tilde{\varepsilon} = \varepsilon$, respectively. Let $\varepsilon > 0$ and $h > 0$ be sufficiently small. Then, choosing $\varepsilon_1 = \varepsilon + h$ and $\varepsilon_2 = \varepsilon$ in the first inequality as well as $\varepsilon_1 = \varepsilon$ and $\varepsilon_2 = \varepsilon - h$ in the second inequality of (2.5.11) yields

$$\frac{1}{p((\varepsilon+h)\varepsilon)^{p-1}}\left(\frac{(\varepsilon+h)^{p-1}-\varepsilon^{p-1}}{h}\right)\|u_\varepsilon - u\|^p \leq -\frac{f_{\varepsilon+h}(u)-f_\varepsilon(u)}{h} \qquad (2.5.13)$$

[9]See, e.g., ELSTRODT [70, Satz 4.5, p. 299]

and

$$-\frac{f_\varepsilon(u) - f_{\varepsilon-h}(u)}{h} \le \frac{1}{p((\varepsilon-h)\varepsilon_1)^{p-1}} \left(\frac{\varepsilon^{p-1} - (\varepsilon-h)^{p-1}}{h} \right) \|u_{\varepsilon-h} - u\|^p \quad (2.5.14)$$
$$\le \frac{1}{p((\varepsilon-h)\varepsilon_1)^{p-1}} \left(\frac{\varepsilon^{p-1} - (\varepsilon-h)^{p-1}}{h} \right) \|u_\varepsilon - u\|^p$$

respectively, where we employed inequality (2.5.12). Finally, letting $h \to 0$ in (2.5.13) and (2.5.14) yields

$$\frac{\mathrm{d}f_\varepsilon}{\mathrm{d}\varepsilon} = -\frac{1}{p^*\varepsilon^p}\|u_\varepsilon - u\|^p \quad \text{for all } \varepsilon > 0.$$

We continue with showing assertion *iii*). Let $u \in \operatorname{dom}(f)$, then the first inequality of (2.5.11) implies

$$\|u_{\varepsilon_2} - u\|^p \le \left(\frac{p(\varepsilon_2\varepsilon_1)^{p-1}}{\varepsilon_1^{p-1} - \varepsilon_2^{p-1}} \right) (f_{\varepsilon_2}(u) - f_{\varepsilon_1}(u)) \quad (2.5.15)$$
$$\le \left(\frac{p(\varepsilon_2\varepsilon_1)^{p-1}}{\varepsilon_1^{p-1} - \varepsilon_2^{p-1}} \right) (f(u) - f_{\varepsilon_1}(u))$$

for all $0 < \varepsilon_2 < \varepsilon_1$. Thus, we obtain $\lim_{\varepsilon_2 \to 0} \|u_{\varepsilon_2} - u\| = 0$. Taking into account the latter convergence and the lower semicontinuity of f, assertion *ii*) yields

$$\begin{aligned} f(u) &\le \liminf_{\varepsilon \to 0} f(u_\varepsilon) \\ &\le \liminf_{\varepsilon \to 0} f_\varepsilon(u) \\ &\le \limsup_{\varepsilon \to 0} f_\varepsilon(u) \le f(u) \quad \text{for all } u \in \operatorname{dom}(f). \end{aligned}$$

If $u \in X \setminus \operatorname{dom}(f)$, we assume that there exists a sequence $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,\infty)$ with $\varepsilon_n \to 0$ as $n \to \infty$ such that $f_{\varepsilon_n}(u) \le C$ for all $n \in \mathbb{N}$ for a constant $C > 0$. However, Inequality (2.5.15) yields $\lim_{n\to\infty} \|u_{\varepsilon_n} - u\| = 0$, and we obtain $f(u) \le \liminf f_{\varepsilon_n}(u) \le C$, which is a contradiction to $u \in X \setminus \operatorname{dom}(f)$. □

The theorem showed us that the MOREAU–YOSIDA regularization has indeed a regularizing effect. In fact, in view of assertion *iv*) and (2.5.6), one can interpret the MOREAU–YOSIDA regularization as a regularizing process described by the following HAMILTON–JACOBI equation supplemented with an initial condition

$$\begin{cases} \frac{\partial}{\partial t}u(t,x) + \frac{1}{p^*}\|d_x u(t,x)\|^p = 0, & x \in X, t > 0 \\ u(0+,x) = f(x), & x \in X, \end{cases} \quad (2.5.16)$$

where a solution $u : [0,\infty) \times X \to \mathbb{R}$ is given by the so-called LAX–OLEINIK formula

$$u(t,x) = f_t(x) = \inf_{y \in X} \left\{ \frac{t}{p} \left\| \frac{x-y}{t} \right\|^p + f(y) \right\},$$

see, e.g., LIONS [109].

Moreover, we have seen to what extent these regularizing and approximating properties depend on the properties of X^* . This, as previously mentioned, becomes more clear when $X = H$ is a Hilbert space. In this case, the Moreau–Yosida regularization is even Fréchet-differentiable and has a Lipschitz continuous derivative with Lipschitz constant equal to the reciprocal of the regularization parameter ε, see, e.g., Barbu & Precupanu [27, Corollary 2.59, p. 99]. Thanks to these nice properties of the regularization and its derivative only available on a Hilbert space, the Moreau–Yosida regularization is more often applied on Hilbert spaces, see, e.g., Bauschke & Combettes [28] for a detailed treatise on Hilbert spaces. The Moreau–Yosida regularization is related to the so-called Yosida approximation, which for a given operator A and $\varepsilon > 0$, refers to the operator $A_\varepsilon = \varepsilon^{-1}(I - S_\varepsilon)$, which is approximative to A, where $S_\varepsilon = (I + \varepsilon A)^{-1}$. The Yosida approximation is successfully employed in the theory of semigroups in order to generate strongly continuous semigroups as in the eminent Hille–Yosida theorem [92, 162] or the nonlinear counterpart [49, 63] as well as in the theory of maximal monotone operators in Brézis [32].

2.6 Parameterized Young measures

In this section, we introduce *parameterized* Young *measures* on infinite-dimensional spaces. The notion of a Young *measure* was invented by the British mathematician Laurence C. Young [163] in 1937 where he introduced them as generalized curves. He introduced generalized curves in order to overcome, for a special class of functionals, the general problem in the theory of calculus of variations that the minimum in the minimization problem (2.1.1) may not be achieved on the space X, but on a larger space $\tilde{X}$, even though, by Ekeland's variational principle [68], one might find a sequence of elements in X that can get f arbitrarily close to the optimal value in (2.1.1). As we mentioned before, solutions to variational problems correlate with weak solutions of differential equations. Therefore, by extending the solution space to generalized curves (Young measures), he generalized the notion of a solution. However, our purpose of introducing parameterized Young measures is not to show existence of measure-valued solutions, but to use it as a tool in order to characterize the weak limits of sequences in terms of Young measures. This section is mainly guided by Stefanelli [155]. For a comprehensive treatise of Young measures, we refer the reader to Castaing, Raynaud de Fitte, & Valadier [42] for Young measures on separable Banach spaces, and to Málek, Nečas, Rokyta, & Růžička [112] and Evans [81] for the classical Young measures on finite-dimensional spaces.

First, we introduce some notions and functional spaces. Here, X is a reflexive and separable Banach space. A $\mathscr{L}_{(0,T)} \otimes \mathscr{B}(X)$-measurable functional $f : [0,T] \times X \to (-\infty, +\infty]$ is called *weakly-normal integrand* if for a.e. $t \in (0,T)$ the mapping $w \mapsto f(t,w)$ is sequentially lower semicontinuous with respect to the weak topology of X. Furthermore, a family $\boldsymbol{\mu} = (\mu_t)_{t\in(0,T)}$ of Borel probability measures on X is called Young measure if the mapping $(0,T) \ni t \mapsto \mu_t(B)$ is $\mathscr{L}_{(0,T)}$-measurable for all $B \in \mathscr{B}(X)$. We denote with $\mathscr{Y}(0,T;X)$ the set of all Young measures in X.

The following theorem, the so-called fundamental theorem for weak topologies, provides an infinite-dimensional and lower semicontinuous version of the classical fundamental theorem for YOUNG measures, see, e.g., Ball [22].

Theorem 2.6.1 (Fundamental theorem for weak topologies) *Let $f_n, f : (0,T)\times X \to (-\infty,+\infty]$ be for all $n \in \mathbb{N}$ a weakly normal integrand such that for all $w \in X$ and for almost every $t \in (0,T)$, there holds*

$$f(t,w) \le \inf\{\liminf_{n\to\infty} f_n(t,w_n) : w_n \rightharpoonup w \quad \text{in } X\}. \tag{2.6.1}$$

For $p \in [1,+\infty]$, let $(w_n)_{n\in\mathbb{N}} \subset \mathrm{L}^p(0,T;X)$ be a bounded sequence. If $p = 1$, we suppose further that $(w_n)_{n\in\mathbb{N}}$ is equi-integrable in $\mathrm{L}^1(0,T;X)$. Then, there exists a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ and a YOUNG *measure $\boldsymbol{\mu} = (\mu_t)_{t\in(0,T)}$ such that for almost every $t \in (0,T)$, there holds*

$$\mathrm{sppt}(\mu_t) \subset \mathrm{Li}(t) := \bigcap_{p=1}^{\infty} \mathrm{clos}_{\mathrm{weak}}\Big(\{w_{n_k}(t) : k \ge p\}\Big),$$

i.e., μ_t is concentrated on the set of all weak limit points of $(w_{n_k})_{k\in\mathbb{N}}$, and, if the sequence $(f^-(\cdot,w_{n_k}(\cdot))_{k\in\mathbb{N}}$, with $f_n^-(t,w_{n_k}(t)) := \max\{-f_n(t,w_{n_k}(t)),0\}$, is equi-integrable, there holds

$$\int_0^T \int_X f(t,w)\,\mathrm{d}\mu_t(w)\,\mathrm{d}t \le \liminf_{k\to} \int_0^T f_{n_k}(t,w_{n_k}(t))\,\mathrm{d}t.$$

Setting

$$\mathrm{W}(t) := \int_X w \;\mathrm{d}\mu_t(w) \quad \text{a.e. } t \in (0,T),$$

there holds

$$w_{n_k} \rightharpoonup \mathrm{W} \quad \text{in } \mathrm{L}^p(0,T;X) \quad \text{as } k \to \infty,$$

with $\rightharpoonup$ replaced by $\overset{*}{\rightharpoonup}$ if $p = \infty$.

Proof. This has been shown in Theorem 4.3 and a subsequent discussion of the same theorem in STEFANELLI [155]. □

For the sake of completeness, we want to introduce the BOCHNER spaces we deal with throughout the thesis: for $k \in \mathbb{N}$ and $p \in [1,+\infty]$, we denote with $\mathrm{W}^{k,p}(0,T;X)$ the space of abstract functions $v : [0,T] \to X$ which are weakly differentiable up to the order k and whose k-th derivative is in the BOCHNER–LEBESGUE space $\mathrm{L}^p(0,T;X)$. If $p = 2$, we write $\mathrm{H}^m(0,T;X) = \mathrm{W}^{p,2}(0,T;X)$. Furthermore, with $\mathrm{AC}([0,T];X)$, $\mathrm{C}([0,T];X)$ and $\mathrm{C}_w([0,T];X)$ we denote the space of abstract functions which are absolutely continuous, continuous, and continuous with respect to the weak topology of X, respectively. All spaces are equipped with the standard norm.

Part I

Evolution Inclusion of First Order

Chapter 3

Perturbed Gradient System

In this chapter, we investigate the abstract CAUCHY problem

$$\begin{cases} \partial \Psi_{u(t)}(u'(t)) + \partial \mathcal{E}_t(u(t)) \ni B(t, u(t)) & \text{in } V^* \quad \text{for a.e. } t \in (0, T), \\ u(0) = u_0 \in \operatorname{dom}(\mathcal{E}_0) \end{cases} \tag{3.0.1}$$

on a separable and reflexive BANACH space $(V, \|\cdot\|)$, where $\Psi : V \times V \to \mathbb{R}, (u, v) \mapsto \Psi_u(v)$ is the *dissipation potential* or *dissipation mechanism*, $\mathcal{E} : [0, T] \times V \to (-\infty, +\infty], (t, u) \mapsto \mathcal{E}_t(u)$ is the *energy* or *driving functional*, and $B : [0, T] \times V \to V^*$ is the *perturbation*. As the name suggests, the dissipation potential describes dissipative or irreversible processes[1] of a physical system modeled by (3.0.1). The free energy or the entropy itself is described by the energy functional that drives the evolution of the system. In a pure gradient system, i.e., if $B = 0$, the dissipation potential and the energy functional completely determine the evolution of the system. The equation, on the other hand, does not uniquely determine the dissipation mechanism and the energy functional of the system, since there might be various choices for them as we will see in some examples below. The perturbation in turn, perturbs the subdifferential of the energy functional and is non-variational, i.e., does not have a potential. The perturbed gradient system does, in general, not possess a *gradient flow structure*, which means that the equations can not be formulated as a generalized gradient system. However, to conclude that a concrete example does not have gradient flow structure can be fairly non-trivial, since it depends on the underlying space V, the choice of the dissipation potential as well as the energy functional. If it can be shown that concrete equations are a perturbed gradient system, this special structure of the equation can be used to characterize and equivalently describe solutions. In order to demonstrate this heuristically, we consider the classical gradient flow equation

$$u'(t) = -\nabla E(u(t)) \quad \text{in } H \text{ for a.e. } t \in (0, T), \tag{3.0.2}$$

with a FRÉCHET differentiable energy functional $E : H \to \mathbb{R}$ defined on a HILBERT space H with norm $|\cdot|$ and inner product $(\cdot, \cdot)$. The gradient of E is related with its FRÉCHET differential $\mathrm{D}E$ via the RIESZ isomorphism by $\langle DE(u), v\rangle_{H^* \times H} =$

[1] Dissipative or irreversible processes are those processes which lead to an irreversible transformation of the free energy or the entropy to thermal energy, e.g., through friction.

$(\nabla E(u), v), v \in H$. Then, by Lemma 2.3.1, an absolutely continuous curve $u : [0,T] \to H$ satisfies (3.0.2) if and only if

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}E(u(t)) = (\nabla E(u(t)), u'(t)) &= -|\nabla E(u(t)|^2 = -|u'(t)|^2 \\ &= -\frac{1}{2}|\nabla E(u(t)|^2 - \frac{1}{2}|u'(t)|^2 \quad \text{for a.e. } t \in (0,T).\end{aligned}$$

The latter equality can be replaced by the inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}E(u(t)) \leq -\frac{1}{2}|\nabla E(u(t)|^2 - \frac{1}{2}|u'(t)|^2, \tag{3.0.3}$$

since the reversed inequality holds true by Young's inequality. Integrating the latter inequality over $[0,T]$ yields the so-called *energy-dissipation balance*

$$E(u(T)) + \int_0^T \left(\frac{1}{2}|\nabla E(u(r))|^2 + \frac{1}{2}|u'(r)|^2\right) \mathrm{d}r \leq E(u(0)). \tag{3.0.4}$$

Conversely, if an absolutely continuous curve $u : [0,T] \to H$ fulfills (3.0.4), then there holds

$$\int_0^T \left(\frac{1}{2}|\nabla E(u(r))|^2 + \frac{1}{2}|u'(r)|^2 + (\nabla E(u(r)), u'(r))\right) \mathrm{d}r \leq 0$$

and by the non-negativity of the integrand,

$$\frac{1}{2}|\nabla E(u(t))|^2 + \frac{1}{2}|u'(t)|^2 = (-\nabla E(u(t)), u'(t)) \quad \text{for a.e. } t \in (0,T).$$

Again, by Lemma 2.3.1, u satisfies (3.0.2). We conclude that u is a classical solution to (3.0.2), thus is in a smooth setting entirely characterized by the energy-dissipation balance (3.0.4). Generalizing the solution concept, one can define a solution to (3.0.2) as any absolutely continuous curve fulfilling the energy-dissipation balance (3.0.4). This allows us to generalize the solution concept even further for complete metric spaces $(\mathscr{S}, d)$ based on metric formulation of gradient flows introduced by De Giorgi et al. [54]. The idea is to replace the norm of the time derivative $|u'(t)|$ by the so-called *metric derivative*

$$|u|'(t) := \lim_{s \to t} \frac{d(u(t), u(s))}{|t-s|}$$

and the norm of the gradient of the energy functional $|\nabla E(u(t)|$ by the so-called *upper gradient*

$$|\partial E|(u(t)) := \limsup_{v \to u(t)} \frac{(E(u(t) - E(v))^+}{d(v, u(t))}.$$

where $(E(u(t) - E(v))^+ = \max\{(E(u(t) - E(v)), 0\}$. In a Banach space setting, we have

$$|u'|(t) = \|u(t)\| \quad \text{and} \quad |\partial E|(u(t)) = \|\partial^\circ E(u(t))\|,$$

where $\partial^\circ E(u(t))$ denotes the subdifferential of E in $u(t)$ with the smallest norm. A solution to a gradient flow on a metric space can then be defined as any absolutely continuous curve $u \in \mathrm{AC}([0,T];\mathscr{S})$ (for which the metric derivative always exists a.e.) satisfying the energy-dissipation balance (3.0.4). Hence, in the case $\mathscr{S}$ is not a linear space, solutions to (3.0.4) do not need to satisfy any vector differentiability property. Therefore, the metric formulation of gradient flows has many advantages. An obvious point is that the spaces can have a nonlinear structure. This implies that nonlinear side conditions can be incorporated into the space or that the initial data can be quite general. A particular case for the metric space $\mathscr{S}$ has revealed to be very fruitful, not only from a theoretical point of view, but also from a numerical point of view: when $\mathscr{S}$ is the WASSERSTEIN space $(\mathscr{P}_p(X, W_p)$, $p \geq 1$, the space of all BOREL probability measures $\mu : X \to [0,1]$ on a separable Hilbert space (X,d) with finite p-moments

$$\int_X d^p(x,x_0)\,\mathrm{d}\mu(x) < \infty \quad \text{for some } x_0 \in X,$$

endowed with the p^{th} WASSERSTEIN distance

$$W_p(\mu,\nu) := \left(\inf_{\gamma\in\Gamma(\mu,\nu)} \int_{X\times X} d(x,y)\,\mathrm{d}\gamma(x,y) \right)^{\frac{1}{p}}, \quad \mu,\nu \in \mathscr{P}_p(X),$$

where $\Gamma(\mu,\nu)$ denotes the set of all couplings of μ and ν. It has been shown that various partial differential equations can be viewed as a gradient flow in a WASSERSTEIN space, e.g., the FOKKER–PLANCK equation, the porous medium equation, the LANDAU equation, the BOLTZMANN equation, and other equations of diffusion type, which was first pointed out by OTTO in a series of seminal works [96, 128–130], see also [11, 41, 80]. Besides, it has also been used to prove and improve functional inequalities as the SOBOLEV, GAGLIARDO–NIRENBERG, BRUNN–MINKOWSKI, PRÉKOPA–LEINDLER, isoperimetric inequality and other inequalities, see, e.g., [1, 46–48, 131] We refer the reader to the monograph AMBROSIO et al. [10] for a comprehensive presentation of the theory of gradient flows in metric spaces and in the space of probability measures and to VILLANI [158] for a description of the interplay with the theory of optimal transportation.

The following simple example illustrates how the dissipation mechanism as well as the driving functional can be chosen in multiple and non-trivial ways in order to describe the same equation as a gradient system. We consider the homogeneous diffusion or heat equation

$$\begin{aligned} \partial_t u(x,t) &= a\Delta u(x,t), && (x,t) \in \mathbb{R}^d \times (0,T), \\ u(x,0) &= u_0(x), && x \in \mathbb{R}^d, \end{aligned}$$

which is a model to describe the heat in a homogeneous and isotropic medium evolving over time for a given initial heat distribution u_0 and possible boundary conditions, where $u(x,t)$ describes the temperature at point $x \in \mathbb{R}^d, d \in \mathbb{N}$, and time $t > 0$, and $a > 0$ is the thermal conductivity. It is well-known that the heat equation is the gradient flow of the DIRICHLET energy $E : \mathrm{L}^2(\mathbb{R}^d) \to [0, +\infty]$ defined by

$$E(v) := \begin{cases} \int_{\mathbb{R}^d} |\nabla v(x)|^2 \,\mathrm{d}x & \text{if } \nabla v \in \mathrm{L}^2(\mathbb{R}^d), \\ +\infty & \text{otherwise,} \end{cases}$$

with respect to the L^2-metric, which is therefore the dissipation mechanism, see, e.g., AMBROSIO et al. [10, Remark 2.3.9., p. 49]. Another choice for the energy functional is $E : \mathrm{H}^{-1}(\mathbb{R}^d) \to [0, +\infty]$ with

$$E(v) := \begin{cases} \int_{\mathbb{R}^d} v^2(x) \,\mathrm{d}x & \text{if } v \in \mathrm{L}^2(\mathbb{R}^d), \\ +\infty & \text{otherwise,} \end{cases}$$

on the HILBERT space $\mathrm{H}^{-1}(\mathbb{R}^d)$ which formally leads to the equation

$$(-\Delta)^{-1}\partial_t u(x,t) = au(x,t), \quad (x,t) \in \mathbb{R}^d \times (0,T),$$

see, e.g., ROSSI & SAVARÉ [143]. Furthermore, JORDAN, KINDERLEHRER & OTTO [96] have shown that the more general FOKKER–PLANCK equation[2]

$$\begin{aligned} \partial_t \varrho(x,t) &= \nabla \cdot (\nabla\Psi(x)\varrho(x,t)) + a\Delta\varrho(x,t), && (x,t) \in \mathbb{R}^d \times (0,T), \\ \varrho(x,0) &= \varrho_0(x), && x \in \mathbb{R}^d, \end{aligned}$$

which describes the time evolution of a probability density functional under drift and diffusion, is the gradient flow of the free energy functional $F : \mathscr{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ with $F(\varrho) = E(\varrho) + aS(\varrho)$ with respect to the WASSERSTEIN metric, where E and S are the energy functional and the negative of the GIBBS–BOLTZMANN entropy functional, respectively, and are, on their effective domains, given by

$$E(\varrho) := \int_{\mathbb{R}^d} \Psi(x)\varrho(x) \,\mathrm{d}x,$$

and

$$S(\varrho) := \int_{\mathbb{R}^d} \varrho(x) \ln(\varrho(x)) \,\mathrm{d}x,$$

and taking the value infinity otherwise. Henceforth, we will simply give the values of the functionals on the effective domain and implicitly assume they are infinity otherwise. In the case $\Psi = 0$, F reduces to S which becomes then the driving functional for the heat equation. These are only three of many examples on how one can rewrite the heat equation gradient flow with different dissipation mechanisms and energy functionals. Each formulation is preferable depending on the system to be modelled. In each case, however, all choices share a common characteristic; the energy functional serves as a LYAPUNOV functional for the gradient flow equation,

[2] The FOKKER–PLANCK equation is also known as the KOLMOGOROV forward equation or the SMOLUCHOWSKI equation, and is also referred to as convection–diffusion equation when the equation models the transfer of mass, energy, temperature or other physical quantities through diffusion and convection.

i.e., the solution to the pure gradient flow equation minimizes the energy functional along the time-trajectory. This fact easily follows from the inequality (3.0.3), which shows that the time derivative of $E(u(t))$ is non-positive. This is still true for the the so-called *generalized gradient flow* equation referred to the equation

$$\mathrm{D}\Psi_{u(t)}(u'(t)) = -\mathrm{D}\mathcal{E}(u(t)) \quad \text{in } V^*,$$

which is also called *force balance*. In the nonsmooth setting, we will replace the derivatives $\mathrm{D}\Psi_{u(t)}$ and $\mathrm{D}\mathcal{E}$ by their subdifferentials $\partial\mathcal{E}$ and $\partial\Psi_{u(t)}$. For the purpose of illustration, we assume for a moment that $\mathcal{E}$ and Ψ are sufficiently smooth. The crucial assumptions on the dissipation potential are the convexity and lower semicontinuity. Then, Lemma 2.3.1 allows us to reformulate this equation equivalently by the *rate equation*

$$u'(t) = \mathrm{D}\Psi^*_{u(t)}(-\mathrm{D}\mathcal{E}(u(t))) \quad \text{in } V,$$

or the *power balance*

$$\Psi_{u(t)}(u'(t)) + \Psi^*_{u(t)}(-\mathrm{D}\mathcal{E}(u(t)) = \langle -\mathrm{D}\mathcal{E}(u(t), u'(t)\rangle_{V^*\times V} \quad \text{in } \mathbb{R}. \tag{3.0.5}$$

A very important example is given by the quadratic case $\Psi_u(v) = \frac{1}{2}\langle G(u)v, v\rangle_{V^*\times V}$ for which the conjugate functional is given by $\Psi^*_u(\xi) = \frac{1}{2}\langle \xi, K(u)\xi\rangle_{V^*\times V}$, where $G(u) : V \to V^*$ is a linear, bounded, symmetric and positive definite operator for each $u \in V$ and $G(u) = K^{-1}(u), u \in V$. In this case, the force balance and the rate equation are given by

$$G(u(t))u'(t) = -\mathrm{D}\mathcal{E}(u(t)) \quad \text{and} \quad u'(t) = -K(u(t))\mathrm{D}\mathcal{E}(u(t)),$$

respectively, which are also known as BIOT's equation and ONSAGER's or GINZBURG–LANDAU equation, see [89, 119].

Assuming the dissipation potential is non-negative and satisfies $\Psi_u(0) = 0$, then by Lemma 2.3.2, there holds $\Psi^*_u \geq 0$, so that equation (3.0.5) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(u(t)) = \langle -D\mathcal{E}(u(t)), u'(t)\rangle = -\Psi(u'(t)) - \Psi^*(-D\mathcal{E}_t(u(t))) \leq 0,$$

i.e., the energy decreases along solutions. Thus, as previously seen, the energy functional serves again as a LYAPUNOV functional for the generalized gradient flow equation. However, this fact does not, in general, hold true in perturbed gradient systems or in gradient systems where the energy functional is explicitly time-dependent, which causes additional external forces. Therefore, the time-trajectory of the energy along the solution of perturbed gradient systems is not minimizing, which is illustrated in Figure 3.1. Nevertheless, the main idea of reformulating the gradient system as a scalar-valued equation still applies for perturbed gradient systems. This can be seen from the following reasoning: let $u : [0, T] \to V$ be an absolutely continuous curve satisfying the perturbed gradient flow equation

$$\mathrm{D}\Psi(u'(t)) = -\mathrm{D}\mathcal{E}(u(t)) + B(t, u(t)) \quad \text{in } V^* \text{for a.e. } t \in (0, T). \tag{3.0.6}$$

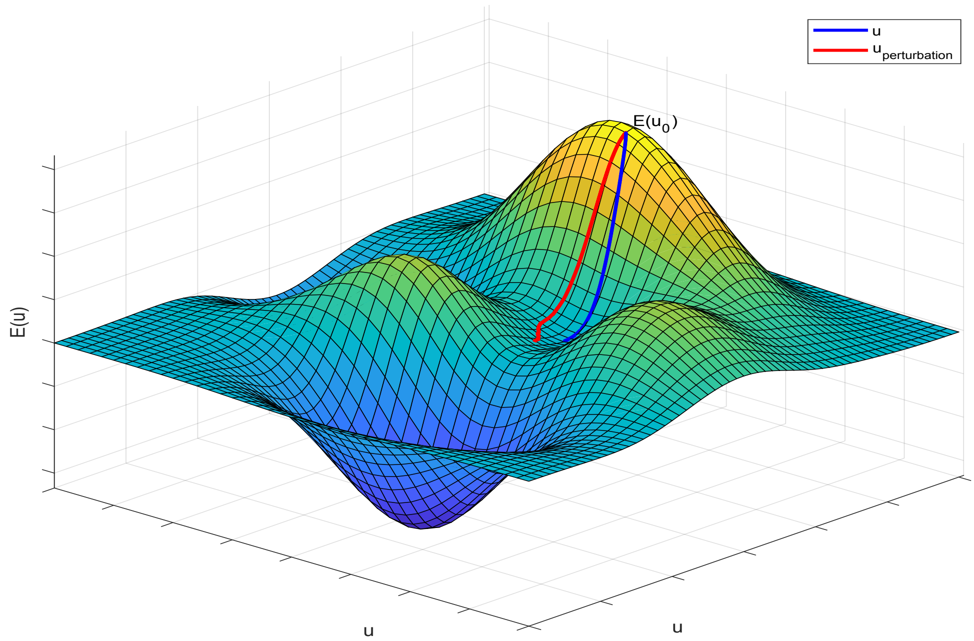

Figure 3.1: Evolution of the energy along solutions u and $u_{\text{perturbation}}$ to the pure and the perturbed gradient flow equation, respectively.

Then, from the equivalent relations in Lemma 2.3.1, the equation (3.0.6) is equivalent to the scalar equation

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(u(t)) &= \langle -D\mathcal{E}(u(t)), u'(t)\rangle \\ &= -\langle B(t,u(t)) - D\mathcal{E}(u(t)), u'(t)\rangle + \langle B(t,u(t)), u'(t)\rangle \\ &= -\Psi(u'(t)) - \Psi^*(B(t,u(t)) - D\mathcal{E}(u(t))) \\ &\quad + \langle B(t,u(t)), u'(t)\rangle \quad \text{a.e. in } (0,T),\end{aligned}$$

and after integration

$$\begin{aligned}&\mathcal{E}(u(t)) + \int_s^t (\Psi(u'(r)) + \Psi^*(B(r,u(r)) - D\mathcal{E}(u(r))))\,\mathrm{d}r \\ &= \mathcal{E}(u(s)) + \int_s^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r\end{aligned} \tag{3.0.7}$$

for all $s,t \in [0,T]$. Hence, again the equation (3.0.6) is in a sufficiently smooth setting equivalent to the energy-dissipation balance (3.0.7).

The question arises why it is interesting to study perturbed gradient systems. Even though it has been shown that gradient flows cover a large class of differential equations, there are still enough important equations that do not possess the gradient flow structure. The probably most famous equations of this class are the Navier-Stokes equations in fluid dynamics, which are for incompressible fluids given by

$$\begin{cases}\partial_t \boldsymbol{u}(\boldsymbol{x},t) + (\boldsymbol{u}(\boldsymbol{x},t)\cdot\nabla)\boldsymbol{u}(\boldsymbol{x},t) - \nu\Delta\boldsymbol{u}(\boldsymbol{x},t) + \nabla p(\boldsymbol{x},t) = \boldsymbol{f}(\boldsymbol{x},t) & \text{on } \Omega\times(0,T), \\ \nabla\cdot\boldsymbol{u}(\boldsymbol{x},t) = 0 \qquad \text{on } \Omega\times(0,T),\end{cases}$$

Choosing $V := \mathrm{L}^2_\sigma(\Omega)^d$ the closure of the test functions $\mathrm{C}^\infty_c(\Omega)^d$ that are divergence free with respect to the $\mathrm{L}^2(\Omega)^d$ norm, we obtain for the energy functional $\mathcal{E}^\varepsilon : V \to [0, +\infty]$ and the dissipation potential $\Psi : V \to \mathbb{R}$

$$\mathcal{E}^\varepsilon(\boldsymbol{u}) := \frac{\nu}{2} \int_\Omega |\nabla \boldsymbol{u}(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x} \quad \text{and} \quad \Psi(\boldsymbol{v}) := \int_\Omega |\boldsymbol{v}(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x}.$$

The perturbation is (formally) given by

$$\langle B(\boldsymbol{u}), \boldsymbol{w} \rangle_{\mathrm{L}^2} = \int_\Omega (\boldsymbol{u}(\boldsymbol{x}) \cdot \nabla) \boldsymbol{u}(\boldsymbol{x}) \cdot \boldsymbol{w}(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}.$$

We refer the reader to Temam [157] for a detailed discussion of the Navier–Stokes equations.
Even in finite dimensions, one can easily construct equations which do not possess the gradient flow structure. Consider, e.g., the coupled system of linear ordinary differential equations

$$\begin{aligned} u_1'(t) &= -u_1(t) + (\eta + \lambda) u_2(t) \\ u_2'(t) &= -u_2(t) + (\eta - \lambda) u_1(t) \end{aligned}$$

with $\eta, \lambda \in \mathbb{R}$ and $\lambda \neq 0$. The dissipation potential and energy functional are given by

$$\Psi(v_1, v_2) = v_1^2 + v_2^2 \quad \text{and} \quad \mathcal{E}(u_1, u_2) = \frac{1}{2}(u_1^2 + u_2^2) - \eta u_1 u_2,$$

respectively. As a result, the perturbation is then given by the term

$$B(t, u_1, u_2) = B(u_1, u_2) = \lambda \begin{pmatrix} u_2 \\ -u_1 \end{pmatrix}.$$

Rewriting the coupled system in the form

$$\begin{aligned} \begin{pmatrix} u_1'(t) \\ u_2'(t) \end{pmatrix} &= \begin{pmatrix} -1 & \eta \\ \eta & -1 \end{pmatrix} \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} + \begin{pmatrix} 0 & \lambda \\ -\lambda & 0 \end{pmatrix} \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} \\ &= \mathrm{D}_{(u_1,u_2)} \mathcal{E}(u_1(t), u_2(t)) + B(u_1(t), u_2(t)) \quad \text{on } (0, T), \end{aligned}$$

we can see that the system can not be cast into a gradient flow formulation. A more physical example is the rescaled fourth order parabolic Swift–Hohenberg equation on the circle $\mathbb{S} := \mathbb{R}/2\pi\mathbb{Z}$ considered in Mielke [118] and given by

$$\partial_t u(x,t) = -\frac{1}{\varepsilon^2}\left(1 + \varepsilon^2 \partial_x^2\right)^2 u(x,t) + \mu u(x,t) + \beta \varepsilon \partial_x u(x,t) - u^3(x,t) \quad \text{on } \mathbb{S} \times (0, T),$$

which is a model to describe pattern formations in a self-organizing nonlinear system where $\varepsilon > 0$ plays the role of a bifurcation parameter, see, e.g., [104, 117–119, 156] for more details and different applications of this model. The equation is an exact gradient flow on the space $V := \mathrm{L}^2(\mathbb{S})$ if $\beta = 0$ and can be treated as a perturbed gradient flow otherwise, see Mielke [118]. In the latter case, the energy functional $\mathcal{E}^\varepsilon : V \to (-\infty, +\infty]$ and the dissipation potential $\Psi : V \to \mathbb{R}$ are given by

$$\mathcal{E}^\varepsilon(u) := \int_{\mathbb{S}} \frac{1}{\varepsilon^2} \left(\left(u(x) + \varepsilon^2 u''(x) \right)^2 - \frac{\mu}{2} u^2(x) - \frac{1}{4} u^4(x) \right) \mathrm{d}x$$

and

$$\Psi(v) := \int_{\mathbb{S}} v^2(x) \, \mathrm{d}x,$$

whereas the perturbation is (formally) given by $\langle B^\varepsilon(u), w \rangle_{\mathrm{L}^2} = \int_{\mathbb{S}} \varepsilon\beta \partial_x u(x) w(x) \, \mathrm{d}x$.

3.1 Variational approximation scheme

Showing the existence of strong solutions, i.e., functions $u \in \mathrm{AC}([0,T];V)$ to the perturbed gradient system (3.0.1), is based on the idea of discretizing the equation (3.0.1) in time via a semi-implicit Euler method. More precisely, we discretize the terms coming from the energy functional and the dissipation potential implicitly in u and u', while the perturbation will be discretized explicitly. This approach is advantageous for our purposes, since this allows us to construct a solution to the discrete problem by a variational approximation scheme even though the system (3.0.1) does not possess the gradient flow structure. To elaborate on this, we define for $N \in \mathbb{N}$ and the associated step size $\tau := \frac{T}{N}$ the partition of the time interval $[0,T]$

$$\begin{aligned} &\mathcal{P}_\tau = \{0 = t_0 < t_1 < \cdots t_N = T\}, \\ &\tau = t_n - t_{n-1}, \end{aligned}$$

where we have suppressed the dependence of t_n on the step size τ for notational convenience. Then, the discretized equation of (3.0.1) reads

$$\partial\Psi_{U_\tau^{n-1}} \left(\frac{U_\tau^n - U_\tau^{n-1}}{\tau} \right) + \partial\mathcal{E}_{t_n}(U_\tau^n) \ni B(t_n, U_\tau^{n-1}), \quad n = 1, \cdots, N, \tag{3.1.1}$$

where the values $U_\tau^n \approx u(t_n)$ for $n = 0, \ldots, N$ shall approximate the values of the exact solution u at time $t = t_n$, and are to be determined. If we assume the energy functional and the dissipation potential to be (Fréchet) differentiable, the differential inclusion (3.1.1) becomes the equation

$$\mathrm{D}\Psi_{U_\tau^{n-1}} \left(\frac{U_\tau^n - U_\tau^{n-1}}{\tau} \right) + \mathrm{D}\mathcal{E}_{t_n}(U_\tau^n) = B(t_n, U_\tau^{n-1}), \quad n = 1, \cdots, N. \tag{3.1.2}$$

This choice of discretization has several advantages. First, the values U_τ^n which for a given U_τ^{n-1} are to be determined, can be characterized as a solution to the Euler-Lagrange equation (3.1.2) associated to the mapping

$$v \mapsto \Phi(\tau, t_{n-1}, U_\tau^{n-1}, B(t_n, U_\tau^{n-1}); v), \tag{3.1.3}$$

where

$$\Phi(r, t, u, w; v) = r\Psi_u \left(\frac{v - u}{r} \right) + \mathcal{E}_{t+r}(v) - \langle w, v \rangle \tag{3.1.4}$$

for $r \in \mathbb{R}^{>0}, t \in [0,T)$ with $r+t \in [0,T]$, $u, v \in V$ and $w \in V^*$. This leads to the so-called *variational approximation scheme*

$$\begin{cases} U_\tau^0 \text{ is given; whenever } U_\tau^1, U_\tau^2, \ldots, U_\tau^{n-1} \text{ are known,} \\ \text{find} \quad U_\tau^n \in \operatorname{argmin}_{v\in V} \Phi(\tau, t_{n-1}, U_\tau^{n-1}, B(t_n, U_\tau^{n-1}); v) \end{cases} \tag{3.1.5}$$

for $n = 1, \ldots, N,$.

The solvability of the variational approximation scheme can be established by virtue of the direct methods of calculus of variation, i.e, those methods where the solvability of the Euler–Lagrange equation relies on the minimization of (3.1.3) under relatively mild assumptions on the functionals $\mathcal{E}_t$ and Ψ_u. The solvability of the discrete problem by minimization would fail to accomplish with a full implicit discretization in time, since the perturbation is not explicitly supposed to be variational; the equation (3.0.1) does not possess the so-called gradient flow structure. In order to solve the discrete problem in the case of a full implicit discretization, one would have to use fixed point arguments for set valued maps as the fixed point theorem of Kakutani[3], which is a set-valued version of the fixed point theorem of Schauder for which the compactness of the images of the set-valued operator has to be assumed. This, however, is not satisfied, in general, by the subdifferential operator $\partial\mathcal{E}_t$, which might be unbounded as we will see in the applications. Besides, a fully implicit discretization would not be useful, since one would not be able to obtain appropriate a priori estimates without making further assumptions on the growth of Ψ and the subgradients of $\partial\mathcal{E}_t$. However, we obtain a priori estimates immediately when we solve the semi-implicit discretized problem by the direct method. This leads us to the last and most important point which is the equivalence between the force balance (3.0.1) and the energy-dissipation balance (3.0.7). As we mentioned before, the main idea of our approach is based on the aforementioned equivalence which allows us to infer the solvability of the perturbed gradient flow equation by proving the energy-dissipation balance.

3.2 Topological assumptions and main result

In this section, we collect all assumptions for the system $(V, \mathcal{E}, \Psi, B)$ to ensure the existence of a solution. We refer to the assumptions by (3.E), (3.Ψ) and (3.B) for the energy functional, the dissipation potential and the perturbation, respectively.

We start with collecting the assumptions for Ψ and emphasize that in contrast to Bacho et al. [21] and Mielke et al. [122], we will not suppose that for all $w_1, w_2 \in \partial\Psi_u(v)$ there holds $\Psi_u^*(w_1) = \Psi_u^*(w_2)$, where Ψ_u^* denotes the Legendre–Fenchel transformation or the conjugate of Ψ_u. We circumvent this condition by regularizing the dissipation potential via the p-Moreau–Yosida regularization and let the regularization parameter ε afterwards tend to zero. To do so, we need to

[3] see, e.g., Aubin & Frankowska [19, Theorem 3.2.3, p. 87].

verify that the following conditions imposed on Ψ_u are also inherited by the MOREAU–YOSIDA regularization, which will be shown in Lemma 3.2.4. Before we collect the assumptions, we define for notational convenience $\mathcal{G}(u) := \sup_{t\in[0,T]} \mathcal{E}_t(u), u \in V$. Furthermore, we denote with $D := \mathrm{dom}(\mathcal{E})$ the time-independent effective domain of $\mathcal{E}_t$, see Condition (3.Ea).

(3.Ψa) **Dissipation potential.** For all $u \in D$, let $\Psi_u : V \to [0, +\infty)$ be lower semicontinuous and convex with $\Psi_u(0) = 0$.

(3.Ψb) **Superlinearity.** The functionals Ψ_u and Ψ_u^* are superlinear, uniformly with respect to u on sublevels of $\mathcal{G} = \sup_{t\in[0,T]} \mathcal{E}_t$, i.e., for all $R > 0$, there holds

$$\lim_{\|\xi\|_*\to+\infty} \frac{1}{\|\xi\|_*}\Big(\inf_{\substack{u\in V\\ \mathcal{G}(u)\leq R}} \Psi_u^*(\xi)\Big) = \infty, \qquad \lim_{\|v\|\to+\infty} \frac{1}{\|v\|}\Big(\inf_{\substack{u\in V\\ \mathcal{G}(u)\leq R}} \Psi_u(v)\Big) = \infty.$$

(3.Ψc) **Mosco-convergence** The state dependence $u \mapsto \Psi_u$ on sublevels of $\mathcal{E}$ is continuous in the sense of MOSCO-convergence, i.e., for all $R > 0$ and all sequences $(u_n)_{n\in\mathbb{N}} \subset V$ with $u_n \to u \in V$ as $n \to \infty$ and $\sup_{n\in\mathbb{N}} \mathcal{G}(u_n) \leq R$, there holds $\Psi_{u_n} \xrightarrow{\mathrm{M}} \Psi_u$.

Before we proceed with the assumptions on the energy functional, we make some important remarks.

Remark 3.2.1

i) We showed in Lemma 2.3.2 that the Conditions (3.Ψa) and (3.Ψb) together imply that the conjugate Ψ^* is lower semicontinuous, convex, and non-negative on V^* with $\Psi_u^*(0) = 0$ for all $u \in D$. Furthermore, it is easy to show by contradiction that Condition (3.Ψb) implies that Ψ_u^* is finite everywhere, i.e., $\mathrm{dom}(\Psi_u^*) = V^*$ for all $u \in D$.

ii) Condition (3.Ψb) is equivalent to say that for all $R > 0$ and $\gamma > 0$, there exists $K_1, K_2 > 0$ such that

$$\Psi_u(v) \geq \gamma\|v\|, \quad \Psi_u^*(\xi) \geq \gamma\|\xi\|_*$$

for all $u \in D$ with $G(u) \leq R$ and all $v \in V$ and $\xi \in V^*$ with $\|v\| \geq K_1$ and $\|\xi\|_* \geq K_2$, respectively.

iii) Condition (3.Ψc) implies in particular that the mapping $(u, v) \mapsto \Psi_u(v)$ is strongly-weakly lower semicontinuous on $J_\alpha \times V$ for each $\alpha \in \mathbb{R}$, where $J_\alpha := \{u \in V : \mathcal{G}(u) \leq \alpha\}$. Furthermore, Lemma 2.4.3 implies that $\Psi_{u_n} \xrightarrow{\mathrm{M}} \Psi_u$ if and only if for all $v \in V$ and $\xi \in V^*$

$$a) \quad \Psi_u(v) \leq \inf\{\liminf_{n\to\infty} \Psi_{u_n}(v_n) : v_n \rightharpoonup v \text{ in } V\},$$
$$b) \quad \Psi_u^*(\xi) \leq \inf\{\liminf_{n\to\infty} \Psi_{u_n}^*(\xi_n) : \xi_n \rightharpoonup \xi \text{ in } V\}.$$

This in turn implies that the mapping $(u, \xi) \mapsto \Psi_u^*(\xi)$ is strongly-weakly lower semicontinuous on $J_\alpha \times V^*$ for each $\alpha \in \mathbb{R}$.

iv) We note that we can consider more general time-dependent dissipation potentials $\Psi_u : [0,T] \times V \to [0,+\infty)$ by assuming that Ψ_u are normal integrands for all $u \in D$, the functional $\Psi_u(t,\cdot)$ satisfies Condition (3.Ψa) and (3.Ψc) for every $t \in [0,T]$ and Condition (3.Ψb) uniformly in $t \in [0,T]$.

Now, we present the assumptions for the energy functional.

(3.Ea) **Lower semicontinuity.** For all $t \in [0,T]$, let the functional $\mathcal{E}_t : V \to [0,+\infty]$ be proper and lower semicontinuous and and have a time-independent effective domain $D := \mathrm{dom}(\mathcal{E}_t)$ for all $t \in [0,T]$.

(3.Eb) **Compactness.** For all $t \in [0,T]$, $\mathcal{E}_t$ has compact sublevel sets in V, i.e., for all $t \in [0,T]$ and $R \geq 0$, the set $\{u \in V : \mathcal{E}_t(u) \leq R\}$ is compact in V.

(3.Ec) **Control of the time derivative** For all $u \in U$, the mapping $t \mapsto \mathcal{E}_t(u)$ is differentiable and its derivative $\partial_t\mathcal{E}_t$ is controlled by $\mathcal{E}_t$, i.e., there exists $C_1 > 0$ such that

$$|\partial_t\mathcal{E}_t(u)| \leq C_1\mathcal{E}_t(u) \quad \text{for all } t \in (0,T) \text{ and } u \in V. \tag{3.2.1}$$

(3.Ed) **Chainrule.** For every absolutely continuous curves $u \in \mathrm{AC}([0,T];V)$ and every integrable functions $\xi \in \mathrm{L}^1(0,T;V^*)$ such that

$$\sup_{t\in[0,T]} |\mathcal{E}_t(u(t))| < +\infty, \quad \xi(t) \in \partial\mathcal{E}_t(u(t)) \quad \text{a.e. in } (0,T),$$
$$\int_0^T \Psi_{u(t)}(u'(t))\,\mathrm{d}t < +\infty \quad \text{and} \quad \int_0^T \Psi^*_{u(t)}(\xi(t))\,\mathrm{d}t < +\infty,$$

the mapping $t \mapsto \mathcal{E}_t(u(t))$ is absolutely continuous on $[0,T]$ and there holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) \geq \langle \xi(t), u'(t)\rangle + \partial_t\mathcal{E}_t(u(t)) \quad \text{a.e. in } (0,T). \tag{3.2.2}$$

(3.Ee) **Strong-weak closedness.** For all $t \in [0,T]$ and all sequences $(u_n,\xi_n)_{n\in\mathbb{N}} \subset V\times V^*$ with $\xi_n \in \partial\mathcal{E}_t(u_n)$ such that

$$u_n \to u \in V, \quad \xi_n \rightharpoonup \xi \in V^*, \quad \mathcal{E}_t(u_n) \to \mathscr{E} \in \mathbb{R} \quad \text{and} \quad \partial_t\mathcal{E}_t(u_n) \to p \in \mathbb{R}$$

as $n \to \infty$, the following relations hold

$$\xi \in \partial\mathcal{E}_t(u), \quad p \leq \partial_t\mathcal{E}_t(u) \quad \text{and} \quad \mathscr{E} = \mathcal{E}_t(u).$$

Before we continue collecting the assumptions for the dissipation potential, we discuss the conditions.

Remark 3.2.2

i) From the estimate (3.2.1), we obtain after integration

$$e^{-C_1|t-s|}\mathcal{E}_s(u) \le \mathcal{E}_t(u) \le e^{C_1|t-s|}\mathcal{E}_s(u) \quad \text{for all } s,t \in [0,T], \tag{3.2.3}$$

and in particular

$$\sup_{t\in[0,T]} \mathcal{E}_t(u) \le e^{C_1 T} \inf_{t\in[0,T]} \mathcal{E}_t(u) \quad \text{for all } u \in D.$$

ii) The compactness condition (3.Eb) in particular implies that there exists a constant $C_0 \in \mathbb{R}$ such that the energy functional is bounded from below by that constant, i.e.,

$$\mathcal{E}_t(u) \ge C_0 \quad \text{for all } u \in V,\ t \in [0,T],$$

see, e.g. Ambrosio et al. [10, Remark 2.1.1, p. 43]. We assume without loss of generality that $C_0 = 0$, since every potential is determined uniquely up to a constant.

iii) It can be shown in the exact same manner to Proposition 4.2 in Mielke et al. [122] that under the Assumptions (3.Ψa)-(3.Ψc), (3.Ea), and (3.Ee), the variational sum rule holds: if for $u_0 \in V$, $r > 0$, and $t \in [0,T)$ such that $r + t \le T$, the point $u \in V$ is a global minimizer of $v \mapsto r\Psi\left(\frac{v-u_0}{r}\right) + \mathcal{E}_{r+t}(v)$, then

$$\text{there exists } \xi \in \partial\mathcal{E}_t(u) \quad \text{such that} \quad w - \xi \in \partial\Psi_{u_0}\left(\frac{u-u_0}{r}\right) \tag{3.2.4}$$

or equivalently $w \in \partial\Psi_{u_0}\left(\dfrac{u-u_0}{r}\right) + \partial\mathcal{E}_{t+r}(u)$. The variational sum rule as it is stated for convex functionals in Lemma 2.2.7, is, in general, not true for non-convex and non-differentiable functionals.

iv) The strong-weak closedness of the graph of $\partial\mathcal{E}_t$ is already satisfied when $(t,u) \mapsto \mathcal{E}_t \in C^1([0,T]\times V)$ or when $\mathcal{E}_t = \mathcal{E}$ is proper, lower semicontinuous, and λ-convex. While the former follows immediately from the continuity of $D\mathcal{E}_t$ and $\partial_t\mathcal{E}_t$, the latter can be seen as follows: from the characterization of the subdifferential of λ-convex functions, there holds

$$\mathcal{E}(u_n) \le \mathcal{E}(v) + \langle \xi_n, u_n - v\rangle + \lambda\|u_n - v\|^2 \quad \text{for all } n \in \mathbb{N}.$$

Then, we obtain from the lower semicontinuity of $\mathcal{E}$ that

$$\begin{aligned}\mathcal{E}(u) \le \mathscr{E} \le \liminf_{n\to\infty} \mathcal{E}(u_n) &\le \liminf_{n\to\infty} \left(\mathcal{E}(v) + \langle \xi_n, u_n - v\rangle + \lambda\|u_n - v\|^2\right) \\ &= \mathcal{E}(v) + \langle \xi, u - v\rangle + \lambda\|u - v\|^2\end{aligned}$$

and hence $\xi \in \partial\mathcal{E}(u)$. Choosing $v = u$ yields $\mathscr{E} = \mathcal{E}(u)$.

Finally, we state the assumptions for perturbation B.

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

(3.Bb) **Control of the growth.** There exist a real number $p > 1$ and constants $c_1 \in (0,1)$ and $\beta > 0$ such that

$$c_1\, \Psi_u^{\varepsilon,*}\left(\frac{B(t,u)}{c_1}\right) \leq \beta\Big(1 + \mathcal{E}_t(u)\Big) \quad \text{for all } u \in D,\ t \in [0,T] \text{ and } \varepsilon \geq 0,$$

where $\Psi_u^{\varepsilon,*}$ denotes the conjugate functional of the p-Moreau–Yosida regularization of Ψ_u.

3.2.1 Discussion of the assumptions

Having collected the assumptions on the perturbed gradient system $(V, \mathcal{E}, \Psi, B)$, we want to discuss several conditions more in detail apart from the assertions and implications made in the remarks. In particular, we want to discuss the practical meaning of the rather abstract and quite general assumptions and provide sufficient conditions for them to hold true.

The advantage of the multivalued setting is that the domain of the functionals Ψ_u and $\mathcal{E}_t$ can, in general, be subsets of different spaces which share a common dense subspace. This stems from the fact that the operators can be unbounded. For example, if Ψ_u and the functional $\tilde{\mathcal{E}}_t$ are finite on the Banach spaces V and W such that $V \cap W$ is densely embedded in V and W, respectively, then one can extend $\tilde{\mathcal{E}}_t$ to the whole V by setting

$$\mathcal{E}_t(v) = \begin{cases} \tilde{\mathcal{E}}_t(v) & \text{if } v \in \mathrm{dom}(\tilde{\mathcal{E}}_t) \cap V, \\ +\infty, & \text{otherwise.} \end{cases}$$

However, the functional $\mathcal{E}_t$ might not be coercive anymore on V, which is not a problem if one assumes the coercivity of the sum $\Psi_u + \tilde{\mathcal{E}}_t$. Since we have assumed the coercivity of Ψ_u and $\mathcal{E}_t$ separately, this case is not covered by the present work. Indeed, it has been shown in Mielke et al. [122] that this case can easily be included under the stronger assumptions that the mapping $t \mapsto \Psi_u(tv)$ is differentiable in $t = 1$ for all $u \in D, v \in V$, which has been addressed in this work by regularizing Ψ_u. The problem which occurs in the present work is that the condition $\Psi_u + \mathcal{E}_t$ to have compact sublevels in V is, in general, not satisfied by $\Psi_u^\varepsilon + \mathcal{E}_t$, with Ψ_u^ε being the Moreau–Yosida regularization of Ψ_u. However, if we make the further assumption that Ψ_u has p-growth, then we can solve this problem by taking the p-Moreau–Yosida regularization of Ψ_u which maintains the p-growth for Ψ_u^ε. The proof of this case would require minor changes of the proof presented here, but will not be discussed more in detail. We refer the reader to Chapter 5 and 6, where we study evolution inclusions of second order where Ψ_u and $\mathcal{E}_t$ are defined on different spaces.

Ad (3.Ψ). In case the dissipation potentials are state-independent, i.e., $\Psi_u = \Psi$ for all $u \in D$, then the Condition (3.Ψc) already follows from Condition (3.Ψa): the liminf estimate *a*) in Definition 2.4.1 reduces to a weak lower semicontinuity Ψ which is clearly satisfied, since Ψ is convex and lower semicontinuous and thus weakly lower semicontinuous on V. The limsup estimate *b*) follows from the fact that by Lemma 2.3.2 Ψ is continuous over its effective domain, which is by assumption $\mathrm{dom}(\Psi) = V$. The prototypical example for state-independent dissipation potentials is

$$\Psi(v) = \int_\Omega \left(\frac{1}{p} |v(\boldsymbol{x})|^p + |v(\boldsymbol{x})| \right) \mathrm{d}\boldsymbol{x}$$

defined on the Lebesgue space $\mathrm{L}^p(\Omega)$ with $p \in (1, +\infty)$. Although Conditions (3.Ψa) and (3.Ψb) are satisfied by many functions with anisotropic and nonpolynomial growth, it can not be treated in this setting, since the corresponding Orlicz spaces are, in general, neither separable nor reflexive, see, e.g., Skaff [153, 154] and to the monograph Rao & Ren [136] for a detailed treatise of the theory of Orlicz spaces. For state-dependent dissipation potentials, an example for a general case is given by

$$\Psi_u(v) = \int_\Omega \psi(u(\boldsymbol{x}), v(\boldsymbol{x}))\, \mathrm{d}\boldsymbol{x}$$

defined on the Lebesgue space $\mathrm{L}^p(\Omega)$ with $p \in (1, +\infty)$ where ψ satisfies certain growth and continuity conditions, see Chapter 4, where we discuss this more in detail. As we mentioned in Remark 3.2.1 *iv*), we can allow time-dependent dissipation potentials $\Psi_u : [0, T] \times V \to [0, +\infty)$. However, we will not consider this for simplicity. See Akagi [4], where doubly nonlinear evolution equations have been investigated with time-dependent dissipation potentials.

Ad (3.E). The assumption that the energy functional $\mathcal{E}_t$ has time-independent effective domain is a natural consequence of the control of its time derivative $\partial_t \mathcal{E}_t$ (Condition (3.Ec)). However, this does not imply that the domain of the subdifferential $\partial \mathcal{E}_t$ is time-independent, because it is generally not. A result on abstract evolution inclusions where the domain of the energy functional is time-dependent has been shown in Yamada [161] and Ôtani [126] with applications to the Navier–Stokes equations in a time-dependent bounded domain.

A sufficient condition for $\mathcal{E}_t$ to have compact sublevel sets is for its domain D to be contained in a Banach space U which is compactly embedded into V. This means in practice that the energy functional contains higher spatial derivatives than the dissipation potential, so that the compact embedding $U \overset{c}{\hookrightarrow} V$ is given by the Rellich–Kondrachov theorem or other compactness theorems. As we mentioned before, we do not impose any convexity condition on $\mathcal{E}_t$. Instead, we have to impose two conditions, which are in particular fulfilled by convex functionals: the chainrule condition (3.Ee) and the closedness condition (3.Ee). These conditions are not only fulfilled by convex functionals, but also by λ-convex functionals, functionals of the form $\mathcal{E}_t = \mathcal{E}_t^1 - \mathcal{E}_t^2$ where $\mathcal{E}_t^1$ is convex and $\mathcal{E}_t^2$ is either convex or continuously differentiable functional, see Mielke et al. [122, 142] and Rossi & Savaré [143].

Ad (3.B). The continuity condition (3.Ba) means that B is a continuous perturbation of $\partial\mathcal{E}_t$. The term B contains in the applications non-variational and non-monotone terms of lower order in terms of only containing lower order derivatives as well as obeying a growth condition of lower order. This is reflected by Condition (3.Bb), where B satisfies a growth condition in terms of the conjugate dissipation potential Ψ_u^* and the energy functional $\mathcal{E}_t$. In fact, the growth condition shows that the higher the order of the growths of Ψ_u and $\mathcal{E}_t$ are, the more we can allow for the growth of the perturbation. Condition (3.Bb) ensures that we are able to derive a priori estimates. Both conditions can be generalized in a framework that instead of a point-wise continuity and a pointwise growth condition, a continuity on suitable Bochner spaces can be imposed as well as a growth condition on the level of time integrals. Furthermore, it would be sufficient to define the perturbation on the domain of the subdifferential of $\mathcal{E}_t$ or more generally to consider set-valued maps. This would allow a broader class of perturbations, see Ôtani [126] and Akagi [4], where this has been considered.

Now, we want to elaborate on why we imposed the growth condition on the regularization Ψ_u^ε instead of Ψ_u and why we regularize Ψ_u by the p-Moreau–Yosida regularization and not the classical Moreau–Yosida regularization where $p = 2$. The reason why it is important to regularize Ψ_u with the order $p > 1$ instead of $p = 2$ is the Condition (3.Bb). Regularizing the dissipation potential with the order $p > 1$ which is larger than the growth rate of Ψ_u might make Condition (3.Bb) too restrictive. This can be seen as follows: assume that Condition (3.Bb) is fulfilled only for Ψ_u, i.e., when $\varepsilon = 0$, and that Ψ_u fulfills the following growth condition: for all $R > 0$, there exist a a real number $p > 1$ and a constant $C_R > 0$ such that

$$\Psi_u(v) \le C_R(\|v\|^p + 1) \quad \text{for all } v, u \in V \ \text{ with } \mathcal{G}(u) \le R. \tag{3.2.5}$$

It is easy to see that the p-Moreau–Yosida regularization of Ψ_u satisfies the same growth condition. Furthermore, by the calculation rules for the Legendre-Fenchel transformation (Lemma 2.3.2), for all $R > 0$, there exist constants $\tilde{c}_R, \tilde{C}_R > 0$ such that

$$\tilde{c}\|\xi\|_*^{q^*} - \tilde{C} \le \Psi_u^*(\xi) \quad \text{for all } \xi \in V^*, u \in V \ \text{ with } \mathcal{G}(u) \le R.$$

Now, let $c_1 \in (0, 1)$ and $\beta > 0$ be from Condition (3.Bb). Then, there holds

$$\begin{aligned} c_1\Psi_u^{\varepsilon,*}\left(\frac{B(t,u)}{c_1}\right) &= \frac{c_1\varepsilon}{p^*}\left\|\frac{B(t,u)}{c_1}\right\|^{p^*} + c\Psi_u^*\left(\frac{B(t,u)}{c_1}\right) \\ &\le \left(\frac{\varepsilon}{p^*\tilde{c}} + 1\right) c_1\Psi_u^*\left(\frac{B(t,u)}{c_1}\right) + \frac{\tilde{C}}{\tilde{c}} \\ &\le \left(\frac{\varepsilon}{p^*\tilde{c}} + 1\right) \beta(1 + \mathcal{E}_t(u)) + \frac{\tilde{C}}{\tilde{c}} \\ &\le \tilde{\beta}(1 + \mathcal{E}_t(u)) \quad \text{for all } u \in D, t \in [0,1] \text{ and } \varepsilon \in [0,1], \end{aligned} \tag{3.2.6}$$

for a constant $\tilde{\beta} > 0$, where the equality (3.2.6) will be shown in the Lemma 3.2.4 below. Therefore, if $p \ge 2$, we can in fact use the classical Moreau–Yosida regularization with $p = 2$. However, if $p < 2$, we need to take the regularization with order p.

Having discussed the assumptions for the perturbed gradient system $(V, \mathcal{E}, \Psi, B)$, we are in the position to state the main existence result, which also contains the notion of a solution to (3.0.1).

Theorem 3.2.3 (Existence result) *Let the perturbed gradient system* $(V, \mathcal{E}, \Psi, B)$ *satisfy the Assumptions* (3.Ψ), (3.E)*and* (3.B). *Then, for every* $u_0 \in D$, *there exists an absolutely continuous curve* $u \in \mathrm{AC}([0,T];V)$ *with* $u(0) = u_0$ *and an integrable function* $\xi \in \mathrm{L}^1(0,T;V)$ *with* $\xi(t) \in \partial\mathcal{E}_t(u(t))$ *for a.e.* $t \in (0,T)$ *such that the following energy-dissipation balance holds:*

$$\begin{aligned} &\mathcal{E}_t(u(t)) + \int_s^t \Big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}\Big(B(r,u(r)) - \xi(r)\Big)\Big)\,\mathrm{d}r \\ &= \mathcal{E}_s(u(s)) + \int_s^t \partial_r\mathcal{E}_r(u(r))\,\mathrm{d}r + \int_s^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r \quad \textit{for all } s,t \in [0,T]. \end{aligned} \tag{3.2.7}$$

We also want to show that for nonsmooth functionals $\mathcal{E}_t$ and Ψ_u, every absolutely continuous function which satisfies the energy-dissipation balance (3.2.7) is a solution to the perturbed gradient system (3.0.1). First, by the chain rule (3.Ed), we obtain after the integration of (3.2.2) that

$$\mathcal{E}_t(u(t)) \geq \mathcal{E}_s(u(s)) + \int_s^t \left(\langle \xi(r), u'(t)\rangle + \partial_r\mathcal{E}_r(u(r))\right)\mathrm{d}r \quad \text{for all } s,t \in [0,T].$$

Plugging in the latter inequality into the energy-dissipation balance (3.2.7), we obtain

$$\begin{aligned} &\int_s^t \Big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}\Big(B(r,u(r)) - \xi(r)\Big)\mathrm{d}r - \langle B(r,u(r)) - \xi(r), u'(r)\rangle\Big)\,\mathrm{d}r \\ &\leq 0 \quad \text{for all } s,t \in [0,T]. \end{aligned}$$

Since the integrand of (3.2.7) is by the Fenchel–Young inequality non-negative, there holds

$$\langle B(t,u(t)) - \xi(t), u'(t)\rangle = \Psi_{u(t)}(u'(t)) + \Psi^*_{u(t)}\Big(B(t,u(t)) - \xi(t)\Big)$$

for almost every $t \in (0,T)$. Finally, by Lemma 2.3.1, there holds

$$\xi(t) \in \partial\mathcal{E}_t(u(t)), \quad \partial\Psi_{u(t)}(u'(t)) \ni B(t,u(t)) - \xi(t) \quad \text{for almost every } t \in (0,T).$$

Since the initial condition $u(0) = u_0$ is fulfilled, we conclude that u solves (3.0.1). We will refer to the just described equivalence between solutions to (3.2.7) and (3.0.1) as the *energy-dissipation principle.*

The proof of Theorem 3.2.3 is divided in two main parts. In the first part, we show the existence of solutions to the regularized problem emerging from the regularization of the dissipation potential via the p-Moreau–Yosida regularization by proving the convergence of a time discretization scheme. In the second part, the solvability of the problem (3.0.1) is obtained by essentially repeating the same arguments, while instead of passing with the time step τ to zero, we let the regularization parameter ε vanish.

Before we start with the proof of Theorem 3.2.3, we need several auxiliary results. In the following lemma, we show that under the Assumption (3.Ψ), the p-Moreau–Yosida regularization of Ψ_u satisfies certain properties including Conditions (3.Ψa)-(3.Ψc). In fact, we will show that the p-Moreau–Yosida regularization satisfies even more properties than it.

Lemma 3.2.4 *Let the family of dissipation potential $\Psi_u : V \to \mathbb{R}$, $u \in D$, be given and satisfy Assumptions* (3.Ψa)-(3.Ψc) *and let $\varepsilon \in (0,1]$ and $p > 1$. Then, the p-Moreau–Yosida regularization Ψ_u^ε fulfills* (3.Ψa)-(3.Ψc) *and the conjugate of Ψ_u^ε is given by*

$$\Psi_u^{\varepsilon,*}(\xi) = \frac{\varepsilon}{p^*}\|\xi\|_*^{p^*} + \Psi_u^*(\xi) \quad \text{for all } \xi \in V^*, u \in D, \tag{3.2.8}$$

where $p^* > 1$ is the conjugate exponent to p. Moreover, Ψ_u^ε and $\Psi_u^{\varepsilon,*}$ are superlinear uniformly with respect to $\varepsilon > 0$ and on sublevels of the energy. Finally, for all $R > 0$ and all sequences $(\varepsilon_n, u_n)_{n\in\mathbb{N}} \subset (0,1]\times V$ with $\sup_{n\in\mathbb{N}} \mathcal{G}(u_n) \le R$ and $(\varepsilon_n, u_n) \to (0,u)$ as $n \to \infty$, there holds $\Psi_{u_n}^{\varepsilon_n} \xrightarrow{\mathrm{M}} \Psi_u$.

Proof. First, for each $\varepsilon > 0$, the regularization Ψ_u^ε is a dissipation potential. This follows immediately from the fact that $\Psi_u^\varepsilon(0) = \inf_{v\in V}\left\{\frac{1}{p\varepsilon^{p-1}}\|v\|^p + \Psi_u(v)\right\} = 0$ and Lemma 2.5.1. The formula (3.2.8) follows from the calculations

$$\begin{aligned}
\Psi_u^{\varepsilon,*}(\xi) &= \sup_{v\in V}\{\langle \xi, v\rangle - \Psi_u^\varepsilon(v)\}\\
&= \sup_{v\in V}\left\{\langle \xi, v\rangle - \inf_{w\in V}\left\{\frac{\varepsilon}{p}\left\|\frac{v-w}{\varepsilon}\right\|^p + \Psi_u(w)\right\}\right\}\\
&= \sup_{v\in V}\sup_{w\in V}\left\{\langle \xi, v\rangle - \frac{\varepsilon}{p}\left\|\frac{v-w}{\varepsilon}\right\|^p - \Psi_u(w)\right\}\\
&= \sup_{w\in V}\sup_{v\in V}\left\{\langle \xi, v\rangle - \frac{\varepsilon}{p}\left\|\frac{v-w}{\varepsilon}\right\|^p - \Psi_u(w)\right\}\\
&= \sup_{w\in V}\left\{\sup_{v\in V}\left\{\langle \xi, v-w\rangle - \frac{\varepsilon}{p}\left\|\frac{v-w}{\varepsilon}\right\|^p\right\} + \langle \xi, w\rangle - \Psi_u(w)\right\}\\
&= \sup_{w\in V}\left\{\varepsilon\sup_{v\in V}\left\{\left\langle \xi, \frac{v-w}{\varepsilon}\right\rangle - \frac{1}{p}\left\|\frac{v-w}{\varepsilon}\right\|^p\right\} + \langle \xi, w\rangle - \Psi_u(w)\right\}\\
&= \sup_{w\in V}\left\{\frac{\varepsilon}{p^*}\|\xi\|_*^{p^*} + \langle \xi, w\rangle - \Psi_u(w)\right\} = \frac{\varepsilon}{p^*}\|\xi\|_*^{p^*} + \Psi_u^*(\xi)
\end{aligned}$$

for all $\xi \in V^*$ and $u \in D$, where we have used the fact $(\frac{1}{p\varepsilon^{p-1}}\|\cdot\|^p)^* = \frac{\varepsilon}{p^*}\|\cdot\|_*^{p^*}$. The expression (3.2.8) also shows the superlinearity of $\Psi_u^{\varepsilon,*}$ uniformly in ε. We proceed by showing the superlinearity of $\Psi_u^{\varepsilon,*}$. To do so, we note that the superlinearity of Ψ_u equivalently says that for all $R > 0$ and $M > 0$, there exists a positive real number $K > 0$ such that

$$\Psi_u(v) \ge M\|v\| \tag{3.2.9}$$

for all $u \in D$ with $\mathcal{G}(u) \leq R$ and all $v \in V$ with $\|v\| \geq K$. The idea is to show that for the regularization Ψ_u^ε for all $\tilde{R} > 0$ and $\tilde{M} > 0$, there exists a positive real number $\tilde{K} > 0$ independent of the parameter $\varepsilon > 0$ with the above mentioned property. So, let $\tilde{R} > 0$ and $\tilde{M} > 0$, then for $R = \tilde{R}$ and $M = 2\tilde{M}$, there exists $K > 0$ such that (3.2.9) holds. We find by YOUNG's inequality and the triangle inequality

$$\begin{aligned}
\Psi_u^\varepsilon(v) &= \inf_{\tilde{v}\in V} \left\{ \frac{1}{p\varepsilon^{p-1}} \|v - \tilde{v}\|^p + \Psi_u(\tilde{v}) \right\} \\
&= \min \left\{ \inf_{\substack{\tilde{v}\in V \\ \|\tilde{v}\|\geq K}} \left\{ \frac{1}{p\varepsilon^{p-1}} \|v - \tilde{v}\|^p + \Psi_u(\tilde{v}) \right\}, \inf_{\substack{\tilde{v}\in V \\ \|\tilde{v}\|\leq K}} \left\{ \frac{1}{p\varepsilon^{p-1}} \|v - \tilde{v}\|^p + \Psi_u(\tilde{v}) \right\} \right\} \\
&\geq \min \left\{ \inf_{\substack{\tilde{v}\in V \\ \|\tilde{v}\|\geq K}} \left\{ \frac{1}{p\varepsilon^{p-1}} \|v - \tilde{v}\|^p + M\|\tilde{v}\| \right\}, \inf_{\substack{\tilde{v}\in V \\ \|\tilde{v}\|\leq K}} \frac{1}{p\varepsilon^{p-1}} \|v - \tilde{v}\|^p \right\} \\
&\geq \min \left\{ \inf_{\substack{\tilde{v}\in V \\ \|\tilde{v}\|\geq K}} \left\{ M\|v - \tilde{v}\| + M\|\tilde{v}\| - \frac{M^{p^*}\varepsilon}{p^*} \right\}, \inf_{\substack{\tilde{v}\in V \\ \|\tilde{v}\|\leq K}} \left\{ M\|v - \tilde{v}\| - \frac{M^{p^*}\varepsilon}{p^*} \right\} \right\} \\
&\geq \min \left\{ \left(M\|v\| - \frac{M^{p^*}}{p^*} \right), \left(M\|v\| - KM - \frac{M^{p^*}}{p^*} \right) \right\} \\
&= M\|v\| - KM - \frac{M^{p^*}}{p^*} \\
&\geq \frac{M}{2}\|v\| = \tilde{M}\|v\|
\end{aligned}$$

for all $v \in V$ with $\|v\| \geq \tilde{K} := 2\left(K + \frac{\tilde{M}^{p^*-1}}{p^* 2^{p^*-1}}\right)$ and $\varepsilon \in (0, 1]$. This implies the uniform superlinearty of Ψ_u^ε uniformly in $\varepsilon > 0$, which in turn implies the superlinearity for a fixed $\varepsilon > 0$. We continue by showing that Ψ_u^ε is continuous in the sense of MOSCO-convergence. In fact, we show that for a fixed $\varepsilon > 0$, the regularization satisfies a stronger version of MOSCO-convergence, meaning that there not only exists a recovery sequence, but that every sequence converging against the same limit is a recovery sequence, or in other words, the mapping $(u, v) \mapsto \Psi_u^\varepsilon(v)$ is continuous with respect to the strong topology for u on sublevels of $\mathcal{G}$ and v in V. Let $(u_n)_{n\in\mathbb{N}} \subset V$ with $\sup_{n\in\mathbb{N}} \mathcal{G}(u_n) < \infty$ be given such that $u_n \to u \in V$ as $n \to \infty$, and $(v_n)_{n\in\mathbb{N}} \subset V$ a weakly convergent sequence with weak limit $v \in V$. Now, let $(n_k)_{k\in\mathbb{N}}$ be a subsequence such that

$$\liminf_{n\to\infty} \Psi_{u_n}^\varepsilon(v_n) = \lim_{k\to\infty} \Psi_{u_{n_k}}^\varepsilon(v_{n_k}).$$

For each $k \in \mathbb{N}$, we denote by v_ε^k the unique minimizer of $v \mapsto \frac{1}{p\varepsilon^{p-1}}\|v - v_{n_k}\|^p + \Psi_{u_{n_k}}(v)$ and note that thanks to the estimate

$$\frac{1}{p\varepsilon^{p-1}} \|v_{n_k} - v_\varepsilon^k\|^p \leq \Psi_{u_{n_k}}^\varepsilon(v_{n_k}) \leq \frac{1}{p\varepsilon^{p-1}} \|v_{n_k}\|^p, \tag{3.2.10}$$

the corresponding sequence of minimizers $(v_\varepsilon^k)_{k\in\mathbb{N}}$ is bounded. Therefore, there exists a subsequence (labeled as before) which is weakly convergent to an element $\tilde{v}_\varepsilon \in V$.

Then, by the Mosco-convergence $\Psi_{u_n} \xrightarrow{\mathrm{M}} \Psi_u$, we observe

$$\begin{aligned}
\Psi_u^\varepsilon(v) &\leq \frac{1}{p\varepsilon^{p-1}}\|v - \tilde{v}_\varepsilon\|^p + \Psi_u(\tilde{v}_\varepsilon) \\
&\leq \liminf_{k\to\infty} \left\{ \frac{1}{p\varepsilon^{p-1}}\|v_{n_k} - v_\varepsilon^k\|^p + \Psi_{u_{n_k}}(v_\varepsilon^k) \right\} \\
&= \lim_{k\to\infty} \Psi_{u_{n_k}}^\varepsilon(v_{n_k}) = \liminf_{n\to\infty} \Psi_{u_n}^\varepsilon(v_n).
\end{aligned}$$

Now, let $v \in V$ be arbitrary and $(v_n)_{n\in\mathbb{N}} \subset V$ any strongly convergent sequence $v_n \to v$ as $n \to \infty$. We extract an arbitrary subsequence $(n_k)_{k\in\mathbb{N}}$, and to each $k \in \mathbb{N}$, we denote the minimizers of $v \mapsto \frac{1}{p\varepsilon^{p-1}}\|v - v_{n_k}\|^p + \Psi_{u_{n_k}}(v)$ again by $v_\varepsilon^k \in V$ and by $\tilde{v}_\varepsilon \in V$ the weak limit of a further subsequence of the very same sequence which we label as before. Once more, by (3.Ψc), for the minimizer v_ε of $\Psi_u^\varepsilon(v)$, there exists a strongly convergent recovery sequence $(\hat{v}_k)_{k\in\mathbb{N}} \subset V$ fulfilling $\hat{v}_k \to v_\varepsilon$ and $\lim_{k\to\infty} \Psi_{u_{n_k}}(\hat{v}_k) = \Psi_u(v_\varepsilon)$. It follows

$$\begin{aligned}
\Psi_u^\varepsilon(v) &\leq \frac{1}{p\varepsilon^{p-1}}\|v - \tilde{v}_\varepsilon\|^p + \Psi_u(\tilde{v}_\varepsilon) \\
&\leq \liminf_{k\to\infty} \left\{ \frac{1}{p\varepsilon^{p-1}}\|v_{n_k} - v_\varepsilon^k\|^p + \Psi_{u_{n_k}}(v_\varepsilon^k) \right\} \\
&= \liminf_{k\to\infty} \Psi_{u_{n_k}}^\varepsilon(v_{n_k}) \\
&\leq \limsup_{k\to\infty} \Psi_{u_{n_k}}^\varepsilon(v_{n_k}) \\
&\leq \limsup_{k\to\infty} \left\{ \frac{1}{p\varepsilon^{p-1}}\|v_{n_k} - \hat{v}_k\|^p + \Psi_{u_{n_k}}(\hat{v}_k) \right\} \\
&= \lim_{k\to\infty} \left\{ \frac{1}{p\varepsilon^{p-1}}\|v_{n_k} - \hat{v}_k\|^p + \Psi_{u_{n_k}}(\hat{v}_k) \right\} \\
&= \frac{1}{p\varepsilon^{p-1}}\|v - v_\varepsilon\|^p + \Psi_u(v_\varepsilon) = \Psi_u^\varepsilon(v).
\end{aligned}$$

Therefore, every subsequence $(n_k)_{k\in\mathbb{N}}$ contains a further subsequence $(n_{k_l})_{l\in\mathbb{N}}$ such that $\lim_{l\to\infty} \Psi_{u_{n_{k_l}}}^\varepsilon(v_{n_{k_l}}) = \Psi_u^\varepsilon(v)$. By the subsequence principle, the convergence of the whole sequence follows. In particular, this shows $v_\varepsilon = \tilde{v}_\varepsilon$.
Finally, we show the Mosco-convergence involving a vanishing sequence of regularization parameters $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,1]$, i.e. satisfying $\varepsilon_n \to 0$ as $n \to \infty$. As before, let the sequences $(u_n)_{n\in\mathbb{N}}, (v_n)_{n\in\mathbb{N}} \subset V$ with $\sup_{n\in\mathbb{N}} \mathcal{G}(u_n) < \infty$ be given such that $u_n \to u \in V$ and $v_n \rightharpoonup v \in V$ as $n \to \infty$, and let $(n_k)_{k\in\mathbb{N}}$ be a subsequence such that

$$\liminf_{n\to\infty} \Psi_{u_n}^{\varepsilon_n}(v_n) = \lim_{k\to\infty} \Psi_{u_{n_k}}^{\varepsilon_{n_k}}(v_{n_k}).$$

By $\tilde{v}_k \in V$, $k \in \mathbb{N}$, we denote again the minimizer of $\Psi_{u_{n_k}}^{\varepsilon_{n_k}}(v_{n_k})$. Due to the same estimate as (3.2.10) for $(\tilde{v}_k)_{k\in\mathbb{N}}$, the sequence of minimizers is bounded and therefore sequentially compact with respect to the weak topology. So, after extracting a subsequence (labeled as before), we obtain a weak limit $\tilde{v} \in V$ such that $\tilde{v}_k \rightharpoonup \tilde{v}$ as $n \to \infty$. Now, we consider two cases:

i) $\frac{1}{p\varepsilon_{n_k}^{p-1}}\|v_{n_k} - \tilde{v}_k\|^p \leq C$ for a constant $C > 0$,

ii) $\frac{1}{p\varepsilon_{n_k}^{p-1}}\|v_{n_k} - \tilde{v}_k\|^p \to \infty$ as $k \to \infty$ after possibly extracting a further subsequence.

Ad *i)*. We immediately find $v = \tilde{v}$ and therefore $\tilde{v_k} \rightharpoonup v$ as $k \to \infty$. By the continuity of Ψ in the sense of Mosco-convergence, it follows

$$\begin{aligned}
\Psi_u(v) &\leq \liminf_{k\to\infty} \Psi_{u_{n_k}}(\tilde{v}_k) \\
&\leq \liminf_{k\to\infty} \left\{ \frac{1}{p\varepsilon_{n_k}^{p-1}}\|v_{n_k} - \tilde{v}_k\|^p + \Psi_{u_{n_k}}(\tilde{v}_k) \right\} \\
&= \liminf_{k\to\infty} \Psi_{u_{n_k}}^{\varepsilon_{n_k}}(v_{n_k}) \\
&= \lim_{k\to\infty} \Psi_{u_{n_k}}^{\varepsilon_{n_k}}(v_{n_k}) \\
&= \liminf_{n\to\infty} \Psi_{u_n}^{\varepsilon_n}(v_n).
\end{aligned}$$

Ad *ii)*. We obtain

$$\begin{aligned}
\Psi_u(v) &\leq \lim_{k\to\infty} \left(\frac{1}{p\varepsilon_{n_k}^{p-1}}\|v_{n_k} - \tilde{v}_k\|^p \right) \\
&\leq \lim_{k\to\infty} \left\{ \frac{1}{p\varepsilon_{n_k}^{p-1}}\|v_{n_k} - \tilde{v}_k\|^p + \Psi_{u_{n_k}}(\tilde{v}_k) \right\} \\
&= \lim_{k\to\infty} \Psi_{u_{n_k}}^{\varepsilon_{n_k}}(v_{n_k}) \\
&= \liminf_{n\to\infty} \Psi_{u_n}^{\varepsilon_n}(v_n).
\end{aligned}$$

It remains to show the existence of a recovery sequence. Let $v \in V$ be arbitrarily chosen. Then, there exists a recovery sequence $(v_n)_{n\in\mathbb{N}} \subset V$ for Ψ_u with $v_n \to v$ as $v \to \infty$ such that $\lim_{n\to\infty} \Psi_{u_n}(v_n) = \Psi_u(v)$. Proceeding as before, we take an arbitrary subsequence $(n_k)_{k\in\mathbb{N}}$ and denote by $(\tilde{v}_k)_{k\in\mathbb{N}} \subset V$ again the minimizing sequence. Then, we consider again the two cases *i)* and *ii)*.
Ad *i)*. Since the recovery sequence is strongly convergent, it follows that the minimizing sequence of Ψ_u is also strongly convergent with limit $v \in V$. We obtain

$$\begin{aligned}
\Psi_u(v) &\leq \liminf_{k\to\infty} \Psi_{u_{n_k}}(\tilde{v}_k) \\
&\leq \liminf_{k\to\infty} \left\{ \frac{1}{p\varepsilon_{n_k}^{p-1}}\|v_{n_k} - \tilde{v}_k\|^p + \Psi_{u_{n_k}}(\tilde{v}_k) \right\} \\
&= \liminf_{k\to\infty} \Psi_{u_{n_k}}^{\varepsilon_{n_k}}(v_{n_k}) \\
&\leq \limsup_{k\to\infty} \Psi_{u_{n_k}}^{\varepsilon_{n_k}}(v_{n_k}) \\
&\leq \limsup_{k\to\infty} \Psi_{u_{n_k}}(v_{n_k}) \\
&= \lim_{k\to\infty} \Psi_{u_{n_k}}(v_{n_k}) = \Psi_u(v),
\end{aligned}$$

which by the same argument as before implies the convergence of the full sequence, i.e., $\lim_{n\to\infty} \Psi_{u_n}^{\varepsilon_n}(v_n) = \Psi_u(v)$.
Ad *ii*). Due to $\Psi_{u_n}^{\varepsilon_n}(v_n) \leq \Psi_{u_n}(v_n)$ and the convergence of the right-hand side, this case can not occur, which completes the proof. □

As mentioned above, the p-Moreau–Yosida regularization can be viewed as a regularization process described by the Hamilton–Jacobi equation (2.5.16). However, introducing the Moreau–Yosida regularization as a solution to the Cauchy problem (2.5.16) does not seem "natural". Interestingly, the regularization arises naturally when one deals with (generalized) gradient flow equations which we have seen already in (3.1.2) and (3.1.4). To demonstrate this more clearly, we consider the generalized gradient flow

$$-|u'(t)|^{p-2}u'(t) \in \partial E(u(t)), \quad t > 0,$$

of a functional $E : H \to (-\infty, +\infty]$ on a Hilbert space H. Discretizing the equation by the implicit Euler scheme leads to

$$-\left|\frac{U_\tau^n - U_\tau^{n-1}}{\tau}\right|^{p-2} \frac{U_\tau^n - U_\tau^{n-1}}{\tau} \in \partial E(U_\tau^n), \quad n = 1, 2, \ldots, N,$$

where, starting with $U_\tau^0 = u_0 \in \mathrm{dom}(E)$, the values $U_\tau^n, n = 1, \ldots, N$, can under certain conditions be obtained by the time-incremental minimization scheme

$$U_\tau^n \in J_\tau(U^{n-1}) := \operatorname*{argmin}_{v\in H} \left\{ \frac{\tau}{p} \left| \frac{v - U_\tau^{n-1}}{\tau} \right|^p + E(v) \right\}, \quad n = 1, 2, \ldots, N. \tag{3.2.11}$$

Here, obviously the Moreau–Yosida regularization occurs naturally after time-discretization of the equation and the approximative values $U_\tau^n \in H$ are defined by the Moreau–Yosida regularization E_τ where the regularization parameter is given by the step size τ of the time-discretization. It is also worth mentioning that the Moreau–Yosida regularization does not only regularize a function itself, but the associated resolvent operator $J_\tau(u)$ regularizes in a certain sense its arguments $u \in H$: the values $U_\tau^n \in \mathrm{dom}(\partial E)$, which are achieved in the minimization scheme, are not only contained in the domain of the functional E, but are also contained in the domain of the subdifferential ∂E. The latter is also referred to as the regularizing or smoothing effect of the gradient flow equation, which means that for a given initial datum $u_0 \in \mathrm{dom}(E)$ (or in some cases even $u_0 \in \overline{\mathrm{dom}(E)}$) the solution does not only belong to the domain of E but also to the domain of its subdifferential ∂E for an infinitesimal larger time step, i.e, $u(t) \in \mathrm{dom}(\partial E)$ for every $t > 0$. It is well-known that in the case $p = 2$ and when $E : H \to (-\infty, \infty]$ is a proper, lower semicontinuous and convex functional, the subdifferential operator ∂E is an infinitesimal generator of a C_0-semigroup such that $S(t)u_0 = u(t)$ is the unique solution to the Cauchy problem

$$\begin{cases} u'(t) \in -\partial E(u(t)), & t > 0, \\ u(0) = u_0 \in \overline{\mathrm{dom}(E)} \end{cases}$$

and which fulfills $S(t)u_0 = \lim_{n\to\infty} J^n_{t/n}(u_0)$, where $J_{t/n}$ denotes again the resolvent operator given by (3.2.11), see, e.g., [24, 32]. This property even holds true in a complete metric space under slightly weaker assumptions on the functional E, see Ambrosio et al. [10] for a detailed discussion.
Since the solutions of the Cauchy problem we aim to study are, in general, not unique (even for the discretized problem), the generation of a semigroup can not be expected. Nevertheless, the nice regularizing effect can be maintained despite the fact that the energy functional is not assumed to be convex, see Section 3.1 and Assumption (3.E) in Section 3.2.

Now, in regard to the main evolution inclusion

$$\partial\Psi_{u(t))}(u'(t)) + \partial\mathcal{E}_t(u(t)) \ni B(t, u(t)),$$

and its time-discretization as performed in (3.1.1), it is worth investigating the properties of the so-called *Ψ^ε-Moreau-Yosida regularization*

$$\begin{aligned}\Phi_{r,t}(w;u) :&= \inf_{v\in V} \Phi(r,t,u,w;v) \\ &= \inf_{v\in V} \left\{ r\Psi^\varepsilon_u\left(\frac{v-u}{r}\right) + \mathcal{E}_{t+r}(v) - \langle w, v\rangle \right\}, \end{aligned} \tag{3.2.12}$$

of $\mathcal{E}_t + \langle w, \cdot\rangle$, where $u \in D, w \in V^*$, $r > 0$ and $t \in [0,T)$ such that $r + t \in [0,T]$. As previously mentioned, the energy $\mathcal{E}_t$ was not supposed to be convex. Hence, it might be at first sight unclear what regularizing properties can be expected from the Ψ^ε-*Moreau-Yosida* regularization and why it is useful to study the function $\Phi_{r,t}(w;u)$. This question will be answered in the following section.

3.3 The Ψ-Moreau-Yosida regularization

In this section, we study the properties of the Ψ^ε-Moreau–Yosida regularization (3.2.12), where for $p > 1$ and fixed $\varepsilon > 0$, Ψ^ε is the p-Moreau-Yosida regularization of the dissipation potential. In particular, we have to ensure that the infimum in formula (3.2.12) is attained so that the resolvent set

$$J_{r,t}(w;u) := \operatorname*{argmin}_{v\in V} \Phi(r,t,u,w;v) = \operatorname*{argmin}_{v\in V} \left\{ r\Psi^\varepsilon_u\left(\frac{v-u}{r}\right) + \mathcal{E}_{t+r}(v) - \langle w, v\rangle \right\}$$

is non-empty and we obtain solvability of the discretized inclusion (5.2.4).

Lemma 3.3.1 *Let the perturbed gradient system $(V, \mathcal{E}, \Psi^\varepsilon, B)$ satisfy the Assumptions* (3.Ea)-(3.Eb) *and* (3.Ψa). *Then, for all $r > 0$, $t \in [0,T)$ with $t + r \le T$, $u \in D$, and $w \in V^*$, the resolvent set $J_{r,t}(w;u)$ is non-empty.*

Proof. The proof is based on the direct method of calculus of variations: let $u \in D, w \in V^*$ and $r > 0, t \in [0,T)$ with $r + t \in [0,T]$ be fixed. First, we note that by the Fenchel–Young inequality and with the boundedness of the energy from below, there holds

$$\begin{aligned}\Phi(r,t,u,w;v) &= r\Psi_u^\varepsilon\left(\frac{v-u}{r}\right) + \mathcal{E}_{t+r}(v) - \langle w, v\rangle \\ &\geq -r\Psi_u^{\varepsilon,*}(w) + \mathcal{E}_{t+r}(v) - \langle w, u\rangle \\ &\geq -r\Psi_u^{\varepsilon,*}(w) + S - \langle w, u\rangle\end{aligned} \tag{3.3.1}$$

and hence $\Phi_{r,t}(w;u) > -\infty$. On the other hand, we have

$$\inf_{v\in V}\left\{r\Psi_u^\varepsilon\left(\frac{v-u}{r}\right) + \mathcal{E}_{t+r}(v) - \langle w, v\rangle\right\} \leq \mathcal{E}_{t+r}(u) - \langle w, u\rangle, \tag{3.3.2}$$

so that we also find $\Phi_{r,t}(w;u) < +\infty$. Now, let $(v_n)_{n\in\mathbb{N}} \subset V$ be a minimizing sequence for $\Phi(r,t,u,w;\cdot)$. From (3.3.1), we deduce with Remark 3.2.2 that $(v_n)_{n\in\mathbb{N}} \subset V$ is contained in a sublevel set of the energy functional. Since the energy has compact sublevels (Assumption (3.Eb)), there exists a subsequence (relabeled as before) which converges strongly in V to a limit $v \in V$. The lower semicontinuity of the energy functional and the dissipation potential yield the lower semicontinuity of the mapping $v \mapsto \Phi(r,t,u,w;v)$. Therefore, there holds

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

and therefore $v \in J_{r,t}(w;u) \neq \emptyset$ is a global minimizer from which $v \in D$ follows. □

The next lemma provides an analogous result to Theorem 2.5.2 for the Ψ^ε-MOREAU–YOSIDA regularization and is crucial for the main existence result and particularly for deriving a priori estimates. The result is an adaptation of the unperturbed case of Lemma 4.2 in ROSSI & SAVARÉ [143] and Lemma 6.1 in MIELKE et al. [122], where $w = 0$ has been considered.

Lemma 3.3.2 *Let the perturbed gradient system $(V, \mathcal{E}, \Psi, B)$ be given and satisfy the Assumptions (3.Ψ) and (3.Ea)-(3.Ed). Then, for every $\mathrm{t} \in [0,T), u \in D$ and $w \in V^*$, there exists a measurable selection $r \mapsto u_r : (0, T-\mathrm{t}) \to J_{r,\mathrm{t}}(w;u)$ such that*

$$w \in D_G\Psi_u^\varepsilon\left(\frac{u_r-u}{r}\right) + \partial\mathcal{E}_{\mathrm{t}+r}(u). \tag{3.3.3}$$

Moreover, there holds

i) $\exists\tilde{C} > 0 :\; G(u_r) \leq \widetilde{C}(G(u) + r\Psi_u^{\varepsilon,*}(w))$ *for all* $r \in (0, T-\mathrm{t})$, (3.3.4)

ii) $\lim_{r\to 0} \Phi_{r,\mathrm{t}}(w;u) = \mathcal{E}_\mathrm{t}(u) - \langle w, u\rangle$, (3.3.5)

iii) $\lim_{r\to 0} \sup_{u_r\in J_{r,\mathrm{t}}(w;u)} \|u_r - u\| = 0$, (3.3.6)

for all $\mathrm{t} \in [0,T), u \in D$ *and* $w \in V^*$.
Finally, the mapping $r \mapsto \Phi_{r,\mathrm{t}}(w;u)$ is differentiable almost everywhere in $(0, T-\mathrm{t})$ and for every measurable selection $r \mapsto u_r : (0, r_0) \to J_{r,\mathrm{t}}(w;u)$, there exists a measurable map $r \mapsto \xi_r : (0, T-\mathrm{t}) \to \partial\mathcal{E}_{\mathrm{t}+r}(u)$ with $\xi_r = w - D_G\Psi^\varepsilon\left(\frac{u_r-u}{r}\right)$ such that

$$\frac{\mathrm{d}}{\mathrm{d}r}\Phi_{r,\mathrm{t}}(w;u) \ \leq \ -\Psi_u^{\varepsilon,*}(w-\xi_r)+\partial_t\mathcal{E}_{\mathrm{t}+r}(u_r) \quad \textit{for a.e. } r\in(0,T{-}\mathrm{t}). \tag{3.3.7}$$

In particular, there holds

$$\begin{aligned}\mathcal{E}_{\mathrm{t}+r_0}(u_{r_0})+r_0\Psi_u^{\varepsilon}\left(\frac{u_{r_0}-u}{r_0}\right)+\int_0^{r_0}\Psi_u^{\varepsilon,*}(w-\xi_r)\,\mathrm{d}r&\\ \leq \mathcal{E}_{\mathrm{t}}(u)+\int_0^{r_0}\partial_t\mathcal{E}_{\mathrm{t}+r}(u_r)\,\mathrm{d}r+\langle w,u_{r_0}-u\rangle&\end{aligned} \tag{3.3.8}$$

for every $r_0\in(0,T-\mathrm{t})$.

Proof. Let $\mathrm{t}\in[0,T), u\in D$ and $w\in V^*$ be arbitrary and fixed. Lemma 3.3.1 guarantees that the resolvent set $J_{r,\mathrm{t}}(w;u)$ is non-empty for all $r\in(0,T-\mathrm{t})$. Then, the existence of a measurable selection $r\mapsto u_r:(0,T-\mathrm{t})\to J_{r,\mathrm{t}}(w;u)$ is ensured by Castaing & Valadier [43, Cor. III.3, Prop. III.4, Thm. III.6, pp. 63]. The inclusion (3.3.3) follows immediately from the variational sum rule (3.2.4). Further, we obtain from the estimates (3.3.1) and (3.3.2) with the choice $v=u_r, r\in(0,T-\mathrm{t})$ the inequality

$$\mathcal{E}_{\mathrm{t}+r}(u_r)\leq\mathcal{E}_{\mathrm{t}+r}(u)+r\Psi_u^{\varepsilon,*}(w).$$

The latter inequality together with the estimate (3.2.3) yields inequality (3.3.4). In order to show the convergences in (3.3.5), we note again that the superlinearity of the dissipation potential (Assumption (3.Ψb)) and Lemma 3.2.4 implies that for all $R>0$ and $\gamma>0$, there exists $K>0$ such that

$$\Psi_u^{\varepsilon}(v)\geq\gamma\|v\|$$

for all $\varepsilon\in[0,1]$, $u\in D$ with $G(u)\leq R$ and all $v\in V$ with $\|v\|\geq K$. Based on this, we infer

$$\gamma\|v\|\leq\Psi_u^{\varepsilon}(v)+\gamma K\quad\text{for all } v\in V$$

and in particular

$$\gamma\left\|\frac{u_r-u}{r}\right\|\leq\Psi^{\varepsilon}\left(\frac{u_r-u}{r}\right)+\gamma K\qquad\text{for every } r>0.$$

In view of (3.3.2) and Remark 3.2.2 ii), this implies

$$\begin{aligned}\gamma\|u_r-u\|&\leq\langle w,u_r-u\rangle+\mathcal{E}_{\mathrm{t}+r}(u)-\mathcal{E}_{\mathrm{t}+r}(u_r)+r\gamma K\\&\leq\|w\|_*\|u_r-u\|+\mathcal{E}_{\mathrm{t}+r}(u)-S+r\gamma K.\end{aligned}$$

This yields

$$(\gamma-\|w\|_*)\|u_r{-}u\|\leq\mathcal{E}_{\mathrm{t}+r}(u)-S+r\gamma K\leq C_1\mathcal{E}_0(u)-S+r\gamma K$$

for all $r \in (0, T - \mathrm{t})$ and $u_r \in J_{r,\mathrm{t}}(w; u)$, where the constant $C_1 > 0$ comes from Remark 3.2.2 i). After taking the supremum over all $u_r \in J_{r,\mathrm{t}}(w; u)$ and passing to the limes superior as $r \to 0$ in the last inequality, we finally obtain

$$(\gamma - \|w\|_*) \limsup_{r\to 0} \sup_{u_r \in J_{r,\mathrm{t}}(w;u)} \|u_r - u\| \leq \mathrm{e}^{CT} \mathcal{E}_0(u) - S \quad \text{for every } \gamma > \|w\|_*.$$

By choosing $\gamma > 0$ sufficiently large, we conclude

$$\limsup_{r\to 0} \sup_{u_r \in J_{r,\mathrm{t}}(w;u)} \|u_r - u\| = 0,$$

which shows (3.3.6). We proceed with showing the convergence (3.3.5). Taking into account the estimate

$$\begin{aligned}
\mathcal{E}_{\mathrm{t}+r}(u_r) - \langle w, u_r \rangle &\leq \Phi_{r,\mathrm{t}}(w; u) \\
&= r\Psi_u^\varepsilon\left(\frac{u_r - u}{r}\right) + \mathcal{E}_{\mathrm{t}+r}(u_r) - \langle w, u_r \rangle \leq \mathcal{E}_{\mathrm{t}+r}(u) - \langle w, u \rangle,
\end{aligned}$$

the lower semicontinuity and the time continuity of the $\mathcal{E}_t$ as well as the fact that $\liminf_{r\to 0} \mathcal{E}_{\mathrm{t}+r}(u_r) = \liminf_{r\to 0} \mathcal{E}_{\mathrm{t}}(u_r)$, which follows from the LIPSCHITZ continuity of the time dependence of the energy functional, the convergence (3.3.5) follows from

$$\begin{aligned}
\mathcal{E}_{\mathrm{t}}(u) - \langle w, u \rangle &\leq \liminf_{r\to 0} \left(\mathcal{E}_{\mathrm{t}+r}(u_r) - \langle w, u_r \rangle\right) \\
&\leq \liminf_{r\to 0} \Phi_{r,t}(w; u) \ \leq \ \limsup_{r\to 0} \Phi_{r,\mathrm{t}}(w; u) \\
&\leq \limsup_{r\to 0} \left(\mathcal{E}_{\mathrm{t}+r}(u) - \langle w, u \rangle\right) \ = \ \mathcal{E}_{\mathrm{t}}(u) - \langle w, u \rangle.
\end{aligned}$$

We complete the proof of this lemma by showing the last assertion. Let for $0 < r_1 < r_2 < T - \mathrm{t}$, $u_{r_i} \in J_{r_i,\mathrm{t}}(w; u), i = 1, 2$. Then, there holds

$$\begin{aligned}
&\Phi_{r_2,\mathrm{t}}(w; u) - \Phi_{r_1,\mathrm{t}}(w; u) - (\mathcal{E}_{\mathrm{t}+r_2}(u_{r_1}) - \mathcal{E}_{\mathrm{t}+r_1}(u_{r_1})) \\
&\leq r_2\Psi_u^\varepsilon\left(\frac{u_{r_1} - u}{r_2}\right) - r_1\Psi_u^\varepsilon\left(\frac{u_{r_1} - u}{r_1}\right) \\
&= (r_2 - r_1)\Psi_u^\varepsilon\left(\frac{u_{r_1} - u}{r_2}\right) + r_1\left(\Psi_u^\varepsilon\left(\frac{u_{r_1} - u}{r_2}\right) - \Psi_u^\varepsilon\left(\frac{u_{r_1} - u}{r_1}\right)\right) \\
&\leq (r_2 - r_1)\left(\Psi_u^\varepsilon\left(\frac{u_{r_1} - u}{r_2}\right) - \left\langle w_2^1, \frac{u_{r_1} - u}{r_2}\right\rangle\right) \\
&= -(r_2 - r_1)\Psi_u^{\varepsilon,*}(w_2^1) \leq 0,
\end{aligned} \tag{3.3.9}$$

where $w_2^1 = D_G\Psi_u^\varepsilon\left(\frac{u_{r_1} - u}{r_2}\right)$ is the GÂTEAUX derivative of the regularized dissipation potential, which exists everywhere in V, and we used in (3.3.9) again Lemma 2.3.1. Employing (3.Ec), (3.3.1), and the inequality (3.3.4), we obtain

$$\begin{aligned}
\Phi_{r_2,\mathrm{t}}(w; u) &\leq \Phi_{r_1,\mathrm{t}}(w; u) + (\mathcal{E}_{\mathrm{t}+r_2}(u_{r_1}) - \mathcal{E}_{\mathrm{t}+r_1}(u_{r_1})) \\
&\leq \Phi_{r_1,\mathrm{t}}(w; u) + (r_2 - r_1)CG(u_{r_1}) \\
&\leq \Phi_{r_1,\mathrm{t}}(w; u) + (r_2 - r_1)CG(u) + (r_2 - r_1)r_1C\Psi_u^{\varepsilon,*}(w) \\
&\leq \Phi_{r_1,\mathrm{t}}(w; u) + (r_2 - r_1)CG(u) + \frac{1}{2}(r_2^2 - r_1^2)C\Psi_u^{\varepsilon,*}(w)
\end{aligned}$$

from which we conclude that the mapping $r \mapsto \Phi_{r,\mathrm{t}}(w;u) - C(rG(u) + \frac{1}{2}r^2\Psi_u^{\varepsilon,*}(w))$ is non-increasing on $(0, T-\mathrm{t})$ and therefore as a real-valued function almost everywhere differentiable. Since the latter function is the sum of the regularization $r \mapsto \Phi_{r,\mathrm{t}}(w;u)$ and a differentiable function (in r), we conclude that the mapping $r \mapsto \Phi_{r,\mathrm{t}}(w;u)$ is also almost everywhere differentiable in $(0, T-\mathrm{t})$. Hence, there exists a negligible set $\mathscr{N} \subset (0, T-\mathrm{t})$ such that the mapping $r \mapsto \Phi_{r,\mathrm{t}}(w;u)$ is differentiable on $(0,T-\mathrm{t})\backslash\mathscr{N}$, and we note that the set depends on u and w, i.e., $\mathscr{N} = \mathscr{N}_{u,w}$. Let $r \in (0, T-\mathrm{t})\backslash\mathscr{N}$ be arbitrary but fixed, and $(h_n)_{n\in\mathbb{N}} \in \mathbb{R}^{>0}$ be a sequence which converges from above to zero and whose elements are sufficiently small. Let the sequence $(w_n^r)_{n\in\mathbb{N}} \subset V^*$ be given by $w_n^r = D_G\Psi_u^\varepsilon\left(\frac{u_r-u}{r+h_n}\right)$ for all $n \in \mathbb{N}$. From the demicontinuity of $D_G\Psi_u^\varphi$, we obtain the weak convergence $w_n^r \rightharpoonup w_r = \left\{D_G\Psi_u^\varepsilon\left(\frac{u_r-u}{r}\right)\right\}$ in V^*. Then, by Lemma 2.3.1 and the continuity of Ψ_u^ε (see Remark 3.2.1), we obtain

$$\begin{aligned}\lim_{n\to\infty}\Psi_u^{\varepsilon,*}(w_n^r) &= \lim_{n\to\infty}\left(\left\langle w_n^r, \frac{u_r-u}{r+h_n}\right\rangle - \Psi^\varepsilon\left(\frac{u_r-u}{r+h_n}\right)\right)\\ &= \left\langle w_r, \frac{u_r-u}{r}\right\rangle - \Psi^\varepsilon\left(\frac{u_r-u}{r}\right) = \Psi_u^{\varepsilon,*}(w_r).\end{aligned}$$

Since the mapping $r \mapsto u_r$ is measurable, inclusion (3.3.3) ensures the measurability of the mapping $r \mapsto \xi_r : (0, T-\mathrm{t}) \to \partial\mathcal{E}_{\mathrm{t}+r}(u)$ given by $\xi_r = w - D_G\Psi_u\left(\frac{u_r-u}{r}\right) = w - w_r$. Finally, by the differentiability of the mapping $r \mapsto \Phi_{r,\mathrm{t}}(w;u)$ in $r \in (0,T-\mathrm{t})\backslash\mathscr{N}_{u,w}$, we deduce from (3.3.9)

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}r}\Phi_{r,\mathrm{t}}(w;u)|_{r=r} + \Psi_u^*(w-\xi_r) \;&=\; \lim_{n\to\infty}\left(\frac{\Phi_{r+h_n,\mathrm{t}}(w;u) - \Phi_{r,\mathrm{t}}(w;u)}{h_n} + \Psi_u^*(w_n^r)\right)\\ \le \liminf_{n\to\infty}\left(\frac{\mathcal{E}_{\mathrm{t}+r+h_n}(u_r) - \mathcal{E}_{\mathrm{t}+r}(u_r)}{h_n}\right) \;&=\; \partial_t\mathcal{E}_{\mathrm{t}+r}(u_r) \quad \text{for a.e. } r \in (0, T-\mathrm{t}),\end{aligned} \tag{3.3.10}$$

where we also employed Assumption (3.Ec). Integrating (3.3.10) from $r = 0$ to $r = r_0$ and making use of (3.3.5) leads to the desired result. ☐

If we compare this lemma to the analogous result of Lemma 2.5.2, we see that the assertions (3.3.3), (3.3.5) and (3.3.6) still hold true while (3.3.4) and (3.3.7) hold true in a weaker form. This is mainly due to the lack of convexity of $\mathcal{E}$. However, the shown properties are sufficient to show existence of solutions. In particular, the inequality (3.3.8) is crucial to obtain the right a priori estimates by choosing $r = \tau$, $U_\tau^n = u_r$, and $U_\tau^{n-1} = u$. It is quite remarkable that one can derive inequality (3.3.8) without any convexity assumption for $\mathcal{E}_t$. It is not obvious how to obtain the term

$$\int_0^{r_0}\Psi_u^{\varepsilon,*}(w-\xi_r)\,\mathrm{d}r$$

by only using the fact that u_{r_0} is a minimizer of $\Phi_{r,\mathrm{t}}(w;u)$, which would only lead to

$$\begin{aligned}\mathcal{E}_{\mathrm{t}+r_0}(u_{r_0}) + r_0\Psi_u^\varepsilon\left(\frac{u_{r_0}-u}{r_0}\right) &\le \mathcal{E}_{\mathrm{t}+r_0}(u) + \langle w, u_{r_0} - u\rangle\\ &= \mathcal{E}_{\mathrm{t}}(u) + \int_0^{r_0}\partial_t\mathcal{E}_{\mathrm{t}+r}(u)\,\mathrm{d}r + \langle w, u_{r_0} - u\rangle.\end{aligned}$$

In the latter inequality, the subgradient $\xi_r = w - D_G\Psi^\varepsilon\left(\frac{u_r-u}{r}\right)$ of the energy functional $\mathcal{E}_{t+r}(u_r)$ does not appear so that we would not succeed in deriving a priori estimates for ξ_r. However, if $\mathcal{E}_t$ is convex, then the relation (3.3.3) implies with Lemma 2.2.2 and 2.3.1 that for $r = r_0$, there holds

$$\begin{aligned}
&r_0\Psi_u^\varepsilon\left(\frac{u_{r_0}-u}{r_0}\right) + r_0\Psi_u^{\varepsilon,*}(w-\xi_{r_0})\\
&= \langle w-\xi_{r_0}, u_{r_0}-u\rangle,\\
&= \langle w, u_{r_0}-u\rangle - \langle \xi_{r_0}, u_{r_0}-u\rangle\\
&\leq \mathcal{E}_{\mathrm{t}+r_0}(u) - \mathcal{E}_{\mathrm{t}+r_0}(u_{r_0}) + \langle w, u_{r_0}-u\rangle\\
&= \mathcal{E}_{\mathrm{t}}(u) - \mathcal{E}_{\mathrm{t}+r_0}(u_{r_0}) + \int_0^{r_0} \partial_t\mathcal{E}_{\mathrm{t}+r}(u)\,\mathrm{d}r + \langle w, u_{r_0}-u\rangle,
\end{aligned}$$

which again leads to appropriate a priori estimates. It seems that the inevitable assumption is the strong-weak closedness of the graph of the subdifferential $\partial\mathcal{E}_t$, i.e., Condition (3.Ee) is sufficient to obtain the right estimates.

3.4 Discrete energy-dissipation inequality and a priori estimates

In this section, we derive with virtue of Lemma 3.3.2 a priori estimates on *approximate solutions*. As a result of this, we need to define appropriate interpolations of the approximative values $U_\tau^k \in V$, $k = 1, \ldots, N$, which we call interchangeably, *discrete solutions*. To this aim, let for a given time step $\tau = T/N$ the initial value $U_\tau^0 \in D$ be given. Furthermore, let $(U_\tau^n)_{n=1}^N \subset D$ be the sequence of approximate values which under the assumptions of Lemma 3.3.1 are always existent and are given by the variational approximation scheme (3.1.5). Then, for a given sequence of approximate values $(U_\tau^n)_{n=0}^N \subset D$, we define the piecewise constant and linear interpolations in the following way

$$\begin{aligned}
&\overline{U}_\tau(0) = \underline{U}_\tau(0) = \widehat{U}_\tau(0) := U_\tau^0 \text{ and}\\
&\underline{U}_\tau(t) := U_\tau^{n-1}, \quad \widehat{U}_\tau(t) := \frac{t_n - t}{\tau}U_\tau^{n-1} + \frac{t-t_{n-1}}{\tau}U_\tau^n \quad \text{for } t \in [t_{n-1}, t_n),\\
&\overline{U}_\tau(t) := U_\tau^n \quad \text{for } t \in (t_{n-1}, t_n] \quad \text{and } n = 1, \ldots, N.
\end{aligned} \tag{3.4.1}$$

It is easy to verify that the piecewise linear interpolation is piecewise differentiable in the classical sense with the piecewise derivative

$$\widehat{U}_\tau'(t) = \frac{U_\tau^n - U_\tau^{n-1}}{\tau} \quad \text{for } t \in [t_{n-1}, t_n),\ n = 1, \ldots, N,$$

Apart from the piecewise constant and piecewise linear interpolations, which by themselves are not target-aimed, we need to introduce the so-called De Giorgi interpolation $\widetilde{U}_\tau$ in order to obtain proper estimates and especially the so-called *energy-dissipation estimate*.

The above mentioned De Giorgi variational interpolation $\widetilde{U}_\tau$ is defined as any function whose values satisfies

$$\begin{cases} \widetilde{U}_\tau(0) := U_\tau^0, \\ \widetilde{U}_\tau(t) \in J_{r,t}(B(t_n, U_\tau^{n-1}); U_\tau^{n-1}) \quad \text{for } t = t_{n-1} + r \in (t_{n-1}, t_n] \end{cases} \tag{3.4.2}$$

and $n = 1, 2, \ldots, N$. Since the De Giorgi scheme (3.4.2) yields the variational scheme given by (3.1.5) for the choice $r = \tau$, we assume without loss of generality that all interpolations coincide on the points t_n, $n = 1, \ldots, N$, i.e.,

$$\widetilde{U}_\tau(t_n) = \overline{U}_\tau(t_n) = \underline{U}_\tau(t_n) = \widehat{U}_\tau(t_n) = U_\tau^n \quad \text{for all } n = 1, \cdots, N.$$

We make the important observation that by virtue of Lemma 3.3.2, the De Giorgi interpolation can be chosen to be measurable, since the same lemma ensures the existence of a measurable selection of the mapping $r \mapsto u_r \in J_{r,\mathrm{t}}(w; u)$.

Furthermore, we introduce for notational reasons the piecewise constant interpolations $\overline{\mathbf{t}}_\tau : [0, T] \to [0, T]$ and $\underline{\mathbf{t}}_\tau : [0, T] \to [0, T]$ of the time points t_n, $n = 1, \ldots, N$:

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

It is easy to verify that $\overline{\mathbf{t}}_\tau(t) \to t$ and $\underline{\mathbf{t}}_\tau(t) \to t$ as $\tau \to 0$.

The following simple example gives a rough idea of the unusually defined De Giorgi interpolation.

Example 3.4.1 We consider the initial value problem

$$\begin{cases} u'(t) = -u(t), \quad t \in (0, 6), \\ u(0) = 2 \end{cases}$$

for which the exact solution is given by $u(t) = 2\mathrm{e}^{-t}$, $t \in [0, 2]$. We discretize the equation via the implicit Euler scheme with an equidistant step size $\tau = T/N = 6/6 = 1$ so that the grid points are given by $t_n = n$, $n = 1, \ldots, 6$. The De Giorgi interpolation is then equivalently given by the corresponding Euler-Lagrange equation

$$\frac{\widetilde{U}_\tau(t) - U_\tau^{n-1}}{t - t_{n-1}} = -\widetilde{U}_\tau(t) \quad \text{or} \quad \widetilde{U}_\tau(t) = \frac{U_\tau^{n-1}}{1 + t - t_{n-1}} \quad \text{for } t \in (t_{n-1}, t_n]$$

and $n = 1, \ldots, N$. Figure 3.2 illustrates the De Giorgi interpolation in comparison with the piecewise linear interpolation $\widehat{U}_\tau$ and the exact solution u.

To understand the principle of the De Giorgi interpolation, we consider the following example of a nonlinear ODE.

Example 3.4.2 We calculate $\widetilde{U}_\tau$ associated to the initial value problem We consider the initial value problem

$$\begin{cases} u'(t) = -u^2(t), \quad t \in (0, 2), \\ u(0) = 2. \end{cases}$$

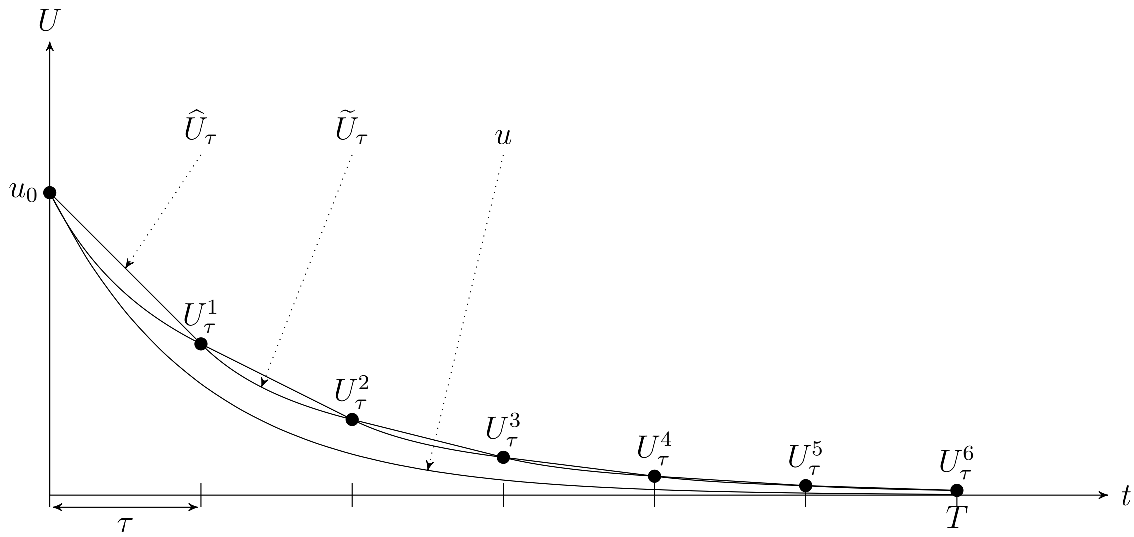

Figure 3.2: The figure shows the graphs of the piecewise linear interpolation $\widehat{U}_\tau$ and the De Giorgi interpolation $\widetilde{U}_\tau$ both approximating the function $u(t) = 2\mathrm{e}^{-t}, t \in [0, 6]$.

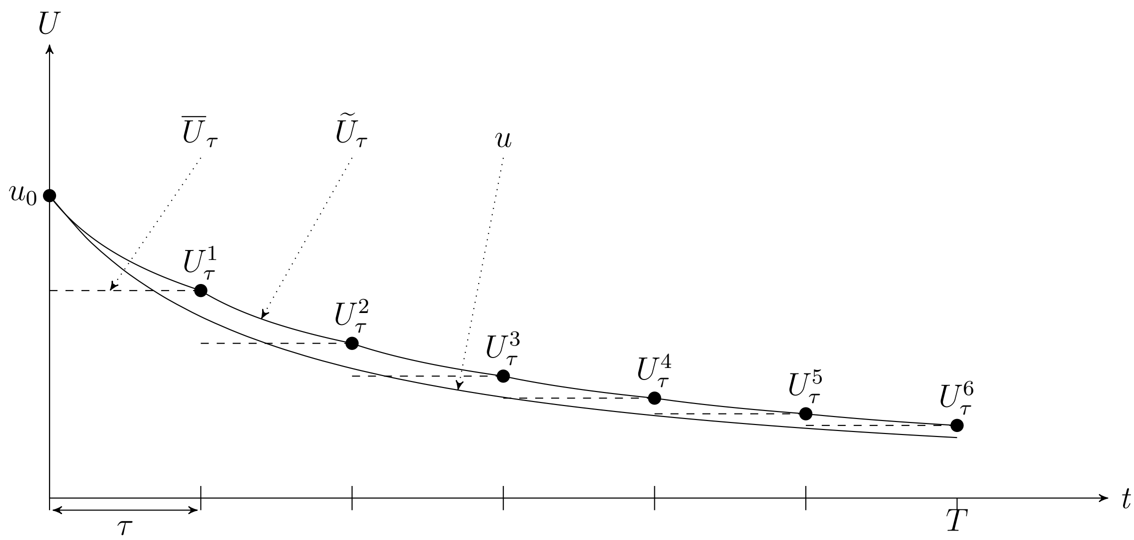

Figure 3.3: The figure shows the graphs of the piecewise constant interpolation $\overline{U}_\tau$ and the De Giorgi interpolation $\widetilde{U}_\tau$ both approximating the function $u(t) = (t + 0.5)^{-1}, t \in [0, 2]$.

Here, the exact solution to this problem is given by $u(t) = (t + 0.5)^{-1}$, $t \in [0, 2]$. With the same discretization as above, the De Giorgi interpolation is equivalently given by

$$\frac{\widetilde{U}_\tau(t) - U_\tau^{n-1}}{t - t_{n-1}} = -\widetilde{U}_\tau^2(t) \quad \text{for } t \in (t_{n-1}, t_n], n = 1, \dots, N. \tag{3.4.3}$$

Choosing the positive solution to (3.4.3), we obtain the explicit formula

$$\widetilde{U}_\tau(t) = \frac{\sqrt{4(t - t_{n-1})U_\tau^{n-1} + 1} - 1}{2(t - t_{n-1})}, \quad \text{for } t \in (t_{n-1}, t_n],\ n = 1, \dots, N.$$

Figure 3.3 compares the exact solution to the De Giorgi interpolation. We can see that $\widetilde{U}_\tau$ is a "better" approximation to the exact solution than $\overline{U}_\tau$ and $\widehat{U}_\tau$ in the sense that it catches more of the dynamic of the exact solution. This is because the De Giorgi interpolation not only approximates, loosely speaking, the exact solution on the grid points t_n but their difference quotient approximates u' at the same time. This stems from the fact that the values of $\widehat{U}_\tau$ are in every point determined by (3.4.3) or more generally by (3.4.4) below. However, the function $\widetilde{U}_\tau$ quickly becomes very difficult for highly nonlinear problems. If we consider the initial value problem

$$\begin{cases} u'(t) = -u^3(t), & t \in (0, 2), \\ u(0) = 2, \end{cases}$$

then one can show that $\widetilde{U}_\tau$ is given by

$$\begin{aligned} \widetilde{U}_\tau(t) = & \frac{\sqrt[3]{9(t - t_{n-1})^2 U_\tau^{n-1} + \sqrt{3}\sqrt{27(t - t_{n-1})^4 (U_\tau^{n-1})^2 + 4(t - t_{n-1})^3}}}{\sqrt[3]{18}(t - t_{n-1})} \\ & - \frac{\sqrt[3]{\frac{2}{3}}}{\sqrt[3]{9(t - t_{n-1})^2 U_\tau^{n-1} + \sqrt{3}\sqrt{27(t - t_{n-1})^4 (U_\tau^{n-1})^2 + 4(t - t_{n-1})^3}}} \end{aligned}$$

for all $t \in (t_{n-1}, t_n]$ and $n = 1, \dots, N$. Here, we did not even consider non-quadratic dissipation potentials. Figure 3.3 illustrates the De Giorgi interpolation in comparison with the piecewise constant interpolation $\overline{U}_\tau$ and the exact solution u.

As we mentioned in the previous example, the De Giorgi variational interpolation $\widetilde{U}_\tau(t)$ solves by definition a minimization problem at every time $t \in (0, T]$ so that $\widetilde{U}_\tau(t)$ satisfies the inclusion

$$\mathrm{D}_G \Psi^\varepsilon_{U_\tau^{n-1}} \left(\frac{\widetilde{U}_\tau(t) - U_\tau^{n-1}}{t - t_{n-1}} \right) + \partial \mathcal{E}_t(\widetilde{U}_\tau(t)) \ni B(t_n, U_\tau^{n-1}), \quad t \in (t_{n-1}, t_n] \tag{3.4.4}$$

for $n = 1, \dots, N$, whereas the piecewise constant and linear interpolations together merely satisfy the inclusion (3.4.4) on the gridpoints t_n for all $n = 1, \dots, N$. So, theoretically, the De Giorgi interpolation may be a better approximation to the solution of perturbed gradient system (3.0.1) than $\overline{U}_\tau$ or $\widehat{U}_\tau$. In fact, the De Giorgi interpolation approximates the solutions just well enough to obtain the right estimate, the so-called discrete upper energy estimate for the discrete solutions. The latter estimate is the discrete analogue to the energy-dissipation balance (3.0.7) which is obtained by plugging in the De Giorgi interpolation into the inequality (3.3.8) substituting $r \mapsto u_r$ in Lemma 3.3.2. We will see this in more detail in the following lemma together with other a priori estimates.

Lemma 3.4.3 (A priori estimates) *Let the perturbed gradient system* $(V, \mathcal{E}, \Psi, B)$ *satisfy the Assumptions* (3.Ψ), (3.E) *and* (3.B). *Furthermore, let* $\widetilde{U}_\tau, \overline{U}_\tau, \underline{U}_\tau$ *and* $\widehat{U}_\tau$ *be the interpolations defined in* (3.4.1)-(3.4.2) *corresponding to an initial datum* $U_\tau^0 \in D$ *and the step size* $\tau > 0$. *Then, defining* $\xi_\tau : [0,T] \to V^*$ *by*

$$\xi_\tau(t) = B(t_n, U_\tau^{n-1}) - D_G \Psi^\varepsilon_{U_\tau^{n-1}} \left(\frac{\widetilde{U}_\tau(t) - U_\tau^{n-1}}{t - t_{n-1}} \right) \tag{3.4.5}$$

for all $t = t_{n-1} + r \in (t_{n-1}, t_n]$ *and* $n = 1, \dots, N$, *there holds*

$$\xi_\tau(t) \in \partial \mathcal{E}_{t_{n-1}+r}(\widetilde{U}_\tau(t)) \quad \textit{for all } t \in (0,T) \tag{3.4.6}$$

and the discrete energy-dissipation inequality

$$\begin{aligned} &\mathcal{E}_{\bar{\mathbf{t}}_\tau(t)}(\overline{U}_\tau(t)) + \int_{\bar{\mathbf{t}}_\tau(s)}^{\bar{\mathbf{t}}_\tau(t)} \left(\Psi^\varepsilon_{\underline{U}_\tau(r)} \left(\widehat{U}'_\tau(r) \right) + \Psi^{\varepsilon,*}_{\underline{U}_\tau(r)} \left(B(\bar{\mathbf{t}}_\tau(r), \underline{U}_\tau(r)) - \xi_\tau(r) \right) \right) \mathrm{d}r \\ &\le \mathcal{E}_{\bar{\mathbf{t}}_\tau(s)}(\overline{U}_\tau(s)) + \int_{\bar{\mathbf{t}}_\tau(s)}^{\bar{\mathbf{t}}_\tau(t)} \partial_r \mathcal{E}_r(\widetilde{U}_\tau(r)) \, \mathrm{d}r + \int_{\bar{\mathbf{t}}_\tau(s)}^{\bar{\mathbf{t}}_\tau(t)} \langle B(\bar{\mathbf{t}}_\tau(r), \underline{U}_\tau(r)), \widehat{U}'_\tau(r) \rangle \, \mathrm{d}r \end{aligned} \tag{3.4.7}$$

holds for all $0 \le s < t \le T$ *and* $\varepsilon > 0$. *If we additionally assume* $\mathcal{E}_0(U_\tau^0) \le \hat{C}$ *for all sequences of step sizes* τ *for a constant* $\hat{C} > 0$, *then there exist positive numbers* $M, \tau^* > 0$ *such that*

$$\sup_{t \in [0,T]} \mathcal{E}_t(\overline{U}_\tau(t)) \le M, \qquad \sup_{t \in [0,T]} \mathcal{E}_t(\widetilde{U}_\tau(t)) \le M, \tag{3.4.8}$$

$$\sup_{t \in [0,T]} \| B(\bar{\mathbf{t}}_\tau(t), \underline{U}_\tau(t)) \|_* \le M, \qquad \sup_{t \in (0,T)} |\partial_t \mathcal{E}_t(\widetilde{U}_\tau(t))| \le M, \tag{3.4.9}$$

$$\int_0^T \left(\Psi^\varepsilon_{\underline{U}_\tau(r)} \left(\widehat{U}'_\tau(r) \right) + \Psi^{\varepsilon,*}_{\underline{U}_\tau(r)} \left(B(\bar{\mathbf{t}}_\tau(r), \underline{U}_\tau(r)) - \xi_\tau(r) \right) \right) \mathrm{d}r \le M \tag{3.4.10}$$

holds for all $0 < \tau \le \tau^*$ *and uniformly in* $\varepsilon > 0$. *Besides,* $(\widehat{U}'_\tau)_{0<\tau\le\tau^*} \subset \mathrm{L}^1(0,T;V)$ *and* $(\xi_\tau)_{0<\tau\le\tau^*} \subset \mathrm{L}^1(0,T;V^*)$ *are equi-integrable with respect to* τ *in* $\mathrm{L}^1(0,T;V)$ *and* $\mathrm{L}^1(0,T;V^*)$, *respectively. Finally, there holds*

$$\sup_{t \in [0,T]} \left(\| \underline{U}_\tau(t) - \overline{U}_\tau(t) \| + \| \widehat{U}_\tau(t) - \overline{U}_\tau(t) \| + + \| \widetilde{U}_\tau(t) - \underline{U}_\tau(t) \| \right) \to 0 \tag{3.4.11}$$

as $\tau \to 0$.

Proof. The mapping $\xi_\tau : [0,T] \to V^*$ satisfying the inclusion (3.4.6) follows from the variational sum rule (3.3.3) with the choice $r = \tau$, $t = t_{n-1}, u_r = \widetilde{U}_\tau(t), u = U_\tau^{n-1}$ and $w = B(t_n, U_\tau^{n-1})$, $n = 1, \dots, N$, satisfying (3.4.6) and (3.4.5). The measurability of $\xi_\tau : (0,T) \to V^*$ is given by Lemma 3.3.2.

We seek now to show the discrete energy-dissipation inequality. For that, we make use of inequality (3.3.8) in Lemma 3.3.2 and make the choice t $= t_{n-1}, u = U_\tau^{n-1}, r_0 = t - t_{n-1}, u_{r_0} = \widetilde{U}_\tau(t), u_r = \widetilde{U}_\tau(t_{n-1}+r)$ and $\xi_r = \xi_\tau(t_{n-1}+r)$, where $t \in (t_{n-1}, t_n]$. Then, after substitution we find

$$
\begin{aligned}
&(t-t_{n-1})\Psi^\varepsilon_{U_\tau^{n-1}}\left(\frac{\widetilde{U}_\tau(t)-U_\tau^{n-1}}{t-t_{n-1}}\right)+\int_{t_{n-1}}^t \Psi^{\varepsilon,*}_{U_\tau^{n-1}}\Big(B(t_n,U_\tau^{n-1})-\xi_\tau(r)\Big)\,\mathrm{d}r+\mathcal{E}_t(\widetilde{U}_\tau(t))\\
&\le \mathcal{E}_{t_{n-1}}(U_\tau^{n-1})+\int_{t_{n-1}}^t \partial_r\mathcal{E}_r(\widetilde{U}_\tau(r))\,\mathrm{d}r+\langle B(t_n,U_\tau^{n-1}),U_\tau^n-U_\tau^{n-1}\rangle \qquad (3.4.12)
\end{aligned}
$$

for any $t \in (t_{n-1}, t_n]$ and all $n = 1, \ldots, N$. By choosing $t = t_n$, we obtain particularly

$$
\begin{aligned}
&\int_{t_{n-1}}^{t_n}\Big(\Psi^\varepsilon_{\underline{U}_\tau(r)}\left(\widehat{U}'_\tau(r)\right)+\Psi^{\varepsilon,*}_{\underline{U}_\tau(r)}\left(B(t_n,\underline{U}_\tau(r))-\xi_\tau(r)\right)\Big)\,\mathrm{d}r+\mathcal{E}_{t_n}(\overline{U}_\tau(t_n))\\
&\le \mathcal{E}_{t_{n-1}}(\underline{U}_\tau(t_{n-1}))+\int_{t_{n-1}}^{t_n}\partial_r\mathcal{E}_r(\widetilde{U}_\tau(r))\,\mathrm{d}r+\int_{t_{n-1}}^{t_n}\langle B(t_n,\underline{U}_\tau(r)),\widehat{U}'_\tau(r)\rangle\,\mathrm{d}r \qquad (3.4.13)
\end{aligned}
$$

for all $n = 1, \cdots, N$. Then, the discrete energy-dissipation inequality is obtainable by summing up the inequalities from $n = 1$ to $n = N$. We continue with deriving the estimates (3.4.8)-(5.3.7). We will establish the estimates by using the discrete version of the GRONWALL Lemma. First, from Assumption (3.Bb) and the FENCHEL-YOUNG inequality, we obtain

$$
\begin{aligned}
&\int_{t_{n-1}}^{t_n}\langle B(t_n,\underline{U}_\tau(r)),\widehat{U}'_\tau(r)\rangle\,\mathrm{d}r\\
&\le c\int_{t_{n-1}}^{t_n}\Psi^\varepsilon_{\underline{U}_\tau(r)}\left(\widehat{U}'_\tau(r)\right)\,\mathrm{d}r+c\int_{t_{n-1}}^{t_n}\Psi^{\varepsilon,*}_{\underline{U}_\tau(r)}\left(\frac{B(t_n,\underline{U}_\tau(r))}{c}\right)\,\mathrm{d}r\\
&\le c\int_{t_{n-1}}^{t_n}\Psi^\varepsilon_{\underline{U}_\tau(r)}\left(\widehat{U}'_\tau(r)\right)\,\mathrm{d}r+\tau\beta(1+\mathcal{E}_{t_{n-1}}(U_\tau^{n-1}))\\
&\le c\int_{t_{n-1}}^{t_n}\Psi^\varepsilon_{\underline{U}_\tau(r)}\left(\widehat{U}'_\tau(r)\right)\,\mathrm{d}r+\tau\beta(1+G(U_\tau^{n-1})),
\end{aligned}
$$

where $c \in (0, 1)$ is from Assumption (3.Bb). Together with the latter inequality, inequality (3.4.13) yields, while employing again Assumption (3.Bb),

$$
\begin{aligned}
&\int_{t_{n-1}}^{t_n}\Big((1-c)\Psi^\varepsilon_{\underline{U}_\tau(r)}\left(\widehat{U}'_\tau(r)\right)+\Psi^{\varepsilon,*}_{\underline{U}_\tau(r)}\left(B(t_n,\underline{U}_\tau(r))-\xi_\tau(r)\right)\Big)\,\mathrm{d}r+\mathcal{E}_{t_n}(U_\tau^n)\\
&\le \mathcal{E}_{t_{n-1}}(U_\tau^{n-1})+\int_{t_{n-1}}^{t_n}\partial_r\mathcal{E}_r(\widetilde{U}_\tau(r))\,\mathrm{d}r+\tau\beta(1+G(U_\tau^{n-1}))\\
&\le \mathcal{E}_{t_{n-1}}(U_\tau^{n-1})+\tau\beta(1+G(U_\tau^{n-1}))+C_1\int_{t_{n-1}}^{t_n}G(U_\tau^{n-1})\,\mathrm{d}r\\
&\quad +C_1\int_{t_{n-1}}^{t_n}(r-t_{n-1})\Psi^{\varepsilon,*}_{U_\tau^{n-1}}(B(t_n,U_\tau^{n-1}))\,\mathrm{d}r \qquad (3.4.14)\\
&\le \mathcal{E}_{t_{n-1}}(U_\tau^{n-1})+\tau\beta(1+G(U_\tau^{n-1}))+C_1\int_{t_{n-1}}^{t_n}G(U_\tau^{n-1})\,\mathrm{d}r\\
&\quad +C_1\int_{t_{n-1}}^{t_n}c\tau\Psi^{\varepsilon,*}_{\underline{U}_\tau(r)}\left(\frac{B(t_n,\underline{U}_\tau(r))}{c}\right)\,\mathrm{d}r \qquad (3.4.15)
\end{aligned}
$$

$$
\begin{aligned}
&\leq \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) + \tau\beta(1+G(U_\tau^{n-1})) + C_1\tau G(U_\tau^{n-1}) \\
&\quad + C_1\tau\beta(1+G(U_\tau^{n-1})) \\
&= \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) + \tau 2(\beta + C_1)G(U_\tau^{n-1}) + 2\tau(\beta + C_1)
\end{aligned} \tag{3.4.16}
$$

for all $n = 1, \ldots, N$ and $0 < \tau \leq 1$, where in (3.4.14) we used the estimates

$$
\begin{aligned}
\partial_t \mathcal{E}_t(\widetilde{U}_\tau(t)) &\leq CG(\widetilde{U}_\tau(t)) \\
&\leq C_1(G(U_\tau^{n-1}) + (t - t_{n-1})\Psi^{\varepsilon,*}_{U_\tau^{n-1}}(B(t_n, U_\tau^{n-1}))), \quad t \in (t_{n-1}, t_n],
\end{aligned} \tag{3.4.17}
$$

coming from Lemma 3.3.2 and Assumption (3.Ec). In (3.4.15), we used the fact that the mapping $r \mapsto r\Psi_v^{\varepsilon,*}\left(\frac{\xi}{r}\right)$ is increasing on $(0, +\infty)$ for every $v \in V, \xi \in V^*$, which follows from the convexity of $\Psi_v^{\varepsilon,*}$ on V^* and the fact that $\Psi_v^{\varepsilon,*}(0) = 0$ for all $v \in U$. We set $A_1 := C_1\hat{C} + 2C_1T(\beta + C_1)$ and $\alpha = 2C_1(\beta + C_1)$, and sum up the inequalities (3.4.16) to obtain

$$
\begin{aligned}
&G(U_\tau^n) \leq C_1\mathcal{E}_{t_n}(U_\tau^n) \\
&\leq C_1\left(\int_0^{t_n}\left((1-c)\Psi^\varepsilon_{\underline{U}_\tau(r)}\left(\widehat{U}'_\tau(r)\right) + \Psi^{\varepsilon,*}_{\underline{U}_\tau(r)}\left(B(t_n, \underline{U}_\tau(r)) - \xi_\tau(r)\right)\right) \mathrm{d}r\right) \\
&\quad + C_1\mathcal{E}_{t_n}(U_\tau^n)
\end{aligned} \tag{3.4.18}
$$

$$
\leq A_1 + \alpha\sum_{k=1}^n \tau G(U_\tau^{k-1}) \tag{3.4.19}
$$

for all $n = 1, \ldots, N$ and $0 < \tau \leq 1$, where we used the non-negativity of the dissipation potential and its conjugate. Now, for sufficiently small step sizes, i.e., $\tau < \min\{\delta/\alpha, \delta\} =: \tau^*$ for a $\delta \in (0, 1)$, the discrete Gronwall lemma (Lemma A.1.2) implies

$$
G(U_\tau^n) \leq A_2 e^{\tilde{\beta}\sum_{k=0}^{n-1}\tau_k} \leq A_2 e^{\tilde{\beta}T} \quad \text{for all } n = 1, \ldots, N,
$$

where we set $\tilde{\beta} = \frac{\alpha}{1-\alpha\tau}$ and $A_2 = \frac{A_1}{1-\alpha\tau}$ as well as $\tau_0 = 0$, whence the uniform bound of $G(U_\tau^n)$ for all $n = 1, \ldots, N$ and $0 < \tau < \tau^*$. In view of (3.4.17), (3.4.18), and Assumption (3.Bb) and (3.Ψb), we deduce also the rest of the bounds in (3.4.8)-(5.3.7). The constant $M > 0$ can be chosen to be the sum of all constants obtained from the shown inequalities. We proceed by showing the equi-integrability of $(\widehat{U}'_\tau)_{0<\tau\leq\tau^*}, (B(\overline{\mathbf{t}}_\tau, \underline{U}_\tau))_{0<\tau\leq\tau^*}$ and $(\xi_\tau)_{0<\tau\leq\tau^*}$ in $\mathrm{L}^1(0, T; V)$ and $\mathrm{L}^1(0, T; V^*)$, respectively. This essentially follows from the superlinear growth of Ψ_u^ε and $\Psi_u^{\varepsilon,*}$ (Assumption (3.Ψb)) with the criterion of De La Vallée Poussin for equi-integrability. We show the equi-integrability exemplarily for $(\widehat{U}'_\tau)_{0<\tau\leq\tau^*}$ without using the weak compactness criterion of De La Vallée Poussin which actually can be proved along the same lines. Let $\delta > 0$ be an arbitrary positive real number. Then, for M and M/δ, there exist by Assumption (3.Ψb) positive numbers $K_1, K_2 > 0$ such that

$$
\Psi_u^\varepsilon(v) \geq \frac{M}{\delta}\|v\| \tag{3.4.20}
$$

for all $v \in V$ with $\|v\| \geq K_1$, and all $u \in D$ with $G(u) \leq M$. Then, by (3.4.20), (3.4.8) and (5.3.7), we have

$$\int_{\{t\in[0,T]:\|\widehat{U}_\tau'(t)\|\geq K_1\}} \|\widehat{U}_\tau'(t)\| \, \mathrm{d}t \leq \frac{\delta}{\widetilde{M}} \int_{\{t\in[0,T]:\|\widehat{U}_\tau'(t)\|\geq K_1\}} \Psi^{\varepsilon}_{\underline{U}_\tau(t)}(\widehat{U}_\tau'(t)) \, \mathrm{d}t \leq \delta$$

for all $0 < \tau \leq \tau^*$, which shows the equi-integrability. Again, in regard to the superlinear growth of $\Psi_u^{\varepsilon,*}$, it can also be shown that $(B(\overline{\mathbf{t}}_\tau, \underline{U}_\tau) - \xi_\tau)_{0<\tau\leq\tau^*}$ is equi-integrable in $\mathrm{L}^1(0,T;V^*)$. Since $B(\overline{\mathbf{t}}_\tau, \underline{U}_\tau)_{0<\tau\leq\tau^*}$ is uniformly bounded, we obtain the equi-integrability of $(\xi_\tau)_{0<\tau\leq\tau^*}$ in $\mathrm{L}^1(0,T;V^*)$. In order to show the convergences (3.4.11), we first note that inequality (3.4.12) together with (3.4.8) and (5.3.7) imply

$$\sup_{t\in[0,T]} (t - \underline{\mathbf{t}}_\tau(t)) \Psi^{\varepsilon}_{\underline{U}_\tau(t)} \left(\frac{\widetilde{U}_\tau(t) - \underline{U}_\tau(t)}{t - \underline{\mathbf{t}}_\tau(t)} \right) \leq C_2$$

for a constant $C_2 > 0$. Employing once again Assumption (3.Ψb), we obtain for every $R > 0$ and $\gamma > 0$ a positive constant $K > 0$ satisfying

$$\begin{aligned} \gamma \|\widetilde{U}_\tau(t) - \underline{U}_\tau(t)\| &\leq (t - \underline{\mathbf{t}}_\tau(t)) \Psi^{\varepsilon}_{\underline{U}_\tau(t)} \left(\frac{\widetilde{U}_\tau(t) - \underline{U}_\tau(t)}{t - \underline{\mathbf{t}}_\tau(t)} \right) + (t - \underline{\mathbf{t}}_\tau(t))\gamma K \\ &\leq M + \tau\gamma K \quad \text{ for all } t \in [0,T] \text{ and all } 0 < \tau < \tau^*. \end{aligned}$$

Finally, taking the supremum over all $t \in [0,T]$ and then taking the limes superior as $\tau \to 0$, we obtain

$$\gamma \limsup_{\tau\to 0} \sup_{t\in[0,T]} \|\widetilde{U}_\tau(t) - \underline{U}_\tau(t)\| \leq M \tag{3.4.21}$$

for any $\gamma > 0$, which necessarily yields $\lim_{\tau\to 0} \sup_{t\in[0,T]} \|\widetilde{U}_\tau(t) - \underline{U}_\tau(t)\| = 0$. Choosing in the inequality (3.4.21) specifically $t = t_n$, $n = 1, \ldots, N$, we also obtain $\lim_{\tau\to 0} \sup_{t\in[0,T]} \|\overline{U}_\tau(t) - \underline{U}_\tau(t)\| = 0$, which in turn implies $\lim_{\tau\to 0} \sup_{t\in[0,T]} \|\widehat{U}_\tau(t) - \overline{U}_\tau(t)\| = 0$ and hence the completion of the proof. □

3.5 Compactness and parameterized Young measures

In this section, we show the compactness of the interpolations in certain spaces.

Lemma 3.5.1 *Let the perturbed gradient system $(V, \mathcal{E}, \Psi, B)$ be given and satisfy the Assumption* (3.Ψ), (3.E) *and* (3.B). *Let $u_0 \in D$ and for a sequence $(\tau_n)_{n\in\mathbb{N}}$ with $\tau_n \to 0$ as $n \to \infty$, let $(U^0_{\tau_n})_{n\in\mathbb{N}} \subset D$ be a sequence of initial values fulfilling $U^0_{\tau_n} \to u_0$ and $\mathcal{E}_0(U^0_{\tau_n}) \to \mathcal{E}_0(u_0)$ as $n \to \infty$. Furthermore, for $n \in \mathbb{N}$, let $\widetilde{U}_{\tau_n}, \overline{U}_{\tau_n}, \underline{U}_{\tau_n}$ and $\widehat{U}_{\tau_n}$ be the interpolations defined in* (3.4.1) *and* (3.4.2) *associated to an initial value $U^0_{\tau_n}$, and ξ_{τ_n} the subgradient of $\mathcal{E}_t$ satisfying* (3.4.6) *and* (3.4.5). *Then, there exists a subsequence $(\tau_{n_k})_{k\in\mathbb{N}}$, an absolutely continuous curve $u \in \mathrm{AC}([0,T];V)$ with $u(0) = u_0$, an integrable function $\xi \in \mathrm{L}^1(0,T;V^*)$, a function $\mathscr{E} : [0,T] \to \mathbb{R}$ of bounded* *variation, an essentially bounded function* $\mathscr{P} \in \mathrm{L}^\infty(0,T)$*, and a parameterized* YOUNG *measure* $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in \mathscr{Y}(0,T;V \times V^* \times \mathbb{R})$ *such that*

$$\widehat{U}_{\tau_{n_k}} \to u \quad \textit{in } \mathrm{C}([0,T];V), \tag{3.5.1a}$$

$$\overline{U}_{\tau_{n_k}}, \underline{U}_{\tau_{n_k}}, \widetilde{U}_{\tau_{n_k}} \to u \quad \textit{in } \mathrm{L}^\infty(0,T;V), \tag{3.5.1b}$$

$$\widehat{U}'_{\tau_{n_k}} \rightharpoonup u' \quad \textit{in } \mathrm{L}^1(0,T;V), \tag{3.5.1c}$$

$$\xi_{\tau_{n_k}} \rightharpoonup \xi \quad \textit{in } \mathrm{L}^1(0,T;V^*), \tag{3.5.1d}$$

$$B(\bar{\mathfrak{t}}_{\tau_{n_k}}, \underline{U}_{\tau_{n_k}}) \to B(\cdot, u(\cdot)) \quad \textit{in } \mathrm{L}^\infty(0,T;V^*), \tag{3.5.1e}$$

$$\partial_t\mathcal{E}_t(\widetilde{U}_{\tau_{n_k}}(t)) \overset{*}{\rightharpoonup} \mathscr{P} \quad \textit{in } \mathrm{L}^\infty(0,T), \tag{3.5.1f}$$

as $k \to \infty$*, and the weak limits satisfy*

$$u'(t) = \int_{V\times V^*\times\mathbb{R}} v \,\mathrm{d}\mu_t(v,\zeta,p) \quad \textit{for a.e. } t \in (0,T), \tag{3.5.2a}$$

$$\xi(t) = \int_{V\times V^*\times\mathbb{R}} \zeta \,\mathrm{d}\mu_t(v,\zeta,p) \quad \textit{for a.e. } t \in (0,T), \tag{3.5.2b}$$

$$\mathscr{P}(t) = \int_{V\times V^*\times\mathbb{R}} p \,\mathrm{d}\mu_t(v,\zeta,p) \quad \textit{for a.e. } t \in (0,T). \tag{3.5.2c}$$

Finally, there holds

$$\begin{cases} \mathcal{E}_t(\overline{U}_{\tau_{n_k}}(t)) \to \mathscr{E}(t) \quad \textit{as } k \to \infty & \textit{for all } t \in [0,T], \\ \mathcal{E}_t(u(t)) \le \mathscr{E}(t) & \textit{for all } t \in [0,T], \\ \mathcal{E}_0(u_0) = \mathscr{E}(0) \quad \textit{and} \quad \mathcal{E}_t(u(t)) = \mathscr{E}(t) & \textit{for a.e. } t \in (0,T), \\ \mathscr{P}(t) \le \partial_t\mathcal{E}_t(u(t)) & \textit{for a.e. } t \in (0,T), \\ \xi(t) \in \partial\mathcal{E}_t(u(t)) & \textit{for a.e. } t \in (0,T), \end{cases} \tag{3.5.3}$$

and the energy-dissipation inequality

$$\begin{aligned} &\int_s^t \Big(\Psi^\varepsilon_{u(r)}(u'(r)) + \Psi^{\varepsilon,*}_{u(r)}(B(r,u(r)) - \xi(r))\Big)\,\mathrm{d}r + \mathcal{E}_t(u(t)) \\ &\le \mathscr{E}(s) + \int_s^t \partial_t\mathcal{E}_t(u(t))\,\mathrm{d}r + \int_s^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r \end{aligned} \tag{3.5.4}$$

holds for all $0 \le s < t \le T$*.*

Proof. We fix an arbitrary initial value $u_0 \in D$, and for a sequence $(\tau_n)_{n\in\mathbb{N}}$ with $\tau_n \to 0$ as $n \to \infty$, let $(U^0_{\tau_n})_{n\in\mathbb{N}} \subset D$ be such that $U^0_{\tau_n} \to u_0$ as $n \to \infty$. We assume without loss of generality that the step sizes are sufficiently small, i.e., $\tau_n < \tau^*$ for all $n \in \mathbb{N}$. The assertion (3.5.1a) then follows by means of the ARZELÀ-ASCOLI theorem applied to the continuous functions $(\widehat{U}_{\tau_n})_{n\in\mathbb{N}} \subset C([0,T];V)$: the equi-continuity of $(\widehat{U}_{\tau_n})_{n\in\mathbb{N}}$ is a consequence of the equi-integrability of $(\widehat{U}'_{\tau_n})_{n\in\mathbb{N}}$ leading to the LIPSCHITZ continuity of $(\widehat{U}_{\tau_n})_{n\in\mathbb{N}}$ with a LIPSCHITZ constant independent of the step size. Due to (3.4.8), showing that $(\overline{U}_{\tau_n}(t))_{t\in[0,T]}$ belongs to a sublevel set of the energy functional $\mathcal{E}_t$ independent of $n \in \mathbb{N}$, which by Assumption (3.Eb), is supposed to be compact in V, MAZUR's lemma implies that the convex hull of $(\overline{U}_{\tau_n}(t))_{t\in[0,T]}$ is by itself compact in V, and therefore also $(\overline{U}_{\tau_n}(t))_{t\in[0,T]}$ for all $n \in \mathbb{N}$. Thus, by the theorem of ARZELÀ-ASCOLI, there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ and a continuous function $u \in C([0,T];V)$ such that $\widehat{U}_{\tau_{n_k}} \to u$ in $C([0,T];V)$ as $k \to \infty$, and in particular $u(0) = u_0$. The convergences in (3.5.1b) follow then from those in (3.4.11). Continuing, the DUNFORD-PETTIS theorem, see, e.g., DUNFORD & SCHWARTZ [65, Corollary 11,p. 294], ensures the compactness of $(\widehat{U}'_{\tau_{n_k}})_{n\in\mathbb{N}}$ and $(\xi_{\tau_{n_k}})_{n\in\mathbb{N}}$ in $\mathrm{L}^1(0,T;V)$ and $\mathrm{L}^1(0,T;V^*)$, respectively, with respect to the weak topology, since both sequences are equi-integrable in the respective spaces. Hence, there exists a subsequence (labeled as before) and weak limits $v \in \mathrm{L}^1(0,T;V)$ and $\xi \in \mathrm{L}^1(0,T;V^*)$ such that $\widehat{U}'_{\tau_{n_k}} \rightharpoonup v$ weakly in $\mathrm{L}^1(0,T;V)$ and $\xi_{\tau_{n_k}} \rightharpoonup \xi$ weakly in $\mathrm{L}^1(0,T,V^*)$ as $k \to \infty$. By standard arguments, one can show that $u' = v$ in the weak sense. This yields $u \in \mathrm{W}^{1,1}(0,T;V)$, and therefore $u \in \mathrm{AC}([0,T];V)$. We continue with showing the convergence of the perturbation in (3.5.1e). We first note that by $\sup_{t\in[0,T]} \mathcal{E}_t(\overline{U}_{\tau_{n_k}})(t) \le M$ for all $k \in \mathbb{N}$, see (3.4.8), the functions $\overline{U}_{\tau_{n_k}}$ and therefore also $\underline{U}_{\tau_{n_k}}$ are contained in a compact subset $\mathcal{K} \subset D \subset V$ uniformly in $k \in \mathbb{N}$ and $t \in [0,T]$, since the energy functional has by Condition ((3.Eb)) compact sublevel sets. By TYCHONOFF'S theorem, the set $[0,T] \times \mathcal{K}$ is compact with respect to the product topology of $[0,T] \times V$. Then, Condition (3.Ba) yields the uniform continuity of the map $(t,u) \mapsto B(t,u)$ on $[0,T] \times \mathcal{K}$. Second, the convergences (3.4.11) and (3.5.1a) together imply $(\underline{\mathbf{t}}_{\tau_{n_k}}(t), \underline{U}_{\tau_{n_k}}(t))) \to (t,u(t))$ uniformly in $t \in (0,T)$ as $k \to \infty$. Finally, we obtain

$$\lim_{n\to\infty} \sup_{t\in(0,T)} \|B(\bar{\mathbf{t}}_{\tau_{n_k}}(t), \underline{U}_{\tau_{n_k}}(t)) - B(t,u(t))\|_* \qquad \text{as } n \to \infty.$$

In order to show the convergence in (3.5.1f), we note that due to (3.4.8), we have the uniform bound $(\partial_t\mathcal{E}_t(\widetilde{U}_{\tau_{n_k}})_{k\in\mathbb{N}} \subset \mathrm{L}^\infty(0,T)$. Thus, since $\mathrm{L}^\infty(0,T)$ is the dual space of the separable BANACH space $\mathrm{L}^1(0,T)$, there exists a weak* limit $\mathscr{P} \in \mathrm{L}^\infty(0,T)$ such that, up to a subsequence, there holds $\partial_t\mathcal{E}_t(\widetilde{U}_{\tau_{n_k}}) \overset{*}{\rightharpoonup} \mathscr{P}$ weakly* in $\mathrm{L}^\infty(0,T)$ as $k \to \infty$. Now, we aim to show that the weak limits can be expressed via a parameterized YOUNG measure. We define the product space $\mathcal{V} := V \times V^* \times \mathbb{R}$ endowed with the product topology, and for $k \in \mathbb{N}$, $w_k := (\widehat{U}'_{\tau_{n_k}}, \xi_{\tau_{n_k}}, \partial_t\mathcal{E}_t(\widetilde{U}_{\tau_{n_k}}))$. Then, since V is a reflexive BANACH space, the space $\mathcal{V}$ also becomes a reflexive BANACH space, and the fundamental theorem for weak topologies (Theorem 2.6.1) ensures, due to the equi-integrability of $(w_k)_{k\in\mathbb{N}}$ in $\mathrm{L}^1(0,T;\mathcal{V})$, the existence of a YOUNG measure $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in \mathscr{Y}(0,T;\mathcal{V})$ such that the (unique!) weak limit $(u',\xi,\mathscr{P}) \in \mathrm{L}^1(0,T;\mathcal{V})$ of the sequence $(w_k)_{k\in\mathbb{N}}$ satisfies

$$(u'(t),\xi(t),\mathscr{P}(t)) = \int_{V\times V^*\times\mathbb{R}} (v,\zeta,p)\,\mathrm{d}\mu_t(v,\zeta,p) \quad \text{for a.e. } t \in (0,T),$$

whence (3.5.2).

We proceed with showing (3.5.3). First, for notational convenience, we define for all $t \in [0,T]$

$$\eta_\tau(t) := \mathcal{E}_{\bar{\mathbf{t}}_\tau(t)}(\overline{U}_\tau(t)) - \int_0^{\bar{\mathbf{t}}_\tau(t)} \partial_r\mathcal{E}_r(\widetilde{U}_\tau(r))\,\mathrm{d}r - \int_0^{\bar{\mathbf{t}}_\tau(t)} \langle B(\bar{\mathbf{t}}_\tau(r), \underline{U}_\tau(r)), \widehat{U}'_\tau(r)\rangle\,\mathrm{d}r.$$

Second, considering the non-negativity of Ψ^ε_u and $\Psi^{\varepsilon,*}_u$, from the discrete energy-dissipation inequality (3.4.7), we deduce that the mapping $t \mapsto \eta_\tau(t) : [0,T] \to \mathbb{R}$ is non-increasing. Then, by HELLY's theorem, see, e.g, AMBROSIO et al. [10, Lemma 3.3.3, p. 70], there exists a non-increasing function $\eta : [0,T] \to \mathbb{R}$ such that (up to a subsequence denoted as before) $\eta_{\tau_{n_k}}(t) \to \eta(t)$ as $k \to \infty$ for any $t \in [0,T]$. Defining

$$\psi_\tau(t) := \int_0^{\bar{\mathbf{t}}_\tau(t)} \langle B(\bar{\mathbf{t}}_\tau(r), \underline{U}_\tau(r)), \widehat{U}'_\tau(r)\rangle \,\mathrm{d}r \quad \text{and}$$
$$\psi(t) := \int_0^t \langle B(r,u(r)), u'(r)\rangle \,\mathrm{d}r \quad \text{for all } t \in [0,T],$$

it is, in view of (3.5.1c) and (3.5.1e), easily seen that

$$\psi_{\tau_{n_k}}(t) \to \psi(t) \quad \text{as } k \to \infty \quad \text{for all } t \in [0,T].$$

The convergence of $\eta_{\tau_{n_k}}$ and $\psi_{\tau_{n_k}}$ together with (3.5.1f) yields the pointwise convergence of the energy functional, i.e.,

$$\mathcal{E}_{\bar{\mathbf{t}}_{\tau_{n_k}}(t)}(\overline{U}_{\tau_{n_k}}(t)) \to \mathscr{E}(t) := \eta(t) + \int_0^t \mathscr{P}(r)\,\mathrm{d}r + \psi(t) \quad \text{as } k \to \infty \quad \text{for all } t \in [0,T].$$

We observe that the real-valued function $\mathscr{E}$ is a sum of a monotone function η and absolutely continuous functions ψ and $\int_0^\cdot \mathscr{P}(r)\,\mathrm{d}r$ differentiable almost everywhere, see, e.g., ELSTRODT [70, Theorems 4.5, p. 299], and hence differentiable almost everywhere on $(0,T)$. Now, we conclude the convergence in (3.5.3) by noting that

$$|\mathcal{E}_{\bar{\mathbf{t}}_{\tau_{n_k}}(t)}(\overline{U}_{\tau_{n_k}}(t)) - \mathcal{E}_t(\overline{U}_{\tau_{n_k}}(t))| \to 0 \quad \text{as } k \to \infty \quad \text{for all } t \in [0,T]$$

which follows from LIPSCHITZ continuity of the time-dependence of the energy stated in (3.2.1) and the bound (3.4.8) as well as $\bar{\mathbf{t}}_{\tau_{n_k}}(t) \to t$ as $k \to \infty$ for all $t \in [0,T]$. Further, from the lower semicontinuity of the energy functional and the convergence (3.4.11), we obtain

$$\mathcal{E}_t(u(t)) \leq \liminf \mathcal{E}_t(\overline{U}_{\tau_{n_k}}(t)) = \mathscr{E}(t) \leq M \quad \text{for all } t \in [0,T],$$

whereas the last inequality is due to the bound (3.4.8). Moreover, by assumption, there also holds $\mathscr{E}(0) = \mathcal{E}_0(u_0)$. By Theorem 2.6.1, the YOUNG measure μ_t is for almost every $t \in (0,T)$ concentrated on the set

$$\mathrm{Li}(t) := \bigcap_{p=1}^{\infty} \mathrm{clos}_{\mathrm{weak}}\Big(\{w_k(t) : k \geq p\}\Big)$$

of all weak limit points of $w_k(t)$, meaning that

$$\mathrm{sppt}(\mu_t) := \mathrm{clos}\{A \in \mathscr{B}(\mathcal{V}) : \mu_t(A) > 0\} \subset \mathrm{Li}(t) \neq \emptyset \quad \text{for a.e. } t \in (0,T),$$

where $\mathscr{B}(\mathcal{V})$ denotes the BOREL σ-algebra of $\mathcal{V}$[4]. Let $\mathcal{N} \subset (0,T)$ be the negligible set such that the above-mentioned property holds on all $(0,T)\backslash\mathcal{N}$. Then, for a fixed $t \in (0,T)\backslash\mathcal{N}$ and for every $w = (v,\zeta,p) \in \mathrm{Li}(t)$, there holds (up to a subsequence)

[4]In fact, since $\mathcal{V} = V \times V^* \times \mathbb{R}$ is a separable metric space, the BOREL σ-algebra on $\mathcal{V}$ coincides with the product σ-algebra $\mathscr{B}(\mathcal{V}) = \mathscr{B}(V) \otimes \mathscr{B}(V^*) \otimes \mathscr{B}(\mathbb{R})$, see, e.g., AMANN & ESCHER [7, Theorem 1.15, p. 12].

$\widehat{U}'_{\tau_{n_k}}(t) \rightharpoonup v$, $\xi_{\tau_{n_k}}(t) \rightharpoonup \zeta$ and $\partial_t \mathcal{E}_t(\widetilde{U}_{\tau_{n_k}}(t)) \to p$ as $k \to \infty$, where in the latter convergence we used the fact that the weak topology and the strong topology coincide on finite-dimensional spaces. As a consequence of Condition (3.Ee), we have

$$\begin{aligned} &\zeta \in \partial \mathcal{E}_t(u(t)), \quad p \le \partial_t \mathcal{E}_t(u(t)) \quad \text{and} \quad \mathscr{E}(t) = \mathcal{E}_t(u(t) \\ &\text{for all } (v,\zeta,p) \in \mathrm{Li}(t) \quad \text{for all } t \in (0,T) \backslash \mathcal{N}. \end{aligned} \tag{3.5.5}$$

Then, in view of (3.5.2), we find after integration with respect to μ_t on $\mathcal{V}$ that

$$\begin{aligned} \mathscr{P}(t) &= \int_{V\times V^*\times\mathbb{R}} p \,\mathrm{d}\mu_t(v,\zeta,p) \le \partial_t \mathcal{E}_t(u(t)) \quad \text{for a.e. } t \in (0,T), \\ \xi(t) &= \int_{V\times V^*\times\mathbb{R}} \zeta \,\mathrm{d}\mu_t(v,\zeta,p) \in \partial \mathcal{E}_t(u(t)) \quad \text{for a.e. } t \in (0,T), \end{aligned}$$

where the last inclusion follows from the fact that the subdifferential $\partial\mathcal{E}_t(u(t))$ is closed and convex for almost every $t \in (0,T)$ and that μ_t is a probability measure for all $t \in [0,T]$. This implies (3.5.3). In order to show the remaining inequalities (3.5.4), let $s,t \in [0,T]$ be chosen fixed with $s < t$. We employ Theorem 2.6.1 by choosing $f, f_k : [0,T] \times \mathcal{V} \to \mathbb{R}$ given by

$$\begin{aligned} f_k(r,w) &= \Psi^{\varepsilon}_{\underline{U}_{\tau_{n_k}}(r)}(v) + \Psi^{\varepsilon,*}_{\underline{U}_{\tau_{n_k}}(r)}(\zeta) \\ f(r,w) &= \Psi^{\varepsilon}_{u(r)}(v) + \Psi^{\varepsilon,*}_{u(r)}(\zeta), \quad w = (v,\zeta,p) \in \mathcal{V}, r \in [s,t], \end{aligned}$$

and $f(r,w), f_k(r,w) = 0$ outside of $[s,t]$, where $M > 0$ is the constant from the boundedness in the a priori estimates. From Remark 3.2.1 *iii*) and the measurability of $\underline{U}_{\tau_{n_k}}$, $k \in \mathbb{N}$, and u, we deduce that the functionals f_n and f are weakly normal integrands for all $n \in \mathbb{N}$ which satisfy the Condition (2.6.1) of Theorem 2.6.1. Furthermore, by the a priori estimates (3.4.9) and (5.3.7), the sequence $(w_k)_{k\in\mathbb{N}}$ is equi-integrable so that all assumptions of Theorem 2.6.1 are satisfied. Consequently, there holds

$$\int_s^t \int_{\mathcal{V}} f(r,w)\,\mathrm{d}\mu_r(w)\,\mathrm{d}r \le \liminf_{k\to\infty} \int_s^t f(r,w_k(r))\,\mathrm{d}r. \tag{3.5.6}$$

On the other hand, we have by JENSEN's inequality

$$\Psi^{\varepsilon}_{u(r)}(u'(r)) \le \int_{V\times V^*\times\mathbb{R}} \Psi^{\varepsilon}_{u(r)}(v)\,\mathrm{d}\mu_r(v,\zeta,p), \tag{3.5.7}$$

$$\Psi^{\varepsilon,*}_{u(r)}(B(r,u(r)) - \xi(r))) \le \int_{V\times V^*\times\mathbb{R}} \Psi^{\varepsilon,*}_{u(r)}(B(r,u(r)) - \zeta)\,\mathrm{d}\mu_r(v,\zeta,p) \tag{3.5.8}$$

for almost every $r \in (s,t)$. Integrating the inequalities (3.5.7) and (3.5.8) over the interval (s,t) and taking into account (3.5.6) yields

$$\int_s^t f(r,w(r))\,\mathrm{d}r \le \liminf_{k\to\infty} \int_s^t f(r,w_k(r))\,\mathrm{d}r.$$

Thus, passing to the limit as $k \to \infty$ in the discrete energy-dissipation inequality

(3.4.7), we obtain

$$
\begin{aligned}
&\int_s^t \Big(\Psi^\varepsilon_{u(r)}(u'(r)) + \Psi^{\varepsilon,*}_{u(r)}(B(r,u(r)) - \xi(r))\Big)\,\mathrm{d}r + \mathcal{E}_t(u(t))\\
&\leq \int_s^t \int_{V\times V^*\times\mathbb{R}} \Big(\Psi^\varepsilon_{u(r)}(v) + \Psi^{\varepsilon,*}_{u(r)}(B(r,u(r)) - \zeta)\Big)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r + \mathcal{E}_t(u(t))\\
&\leq \int_s^t \int_{V\times V^*\times\mathbb{R}} \Big(\Psi^\varepsilon_{u(r)}(v) + \Psi^{\varepsilon,*}_{u(r)}(B(r,u(r)) - \zeta)\Big)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r + \mathscr{E}(t)\\
&\leq \liminf_{k\to\infty}\left(\int_s^t \int_{V\times V^*\times\mathbb{R}} f(r,w_k)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r + \mathcal{E}_{\bar{\mathbf{t}}_{\tau_{n_k}}(t)}(\overline{U}_\tau(t))\right)\\
&\leq \liminf_{k\to\infty}\left(\int_{\bar{\mathbf{t}}_{\tau_{n_k}}(s)}^{\bar{\mathbf{t}}_{\tau_{n_k}}(t)} \int_{V\times V^*\times\mathbb{R}} f(r,w_k)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r + \mathcal{E}_{\bar{\mathbf{t}}_{\tau_{n_k}}(t)}(\overline{U}_\tau(t))\right)\\
&\leq \liminf_{k\to\infty}\left(\mathcal{E}_{\bar{\mathbf{t}}_\tau(s)}(\overline{U}_\tau(s)) + \int_{\bar{\mathbf{t}}_\tau(s)}^{\bar{\mathbf{t}}_\tau(t)} \Big(\partial_r\mathcal{E}_r(\widetilde{U}_\tau(r)) + \langle B(\bar{\mathbf{t}}_\tau(r), \underline{U}_\tau(r)), \widehat{U}'_\tau(r)\rangle\Big)\,\mathrm{d}r\right)\\
&= \mathscr{E}(s) + \int_s^t \mathscr{P}(r)\,\mathrm{d}r + \int_s^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r\\
&\leq \mathscr{E}(s) + \int_s^t \partial_r\mathcal{E}_r(u(r))\,\mathrm{d}r + \int_s^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r
\end{aligned}
\tag{3.5.9}
$$

for all $0 \leq s < t \leq T$, which proves the assertion of this lemma. □

Although we know that

$$
\xi_\tau(t) = B(\bar{\mathbf{t}}_\tau(t), \underline{U}_\tau(t)) - D_G\Psi^\varepsilon_{\underline{U}_\tau(t)}\left(\frac{\widetilde{U}_\tau(t) - \underline{U}_\tau(t)}{t - \underline{\mathbf{t}}_\tau(t)}\right), \quad \text{for all } t \in [0,T],
$$

the strong convergence of the perturbation B and the demicontinuity of $\mathrm{D}_G\Psi^\varepsilon_{\underline{U}_\tau}$ are not sufficient to conclude $\xi(t) = B(t,u(t)) - \mathrm{D}_G\Psi^\varepsilon_{u(t)}(u'(t))$. However, characterizing the weak limits as parameterized Young measures, we can make this conclusion as we will see in the following proof of the main result.

3.6 Proof of Theorem 3.2.3

In order to show that the curve $u \in \mathrm{AC}([0,T];V)$ obtained from Lemma 3.5.1 is a solution to the differential inclusion (3.0.1), we employ the chain rule condition (3.Ed), which is justified by (3.5.1f), (3.5.2a), (3.5.5), (3.5.7) and (3.5.8), where $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in \mathscr{Y}(0,T;V\times V^*\times\mathbb{R})$ is to be chosen as in Lemma 3.5.1. Hence, by the chain rule condition (3.Ed), the mapping $t \mapsto \mathcal{E}_t(u(t))$ is absolutely continuous on $(0,T)$ and there holds

$$
\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) \geq \langle \xi(t), u'(t)\rangle - \partial_t\mathcal{E}_t(u(t)) \quad \text{for a.e. } t \in (0,T),
$$

where we have used the characterization (3.5.2b) and (3.5.3). Thus, together with (3.5.3), (3.5.2c) and (3.5.9), there holds for $s = 0$

$$\begin{aligned}
&\int_0^t \Big(\Psi^{\varepsilon}_{u(r)}(u'(r)) + \Psi^{\varepsilon,*}_{u(r)}(B(r,u(r)) - \xi(r))\Big)\,\mathrm{d}r + \mathcal{E}_t(u(t)) \\
&\leq \mathcal{E}_0(u_0) + \int_0^t \partial_r \mathcal{E}_r(u(r))\,\mathrm{d}r + \int_0^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r \\
&\leq \mathcal{E}_t(u(t)) - \int_0^t \langle \xi(r), u'(r)\rangle\,\mathrm{d}r + \int_0^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r \\
&= \mathcal{E}_t(u(t)) + \int_0^t \langle B(r,u(r)) - \xi(r), u'(r)\rangle\,\mathrm{d}r \quad \text{for all } t \in (0,T).
\end{aligned}$$

Therefore, we obtain

$$\int_0^t \Psi^{\varepsilon}_{u(r)}(u'(r)) + \Psi^{\varepsilon,*}_{u(r)}(B(r,u(r)) - \xi(r)) - \langle B(r,u(r)) - \xi(r), u'(r)\rangle)\,\mathrm{d}r \leq 0 \tag{3.6.1}$$

for all $t \in (0,T)$. Then, from the FENCHEL-YOUNG inequality we deduce the non-negativity of the integrand in (3.6.1) and infer

$$\Psi^{\varepsilon}_{u(t)}(u'(t)) + \Psi^{\varepsilon,*}_{u(t)}(B(t,u(t)) - \xi(t)) - \langle B(t,u(t)) - \xi(t), u'(t)\rangle = 0 \quad \text{for a.e. } t \in (0,T).$$

By Lemma 2.3.1, this implies in fact that

$$\xi(t) = B(t,u(t)) - \mathrm{D}_G \Psi^{\varepsilon}_{u(t)}(u'(t)) \quad \text{for a.e. } t \in (0,T).$$

Furthermore, by Lemma 3.5.1, there holds

$$\xi(t) = \int_{V\times V^*\times\mathbb{R}} \zeta \,\mathrm{d}\mu_t(v,\zeta,p) \in \partial\mathcal{E}_t(u(t)) \quad \text{for a.e. } t \in (0,T),$$

which shows that the couple (u,ξ) is a solution of the regularized perturbed gradient system $(V,\mathcal{E},\Psi^{\varepsilon},B)$ and in particular fulfills the energy-dissipation balance (3.2.7). For each $\varepsilon \in (0,1]$, we denote with $(u_\varepsilon,\xi_\varepsilon)$ the couple of solutions of $(V,\mathcal{E},\Psi^{\varepsilon},B)$. Now, we want to pass to the limit with $\varepsilon \searrow 0$ and want to show that the couple $(u_\varepsilon,\xi_\varepsilon)$ converge to a solution to the limiting system $(V,\mathcal{E},\Psi^0,B) = (V,\mathcal{E},\Psi,B)$. The steps are essentially the same as before:

1. We derive a priori estimates based on the energy-dissipation balance (3.2.7),

2. We show compactness of the solutions u_ε and the pointwise subgradients $\xi_\varepsilon = B(\cdot,u_\varepsilon) - \mathrm{D}_G\Psi^{\varepsilon}(u_\varepsilon')$ of $\mathcal{E}(u_\varepsilon)$ in appropriate spaces,

3. With the aid of YOUNG measures, we pass to the limit as $\varepsilon \searrow 0$.

Therefore, we do not give all of the details of the proof and refer to the full proof of the aforementioned lemmas. Instead, we highlight the difference from the previous steps which mostly relies on Lemma 3.2.4 and the continuity of the dissipation

potential in the sense of Mosco-convergence.
Ad 1. Starting from the energy-dissipation balance

$$\begin{aligned}\mathcal{E}_t(u_\varepsilon(t)) &+ \int_0^t \Big(\Psi^\varepsilon_{u_\varepsilon(r)}(u_\varepsilon'(r)) + \Psi^{\varepsilon,*}_{u_\varepsilon(r)}\Big(B(r,u_\varepsilon(r)) - \xi_\varepsilon(r)\Big)\Big)\,\mathrm{d}r \\ &= \mathcal{E}_0(u_\varepsilon(0)) + \int_0^t \Big(\partial_r \mathcal{E}_t(u_\varepsilon(r)) + \Big\langle B(r,u_\varepsilon(r)), u_\varepsilon'(t)\Big\rangle\Big)\,\mathrm{d}r\end{aligned}$$

and proceeding in the exact same way as before, we obtain with the Gronwall lemma (Lemma A.1.1) a constant $M = M(\mathcal{E}_0(u_0), T) > 0$ such that the bounds

$$\begin{aligned}&\sup_{t\in[0,T]} \mathcal{E}^\varepsilon_t(u_\varepsilon(t)) \le M,\\ &\sup_{t\in[0,T]} \|B^\varepsilon(t,u_\varepsilon(t))\|_* \le M, \qquad \sup_{t\in(0,T)} |\partial_t \mathcal{E}^\varepsilon_t(u_\varepsilon(t))| \le M,\\ &\int_0^T \Big(\Psi^\varepsilon_{u_\varepsilon(r)}(u_\varepsilon(r))) + \Psi^{\varepsilon,*}_{u_\varepsilon(r)}(B^\varepsilon(r,u_\varepsilon(r)) - \xi_\varepsilon(r))\Big)\,\mathrm{d}r \le M\end{aligned}$$

hold for all $0 \le \varepsilon \le 1$. Besides, $(u'_\varepsilon)_{0\le\varepsilon\le1} \subset \mathrm{L}^1(0,T;V)$ and $(\xi_\varepsilon)_{0\le\varepsilon\le1} \subset \mathrm{L}^1(0,T;V^*)$ are equi-integrable with respect to ε in $\mathrm{L}^1(0,T;V)$ and $\mathrm{L}^1(0,T;V^*)$, respectively. The equi-integrability follows from the fact that the dissipation potential and its conjugate are superlinear uniformly in $\varepsilon > 0$ and on sublevels of the energy, which follows from Lemma 3.2.4, and the criterion of de la Vallée-Poussin for equi-integrability.
Ad 2. For every vanishing sequence $\varepsilon_k \to 0$, we find in the same manner as Lemma 3.5.1, there exists a subsequence (labeled as before), an absolutely continuous curve $u \in \mathrm{AC}([0,T];V)$ with $u(0) = u_0$, an integrable function $\xi \in \mathrm{L}^1(0,T;V^*)$, a function $\mathscr{E}^0 : [0,T] \to \mathbb{R}$ of bounded variation, an essentially bounded function $\mathscr{P}^0 \in \mathrm{L}^\infty(0,T^*)$, and a parameterized Young measure $\boldsymbol{\nu} = (\nu_t)_{t\in[0,T]} \in \mathscr{Y}(0,T;V\times V^*\times\mathbb{R})$ such that

$$\begin{aligned} u_{\varepsilon_k} &\to u && \text{in } \mathrm{C}([0,T];V),\\ u'_{\varepsilon_k} &\rightharpoonup u' && \text{in } \mathrm{L}^1(0,T;V),\\ \xi_{\varepsilon_k} &\rightharpoonup \xi && \text{in } \mathrm{L}^1(0,T;V^*),\\ B(\cdot,u_{\varepsilon_k}(\cdot)) &\to B^0(\cdot,u(\cdot)) && \text{in } \mathrm{L}^\infty(0,T;V^*),\\ \partial_t\mathcal{E}_t(u_{\varepsilon_k}) &\overset{*}{\rightharpoonup} \mathscr{P}^0 && \text{in } \mathrm{L}^\infty(0,T),\end{aligned}$$

as $k \to \infty$, and the weak limits satisfy

$$\begin{aligned} u'(t) &= \int_{V\times V^*\times\mathbb{R}} v\,\mathrm{d}\nu_t(v,\zeta,p) && \text{for a.e. } t\in(0,T),\\ \xi(t) &= \int_{V\times V^*\times\mathbb{R}} \zeta\,\mathrm{d}\nu_t(v,\zeta,p) && \text{for a.e. } t\in(0,T),\\ \mathscr{P}^0(t) &= \int_{V\times V^*\times\mathbb{R}} p\,\mathrm{d}\nu_t(v,\zeta,p) && \text{for a.e. } t\in(0,T).\end{aligned}$$

Furthermore, there holds

$$\begin{cases} \mathcal{E}_t(u_{\varepsilon_k}(t)) \to \mathscr{E}^0(t) \quad \text{as } k\to\infty & \text{for all } t\in[0,T],\\ \mathcal{E}_t(u(t)) \le \mathscr{E}^0(t) & \text{for all } t\in[0,T],\\ \mathcal{E}_0(u_0) = \mathscr{E}^0(0) \quad\text{and}\quad \mathcal{E}_t(u(t)) = \mathscr{E}^0(t) & \text{for a.e. } t\in(0,T),\\ \mathscr{P}^0(t) \le \partial_t\mathcal{E}_t(u(t)) & \text{for a.e. } t\in(0,T).\end{cases}$$

and the inequality

$$\begin{aligned}
&\int_s^t \Big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}(B(r,u(r)) - \xi(r))\Big)\,\mathrm{d}r + \mathcal{E}_t(u(t)) \\
&\le \int_s^t \int_{V\times V^*\times\mathbb{R}} \Big(\Psi_{u(r)}(v) + \Psi^*_{u(r)}(B(r,u(r)) - \zeta)\Big)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r + \mathcal{E}_t(u(t)) \\
&\le \mathscr{E}^0(s) + \int_s^t \partial_t\mathcal{E}_t(u(t))\,\mathrm{d}r + \int_s^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r
\end{aligned} \tag{3.6.4}$$

holds for all $0 \le s < t \le T$. Here, in order to establish the inequality (3.6.4), we use the Mosco-convergence $\Psi^{\varepsilon_k}_{u_{\varepsilon_k}} \xrightarrow{\mathrm{M}} \Psi_u$ as $k \to \infty$ and Theorem 2.6.1 by choosing $f, f_k : [0,T^*] \times \mathcal{V} \to \mathbb{R}$ by

$$\begin{aligned}
f_k(r,w) &= \Psi^{\varepsilon_k}_{u_{\varepsilon_k}(r)}(v) + \Psi^{\varepsilon_k,*}_{u_{\varepsilon_k}(r)}(\zeta) \\
f(r,w) &= \Psi_{u(r)}(v) + \Psi^*_{u(r)}(\zeta), \quad w = (v,\zeta,p) \in \mathcal{V}, r \in [s,t],
\end{aligned}$$

and $f(r,w), f_k(r,w) = 0$ outside of $[s,t]$.
Ad 3. This part of the proof follows the same steps as the part where we show that u_{ε_k} is a solution to the regularized perturbed gradient system $(V, \mathcal{E}, \Psi^\varepsilon, B)$. □

Remark 3.6.1 Along the same lines as the proof of Theorem 4.4 in Mielke et al. [122], it can be proven that (up to a subsequence) the following convergences hold:

$$\begin{aligned}
\mathcal{E}_t(u_{\varepsilon_k}(t)) &\to \mathcal{E}_t(u(t)), \\
\int_r^s \Psi^{\varepsilon_k}_{u_{\varepsilon_k}(t)}(u'_{\varepsilon_k}(t))\,\mathrm{d}t &\to \int_r^s \Psi_{u(t)}(u'(t))\,\mathrm{d}t, \\
\int_r^s \Psi^*_{u_{\varepsilon_k}(t)}\Big(B(t,u_{\varepsilon_k}(t)) - \xi_{\varepsilon_k}(t)\Big)\,\mathrm{d}t &\to \int_r^s \Psi^*_{u(t)}\Big(B(t,u(t)) - \xi(t)\Big)\,\mathrm{d}t
\end{aligned}$$

as $k \to \infty$ for all $0 \le s < t \le T$. Furthermore, if we additionally assume that the dissipation potential Ψ_u and its conjugate Ψ^*_u are strictly convex for all $u \in V$, then we obtain the pointwise weak convergences

$$u'_{\varepsilon_k}(t) \rightharpoonup u'(t) \quad \text{and} \quad \xi_{\varepsilon_{n_k}}(t) \rightharpoonup \xi(t) \quad \text{for a.e. } t \in (0,T).$$

In fact, it is feasible to show a more general existence result based on the so-called evolutionary Γ-convergence where one shows that solutions to a perturbed gradient system

$$B^\varepsilon(t,u(t)) \in \partial\Psi^\varepsilon_{u(t)}(u'(t)) + \partial\mathcal{E}^\varepsilon_t(u(t))$$

which depends on a parameter ε, converge to a solution of the limiting (effective) system $(V, \mathcal{E}^0, \Psi^0, B^0)$ under the assumption that $\mathcal{E}^\varepsilon_t \to \mathcal{E}_t$ in the sense of Γ-convergence, $\Psi^\varepsilon_u \to \Psi^0_u$ in the sense of Mosco-convergence,$B^\varepsilon \to B$ uniformly on $[0,T] \times V$, see [21, 122].

Chapter 4

Application

In this chapter, we provide a nontrivial example of our abstract existence result formulated in Theorem 3.2.3, which was developed and proven in Chapter 3. Before we start with the example, we want to fix the notation.

In the following example, let $d, m \in \mathbb{N}$ and $\Omega \subset \mathbb{R}^d$ be a bounded domain with a LIPSCHITZ boundary $\partial\Omega$ with the outward-pointing unit normal vector $\boldsymbol{\nu}$ on the boundary, $T > 0$ and $\Omega_T := \Omega \times (0, T)$. We denote the multi-dimensional vectors and matrices with bold letters and the one-dimensional objects with small letters. For two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ and two matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{d,m}$, the EUCLIDian and the FROBENIUS scalar product are given by

$$\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{d} x_i y_i \quad \text{and} \quad \boldsymbol{A} : \boldsymbol{B} = \sum_{i,j=1}^{d,m} A_{i,j} B_{i,j}, \quad \text{respectively.}$$

The norms on $\mathbb{R}^d$ and $\mathbb{R}^{d,m}$ induced by the EUCLIDian and the FROBENIUS scalar product, respectively, are both denoted by $|\cdot|$. Furthermore, for a real valued function $h : \Omega \to \mathbb{R}$ and a vector valued function $\boldsymbol{h} : \Omega \to \mathbb{R}^m, \boldsymbol{x} \mapsto \boldsymbol{h}(\boldsymbol{x}) := (h_1(\boldsymbol{x}), \ldots, h_m(\boldsymbol{x}))$, the nabla operator ∇ is defined as

$$\nabla h(\boldsymbol{x}) = \left(\frac{\partial h}{\partial x_i}(\boldsymbol{x}) \right)_{i=1}^{d} \quad \text{and} \quad \nabla \boldsymbol{h}(\boldsymbol{x}) = \left(\frac{\partial h_i}{\partial x_j}(\boldsymbol{x}) \right)_{i,j=1}^{m,d}, \quad \boldsymbol{x} \in \Omega.$$

For a vector valued function $\boldsymbol{g} : \Omega \to \mathbb{R}^d, \boldsymbol{x} \mapsto \boldsymbol{g}(\boldsymbol{x}) := (g_1(\boldsymbol{x}), \ldots, g_d(\boldsymbol{x}))$ and a matrix valued function $\boldsymbol{A} : \Omega \to \mathbb{R}^{m \times d}, \boldsymbol{x} \mapsto \boldsymbol{A}(\boldsymbol{x}) := (A_{ij}(\boldsymbol{x}))_{i,j=1}^{m,d}$, the divergence is defined as

$$\nabla \cdot \boldsymbol{g}(\boldsymbol{x}) = \operatorname{div}(\boldsymbol{g}(\boldsymbol{x})) = \sum_{i=1}^{d} \frac{\partial g_i}{\partial x_i}(\boldsymbol{x}) \quad \text{and} \quad \nabla \cdot \boldsymbol{A}(\boldsymbol{x}) = \sum_{i,j=1}^{m,d} \frac{\partial A_{i,j}}{\partial x_j}(\boldsymbol{x}) \boldsymbol{e}_j,$$

where $\boldsymbol{e}_j \in \mathbb{R}^d$ is the j-th standard unit vector. Finally, the LAPLACE operator is defined by $\Delta = \nabla \cdot \nabla = \nabla^2$. Higher order LAPLACIAN's are also denoted by Δ^k and we denote $\nabla^k = \Delta^{k/2}$ if $k \in 2\mathbb{N}$ or $\nabla^k = \nabla \Delta^{(k-1)/2}$ otherwise. For $p \geq 1$, the p-LAPLACE of the vector valued function $\boldsymbol{h} : \Omega \to \mathbb{R}^m$ is defined by $\Delta_p \boldsymbol{h}(\boldsymbol{x}) = \nabla \cdot \left(|\nabla \boldsymbol{h}(\boldsymbol{x})|^{p-2} \nabla \boldsymbol{h}(\boldsymbol{x}) \right), \boldsymbol{x} \in \Omega$. For notational convenience, we use the short-hand notations $\partial_t = \frac{\partial}{\partial t}$ and $\partial_{tt} = \frac{\partial^2}{\partial t^2}$ for the first and second time derivatives, respectively. For the LEBESGUE and SOBOLEV spaces[1], we use the usual notation $\mathrm{L}^p(\Omega)^m$ and $\mathrm{W}^{k,p}(\Omega)^m$ for $p \in [1, +\infty]$ and $k \in \mathbb{N}$ equipped with the standard norms, respectively. The space of functions in $\mathrm{W}^{k,p}(\Omega)^m$ with zero trace is denoted by $\mathrm{W}_0^{k,p}(\Omega)^m$. For $p = 2$, we use the notation $\mathrm{H}^k(\Omega)^m = \mathrm{W}^{k,p}(\Omega)^m$ and $\mathrm{H}_0^k(\Omega)^m = \mathrm{W}_0^{k,p}(\Omega)^m$. Furthermore, we will not distinguish between the abstract function $\tilde{u}$ and the concrete function u, which are related to each other via $[\tilde{u}(t)](\boldsymbol{x}) = u(\boldsymbol{x}, t)$. Finally, $C > 0$ denotes a generic constant.

We consider an initial-boundary value problem supplemented with nonlinear constraints which has, in a modified version and without perturbation, been studied in MIELKE et al. [122]. The governing equations are given

$$
\text{(P1)} \begin{cases} \mathrm{D}_{\boldsymbol{v}}\psi(\boldsymbol{x}, \boldsymbol{u}, \partial_t \boldsymbol{u}) + \boldsymbol{p} - \Delta_p \boldsymbol{u} + \mathrm{D}W(\boldsymbol{u}) + \partial\imath_K(\boldsymbol{u}) + \boldsymbol{b}(\boldsymbol{x}, t, \boldsymbol{u}) \ni \boldsymbol{f} \quad \text{in } \Omega_T, \\ \boldsymbol{p}(\boldsymbol{x}, t) \in \mathbf{Sgn}\,(\partial_t \boldsymbol{u}(\boldsymbol{x}, t)) \quad \text{a.e. in } \Omega_T, \\ \boldsymbol{u}(\boldsymbol{x}, t) \in K \quad \text{a.e. in } \Omega_T, \\ \boldsymbol{u}(\boldsymbol{x}, 0) \;= \boldsymbol{u}_0(\boldsymbol{x}) \quad \text{on } \Omega, \\ \boldsymbol{u}(\boldsymbol{x}, t) \;= 0 \qquad\quad \text{on } \partial\Omega \times [0, T], \end{cases}
$$

where $p \geq 2$, $K \subset \mathbb{R}^m$ is a compact and convex set, $\mathbf{Sgn} : \mathbb{R}^{d\times m} \rightrightarrows \mathbb{R}^{d\times m}$

$$
\mathbf{Sgn}(\boldsymbol{A}) = \begin{cases} B_{\mathbb{R}^{d\times m}}(0,1) & \text{if } \boldsymbol{A} = 0 \\ \frac{\boldsymbol{A}}{|\boldsymbol{A}|} & \text{otherwise,} \end{cases} \tag{4.0.1}
$$

is the multi-valued and multi-dimensional sign function, and $\imath_K \to \{0, +\infty\}$ denotes the indicator function on K defined by

$$
\imath_K(\boldsymbol{A}) = \begin{cases} 0 & \text{if } \boldsymbol{A} \in K \\ +\infty & \text{otherweise.} \end{cases}
$$

We could have also imposed other types of boundary conditions as non-homogeneous DIRICHLET, NEUMANN, or mixed boundary conditions which can be incorporated into the energy functional or into the space, see, e.g., [69, 122, 143, 144], where these cases have been considered.

Furthermore, we impose the following conditions on ψ, W, $\boldsymbol{b}$ and $\boldsymbol{f}$. We start with the assumptions on ψ.

(4.0.a) The function $\psi : \Omega \times \mathbb{R}^m \times \mathbb{R}^m \to [0, +\infty)$ is a CARATHÉODORY function such that $\psi(\boldsymbol{x}, \boldsymbol{y}, \cdot)$ is a proper, convex, GÂTEAUX differentiable functional with derivative $\mathrm{D}_{\boldsymbol{z}}\psi$ with respect to the third variable, and $\psi(\boldsymbol{x}, \boldsymbol{y}, 0) = 0$ for almost every $\boldsymbol{x} \in \Omega$ and all $\boldsymbol{y} \in K$.

(4.0.b) The functional ψ satisfies the following growth condition: there exists a number $q > 1$ and positive constants $c_\psi, C_\psi > 0$ such that

$$
c_\psi\left(|\boldsymbol{z}|^q - 1\right) \leq \psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \leq C_\psi\left(1 + |\boldsymbol{z}|^q\right) \tag{4.0.2}
$$

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

[1]See BRÉZIS [35, Chapter 4 & 9] or DOBROWOLSKI [62, Kapitel 4 & 5] for a definition and a detailed discussion of the LEBESGUE and SOBOLEV spcaes.

(4.0.c) The function $W \in \mathrm{C}^1(\mathbb{R}^m;\mathbb{R})$ is λ-convex and bounded from below.

(4.0.d) The function $\boldsymbol{b} : \Omega \times [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ is a CARATHÉODORY function in the sense that $\boldsymbol{b}(\boldsymbol{x},\cdot,\cdot)$ is continuous for almost every $\boldsymbol{x} \in \Omega$ and that $\boldsymbol{b}(\cdot,t,\boldsymbol{y})$ is measurable for all $t \in [0,T]$ and $\boldsymbol{y} \in \mathbb{R}^m$.

(4.0.e) There exists a function $h \in \mathrm{L}^{p^*}(\Omega)$ and a constant $C_{\boldsymbol{b}} > 0$ such that

$$|\boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{y})| \le h(\boldsymbol{x}) + C_{\boldsymbol{b}} \quad \text{for a.e. } \boldsymbol{x} \in \Omega, \text{ and all } t \in [0,T], \boldsymbol{y} \in K. \tag{4.0.3}$$

(4.0.f) There holds $\boldsymbol{f} \in \mathrm{C}^1([0,T];\mathrm{W}^{-1,p^*}(\Omega)^d)$.

Here, we assume for simplicity the (GÂTEAUX) differentiability of W and the λ-convexity of W. More general nonsmooth functions in the form $W = W_1 - W_2$ with W_1 being convex and W_2 being convex or continuously differentiable where both functionals satisfying certain growth conditions, see, e.g., [122, 142, 143, 148]. For the external force, we could in fact assume $\boldsymbol{f} \in \mathrm{C}^1([0,T];\mathrm{W}^{-1,p'}(\Omega)^d) + \mathrm{C}([0,T];\mathrm{L}^{p^*}(\Omega)^m)$ by treating the part from $\mathrm{C}([0,T];\mathrm{L}^{p^*}(\Omega)^m)$ as perturbation.

Simple examples for ψ and $\boldsymbol{b}$ might be $\psi(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = \psi(\boldsymbol{z}) = \frac{1}{q}|\boldsymbol{z}|^q, \boldsymbol{z} \in \mathbb{R}^m$, $\boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{y}) = \boldsymbol{b}(\boldsymbol{y}) = g(|\boldsymbol{y}|), \boldsymbol{y} \in \mathbb{R}^m$ for any continuous function $g \in \mathrm{C}(\mathbb{R})$. Admissible choices for W are the double-well potential $W(\boldsymbol{z}) = \frac{1}{4}(|\boldsymbol{z}|^2-1)^2 = \frac{1}{4}\left(|\boldsymbol{z}|^4+1\right) - \frac{1}{2}|\boldsymbol{z}|^2$, or in a more general setting, the logarithmic potential

$$W(z) = \begin{cases} (z-z_1)\ln(z-z_1) + (z_2-z)\ln(z_2-z) - \frac{\lambda}{2}z^2 & \text{if } z_1 < z < z_2 \\ +\infty, \quad \text{otherwise} \end{cases}$$

if $m = 1$, where $-\infty < z_1 < z_2 < +\infty$ are real numbers. It is easy to verify that both functions are λ-convex.

Accordingly, we have $V = \mathrm{L}^q(\Omega)^m$. Then, the energy functional $\mathcal{E}_t : V \to (-\infty,+\infty]$ is given by

$$\mathcal{E}_t(u) = \begin{cases} \int_\Omega \left(\frac{1}{p}|\nabla \boldsymbol{u}(\boldsymbol{x})|^p + W(\boldsymbol{u}(\boldsymbol{x})) + \imath_K(\boldsymbol{u}(\boldsymbol{x})) - \langle \boldsymbol{f}(t), \boldsymbol{u}\rangle_{\mathrm{W}_0^{1,p}}\right)\mathrm{d}\boldsymbol{x} & \text{if } \boldsymbol{u} \in D, \\ +\infty \quad \text{otherwise,} \end{cases}$$

where $D := \mathrm{dom}(\mathcal{E}_t)$ and $\langle \cdot,\cdot\rangle_{\mathrm{W}_0^{1,p}}$ is a shorthand notation for the duality pairing between $\mathrm{W}_0^{1,p}(\Omega)^m$ and $\mathrm{W}^{-1,p^*}(\Omega)^m$. In order for the energy functional to be finite, it must be true that $\boldsymbol{u} \in K$ a.e. in Ω, which implies $\boldsymbol{u} \in \mathrm{L}^\infty(\Omega)^d$. Therefore, by the continuity of W, there holds $\int_\Omega W(\boldsymbol{u}(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} < +\infty$ for all $\boldsymbol{u} \in D$. Hence, we have the characterization $D = \{u \in \mathrm{W}_0^{1,p}(\Omega)^d \cap \mathrm{L}^\infty(\Omega)^d : \boldsymbol{u}(\boldsymbol{x}) \in K \text{ a.e. in } \Omega\}$. The dissipation potential $\Psi : V \to \mathbb{R}$ is given by

$$\Psi_{\boldsymbol{u}}(\boldsymbol{v}) = \int_\Omega \left(\psi(\boldsymbol{x},\boldsymbol{u}(\boldsymbol{x}),\boldsymbol{v}(\boldsymbol{x})) + |\boldsymbol{v}(\boldsymbol{x})|\right)\mathrm{d}\boldsymbol{x}, \quad \boldsymbol{v} \in V, \boldsymbol{u} \in D,$$

and the perturbation $B : D \to V^*$ by

$$\langle B(t,\boldsymbol{u}), \boldsymbol{w}\rangle_{V^* \times V} = \int_\Omega -\boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{u}(\boldsymbol{x})) \cdot \boldsymbol{w}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \quad \boldsymbol{u} \in D, \boldsymbol{w} \in V.$$

From the Assumptions (4.0.b) and (4.0.d), the functional $\Psi_{\boldsymbol{u}}$ and the operator B are well-defined. The conjugate $\Psi_{\boldsymbol{u}}^* : V^* \to \mathbb{R}$ is with Lemma 2.3.5 and Ekeland & Temam [69, Proposition 1.2, p. 78] given by the formula

$$\Psi_{\boldsymbol{u}}^*(\boldsymbol{\xi}) = \int_\Omega \min_{\boldsymbol{\eta} \in \overline{B}(0,1)} \left(\psi^*(\boldsymbol{x}, \boldsymbol{u}(\boldsymbol{x}), \boldsymbol{\eta} - \boldsymbol{\xi}(\boldsymbol{x}))\right) \mathrm{d}\boldsymbol{x}, \quad \boldsymbol{\xi} \in V^*, \boldsymbol{u} \in D,$$

where we have used the fact that $(|\cdot|)^* = \imath_{\overline{B}(0,1)}$. Once again, we want to prove that under (4.0.a)-(4.0.f) the Conditions (4.Ψ), (4.$\mathcal{E}$) and (4.B) are fulfilled.

We start with the dissipation potential and observe that it is readily seen by the assumptions that $\Psi_{\boldsymbol{u}}$ is a lower semicontinuous and convex functional with $\Psi_{\boldsymbol{u}}(0)$, which in turn implies these properties for $\Psi_{\boldsymbol{u}}^*$ for all $\boldsymbol{u} \in D$ as we pointed out in Remark 3.2.1 *i)* and thus (3.Ψa). By Assumption (4.0.2), for all $R > 0$, there exist constants $\tilde{c}_\psi^R, \tilde{C}_\psi^R > 0$ such that

$$\tilde{c}_\psi(\|\boldsymbol{v}\|_V^q - 1) \leq \Psi_{\boldsymbol{u}}(\boldsymbol{v}) \leq \tilde{C}_\psi(\|\boldsymbol{v}\|_V^q + 1) \quad \text{for all } \boldsymbol{v}, \in V, \boldsymbol{u} \in D, \mathcal{G}(\boldsymbol{u}) \leq R,$$

where $\mathcal{G} = \sup_{t\in[0,T]} \mathcal{E}_t$. Thus, we obtain for the conjugate

$$c_*(\|\boldsymbol{v}\|_{V^*}^{q*} - 1) \leq \Psi_{\boldsymbol{u}}^*(\boldsymbol{v}) \leq C_*(\|\boldsymbol{v}\|_{V^*}^{q*} + 1) \quad \text{for all } \boldsymbol{v}, \in V, \boldsymbol{u} \in D$$

for constants $c_*, C_* > 0$, where $q* = q/(q-1) > 1$ is the conjugate exponent to q. Thus, Condition (3.Ψb) is fulfilled. The sequential lower semicontinuity of the integrals $\Psi_{\boldsymbol{u}}$ and $\Psi_{\boldsymbol{u}}^*$ follows from the assumptions on ψ, the compact embedding (4.0.4), and Ioffe [95, Theorem 3], which implies (3.Ψc). The subdifferential of $\Psi_{\boldsymbol{u}}$ is according to Lemma 2.3.1 characterized by

$$\boldsymbol{\eta} \in \partial\Psi_{\boldsymbol{u}}(\boldsymbol{v}) \quad \text{iff} \quad \boldsymbol{\eta}(\boldsymbol{x}) \in \mathrm{D}_{\boldsymbol{v}}\psi(\boldsymbol{x}, \boldsymbol{u}(\boldsymbol{x}), \boldsymbol{v}(\boldsymbol{x})) + \mathbf{Sgn}(\boldsymbol{v}(\boldsymbol{x})) \quad \text{for a.e. } \boldsymbol{x} \in \Omega$$

for all $\boldsymbol{u} \in D, \boldsymbol{v} \in V$.

We continue with showing the conditions for the energy functional. In order to show the sequential lower semicontinuity of $\mathcal{E}_t$, we show that all sublevel sets $J_a := \{v \in V : \mathcal{E}_t(v) \leq a\}$ are closed in V for all $a \in \mathbb{R}$ and $t \in [0,T]$. So, let $t \in [0,T]$, $a \in \mathbb{R}$ and $u_n \to u$ in V as $n \to \infty$ be a strongly converging sequence in V with $u_n \in J_a$ for all $n \in \mathbb{N}$. Then, obviously $u_n \in D$ for all $n \in \mathbb{N}$ and the sequence $(u_n)_{n\in\mathbb{N}}$ is bounded in $\mathrm{W}_0^{1,p}(\Omega)^m$. Hence, there exists a subsequence (relabeled as before) such that $u_n \rightharpoonup u$ in $\mathrm{W}_0^{1,p}(\Omega)^m$ and $\boldsymbol{u}_n(\boldsymbol{x}) \to \boldsymbol{u}(\boldsymbol{x})$ a.e. in Ω as $u_n \to u$, where the latter convergence follows from the converse of the dominated convergence theorem, see, e.g., Brézis [35, Theorem 4.9, p. 94]. Since K is compact and $\boldsymbol{u}_n(\boldsymbol{x}) \in K$ a.e. in Ω for all $n \in \mathbb{N}$, there holds $\boldsymbol{u}(\boldsymbol{x}) \in K$ a.e. in Ω as well. We obtain with the lemma of Fatou (see, e.g., Brézis [35, Lemma 4.1, p. 90])

$$\begin{aligned}\mathcal{E}_t(u) &\leq \liminf_{n\to\infty} \int_\Omega \Big(\frac{1}{p}|\nabla \boldsymbol{u}_n(\boldsymbol{x})|^p + W(\boldsymbol{u}_n(\boldsymbol{x})) + \imath_K(\boldsymbol{u}_n(\boldsymbol{x})) - \langle \boldsymbol{f}(t), \boldsymbol{u}_n\rangle_{\mathrm{W}_0^{1,p}}\Big) \mathrm{d}\boldsymbol{x} \\ &\leq a\end{aligned}$$

from which $u \in J_a$ follows. Together with the compact embedding

$$\mathrm{L}^\infty(\Omega)^m \cap \mathrm{W}^{1,p}(\Omega)^m \overset{c}{\hookrightarrow} \mathrm{L}^s(\Omega)^m \quad \text{for all } s \in [1, +\infty), \tag{4.0.4}$$

which follows from the Rellich–Kondrachov theorem (see, e.g., Brézis [35, Theorem 9.16, p. 285]) and an interpolation between the Lebesgue spaces, this implies that $\mathcal{E}_t$ also has compact sublevels sets in V for every $t \in [0,T]$. Hence, (3.Ea) and (3.Eb) are fulfilled. The condition (3.Ec) is due to (4.0.f) obviously fulfilled. Now, we want to verify the chain rule condition (3.Ed) and the strong-weak closedness condition (3.Ee). To do so, we show that $\mathcal{E}_t$ is Λ-convex uniformly in $t \in [0,T]$, since in that case the energy functional complies with (3.Ed) and (3.Ee) by Remark 3.2.2. First, the λ-convexity of W yields

$$\begin{aligned}\mathcal{E}_t(\theta\boldsymbol{v} + (1-\theta)\boldsymbol{w}) &\le \theta\mathcal{E}_t(\boldsymbol{v}) + (1-\theta)\mathcal{E}_t(\boldsymbol{w}) + \lambda(1-\theta)\theta\|\boldsymbol{v}-\boldsymbol{w}\|_{\mathrm{L}^2}^2 \\ &\le \theta\mathcal{E}_t(\boldsymbol{v}) + (1-\theta)\mathcal{E}_t(\boldsymbol{w}) + \lambda C(1-\theta)\theta\|\boldsymbol{v}-\boldsymbol{w}\|_{\mathrm{L}^p}^2\end{aligned}$$

for all $\boldsymbol{v},\boldsymbol{w} \in D$, $t \in [0,T]$, and $\theta \in (0,1)$, where we used the Hölder inequality. Therefore, there exists a $\Lambda > 0$ such that $\mathcal{E}_t$ is Λ-convex uniformly in $t \in [0,T]$. Since the λ-convex part of the energy functional is Fréchet differentiable, we obtain with Lemma 2.2.5 and Lemma 2.2.7 that

$$\boldsymbol{\xi} \in \partial\mathcal{E}_t(\boldsymbol{u}) \quad \text{iff} \quad \boldsymbol{\xi}(\boldsymbol{x}) = -\Delta_p\boldsymbol{u}(\boldsymbol{x}) + \mathrm{D}W(\boldsymbol{u}(\boldsymbol{x})) + \partial\imath_K(\boldsymbol{u}(\boldsymbol{x})) \quad \text{for a.e. } \boldsymbol{x} \in \Omega$$

for all $\boldsymbol{u} \in D$, where in turn $\boldsymbol{\eta}(\boldsymbol{x}) \in \partial\imath_K(\boldsymbol{u}(\boldsymbol{x})) \subset V^* = \mathrm{L}^{p^*}(\Omega)^m$ if and only if

$$\int_\Omega \boldsymbol{\eta}(\boldsymbol{x})\boldsymbol{w}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \le \int_\Omega \boldsymbol{\eta}(\boldsymbol{x})\boldsymbol{v}(x)\,\mathrm{d}\boldsymbol{x}$$

for all $\boldsymbol{w} \in V$ with $\boldsymbol{w}(\boldsymbol{x}) \in K$ a.e. in Ω, which follows from (2.2.3).

Finally, we show that the perturbation B fulfills Conditions (3.Ba) and (3.Bb). We first show that B is continuous on sublevel sets of $\mathcal{E}_t$. Therefore, let $t_n \to t$ in $[0,T]$ and $\boldsymbol{u}_n \to \boldsymbol{u}$ in V as $n \to \infty$ and $\sup_{n\in\mathbb{N},t\in[0,T]}\mathcal{E}_t(\boldsymbol{u}_n) < +\infty$. Therefore, there exists a subsequence (labeled as before) such that $\boldsymbol{u}_n(\boldsymbol{x}) \to \boldsymbol{u}(\boldsymbol{x})$ as $n \to \infty$ for a.e. $\boldsymbol{x} \in \Omega$. Since $\boldsymbol{b}$ is a Carathéodory function, we infer that

$$\lim_{n\to\infty} |\boldsymbol{b}(\boldsymbol{x},t_n,\boldsymbol{u}_n(\boldsymbol{x})) - \boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{u}(\boldsymbol{x}))| = 0 \quad \text{for a.e. } \boldsymbol{x} \in \Omega$$

and by (4.0.3)

$$\begin{aligned}|\boldsymbol{b}(\boldsymbol{x},t_n,\boldsymbol{u}_n(\boldsymbol{x})) - \boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{u}(\boldsymbol{x}))| &\le |\boldsymbol{b}(\boldsymbol{x},t_n,\boldsymbol{u}_n(\boldsymbol{x}))| + |\boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{u}(\boldsymbol{x}))| \\ &\le 2h(\boldsymbol{x})2C_{\boldsymbol{b}}, \quad \text{for a.e. } \boldsymbol{x} \in \Omega,\end{aligned}$$

where we have taken into account that $(\boldsymbol{u}_n)_{n\in\mathbb{N}}$ and u are in the domain of $\mathcal{E}_t$ and therefore takes their values in K almost everywhere. Thus, by the dominated convergence theorem, there holds

$$\begin{aligned}&\lim_{n\to\infty}\|B(t_n,\boldsymbol{u}_n) - B(t,\boldsymbol{u})\|_V \\ &\lim_{n\to\infty} = \sup_{\boldsymbol{w}\in V^*,\|\boldsymbol{w}\|_{V^*}\le 1} \int_\Omega \left(-(\boldsymbol{b}(\boldsymbol{x},t_n,\boldsymbol{u}_n(\boldsymbol{x})) - \boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{u}(\boldsymbol{x})))\cdot\boldsymbol{w}(\boldsymbol{x})\right)\mathrm{d}\boldsymbol{x} \\ &\lim_{n\to\infty} \le \left(\int_\Omega |\boldsymbol{b}(\boldsymbol{x},t_n,\boldsymbol{u}_n(\boldsymbol{x})) - \boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{u}(\boldsymbol{x}))|^{q^*}\,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{q^*}}.\end{aligned}$$

We continue by verifying that B is controlled in terms of $\Psi_{\boldsymbol{u}}$ and $\mathcal{E}_t$. Let $c \in (0,1)$, then employing HÖLDER's and YOUNG's inequality with $\varepsilon \in (0, \frac{1}{p})$

$$\begin{aligned}
c\,\psi_{\boldsymbol{u}}^*\left(\frac{B(t,\boldsymbol{u})}{c}\right) &\le c\,C_*\left(\left\|\frac{B(t,\boldsymbol{u})}{c}\right\|_{\mathrm{L}^{q^*}(\Omega)^m}^{q^*} + 1\right)\\
&\le C\left(\|h + C_{\boldsymbol{b}}\|_{\mathrm{L}^{q^*}(\Omega)^m}^{q^*} + 1\right)\\
&\le C\left(\|h\|_{\mathrm{L}^{q^*}(\Omega)}^{q^*} + 1\right)\\
&\le C\Bigg(\left(\frac{1}{p} - \varepsilon\right)\|\nabla\boldsymbol{u}\|_{\mathrm{L}^p(\Omega)^{d\times m}}^p + \|h\|_{\mathrm{L}^{q^*}(\Omega)}^{q^*}\\
&\quad - C_{\varepsilon,p,p^*}\|\boldsymbol{f}(t)\|_{\mathrm{W}^{-1,p^*}(\Omega)^m}^{p^*} + C_{\varepsilon,p,p^*}\|\boldsymbol{f}(t)\|_{\mathrm{W}^{-1,p^*}(\Omega)^m}^{p^*} + 1\Bigg)\\
&\le C\Bigg(\frac{1}{p}\|\nabla\boldsymbol{u}\|_{\mathrm{L}^p(\Omega)^{d\times m}}^p + \langle f(t),\boldsymbol{u}\rangle_{\mathrm{W}_0^{1,p}} + \|h\|_{\mathrm{L}^{q^*}(\Omega)}^{q^*}\\
&\quad + c_{\varepsilon,p,p^*}\|\boldsymbol{f}\|_{\mathrm{C}([0,T];\mathrm{W}^{-1,p^*}(\Omega)^m)}^{p^*} + 1\Bigg)\\
&\le \beta\left(\mathcal{E}_t(\boldsymbol{u}) + 1\right) \quad \text{for all } \boldsymbol{u} \in D, t \in [0,T],
\end{aligned}$$

for a constant $\beta = \beta(\boldsymbol{f}, h, p, q, K, \Omega) > 0$, where $C_{\varepsilon,p,p^*} = (p^*(\varepsilon p)^{\frac{1}{(p-1)}})^{-1}$. Hence, Condition (3.Ba) and (3.Bb) are fulfilled as well. Therefore, by Theorem 3.2.3, for all $\boldsymbol{u}_0 \in D$, there exists an absolutely continuous function $\boldsymbol{u} \in \mathrm{AC}([0,T];V)$ solving (P1) in $V^* = \mathrm{L}^{p^*}(\Omega)^m$ such that the mapping $t \mapsto \mathcal{E}_t(\boldsymbol{u}(t))$ is absolutely continuous, and the energy-dissipation balance holds

$$\begin{aligned}
&\frac{1}{p}|\boldsymbol{u}(t)|_{\mathrm{W}_0^{1,p}(\Omega)^m} + \int_\Omega W(\boldsymbol{u}(t))\,\mathrm{d}\boldsymbol{x} + \langle f(t),\boldsymbol{u}(t)\rangle_{\mathrm{W}_0^{1,p}}\\
&+ \int_0^t\int_\Omega \left(\psi(\boldsymbol{x},\boldsymbol{u}(\boldsymbol{x},r),\partial_t\boldsymbol{u}(\boldsymbol{x},r)) + |\partial_t\boldsymbol{u}(\boldsymbol{x},r)|\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}r\\
&+ \int_0^t\int_\Omega \min_{\boldsymbol{\eta}\in\overline{B}(0,1)}\left(\psi^*(\boldsymbol{x},\boldsymbol{u}(\boldsymbol{x},r),\boldsymbol{\eta} + \boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{u}(\boldsymbol{x},t) + \boldsymbol{\xi}(\boldsymbol{x},t)))\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}r\\
&= \frac{1}{p}|\boldsymbol{u}(s)|_{\mathrm{W}_0^{1,p}(\Omega)^m} + \int_\Omega W(\boldsymbol{u}(s))\,\mathrm{d}\boldsymbol{x} + \langle f'(s),\boldsymbol{u}(s)\rangle_{\mathrm{W}_0^{1,p}}\\
&+ \int_s^t \langle\boldsymbol{f}(r),\boldsymbol{u}'(r)\rangle_{\mathrm{W}_0^{1,p}}\,\mathrm{d}r - \int_s^t\int_\Omega \boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{u}(\boldsymbol{x},t))\cdot\partial_r\boldsymbol{u}(\boldsymbol{x},r)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}r
\end{aligned}$$

for all $s,t \in [0,T]$, where $\boldsymbol{\xi}(t) \in \partial\mathcal{E}_t(\boldsymbol{u}(t)) = \Delta_p\boldsymbol{u}(t) + \mathrm{D}W(\boldsymbol{u}(t)) + \partial\imath_K(\boldsymbol{u}(t))$ a.e. in $(0,T)$.

Part II

Evolution Inclusion of Second Order

Chapter 5

Linearly damped Inertial System

In this chapter, we investigate the abstract CAUCHY problem

$$\begin{cases} u''(t) + \partial\Psi(u'(t)) + \partial\mathcal{E}_t(u(t)) + B(t, u(t), u'(t)) \ni f(t), & t \in (0, T), \\ u(0) = u_0, \quad u'(0) = v_0, \end{cases} \tag{5.0.1}$$

where Ψ is the dissipation potential, $\mathcal{E}_t$ the energy functional, B the perturbation, and f the external force. The functionals and operators are defined on suitable spaces, which will be specified below. Here, the main assumptions are that the leading part of Ψ is defined by a strongly positive, symmetric, and bounded bilinear form a, the energy functional $\mathcal{E}_t$ is λ-convex, and the perturbation B is a strongly continuous perturbation of $\partial\Psi$ and $\partial\mathcal{E}_t$. Within the above-mentioned class of dissipation potentials, we consider the following two cases separately: in the first case (Case **(a)**), we assume that $\Psi(v) = a(v, v)$ and in the second case (Case **(b)**), we assume that $\Psi = \Psi_1 + \Psi_2$, where $\Psi_1(v) = a(v, v)$ and Ψ_2 is a strongly continuous and convex perturbation. Furthermore, we will specifically consider the case when $\mathcal{E}_t$ is convex. As already mentioned in Section 1.2, the energy functional and the dissipation potential are in general, defined on different spaces. An illustrative example in the smooth setting that satisfies all assumptions above is given by

$$\partial_{tt}u - \nabla \cdot \Big(A\nabla\partial_t u\Big) + \nu\, |\partial_t u|^{q-2}\, \partial_t u - \nabla \cdot \Big(|\nabla u|^{p-2}\nabla u\Big) + W'(u) + b(u, \partial_t u) = f,$$

where $p, q > 1$ are to be chosen suitably, $\nu \geq 0$, $A : \mathbb{R}^d \to \mathbb{R}^d$ is a linear, symmetric, and elliptic operator, $W : \mathbb{R} \to \mathbb{R}$ is a double-well potential given by $W(u) = \frac{1}{4}(u^2 - 1)^2$, $b : \mathbb{R} \to \mathbb{R}$ a lower order perturbation, and $f : \mathbb{R} \to \mathbb{R}$ an external force. The energy functional and the dissipation potential are given by

$$\mathcal{E}(u) = \int_\Omega \left(\frac{1}{p}|\nabla u(x)|^p + \frac{1}{4}(u^2(x) - 1)^2\right) \mathrm{d}x$$

and

$$\Psi(v) = \int_\Omega \left(A(x)\nabla v(x) \cdot \nabla v(x) + \frac{\nu}{q}|v(x)|^q\right) \mathrm{d}x$$

and the perturbation is (formally) given by

$$\langle B(u, v), w\rangle_{\mathrm{L}^2} = \int_\Omega b(u(x), v(x))w(x)\,\mathrm{d}x.$$

Note that if $\nu = 0$, we are in Case **(a)** and if $\nu > 0$, we are in Case **(b)**. More, in particular, multi-valued applications will be discussed in Section 7.1 and 7.2.

5.1 Topological assumptions and main result

In the following, let $(U, \|\cdot\|_U)$, $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ and $(\widetilde{W}, \|\cdot\|_{\widetilde{W}})$ be real, separable, and reflexive BANACH spaces and let $(H, |\cdot|, (\cdot,\cdot))$ be a HILBERT space with norm $|\cdot|$ induced by the inner product $(\cdot,\cdot)$.
We will assume the dense, continuous and compact embeddings

$$\begin{cases} U \cap V \overset{d}{\hookrightarrow} U \overset{c,d}{\hookrightarrow} \widetilde{W} \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} \widetilde{W}^* \overset{d}{\hookrightarrow} U^* \overset{d}{\hookrightarrow} V^* + U^* \\ U \cap V \overset{d}{\hookrightarrow} V \overset{c,d}{\hookrightarrow} W \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} W^* \overset{d}{\hookrightarrow} V^* \overset{d}{\hookrightarrow} V^* + U^* \end{cases}$$

and if the perturbation does not explicitly depend on u or u', then we do not need to assume $U \overset{c}{\hookrightarrow} \widetilde{W}$ or $V \overset{c}{\hookrightarrow} W$, respectively, but instead that $V \overset{c}{\hookrightarrow} H$. We stress that we neither assume $U \hookrightarrow V$ nor $V \hookrightarrow U$. The spaces can coincide if a certain embedding is not assumed to be compact. For instance, the cases $V = U$, $\widetilde{W} = H$ or $W = H$ are admissible. Introducing the spaces W and $\widetilde{W}$ allows us to make use of the finer structure of the spaces which enables us to treat additional nonlinearities of lower order. As examples for the appearing spaces, we can think of the SOBOLEV spaces $U = \mathrm{W}^{k,p}(\Omega), V = \mathrm{H}^l(\Omega)$ and the LEBESGUE spaces $W = \mathrm{L}^q(\Omega)$ and $H = \mathrm{L}^2(\Omega)$ or $U = \mathrm{W}^{k,p}(\Omega), V = \mathrm{W}^{s,p}(\Omega)$, $W = \mathrm{H}^l(\Omega)$ and $H = \mathrm{L}^2(\Omega)$ for suitably chosen numbers $k, l \in \mathbb{N}$ and real values $s, p, q > 0$.

Before we present the precise assumptions on the functionals and the operators, we recall some functional analytical facts. First, the space $U \cap V$ equipped with the norm $\|\cdot\|_{U\cap V} = \|\cdot\|_U + \|\cdot\|_V$ is a separable and reflexive BANACH space and the dual space is given by $(U \cap V)^* = U^* + V^*$ with the norm $\|\xi\|_{U^*+V^*} = \inf_{\substack{\xi_1 \in U^*, \xi_2 \in V^* \\ \xi = \xi_1 + \xi_2}} \max\{\|\xi_1\|_{U^*}, \|\xi_2\|_{V^*}\}$, see Example 2.3.6. Furthermore, the duality pairing between $U \cap V$ and $U^* + V^*$ is given by

$$\langle f, v\rangle_{(U^*+V^*)\times(U\cap V)} = \langle f_1, u\rangle_{U^*\times U} + \langle f_2, u\rangle_{V^*\times V}, \quad u \in U \cap V,$$

for all $v \in U \cap V$ and any partition $f = f_1 + f_2$ with $f_1 \in U$ and $f_2 \in V$. Second, for any $p \in [1, +\infty]$, there holds $\mathrm{L}^p(0,T;U) \cap \mathrm{L}^p(0,T;V) = \mathrm{L}^p(0,T;U \cap V)$, where the measurability immediately follows from the PETTIS theorem, see, e.g., DIESTEL & UHL [58, Theorem 2, p. 42]. And third, for the BANACH spaces $X, Y \in \{U \cap V, U, V, W, \widetilde{W}, H\}$ satisfying the embedding $X \hookrightarrow Y$, there holds

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

see, e.g, BRÉZIS [35, Remark 3, pp. 136] and GAJEWSKI et al. [84, Kapitel 1, §5].
Now, we want to collect all assumptions concerning the dissipation potential Ψ, the energy functional $\mathcal{E}$, the perturbation B as well as the external force f. Since the the subdifferential of the main part of Ψ is linear, we refer to the inclusion (5.0.1) in the given framework as *linearly damped inertial system* $(U, V, W, \widetilde{W}, H, \mathcal{E}, \Psi, B, f)$. The

assumptions we make for the linearly damped inertial system resembles the structure to those we made for the perturbed gradient system where the same evolution inclusion has been investigated after neglecting the inertial term $u''(t)$. Involving inertia makes the situation much more delicate. As a consequence, we will impose, in general, stronger conditions on the functionals and operator in order to ensure solvability of the problem. Hereinafter, we collect the assumptions for the dissipation potential Ψ and remind the reader that we distinguish two cases **(a)** and **(b)**.

(5.Ψ) **Dissipation potential.**
Case **(a)**: we assume that there exists a strongly positive, symmetric, and continuous bilinear form $a : V \times V \to \mathbb{R}$ such that $\Psi(v) = \frac{1}{2}a(v, v)$, i.e., there is a constant $\mu > 0$ such that

$$\mu\|v\|_V^2 \leq \Psi(v) \quad \text{for all } v \in V. \tag{5.1.1}$$

Case **(b)**: we assume that $\Psi = \Psi_1 + \Psi_2$, where $\Psi_1(v) = \frac{1}{2}a(v, v)$ with the bilinear form $a : V \times V \to \mathbb{R}$ as above and $\Psi_2 : W \to \mathbb{R}$ to be a lower semicontinuous and convex functional with $\Psi_2(0) = 0$ satisfying the following growth condition: there exists a positive number $q > 1$ and constants $\hat{c}, \hat{C} > 0$ such that

$$\hat{c}(\|v\|_W^q - 1) \leq \Psi_2(v) \leq \hat{C}(\|v\|_W^q + 1) \quad \text{for all } v \in W. \tag{5.1.2}$$

In addition, we assume that Ψ_2 is GÂTEAUX differentiable on V with derivative $\mathrm{D}_G\Psi_2$ being continuous as mapping from W to $U^* + V^*$ and satisfying the following growth condition: for all $R > 0$, there exists a positive real constant $C_R > 0$ such that

$$\|D_G\Psi_2(v)\|_{U^*+V^*} \leq C_R(1 + \|v\|_W^{q-1}) \quad \text{for all } v \in W \text{ with } |v| \leq R. \tag{5.1.3}$$

Remark 5.1.1

i) Assumption (5.Ψ) yields the convexity and continuity of the mapping $v \mapsto \Psi(v)$. Furthermore, Ψ is GÂTEAUX differentiable with the GÂTEAUX derivative given by a positive, linear bounded and symmetric operator $A : V \to V^*$ such that $\partial\Psi(v) = \{Av\}$ and the potential can be expressed by $\Psi(v) = \frac{1}{2}\langle Av, v\rangle$. Assumption (5.Ψ) implies that the LEGENDRE–FENCHEL transform Ψ^* of Ψ is convex, continuous, finite everywhere, i.e., $\mathrm{dom}(\Psi^*) = V^*$, and can be 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 positive, which follows from the LAX–MILGRAM theorem, see, e.g., BRÉZIS [35, Corollary 5.8, p. 140].

ii) From the properties of the conjugate, we obtain from (5.1.1) and (5.1.2) the following growth condition for the conjugates $\Psi_1^* : V^* \to \mathbb{R}$ and $\Psi_2^* : V^* \to (-\infty, +\infty]$: there exist positive constants $\bar{c}, \bar{C} >$ such that

$$\begin{aligned} \bar{c}\|\xi\|_{V^*}^2 \leq{} & \Psi_1^*(\xi) \leq \bar{C}\|\xi\|_{V^*}^2 \quad \text{for all } \xi \in V^*, \\ \left.\begin{matrix} \bar{c}(\|\xi\|_{W^*}^{q^*} - 1) \\ +\infty \end{matrix}\right\} \leq{} & \Psi_2^*(\xi) \leq \begin{cases} \bar{C}(\|\xi\|_{W^*}^{q^*} + 1) & \text{if } \xi \in W^* \\ +\infty & \text{otherwise ,} \end{cases} \end{aligned} \tag{5.1.4}$$

where $q^* > 1$ denotes the conjugate exponent of q. In order to justify the formula (5.1.4), it is not restrictive to show it for $\Psi^2(v) = \frac{1}{q}\|v\|_W^q, v \in V$. To do so, we employ Lemma 2.3.2, which shows that the conjugate function f^* of any proper, convex, and lower semicontinuous function $f : V \to \overline{\mathbb{R}}$ is also proper, convex, and lower semicontinuous, and that $f^{**} = f$. Thus, defining $\widetilde{\Psi} : V^* \to \overline{\mathbb{R}}$ through

$$\widetilde{\Psi}(\xi) = \begin{cases} \frac{1}{q^*}\|\xi\|_{W^*}^{q^*} & \text{if } \xi \in W^* \\ +\infty & \text{otherwise }, \end{cases}$$

it follows that $\widetilde{\Psi}$ is a proper, convex, and lower semicontinuous function on V^* which easily follows from the fact that a function is convex and lower semicontinuous if and only if its epigraph is convex and closed, see Lemma 2.1.2. Then, we show that $\widetilde{\Psi}^* = \Psi_2$ which in turn implies $\Psi_2^* = \widetilde{\Psi} = \widetilde{\Psi}^{**}$ where the first equality follows from

$$\begin{aligned} \widetilde{\Psi}^*(v) &= \sup_{\xi \in V^*} \left\{ \langle \xi, v \rangle_{V^* \times V} - \widetilde{\Psi}^*(\xi) \right\} \\ &= \sup_{\xi \in W^*} \left\{ \langle \xi, v \rangle_{W^* \times W} - \frac{1}{q^*}\|\xi\|_{W^*}^{q^*} \right\} \\ &= \frac{1}{q}\|v\|_W^q \\ &= \Psi_2(v) \quad \text{for all } v \in W, \end{aligned}$$

where we have used that $(\frac{1}{q^*}\|\cdot\|_W^{q^*})^* = \frac{1}{q}\|v\|_W^q$ on W, see Example 2.3.4.

iii) We remark that we could also allow for a time-dependent dissipation potential $\Psi_t = \Psi_t^1 + \Psi_t^2$ when we assume that $t \mapsto a(t, u, v) \in \mathrm{C}([0, T]) \cap \mathrm{C}^1(0, T)$ for all $u, v \in V$ and a strong monotonicity and boundedness of $A(t) : V \to V^*$ uniformly in time as well as a slight modification of Assumption (3.Ec) and (3.Bb), whereas for Ψ_t^2 we would assume that for all $t \in [0, T]$, the functional Ψ_t^2 is lower semicontinuous, convex and Gâteaux differentiable with continuous Gâteaux $D_G\Psi_t^2$ being continuous on $[0, T] \times W$ and satisfying the Conditions (5.1.2) and (5.1.3) uniformly in time. For simplicity, we will not consider this case here.

We proceed with collecting the assumptions for the energy functional $\mathcal{E}$. To do so, we define $V_\lambda = U$ if $\lambda = 0$ and $V_\lambda = U \cap V$ if $\lambda > 0$. We make this distinction because for the convex case, i.e. when $\lambda = 0$, we will obtain a stronger result meaning that the initial value u_0 can be chosen to be in $\mathrm{dom}(\mathcal{E}_t)$ instead of $\mathrm{dom}(\mathcal{E}_t) \cap V$ as in the λ-convex case with $\lambda \neq 0$, and that the subgradient of $\mathcal{E}_t$ is in U^* instead of $U^* + V^*$, see Theorem 5.1.4.

(5.Ea) **Lower semicontinuity.** For all $t \in [0, T]$, the functional $\mathcal{E}_t : U \to (-\infty, +\infty]$ is proper and sequentially weakly lower semicontinuous with time-independent effective domain $D := \mathrm{dom}(\mathcal{E}_t) \subset U$ for all $t \in [0, T]$. Furthermore, the set $D \cap V$ is dense in D in the topology of U, and if $\mathcal{E}_t$ is convex, the interior of D is non-empty.

(5.Eb) **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

$$\mathcal{E}_t(u) \geq C_0 \quad \text{for all } u \in U \text{ and } t \in [0,T].$$

Since a potential is unique up to a constant, we assume without loss of generality $C_0 = 0$.

(5.Ec) **Coercivity.** For every $t \in [0,T]$, $\mathcal{E}_t$ has bounded sublevel sets in U.

(5.Ed) **Control of the time derivative.** For all $u \in U$, the mapping $t \mapsto \mathcal{E}_t(u)$ is in $\mathrm{C}([0,T]) \cap \mathrm{C}^1(0,T)$ 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

$$|\partial_t \mathcal{E}_t(u)| \leq C_1 \mathcal{E}_t(u) \quad \text{for all } t \in (0,T) \text{ and } u \in U.$$

(5.Ee) **Closedness of** $\mathrm{Gr}(\partial\mathcal{E})$**.** For all sequences of measurable functions $(\mathbf{t}_n)_{n\in\mathbb{N}}$ with $\mathbf{t}_n : [0,T] \to [0,T]$, $n \in \mathbb{N}$, $(u_n)_{n\in\mathbb{N}}$, $(\xi_n)_{n\in\mathbb{N}}$, and measurable functions u, ξ satisfying

a) $\mathbf{t}_n(t) \to t$ for a.e. $\mathrm{t} \in (0,T)$, as $n \to \infty$,

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

c) $\xi_n(t) \in \partial_{V_\lambda} \mathcal{E}_{\mathbf{t}_n(t)}(u_n(t))$ a.e. in $(0,T), n \in \mathbb{N}$,

d) $u_n - \tilde{u}_0 \overset{*}{\rightharpoonup} u - \tilde{u}_0$ in $\mathrm{L}^\infty(0,T;V_\lambda)$ and $u_n - \tilde{u}_0 \to u - \tilde{u}_0$ in $\mathrm{L}^2(0,T;V)$ for any $\tilde{u}_0 \in D$ and $\xi_n \rightharpoonup \xi$ in $\mathrm{L}^2(0,T;V_\lambda^*)$ as $n \to \infty$. Additionally, there exists a constant $C_3 > 0$ such that for sufficiently small $h > 0$, there holds Case **(a)**:

$$\sup_{n\in\mathbb{N}} \|\sigma_h u_n - u_n\|_{\mathrm{L}^2(0,T-h,V)} \leq C_3 h \tag{5.1.5}$$

Case **(b):**

$$\sup_{n\in\mathbb{N}} \|\sigma_h u_n - u_n\|_{\mathrm{L}^2(0,T-h,V)\cap\mathrm{L}^r(0,T-h,W)} \leq C_3 h, \tag{5.1.6}$$

where $\sigma_h v := \chi_{[0,T-h]} v(\cdot + h)$ for any function $v : [0,T] \to V$,

e) $\limsup_{n\to\infty} \int_0^T \langle \xi_n(t) - \xi(t), u_n(t) - u(t) \rangle_{V_\lambda^* \times V_\lambda} \,\mathrm{d}t \leq 0$,

we have the relations

$$\xi(t) \in \partial_{V_\lambda} \mathcal{E}_t(u(t)) \subset V_\lambda^*, \quad \mathcal{E}_{\mathbf{t}_n(t)}(u_n(t)) \to \mathcal{E}_t(u(t)) \quad \text{as } n \to \infty$$
$$\text{and} \quad \limsup_{n\to\infty} \partial_t \mathcal{E}_t(u_n(t)) \leq \partial_t \mathcal{E}_t(u(t)) \quad \text{for a.e. } t \in (0,T).$$

(5.Ef) **λ-convexity.** There exists $\lambda \geq 0$ such that for every $t \in [0,T]$, the energy functional $\mathcal{E}_t$ is λ-convex on V (by extending $\mathcal{E}$ on V), i.e., for all $u, v \in U \cap V$ and $\vartheta \in (0,1)$, there holds

$$\mathcal{E}_t(\vartheta v + (1-\vartheta)u) \leq \vartheta \mathcal{E}_t(v) + (1-\vartheta)\mathcal{E}_t(u) + \lambda\vartheta(1-\vartheta)\|v-u\|_V^2. \tag{5.1.7}$$

(5.Eg) **Control of the subgradient.** There exist constants $C_4 > 0$ and $\sigma > 0$ such that

$$\|\xi\|_{V_\lambda^*}^\sigma \leq C_4(1 + \mathcal{E}_t(u) + \|u\|_{V_\lambda}) \quad \forall t \in [0,T], u \in D(\partial_{V_\lambda}\mathcal{E}_t),\ \xi \in \partial_{V_\lambda}\mathcal{E}_t(u).$$

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

Remark 5.1.2

i) From Assumption (5.Ed), we deduce again with GRONWALL's lemma (Lemma A.1.1) the chain of inequalities

$$\mathrm{e}^{-C_1|t-s|}\mathcal{E}_s(u) \leq \mathcal{E}_t(u) \leq \mathrm{e}^{C_1|t-s|}\mathcal{E}_s(u) \quad \text{for all } s,t \in [0,T],\ u \in D.$$

In particular, there holds

$$\mathcal{G}(u) = \sup_{t\in[0,T]} \mathcal{E}_t(u) \leq \mathrm{e}^{C_1 T} \inf_{t\in[0,T]} \mathcal{E}_t(u) \quad \text{for all } u \in D.$$

ii) In Case **(b)** it is possible to improve the assumption of λ-convexity in the following way: there exist positive constants $\lambda_1, \lambda_2 > 0$ such that

$$\begin{aligned}\mathcal{E}_t(\vartheta u + (1-\vartheta)v) \leq& \vartheta\mathcal{E}_t(u) + (1-\vartheta)\mathcal{E}_t(v)\\ &+ \vartheta(1-\vartheta)\left(\lambda_1\|u-v\|_V^2 + \lambda_2\Psi(u-v)^{\frac{1}{q}}|u-v|\right)\end{aligned}$$

for all $u \in D, v \in V$ and $\vartheta \in (0,1)$, where $q > 1$ comes from Assumption (3.Ψa).

Finally, we present the assumptions on the non-variational non-monotone perturbation B and the external force f.

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

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

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

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

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

Remark 5.1.3 *i)* In fact, the continuity of the perturbation with image in V^* is only needed to show the energy-dissipation inequality (5.1.10). If we only address the existence of solutions to the inclusion (5.1.9) without the energy-dissipation inequality, then it is sufficient to suppose that $B : [0,T]\times\widetilde{W}\times W \to U^* + V^*$ is a mapping with values in $U^* + V^*$ which is continuous on sublevel sets of the energy, see Example 7.3 for such an instance.

ii) The condition (5.Bb) can be relaxed to $f \in \mathrm{L}^1(0,T;H) + \mathrm{L}^2(0,T;V^*)$ in the Case **(a)** and to $f \in \mathrm{L}^1(0,T;H) + \mathrm{L}^2(0,T;V^*) + \mathrm{L}^{q^*}(0,T;W^*)$ in the case **(b)**, where $q^* > 1$ is again the conjugate exponent to $q > 1$ from Assumption (5.Ψ).

5.1.1 Discussion of the assumptions

Having collected the assumptions on the system $(U, V, W, H, \mathcal{E}, \Psi, B, f)$ system, we want to discuss several conditions more in detail apart from the assertions and implications made in the remarks. As for perturbed gradient systems, we want to discuss the practical meaning of the assumptions and provide sufficient conditions for them to hold true.

As we already mentioned in Section 1.2, evolution equations of second order are, in general, more delicate than evolution equations of first order because of the nonsmoothing effect in time caused by the term $\partial_{tt}u$. This leads to a formation of discontinuities or a blow-up of a solution in finite time despite having smooth initial data which makes it more difficult to prove strong solutions, see, e.g., Zeidler [164, Section 33.7] for a discussion of these phenomena in connection with problems arising in physics. Therefore, we need well-adjusted assumptions which are, in general, stronger than for perturbed gradient systems. However, here we deal with the case where the energy functional and the dissipation potential are defined on different spaces which has not been considered in Chapter 3.

Ad (5.Ψ). Here, the dissipation potential is (in Case **(b)**) the sum of a leading part Ψ_1, which is defined by a strongly positive and bounded bilinear form, and a strongly continuous perturbation Ψ_2 of polynomial growth. As mentioned in Remark 5.1.1 *i*), the subdifferential $\partial\Psi = \{A + \mathrm{D}_G\Psi^2\}$ is given by a positive, linear, bounded, and symmetric operator A and a strongly continuous perturbation $\mathrm{D}_G\Psi^2$. The important assumption here is that A is a positive, linear, bounded, and symmetric operator, which is crucial in identifying the limit of a sequence of approximate solutions stemming from a discretization scheme, see Section 5.2. As we mentioned in Remark 3.2.1 *iii*), we can allow more general time-dependent dissipation potentials. However, we will not assume that for simplicity.

Admissible examples of dissipation potentials are, e.g.,

$$\Psi_1(v) = \frac{1}{2}\int_\Omega |\nabla^n v(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x} \quad \text{and} \quad \Psi_2(v) = \frac{1}{p}\int_\Omega |\nabla^m v(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x}$$

on the Sobolev space $\mathrm{H}^n(\Omega)$ for any $m, n \in \mathbb{N}$ with $m < n$ and $p \in (1, +\infty)$ such that the compact embedding $\mathrm{H}^n(\Omega) \overset{c}{\hookrightarrow} \mathrm{L}^p(\Omega)$ holds.

Ad (5.E). As for perturbed gradient systems, we assume that the effective domain of $\mathcal{E}_t$ is time-independent, which in fact is already implied by Condition (3.Ed), see Remark 5.1.2 *i*). The assumption that $D \cap V$ is dense in D in Condition (5.Ea) ensures the existence of an approximating sequence in $D \cap V$ to any initial value $u_0 \in D$, which is needed in order to obtain a priori estimates for the interpolations in Lemma 5.3.1. The non-emptiness of D in the case that $\mathcal{E}_t$ is convex ensures the existence of a continuity point for $\mathcal{E}_t$, see Ekeland & Temam [69, Corollary 2.5., p. 13], which in turn allows us to use the variational sum rule in Lemma 2.2.7.

The Assumption (5.Ee) replaces the strong-weak closedness condition (3.Ee) for $\partial\mathcal{E}_t$ in Chapter 3. The same assumption can not be made here, since we do not obtain a pointwise strong convergence $\overline{U}_{\tau_n}(t) \to u(t)$ in V as $n \to \infty$ later in the proof of the existence result due to the fact that U is not assumed to be compactly embedded in V. Instead, we have weak convergence on certain BOCHNER spaces. As we have seen by Lemma 2.4.2 and Lemma 2.6.1, a sufficient condition for the subdifferential of a convex functionals to be weak-weak closedness on a suitable BOCHNER space is given by Condition (5.Ee) e). However, since $\mathcal{E}_t$ is not convex but λ-convex, the above-mentioned lemmas can not be employed directly and therefore we enforce the implication by imposing this condition which also encounters for the time dependence of the energy functional. The condition is formulated in such a way that it can be applied to the piecewise constant interpolations $u_n(t) := U_\tau^n$ and $\mathbf{t}_n(t) = t_n$ for $t \in (t_{n-1}, t_n]$ arising from a discretization scheme, see Section 5.2. For a sequence of weakly differentiable functions $(u_n)_{n\in\mathbb{N}}$ the conditions (5.1.5) and (5.1.6) are satisfied when it is bounded in $\mathrm{H}^\sigma(0,T;V)$ and $\mathrm{H}^\sigma(0,T;V) \cap \mathrm{W}^{1,r}(0,T;W)$, respectively.

The Condition (5.Eg) is necessary to obtain appropriate a priori estimates for the subgradients of $\mathcal{E}_t$, which in turn is needed to obtain a priori estimates for u''. The situation was different for perturbed gradient systems, since the subdifferential of $\mathcal{E}_t$ could be controlled by Ψ_u^*. In this situation, the sum of the subgradient of $\mathcal{E}_t$ and u'' are controlled by Ψ^* which necessitates independent estimates. Condition (5.Eg) could be replaced by the more general condition that $\partial\mathcal{E}_t$ is a bounded operator.

Ad (5.B). Due to the same structure of the conditions for the perturbation in perturbed gradient systems and here, the same conclusions hold here as well. The difference is that here the perturbation also depends nonlinearly on u'.

Having discussed all assumptions, we are in a position to state the main result which includes the notion of solution to (5.0.1).

Theorem 5.1.4 (Existence result) *Let the linearly damped inertial system* $(U, V, W, H, \mathcal{E}, \Psi, B, f)$ *be given and fulfill Assumptions* (5.E), (5.Ψ) *and* (5.B) *as well as Assumption* (5.f). *Let the initial values* $u_0 \in D \cap V_\lambda$ *and* $v_0 \in H$ *be given and assume that there exists a sequence* $u_0^k \in D \cap V$ *such that*

$$u_0^k \to u_0 \quad \text{in } V_\lambda \text{ as } k \to \infty \qquad \text{and} \quad \sup_{k\in\mathbb{N}} \mathcal{E}_0(u_0^k) < +\infty.$$

Then, there exists a strong solution to (5.0.1)*, i.e., there exist functions*

$$u \in \mathrm{C}_w([0,T];U) \cap \mathrm{W}^{1,\infty}(0,T;H) \text{ with } u' \in \mathrm{L}^2(0,T;V), \quad \xi \in \mathrm{L}^\infty(0,T;V_\lambda^*), \tag{5.1.8}$$

additionally satisfying $u \in \mathrm{H}^2(0,T;U^* + V^*)$ *in Case* ***(a)*** *and* $u \in \mathrm{W}^{1,q}(0,T;W) \cap \mathrm{W}^{2,\min\{2,q^*\}}(0,T;U^* + V^*)$ *in Case* ***(b)*** *such that the initial conditions* $u(0) = u_0$ *in* V_λ *and* $u'(0) = v_0$ *in* H *as well as the inclusions*

$$\xi(t) \in \partial_{V_\lambda}\mathcal{E}_t(u(t)), \; f(t) \in u''(t) + \partial\Psi(u'(t)) + \xi(t) + B(t, u(t), u'(t)) \quad \text{in } U^* + V^* \tag{5.1.9}$$

are fulfilled for almost every $t \in (0,T)$*. Furthermore, the energy-dissipation inequality*

$$\begin{aligned}&\frac{1}{2}|u'(t)|^2 + \mathcal{E}_t(u(t)) + \int_s^t \left(\Psi(u'(r)) + \Psi^*(S(r) - \xi(r) - u''(r))\right) \mathrm{d}r \\ &\le \frac{1}{2}|u'(s)|^2 + \mathcal{E}_s(u(s)) + \int_s^t \partial_r\mathcal{E}_r(u(r))\,\mathrm{d}r + \int_s^t \langle S(r), u'(r)\rangle_{V^*\times V}\,\mathrm{d}r,\end{aligned} \tag{5.1.10}$$

holds for all $0 < t \leq T$ *for* $s = 0$, *and almost every* $s \in (0,t)$, *where* $S(r) := f(r) - B(r, u(r), u'(r))$, $r \in [0,T]$, *and* $V_\lambda = U$ *if* $\mathcal{E}_t$ *is convex, i.e.,* $\lambda = 0$, *and* $V_\lambda = U \cap V$ *if* $\mathcal{E}_t$ *is* λ*-convex with* $\lambda > 0$.

5.2 Variational approxiomation scheme

The proof of Theorem 5.1.4 is based on the construction of strong solutions to (5.0.1) via a semi-implicit time discretization scheme similar to Chapter 3. 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 (5.0.1). Therefore, let for $N \in \mathbb{N}\backslash\{0\}$

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

be an equidistant partition of the time interval $[0,T]$ with step size $\tau := T/N$, where we omit the dependence of the nodes of the partition on the step size for simplicity. The discretization of (5.0.1) is then given by

$$\frac{V_\tau^n - V_\tau^{n-1}}{\tau} + \partial_{V_\lambda}\Psi\left(V_\tau^n\right) + \partial_{V_\lambda}\mathcal{E}_{t_n}(U_\tau^n) + B\left(t_n, U_\tau^{n-1}, V_\tau^{n-1}\right) \ni f_\tau^n \quad \text{in } U^* + V^* \tag{5.2.1}$$

for $n = 1, \ldots, N$, where $V_\tau^0 = v_0$, $V_\tau^n := \frac{U_\tau^n - U_\tau^{n-1}}{\tau}$, and $f_\tau^n := \frac{1}{\tau}\int_{t_{n-1}}^{t_n} f(\sigma)\,\mathrm{d}\sigma, n = 1, \ldots, n$, where ∂_{V_λ} denotes the subdifferential operator with respect to the strong topology of V_λ. The inclusion (5.2.1) is equivalent to saying that there exists a subgradient $\xi_\tau^n \in \partial_{V_\lambda}\mathcal{E}_t(U_\tau^n)$ such that

$$\frac{V_\tau^n - V_\tau^{n-1}}{\tau} + \mathrm{D}_G\Psi(V_\tau^n) + \xi_\tau^n + B\left(t_n, U_\tau^{n-1}, V_\tau^{n-1}\right) = f_\tau^n \quad \text{in } U^* + V^*,$$

where $\mathrm{D}_G\Psi(V_\tau^n) = AV_\tau^n$ in Case **(a)** and $\mathrm{D}_G\Psi(V_\tau^n) = AV_\tau^n + \mathrm{D}_G\Psi_2(V_\tau^n)$ in Case **(b)**.

The values $U_\tau^n \approx u(t_n)$ and $V_\tau^n \approx u'(t_n)$ shall approximate the exact solution and its time derivative, and are to be determined successively from (5.2.1). By Lemma 2.2.5, it follows that the approximate value U_τ^n is characterized as the solution to the Euler–Lagrange equation associated with the mapping

$$u \mapsto \Phi(\tau, t_{n-1}, U_\tau^{n-1}, U_\tau^{n-2}, B(t_n, U_\tau^{n-1}, V_\tau^{n-1}) - f_\tau^n; u),$$

where

$$\Phi(r, t, v, w, \eta; u) = \frac{1}{2r^2}|u - 2v - w|^2 + r\Psi\left(\frac{u - v}{r}\right) + \mathcal{E}_{t+r}(u) + \langle \eta, u \rangle_{V^* \times V}$$

for $r \in \mathbb{R}^{>0}, t \in [0,T)$ with $r + t \in [0,T]$, $u \in D, v \in V, w \in H$, and $\eta \in V^*$.
We end up with the recursive scheme

$$\begin{cases} U_\tau^0 \in D \cap V \text{ and } V_\tau^0 \in V \text{ are given; whenever } U_\tau^1, \ldots, U_\tau^{n-1} \in D \cap V \text{ are known,} \\ \text{find } U_\tau^n \in J_{\tau, t_{n-1}}(U_\tau^{n-1}, U_\tau^{n-2}; B(t_n, U_\tau^{n-1}, V_\tau^{n-1}) - f_\tau^n) \end{cases} \tag{5.2.2}$$

for $n = 1, \dots, N$, where $J_{r,t}(v, w; \eta) := \operatorname{argmin}_{u \in U \cap V} \Phi(r, t, v, w, \eta; u)$ and $U_\tau^{-1} = U_\tau^0 - V_\tau^0 \tau$.

The following lemma ensures the solvability of the variational scheme (5.2.2).

Lemma 5.2.1 *Let the linearly damped inertial system* $(U, V, W, W, H, \mathcal{E}, \Psi)$ *be given and let the Conditions* (5.Ea)-(5.Ec), (5.Ef), *and* (5.Ψ) *be fulfilled. Furthermore, let* $r \in (0, T)$ *and* $t \in [0, T)$ *with* $r + t \le T$ *as well as* $v \in V, w \in H$ *and* $\eta \in V^*$. *Then, the set* $J_{r,t}(v, w; \eta)$ *is non-empty and single valued if* $r \le \frac{\mu}{4\lambda}$, *where* μ *and* λ *are from* (5.Ψ) and (5.Ef), *respectively. Furthermore, to every* $u \in J_{r,t}(v, w; \eta)$ *there exists* $\xi \in \partial_{V_\lambda} \mathcal{E}_t(u) \subset V_\lambda^*$ *such that*

$$\frac{u - 2v - w}{r^2} + \mathrm{D}_G \Psi\left(\frac{u - v}{r}\right) + \xi + \eta = 0 \qquad \text{in } U^* + V^*.$$

Proof. Since the proof is similar for the cases **(a)** and **(b)**, we restrict the proof by showing the assertion for the case **(b)**. Let $u \in D \cap V, v \in V, w \in H, \eta \in V^*$, and $r \in (0, T), t \in [0, T)$ with $r + t \le T$ be given. Employing the Fenchel–Young inequality, we obtain

$$\begin{aligned}
\Phi(r, t, v, w, \eta; u) &= \frac{1}{2r^2}|u - 2v + w|^2 + r\Psi\left(\frac{u - v}{r}\right) + \mathcal{E}_{t+r}(u) + \langle \eta, u \rangle_{V^* \times V} \\
&= \frac{1}{2r^2}|u - 2v + w|^2 + r\Psi_1\left(\frac{u - v}{r}\right) + r\Psi_2\left(\frac{u - v}{r}\right) + \mathcal{E}_{t+r}(u) \\
&\quad + \langle \eta, v \rangle_{V^* \times V} \\
&\ge \frac{1}{2r^2}|u - 2v + w|^2 + \frac{1}{r}\Psi_1(u - v) + r\tilde{c}\left(\left\|\frac{u - v}{r}\right\|_W^q - 1\right) + \mathcal{E}_{t+r}(u) \\
&\quad + \langle \eta, u - v \rangle_{V^* \times V} + \langle \eta, v \rangle_{V^* \times V} \\
&\ge \frac{1}{2r^2}|u - 2v + w|^2 + \left(\frac{1}{r} - \varepsilon\right)\Psi_1(u - v) - r\tilde{c} + \mathcal{E}_{t+r}(u) \\
&\quad - \varepsilon\Psi_1^*\left(-\frac{\eta}{\varepsilon}\right) - \langle \eta, v \rangle_{V^* \times V} \qquad (5.2.3) \\
&\ge \frac{1}{2r^2}|u - 2v + w|^2 + \left(\frac{1}{r} - \varepsilon\right)\Psi_1(u - v) - \varepsilon\Psi_1^*\left(\frac{\eta}{\varepsilon}\right) - r\tilde{c} \\
&\quad - \langle \eta, v \rangle_{V^* \times V},
\end{aligned}$$

where $0 < \varepsilon < \frac{1}{r}$. This implies, on the one hand, $\inf_{u \in U \cap V} \Phi(r, t, v, w, \eta; u) > -\infty$. On the other hand, we observe that

$$\inf_{u \in U \cap V} \Phi(r, t, v, w, \eta; u) \le \frac{1}{2r^2}|\bar{u} - 2v + w|^2 + r\Psi\left(\frac{\bar{u} - v}{r}\right) + \mathcal{E}_{t+r}(\bar{u}) - \langle \eta, \bar{u} \rangle_{V^* \times V} \tag{5.2.4}$$

for any $u_0 \in D \cap V$, so that $\inf_{u \in U \cap V} \Phi(r, t, v, w, \eta; u) < +\infty$ holds as well. It remains to show that the global minimum is attained by an element in $D \cap V$. In order to show that, let $(u_n)_{n \in \mathbb{N}} \subset U \cap V$ be a minimizing sequence for $\Phi(r, t, v, w, \eta; \cdot)$. From (5.2.3) and the coercivity of Ψ_1 and $\mathcal{E}$, we deduce that $(u_n)_{n \in \mathbb{N}} \subset U \cap V$ is contained in a sublevel set of Ψ_1 and $\mathcal{E}$, and thus bounded in $U \cap V$. Hence, by reflexivity of $U \cap V$, there exists a subsequence (not relabeled) which converges weakly in $U \cap V$

to a limit $\tilde{u} \in U \cap V$. By the sequential weak lower semicontinuity of the mapping $u \mapsto \Phi(r,t,v,w,\eta;u)$ on $U \cap V$, we have

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

and therefore, $u \in J_{r,t}(v,w;\eta) \neq \emptyset$ and $u \in D \cap V$. If $r > 0$ is sufficiently small, then there is a unique global minimizer. Indeed, assuming there are two different global minimizer $\tilde{u}_0, \tilde{u}_1 \in D \cap V$, then in view of the λ-convexity of $\mathcal{E}_t$, the convexity of Ψ_2 and the fact that $|\cdot|^2$ and Ψ_1 fulfil a parallelogram identity, we obtain for every $s \in (0,1)$

$$\begin{aligned}
&\Phi(r,t,v,w,\eta;s\tilde{u}_0+(1-s)\tilde{u}_1)\\
&\leq s\Phi(r,t,v,w,\eta;\tilde{u}_0)+(1-s)\Phi(r,t,v,w,\eta;\tilde{u}_1)-s(1-s)|\tilde{u}_0-\tilde{u}_1|^2+\lambda\|\tilde{u}_0-\tilde{u}_1\|_V^2\\
&\quad-\frac{s(1-s)}{r}\Psi_1(\tilde{u}_0-\tilde{u}_1)\\
&=\min_{\tilde{v}\in U\cap V}\Phi(r,t,v,w,\eta;\tilde{v})-s(1-s)|\tilde{u}_0-\tilde{u}_1|^2-\frac{s(1-s)}{r}\Psi_1(\tilde{u}_0-\tilde{u}_1)+\lambda\|\tilde{u}_0-\tilde{u}_1\|_V^2\\
&\leq\min_{\tilde{v}\in U\cap V}\Phi(r,t,v,w,\eta;\tilde{v})-s(1-s)|\tilde{u}_0-\tilde{u}_1|^2-\left(\frac{s(1-s)}{r}-\frac{\lambda}{\mu}\right)\Psi_1(\tilde{u}_0-\tilde{u}_1),
\end{aligned}$$

where we also used the strong positivity of Ψ_1 with constant $\mu > 0$. Choosing $s = \frac{1}{2}$, the uniqueness follows. In order to prove the last assertion, we first assume that the energy functional is λ-convex with $\lambda > 0$. Then, from FERMAT's theorem, we know that for any minimizer $u \in J_{r,t}(v,w;\eta)$, the functional $\Phi(r,t,v,w,\eta;\cdot)$ is subdifferentiable in u and there holds

$$\begin{aligned}
0 &\in \partial_{U\cap V}\Phi(r,t,v,w,\eta;u)\\
&=\partial_{U\cap V}\left(\frac{1}{2r^2}|u-2v+w|^2+r\Psi\left(\frac{u-v}{r}\right)+\mathcal{E}_{t+r}(u)+\langle\eta,u\rangle_{V^*\times V}\right)
\end{aligned}$$

Since all terms expect from the energy functional are convex and GÂTEAUX differentiable on the space $U \cap V$, we obtain with Lemma 2.2.5 that $\mathcal{E}_t$ is subdifferentiable in u and there holds

$$\frac{u-2v-w}{r^2}+\mathrm{D}_G\Psi\left(\frac{u-v}{r}\right)+\eta\in\partial_{U\cap V}\mathcal{E}_t(u).$$

Thus, we define $\xi := \frac{u-2v-w}{r^2} + \mathrm{D}_G\Psi\left(\frac{u-v}{r}\right) + \eta \in U^* + V^*$. Now, we consider the case when $\lambda = 0$. Then, we define the functionals $\widetilde{\Psi} : U \to (-\infty,+\infty]$ and $h : U \to (-\infty,+\infty]$ by

$$\widetilde{\Psi}(\bar{v})=\begin{cases}\Psi(\bar{v}) & \text{if } \bar{v}\in V\cap U\\ +\infty & \text{otherwise.}\end{cases}\quad\text{and}\quad h(\bar{v})=\begin{cases}\langle\eta,\bar{v}\rangle_{V^*\times V} & \text{if } \bar{v}\in V\cap U\\ +\infty & \text{otherwise.}\end{cases}$$

It can be shown that $\widetilde{\Psi} + h$ is proper, convex, and lower semicontinuous on U. The first two properties are readily seen. For the lower semicontinuity, we make use of the equivalent characterization of the lower semicontinuity which states that all sublevel sets are closed in the strong topology of U. Thus, let $\alpha \in \mathbb{R}$ and $(u_n)_{n\in\mathbb{N}} \subset J_\alpha := \{\tilde{v} \in U : \widetilde{\Psi}(\tilde{v}) + h(\tilde{v}) \leq \alpha\}$ such that $u_n \to u \in U$ as $n \to \infty$. We want to show that $u \in J_\alpha$. From the definition of $\widetilde{\Psi}$ and h, there holds $J_\alpha = \{\tilde{v} \in U : \Psi(\tilde{v}) + \langle \eta, \tilde{v}\rangle_{V^*\times V} \leq \alpha\}$ and that $(u_n)_{n\in\mathbb{N}}$ is bounded in V by the coercivity of Ψ on V. Hence, there exists a weakly convergent subsequence (labeled as before) such that $u_n \rightharpoonup \tilde{u} \in V$ as $n \to \infty$. Therefore, $(u_n)_{n\in\mathbb{N}}$ is bounded in $U \cap V$ and from the reflexivity of $U \cap V$, we can extract a further weakly convergent subsequence (labeled as before) such that $u_n \rightharpoonup \hat{u} \in U \cap V$ as $n \to \infty$. We obtain,

$$\begin{aligned}\langle f, \hat{u}\rangle_{(U^*+V^*)\times(U\cap V)} &= \langle f_1, \hat{u}\rangle_{U^*\times U} + \langle f_2, \hat{u}\rangle_{V^*\times V}\\ &= \lim_{n\to\infty} \left(\langle f_1, u_n\rangle_{U^*\times U} + \langle f_2, u_n\rangle_{V^*\times V}\right)\\ &= \langle f_1, u\rangle_{U^*\times U} + \langle f_2, \tilde{u}\rangle_{V^*\times V}\end{aligned}$$

for all $f = f_1 + f_2 \in U^* + V^*$ and in particular for all $f \in U^*$ and $f \in V^*$ whence $u = \bar{u} = \hat{u}$ in $U \cap V$. From the weak lower semicontinuity of Ψ_1 on V, we obtain

$$\Psi(u) + \langle \eta, u\rangle_{V^*\times V} \leq \liminf_{n\to\infty} \left(\Psi(u_n) + \langle \eta, u\rangle_{V^*\times V}\right) \leq \alpha,$$

and thus, $u \in J_\alpha$, from which the lower semicontinuity on U follows. Noting that

$$\begin{aligned}\min_{\tilde{v}\in U\cap V} \Phi(r,t,v,w,\eta;\tilde{v}) &= \min_{\tilde{v}\in U\cap V} \widetilde{\Phi}(r,t,v,w,\eta;\tilde{v})\\ &= \min_{\tilde{v}\in U} \left(\frac{1}{2r^2}|\tilde{v} - 2v + w|^2 + r\widetilde{\Psi}\left(\frac{\tilde{v}-v}{r}\right) + \mathcal{E}_{t+r}(u) + h(\tilde{v})\right),\end{aligned}$$

we obtain again by Fermat's theorem

$$\begin{aligned}0 &\in \partial_U \widetilde{\Phi}(r,t,v,w,\eta;u)\\ &= \partial_U \left(\frac{1}{2r^2}|u - 2v + w|^2 + r\widetilde{\Psi}\left(\frac{u-v}{r}\right) + \mathcal{E}_{t+r}(u) + h(u)\right)\end{aligned}$$

for any global minimizer $u \in J_{r,t}(v,w;\eta)$. In order to decompose the elements of the subdifferential of the sum of the functionals in terms of the subgradients of each functional, we employ Lemma 2.2.7 and note that with Remark 5.1.2 *i)* all assumptions of that lemma are satisfied. Hence, there holds

$$\begin{aligned}&\partial_U \left(\frac{1}{2r^2}|u - 2v + w|^2 + r\widetilde{\Psi}\left(\frac{u-v}{r}\right) + \mathcal{E}_{t+r}(u) + h(u)\right)\\ &= \partial_U \left(\frac{1}{2r^2}|u - 2v + w|^2\right) + \partial_U \left(r\widetilde{\Psi}\left(\frac{u-v}{r}\right) + h(u)\right) + \partial_U \mathcal{E}_{t+r}(u)\end{aligned}$$

and therefore there exists a subgradient $\xi \in \partial_U \mathcal{E}_{t+r}(u)$ such that

$$-\frac{u - 2v - w}{r^2} - \xi \in \partial_U \left(r\widetilde{\Psi}\left(\frac{u-v}{r}\right) + h(u)\right).$$

Unfortunately, we are not allowed to decompose the right subdifferential further, since the functional h is, in general, not lower semicontinuous on U. However, since the sum is proper, convex, and lower semicontinuous on U, we can make use of the equivalent description of the subdifferential by the inequality

$$r\widetilde{\Psi}\left(\frac{u-v}{r}\right)+h(u)-r\widetilde{\Psi}\left(\frac{\tilde{v}-v}{r}\right)+h(\tilde{v})\leq\left\langle-\frac{u-2v-w}{r^2}-\xi,\tilde{v}-u\right\rangle_{U^*\times U}$$

for all $\tilde{v}\in U$ and in particular

$$\begin{aligned}&r\Psi\left(\frac{u-v}{r}\right)+\langle\eta,u\rangle_{V^*\times V}-r\Psi\left(\frac{\tilde{v}-v}{r}\right)-\langle\eta,\tilde{v}\rangle_{V^*\times V}\\&\leq\left\langle-\frac{u-2v-w}{r^2}-\xi,\tilde{v}-u\right\rangle_{(U^*+V^*)\times(U\cap V)}\end{aligned}$$

for all $\tilde{v}\in U\cap V$, which in turn implies

$$\begin{aligned}-\frac{u-2v-w}{r^2}-\xi&\in\partial_{U\cap V}\left(\Psi\left(\frac{u-v}{r}\right)+\langle\eta,u\rangle_{V^*\times V}\right)\\&=\mathrm{D}_G\Psi\left(\frac{u-v}{r}\right)+\eta\in U^*+V^*,\end{aligned}$$

which means that we can decompose the elements of the subdifferential in the weaker space U^*+V^*. We finally obtain

$$-\frac{u-2v-w}{r^2}+\mathrm{D}_G\Psi\left(\frac{u-v}{r}\right)+\xi+\eta\quad\text{in } U^*+V^*,$$

and hence the completion of the proof.

□

5.3 Discrete Energy-Dissipation inequality and a priori estimates

Since the previous lemma ensures the solvability of the approximation scheme (5.2.2), we are now able to define piecewise linear and constant interpolations which will interpolate the values $(U_\tau^n)_{n=0}^N$ and $(V_\tau^n)_{n=0}^N$ for every $\tau>0$, respectively, and we will derive a priori estimates for them. The interpolations shall approximate the desired solution to (5.0.1) and its derivative, and are therefore also referred to as approximate solutions to (5.0.1). In order to define the approximate solutions, we assume for the moment that $u_0\in D\cap V$ and $v_0\in V$. In the main existence proof, we will then approximate the initial values from $D\cap V_\lambda$ and H by sequences from $V\cap V$ and V, respectively. For $\tau>0$, let $(U_\tau^n)_{n=1}^N\subset D\cap V$ be the sequence of approximate values obtained from the variational approximation scheme (5.2.2) for $U_\tau^0:=u_0$ and $V_\tau^0:=v_0$. Moreover, let $(\xi_\tau^n)_{n=1}^N\subset V_\lambda^*$ be a sequence of subgradients of the energy determined by the preceding lemma and satisfying $\xi_\tau^n\in\partial_{V_\lambda}\mathcal{E}_{t_n}(U_\tau^n)$, $i=1,\dots,N$ and (5.2.1). The piecewise constant and linear interpolations are defined by

$$\begin{aligned}&\overline{U}_\tau(0)=\underline{U}_\tau(0)=\widehat{U}_\tau(0):=U_\tau^0=u_0\ \text{ and}\\&\underline{U}_\tau(t):=U_\tau^{n-1},\quad\widehat{U}_\tau(t):=\frac{t_n-t}{\tau}U_\tau^{n-1}+\frac{t-t_{n-1}}{\tau}U_\tau^n\quad\text{for } t\in[t_{n-1},t_n),\qquad(5.3.1)\\&\overline{U}_\tau(t):=U_\tau^n\quad\text{for } t\in(t_{n-1},t_n]\quad\text{and}\quad\underline{U}_\tau(T)=U_\tau^N,\ n=1,\dots,N,\end{aligned}$$

as well as

$$\begin{aligned}
&\overline{V}_\tau(0) = \underline{V}_\tau(0) = \widehat{V}_\tau(0) := V_\tau^0 = v_0 \text{ and}\\
&\underline{V}_\tau(t) := V_\tau^{n-1}, \quad \widehat{V}_\tau(t) := \frac{t_n - t}{\tau} V_\tau^{n-1} + \frac{t - t_{n-1}}{\tau} V_\tau^n \quad \text{for } t \in [t_{n-1}, t_n), \qquad (5.3.2)\\
&\overline{V}_\tau(t) := V_\tau^n \quad \text{for } t \in (t_{n-1}, t_n] \quad \text{and} \quad \underline{V}_\tau(T) = V_\tau^N, \;\; n = 1, \ldots, N,
\end{aligned}$$

where $V_\tau^n = \frac{U_\tau^n - U_\tau^{n-1}}{\tau}$ for $n = 1, \ldots, N$. We note that $\widehat{U}_\tau' = \overline{V}_\tau$ in the weak sense. Furthermore, we define the functions $\xi_\tau : [0,T] \to V_\lambda^*$ and $f_\tau : [0,T] \to H$ by

$$\begin{aligned}
&\xi_\tau(t) = \xi_\tau^n, \quad f_\tau(t) = f_\tau^n = \frac{1}{\tau}\int_{t_{n-1}}^{t_n} f(\sigma)\,\mathrm{d}\sigma \qquad \text{for } t \in (t_{n-1}, t_n],\, n = 1, \ldots, N,\\
& \hspace{30em} (5.3.3)\\
&\xi_\tau(T) = \xi_\tau^N \quad \text{and} \quad f_\tau(T) = f_\tau^N.
\end{aligned}$$

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

$$\begin{aligned}
&\bar{\mathbf{t}}_\tau(0) := 0 \quad \text{and} \quad \bar{\mathbf{t}}_\tau(t) := t_n \quad \text{for } t \in (t_{n-1}, t_n],\\
&\underline{\mathbf{t}}_\tau(T) := T \quad \text{and} \quad \underline{\mathbf{t}}_\tau(t) := t_n \quad \text{for } t \in [t_{n-1}, t_n), \quad n = 1, \ldots, N.
\end{aligned} \qquad (5.3.4)$$

Obviously, there holds $\bar{\mathbf{t}}_\tau(t) \to t$ and $\underline{\mathbf{t}}_\tau(t) \to t$ as $\tau \to 0$.

At last, we are in a position to show useful a priori estimates.

Lemma 5.3.1 (A priori estimates) *Let the system LDS* $(U, V, W, W, H, \mathcal{E}, \Psi, B, f)$ *be given and satisfy the Assumptions* (5.E), (5.Ψ), (5.B) *as well as Assumption* (5.f). *Furthermore, let* $\overline{U}_\tau, \underline{U}_\tau, \widehat{U}_\tau, \overline{V}_\tau, \underline{V}_\tau, \widehat{V}_\tau, \xi_\tau$ *and* f_τ *be the interpolations defined in* (5.3.1)-(5.3.3) *associated with the given initial values* $u_0 \in D \cap V, v_0 \in V$ *and the step size* $\tau > 0$. *Then, the discrete energy-dissipation inequality*

$$\begin{aligned}
&\frac{1}{2}\left|\overline{V}_\tau(t)\right|^2 + \mathcal{E}_{\bar{\mathbf{t}}_\tau(t)}(\overline{U}_\tau(t)) + \int_{\bar{\mathbf{t}}_\tau(s)}^{\bar{\mathbf{t}}_\tau(t)} \Big(\Psi(\overline{V}_\tau(r)) + \Psi^*\left(S_\tau(r) - \widehat{V}_\tau'(r) - \xi_\tau(r)\right)\Big)\,\mathrm{d}r\\
&\leq \frac{1}{2}\left|\overline{V}_\tau(s)\right|^2 + \mathcal{E}_{\bar{\mathbf{t}}_\tau(s)}(\overline{U}_\tau(s)) + \int_{\bar{\mathbf{t}}_\tau(s)}^{\bar{\mathbf{t}}_\tau(t)} \partial_r\mathcal{E}_r(\underline{U}_\tau(r))\,\mathrm{d}r + \int_{\bar{\mathbf{t}}_\tau(s)}^{\bar{\mathbf{t}}_\tau(t)} \langle S_\tau(r), \overline{V}_\tau(r)\rangle_{V^*\times V}\,\mathrm{d}r\\
&\quad + \tau\lambda \int_{\bar{\mathbf{t}}_\tau(s)}^{\bar{\mathbf{t}}_\tau(t)} \|\overline{V}_\tau(r)\|_V^2\,\mathrm{d}r
\end{aligned} \qquad (5.3.5)$$

holds for all $0 \leq s < t \leq T$, *where* $S_\tau(r) := f_\tau(r) - B(\bar{\mathbf{t}}_\tau(t), \underline{U}_\tau(t), \underline{V}_\tau(t)), r \in [0,T]$. *Moreover, there exist positive constants* $M, \tau^* > 0$ *such that the estimates*

$$\sup_{t\in[0,T]} \left|\overline{V}_\tau(t)\right| \leq M, \qquad \sup_{t\in[0,T]} \mathcal{E}_t(\overline{U}_\tau(t)) \leq M, \qquad \sup_{t\in[0,T]} |\partial_t\mathcal{E}_t(\underline{U}_\tau(t))| \leq M, \qquad (5.3.6)$$

$$\int_0^T \left(\Psi\left(\overline{V}_\tau(r)\right) + \Psi^*\left(S_\tau(r) - \widehat{V}_\tau'(r) - \xi_\tau(r)\right)\right)\mathrm{d}r \leq M \qquad (5.3.7)$$

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

$$(\overline{U}_\tau)_{0<\tau\leq\tau^*} \subset \mathrm{L}^\infty(0,T;U), \tag{5.3.8a}$$

$$(\xi_\tau)_{0<\tau\leq\tau^*} \subset \mathrm{L}^\infty(0,T;V_\lambda^*), \tag{5.3.8b}$$

in Case ***(a)***

$$(\overline{V}_\tau)_{0<\tau\leq\tau^*} \subset \mathrm{L}^2(0,T;V) \cap \mathrm{L}^\infty(0,T;H), \tag{5.3.8c}$$

$$(\widehat{V}_\tau')_{0<\tau\leq\tau^*} \subset \mathrm{L}^2(0,T;U^*+V^*), \tag{5.3.8d}$$

$$(B_\tau)_{0<\tau\leq\tau^*} \subset \mathrm{L}^{\frac{2}{\nu}}(0,T;V^*), \tag{5.3.8e}$$

in Case ***(b)***

$$(\overline{V}_\tau)_{0<\tau\leq\tau^*} \subset \mathrm{L}^2(0,T;V) \cap \mathrm{L}^q(0,T;W) \cap \mathrm{L}^\infty(0,T;H), \tag{5.3.8f}$$

$$(\widehat{V}_\tau')_{0<\tau\leq\tau^*} \subset \mathrm{L}^{\min\{2,q\}}(0,T;U^*+V^*), \tag{5.3.8g}$$

$$(B_\tau)_{0<\tau\leq\tau^*} \subset \mathrm{L}^{\frac{2}{\nu}}(0,T;V^*) + \mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*), \tag{5.3.8h}$$

are uniformly bounded with respect to τ *in the respective spaces, where* $B_\tau(t) := B(\overline{\mathbf{t}}_\tau(t), \underline{U}_\tau(t), \underline{V}_\tau(t))$, $t \in [0,T]$, $q^* > 0$ *is the conjugate exponent of* $q > 1$, *and* $\nu \in (0,1)$ *stemming from Assumption* (5.Bb). *Finally, there holds*

$$\begin{aligned} \sup_{t\in[0,T]} \Big(\|\underline{U}_\tau(t) - \overline{U}_\tau(t)\|_V + \|\widehat{U}_\tau(t) - \overline{U}_\tau(t)\|_V \Big) &\to 0 \\ \sup_{t\in[0,T]} \Big(\|\underline{V}_\tau(t) - \overline{V}_\tau(t)\|_{U^*+V^*} + \|\widehat{V}_\tau(t) - \overline{V}_\tau(t)\|_{U^*+V^*} \Big) &\to 0 \end{aligned} \tag{5.3.9}$$

as $\tau \to 0$.

Proof. Let $(U_\tau^n)_{n=1}^N \subset D \cap V$ be the approximative values obtained from the variational approximation scheme (5.2.2) and let $(\xi_\tau^n)_{n=1}^N \subset U^* + V^*$ be the associated subgradients. Then, by Lemma 2.2.5, the approximate value U_τ^n solves the Euler–Lagrange equation (5.2.1), i.e.,

$$S_\tau^n - \frac{V_\tau^n - V_\tau^{n-1}}{\tau} - \xi_\tau^n \in \partial_{V\cap U}\Psi(V_\tau^n) = \{D_G\Psi(V_\tau^n)\} \quad \text{and} \quad \xi_\tau^n \in \partial_{V_\lambda}\mathcal{E}_{t_n}(U_\tau^n), \tag{5.3.10}$$

where $S_\tau^n := f_\tau^n - B(t_n, U_\tau^{n-1}, V_\tau^{n-1})$. Due to Lemma 2.3.1, the first inclusion is equivalent to

$$\Psi(V_\tau^n) + \Psi^*\left(S_\tau^n - \frac{V_\tau^n - V_\tau^{n-1}}{\tau} - \xi_\tau^n \right) = \left\langle S_\tau^n - \frac{V_\tau^n - V_\tau^{n-1}}{\tau} - \xi_\tau^n, V_\tau^n \right\rangle_{V^*\times V}$$

and the second one implies

$$\begin{aligned} -\left\langle \xi_\tau^n, U_\tau^n - U_\tau^{n-1} \right\rangle_{V_\lambda^*\times V_\lambda} &\leq \mathcal{E}_{t_n}(U_\tau^{n-1}) - \mathcal{E}_{t_n}(U_\tau^n) + \lambda\|U_\tau^n - U_\tau^{n-1}\|_V^2 \\ &= \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) - \mathcal{E}_{t_n}(U_\tau^n) + \int_{t_{n-1}}^{t_n} \partial_r E_r(U_\tau^{n-1})\,\mathrm{d}r \\ &\quad + \lambda\|U_\tau^n - U_\tau^{n-1}\|_V^2 \end{aligned}$$

for all $n = 1, \dots, N$. Using the identity

$$(u - v, u) = \frac{1}{2}\left(|u|^2 - |v|^2 + |u - v|^2\right) \quad \text{for all } u, v \in H \tag{5.3.11}$$

and the fact that $\langle w, v\rangle_{V^*\times V} = (w, v)$ for $v \in V$ and $w \in H$, we obtain

$$\begin{aligned}
&\frac{1}{2}|V_\tau^n|^2 + \mathcal{E}_{t_n}(U_\tau^n) + \tau\Psi(V_\tau^n) + \tau\Psi^*\left(S_\tau^n - \frac{V_\tau^n - V_\tau^{n-1}}{\tau} - \xi_\tau^n\right) - \tau\left\langle S_\tau^n, V_\tau^n\right\rangle_{V^*\times V} \\
&\le \frac{1}{2}|V_\tau^{n-1}|^2 + \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) + \int_{t_{n-1}}^{t_n} \partial_r\mathcal{E}_r(U_\tau^{n-1})\,\mathrm{d}r + \lambda\|U_\tau^n - U_\tau^{n-1}\|_V^2
\end{aligned} \tag{5.3.12}$$

for all $n = 1, \dots, N$, which, by summing up the inequalities, implies (5.3.5). In order to show the bounds (5.3.6) and (5.3.7), we make use of the following estimates: first, from Assumption (5.Bb) and the Fenchel–Young inequality, we obtain

$$\begin{aligned}
\tau\langle S_\tau^n, V_\tau^n\rangle &= \langle -B(t_n, U_\tau^{n-1}, V_\tau^{n-1}) + f_\tau^n, V_\tau^n\rangle_{V^*\times V} \\
&= \langle -B(t_n, U_\tau^{n-1}, V_\tau^{n-1}), V_\tau^n\rangle_{V^*\times V} + \langle f_\tau^n, V_\tau^n\rangle_{V^*\times V} \\
&\le c\tau\Psi(V_\tau^n) + c\tau\Psi^*\left(\frac{-B(t_n, U_\tau^{n-1}, V_\tau^{n-1})}{c}\right) + \frac{\tau}{2}(|f_\tau^n|^2 + |V_\tau^n|^2) \\
&\le c\tau\Psi(V_\tau^n) + \tau\beta(1 + \mathcal{E}_{t_n}(U_\tau^{n-1}) + |V_\tau^{n-1}|^2 + \Psi(V_\tau^{n-1})^\nu) \\
&+ \frac{\tau}{2}(|f_\tau^n|^2 + |V_\tau^n|^2), \\
&\le c\tau\Psi(V_\tau^n) + \tau\beta(1 + \mathcal{E}_{t_n}(U_\tau^{n-1}) + |V_\tau^{n-1}|^2) + \tau\varepsilon\Psi(V_\tau^{n-1}) + \tau C \\
&+ \frac{\tau}{2}(|f_\tau^n|^2 + |V_\tau^n|^2),
\end{aligned}$$

for positive constants $\varepsilon, C = C(\varepsilon, \beta) > 0$ such that $\varepsilon < \frac{1-c}{2}$ and $C = \frac{\beta^{\frac{1}{1-\nu}}}{\varepsilon^{\frac{\nu}{1-\nu}}}$. Second, by the strong positivity of Ψ_1 and the growth condition for Ψ_2, we have in Case **(b)**

$$\mu\|U_\tau^n - U_\tau^{n-1}\|_V^2 = \mu\tau^2\|V_\tau^n\|_V^2 \le \tau^2\Psi_1\left(V_\tau^n\right) + \tau^2\Psi_2\left(V_\tau^n\right) + \tau^2\tilde{c} = \tau^2\Psi\left(V_\tau^n\right) + \tau^2\tilde{c}$$

where $\tilde{c} > 0$ is from Condition 5.1.2. In the Case **(a)**, we only employ the strong positivity of Ψ obtaining

$$\mu\|U_\tau^n - U_\tau^{n-1}\|_V^2 = \mu\tau^2\|V_\tau^n\|_V^2 \le \tau^2\Psi\left(V_\tau^n\right).$$

Finally, we use the estimate following from Condition (5.Ed),

$$\int_{t_{n-1}}^{t_n} \partial_r\mathcal{E}_r(U_\tau^{n-1})\,\mathrm{d}r \le \int_{t_{n-1}}^{t_n} C_1\mathcal{E}_r(U_\tau^{n-1})\,\mathrm{d}r \le C_1\int_{t_{n-1}}^{t_n} \mathcal{G}(U_\tau^{n-1})\,\mathrm{d}r.$$

Inserting all preceding inequalities in (5.3.12) and summing up all inequalities from 1 to n, we find a positive constant $C > 0$ such that

$$\begin{aligned}
&\frac{1}{2}|V_\tau^n|^2 + \frac{1}{C_1}\mathcal{G}(U_\tau^n) + \int_0^{t_n}\left((1 - \alpha(\tau))\Psi(\overline{V}_\tau(r)) + \Psi^*\left(S_\tau(r) - \widehat{V}_\tau'(r) - \xi_\tau(r)\right)\right)\mathrm{d}r \\
&\le C\left(|v_0|^2 + \mathcal{E}_0(u_0) + T + \|f\|_{\mathrm{L}^2(0,T;H)}^2 + \Psi(v_0)\right) \\
&+ C\int_0^{t_n}\left(|\overline{V}_\tau(r)|^2 + \mathcal{G}(\overline{U}_\tau(r))\right)\mathrm{d}r,
\end{aligned} \tag{5.3.13}$$

where $\alpha(\tau) := c + \tilde{c} + \tau\frac{\lambda}{\mu} < 1$ for all $\tau < \tau^* := \min\{\frac{\mu}{\lambda}(1 - c - \tilde{c}), 1\}$ and $\alpha(\tau)$ is decreasing for decreasing τ. In the step (5.3.13), we made use of the estimate for the interpolation f_τ

$$\begin{aligned}\|f_\tau\|^2_{\mathrm{L}^2(0,T;H)} &= \sum_{k=1}^{n} \tau |f^k_\tau|^2 \\ &= \sum_{k=1}^{n} \frac{1}{\tau} \Big| \int_{t_{k-1}}^{t_k} f(\sigma)\,\mathrm{d}\sigma \Big|^2 \\ &\le \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} |f(\sigma)|^2\,\mathrm{d}\sigma = \int_0^{t_n} |f(\sigma)|^2\,\mathrm{d}\sigma \le \|f\|^2_{\mathrm{L}^2(0,T;H)}. \end{aligned} \tag{5.3.14}$$

Then, by the discrete version of Gronwall's lemma (Lemma A.1.2), there exists a constant $M > 0$ such that (5.3.6) and (5.3.7) are satisfied. Now, we seek to show the bounds in (5.3.8) by distinguishing the two cases **(a)** and **(b)**.

Ad **(a)**. Due to the coercivity of Ψ and Ψ^*, the uniform boundedness of $(\overline{V}_\tau)_{0<\tau\le\tau^*} \subset \mathrm{L}^2(0,T;V)$ and $(S_\tau - \widehat{V}'_\tau - \xi_\tau)_{0<\tau\le\tau^*} \subset \mathrm{L}^2(0,T;V^*) \subset \mathrm{L}^2(0,T;V^*)$ follow immediately from the a priori estimate (5.3.7). The boundedness of $(B_\tau)_{0<\tau\le\tau^*} \subset \mathrm{L}^2(0,T;V^*)$ uniformly in τ is a consequence of Assumption (5.Bb) and the coercivity of Ψ^*: there holds

$$\begin{aligned}\bar{c}\int_0^T \|B_\tau(r)\|_*^{\frac{2}{\nu}}\,\mathrm{d}r &\le \int_0^T \Psi^*\left(B(\bar{\mathbf{t}}_\tau(r), \underline{U}_\tau(r), \underline{V}_\tau(r))\right)^{\frac{1}{\nu}}\,\mathrm{d}r \\ &\le \int_0^T c\Psi^*\left(\frac{B(\bar{\mathbf{t}}_\tau(r), \underline{U}_\tau(r), \underline{V}_\tau(r))}{c}\right)^{\frac{1}{\nu}}\,\mathrm{d}r \\ &\le \int_0^T \Big(C((1 + \mathcal{E}_{\underline{\mathbf{t}}_\tau(r)}(\underline{U}_\tau(r))^{\frac{1}{\nu}} + |\underline{V}_\tau(r)|^2)^{\frac{1}{\nu}} + \Psi\left(\underline{V}_\tau(r)\right)\Big)\,\mathrm{d}r \\ &\le N \end{aligned} \tag{5.3.15}$$

for positive constants $C, N > 0$ independent of τ, where $c \in (0,1)$ is from Assumption (5.Bb) and where we have used the fact that for all $\zeta \in V^*$ the mapping $r \mapsto r\Psi^*(\zeta/r)$ is monotonically decreasing on $(0, +\infty)$ which follows from the convexity of Ψ^* and $\Psi^*(0) = 0$. Since $(f_\tau)_{0<\tau\le\tau^*}$ is uniformly bounded in $\mathrm{L}^2(0,T;H)$, we infer that $(\widehat{V}'_\tau + \xi_\tau)_{0<\tau\le\tau^*}$ is uniformly bounded in $\mathrm{L}^2(0,T;V^*)$ with respect to τ as well. Finally, Assumption (5.Eg) implies a uniform bound for $(\xi_\tau)_{0<\tau\le\tau^*}$ in $\mathrm{L}^\infty(0,T;V^*_\lambda)$. Since all previous families of functions are bound in the common space $\mathrm{L}^2(0,T;U^* + V^*)$, we deduce that $(\widehat{V}'_\tau)_{0<\tau\le\tau^*}$ is uniformly bounded in $\mathrm{L}^\infty(0,T;U^* + V^*)$ with respect to τ.

Ad **(b)**. Again, the coercivity of the dissipation potential Ψ leads to the boundedness of the sequence of discrete derivatives $(\overline{V}_\tau)$ in $\mathrm{L}^2(0,T;V) \cap \mathrm{L}^q(0,T;W)$ uniformly in $\tau \in (0,\tau^*)$. In order to show that $(S_\tau - \widehat{V}'_\tau - \xi_\tau)_{0<\tau\le\tau^*} \subset \mathrm{L}^2(0,T;V^*) + \mathrm{L}^{q^*}(0,T;W^*)$ is uniformly bounded with respect to τ, we make the following observation: let $\zeta : [0,T] \to V^*$ be any measurable function such that

$$\int_0^T \Psi^*(\zeta(t))\,\mathrm{d}t \le M.$$

We want to show that there exists a positive constant $\widetilde{M} > 0$ such that

$$\tilde{M} \ge \|\zeta\|_{\mathrm{L}^2(0,T;V^*) + \mathrm{L}^{q^*}(0,T;W^*)}.$$

First, by the formula (2.3.4) from Lemma 2.3.5 and the growth conditions for the conjugate in Remark 5.1.1, there exists a constant $M_1 > 0$ such that

$$M_1 \geq \int_0^T \min_{\eta \in W^*} \left(\|\zeta(t) - \eta\|_{V^*}^2 + \|\eta\|_{W^*}^{q^*} \right) \mathrm{d}t.$$

Second, the mapping $t \mapsto \alpha(t) := \min_{\eta \in W^*} \left(\|\zeta(t) - \eta\|_{V^*}^2 + \|\eta\|_{W^*}^{q^*} \right)$ is LEBESGUE measurable: since V^* is separable, there exists a countable dense subset $(\eta_n)_{n \in \mathbb{N}} \subset V^*$. Then, there holds $\alpha(t) = \inf_{n \in \mathbb{N}} \left(\|\zeta(t) - \eta_n\|_{V^*}^2 + \|\eta_n\|_{W^*}^{q^*} \right)$ and from the measurability of the function $\alpha_n(t) := \left(\|\zeta(t) - \eta_n\|_{V^*}^2 + \|\eta_n\|_{W^*}^{q^*} \right)$ for each $n \in \mathbb{N}$, the measurability of α follows. Further, we note that the mapping $g : [0,T] \times W^* \to \mathbb{R}$, $(t,\eta) \mapsto g(t,\eta) = \|\zeta(t) - \eta\|_{V^*}^2 + \|\eta\|_{W^*}^{q^*}$ is a CARATHÉODORY function and therefore, by the Inverse Image Theorem, see, e.g., AUBIN & FRANKOWSKA [19, Theorem 8.2.9, p. 315], the set-valued map

$$H(t) := \{\eta \in W^* : g(t,\eta) = \alpha(t)\}$$

is measurable and there exists a measurable selection $\omega : [0,T] \to W^*$ with $\omega(t) \in H(t)$ and $g(t,\omega(t)) = \alpha(t)$ for all $t \in [0,T]$. We obtain

$$\begin{aligned} M_1 &\geq \int_0^T \min_{\eta \in W^*} \left(\|\zeta(t) - \eta\|_*^2 + \|\eta\|_{W^*}^{q^*} \right) \mathrm{d}t \\ &= \int_0^T \left(\|\zeta(t) - \omega(t)\|_*^2 + \|\omega(t)\|_{W^*}^{q^*} \right) \mathrm{d}t, \end{aligned}$$

whence $\omega \in \mathrm{L}^{q^*}(0,T;W^*)$ and $\zeta - \omega \in \mathrm{L}^2(0,T;V^*)$. It follows

$$\begin{aligned} M_1 &\geq \int_0^T \left(\|\zeta(t) - \omega(t)\|_*^2 + \|\omega(t)\|_{W^*}^{q^*} \right) \mathrm{d}t, \\ &\geq \inf_{\substack{\xi_1 \in \mathrm{L}^2(0,T;V^*), \xi_2 \in \mathrm{L}^{q^*}(0,T;W^*) \\ \zeta = \xi_1 + \xi_2}} \int_0^T \left(\|\xi_1(t)\|_*^2 + \|\xi_2(t)\|_{W^*}^{q^*} \right) \mathrm{d}t \\ &\geq \inf_{\substack{\xi_1 \in \mathrm{L}^2(0,T;V^*), \xi_2 \in \mathrm{L}^{q^*}(0,T;W^*) \\ \zeta = \xi_1 + \xi_2}} \left(\|\xi_1\|_{\mathrm{L}^2(0,T;V^*)} + \|\xi_2\|_{\mathrm{L}^{q^*}(0,T;W^*)} \right) - M_2 \\ &\geq \inf_{\substack{\xi_1 \in \mathrm{L}^2(0,T;V^*), \xi_2 \in \mathrm{L}^{q^*}(0,T;W^*) \\ \zeta = \xi_1 + \xi_2}} \max\{\|\xi_1\|_{\mathrm{L}^2(0,T;V^*)}, \|\xi_2\|_{\mathrm{L}^{q^*}(0,T;W^*)}\} - M_2 \\ &= \|\zeta\|_{\mathrm{L}^2(0,T;V^*) + \mathrm{L}^{q^*}(0,T;W^*)} - M_2 \end{aligned}$$

for a constant $M_2 > 0$ coming from YOUNG's inequality. Since the constant $\widetilde{M} := M_1 + M_2 > 0$ was obtained independently of the function ζ, the uniform bound of the sequence $(S_\tau - \widehat{V}_\tau' - \xi_\tau)_{0 < \tau \leq \tau^*}$ in $\mathrm{L}^2(0,T;V^*) + \mathrm{L}^{q^*}(0,T;W^*)$ follows. Employing Condition (5.Bb) for the perturbation B as in (5.3.15) and noting that

$$\begin{aligned} \int_0^T \left(\min_{\eta \in W^*} \left(\|\zeta(t) - \eta\|_*^2 + \|\eta\|_{W^*}^{q^*} \right) \right)^{\frac{1}{\nu}} \mathrm{d}t &= \int_0^T \min_{\eta \in W^*} \left(\|\zeta(t) - \eta\|_*^2 + \|\eta\|_{W^*}^{q^*} \right)^{\frac{1}{\nu}} \mathrm{d}t \\ &\geq \int_0^T \min_{\eta \in W^*} \left(\|\zeta(t) - \eta\|_*^{\frac{2}{\nu}} + \|\eta\|_{W^*}^{\frac{q^*}{\nu}} \right) \mathrm{d}t, \end{aligned}$$

we obtain the uniform boundedness of $(B_\tau)_{0<\tau\leq\tau^*}$ in $\mathrm{L}^{\frac{2}{\nu}}(0,T;V^*)+\mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*)$ by arguing in the same way as for Case **(a)**. This, together with the uniform bounds of $(f_\tau)_{0<\tau\leq\tau^*}$ in $\mathrm{L}^2(0,T;H)$ and $(\xi_\tau)_{0<\tau\leq\tau^*}$ in $\mathrm{L}^\infty(0,T;U^*+V^*)$ yields the uniform bound of $(\widehat{V}_\tau')_{0<\tau\leq\tau^*}$ in $\mathrm{L}^{\min\{2,q^*\}}(0,T;U^*+V^*)$ with respect to τ. It remains to show the uniform convergences (5.3.9), which follow immediately from the uniform bounds of $(\widehat{V}_\tau')_{0<\tau\leq\tau^*}\subset\mathrm{L}^{\min\{2,q^*\}}(0,T;U^*+V^*)$ and $(\overline{V}_\tau)_{0<\tau\leq\tau^*}\subset\mathrm{L}^2(0,T;V)$ in the respective spaces together with the estimates

$$\|\widehat{U}_\tau(t)-\overline{U}_\tau(t)\|_V\leq\|\underline{U}_\tau(t)-\overline{U}_\tau(t)\|_V=\int_{\underline{\mathbf{t}}(t)}^{\overline{\mathbf{t}}(t)}\|\overline{V}_\tau(r)\|_V\,\mathrm{d}r\quad\text{and}$$
$$\|\widehat{V}_\tau(t)-\overline{V}_\tau(t)\|_{U^*+V^*}\leq\|\underline{V}_\tau(t)-\overline{V}_\tau(t)\|_{U^*+V^*}=\int_{\underline{\mathbf{t}}(t)}^{\overline{\mathbf{t}}(t)}\|\widehat{V}_\tau'(r)\|_{U^*+V^*}\,\mathrm{d}r$$

for all $t\in[0,T]$. □

5.4 Compactness

This section is devoted to the existence of convergent subsequences of the sequence of approximate solutions in some proper Bochner spaces in order to pass to the limit in the discrete inclusion (5.2.1) as the step size vanishes. As we will see, we will indeed obtain in the limit a solution to the Cauchy problem (5.0.1). 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 5.4.1 (Compactness) *Under the assumptions of Lemma 5.3.1, let $(\tau_n)_{n\in\mathbb{N}}$ be a vanishing sequence of positive numbers and let $u_0\in D\cap V$ and $v_0\in V$. Then, there exists a subsequence, still denoted by $(\tau_n)_{n\in\mathbb{N}}$, a pair of functions (u,ξ) with*

$$u\in\mathrm{C}_w([0,T];U)\cap\mathrm{H}^1(0,T;V)\cap\mathrm{W}^{1,\infty}(0,T;H)\text{ and }\xi\in\mathrm{L}^\infty(0,T;U^*+V^*)$$

*that satisfies $u\in\mathrm{H}^2(0,T;U^*+V^*)$ in the case **(a)** and $u\in\mathrm{W}^{1,q}(0,T;W)\cap\mathrm{W}^{2,\min\{2,q^*\}}(0,T;U^*+V^*)$ in the case **(b)** while fulfilling the initial values $u(0)=u_0$ in U and $u'(0)=v_0$ in H such that the following convergences hold*

$$\underline{U}_{\tau_n},\overline{U}_{\tau_n},\widehat{U}_{\tau_n}\overset{*}{\rightharpoonup}u\quad\text{in }\mathrm{L}^\infty(0,T;U\cap V),\tag{5.4.1a}$$
$$\widehat{U}_{\tau_n}(t),\underline{U}_{\tau_n}(t),\overline{U}_{\tau_n}(t)\rightharpoonup u(t)\quad\text{in }U\text{ for all }t\in[0,T],\tag{5.4.1b}$$
$$\underline{U}_{\tau_n}(t)\rightharpoonup u(t)\quad\text{in }V\text{ for all }t\in[0,T],\tag{5.4.1c}$$
$$\underline{U}_{\tau_n}\to u\quad\text{in }\mathrm{L}^r(0,T;\widetilde{W})\quad\text{for any }r\geq1,\tag{5.4.1d}$$
$$\widehat{U}_{\tau_n}(t),\underline{U}_{\tau_n}(t),\overline{U}_{\tau_n}(t)\to u(t)\quad\text{in }\widetilde{W}\text{ for all }t\in[0,T],\tag{5.4.1e}$$
$$\overline{V}_{\tau_n},\underline{V}_{\tau_n}\overset{*}{\rightharpoonup}u'\quad\text{in }\mathrm{L}^2(0,T;V)\cap\mathrm{L}^\infty(0,T;H),\tag{5.4.1f}$$
$$\overline{V}_{\tau_n},\underline{V}_{\tau_n}\to u'\quad\text{in }\mathrm{L}^p(0,T;H)\quad\text{for all }p\geq1,\tag{5.4.1g}$$
$$\overline{V}_{\tau_n}(t),\underline{V}_{\tau_n}(t)\to u'(t)\quad\text{in }H\text{ for a.e. }t\in(0,T),\tag{5.4.1h}$$
$$\underline{V}_{\tau_n}(t),\overline{V}_{\tau_n}(t)\rightharpoonup u'(t)\quad\text{in }H\text{ for all }t\in[0,T],\tag{5.4.1i}$$

$$\xi_{\tau_n} \overset{*}{\rightharpoonup} \xi \quad \textit{in } \mathrm{L}^\infty(0,T;V_\lambda^*), \tag{5.4.1j}$$
$$f_{\tau_n} \to f \quad \textit{in } \mathrm{L}^2(0,T;H), \tag{5.4.1k}$$

*and in Case **(a)***

$$\widehat{V}_{\tau_n}' \rightharpoonup u'' \quad \textit{in } \mathrm{L}^2(0,T;U^*+V^*), \tag{5.4.1l}$$
$$B_{\tau_n} \to B(\cdot,u(\cdot),u'(\cdot)) \quad \textit{in } \mathrm{L}^2(0,T;V^*), \tag{5.4.1m}$$

*and in Case **(b)***

$$\overline{V}_{\tau_n} \rightharpoonup u' \quad \textit{in } \mathrm{L}^q(0,T;W), \tag{5.4.1n}$$
$$\overline{V}_{\tau_n} \to u' \quad \textit{in } \mathrm{L}^{\max\{2,r\}}(0,T;W) \quad \textit{for any } r\in[1,q), \tag{5.4.1o}$$
$$D_G\Psi_2(\overline{V}_{\tau_n}) \to D_G\Psi_2(u') \quad \textit{in } \mathrm{L}^r(0,T;U^*+V^*) \quad \textit{for any } r\in[1,q^*), \tag{5.4.1p}$$
$$\widehat{V}_{\tau_n}' \rightharpoonup u'' \quad \textit{in } \mathrm{L}^{\min\{2,q^*\}}(0,T;U^*+V^*), \tag{5.4.1q}$$
$$B_{\tau_n} \to B(\cdot,u(\cdot),u'(\cdot)) \quad \textit{in } \mathrm{L}^2(0,T;V^*)+\mathrm{L}^{q^*}(0,T;W^*), \tag{5.4.1r}$$

where $B_\tau(t) := B(\overline{\mathfrak{t}}_\tau(t), \underline{U}_\tau(t), \underline{V}_\tau(t)),\ t\in[0,T]$.

Proof. Let $\overline{U}_\tau, \underline{U}_\tau, \widehat{U}_\tau, \overline{V}_\tau, \underline{V}_\tau, \widehat{V}_\tau, \xi_\tau$ as well as f_τ be the interpolations with the initial values $u_0 \in D\cap V$ and $v_0\in V$ as defined in (5.3.1)-(5.3.3). Since all spaces are supposed to be separable and reflexive, we note that if a Banach space X is separable and reflexive, the spaces $\mathrm{L}^p(0,T;X)$ for $1<p<\infty$ are also separable and reflexive, whereas $\mathrm{L}^\infty(0,T;X)$ is the dual of the separable space $\mathrm{L}^1(0,T;X^*)$. So, as a consequence of the Banach–Alaoglu theorem, bounded sets in $\mathrm{L}^p(0,T;X)$, $1<p<\infty$, and $\mathrm{L}^\infty(0,T;X)$ are relatively compact with respect to the weak and weak* topology, respectively. In view of the a priori estimates (5.3.6) and (5.3.7), Assumption (5.Ec) implies that the sequence $(\overline{U}_{\tau_n})_{n\in\mathbb{N}}$ is bounded in $\mathrm{L}^\infty(0,T;U)$. Together with the bounds (5.3.8), this already yields the existence of converging subsequences (denoted as before) fulfilling (5.4.1a), (5.4.1f), and (5.4.1j). We remark that the limit functions can be identified with u and u' by standard arguments. In order to show (5.4.1d), we make use of the Lions–Aubin–Dubinskiĭ lemma (Lemma A.2.1). The boundedness of the sequence of piecewise linear interpolations $(\underline{U}_{\tau_n})_{n\in\mathbb{N}}$ and the discrete derivatives $(\underline{V}_{\tau_n})_{n\in\mathbb{N}}$ uniformly in $\mathrm{L}^\infty(0,T;U)$ and $\mathrm{L}^\infty(0,T;H)$, respectively, yields directly the relative compactness in $\mathrm{L}^r(0,T;\widetilde{W})$ for all $r\geq 1$. In view of (5.4.1b), this implies the convergence (5.4.1e). The just proven convergence is indeed only needed to deduce the strong convergence of the perturbation B, i.e., convergence (5.4.1m) and (5.4.1r). Before showing this convergence, we first proceed with proving the pointwise weak convergence as stated in (5.4.1i). First, we note that from $\widehat{V}_{\tau_n} \in \mathrm{W}^{1,1}(0,T;U^*+V^*) \hookrightarrow \mathrm{C}([0,T];U^*+V^*)$ and (5.4.1l) or (5.4.1q), there holds $\widehat{V}_{\tau_n}(t) \rightharpoonup u'(t)$ in U^*+V^* as $n\to\infty$ for all $t\in[0,T]$. Since $\widehat{V}_{\tau_n}(t)$ is uniformly bounded in H for all $t\in[0,T]$, it is (up to a subsequence) weakly convergent in H to $u'(t)$. Since the weak limit is unique in U^*+V^*, we obtain with the subsequence principle, the convergence of the whole sequence. Together with the strong convergence in (5.3.9), this implies (5.4.1i). With the same argument, we deduce the pointwise weak convergences (5.4.1b) and (5.4.1c) where in the latter convergence we use the fact that $u_0\in D\cap V$. Further, we recall that $\mathrm{L}^\infty(0,T;X)\cap \mathrm{C}_w([0,T];Y) = \mathrm{C}_w([0,T];X)$ for two Banach spaces X and Y with X being reflexive and such that the continuous and dense embedding $X \overset{d}{\hookrightarrow} Y$

holds, see, e.g., in LIONS & MAGENES [107, Lemma 8.1, p. 275]. Applying the latter result to $X = U$ and $Y = H$, there holds $u \in \mathrm{C}_w([0,T];U)$. Now, we seek to apply Lemma A.2.1 to the sequence $(\overline{V}_{\tau_n})_{n\in\mathbb{N}}$ with $X = V$, $B = H$ and $Y = U^* + V^*$ in order to show the strong convergence in $\mathrm{L}^2(0,T;H)$ to the limit $u' \in \mathrm{L}^2(0,T;H)$. The Assumption (A.2.1) of Lemma A.2.1 follows for $p = 2$ and $r = \min\{2, q^*\} > 1$ directly from the a priori estimate (5.3.6) and the following estimate

$$\|\sigma_{\tau_n}\overline{V}_{\tau_n} - \overline{V}_{\tau_n}\|_{\mathrm{L}^q(0,T-\tau_n;U^*+V^*)} = \tau_n\|\widehat{V}'_{\tau_n}\|_{\mathrm{L}^q(0,T;U^*+V^*)} \leq \tau_n M \quad \text{for all } n \in \mathbb{N},$$

from which the strong convergence $\overline{V}_{\tau_n} \to u'$ in $\mathrm{L}^2(0,T;H)$ as $n \to \infty$ follows. Taking into account the boundedness of the very same sequence in $\mathrm{L}^\infty(0,T;H)$, we obtain by a well-known interpolation inequality the strong convergence in $\mathrm{L}^r(0,T;H)$ for all $r \geq 1$, i.e. (5.4.1g). This, in turn, implies pointwise convergence of the very sequence almost everywhere in $(0,T)$, i.e., (5.4.1h). The assertion for $\widehat{V}_{\tau_n}$ can be shown analogously. Recalling the fact that the space of continuous functions $\mathrm{C}([0,T];H)$ is dense in $\mathrm{L}^2(0,T;H)$, for every $\epsilon > 0$ there exists a function $f^\varepsilon \in \mathrm{C}([0,T];H)$ such that $\|f^\varepsilon - f\|_{\mathrm{L}^2(0,T;H)} < \varepsilon/3$. In view of this approximation property and defining $f^\varepsilon_{\tau_n}(t) = \frac{1}{\tau_n}\int_{t_{n_1}}^{n} f^\varepsilon(\sigma)\,\mathrm{d}\sigma$, $t \in [t_{n-1}, t_n)$, $n = 1, \dots, N$, we find

$$\begin{aligned}\|f_{\tau_n} - f\|_{\mathrm{L}^2(0,T;H)} &\leq \|f_{\tau_n} - f^\varepsilon_{\tau_n}\|_{\mathrm{L}^2(0,T;H)} + \|f^\varepsilon_{\tau_n} - f^\varepsilon\|_{\mathrm{L}^2(0,T;H)} + \|f^\varepsilon - f\|_{\mathrm{L}^2(0,T;H)} \\ &\leq \|f - f^\varepsilon\|_{\mathrm{L}^2(0,T;H)} + \|f^\varepsilon_{\tau_n} - f^\varepsilon\|_{\mathrm{L}^2(0,T;H)} + \|f^\varepsilon - f\|_{\mathrm{L}^2(0,T;H)} \\ &\leq \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon\end{aligned}$$

for sufficiently small step sizes τ_n, where we also used the estimate (5.3.14) for the first term, and where we made the second term smaller than $\varepsilon/3$ for sufficiently small step sizes which follows from the uniform continuity of f^ε. We proceed by showing the convergences which differ from each other in Case **(a)** and in Case **(b)**.
Ad case **(a)**. The weak convergence $\widehat{V}'_{\tau_n} \rightharpoonup u''$ as $n \to \infty$ in $\mathrm{L}^2(0,T;U^* + V^*)$ follows immediately from the reflexivity of the space $\mathrm{L}^2(0,T;U^* + V^*)$ and the uniform bound of the sequence $(\widehat{V}'_{\tau_n})_{n\in\mathbb{N}}$ in the very same space with respect to $n \in \mathbb{N}$. Further, we denote by $\mathcal{B}(u)(t) = B(t, u(t), u'(t))$, $t \in [0,T]$, the associated NEMITSKIĬ operator and recall that $B_{\tau_n}(t) = B(\bar{\mathbf{t}}_{\tau_n}(t), \underline{U}_{\tau_n}(t), \underline{V}_{\tau_n}(t))$, $t \in [0,T]$. In order to show the strong convergence of the perturbation, we first note that from the uniform convergence (5.4.1e) and the pointwise convergence (5.4.1h) together with the continuity condition (5.Ba) implies

$$\|\mathcal{B}_{\tau_n}(t) - \mathcal{B}(u)(t)\|V^* \to 0 \quad \text{a.e. in } (0,T) \tag{5.4.2}$$

as $n \to \infty$. By the growth condition (5.Bb), we also obtain $\mathcal{B}(u) \in \mathrm{L}^{\frac{2}{\nu}}(0,T;V^*)$ so that we have $B_{\tau_n} - \mathcal{B}(u) \in \mathrm{L}^{\frac{2}{\nu}}(0,T;V^*)$ being uniformly bounded with respect to $n \in \mathbb{N}$. Using EGOROV's theorem, it is easy to deduce the strong convergence of $B_{\tau_n} \to \mathcal{B}(u)$ in $\mathrm{L}^q(0,T;V^*)$ as $n \to \infty$ for all $0 < q < \frac{2}{\nu}$, and since $2 < \frac{2}{\nu}$, we can choose $q = 2$, i.e., (5.4.1m).
Ad case **(b)**. From the boundedness of the sequences of $(\overline{V}_{\tau_n})_{n\in\mathbb{N}}$ and $(\widehat{V}'_{\tau_n})_{n\in\mathbb{N}}$ in $\mathrm{L}^r(0,T;W)$ and $\mathrm{L}^{\min\{2,q^*\}}(0,T;U^* + V^*)$, respectively, we obtain the weak convergences (5.4.1n) and (5.4.1q). Applying again Lemma A.2.1 to the sequence $(\overline{V}_{\tau_n})_{n\in\mathbb{N}}$ with the choices $X = V$, $B = W$ and $Y = U^*$ with $p = 2$ and $r = 1$ yields compactness

of the sequence in $\mathrm{L}^2(0,T;W)$, and if $q>2$, we obtain compactness of the sequence in every intermediate space $\mathrm{L}^s(0,T;W)$ with $2\le s<q$ between $\mathrm{L}^2(0,T;W)$ and $\mathrm{L}^q(0,T;W)$ by an interpolation inequality, and hence (5.4.1o). With the same reasoning as for the perturbation, the latter convergence yields (5.4.1p) employing EGOROV's theorem and the growth and continuity condition for $D_G\Psi_2$ on W. The strong convergence of the perturbation in the space $\mathrm{L}^2(0,T;V^*)+\mathrm{L}^{q^*}(0,T;W^*)$ is more delicate and is established as follows: first, by the a priori estimate (5.3.8h), the sequence B_{τ_n} is bounded in $\mathrm{L}^{\frac{2}{\nu}}(0,T;V^*)+\mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*)\subset\mathrm{L}^2(0,T;V^*)+\mathrm{L}^{q^*}(0,T;W^*)$ by a constant denoted by $\tilde{M}>0$. With the same reasoning as for the first case, we obtain the convergence (5.4.2) and $\mathcal{B}(u)\in\mathrm{L}^{\frac{2}{\nu}}(0,T;V^*)+\mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*)$. We choose the constant $\tilde{M}>0$ such that $\tilde{M}\ge\|\mathcal{B}(u)\|_{\mathrm{L}^{\frac{2}{\nu}}(0,T;V^*)+\mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*)}$. Second, defining the set

$$G_n:=\{\eta\in\mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*):\|B_{\tau_n}-\mathcal{B}(u)-\eta\|_{\mathrm{L}^{\frac{2}{\nu}}(0,T;V^*)}\le 2\tilde{M},\\ \|\eta\|_{\mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*)}\le 2\tilde{M}\},$$

there holds

$$\begin{aligned}2\tilde{M}\ge&\|B_{\tau_n}-\mathcal{B}(u)\|_{\mathrm{L}^{\frac{2}{\nu}}(0,T;V^*)+\mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*)}\\ &=\inf_{\eta\in\mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*)}\max\{\|B_{\tau_n}-\mathcal{B}(u)-\eta\|_{\mathrm{L}^{\frac{2}{\nu}}(0,T;V^*)},\|\eta\|_{\mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*)}\}\\ &=\inf_{\eta\in G_n}\max\{\|B_{\tau_n}-\mathcal{B}(u)-\eta\|_{\mathrm{L}^{\frac{2}{\nu}}(0,T;V^*)},\|\eta\|_{\mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*)}\}\quad\text{for all } n\in\mathbb{N},\end{aligned}$$

which restricts the set of functions where the infimum is taken over. Then, by EGOROV's theorem, for every $\varepsilon>0$ there exists a subset $E\subset[0,T]$ with measure $\mu(E)<\varepsilon$ such that the uniform convergence

$$\lim_{n\to\infty}\sup_{t\in[0,T]\setminus E}\|B_{\tau_n}(t)-\mathcal{B}(u)(t)\|_*=0.$$

holds. Now, let $\eta:[0,T]\to V^*$ be any measurable function chosen to be fixed. Then, for every $\varepsilon>0$ there exists an index $N\in\mathbb{N}$ such that for all $n\ge N$, there holds

$$\|B_{\tau_n}(t)-\mathcal{B}(u)(t)-\eta(t)\|_*\le\varepsilon+\|\eta(t)\|_*\quad\text{for all } t\in[0,T]\setminus E.$$

Invoking the latter estimate, we obtain

$$\begin{aligned}&\|B_{\tau_n}-\mathcal{B}(u)-\eta\|_{\mathrm{L}^2(0,T;V^*)}\\ &\le\left(\int_E\|B_{\tau_n}(t)-\mathcal{B}(u)(t)-\eta(t)\|_*^2\,\mathrm{d}t\right)^{\frac{1}{2}}+\left(\int_{[0,T]\setminus E}\|B_{\tau_n}(t)-\mathcal{B}(u)(t)-\eta(t)\|_*^2\,\mathrm{d}t\right)^{\frac{1}{2}}\\ &\le\mu(E)^{1-\nu}\left(\int_E\|B_{\tau_n}(t)-\mathcal{B}(u)(t)-\eta(t)\|_*^{\frac{2}{\nu}}\,\mathrm{d}t\right)^{\frac{\nu}{2}}+\left(\int_{[0,T]\setminus E}(\varepsilon+\|\eta(t)\|_*)^2\,\mathrm{d}t\right)^{\frac{1}{2}}\\ &\le\varepsilon^{1-\nu}2\tilde{M}+(2T)^{\frac{1}{2}}\varepsilon+\left(\int_{[0,T]\setminus E}2\|\eta(t)\|_*^2\,\mathrm{d}t\right)^{\frac{1}{2}}\\ &\le\varepsilon^{1-\nu}2\tilde{M}+(2T)^{\frac{1}{2}}\varepsilon+2\|\eta\|_{\mathrm{L}^2(0,T;V^*)}\end{aligned}$$

for all $\eta \in \mathrm{L}^{\frac{q^*}{\nu}}(0,T;W^*) \subset \mathrm{L}^{q^*}(0,T;W^*)$ with $\|B_{\tau_n} - \mathcal{B}(u) - \eta\|_{\mathrm{L}^{\frac{2}{\nu}}(0,T;V^*)} \leq 2\tilde{M}$. Finally, we end up with

$$\begin{aligned}
&\|B_{\tau_n} - \mathcal{B}(u)\|_{\mathrm{L}^2(0,T;V^*)+\mathrm{L}^{q^*}(0,T;W^*)}\\
&= \inf_{\eta\in \mathrm{L}^{q^*}(0,T;W^*)} \max\{\|B_{\tau_n} - \mathcal{B}(u) - \eta\|_{\mathrm{L}^2(0,T;V^*)}, \|\eta\|_{\mathrm{L}^{q^*}(0,T;W^*)}\}\\
&= \inf_{\eta\in G_n} \max\{\|B_{\tau_n} - \mathcal{B}(u) - \eta\|_{\mathrm{L}^2(0,T;V^*)}, \|\eta\|_{\mathrm{L}^{q^*}(0,T;W^*)}\}\\
&\leq \inf_{\eta\in G_n} \max\{\varepsilon^{1-\nu}2\tilde{M} + (2T)^{\frac{1}{2}}\varepsilon + 2\|\eta\|_{\mathrm{L}^2(0,T;V^*)}, \|\eta\|_{\mathrm{L}^{q^*}(0,T;W^*)}\}\\
&\leq \varepsilon^{1-\nu}2\tilde{M} + (2T)^{\frac{1}{2}}\varepsilon \quad \text{for all } n \geq N,
\end{aligned}$$

and hence (5.4.1r). Finally, thanks to (5.4.1b) and (5.4.1i), the initial conditions are also fulfilled by u and u', and since $u_0 \in D \cap V$, there holds $u \in \mathrm{H}^1(0,T;V)$, which completes the proof. □

5.5 Proof of Theorem 5.1.4

We first show that the limit function obtained from the previous lemma is indeed a solution to the Cauchy problem. Let $u_0 \in D \cap V_\lambda$, $v_0 \in H$, and a vanishing sequence of step sizes $(\tau_n)_{n\in N}$ be given. We remark that for the estimate (5.3.13) and the solvability of the variational approximation scheme, we needed necessarily the initial data u_0 and v_0 to be in $U \cap V$ in order to solve the variational approximation scheme (5.2.2) and to make use of the growth condition of B in (5.Bb) for the a priori estimates, since the energy functional and the dissipation potential are defined on different spaces. We circumvent this problem via approximating $u_0 \in D \cap V_\lambda$ and $v_0 \in H$ by approximating sequences $(u_0^k)_{k\in\mathbb{N}} \subset D \cap V$ and $(v_0^k)_{k\in\mathbb{N}} \subset V$ such that $u_0^k \to u_0$ in U and $v_0^k \to v_0$ in H as $k \to \infty$, which exists by Condition (5.Ea). Henceforth, we assume $k \in \mathbb{N}$ to be fixed and we define the interpolations associated to the initial values u_0^k and v_0^k as in the previous lemma while omitting the dependence on k for notational convenience. Then, again by the previous lemma, we obtain after selecting a subsequence (not relabeled) of the interpolations, the existence of a limit function $u \in \mathrm{L}^\infty(0,T;U) \cap \mathrm{H}^1(0,T;V) \cap \mathrm{W}^{1,\infty}(0,T;H)$ with $u(0) = u_0$ in U and $u'(0) = v_0^k$ in H that satisfies $u \in \mathrm{H}^2(0,T;U^* + V^*)$ in Case **(a)** and $u \in \mathrm{W}^{1,q}(0,T;W) \cap \mathrm{W}^{\min\{2,q^*\}}(0,T;U^* + V^*)$ in Case **(b)**, where again we omit the dependence of the limit function on k. Now, the inclusion (5.3.10) fulfilled by the interpolations reads in the weak formulation

$$\int_0^T \langle f_{\tau_n}(r) - B_{\tau_n}(r) - \widehat{V}'_{\tau_n}(r) - \xi_{\tau_n}(r) - D_G\Psi(\overline{V}_{\tau_n}(r)), w(r)\rangle_{(U^*+V^*)\times(U\cap V)}\,\mathrm{d}r = 0$$

for all $w \in \mathrm{L}^r(0,T;U \cap V)$ with $r = 2$ in Case **(a)** and $r = \max\{2, 1 + (q-1)/(1 - \delta(q-1))\}$ for a fixed $\delta \in (0, q^* - 1)$ in Case **(b)**, where again $B_{\tau_n}(r) = B(\overline{\mathbf{t}}(r), \overline{U}_{\tau_n}(r), \overline{V}_{\tau_n}(r))$, $r \in [0,T]$ and $D_G\Psi(\overline{V}_{\tau_n}(r)) = A\overline{V}_{\tau_n}(r)$ in Case **(a)** and $D_G\Psi(\overline{V}_{\tau_n}(r)) = A\overline{V}_{\tau_n}(r) + D_G\Psi_2(\overline{V}_{\tau_n}(r))$ in Case **(b)**. For the readers convenience, we confine ourselves to Case **(b)**, but remark that Case **(a)** can be treated in the exact same manner.

Ad case **(b)**. Since $\Psi_1(v) = a(v,v)$ is defined by a strongly positive quadratic form, the Fréchet derivative is a linear bounded and strongly positive operator $A : V \to V^*$, which implies that the associated Nemitskiĭ operator $\mathcal{A} : \mathrm{L}^2(0,T;V) \to \mathrm{L}^2(0,T;V^*) \hookrightarrow \mathrm{L}^{\min\{2,q^*\}}(0,T;U^*)$ is well defined, linear, bounded, and strongly positive. Therefore, the Nemitskiĭ operator is weak-to-weak continuous so that we can pass with $\tau_n \searrow 0$ to the limit as $n \to \infty$. The Gâteaux derivative $D_G\Psi_2(\overline{V}_{\tau_n})$ is strongly convergent to $D_G\Psi_2(u')$ in $\mathrm{L}^{\min\{2,q^*-\delta\}}(0,T;U^*+V^*)$ so that passing to the limit is also justified in this term. We are also allowed to pass to the limit as the step size vanishes in the terms f_{τ_n} and B_{τ_n} which, according to the previous lemma, converge to f and $B(\cdot,u(\cdot),u'(\cdot))$ strongly in $\mathrm{L}^2(0,T;H) \hookrightarrow \mathrm{L}^{\min\{2,q^*\}}(0,T;V^*)$ and $\mathrm{L}^2(0,T;V^*) + \mathrm{L}^{q^*}(0,T;W^*) \hookrightarrow \mathrm{L}^{\min\{2,q^*\}}(0,T;V^*)$ as $n \to \infty$, respectively. Also by the previous lemma, there holds $\widehat{V}'_{\tau_n} \rightharpoonup u''$ in $\mathrm{L}^{\min\{2,q^*\}}(0,T;U^*+V^*)$ and $\xi_{\tau_n} \rightharpoonup \xi$ in $\mathrm{L}^\infty(0,T;U^*+V^*)$. Thus, we are allowed to pass to the limit in the weak formulation in these terms as well. Then, by a well-known density argument and by the fundamental lemma of calculus of variations, we deduce

$$u''(t) + D_G\Psi(u'(t)) + \xi(t) + B(t,u(t),u'(t)) = f(t) \quad \text{in } U^*+V^* \text{ a.e. in } (0,T). \tag{5.5.1}$$

We proceed by showing that $\xi(t) \in \partial_{U\cap V}\mathcal{E}_t(u(t))$ in U^*+V^* for almost every $t \in (0,T)$. To do so, we employ the closedness condition (5.Ee). Since we have already shown that the conditions a)-c) are satisfied, it remains to show the conditions d) and e). Condition d) follows immediately from

$$\|\sigma_{\tau_n}\overline{U}_{\tau_n} - \overline{U}_{\tau_n}\|_{\mathrm{L}^2(0,T-\tau_n;V)} = \tau_n\|\widehat{U}'_{\tau_n}\|_{\mathrm{L}^2(0,T;V)} \le \tau_n M,$$

within Case **(a)** and

$$\begin{aligned}
&\|\sigma_{\tau_n}\overline{U}_{\tau_n} - \overline{U}_{\tau_n}\|_{\mathrm{L}^2(0,T-\tau_n;V)\cap\mathrm{L}^r(0,T-\tau_n;W)} \\
&= \|\sigma_{\tau_n}\overline{U}_{\tau_n} - \overline{U}_{\tau_n}\|_{\mathrm{L}^2(0,T-\tau_n;V)} + \|\sigma_{\tau_n}\overline{U}_{\tau_n} - \overline{U}_{\tau_n}\|_{\mathrm{L}^r(0,T-\tau_n;W)} \\
&\le \tau_n\|\widehat{U}'_{\tau_n}\|_{\mathrm{L}^2(0,T-\tau_n;V)\cap\mathrm{L}^r(0,T-\tau_n;W)} \le \tau_n M,
\end{aligned}$$

in Case **(b)**. Condition e) in turn is verified by the following calculations: let $t \in [0,T]$, then we have

$$\begin{aligned}
&\int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle \xi_{\tau_n}(r), \overline{U}_{\tau_n}(r)\rangle_{V_\lambda^*\times V_\lambda}\,\mathrm{d}r \\
&= \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle S_{\tau_n}(r) - A\overline{V}_{\tau_n}(r) - \widehat{V}'_{\tau_n}(r) - D_G\Psi(\overline{V}_{\tau_n}(r)), \overline{U}_{\tau_n}(r)\rangle_{V_\lambda^*\times V_\lambda}\,\mathrm{d}r \\
&= \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle S_{\tau_n}(r), \overline{U}_{\tau_n}(r)\rangle_{V^*\times V}\,\mathrm{d}r \\
&- \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle \widehat{V}'_{\tau_n}(r), \overline{U}_{\tau_n}(r)\rangle_{(U^*+V^*)\times(U\cap V)}\,\mathrm{d}r \\
&- \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle A\overline{V}_{\tau_n}(r), \overline{U}_{\tau_n}(r)\rangle_{V^*\times V}\,\mathrm{d}r \\
&- \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle D_G\Psi_2(\overline{V}_{\tau_n}(r)), \overline{U}_{\tau_n}(r)\rangle_{V^*\times V}\,\mathrm{d}r \\
&=: I_1^n(t) + I_2^n(t) + I_3^n(t) + I_4^n(t).
\end{aligned}$$

The convergence of the first integral is due to the strong convergence of $S_{\tau_n} = f_{\tau_n} - B_{\tau_n}$ to $f - \mathcal{B}(u)$ in $\mathrm{L}^{\min\{2,q^*\}}(0,T;V^*)$ and the weak* convergence of $\overline{U}_{\tau_n} \rightharpoonup u$ in $\mathrm{L}^\infty(0,T;U\cap V)$ as $n \to \infty$. For the second integral, we recall the discrete integration by parts formula: let $n \in \mathbb{N}$ and $v^k, u^k \in H, k = 0, \dots, n$. Then, there holds

$$\sum_{k=1}^{n}(v^k - v^{k-1}, u^k) = (v^n, u^n) - (v^0, u^0) - \sum_{k=1}^{n}(v^{k-1}, u^k - u^{k-1}).$$

Employing the discrete integration by parts formula, we obtain

$$\begin{aligned} &- \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle \widehat{V}'_{\tau_n}(r), \overline{U}_{\tau_n}(r)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}r \\ &= \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} (\underline{V}_{\tau_n}(r), \overline{V}_{\tau_n}(r)) \,\mathrm{d}r - (\overline{V}_{\tau_n}(t), \overline{U}_{\tau_n}(t)) + (v_0, u_0). \end{aligned} \tag{5.5.2}$$

Thus, by (5.3.9), (5.4.1b), (5.4.1g) and (5.4.1i)

$$\lim_{n\to\infty} I_2^n(t) = \int_0^t (u'(r), u'(r)) \,\mathrm{d}r - (u'(t), u(t)) + (v_0, u_0) \quad \text{for all } t \in [0,T].$$

Employing the more general integration by parts formula for Bochner spaces from Lemma A.1 in Emmrich & Šiška [74] with $a = u$ and $b = u'$, we obtain

$$\int_0^t (u'(r), u'(r)) \,\mathrm{d}r - (u'(t), u(t)) + (v_0, u_0) = - \int_0^t \langle u''(r), u(r)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}r$$

for all $t \in [0,T]$. We proceed with showing the convergence of the third integral $I_3^n(t)$. To do so, we use the symmetry of A and the convexity of Ψ_1, to obtain

$$\begin{aligned} - \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle A\overline{V}_{\tau_n}(r), \overline{U}_{\tau_n}(r)\rangle_{V^*\times V} \,\mathrm{d}r &= - \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle A\overline{U}_{\tau_n}(r), \overline{V}_{\tau_n}(r)\rangle_{V^*\times V} \,\mathrm{d}r \\ &= - \sum_{k=1}^{m} \langle AU_{\tau_n}^k, U_{\tau_n}^k - U_{\tau_n}^{k-1}\rangle_{V^*\times V} \\ &\le - \sum_{k=1}^{m} \Big(\Psi_1(U_{\tau_n}^k) - \Psi_1(U_{\tau_n}^{k-1})\Big) \\ &= \Psi_1(u_0) - \Psi_1(U_{\tau_n}^m) \\ &= \Psi_1(u_0) - \Psi_1(\overline{U}_{\tau_n}(t)). \end{aligned} \tag{5.5.3}$$

for $m \in \{1, \dots, N\}$. Furthermore, we observe that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi_1(u(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}a(u(t), u(t)) = \langle Au(t), u'(t)\rangle_{V^*\times V} \quad \text{for a.e. } t \in (0,T), \tag{5.5.4}$$

which follows from the properties of A and the fact that $u \in \mathrm{H}^1(0,T;V)$. Then, taking into account (5.5.3), (5.5.4), the weak lower semicontinuity of Ψ_1, the pointwise weak convergence (5.4.1c) as well as the symmetry of A, we obtain

$$\begin{aligned}\limsup_{n\to\infty} I_3^n(t) &\le \limsup_{n\to\infty} \Big(\Psi_1(u_0) - \Psi_1(\overline{U}_{\tau_n}(t))\Big)\\ &= -\liminf_{n\to\infty} \Big(\Psi_1(\overline{U}_{\tau_n}(t)) - \Psi_1(u_0)\Big)\\ &\le \Psi_1(u_0) - \Psi_1(u(t))\\ &= \int_0^t \langle Au(r), u'(r)\rangle_{V^*\times V}\,\mathrm{d}r\\ &= \int_0^t \langle Au'(r), u(r)\rangle_{V^*\times V}\,\mathrm{d}r.\end{aligned}$$

In view of (5.4.1a) and (5.4.1p), we obtain for the last integral

$$\begin{aligned}\lim_{n\to\infty} I_4^n(t) &= -\lim_{n\to\infty} \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle D_G\Psi_2(\overline{V}_{\tau_n}(r)), \overline{U}_{\tau_n}(r)\rangle_{(U^*+V^*)\times(U\cap V)}\,\mathrm{d}r\\ &= \int_0^t \langle D_G\Psi_2(u'(r)), u(r)\rangle_{(U^*+V^*)\times(U\cap V)}\,\mathrm{d}r.\end{aligned}$$

We end up with

$$\limsup_{n\to\infty} \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle \xi_{\tau_n}(r), \overline{U}_{\tau_n}(r)\rangle_{V_\lambda^*\times V_\lambda}\,\mathrm{d}r \le \int_0^t \langle \xi(r), u(r)\rangle_{V_\lambda^*\times V_\lambda}\,\mathrm{d}r$$

and thus

$$\limsup_{n\to\infty} \int_0^T \langle \xi_{\tau_n}(r) - \xi(r), \overline{U}_{\tau_n}(r) - u(r)\rangle_{V_\lambda^*\times V_\lambda}\,\mathrm{d}r \le 0.$$

It remains to show the strong convergence $\overline{U}_{\tau_n} - u_0^n \to u - u_0$ in $\mathrm{L}^2(0,T;V)$ as $n\to\infty$ in order to obtain the conclusions of Assumption (5.Ee). We show equivalently that $(\overline{U}_{\tau_n} - u_0^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathrm{L}^2(0,T;V)$. To do so, we follow the idea of the proof of Lemma 4.6 in Emmrich & Šiška [74] and consider in the first step

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\Psi_1\Big(\hat{U}_{\tau_l}(t) - \hat{U}_{\tau_m}(t)\Big)\\ &= \langle A(\hat{U}_{\tau_l}(t) - \hat{U}_{\tau_m}(t)), \overline{V}_{\tau_l}(t) - \overline{V}_{\tau_m}(t)\rangle_{V^*\times V}\\ &= \langle A(\overline{V}_{\tau_m}(t) - \overline{V}_{\tau_l}(t)), \hat{U}_{\tau_l}(t) - \hat{U}_{\tau_m}(t)\rangle_{V^*\times V}\\ &= \langle A(\overline{V}_{\tau_m}(t) - \overline{V}_{\tau_l}(t)), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V}\\ &\quad + \langle A(\overline{V}_{\tau_m}(t) - \overline{V}_{\tau_l}(t)), \hat{U}_{\tau_l}(t) - \overline{U}_{\tau_l}(t) - \hat{U}_{\tau_m}(t) + \overline{U}_{\tau_m}(t)\rangle_{V^*\times V}\\ &= \langle \xi_{\tau_m}(t) - \xi_{\tau_l}(t) + \hat{V}'_{\tau_m}(t) - \hat{V}'_{\tau_l}(t) + S_{\tau_m}(t) - S_{\tau_l}(t)\\ &\quad - D_G\Psi_2(\overline{V}_{\tau_l}(t)) + D_G\Psi_2(\overline{V}_{\tau_m}(t)), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V} + b_{l,m}(t)\\ &= \langle \xi_{\tau_m}(t) - \xi_{\tau_l}(t), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V} + \langle \hat{V}'_{\tau_m}(t) - \hat{V}'_{\tau_l}(t), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V}\\ &\quad + \langle -D_G\Psi_2(\overline{V}_{\tau_l}(t)) + D_G\Psi_2(\overline{V}_{\tau_m}(t)), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V}\\ &\quad + \langle S_{\tau_m}(t) - S_{\tau_l}(t), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V} + b_{l,m}(t)\\ &\le \lambda\|\overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\|_V^2 + \langle \hat{V}'_{\tau_m}(t) - \hat{V}'_{\tau_l}(t), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V}\\ &\quad + \langle -D_G\Psi_2(\overline{V}_{\tau_l}(t)) + D_G\Psi_2(\overline{V}_{\tau_m}(t)), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V}\end{aligned}$$

$$
\begin{aligned}
&+ \langle S_{\tau_m}(t) - S_{\tau_l}(t), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_l}(t)\rangle_{V^*\times V} + b_{l,m}(t) \\
&\leq 2\lambda \|\hat{U}_{\tau_l}(t) - \hat{U}_{\tau_m}(t)\|_V^2 + 2\lambda\|\overline{U}_{\tau_l}(t) - \hat{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t) + \hat{U}_{\tau_m}(t)\|_V^2 \\
&+ \langle \hat{V}'_{\tau_m}(t) - \hat{V}'_{\tau_l}(t), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V} \\
&+ \langle -D_G\Psi_2(\overline{V}_{\tau_l}(t)) + D_G\Psi_2(\overline{V}_{\tau_m}(t)), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V} \\
&+ \langle S_{\tau_m}(t) - S_{\tau_l}(t), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V} + b_{l,m}(t) \\
&\leq \frac{2\lambda}{\mu}\Psi_1(\hat{U}_{\tau_l}(t) - \hat{U}_{\tau_m}(t)) + 2\lambda\|\overline{U}_{\tau_l}(t) - \hat{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t) + \hat{U}_{\tau_m}(t)\|_V^2 \\
&+ \langle \hat{V}'_{\tau_m}(t) - \hat{V}'_{\tau_l}(t), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V} \\
&+ \langle -D_G\Psi_2(\overline{V}_{\tau_l}(t)) + D_G\Psi_2(\overline{V}_{\tau_m}(t)), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V} \\
&+ \langle S_{\tau_m}(t) - S_{\tau_l}(t), \overline{U}_{\tau_l}(t) - \overline{U}_{\tau_m}(t)\rangle_{V^*\times V} + b_{l,m}(t) \\
&= \frac{2\lambda}{\mu}\Psi_1\left(\hat{U}_{\tau_l}(t) - \hat{U}_{\tau_m}(t)\right) + c_{l,m}(t)
\end{aligned}
$$

for almost every $t \in (0,T)$, where we have used the symmetry and strong positivity of A, the λ-convexity of $\mathcal{E}$, and that (5.3.10) is fulfilled. Then, by GRONWALL's lemma (Lemma A.1.1), there holds

$$
\Psi_1(\hat{U}_{\tau_l}(t) - \hat{U}_{\tau_m}(t)) \leq c_{l,m}(t) + \int_0^t \frac{2\lambda}{\mu} c_{l,m}(r) e^{\frac{2\lambda}{\mu}(t-r)}\,\mathrm{d}r.
$$

Integrating the latter inequality from $t = 0$ to $t = T$ and using the strong positivity of Ψ yields

$$
\mu \int_0^T \|\hat{U}_{\tau_l}(t) - \hat{U}_{\tau_m}(t)\|_V^2\,\mathrm{d}t \leq \int_0^T c_{l,m}(t)\,\mathrm{d}t + \int_0^T \int_0^t \frac{2\lambda}{\mu} c_{l,m}(r) e^{\frac{2\lambda}{\mu}(t-r)}\,\mathrm{d}r\,\mathrm{d}t.
$$

Employing again the convergences (5.3.9), (5.4.1a), (5.4.1g), (5.4.1i), (5.4.1k), and (5.4.1p)-(5.4.1r), as well as the discrete integration by parts formula (5.5.2), we obtain $\lim_{l,m\to\infty} \int_0^t c_{l,m}(r)\,\mathrm{d}r = 0$ for all $t \in [0,T]$ and that $\int_0^t c_{l,m}(r)\,\mathrm{d}r \leq C$ for all $l, m \in \mathbb{N}$. Therefore, by the dominated convergence theorem, $(\hat{U}_{\tau_n} - u_0^n)_{n\in\mathbb{N}}$ is a CAUCHY sequence in $\mathrm{L}^2(0,T;V)$. By the convergence (5.3.9), we obtain that $(\overline{U}_{\tau_n} - u_0^n)_{n\in\mathbb{N}}$ is a CAUCHY sequence in $\mathrm{L}^2(0,T;V)$ as well and thus convergent. Hence, by the closedness condition (5.Ee), there holds $\xi(t) \in \partial_{V_\lambda}\mathcal{E}(u(t))$ as well as

$$
\mathcal{E}_{\bar{\mathbf{t}}_{\tau_n}(t)}(\overline{U}_{\tau_n}(t)) \to \mathcal{E}_t(u(t)) \quad \text{and} \quad \limsup_{n\to\infty} \partial_t \mathcal{E}_{\mathbf{t}_n(t)}(\overline{U}_{\tau_n}(t)) \leq \partial_t\mathcal{E}_t(u(t)) \tag{5.5.5}
$$

for a.e. $t \in (0,T)$. Now, we show that the energy-dissipation inequality holds. Let $t \in [0,T]$ and $\mathcal{N} \subset (0,T]$ be a set of measure zero such that $\mathcal{E}_{\bar{\mathbf{t}}_{\tau_n}(s)}(\overline{U}_{\tau_n}(s)) \to \mathcal{E}_t(u(s))$ and $\underline{V}_{\tau_n}(s) \to u'(s)$ for each $s \in [0,T]\backslash\mathcal{N}$. Then, exploiting the convergences (5.4.1) and (5.5.5) as well as the condition (5.Ed) and Theorem 2.6.1, we obtain from the discrete energy-dissipation inequality,

$$
\begin{aligned}
&\frac{1}{2}|u'(t)|^2 + \mathcal{E}_t(u(t)) + \int_s^t \left(\Psi(u'(r)) + \Psi^*(S(r) - \xi(r) - u''(r))\right) \mathrm{d}r \\
&\le \liminf_{n\to\infty} \Bigg(\frac{1}{2}\left|\overline{V}_{\tau_n}(t)\right|^2 + \mathcal{E}_{\bar{\mathbf{t}}_{\tau_n}(t)}(\overline{U}_{\tau_n}(t)) \\
&\qquad\qquad + \int_s^t \left(\Psi(\overline{V}_{\tau_n}(r)) + \Psi^*\left(S_{\tau_n}(r) - \widehat{V}'_{\tau_n}(r) - \xi_{\tau_n}(r)\right)\right) \mathrm{d}r \Bigg) \\
&\le \limsup_{n\to\infty} \Bigg(\frac{1}{2}\left|\overline{V}_{\tau_n}(s)\right|^2 + \mathcal{E}_{\bar{\mathbf{t}}_{\tau_n}(s)}(\overline{U}_{\tau_n}(s)) + \int_{\bar{\mathbf{t}}_{\tau_n}(s)}^{\bar{\mathbf{t}}_{\tau_n}(t)} \partial_r \mathcal{E}_r(\underline{U}_{\tau_n}(r)) \,\mathrm{d}r \\
&\qquad\qquad + \int_{\bar{\mathbf{t}}_{\tau_n}(s)}^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle S_{\tau_n}(r), \overline{V}_{\tau_n}(r)\rangle_{V^*\times V} \,\mathrm{d}r + \tau\lambda \int_{\bar{\mathbf{t}}_{\tau}(s)}^{\bar{\mathbf{t}}_{\tau}(t)} \|\overline{V}_{\tau}(r)\|_V^2 \,\mathrm{d}r \Bigg) \\
&= \frac{1}{2}|u'(s)|^2 + \mathcal{E}_s(u(s)) + \int_s^t \partial_r \mathcal{E}_r(u(r)) \,\mathrm{d}r + \int_s^t \langle S(r), u'(r)\rangle_{V^*\times V} \,\mathrm{d}r,
\end{aligned}
$$

for all $t \in [0,T]$ if $s = 0$ and almost every $s \in (0,t)$, where $S(r) = f(r) - B(r, u(r), u'(r))$. This shows that u is a strong solution to (5.0.1) satisfying the initial conditions $u_k(0) = u_0^k \in D \cap V$ and $u_k'(0) = v_0^k \in V, k \in \mathbb{N}$. We denote with $(u_k)_{k\in\mathbb{N}}$ and $(\xi_k)_{k\in\mathbb{N}}$ the associated solutions and subgradients of $\mathcal{E}_t$ which satisfy (5.1.8)-(5.1.10). We recall that $u_0^k \to u_0$ in $U \cap V_\lambda$ and $v_0^k \to v_0$ in H as $k \to \infty$. The next steps are the same as before:

1. We derive a priori estimates based on the energy-dissipation inequality (5.1.10),
2. We show compactness of the sequences $(u_k)_{k\in\mathbb{N}}$ and $(\xi_k)_{k\in\mathbb{N}}$ in appropriate spaces,
3. We pass to the limit in the inclusion 5.1.9 as $k \to \infty$.

Ad 1. From the energy-dissipation inequality (5.1.10) for $t \in [0,T]$ and $s = 0$ while using the Fenchel–Young inequality, Condition (5.Bb) and (5.Ed), we obtain

$$
\begin{aligned}
&\frac{1}{2}|u_k'(t)|^2 + \mathcal{E}_t(u_k(t)) + \int_0^t \left(\Psi(u_k'(r)) + \Psi^*(S_k(r) - \xi_k(r) - u_k''(r))\right) \mathrm{d}r \\
&\le \frac{1}{2}|v_0^k|^2 + \mathcal{E}_0(u_0^k) + \int_0^t \partial_r \mathcal{E}_r(u_k(r)) \,\mathrm{d}r + \int_0^t \langle S_k(r), u_k'(r)\rangle_{V^*\times V} \,\mathrm{d}r \\
&\le \frac{1}{2}|v_0^k|^2 + \mathcal{E}_0(u_0^k) + C_1 \int_0^t \mathcal{E}_r(u_k(r)) \,\mathrm{d}r \\
&\quad + \int_0^t \langle f(r) - B(r, u_k(r), u_k'(r)), u_k'(r)\rangle_{V^*\times V} \,\mathrm{d}r \\
&\le \frac{1}{2}|v_0^k|^2 + \mathcal{E}_0(u_0^k) + C_1 \int_0^t \mathcal{E}_r(u_k(r)) \,\mathrm{d}r + \int_0^t \left(\frac{1}{2}|f(r)|^2 + \frac{1}{2}|u_k'(r)|^2\right) \mathrm{d}r \\
&\quad + \int_0^t \left(c\,\Psi(u_k'(r)) + c\,\Psi^*(-B(r, u_k(r), u_k'(r))/c)\right) \mathrm{d}r \\
&\le \frac{1}{2}|v_0^k|^2 + \mathcal{E}_0(u_0^k) + \frac{1}{2}\|f\|^2_{\mathrm{L}^2(0,T;H)} + C_1 \int_0^t \mathcal{E}_r(u_k(r)) \,\mathrm{d}r + \frac{1}{2}\int_0^t |u_k'(r)|^2 \,\mathrm{d}r \\
&\quad + \int_0^t \left(c\Psi(u_k'(r)) + \beta(1 + \mathcal{E}_r(u_k(r)) + |u_k'(r)|^2 + \Psi^\nu(u_k'(r)))\right) \mathrm{d}r
\end{aligned}
$$

$$\leq \frac{1}{2}|v_0^k|^2 + \mathcal{E}_0(u_0^k) + \frac{1}{2}\|f\|^2_{\mathrm{L}^2(0,T;H)} + CT + C\int_0^t \left(\mathcal{E}_r(u_k(r)) + \frac{1}{2}|u_k'(r)|^2\right) \mathrm{d}r$$
$$+ (c+\tilde{c})\int_0^t \Psi(u_k'(r))\,\mathrm{d}r$$

for a constant $C = C(\nu, C_1, \beta) > 0$, where $S_k(r) := f(r) - B(r, u_k(r), u_k'(r))$, $r \in [0,T]$ and $\beta \geq 0, c \in (0,1)$, and $\tilde{c} > 0$ such that $c + \tilde{c} \in (0,1)$. Taking into account the non-negativity of Ψ, Ψ^*, by the lemma of GRONWALL (Lemma A.1.1), there exists a constant $C_B > 0$ such that

$$\frac{1}{2}|u_k'(t)|^2 + \mathcal{E}_t(u_k(t)) + \int_0^t \left(\Psi(u_k'(r)) + \Psi^*(S_k(r) - \xi_k(r) - u_k''(r))\right) \mathrm{d}r \leq C_B \quad (5.5.6)$$

for all $t \in [0,T]$.
Ad 2. With the same reasoning as in Lemma 5.4.1, we find (up to a subsequence) the following convergences

$$u_k \overset{*}{\rightharpoonup} u \quad \text{in } \mathrm{L}^\infty(0,T;U), \tag{5.5.7a}$$
$$u_k - u_0^k \overset{*}{\rightharpoonup} u - u_0 \quad \text{in } \mathrm{L}^\infty(0,T;V), \tag{5.5.7b}$$
$$u_k - u_0^k \to u - u_0 \quad \text{in } \mathrm{L}^2(0,T;V), \tag{5.5.7c}$$
$$u_k(t) \rightharpoonup u(t) \quad \text{in } U \text{ for all } t \in [0,T], \tag{5.5.7d}$$
$$u_k \to u \quad \text{in } \mathrm{L}^r(0,T;\widetilde{W}) \quad \text{for any } r \geq 1, \tag{5.5.7e}$$
$$u_k(t) \to u(t) \quad \text{in } \widetilde{W} \text{ for all } t \in [0,T], \tag{5.5.7f}$$
$$u_k' \overset{*}{\rightharpoonup} u' \quad \text{in } \mathrm{L}^2(0,T;V) \cap \mathrm{L}^\infty(0,T;H), \tag{5.5.7g}$$
$$u_k' \to u' \quad \text{in } \mathrm{L}^p(0,T;H) \quad \text{for all } p \geq 1, \tag{5.5.7h}$$
$$u_k'(t) \to u'(t) \quad \text{in } H \text{ for a.e. } t \in (0,T), \tag{5.5.7i}$$
$$u_k'(t) \rightharpoonup u'(t) \quad \text{in } H \text{ for all } t \in [0,T], \tag{5.5.7j}$$
$$\xi_k \overset{*}{\rightharpoonup} \xi \quad \text{in } \mathrm{L}^\infty(0,T;U^* + V^*), \tag{5.5.7k}$$

and in Case **(a)**

$$u_k'' \rightharpoonup u'' \quad \text{in } \mathrm{L}^2(0,T;U^* + V^*), \tag{5.5.7l}$$
$$B(\cdot, u_k, u_k') \to B(\cdot, u, u') \quad \text{in } \mathrm{L}^2(0,T;V^*), \tag{5.5.7m}$$

in Case **(b)**

$$u_k' \rightharpoonup u' \quad \text{in } \mathrm{L}^q(0,T;W), \tag{5.5.7n}$$
$$u_k' \to u' \quad \text{in } \mathrm{L}^{\max\{2,q-\varepsilon\}}(0,T;W) \quad \text{for any } \varepsilon \in [1,q), \tag{5.5.7o}$$
$$D_G\Psi_2(u_k') \to D_G\Psi_2(u') \quad \text{in } \mathrm{L}^r(0,T;U^* + V^*) \quad \text{for any } r \in [1,q^*), \tag{5.5.7p}$$
$$u_k'' \rightharpoonup u'' \quad \text{in } \mathrm{L}^{\min\{2,q^*\}}(0,T;U^* + V^*), \tag{5.5.7q}$$
$$B(\cdot, u_k, u_k') \to B(\cdot, u, u') \quad \text{in } \mathrm{L}^2(0,T;V^*) + \mathrm{L}^{q^*}(0,T;W^*), \tag{5.5.7r}$$

except from the strong convergence (5.5.7c), which needs to be proven. Thus, we show that $(u_k - u_0^k)_{k\in\mathbb{N}}$ is a CAUCHY sequence in $\mathrm{L}^2(0,T;V)$. To do so, we consider

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi_1(u_l(t) - u_0^l - u_m(t) + u_0^m)$$
$$= \langle A(u_l(t) - u_0^l - u_m(t) + u_0^m), u_l'(t) - u_m'(t)\rangle_{V^*\times V}$$

$$
\begin{aligned}
&= \langle A(u_l'(t) - u_m'(t)), u_l(t) - u_0^l - u_m(t) + u_0^m\rangle_{V^*\times V} \\
&= \langle \xi_m(t) - \xi_l(t) + u_m''(t) - u_l''(t) + S_m(t) - S_l(t), u_l(t) - u_0^l - u_m(t) + u_0^m\rangle_{V^*\times V} \\
&= \langle \xi_m(t) - \xi_l(t), u_l(t) - u_0^l - u_m(t) + u_0^m)\rangle_{U^*\times U} \\
&\quad + \langle u_m''(t) - u_l''(t), u_l(t) - u_0^l - u_m(t) + u_0^m\rangle_{(U^*+V^*)\times(U\cap V)} \\
&\quad + \langle S_m(t) - S_l(t), u_l(t) - u_0^l - u_m(t) + u_0^m\rangle_{V^*\times V} \\
&\le \langle \xi_m(t) - \xi_l(t), u_0^m - u_0^l\rangle_{U^*\times U} \\
&\quad + \langle u_m''(t) - u_l''(t), u_l(t) - u_0^l - u_m(t) + u_0^m\rangle_{(U^*+V^*)\times(U\cap V)} \\
&\quad + \langle S_m(t) - S_l(t), u_l(t) - u_0^l - u_m(t) + u_0^m\rangle_{V^*\times V}.
\end{aligned}
$$

where we have taken into account that u_k is a solution of (5.0.1) and that the subdifferential operator $\partial\mathcal{E}_t$ is monotone. Integrating the latter inequality and using the integration by parts rule yields

$$
\begin{aligned}
&\mu\|u_l(t) - u_0^l - u_m(t) + u_0^m\|_V^2 \\
&\le \Psi_1(u_l(t) - u_0^l - u_m(t) + u_0^m) \\
&\le \int_0^t \langle \xi_m(r) - \xi_l(r), u_0^m - u_0^l\rangle_{U^*\times U}\,\mathrm{d}r \\
&\quad + \int_0^t \langle u_m''(r) - u_l''(r), u_l(r) - u_0^l - u_m(r) + u_0^m\rangle_{(U^*+V^*)\times(U\cap V)}\,\mathrm{d}r \\
&\quad + \int_0^t \langle S_m(r) - S_l(r), u_l(r) - u_0^l - u_m(r) + u_0^m\rangle_{V^*\times V}\,\mathrm{d}r \\
&= \int_0^t \langle \xi_m(r) - \xi_l(r), u_0^m - u_0^l\rangle_{U^*\times U}\,\mathrm{d}r \\
&\quad + \int_0^t |u_m'(r) - u_l'(r)|^2\,\mathrm{d}r + (u_m'(t) - u_l'(t), u_l(t) - u_0^l - u_m(t) + u_0^m) \\
&\quad + \int_0^t \langle S_m(r) - S_l(r), u_l(r) - u_0^l - u_m(r) + u_0^m\rangle_{V^*\times V}\,\mathrm{d}r.
\end{aligned}
$$

From the strong convergence $u_0^k \to u_0$ in U as $k\to\infty$ and in view of the convergences (5.5.7) and the a priori bound (5.5.6), the right-hand side is uniformly bounded and convergent to zero for every $t\in[0,T]$ as $m,l\to\infty$. Thus, by the dominated convergence theorem, we conclude that $(u_k - u_0^k)_{k\in\mathbb{N}}$ is a CAUCHY sequence in $\mathrm{L}^2(0,T;V)$, and therefore strongly convergent in $\mathrm{L}^2(0,T;V)$ with the limit $u - u_0$.
Ad 3. With the same argument as before, we show that the equation (5.5.1) is fulfilled. However, it remains to identify $\xi(t)\in\partial_U\mathcal{E}_t(u(t))$ a.e. in $(0,T)$. But this follows from the following limsup estimate and the closedness condition (5.Ee)

$$
\begin{aligned}
&\limsup_{k\to\infty}\int_0^t \langle \xi_k(r) - \xi(t), u_k(t) - u(t)\rangle_{V_\lambda^*\times V_\lambda}\,\mathrm{d}r \\
&\limsup_{k\to\infty}\int_0^t \langle \xi_k(r) - \xi(t), u_k(t) - u_0^k - u(t) + u_0\rangle_{V_\lambda^*\times V_\lambda}\,\mathrm{d}r \\
&= \limsup_{k\to\infty}\Bigg(\int_0^t \langle u''(r) - u_k''(r), u_k(t) - u_0^k - u(t) + u_0\rangle_{(U^*+V^*)\times(U\cap V)}\,\mathrm{d}r \\
&\quad + \int_0^t \langle B(r,u(r),u'(r)) - B(r,u_k(r),u_k'(r)), u_k(t) - u_0^k - u(t) + u_0\rangle_{V^*\times V}\,\mathrm{d}r
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \int_0^t \langle Au'(r) - Au_k'(r), u_k(t) - u_0^k - u(t) + u_0 \rangle_{V^* \times V} \, \mathrm{d}r \Big) \\
&= \limsup_{k \to \infty} \Big(\int_0^t |u'(r) - u_k'(r)|^2 \, \mathrm{d}r + (u'(t) - u_k'(t), u_k(t) - u_0^k - u(t) + u_0) \\
&\quad + \int_0^t \langle B(r, u(r), u'(r)) - B(r, u_k(r), u_k'(r)), u_k(t) - u_0^k - u(t) + u_0 \rangle_{V^* \times V} \, \mathrm{d}r \\
&\quad + \int_0^t \langle A(u_k(t) - u_0^k - u(t) + u_0), u'(r) - u_k'(r) \rangle_{V^* \times V} \, \mathrm{d}r \Big) = 0
\end{aligned}
$$

which again follows from the convergences (5.5.7). Thus, there holds $\xi(t) \in \partial_{V_\lambda} \mathcal{E}_t(u(t))$ a.e. in $(0, T)$, and hence the completion of this proof. □

Remark 5.5.1 If we take a closer look into the proof, we see that the assumption that $\mathcal{E}_t$ is sequentially weakly lower semicontinuous has only been used to show the existence of solutions to the discrete problem and to show the energy-dissipation inequality. If we only address the existence of solutions without the energy-dissipation inequality, we can relax the condition by assuming (in both cases) that there exists $r_0 > 0$ such that $u \mapsto \frac{1}{r_0} a(u, u) + \mathcal{E}_t(u)$ is sequentially weakly lower semicontinuous. The existence of discrete solutions under this assumption follows from the fact that

$$
\frac{1}{\tau} a(u - u_0, u - u_0) + \mathcal{E}_t(u) = \frac{1}{\tau} a(u, u) + \mathcal{E}_t(u) - \frac{2}{\tau} a(u, u_0) + \frac{1}{\tau} a(u_0, u_0),
$$

so that the first two terms are sequentially weakly lower semicontinuous and that the remaining terms are weak-to-weak continuous. In Section 7.3, we will see that this small difference in the proof makes a significant difference in the applications.

Chapter 6

Nonlinearly damped Inertial System

In this chapter, we investigate the abstract CAUCHY problem

$$\begin{cases} u''(t) + \partial\Psi_{u(t)}(u'(t)) + \partial\mathcal{E}_t(u(t)) + B(t,u(t),u'(t)) \ni f(t), & \text{for a.e. } t \in (0,T), \\ u(0) = u_0, \quad u'(0) = v_0, \end{cases} \tag{6.0.1}$$

where again Ψ_u denotes the dissipation potential, $\mathcal{E}_t$ the energy functional, B the perturbation, and f the external force. In the second case, we essentially deal with the case where Ψ_u is nonlinear and non-quadratic and $\mathcal{E}_t = \mathcal{E}^1 + \mathcal{E}^2_t$ is the sum of a functional $\mathcal{E}^1$ that is defined by a strongly positive, symmetric, and bounded bilinear form and a strongly continuous λ-convex functional $\mathcal{E}^2_t$. The perturbation B is again 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

$$\partial_{tt}u - \nabla\cdot\Big(g(u)\,|\nabla\partial_t u|^{q-2}\,\nabla\partial_t u\Big) - \Delta u + W'(u) + b(u,\partial_t u) = f,$$

where $q > 1$, $W : \mathbb{R} \to \mathbb{R}$ is a λ-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

$$\mathcal{E}(u) = \int_\Omega \Big(\frac{1}{2}|\nabla u(x)|^2 + W(u(x))\Big)\,\mathrm{d}x \quad \text{and} \quad \Psi_u(v) = \frac{1}{q}\int_\Omega g(u(x))|\nabla v(x)|^q\,\mathrm{d}x,$$

and the perturbation is given by

$$\langle B(u,v),w\rangle_{\mathrm{L}^2} = \int_\Omega b(u(x),v(x))w(x)\,\mathrm{d}x.$$

In Section 7.4 and 7.5, we discuss multi-valued equations.

6.1 Topological assumptions and main result

As in Chapter 5, we assume that $(U,\|\cdot\|_U), (V,\|\cdot\|_V)$, $(W,\|\cdot\|_W)$ and $(\widetilde{W},\|\cdot\|_{\widetilde{W}})$ are real, reflexive, and separable BANACH spaces such that $U \cap V$ is separable and

reflexive and that $(H, |\cdot|, (\cdot,\cdot))$ is a HILBERT space with norm $|\cdot|$ induced by the inner product $(\cdot,\cdot)$.
Similarly, we assume again the following dense, continuous and compact embeddings

$$\begin{cases} U \cap V \overset{d}{\hookrightarrow} U \overset{c,d}{\hookrightarrow} \widetilde{W} \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} \widetilde{W}^* \overset{d}{\hookrightarrow} U^* \overset{d}{\hookrightarrow} V^* + U^* \\ U \cap V \overset{d}{\hookrightarrow} V \overset{c,d}{\hookrightarrow} W \overset{d}{\hookrightarrow} H \cong H^* \overset{d}{\hookrightarrow} W^* \overset{d}{\hookrightarrow} V^* \overset{d}{\hookrightarrow} V^* + U^*, \end{cases}$$

and if the perturbation does not explicitly depend on u or u', then we do not assume $U \overset{c}{\hookrightarrow} \widetilde{W}$ or $V \overset{c}{\hookrightarrow} W$, respectively. We further assume $V \hookrightarrow W$ if $\mathcal{E}_t^2 \neq 0$, see Condition (6.Ea). We note that we neither assume $U \hookrightarrow V$ nor $V \hookrightarrow U$ as in Chapter 5. Since in this case the subdifferential of Ψ_u is nonlinear, we refer to the inclusion (6.0.1) in the given framework as *nonlinearly damped inertial system* $(U, V, W, \widetilde{W}, H, \mathcal{E}, \Psi, B, f)$.

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

(6.Ψa) **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 $\mathscr{B}(U) \otimes \mathscr{B}(V)$-measurable.

(6.Ψb) **Superlinearity.** The functional Ψ 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

$$c_R(\|v\|_V^q - 1) \leq \Psi_u(v) \leq C_R(\|v\|_V^q + 1) \quad \text{for all } v \in V, t \in [0,T]. \tag{6.1.1}$$

(6.Ψc) **Lower semicontinuity of $\Psi_u + \Psi_u^*$.** For all sequences $v_n \rightharpoonup v$ in $\mathrm{L}^q(0,T;V)$, $\eta_n \rightharpoonup \eta$ in $\mathrm{L}^{q^*}(0,T;V^*)$, and $u_n(t) \rightharpoonup u(t)$ in U for all $t \in [0,T]$ as $n \to \infty$ such that $\sup_{t\in[0,T], n\in\mathbb{N}} \mathcal{E}_t(u(t)) < +\infty$ and $\eta_n(t) \in \partial\Psi_{u_n(t)}(v_n(t))$ a.e. in $t \in (0,T)$ for all $n \in \mathbb{N}$, there holds

$$\int_0^T \Big(\Psi_{u(t)}(v(t)) + \Psi_{u(t)}^*(\xi(t))\Big)\,\mathrm{d}t \leq \liminf_{n\to\infty} \int_0^T \Big(\Psi_{u_n(t)}(v_n(t)) + \Psi_{u_n(t)}^*(\eta_n(t))\Big)\,\mathrm{d}t.$$

For the solvability of problem (5.0.1), 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 $\overline{V}_{\tau_n}$ in the space $\mathrm{L}^q(0,T;V)$, see Lemma 6.4.1.

(6.Ψd) **Uniform monotonicity of $\partial\Psi$.** For all $R > 0$, there exists a constant $\mu_R > 0$ such that

$$\langle \xi - \eta, v - w\rangle_{V^*\times V} \geq \mu_R \|v - w\|_V^{\max\{2,q\}}$$

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

Remark 6.1.1

i) We recall that the conjugate $\Psi_u^* : V^* \to \mathbb{R}$ is lower semicontinuous and convex itself and that the growth condition (6.1.1) implies the following growth condition for the conjugate Ψ^*: for all $R > 0$, there exist positive numbers $\bar{c}_R, \bar{C}_R > 0$ such that for all $u \in U$ with $\sup_{t\in[0,T]} \mathcal{E}_t(u) \le R$, there holds

$$\bar{c}_R(\|\xi\|_{V^*}^{q^*} - 1) \le \Psi_u^*(\xi) \le \bar{C}_R(\|\xi\|_{V^*}^{q^*} + 1) \quad \text{for all } \xi \in V^*,$$

where $q^* = q/(q-1)$.

ii) Also here, we can allow more general time-dependent dissipation potentials $\Psi_u : [0,T] \times V \to [0,+\infty)$ by making the same assumptions specified in Remark 3.2.1 *iv*).

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

(6.Ea) **Basic properties.** For all $t \in [0,T]$, the functional $\mathcal{E}_t : U \to \mathbb{R}$ is the sum of functionals $\mathcal{E}^1 : U \to \mathbb{R}$ and $\mathcal{E}_t^2 : \widetilde{W} \to \mathbb{R}$. The functional $\mathcal{E}^1(\cdot) = \frac{1}{2}b(\cdot,\cdot)$ is induced by a bounded, symmetric, and strongly positive bilinear form $b : U \times U \to \mathbb{R}$, i.e., there exist constants $\mu, \alpha > 0$ such that

$$\begin{aligned} &b(u,v) \le \alpha\|u\|_U\|v\|_U && \text{for all } u,v \in U \\ &\mu\|u\|_U^2 \le b(u,u) && \text{for all } u \in U. \end{aligned}$$

(6.Eb) **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

$$\mathcal{E}_t(u) \ge C_0 \quad \text{for all } u \in U \text{ and } t \in [0,T].$$

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

(6.Ec) **Coercivity.** For every $t \in [0,T]$, $\mathcal{E}_t$ has bounded sublevel sets in U.

(6.Ed) **Control of the time derivative.** For all $u \in U$, the mapping $t \mapsto \mathcal{E}_t^2(u)$ is in $\mathrm{C}([0,T]) \cap \mathrm{C}^1(0,T)$ 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

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

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

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

(6.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 $\mathrm{D}\mathcal{E}_t^2$ such that the mapping $(t,u) \mapsto \mathrm{D}\mathcal{E}_t^2(u)$ 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

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

(6.Ef) **λ-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 $u, v \in U$.

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

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

Again, several remarks are in order.

Remark 6.1.2

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

ii) From Assumption (6.Ed), it follows after integration

$$\sup_{t\in[0,T]} \mathcal{E}_t^2(u) \leq \mathrm{e}^{C_1 T} \inf_{t\in[0,T]} \mathcal{E}_t^2(u),$$
$$|\mathcal{E}_t^2(u) - \mathcal{E}_s^2(u)| \leq e^{C_1 T} \sup_{r\in[0,T]} \mathcal{E}_r^2(u)|s-t| \quad \text{for all } u \in U, s,t \in [0,T].$$

iii) The derivative of the λ-convex energy functional is characterized by the inequality

$$\mathcal{E}_t^2(u) - \mathcal{E}_t^2(v) \leq \langle \mathrm{D}\mathcal{E}^2(u), u-v\rangle_{U^*\times U} + \lambda|u-v|^2 \tag{6.1.2}$$

for all $t \in [0,T]$, $u,v \in U$. In fact, the λ-convexity can be replaced by the latter inequality, since we only make use of (6.1.2) in order to obtain a priori estimates, see Lemma 6.3.1.

We recall that the Fréchet differentiability of $\mathcal{E}_t$ implies the subdifferentiability of $\mathcal{E}_t$ and the subdifferential is a singleton with $\partial\mathcal{E}(u) = \{\mathrm{D}\mathcal{E}(u)\}$.

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

(6.Ba) **Continuity.** The mapping $B : [0,T] \times \widetilde{W} \times W \to V^*$ 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 $[0,T]\times\widetilde{W}\times W$ with $\sup_{n\in\mathbb{N}} \mathcal{G}(u_n) < +\infty$, there holds $B(t_n,u_n,v_n) \to B(t,u,v)$ in V^* as $n \to \infty$.

(6.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 $f \in \mathrm{L}^2(0,T;H)$.

Remark 6.1.3 If the growth condition (6.Ψb) for Ψ_u holds uniformly in $u \in U$, then more general external forces $f \in \mathrm{L}^1(0,T;H) + \mathrm{L}^{q^*}(0,T;V^*)$ can be considered.

6.1.1 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 literature review (Section 1.2), this has not been studied before. However, for single valued operators, a similar case has been investigated by Lions & Strauss [108] and Emmrich & Thalhammer [77].

Ad (6.Ψ). The conditions for the dissipation potentials are similar to those in Section 3.2.1 for perturbed gradient systems. In contrast to the superlinearity condition (3.Ψb), we assume here that Ψ_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 [77], see Lemma 6.4.1 below.

As we mentioned in Remark 2.4.5, the liminf estimate in Condition (6.Ψc) is already implied by the Mosco-convergence $\Psi_{u_n} \xrightarrow{\mathrm{M}} \Psi_u$ for all sequences $u_n \rightharpoonup u$. The prototypical examples for state-independent dissipation potential which fulfill Condition (6.Ψa)-(6.Ψd) are

$$\Psi(v) = \int_\Omega \left(\frac{1}{p}|\nabla v(\boldsymbol{x})|^p + |\nabla v(\boldsymbol{x})|\right) \mathrm{d}\boldsymbol{x} \quad \text{or} \quad \Psi(v) = \int_\Omega \left(\frac{1}{p}|v(\boldsymbol{x})|^p + |v(\boldsymbol{x})|\right) \mathrm{d}\boldsymbol{x}$$

on $V = \mathrm{W}_0^{1,p}(\Omega)$ or $V = \mathrm{L}^p(\Omega)$ with $p \in (1, +\infty)$, respectively. For state-independent dissipation potentials more general integral functionals of the form

$$\Psi_u(v) = \int_\Omega \psi(\boldsymbol{x}, u(\boldsymbol{x}), v(\boldsymbol{x}), \nabla u(\boldsymbol{x}), \nabla v(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}$$

can be considered on appropriate Sobolev spaces, where ψ is a proper, lower semicontinuous and convex function satisfying certain growth and continuity conditions, see Chapter 4, where we discuss this more in detail. Similar to Chapter 3 and 5, we could also consider here more general time-dependent dissipation functionals.

Ad (6.E). The crucial assumption we make for the energy functional $\mathcal{E}_t = \mathcal{E}^1 + \mathcal{E}_t^2$ is that the leading part $\mathcal{E}^1$ 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 $\mathcal{E}_t^2$ 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 $\mathcal{E}_t^2$, see Section 5.1.1 for a discussion of the assumptions made in Chapter 6. Ad (6.B). Since we have exactly the same conditions on B, we have the same remarks as in Section 5.1.1 for linear damping inertial systems.

Having discussed all assumptions, we are in a position to state the main result which again includes the notion of solution to (6.0.1).

Theorem 6.1.4 (Existence result) *Let the nonlinearly damped inertial system* $(U, V, W, \widetilde{W}, H, \mathcal{E}, \Psi, B, f)$ *be given and fulfill Assumptions* (6.E), (6.Ψa)-(6.Ψc) *as well as* (6.B) *and Assumption* (6.f). *Then, for every* $u_0 \in U$ *and* $v_0 \in H$, *there exists a solution to* (5.0.1), *i.e., there exist functions* $u \in \mathrm{C}_w([0,T]; U) \cap \mathrm{W}^{1,\infty}(0,T; H) \cap \mathrm{W}^{2,q^*}(0,T; U^* + V^*)$ *with* $u - u_0 \in \mathrm{W}^{1,q}(0,T; V)$ *and* $\eta \in \mathrm{L}^{q^*}(0,T; V^*)$ *satisfying the initial conditions* $u(0) = u_0$ *in* U *and* $u'(0) = v_0$ *in* H *such that*

$$\eta(t) \in \partial\Psi_{u(t)}(u'(t)) \quad \textit{and} \quad u''(t) + \eta(t) + \mathrm{D}\mathcal{E}_t(t) + B(t, u(t), u'(t)) = f(t) \quad \textit{in } U^* + V^* \tag{6.1.3}$$

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

$$\begin{aligned} &\frac{1}{2}|u'(t)|^2 + \mathcal{E}_t(u(t)) + \int_s^t \Big(\Psi_{u(t)}(u'(r)) + \Psi^*_{u(t)}(S(r) - \mathrm{D}\mathcal{E}_r(r) - u''(r)\Big)\,\mathrm{d}r \\ &= \frac{1}{2}|v_0|^2 + \mathcal{E}_0(u_0) + \int_0^t \partial_r\mathcal{E}_r(u(r))\,\mathrm{d}r + \int_0^t \langle S(r), u'(r)\rangle_{V^*\times V}\,\mathrm{d}r \end{aligned} \tag{6.1.4}$$

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* (6.1.4) *holds for all* $t \in [0,T]$.

6.2 Variational approxiomation scheme

The proof of Theorem 6.1.4 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 $\mathrm{D}\mathcal{E}$ and $\partial\Psi$. Again, for $N \in \mathbb{N}\backslash\{0\}$, 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 (5.0.1) in a semi-implicit manner yields

$$\frac{V_\tau^n - V_\tau^{n-1}}{\tau} + \partial_V\Psi_{U_\tau^{n-1}}(V_\tau^n) + \mathrm{D}\mathcal{E}_{t_n}(U_\tau^n) + B\left(t_n, U_\tau^{n-1}, V_\tau^{n-1}\right) \ni f_\tau^n \quad \text{in } U^* + V^* \tag{6.2.1}$$

for $n = 1, \ldots, N$ with $V_\tau^n = \frac{U_\tau^n - U_\tau^{n-1}}{\tau}$. The value U_τ^n is to be determined recursively from the variational approximation scheme

$$\begin{cases} U_\tau^0 \in U \cap V \text{ and } V_\tau^0 \in V \text{ are given; whenever } U_\tau^1, \ldots, U_\tau^{n-1} \in D \cap V \text{ are known,} \\ \text{find } U_\tau^n \in J_{\tau, t_{n-1}}(U_\tau^{n-1}, U_\tau^{n-2}; B(t_n, U_\tau^{n-1}, V_\tau^{n-1}) - f_\tau^n) \end{cases} \tag{6.2.2}$$

for $n = 1, \ldots, N$, where $J_{r,t}(v, w; \eta) := \operatorname{argmin}_{u \in U \cap V} \Phi(r, t, v, w, \eta; u)$ and $U_\tau^{-1} = U_\tau^0 - V_\tau^0 \tau$ with

$$\Phi(r, t, v, w, \eta; u) = \frac{1}{2r^2}|u - 2v + w|^2 + r\Psi_u\left(\frac{u - v}{r}\right) + \mathcal{E}_{t+r}(u) - \langle \zeta, u \rangle_{V^* \times V}$$

for $r \in \mathbb{R}^{>0}, t \in [0, T)$ with $r + t \in [0, T]$, $u \in U \cap V, v \in V, w \in H$ and $\zeta \in V^*$. The solvability of the discrete problem (6.2.2) and that every solution fulfills the Euler–Lagrange equation (6.2.1) is ensured by the following lemma.

Lemma 6.2.1 *Let the nonlinearly damped inertial system* $(U, V, W, H, \mathcal{E}, \Psi)$ *be given and fulfill the Conditions* (6.Ea)-(6.Ec)*,* (6.Ee)*,* (6.Ef) *and* (6.Ψa)-(6.Ψb)*. 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* λ *is from* (6.Ef)*. Furthermore, to every* $u \in J_{r,t}(v, w; \eta)$*, there exists* $\eta \in \partial_V \Psi_u\left(\frac{u-v}{r}\right) \subset V^*$ *such that*

$$\frac{u - 2v - w}{r^2} + \eta + \mathrm{D}\mathcal{E}_t(u) + \zeta = 0 \qquad \text{in } U^* + V^*.$$

Proof. The proof follows along the same lines as the proof to Lemma 5.2.1 by employing the direct methods of the calculus of variations as well as Lemma 2.2.7. ☐

Thus, Lemma 2.2.5 ensures that minimizer of the mapping

$$u \mapsto \Phi(\tau, t_{n-1}, U_\tau^{n-1}, U_\tau^{n-2}, B(t_n, U_\tau^{n-1}, V_\tau^{n-1}) - f_\tau^n; u),$$

fulfil the Euler–Lagrange equation (6.2.1).

6.3 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 intial values in the main existence result and approximate by suitable sequences of values. Then, for given approximate values $(U_\tau^n)_{n=0}^N$ with $U_\tau^0 := u_0$ and $V_\tau^0 = v_0$ obtained from the variational approximation scheme (6.2.2), we define again the piecewise constant and linear interpolations $\overline{U}_\tau, \underline{U}_\tau, \widehat{U}_\tau, \overline{V}_\tau, \underline{V}_\tau, \widehat{V}_\tau, \xi_\tau, f_\tau$ as well as $\overline{\mathbf{t}}_\tau$ and $\underline{\mathbf{t}}_\tau$ as in (5.3.1)-(5.3.4).

Furthermore, by Lemma 6.2.1, there exists a sequence $(\eta_\tau^n)_{n=1}^N \subset V^*$ of subgradients fulfilling $\eta_\tau^n \in \partial_V \Psi_{U_\tau^{n-1}}(V_\tau^n), n = 1, \dots, N$, such that

$$\frac{V_\tau^n - V_\tau^{n-1}}{\tau} + \eta_\tau^n + \mathrm{D}\mathcal{E}_{t_n}(U_\tau^n) + B\left(t_n, U_\tau^{n-1}, V_\tau^{n-1}\right) = f_\tau^n \quad \text{in } U^* + V^*, n = 1, \dots, N.$$

Then, we define the measurable function $\eta_\tau : [0,T] \to V^*$ by

$$\eta_\tau(t) = \eta_\tau^n \quad \text{for } t \in (t_{n-1}, t_n],\ n = 1, \dots, N, \quad \text{and} \quad \eta_\tau(T) = \eta_\tau^N. \tag{6.3.1}$$

Having defined the interpolations, we are in the position to show the a priori estimates in the following lemma.

Lemma 6.3.1 (A priori estimates) *Let the system* $(U, V, W, H, \mathcal{E}, \Psi, B, f)$ *be given and satisfy the Assumptions* (6.E), (6.Ψ), (6.B) *as well as Assumption* (6.f). *Furthermore, let* $\overline{U}_\tau, \underline{U}_\tau, \widehat{U}_\tau, \overline{V}_\tau, \underline{V}_\tau, \widehat{V}_\tau, \eta_\tau$ *and* f_τ *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*

$$\begin{aligned}
&\int_{\bar{\mathfrak{t}}_\tau(s)}^{\bar{\mathfrak{t}}_\tau(t)} \left(\Psi_{\underline{U}_\tau(r)}(\overline{V}_\tau(r)) + \Psi^*_{\underline{U}_\tau(r)}\left(S_\tau(r) - \widehat{V}_\tau'(r) - \mathrm{D}\mathcal{E}_{\bar{\mathfrak{t}}(r)}(\overline{U}_\tau(r))\right)\right) \mathrm{d}r \\
&\quad + \frac{1}{2}\left|\overline{V}_\tau(t)\right|^2 + \mathcal{E}_{\bar{\mathfrak{t}}_\tau(t)}(\overline{U}_\tau(t)) \\
&\le \frac{1}{2}\left|\overline{V}_\tau(s)\right|^2 + \mathcal{E}_{\bar{\mathfrak{t}}_\tau(s)}(\overline{U}_\tau(s)) + \int_{\bar{\mathfrak{t}}_\tau(s)}^{\bar{\mathfrak{t}}_\tau(t)} \partial_r \mathcal{E}_r(\underline{U}_\tau(r))\,\mathrm{d}r + \int_{\bar{\mathfrak{t}}_\tau(s)}^{\bar{\mathfrak{t}}_\tau(t)} \langle S_\tau(r), \overline{V}_\tau(r)\rangle_{U^* \times U}\,\mathrm{d}r \\
&\quad + \tau\lambda \int_{\bar{\mathfrak{t}}_\tau(s)}^{\bar{\mathfrak{t}}_\tau(t)} |\overline{V}_\tau(r)|^2\,\mathrm{d}r
\end{aligned} \tag{6.3.2}$$

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

$$\sup_{t\in[0,T]} \left|\overline{V}_\tau(t)\right| \le M, \quad \sup_{t\in[0,T]} \mathcal{E}_t(\overline{U}_\tau(t)) \le M, \quad \sup_{t\in[0,T]} |\partial_t \mathcal{E}_t(\underline{U}_\tau(t))| \le M, \tag{6.3.3}$$

$$\int_0^T \left(\Psi_{\underline{U}_\tau(r)}\left(\overline{V}_\tau(r)\right) + \Psi^*_{\underline{U}_\tau(r)}\left(S_\tau(r) - \widehat{V}_\tau'(r) - \mathrm{D}\mathcal{E}_{\bar{\mathfrak{t}}(r)}(\overline{U}_\tau(r))\right)\right) \mathrm{d}r \le M \tag{6.3.4}$$

hold for all $0 < \tau \le \tau^*$. *In particular, the families of functions*

$$(\overline{U}_\tau)_{0<\tau\le\tau^*} \subset \mathrm{L}^\infty(0,T;U), \tag{6.3.5a}$$

$$(\overline{V}_\tau)_{0<\tau\le\tau^*} \subset \mathrm{L}^q(0,T;V), \tag{6.3.5b}$$

$$(\eta_\tau)_{0<\tau\le\tau^*} \subset \mathrm{L}^{q^*}(0,T;V^*), \tag{6.3.5c}$$

$$(\widehat{V}_\tau')_{0<\tau\le\tau^*} \subset \mathrm{L}^{q^*}(0,T;U^* + V^*), \tag{6.3.5d}$$

$$(B_\tau)_{0<\tau\le\tau^*} \subset \mathrm{L}^{\frac{q}{\nu}}(0,T;V^*), \tag{6.3.5e}$$

$$(\mathrm{D}\mathcal{E}^2_{\bar{\mathfrak{t}}}(\overline{U}_\tau))_{0<\tau\le\tau^*} \subset \mathrm{L}^\infty(0,T;\widetilde{W}^*), \tag{6.3.5f}$$

are uniformly bounded with respect to τ in the respective spaces, where $q^* > 0$ is the conjugate exponent to $q > 1$ and $\nu \in (0,1)$ being from Assumption (6.Bb). Finally, there holds

$$\begin{aligned} \sup_{t\in[0,T]} \left(\|\underline{U}_\tau(t) - \overline{U}_\tau(t)\|_V + \|\widehat{U}_\tau(t) - \overline{U}_\tau(t)\|_V \right) \to 0 \\ \sup_{t\in[0,T]} \left(\|\overline{V}_\tau(t) - \widehat{V}_\tau(t)\|_{U^*+V^*} + \|\underline{V}_\tau(t) - \overline{V}_\tau(t)\|_{U^*+V^*} \right) \to 0 \end{aligned} \tag{6.3.6}$$

as $\tau \to 0$.

Proof. Let $(U_\tau^n)_{n=1}^N \subset U \cap V$ be the approximative values obtained from the variational approximation scheme (6.2.2) which satisfy by Lemma 2.2.7 the Euler–Lagrange equation

$$f_\tau^n - B(t_n, U_\tau^{n-1}, V_\tau^{n-1}) - \frac{V_\tau^n - V_\tau^{n-1}}{\tau} - \mathrm{D}\mathcal{E}_{t_n}(U_\tau^n) = \eta_\tau^n \in \partial_V \Psi_{U_\tau^{n-1}}(V_\tau^n) \tag{6.3.7}$$

for all $n = 1, \dots, N$. According to Lemma 2.3.1, inclusion (6.3.7) is equivalent to

$$\begin{aligned} &\Psi_{U_\tau^{n-1}}(V_\tau^n) + \Psi^*_{U_\tau^{n-1}} \left(f_\tau^n - B(t_n, U_\tau^{n-1}, V_\tau^{n-1}) - \frac{V_\tau^n - V_\tau^{n-1}}{\tau} - \mathrm{D}\mathcal{E}_{t_n}(U_\tau^n) \right) \\ &= \left\langle f_\tau^n - B(t_n, U_\tau^{n-1}, V_\tau^{n-1}) - \frac{V_\tau^n - V_\tau^{n-1}}{\tau} - \mathrm{D}\mathcal{E}_{t_n}(U_\tau^n), V_\tau^n \right\rangle_{V^*\times V}, \quad n = 1, \dots, N. \end{aligned}$$

Furthermore, the enhanced Fréchet subdifferentiability (6.Ef) yields

$$\begin{aligned} -\left\langle \mathrm{D}\mathcal{E}_{t_n}(U_\tau^n), U_\tau^n - U_\tau^{n-1} \right\rangle_{(U^*+V^*)\times(U\cap V)} &\le \mathcal{E}_{t_n}(U_\tau^{n-1}) + \lambda |U_\tau^n - U_\tau^{n-1}|^2 \\ &\quad - \mathcal{E}_{t_n}(U_\tau^n) \\ &= \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) - \mathcal{E}_{t_n}(U_\tau^n) + \lambda |U_\tau^n - U_\tau^{n-1}|^2 \\ &\quad + \int_{t_{n-1}}^{t_n} \partial_r \mathcal{E}_r(U_\tau^{n-1})\,\mathrm{d}r \end{aligned}$$

for all $n = 1, \dots, N$. Employing (5.3.11), we obtain

$$\begin{aligned} &\frac{1}{2}|V_\tau^n|^2 + \mathcal{E}_{t_n}(U_\tau^n) + \tau \Psi_{U_\tau^{n-1}}(V_\tau^n) + \tau \Psi^*_{U_\tau^{n-1}} \left(S_\tau^n - \frac{V_\tau^n - V_\tau^{n-1}}{\tau} - \mathrm{D}\mathcal{E}_{t_n}(U_\tau^n) \right) \\ &\le \frac{1}{2}|V_\tau^{n-1}|^2 + \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) + \int_{t_{n-1}}^{t_n} \partial_r \mathcal{E}_r(U_\tau^{n-1})\,\mathrm{d}r + \lambda \int_{t_{n-1}}^{t_n} |V_\tau^n|^2\,\mathrm{d}r + \tau \left\langle S_\tau^n, V_\tau^n \right\rangle_{V^*\times V} \end{aligned}$$

for all $n = 1, \dots, N$, where $S_\tau^n := f_\tau^n - B(t_n, U_\tau^{n-1}, V_\tau^{n-1}), n = 1, \dots, N$. Summing up the inequalities over n yields (6.3.2). Analogously to the proof of Lemma 5.3.1, the estimates (6.3.3) and (6.3.4) are obtained by employing the discrete version of Gronwall's lemma (Lemma A.1.2) taking further into account that by Condition (6.Ed), there holds

$$\partial_r \mathcal{E}_r^2(U_\tau^{n-1}) \le C_1 \mathcal{E}_r^2(U_\tau^{n-1})$$

for all $r \in (0, T)$ and $n = 1, \dots, N$. The estimates (6.3.3) and (6.3.4) in turn imply in view of Assumption (6.Ea),(6.Ec),(6.Eg) as well as Assumption (6.Ψb) and Remark 6.1.1 the uniform bounds (6.3.5a)-(6.3.5f) and the convergences (6.3.6). □

6.4 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 (6.2.1) as the step size vanishes. After identifying all the weak limits, we will indeed obtain a solution to the Cauchy problem (5.0.1). The compactness result is given in the following lemma whose proof follow along the same line as Lemma (5.4.1) for linearly damped inertial systems. Therefore, we will prove the assertions which differ from the previously mentioned lemma.

Lemma 6.4.1 (Compactness) *Under the same assumptions of Lemma 6.3.1, 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, η) with*

$$u \in \mathrm{C}_w([0,T];U) \cap \mathrm{W}^{1,q}(0,T;V) \cap \mathrm{W}^{1,\infty}(0,T;H) \cap \mathrm{W}^{2,q^*}(0,T;U^* + V^*) \text{ and}$$
$$\eta \in \mathrm{L}^{q^*}(0,T;V^*),$$

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

$$\begin{aligned}
\underline{U}_{\tau_n}, \overline{U}_{\tau_n}, \widehat{U}_{\tau_n} &\overset{*}{\rightharpoonup} u && \text{in } \mathrm{L}^\infty(0,T;U \cap V), && (6.4.1a)\\
\widehat{U}_{\tau_n}(t), \underline{U}_{\tau_n}(t), \overline{U}_{\tau_n}(t) &\rightharpoonup u(t) && \text{in } U \quad \text{for all } t \in [0,T], && (6.4.1b)\\
\underline{U}_{\tau_n}(t) &\rightharpoonup u(t) && \text{in } V \quad \text{for all } t \in [0,T], && (6.4.1c)\\
\underline{U}_{\tau_n} &\to u && \text{in } \mathrm{L}^r(0,T;\widetilde{W}) \quad \text{for any } r \geq 1, && (6.4.1d)\\
\widehat{U}_{\tau_n}(t), \underline{U}_{\tau_n}(t), \overline{U}_{\tau_n}(t) &\to u(t) && \text{in } \widetilde{W} \quad \text{for all } t \in [0,T], && (6.4.1e)\\
\overline{V}_{\tau_n}, \underline{V}_{\tau_n} &\overset{*}{\rightharpoonup} u' && \text{in } \mathrm{L}^q(0,T;V) \cap \mathrm{L}^\infty(0,T;H), && (6.4.1f)\\
\overline{V}_{\tau_n}, \underline{V}_{\tau_n} &\to u' && \text{in } \mathrm{L}^p(0,T;H) \quad \text{for all } p \geq 1, && (6.4.1g)\\
\widehat{V}_{\tau_n}(t), \underline{V}_{\tau_n}(t), \overline{V}_{\tau_n}(t) &\rightharpoonup u'(t) && \text{in } H \quad \text{for all } t \in [0,T], && (6.4.1h)\\
\eta_{\tau_n} &\rightharpoonup \eta && \text{in } \mathrm{L}^{q^*}(0,T;V^*), && (6.4.1i)\\
\mathrm{E}\overline{U}_{\tau_n} &\rightharpoonup \mathrm{E}u && \text{in } \mathrm{L}^2(0,T;U^*), && (6.4.1j)\\
\mathrm{D}\mathcal{E}^2_{\bar{\mathbf{t}}_{\tau_n}}(\overline{U}_{\tau_n}) &\to \mathrm{D}\mathcal{E}^2_t(u) && \text{in } \mathrm{L}^r(0,T;\widetilde{W}^*) \quad \text{for any } r \geq 1, && (6.4.1k)\\
\widehat{V}'_{\tau_n} &\rightharpoonup u'' && \text{in } \mathrm{L}^{\min\{2,q^*\}}(0,T;U^* + V^*), && (6.4.1l)\\
f_{\tau_n} &\to f && \text{in } \mathrm{L}^2(0,T;H), && (6.4.1m)\\
B_{\tau_n} &\to B(\cdot, u(\cdot), u'(\cdot)) && \text{in } \mathrm{L}^{r^*}(0,T;V^*), && (6.4.1n)
\end{aligned}$$

Furthermore, if the dissipation potential satisfies in addition Assumption (6.Ψd), then, there holds

$$\begin{aligned}
\overline{V}_{\tau_n} &\to u' && \text{in } \mathrm{L}^{\max\{2,q\}}(0,T;U), && (6.4.2a)\\
\widehat{U}_{\tau_n} &\to u && \text{in } \mathrm{C}([0,T];U). && (6.4.2b)
\end{aligned}$$

Finally, the function u satisfies the inequality

$$\begin{aligned}
&\frac{1}{2}|v_0|^2 - \frac{1}{2}|u'(t)|^2 + \mathcal{E}_0(u_0) - \mathcal{E}_t(u(t)) + \int_0^t \partial_r \mathcal{E}_r(u(r))\,\mathrm{d}r \\
&\leq -\int_0^t \langle u''(r) + \mathrm{D}\mathcal{E}_r(u(r)), u'(r)\rangle_{V^* \times V} \qquad (6.4.3)
\end{aligned}$$

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

Proof. We restrict the proof by only showing the convergence (6.4.1j),(6.4.1k), (6.4.2a), and (6.4.2b) and note that the remainder of the proof can be proved in the same manner as the proof of Lemma 5.4.1. First, convergence (6.4.1j) follows from Remark 6.1.2 *i)* and the weak convergence (6.4.1a). Further, from the growth condition (6.Eg), we obtain

$$\|\mathrm{D}\mathcal{E}^2_{\mathbf{t}_{\tau_n}(t)}(\overline{U}_{\tau_n}(t))\|^{\sigma}_{\widetilde{W}^*} \leq C_3(1+\mathcal{E}^2(\overline{U}_{\tau_n}(t)) + \|\overline{U}_{\tau_n}(t)\|_{\widetilde{W}})$$

and in view of the a priori estimates (6.3.3),

$$\|\mathrm{D}\mathcal{E}^2_{\mathbf{t}_{\tau_n}(t)}(\overline{U}_{\tau_n}(t))\|_{W^*} \leq C \quad \text{for all } t \in [0,T].$$

Together with the convergence (6.4.1e) and the continuity condition (6.Ee), this leads to (6.4.1k). The last assertions (6.4.2a) and (6.4.2b) follow immediately from Assumption (6.Ψd) and

$$\begin{aligned}
&\limsup_{n\to\infty} \int_0^T \|\overline{V}_{\tau_n}(r) - u'(r)\|_V^{\max\{p,2\}}\,\mathrm{d}r \\
&\leq \limsup_{n\to\infty} \int_0^T \langle \eta_n(r) - \eta(r), \overline{V}_{\tau_n}(r) - u'(r)\rangle_{V^*\times V}\,\mathrm{d}r \leq 0
\end{aligned}$$

and $\eta(t) \in \partial_V \Psi_{u(t)}(u'(t))$ a.e. in $(0,T)$, which we will show in the proof of the main result. It remains to show the inequality (6.4.3). 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 (6.4.3) is a slight modification of Lemma 6 in Emmrich & Thalhammer [77], we follow their proof and regularize the function u' by its so-called Steklov average. For a function $v \in \mathrm{L}^p(0,T;X), p \geq 1$, 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

$$S_h v(t) := \frac{1}{2h}\int_{t-h}^{t+h} v(r)\,\mathrm{d}r.$$

It is readily seen that $S_h v \in \mathrm{L}^p(0,T;X)$ and $\|S_h v\|_{\mathrm{L}^p(0,T;X)} \leq \|v\|_{\mathrm{L}^p(0,T;X)}$. Furthermore, it can be shown by a regularization argument that $S_h v \to v$ in $\mathrm{L}^p(0,T;X)$ as $h \to 0$, see , e.g., Diestel & Uhl [58, Theorem 9, p. 49].
Defining $Kv(t) = \int_0^t v(r)\,\mathrm{d}r$, we commence with calculating

$$\begin{aligned}
&-\int_s^t \langle (S_h u')'(r) + \mathrm{D}\mathcal{E}_r(u_0 + (KS_h u')(r)), (S_h u')(r)\rangle_{V^*\times V}\,\mathrm{d}r \\
&-\int_s^t \langle (S_h u')'(r) + \mathrm{E}(u_0 + (KS_h u')(r)) + \mathrm{D}\mathcal{E}^2_r(u_0 + (KS_h u')(r)), (S_h u')(r)\rangle_{V^*\times V}\,\mathrm{d}r \\
&= \frac{1}{2}|(S_h u')(s)|^2 - \frac{1}{2}|(S_h u)'(t)|^2 + \mathcal{E}^1(u_0 + (KS_h u')(s)) - \mathcal{E}^1(u_0 + (KS_h u')(t)) \\
&+ \quad \mathcal{E}^2_s(u_0 + (KS_h u')(s)) - \mathcal{E}^2_t(u_0 + (KS_h u')(t))
\end{aligned}$$

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 $(S_h u')(t) = \frac{1}{2h}(\tilde{u}(t+h) - \tilde{u}(t-h))$, where $\tilde{u}$ is a continuous extension of u outside $[0,T]$ which makes sense, since $u \in \mathrm{L}^\infty(0,T;U) \cap \mathrm{W}^{1,1}(0,T;H) \subset \mathrm{C}_w([0,T];U)$ and therefore $S_h u' \in \mathrm{L}^2(0,T;U)$. 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 $u'' + \mathrm{D}\mathcal{E}_t(u) \in \mathrm{L}^{q^*}(0,T;V^*)$. Nevertheless, since we have assumed $V \hookrightarrow \widetilde{W}$, we can treat the term involving $\mathrm{D}\mathcal{E}_t^2 : \widetilde{W} \to \widetilde{W}^* \hookrightarrow V^*$ separately. First, taking into account

$$u_0 + (KS_h u')(t) = u_0 + \frac{1}{2h}\int_{t-h}^{t+h} \tilde{u}(r)\,\mathrm{d}r - \frac{1}{2h}\int_{-h}^{+h} \tilde{u}(r)\,\mathrm{d}r$$

and that $u \in \mathrm{C}_w([0,T];U) \subset \mathrm{C}([0,T];\widetilde{W})$ since $U \overset{c}{\hookrightarrow} \widetilde{W}$, there holds

$$\lim_{h\to 0} (u_0 + (KS_h u')) = u \quad \text{in } \mathrm{C}([0,T];\widetilde{W}). \tag{6.4.4}$$

Finally, by the continuity of $\mathcal{E}_t^2$ and $\mathrm{D}\mathcal{E}_t^2$, the convergences (6.4.4) and $S_h u' \to u'$ in $\mathrm{L}^q(0,T;V)$ as $h \to 0$, there holds

$$\begin{aligned}
&= -\int_s^t \langle \mathrm{D}\mathcal{E}_r^2(u(r)), u'(r)\rangle_{V^*\times V}\,\mathrm{d}r \\
&= \lim_{h\to 0} -\int_s^t \langle \mathrm{D}\mathcal{E}_r^2(u_0 + (KS_h u')(r)), (S_h u')(r)\rangle_{V^*\times V}\,\mathrm{d}r \\
&= \lim_{h\to 0} \Big(\mathcal{E}_s^2(u_0 + (KS_h u')(s)) - \mathcal{E}_t^2(u_0 + (KS_h u')(t))\Big) \\
&= \mathcal{E}_s^2(u(s)) - \mathcal{E}_t^2(u(t))
\end{aligned} \tag{6.4.5}$$

for all $s, t \in [0,T]$. Second, it has been shown in Emmrich & Thalhammer [77, Lemma 6] that

$$\begin{aligned}
&-\int_0^t \langle u''(r) + \mathrm{E}(u(r)), u'(r)\rangle_{V^*\times V}\,\mathrm{d}r \\
&\leq \frac{1}{2}|v_0|^2 - \frac{1}{2}|u'(t)|^2 + \mathcal{E}^1(u_0) - \mathcal{E}^1(u(t))
\end{aligned}$$

for almost every $t \in (0,T)$. The latter inequality together with (6.4.5) implies (6.4.3), which completes the proof.

□

6.5 Proof of Theorem 6.1.4

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 again by (5.3.1)-(5.3.3) and (6.3.1). 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 $u \in \mathrm{C}_w([0,T];U) \cap \mathrm{W}^{1,\infty}(0,T;H) \cap \mathrm{W}^{1,q}(0,T;V^*) \cap \mathrm{W}^{2,r^*}(0,T;U^*+V^*)$ (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 (6.4.1) hold, where we again suppress the dependence of the limit functions on k. First, we prove that the inclusion (6.1.3) holds. To do so, we note that the Euler–Lagrange equation (6.3.7) reads

$$\begin{aligned} &\widehat{V}_{\tau_n}'(t) + \eta_{\tau_n}(t) + \mathrm{D}\mathcal{E}_{\bar{\mathbf{t}}_{\tau_n}(t)}(\overline{U}_{\tau_n}(t)) + S_{\tau_n}(t) = 0 \quad \text{in } U^* + V^*, \\ &\eta_n(t) \in \partial_V \Psi_{\underline{U}_{\tau_n}(t)}(\overline{V}_{\tau_n}(t)) \end{aligned} \tag{6.5.1}$$

for all $t \in (0,T)$, where $S_{\tau_n}(t) = B(\bar{\mathbf{t}}_{\tau_n}(t), \underline{V}_{\tau_n}(t), \underline{U}_{\tau_n}(t)) - f_{\tau_n}(t), t \in [0,T]$. Testing equation (6.5.1) with $w \in \mathrm{L}^{\max\{2,q\}}(0,T;U \cap V)$, we obtain

$$\int_0^T \langle \widehat{V}_{\tau_n}'(r) + \eta_{\tau_n}(r) + \mathrm{D}\mathcal{E}_{\bar{\mathbf{t}}_{\tau_n}(s)}(\overline{U}_{\tau_n}(s)) + S_{\tau_n}(r), w(r)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}r = 0.$$

Then, with the aid of the convergences (6.4.1), we are allowed to pass to the limit in the weak formulation obtaining

$$\int_0^T \langle u''(r) + \eta(r) + \mathrm{D}\mathcal{E}_s(u(s)) + B(t, u(r), u'(r)) - f(r), w(r)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}r = 0$$

for all $w \in \mathrm{L}^{\max\{2,q\}}(0,T;U \cap V)$. Then, by a density argument and the fundamental lemma of calculus of variations, we deduce

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

for a.e. $t \in (0,T)$. We shall identify the weak limit η 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 2.4.4 with $f_n(t,v) = \Psi_{\underline{U}_{\tau_n}(t)}(v)$ and $f(t,v) = \Psi_{u(t)}(v)$ for all $v \in X = V$ and $n \in \mathbb{N}$. Assumption (2.4.1) is already fulfilled by Condition (6.Ψc). Hence, it remains to show that

$$\limsup_{n\to\infty} \int_0^T \langle \eta_n(t), \overline{V}_{\tau_n}(t)\rangle_{V^*\times V} \,\mathrm{d}t \le \int_0^T \langle \eta(t), u'(t)\rangle_{V^*\times V} \,\mathrm{d}t. \tag{6.5.2}$$

In order to show the latter limes superior estimate, we use the fact that η_{τ_n} can be expressed through the remaining terms of the Euler–Lagrange equation (6.5.1). Therefore, we will split the integral on the left-hand side of (6.5.2) and note first that

$$\begin{aligned} &-\int_0^t \langle \widehat{V}_{\tau_n}'(r), \overline{V}_{\tau_n}(r)\rangle_{V^*\times V} \,\mathrm{d}r \\ &= -\int_0^t \langle \widehat{V}_{\tau_n}'(r), \widehat{V}_{\tau_n}(r)\rangle_{V^*\times V} \,\mathrm{d}r + \int_0^t \langle \widehat{V}_{\tau_n}'(r), \widehat{V}_{\tau_n}(r) - \overline{V}_{\tau_n}(r)\rangle_{V^*\times V} \,\mathrm{d}r \\ &= \frac{1}{2}|v_0|^2 - \frac{1}{2}|\widehat{V}_{\tau_n}(t)|^2 + \int_0^t \langle \widehat{V}_{\tau_n}'(r), \widehat{V}_{\tau_n}(r) - \overline{V}_{\tau_n}(r)\rangle_{V^*\times V} \,\mathrm{d}r \\ &\le \frac{1}{2}|v_0|^2 - \frac{1}{2}|\widehat{V}_{\tau_n}(t)|^2, \end{aligned}$$

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

$$
\begin{aligned}
&\int_0^t \langle \widehat{V}'_{\tau_n}(r), \widehat{V}_{\tau_n}(r) - \overline{V}_{\tau_n}(r)\rangle_{V^*\times V}\,\mathrm{d}r \\
&= \sum_{i=1}^{m-1}\int_{t_{i-1}}^{t_i}\left(\frac{V_{\tau_n}^i - V_{\tau_n}^{i-1}}{\tau_n}, V_{\tau_n}^i\frac{r-t_{i-1}}{\tau_n} + V_{\tau_n}^{i-1}\frac{t_i - r}{\tau_n} - V_{\tau_n}^i\right)\mathrm{d}r \\
&\quad + \int_{t_{m-1}}^{t}\left(\frac{V_{\tau_n}^m - V_{\tau_n}^{m-1}}{\tau_n}, V_{\tau_n}^m\frac{r-t_{m-1}}{\tau_n} + V_{\tau_n}^{m-1}\frac{t_m - r}{\tau_n} - V_{\tau_n}^m\right)\mathrm{d}r \\
&= -\sum_{i=1}^{m-1}\int_{t_{i-1}}^{t_i}\left(\frac{V_{\tau_n}^i - V_{\tau_n}^{i-1}}{\tau_n}, \left(V_{\tau_n}^i - V_{\tau_n}^{i-1}\right)\frac{t_i - r}{\tau_n}\right)\mathrm{d}r \\
&\quad - \int_{t_{m-1}}^{t}\left(\frac{V_{\tau_n}^m - V_{\tau_n}^{m-1}}{\tau_n}, \left(V_{\tau_n}^m - V_{\tau_n}^{m-1}\right)\frac{t_m - r}{\tau_n}\right)\mathrm{d}r \\
&= -\sum_{i=1}^{m-1}\int_{t_{i-1}}^{t_i}\frac{t_i - r}{\tau_n^2}\left|V_{\tau_n}^i - V_{\tau_n}^{i-1}\right|^2\mathrm{d}r - \int_{t_{m-1}}^{t}\frac{t_m - r}{\tau_n^2}\left|V_{\tau_n}^m - V_{\tau_n}^{m-1}\right|^2\mathrm{d}r \le 0
\end{aligned}
$$

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

We continue with the term involving the derivative of the energy functional and start with the linear part:

$$
\begin{aligned}
&-\int_0^t \langle \mathrm{E}\overline{U}_{\tau_n}(r), \overline{V}_{\tau_n}(r)\rangle_{U^*\times U}\,\mathrm{d}r \\
&= -\int_0^t \langle \mathrm{E}\widehat{U}_{\tau_n}(r), \overline{V}_{\tau_n}(r)\rangle_{U^*\times U}\,\mathrm{d}r + \int_0^t \langle \mathrm{E}\widehat{U}_{\tau_n}(r) - \mathrm{E}\overline{U}_{\tau_n}(r), \overline{V}_{\tau_n}(r)\rangle_{U^*\times U}\,\mathrm{d}r \\
&= \mathcal{E}^1(u_0) - \mathcal{E}^1(\widehat{U}_{\tau_n}(t)) + \int_0^t \langle \mathrm{E}(\widehat{U}_{\tau_n}(r) - \overline{U}_{\tau_n}(r)), \overline{V}_{\tau_n}(r)\rangle_{U^*\times U}\,\mathrm{d}r \\
&\le \mathcal{E}^1(u_0) - \mathcal{E}^1(\widehat{U}_{\tau_n}(t)),
\end{aligned}
$$

where we used

$$
\int_0^t \langle \mathrm{E}(\widehat{U}_{\tau_n}(r) - \overline{U}_{\tau_n}(r)), \overline{V}_{\tau_n}(r)\rangle_{U^*\times U}\,\mathrm{d}r \le 0,
$$

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 λ-convexity of $\mathcal{E}_t^2$ that

$$
\begin{aligned}
&-\int_0^t \langle \mathrm{D}\mathcal{E}^2_{\mathbf{t}_{\tau_n}(r)}(\overline{U}_{\tau_n}(r)), \overline{V}_{\tau_n}(r)\rangle_{U^*\times U}\,\mathrm{d}r \\
&= -\sum_{i=1}^{m-1}\langle \mathrm{D}\mathcal{E}^2_{t_i}(U_{\tau_n}^i), U_{\tau_n}^i - U_{\tau_n}^{i-1}\rangle_{U^*\times U} - \frac{t - t_{m-1}}{\tau_n}\langle \mathrm{D}\mathcal{E}^2_{t_m}(U_{\tau_n}^m), U_{\tau_n}^m - U_{\tau_n}^{m-1}\rangle_{U^*\times U} \\
&\le -\sum_{i=1}^{m-1}\left(\mathcal{E}^2_{t_i}(U_{\tau_n}^{i-1}) - \mathcal{E}^2_{t_i}(U_{\tau_n}^i) - \lambda\left|U_{\tau_n}^i - U_{\tau_n}^{i-1}\right|^2\right) \\
&\quad - \frac{t - t_{m-1}}{\tau_n}\left(\mathcal{E}^2_{t_m}(U_{\tau_n}^{m-1}) - \mathcal{E}^2_{t_m}(U_{\tau_n}^m) - \lambda\left|U_{\tau_n}^m - U_{\tau_n}^{m-1}\right|^2\right) \\
&= -\sum_{i=1}^{m}\left(\mathcal{E}^2_{t_{i-1}}(U_{\tau_n}^{i-1}) - \mathcal{E}_{t_i}(U_{\tau_n}^i) + \int_{t_{i-1}}^{t_i}\partial_r\mathcal{E}_r^2(U_{\tau_n}^i)\,\mathrm{d}r + \lambda\tau_n^2\left|V_{\tau_n}^i\right|^2\right)
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{t_m-t}{\tau_n}\left(\mathcal{E}^2_{t_m}(U^{m-1}_{\tau_n})-\mathcal{E}^2_{t_m}(U^m_{\tau_n})-\lambda\left|U^m_{\tau_n}-U^{m-1}_{\tau_n}\right|^2\right)\\
&=\mathcal{E}^2_0(u_0)-\mathcal{E}^2_{\bar{\mathbf{t}}_{\tau_n}(t)}(\overline{U}_{\tau_n}(t))+\int_0^{\bar{\mathbf{t}}_{\tau_n}(t)}\partial_r\mathcal{E}^2_r(\overline{U}_{\tau_n}(r))\,\mathrm{d}r+I_n(t),
\end{aligned}
$$

where

$$
I_n(t)=\frac{t_m-t}{\tau_n}\left(\mathcal{E}^2_{t_m}(U^{m-1}_{\tau_n})-\mathcal{E}^2_{t_m}(U^m_{\tau_n})-\lambda\left|U^m_{\tau_n}-U^{m-1}_{\tau_n}\right|^2\right)+\lambda\tau_n\int_0^{\bar{\mathbf{t}}_{\tau_n}(t)}|\overline{V}_{\tau_n}(r)|^2\,\mathrm{d}r.
$$

Now, we want to make use of the inequality (6.4.3). However, the aforementioned inequality only holds true for almost every $t\in(0,T)$. Therefore, we take a sequence of increasing values $(\beta_l)_{l\in\mathbb{N}}$, $\beta_i\in(0,T)$ for all $i\in\mathbb{N}$, converging to T for which (6.4.3) holds true. Then, choosing $t=\beta_l$, we obtain with the convergences (6.4.1b), (6.4.1h), (6.4.1e), and (6.4.1g), the sequential weak lower semicontinuity of $\mathcal{E}^1_t$ and $|\cdot|$ and the continuity of $\mathcal{E}^1_t$, the limes superior condition and growth condition (6.Ed) on $\partial_t\mathcal{E}^2_t$ and FATOU's Lemma that

$$
\begin{aligned}
&\limsup_{n\to\infty}-\int_0^{\beta_l}\langle\widehat{V}'_{\tau_n}(r)+\mathrm{D}\mathcal{E}_{\bar{\mathbf{t}}_{\tau_n}(r)}(\overline{U}_{\tau_n}(r)),\overline{V}_{\tau_n}(r)\rangle_{V^*\times V}\,\mathrm{d}r\\
&\le\limsup_{n\to\infty}\left(\frac12|v_0|^2-\frac12|\widehat{V}_{\tau_n}(\beta_l)|^2+\mathcal{E}_0(u_0)-\mathcal{E}^1(\widehat{U}_{\tau_n}(\beta_l))-\mathcal{E}^2_{\bar{\mathbf{t}}_{\tau_n}(\beta_l)}(\overline{U}_{\tau_n}(\beta_l))\right.\\
&\quad\left.+\int_0^{\bar{\mathbf{t}}_{\tau_n}(\beta_l)}\partial_r\mathcal{E}_r(\overline{U}_{\tau_n}(r))\,\mathrm{d}r+I_n(t)\right)\\
&\le\frac12|v_0|^2-\frac12|u'(\beta_l)|^2+\mathcal{E}_0(u_0)-\mathcal{E}_{\beta_l}(u(\beta_l))+\int_0^{\beta_l}\partial_r\mathcal{E}_r(u(r))\,\mathrm{d}r.
\end{aligned}
$$

Since $u\in\mathrm{C}_w([0,T];U)$ and $u'\in\mathrm{L}^\infty(0,T;H)\cap\mathrm{W}^{1,1}(0,T;U^*+V^*)\subset\mathrm{C}_w([0,T];H)$, Lemma 6.4.1 then yields

$$
\begin{aligned}
&\frac12|v_0|^2-\frac12|u'(\beta_l)|^2+\mathcal{E}_0(u_0)-\mathcal{E}_{\beta_l}(u(\beta_l))+\int_0^{\beta_l}\partial_r\mathcal{E}_r(u(r))\,\mathrm{d}r\\
&\le-\int_0^{\beta_l}\langle u''(r)+\mathrm{D}\mathcal{E}_r(u(r)),u'(r)\rangle_{V^*\times V}.
\end{aligned}
$$

Then, in view of the convergences (6.4.1m) and (6.4.1n), the EULER–LAGRANGE equation (6.5.1), we obtain

$$
\begin{aligned}
&\limsup_{n\to\infty}\int_0^{\beta_l}\langle\eta_n(t),\overline{V}_{\tau_n}(t)\rangle_{V^*\times V}\,\mathrm{d}t\\
&=\limsup_{n\to\infty}\int_0^{\beta_l}\langle S_{\tau_n}(t)-\widehat{V}'_{\tau_n}(t)-\mathrm{D}\mathcal{E}_{\bar{\mathbf{t}}_{\tau_n}(t)}(\overline{U}_{\tau_n}(t)),\overline{V}_{\tau_n}(t)\rangle_{V^*\times V}\,\mathrm{d}t\\
&\le\int_0^{\beta_l}\langle f(t)-B(t,u(t),u'(t))-u''(t)-\mathrm{D}\mathcal{E}_t(u(t)),u'(t)\rangle_{V^*\times V}\,\mathrm{d}t\\
&=\int_0^{\beta_l}\langle\eta(t),u'(t)\rangle_{U^*\times U}\,\mathrm{d}t.
\end{aligned}
$$

Together with Condition (6.Ψc) and Lemma 2.4.4, this implies $\eta(t)\in\partial_V\Psi_{u(t)}(u'(t))$ for almost every $t\in(0,\beta_l)$ for all $l\in\mathbb{N}$. Letting $l\to\infty$ leads to $\eta(t)\in\partial\Psi_{u(t)}(u'(t))$ for almost every $t\in(0,T)$. This shows for each $k\in\mathbb{N}$ the existence of a function

u satisfying the inclusion (6.1.3), 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 $\Psi_{u_k(t)}(u_k'(t))$. In the last step, we want to show that there exists a limit function u which satisfies (6.1.3) and (6.1.4) as well as the intial 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 $k \to \infty$. As in Chapter 5, the next steps are the following.

1. We derive a priori estimates based on the energy-dissipation inequality (6.1.4),
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 6.1.3 as $k \to \infty$.

Ad 1. Let $t \in [0,T]$ and $\mathcal{N} \subset (0,T]$ a set of measure zero such that $\mathcal{E}_{\bar{\mathbf{t}}_{\tau_n}(s)}(\overline{U}_{\tau_n}(s)) \to \mathcal{E}_t(u(s))$ and $\overline{V}_{\tau_n}(s) \to u'(s)$ for each $s \in [0,T]\backslash\mathcal{N}$. Then, employing the convergences (6.4.1), we obtain

$$
\begin{aligned}
&\frac{1}{2}|u_k'(t)|^2 + \mathcal{E}_t(u_k(t)) + \int_0^t \Big(\Psi_{u_k(r)}(u_k'(r)) + \Psi^*_{u_k(r)}(S_k(r) - \mathrm{D}\mathcal{E}_r(u_k(r)) - u_k''(r))\Big)\,\mathrm{d}r \\
&\le \liminf_{n\to\infty}\Bigg(\frac{1}{2}\left|\overline{V}_{\tau_n}(t)\right|^2 + \mathcal{E}_{\bar{\mathbf{t}}_{\tau_n}(t)}(\overline{U}_{\tau_n}(t)) \\
&\quad + \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \Big(\Psi_{\underline{U}_{\tau_n}(r)}(\overline{V}_{\tau_n}(r)) + \Psi^*_{\underline{U}_{\tau_n}(t)}\Big(S_{\tau_n}(r) - \mathrm{D}\mathcal{E}_{\bar{\mathbf{t}}_{\tau_n}(r)}(\overline{U}_{\tau_n}(r)) - \widehat{V}_{\tau_n}'(r)\Big)\Big)\,\mathrm{d}r\Bigg) \\
&\le \limsup_{n\to\infty}\Bigg(\frac{1}{2}\left|v_0^k\right|^2 + \mathcal{E}_0(u_0^k) + \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \partial_r\mathcal{E}_r(\underline{U}_{\tau_n}(r))\,\mathrm{d}r \\
&\quad + \int_0^{\bar{\mathbf{t}}_{\tau_n}(t)} \langle S_{\tau_n}(r), \overline{V}_{\tau_n}(r)\rangle_{V^*\times V}\,\mathrm{d}r + \tau\lambda\int_0^{\bar{\mathbf{t}}_\tau(t)} \|\overline{V}_\tau(r)\|_V^2\,\mathrm{d}r\Bigg) \\
&= \frac{1}{2}|v_0^k|^2 + \mathcal{E}_0(u_0^k) + \int_0^t \partial_r\mathcal{E}_r(u_k(r))\,\mathrm{d}r + \int_0^t \langle S_k(r), u_k'(r)\rangle_{V^*\times V}\,\mathrm{d}r
\end{aligned}
$$

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 (6.Ed), (6.Bb), and (6.Bb), we obtain with the lemma of Gronwall (Lemma A.1.1)

$$
\frac{1}{2}|u_k'(t)|^2 + \mathcal{E}_t(u_k(t)) + \int_0^t (\Psi(u_k'(r)) + \Psi^*(S_k(r) - \mathrm{D}\mathcal{E}_r(u_k(r)) - u_k''(r)))\,\mathrm{d}r \le C_B.
$$

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

$$\begin{aligned}
u_k &\overset{*}{\rightharpoonup} u && \text{in } \mathrm{L}^\infty(0,T;U), && \text{(6.5.3a)}\\
u_k - u_0^k &\overset{*}{\rightharpoonup} u - u_0 && \text{in } \mathrm{L}^\infty(0,T;V), && \text{(6.5.3b)}\\
u_k(t) &\rightharpoonup u(t) && \text{in } U \ \text{ for all } t \in [0,T], && \text{(6.5.3c)}\\
u_k(t) - u_0^k &\rightharpoonup u(t) - u_0 && \text{in } V \ \text{ for all } t \in [0,T], && \text{(6.5.3d)}\\
u_k &\to u && \text{in } \mathrm{L}^r(0,T;\widetilde{W}) \quad \text{for any } r \geq 1, && \text{(6.5.3e)}\\
u_k(t) &\to u(t) && \text{in } \widetilde{W} \ \text{ for all } t \in [0,T], && \text{(6.5.3f)}\\
u_k'(t) &\overset{*}{\rightharpoonup} u' && \text{in } \mathrm{L}^q(0,T;V) \cap \mathrm{L}^\infty(0,T;H), && \text{(6.5.3g)}\\
u_k'(t) &\to u' && \text{in } \mathrm{L}^p(0,T;H) \quad \text{for all } p \geq 1, && \text{(6.5.3h)}\\
u_k'(t) &\rightharpoonup u'(t) && \text{in } H \ \text{ for all } t \in [0,T], && \text{(6.5.3i)}\\
\eta_{\tau_n}^k &\rightharpoonup \eta && \text{in } \mathrm{L}^{q^*}(0,T;V^*), && \text{(6.5.3j)}\\
\mathrm{E}u_k &\rightharpoonup \mathrm{E}u && \text{in } \mathrm{L}^2(0,T;U^*), && \text{(6.5.3k)}\\
\mathrm{D}\mathcal{E}_t^2(u_k) &\to \mathrm{D}\mathcal{E}_t^2(u) && \text{in } \mathrm{L}^r(0,T;U^*) \quad \text{for any } r \geq 1, && \text{(6.5.3l)}\\
u_k'' &\rightharpoonup u'' && \text{in } \mathrm{L}^{\min\{2,q^*\}}(0,T;U^* + V^*), && \text{(6.5.3m)}\\
B(\cdot, u_k, u_k') &\to B(\cdot, u, u') && \text{in } \mathrm{L}^{r^*}(0,T;V^*), && \text{(6.5.3n)}
\end{aligned}$$

and if Ψ_u satisfies (6.Ψd), then

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

Ad 3. Therefore, $u \in \mathrm{C}_w([0,T];U) \cap \mathrm{W}^{1,\infty}([0,T];H) \cap \mathrm{W}^{2,q^*}(0,T;U^* + V^*)$ with $u - u_0 \in \mathrm{W}^{1,q}(0,T;V)$ and $\eta \in \mathrm{L}^{q^*}(0,T;V^*)$ 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 (6.Ψc) and Lemma 2.4.4 that u and η satisfy the inclusions (6.1.3). It remains to show the energy-dissipation balance (6.1.4). The inequality

$$\begin{aligned}
&\frac{1}{2}|u'(t)|^2 + \mathcal{E}_t(u(t)) + \int_0^t \Big(\Psi_{u(r)}(u'(r)) + \Psi_{u(r)}^*(S(r) - \mathrm{D}\mathcal{E}_r(u(r)) - u''(r))\Big)\,\mathrm{d}r\\
&\leq \frac{1}{2}|v_0|^2 + \mathcal{E}_0(u_0) + \int_0^t \partial_r \mathcal{E}_r(u(r))\,\mathrm{d}r + \int_0^t \langle S(r), u'(r)\rangle_{V^* \times V}\,\mathrm{d}r,
\end{aligned}$$

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 (6.4.1). Then, employing again (6.4.3) and the Fenchel–Young inequality, we obtain

$$\begin{aligned}
&\int_0^t \Big(\Psi_{u(r)}(u'(r)) + \Psi_{u(r)}^*(S(r) - \mathrm{D}\mathcal{E}_r(u(r)) - u''(r))\Big)\,\mathrm{d}r\\
&\leq \frac{1}{2}|v_0|^2 - \frac{1}{2}|u'(t)|^2 + \mathcal{E}_0(u_0) - \mathcal{E}_T(u(t)) + \int_0^t \partial_r \mathcal{E}_r(u(r))\,\mathrm{d}r\\
&\quad + \int_0^t \langle S(r), u'(r)\rangle_{V^* \times V}\,\mathrm{d}r\\
&\leq \int_0^t \langle \mathrm{D}\mathcal{E}_r(u(r)) - u''(r), u'(r)\rangle_{V^* \times V}\,\mathrm{d}r + \int_0^t \langle S(r), u'(r)\rangle_{V^* \times V}\,\mathrm{d}r
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^t \langle S(r) - \mathrm{D}\mathcal{E}_r(u(r)) - u''(r), u'(r)\rangle_{V^*\times V}\,\mathrm{d}r \\
&\le \int_0^t \Big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}(S(r) - \mathrm{D}\mathcal{E}_r(u(r)) - u''(r))\Big)\,\mathrm{d}r
\end{aligned}
$$

for almost every $t \in (0,T)$. Now, if $V \hookrightarrow U$, then the inequality (6.4.3) indeed holds as equality for all $t \in [0,T]$ by the classical integration by parts formula. This shows (6.1.4), and hence the completion of the proof. □

Remark 6.5.1 The proof of Theorem 6.1.4 reveals that one can consider dissipation potentials that depend on a parameter ε. In this case, the Condition (6.Ψa) is assumed to hold for every $\varepsilon \ge 0$ while Condition (6.Ψb) holds uniformly in $\varepsilon \ge 0$. Condition (6.Ψc) can either be replaced with the MOSCO-convergence $\Psi^{\varepsilon_n}_{u_n} \xrightarrow{\mathrm{M}} \Psi^0_u$ for every sequence $u_n \rightharpoonup u$ as $\varepsilon \searrow 0$, or with a more general liminf estimate (2.4.1).

Chapter 7

Applications

In this section, we want to apply the abstract results on linear and nonlinear inertial systems developed and proven in Chapter 5 and 6, respectively, to concrete examples. We will give a sufficient number of examples to cover the range of applications from the abstract results. Since our main results are established in a nonsmooth setting, the examples with nonsmooth functionals, in particular, can not be cast into the framework of the existing results. Those nonsmooth functionals correspond to multi-valued equations or nonlinear constraints. We first start with examples for linearly damped inertial systems and continue with examples for nonlinearly damped inertial systems. We assume the same notation and function spaces as in Chapter 4.

7.1 Differential inclusion I A

In the first example, we consider a system, which can be treated in the Case **(a)** of the linearly damped intertial system, where the dissipation potential is given by the DIRICHLET energy and the energy functional is a nonsmooth λ-convex function which to the best of the authors' knowledge can not be treated with the abstract results known thus far. More precisely, we consider the initial-boundary value problem

$$\text{(P1)}\begin{cases} \partial_{tt}\boldsymbol{u} - \Delta\partial_t\boldsymbol{u} - \Delta_p\boldsymbol{u} + (|\boldsymbol{u}|^2 - 1)\boldsymbol{u} - \nabla\cdot\boldsymbol{p} + \boldsymbol{b}\,(\boldsymbol{x}, t, \boldsymbol{u}, \partial_t\boldsymbol{u}) = \boldsymbol{f} \quad \text{in } \Omega_T, \\ \boldsymbol{p}(\boldsymbol{x}, t) \in \mathbf{Sgn}\,(\nabla\boldsymbol{u}(\boldsymbol{x}, t)) \quad \text{a.e. in } \Omega_T, \\ \boldsymbol{u}(\boldsymbol{x}, 0) \ = \boldsymbol{u}_0(\boldsymbol{x}) \quad \text{on } \Omega, \\ \boldsymbol{u}'(\boldsymbol{x}, 0) = \boldsymbol{v}_0(\boldsymbol{x}) \quad \text{on } \Omega, \\ \boldsymbol{u}(\boldsymbol{x}, t) \ = 0 \qquad \text{on } \partial\Omega \times [0, T], \end{cases}$$

where $\mathbf{Sgn} : \mathbb{R}^{d\times m} \rightrightarrows \mathbb{R}^{d\times m}$ is the sign function defined in (4.0.1), $\boldsymbol{f} : \Omega \to \mathbb{R}^m$, $\boldsymbol{b} : \Omega \times [0, T] \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ a CARATHÉODORY function in the sense that $\boldsymbol{b}(\boldsymbol{x}, \cdot, \cdot, \cdot)$ is continuous for almost every $\boldsymbol{x} \in \Omega$ and $\boldsymbol{b}(\cdot, t, \boldsymbol{y}, \boldsymbol{z})$ is measurable for all $t \in [0, T]$ and $\boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^m$. Furthermore, $\boldsymbol{b}$ is assumed to satisfy the following growth condition: there exists a constant $C_b > 0$ and numbers $q, r > 1$ such that

$$|\boldsymbol{b}(\boldsymbol{x}, t, \boldsymbol{u}, \boldsymbol{v})| \leq C_b(1 + |\boldsymbol{u}|^{q-1} + |\boldsymbol{v}|^{r-1}) \quad \text{for a.e. } \boldsymbol{x} \in \Omega, t \in [0, T] \text{ and all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^m.$$

Here, $p, q, r \geq 1$ are to be chosen in accordance with the assumptions.

Choosing the spaces $U = \mathrm{W}_0^{1,p}(\Omega)^m \cap \mathrm{L}^4(\Omega)^m, V = \mathrm{H}_0^1(\Omega)^m$, $W = \mathrm{L}^{\max\{2,q\}}(\Omega)^m$ and $H = \mathrm{L}^2(\Omega)^m$ equipped with the standard norms, we assume $\boldsymbol{f} \in \mathrm{L}^2(0,T;V^*)$. The energy functional $\mathcal{E} : V \to (-\infty, +\infty]$ and the dissipation potential $\Psi : V \to \mathbb{R}$ are given by

$$\mathcal{E}(u) = \begin{cases} \int_\Omega \left(\frac{1}{p}|\nabla \boldsymbol{u}(\boldsymbol{x})|^p + |\nabla \boldsymbol{u}(\boldsymbol{x})| + \frac{1}{4}(|\boldsymbol{u}(\boldsymbol{x})|^2 - 1)^2\right) \mathrm{d}\boldsymbol{x} & \text{if } \boldsymbol{u} \in \mathrm{dom}(\mathcal{E}), \\ +\infty \quad \text{otherwise}, \end{cases}$$

and

$$\Psi(v) = \frac{1}{2}\int_\Omega |\nabla \boldsymbol{v}(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x},$$

respectively, whereas the perturbation $B : [0,T] \times W \times H \to V^*$ is defined by

$$\langle B(t,\boldsymbol{u},\boldsymbol{v}),\boldsymbol{w}\rangle_{V^*\times V} = \langle B(t,\boldsymbol{u},\boldsymbol{v}),\boldsymbol{w}\rangle_{W^*\times W} = \int_\Omega \boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{u}(\boldsymbol{x}),\boldsymbol{v}(\boldsymbol{x}))\cdot\boldsymbol{w}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

The Legendre–Fenchel transformation $\Psi^* : \mathrm{H}^{-1}(\Omega)^m \to \mathbb{R}$ of Ψ is obviously given by $\Psi^*(\boldsymbol{\xi}) = \frac{1}{2}\|\boldsymbol{\xi}\|_{-1,2}^2$. Furthermore, it is readily seen that the energy functional is not Gâteaux differentiable and its effective domain is given by $\mathrm{dom}(\mathcal{E}) = \mathrm{W}_0^{1,p}(\Omega)^m \cap \mathrm{L}^4(\Omega)^m$. The values $p,q,r \geq 1$ are to be chosen such that all assumptions are fulfilled. We can choose, e.g.,

$$\begin{aligned} &d = 1,\ p \in (1,+\infty), r \in [1,2], q \in [1,p/2+1], \\ &d = 2,\ p \in (1,+\infty), r \in [1,2], q \in \begin{cases} [1,pd/(p-d)) \cap [1,p/2+1] & \text{if } p \in (1,2), \\ [1,p/2+1] \quad \text{if } p \geq 2, \end{cases} \\ &d \geq 3,\ p \in (1,+\infty), r \in [1,2], q \in \begin{cases} [1,q^*) \quad \text{if } p \in (1,2), \\ [1,p/2+1] \quad \text{if } p \geq 3, \end{cases} \end{aligned}$$

where $q^* = \min\{\frac{d(p+2)}{2(d-p)}, \frac{3d+4}{d}, p+1\}$. Then, by the Sobolev embedding theorem and the Rellich–Kondrachov theorem, U and V are densely, continuously and compactly embedded in W and H, respectively. We will verify for illustration the assumptions for the case $d \geq 3$. Since the dissipation potential is state-independent, it is induced by the bilinear form $a : V \times V \to \mathbb{R}$,

$$a(\boldsymbol{v},\boldsymbol{w}) := \frac{1}{2}\int_\Omega \nabla\boldsymbol{v}\cdot\nabla\boldsymbol{w}\,\mathrm{d}\boldsymbol{x},$$

and therefore satisfies all conditions. The conditions (5.Eb)-(5.Ed) are obviously fulfilled by the energy functional. In order to verify (5.Ea), we note that every convex and lower semicontinuous functional on a Banach space is weakly lower semicontinuous. Taking the latter into account, we observe that for $u \in \mathrm{dom}(\mathcal{E})$, the energy functional

$$\begin{aligned}\mathcal{E}(\boldsymbol{u}) &= \int_{\Omega}\left(\frac{1}{p}|\nabla\boldsymbol{u}(\boldsymbol{x})|^p + |\nabla\boldsymbol{u}(\boldsymbol{x})| + \frac{1}{4}(|\boldsymbol{u}(\boldsymbol{x})|^2-1)^2\right)\mathrm{d}\boldsymbol{x}\\ &= \int_{\Omega}\left(\frac{1}{p}|\nabla\boldsymbol{u}(\boldsymbol{x})|^p + |\nabla\boldsymbol{u}(\boldsymbol{x})| + \frac{1}{4}(|\boldsymbol{u}(\boldsymbol{x})|^4-2|\boldsymbol{u}(\boldsymbol{x})|^2+1)\right)\mathrm{d}\boldsymbol{x}\\ &= \int_{\Omega}\left(\frac{1}{p}|\nabla\boldsymbol{u}(\boldsymbol{x})|^p + |\nabla\boldsymbol{u}(\boldsymbol{x})| + \frac{1}{4}(|\boldsymbol{u}(\boldsymbol{x})|^4+1)\right)\mathrm{d}\boldsymbol{x} - \frac{1}{2}\int_{\Omega}|\boldsymbol{u}(\boldsymbol{x})|^2\,\mathrm{d}\boldsymbol{x}\\ &= \mathcal{W}(\boldsymbol{u}) - \frac{1}{2}\int_{\Omega}|\boldsymbol{u}(\boldsymbol{x})|^2\,\mathrm{d}\boldsymbol{x}\end{aligned}$$

is the sum of a convex function $\mathcal{W}$ and a concave function $u \mapsto -\frac{1}{2}\int_{\Omega}|\boldsymbol{u}(\boldsymbol{x})|^2\,\mathrm{d}\boldsymbol{x}$ on V. The lower semicontinuity of $\mathcal{W}$ on V follows immediately from the converse of the dominated convergence theorem (see, e.g., Brézis [35, Theorem 4.9, p. 94]) and Fatou's lemma. Further, due to the compact embedding of V in H, the concave function is continuous on V 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 H squared which by the parallelogram law and the embedding $V \hookrightarrow H$ leads to the λ-convexity of $\mathcal{E}$ with $\lambda := C$ being the constant of the very same embedding. Now, we show the closedness property (5.Ee). First, we note that for each $\boldsymbol{u} \in D(\partial\mathcal{E})$, there holds $\boldsymbol{\xi} \in \partial_U\mathcal{E}(\boldsymbol{u}) = \partial_U\mathcal{W}(\boldsymbol{u}) - \boldsymbol{u}$ if and only if $\boldsymbol{\xi} = -\Delta_p\boldsymbol{u} + \nabla\cdot\boldsymbol{p} + (|\boldsymbol{u}|^2-1)\boldsymbol{u} \in U^*$ for a measurable selection $\boldsymbol{p} \in \mathrm{L}^\infty(\Omega)^{d\times m}$ satisfying $\boldsymbol{p} \in \mathbf{Sgn}(\nabla\boldsymbol{u})$ a.e. in Ω. This can be seen as follows: we define the functionals $\mathcal{W}_1 : W_0^{1,p}(\Omega)^m \to [0,+\infty]$ and $\mathcal{W}_2 : \mathrm{L}^1(\Omega)^{d\times m} \to \mathbb{R}$ as well as the operator $\Lambda : W_0^{1,p}(\Omega)^m \to \mathrm{L}^1(\Omega)^{d\times m}$ via

$$\mathcal{W}_1(u) = \begin{cases}\int_{\Omega}\left(\frac{1}{p}|\nabla\boldsymbol{u}(\boldsymbol{x})|^p + \frac{1}{4}(|\boldsymbol{u}(\boldsymbol{x})|^2-1)^2\right)\mathrm{d}\boldsymbol{x} & \text{if } \boldsymbol{u} \in \mathrm{dom}(\mathcal{E}),\\ +\infty \quad \text{otherwise,}\end{cases}$$

$$\mathcal{W}_2(\boldsymbol{A}) = \int_{\Omega}|\boldsymbol{A}(x)|\,\mathrm{d}\boldsymbol{x}$$

and $\Lambda u = \nabla u$. We note that Λ is linear and bounded and has as adjoint operator $\Lambda^* : \mathrm{L}^\infty(\Omega)^{d\times m} \to W^{-1,p^*}(\Omega)^m, \boldsymbol{A} \mapsto -\nabla\cdot\boldsymbol{A}$ the divergence operator. Let $\boldsymbol{u} \in \mathrm{dom}(\partial_U\mathcal{W})$, then by the variational sum rule (Lemma2.2.7) and Lemma 2.2.8, there holds

$$\begin{aligned}\boldsymbol{\xi} &\in \partial_U\left(\mathcal{W}_1(\boldsymbol{u}) + \mathcal{W}_2(\Lambda\boldsymbol{u})\right)\\ &= \partial_U\mathcal{W}_1(\boldsymbol{u}) + \partial_U\mathcal{W}_2(\Lambda\boldsymbol{u})\\ &= \partial_U\mathcal{W}_1(\boldsymbol{u}) + \Lambda^*\partial_X\mathcal{W}_2(\Lambda\boldsymbol{u}),\end{aligned}$$

where $X = \mathrm{L}^1(\Omega)^{m\times d}$. Thus, there exists $\boldsymbol{\xi}_1 \in \partial_U\mathcal{W}_1(\boldsymbol{u})$ and $\boldsymbol{\xi}_2 \in +\Lambda^*\partial_X\mathcal{W}_2(\Lambda\boldsymbol{u})$ such that $\boldsymbol{\xi} = \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2$. Now, we shall determine $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$. Since $\mathcal{W}_1$ is Gâteaux differentiable on $\mathrm{W}^{1,r}(\Omega)^m \cap \mathrm{L}^4(\Omega)^m$, we deduce immediately $\boldsymbol{\xi}_1 = -\Delta_p\boldsymbol{u} + (|\boldsymbol{u}|^2-1)\boldsymbol{u}$ a.e. in Ω. In order to determine ξ_2, we note first that $\boldsymbol{\xi}_2 = \nabla\cdot\boldsymbol{p}$ with $\boldsymbol{p} \in \partial_X\mathcal{W}_2(\Lambda\boldsymbol{u})$.

Second, we express $\partial_X \mathcal{W}_2(\Lambda \boldsymbol{u})$ with the aid of Lemma 2.3.1 equivalently through the equation

$$\langle \boldsymbol{p}, \Lambda \boldsymbol{u} \rangle_{X^* \times X} = \mathcal{W}_2(\Lambda \boldsymbol{u}) + \mathcal{W}_2^*(\boldsymbol{p}). \tag{7.1.1}$$

Third, by Ekeland & Temam [69, Proposition 1.2, p. 87], the conjugate $\mathcal{W}_2^*$ is given by

$$\mathcal{W}_2^*(\boldsymbol{B}) = \int_\Omega \imath_{\overline{B}_{\mathbb{R}^{m\times d}}(0,1)}(\boldsymbol{B}(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x},$$

with the indicator function $\imath_{\overline{B}_{\mathbb{R}^{m\times d}}(0,1)} \to \{0, +\infty\}$ defined by

$$\imath_{\overline{B}_{\mathbb{R}^{m\times d}}(0,1)}(\boldsymbol{A}) = \begin{cases} 0 & \text{if } |\boldsymbol{A}| \leq 1 \\ +\infty & \text{otherweise.} \end{cases}$$

This implies

$$\mathcal{W}_2^*(\boldsymbol{B}) = \begin{cases} 0 & \text{if } |\boldsymbol{B}(\boldsymbol{x})| \leq 1 \quad \text{a.e. in } \Omega \\ +\infty & \text{otherweise.} \end{cases}$$

Inserting the latter expression into the equality (7.1.1), we obtain

$$\int_\Omega \boldsymbol{p}(\boldsymbol{x}) : \nabla \boldsymbol{u}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_\Omega |\boldsymbol{p}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}$$

and $|\boldsymbol{p}(\boldsymbol{x})| \leq 1$ a.e. in Ω. Since $\boldsymbol{p}(\boldsymbol{x}) : \nabla \boldsymbol{u}(\boldsymbol{x}) \leq |\boldsymbol{p}(\boldsymbol{x})|$ by the Fenchel–Young (or Cauchy–Schwarz) inequality, we deduce

$$\boldsymbol{p}(\boldsymbol{x}) : \nabla \boldsymbol{u}(\boldsymbol{x}) = |\boldsymbol{p}(\boldsymbol{x})| \quad \text{a.e. in } \Omega.$$

Therefore, $\boldsymbol{p}(\boldsymbol{x}) \in \overline{B}_{\mathbb{R}^{m\times d}}(0,1)$ if $|\nabla \boldsymbol{u}(\boldsymbol{x})| = 0$ and $\boldsymbol{p}(\boldsymbol{x}) = \frac{\nabla \boldsymbol{u}(\boldsymbol{x})}{|\nabla \boldsymbol{u}(\boldsymbol{x})|}$ otherwise. We obtain $\boldsymbol{p}(\boldsymbol{x}) \in \mathbf{Sgn}(\nabla \boldsymbol{u}(\boldsymbol{x}))$ a.e. in Ω. Now let $\boldsymbol{u}_n \overset{*}{\rightharpoonup} \boldsymbol{u}$ in $\mathrm{L}^\infty(0,T;U) \cap \mathrm{H}^1(0,T;V)$, $\boldsymbol{u}_n \to \boldsymbol{u}$ in $\mathrm{L}^2(0,T;V)$, and $\boldsymbol{\xi}_n \overset{*}{\rightharpoonup} \boldsymbol{\xi}$ in $\mathrm{L}^2(0,T;U^*)$ as $n \to \infty$ such that $\boldsymbol{\xi}_n(t) \in \partial \mathcal{E}(\boldsymbol{u}_n(t))$ for almost every $t \in (0,T)$, $\sup_{n\in\mathbb{N}, t\in[0,T]} \mathcal{E}(\boldsymbol{u}_n(t)) \leq C_2$, and

$$\limsup_{n\to\infty} \int_0^T \langle \boldsymbol{\xi}_n(t), \boldsymbol{u}_n(t) \rangle_{U^*\times U}\,\mathrm{d}t \leq \int_0^T \langle \boldsymbol{\xi}(t), \boldsymbol{u}(t) \rangle_{U^*\times U}\,\mathrm{d}t. \tag{7.1.2}$$

We note that we can decompose $\boldsymbol{\xi} = \boldsymbol{\zeta} - \boldsymbol{u} \in V^*$ with $\boldsymbol{\zeta} \in \partial \mathcal{W}(\boldsymbol{u})$. Then, defining $\boldsymbol{\zeta}_n := \boldsymbol{\xi}_n + \boldsymbol{u}_n$, there holds $\boldsymbol{\zeta}_n \in \partial W(\boldsymbol{u}_n)$ and $\boldsymbol{\zeta}_n \rightharpoonup \boldsymbol{\zeta} := \boldsymbol{\xi} + \boldsymbol{u}$ in $\mathrm{L}^2(0,T;U^*)$. By the Lions–Aubin lemma, we obtain the strong convergence of $\boldsymbol{u}_n \to \boldsymbol{u}$ in $\mathrm{C}([0,T];H)$. Thus, in view of (7.1.2), we deduce

$$\limsup_{n\to\infty} \int_0^T \langle \boldsymbol{\zeta}_n(t), \boldsymbol{u}_n(t) \rangle_{U^*\times U}\,\mathrm{d}t \leq \int_0^T \langle \boldsymbol{\zeta}(t), \boldsymbol{u}(t) \rangle_{U^*\times U}\,\mathrm{d}t.$$

Since W is convex, by Theorem 2.3.7, there holds $\boldsymbol{\zeta}(t) \in \partial_U W(\boldsymbol{u}(t))$ in U^* and $W(\boldsymbol{u}_n(t)) \to W(\boldsymbol{u}(t))$ as $n \to \infty$ a.e. in $(0,T)$, whence $\boldsymbol{\xi}(t) \in \partial_U \mathcal{E}(\boldsymbol{u}(t))$ a.e. in $(0,T)$ and $\mathcal{E}(\boldsymbol{u}_n(t)) \to \mathcal{E}(\boldsymbol{u}(t))$ as $n \to \infty$ a.e. in $(0,T)$. We proceed with showing the control of the subgradient of $\mathcal{E}$, i.e., Condition (5.Eg). Let $\boldsymbol{u} \in D(\partial\mathcal{E})$ and $\boldsymbol{\xi} \in \partial_U \mathcal{E}(\boldsymbol{u})$. Then, by HÖLDER's and YOUNG's inequality, the SOBOLEV embedding theorem, we obtain

$$\begin{aligned}
\langle \boldsymbol{\xi}, \boldsymbol{v}\rangle_{U^*\times U} &= \int_\Omega \Big(|\nabla \boldsymbol{u}|^{p-2}\nabla\boldsymbol{u} : \nabla \boldsymbol{v} + \boldsymbol{p} : \nabla\boldsymbol{v} + (|\boldsymbol{u}|^2 - 1)\boldsymbol{u}\cdot\boldsymbol{v}\Big)\,\mathrm{d}\boldsymbol{x} \\
&\le C\left(\|\boldsymbol{u}\|^{p-1}_{\mathrm{W}^{1,p}_0(\Omega)^m} + \|\boldsymbol{p}\|_{\mathrm{L}^\infty(\Omega)^{m\times d}}\right)\|\boldsymbol{v}\|_{\mathrm{W}^{1,p}_0(\Omega)^m} \\
&\quad + \|\boldsymbol{u}\|^3_{\mathrm{L}^4(\Omega)^m}\|\boldsymbol{v}\|_{\mathrm{L}^4(\Omega)^m} + \|\boldsymbol{u}\|_{\mathrm{L}^2(\Omega)^m}\|\boldsymbol{v}\|_{\mathrm{L}^2(\Omega)^m} \\
&\le C\left(\|\boldsymbol{u}\|^{p-1}_{\mathrm{W}^{1,p}_0(\Omega)^m} + \|\boldsymbol{p}\|_{\mathrm{L}^\infty(\Omega)^{m\times d}} + \|\boldsymbol{u}\|^3_{\mathrm{L}^4(\Omega)^m} + \|\boldsymbol{u}\|_{\mathrm{L}^2(\Omega)^m}\right)\|\boldsymbol{v}\|_U \\
&\le C\left(1 + \frac{1}{p}\|\boldsymbol{u}\|^p_{\mathrm{W}^{1,p}_0(\Omega)^m} + \frac{1}{4}\|\boldsymbol{u}\|^4_{\mathrm{L}^4(\Omega)^m} + \frac{1}{2}\|\boldsymbol{u}\|^2_{\mathrm{L}^2(\Omega)^m}\right)\|\boldsymbol{v}\|_U \\
&\le C\left(1 + \frac{1}{p}\|\boldsymbol{u}\|^p_{\mathrm{W}^{1,p}_0(\Omega)^m} + \frac{1}{4}\|\boldsymbol{u}\|^4_{\mathrm{L}^4(\Omega)^m} - \frac{1}{2}\|\boldsymbol{u}\|^2_{\mathrm{L}^2(\Omega)^m} + \|\boldsymbol{u}\|^2_{\mathrm{L}^2(\Omega)^m}\right)\|\boldsymbol{v}\|_U \\
&\le C\left(1 + \mathcal{E}(\boldsymbol{u}) + \|\boldsymbol{u}\|_{\mathrm{L}^2(\Omega)^m}\right)\|\boldsymbol{v}\|_U \\
&\le C\left(1 + \mathcal{E}(\boldsymbol{u}) + \|\boldsymbol{v}\|_{\mathrm{W}^{1,p}_0(\Omega)^m}\right)\|\boldsymbol{v}\|_U
\end{aligned}$$

for all $\boldsymbol{v} \in U = \mathrm{W}^{1,p}_0(\Omega)^m \cap \mathrm{L}^4(\Omega)^m$, where we also used the fact that **Sgn** is uniformly bounded, from which (5.Eg) follows. Finally, we verify the assumptions on the perturbation B. The continuity condition (5.Ba) can easily be checked with the dominated convergence theorem.
Ad (5.Bb). Let $\boldsymbol{u} \in \mathrm{dom}(\mathcal{E})$ and $v, w \in V$. Then, by the HÖLDER & YOUNG inequalities as well as the SOBOLEV embedding theorem, there holds

$$\begin{aligned}
\langle B(\boldsymbol{u},\boldsymbol{v}), \boldsymbol{w}\rangle_{V^*\times V} &= \int_\Omega \boldsymbol{b}(\boldsymbol{x}, t, \boldsymbol{u}(\boldsymbol{x}), \boldsymbol{v}(\boldsymbol{x}))\cdot\boldsymbol{w}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&\le C_b\int_\Omega (|\boldsymbol{u}(\boldsymbol{x})|^{q-1} + |\boldsymbol{v}(\boldsymbol{x})|^{r-1})|\boldsymbol{w}(\boldsymbol{x})|\,\mathrm{d}x \\
&\le C\left(\|\boldsymbol{u}\|^{q-1}_{\mathrm{L}^{(q-1)2d/(d+2)}(\Omega)^m} + \|\boldsymbol{v}\|^{r-1}_{\mathrm{L}^{(r-1)2d/(d+2)}(\Omega)^m}\right)\|\boldsymbol{w}\|_{\mathrm{L}^{(2d/(d-2)}(\Omega)^m} \\
&\le C\left(\|\boldsymbol{u}\|^{q-1}_{\mathrm{W}^{1,p}_0(\Omega)^m} + \|\boldsymbol{v}\|^{r-1}_{\mathrm{L}^2(\Omega)^m}\right)\|\boldsymbol{w}\|_{\mathrm{H}^1_0(\Omega)^m} \\
&\le C\left((1+\mathcal{E}(\boldsymbol{u})) + \|\boldsymbol{v}\|^2_{\mathrm{L}^2(\Omega)^m}\right)^{\frac{1}{2}}\|\boldsymbol{w}\|_{\mathrm{H}^1_0(\Omega)^m},
\end{aligned} \tag{7.1.3}$$

where $\tilde{c} \in (0,1)$. Recalling that the conjugate is given by $\Psi^*(\boldsymbol{\xi}) = \frac{1}{2}\|\boldsymbol{\xi}\|^2_{V^*}$ for all $\boldsymbol{\xi} \in V^* = \mathrm{H}^{-1}(\Omega)^m$, we conclude (5.Bb). Therefore, for every initial datum $\boldsymbol{u}_0 \in \mathrm{dom}(\mathcal{E}) = \mathrm{W}^{1,p}_0(\Omega)^m \cap \mathrm{L}^4(\Omega)^m$ and $\boldsymbol{v}_0 \in \mathrm{L}^2(\Omega)^m$, there exists a weak solution

$$\boldsymbol{u} \in \mathrm{L}^\infty(0,T;U) \cap \mathrm{H}^1(0,T;V) \cap \mathrm{W}^{1,\infty}(0,T;H) \cap \mathrm{H}^2(0,T;U^*)$$

to (P1) in the sense that

$$\begin{aligned}
\int_0^T \Big(\langle \boldsymbol{u}'', \boldsymbol{v}\rangle_{U^*\times U} + \int_\Omega \Big(&\nabla\partial_t\boldsymbol{u} : \nabla\boldsymbol{v} + |\nabla\boldsymbol{u}|^{p-2}\nabla\boldsymbol{u} : \nabla\boldsymbol{v} + (|\boldsymbol{u}|^2-1)\boldsymbol{u}\cdot\boldsymbol{v} + \boldsymbol{p} : \nabla\boldsymbol{v} \\
&+ \boldsymbol{b}(\boldsymbol{x}, t, \partial_t\boldsymbol{u}, \boldsymbol{u})\Big)\,\mathrm{d}x + \langle \boldsymbol{f}, \boldsymbol{v}\rangle_{U^*\times U}\Big)\,\mathrm{d}t \quad \text{for all } \boldsymbol{v} \in \mathrm{L}^2(0,T;U)
\end{aligned}$$

with $\boldsymbol{p}(\boldsymbol{x},t) \in \mathbf{Sgn}(\nabla \boldsymbol{u}(\boldsymbol{x},t))$ a.e. in Ω_T, and the energy-dissipation inequality

$$\begin{aligned}
&\frac{1}{2}\|\boldsymbol{u}'(t)\|^2_{\mathrm{L}^2(\Omega)} + \mathcal{E}(\boldsymbol{u}(t)) + \int_s^t \Psi(\boldsymbol{u}'(r))\,\mathrm{d}r \\
&+ \int_s^t \Psi^*(\boldsymbol{f}(r) - \boldsymbol{b}(r,\boldsymbol{u}(r),\boldsymbol{u}'(r)) - \boldsymbol{u}''(r) - \boldsymbol{\xi}(r))\,\mathrm{d}r \\
&\le \frac{1}{2}\|\boldsymbol{u}'(s)\|^2_{\mathrm{L}^2(\Omega)} + \mathcal{E}(\boldsymbol{u}(s)) + \int_s^t \langle \boldsymbol{f}(r) - \boldsymbol{b}(r,\boldsymbol{u}(r),\boldsymbol{u}'(r)), \boldsymbol{u}'(r)\rangle_{V^*\times V}\,\mathrm{d}r
\end{aligned}$$

holds for all $t \in [0,T]$ if $s = 0$ and a.e. $s \in (0,t)$, where $\boldsymbol{\xi} \in \mathrm{L}^\infty(0,T;U^*)$ and $\boldsymbol{\xi}(t) = -\Delta_p \boldsymbol{u}(t) - (|\boldsymbol{u}(t)|^2 - 1)\boldsymbol{u}(t)$ in $U^* = \mathrm{W}^{-1,p^*}(\Omega)^m + \mathrm{L}^{\frac{4}{3}}(\Omega)^m$ a.e. in $t \in (0,T)$.

7.2 Differential inclusion I B

In the following example, we will cover Case **(b)** while also highlighting the difference between Case **(a)** and **(b)**. To do so, we consider the initial-boundary value problem

$$\text{(P2)}\begin{cases}
\partial_{tt}\boldsymbol{u} - \Delta\partial_t \boldsymbol{u} + \psi'(\boldsymbol{x},\partial_t\boldsymbol{u}) - \Delta_p \boldsymbol{u} + (|\boldsymbol{u}|^2 - 1)\boldsymbol{u} - \nabla\cdot\boldsymbol{p} + \boldsymbol{b}(\boldsymbol{x},t,\boldsymbol{u}) = \boldsymbol{f} & \text{in } \Omega_T, \\
\boldsymbol{p}(\boldsymbol{x},t) \in \mathbf{Sgn}(\nabla\boldsymbol{u}(\boldsymbol{x},t)) \quad \text{a.e. in } \Omega_T, \\
\boldsymbol{u}(\boldsymbol{x},0) = \boldsymbol{u}_0(\boldsymbol{x}) \quad \text{on } \Omega, \\
\boldsymbol{u}'(\boldsymbol{x},0) = \boldsymbol{v}_0(\boldsymbol{x}) \quad \text{on } \Omega, \\
\boldsymbol{u}(\boldsymbol{x},t) = 0 \quad \text{on } \partial\Omega \times [0,T].
\end{cases}$$

where we assumed that the contribution of $\partial_t \boldsymbol{u}$ in the perturbation from the first example is variational, i.e., it has a potential $\psi : \Omega\times\mathbb{R}^m \to \mathbb{R}$ so that the perturbation $\boldsymbol{b}$ has only a contribution from $\boldsymbol{u}$. Here, $\psi(\boldsymbol{x},\cdot)$ is for almost every $\boldsymbol{x} \in \Omega$ a proper and convex and Gâteaux differentiable function with derivative $\psi' : \mathbb{R}^m \to \mathbb{R}^m$ such that $\psi'(\boldsymbol{x},\cdot)$ is a Carathéodory function and satisfies the following growth conditions: there exists a number $r > 1$ and constants $c_1, c_2, \tilde{C}_1 > 0$ such that

$$\begin{aligned}
c_1(|\boldsymbol{z}|^r - 1) \le{}& \psi(\boldsymbol{x},\boldsymbol{z}) \le \tilde{C}_1(1 + |\boldsymbol{z}|^r), \\
&|\psi'(\boldsymbol{x},\boldsymbol{z})| \le c_1\left(1 + |\boldsymbol{z}|^{r-1}\right)
\end{aligned}$$

for almost every $\boldsymbol{x} \in \Omega$ and all $\boldsymbol{z} \in \mathbb{R}^m$. As we discussed in Section 5.1.1 a prototypical example is $\psi(\boldsymbol{z}) = \frac{1}{r}|\boldsymbol{z}|^r$. We choose the same function spaces for $U = \mathrm{W}^{1,p}_0(\Omega)^m \cap \mathrm{L}^4(\Omega)^m$, $V = \mathrm{H}^1_0(\Omega)^m$, $W = \mathrm{L}^{\max\{2,r\}}(\Omega)^m$, $\widetilde{W} = \mathrm{L}^{\max\{2,q\}}(\Omega)^m$, and $H = \mathrm{L}^2(\Omega)^m$ as above. Again $p, q, r \ge 1$ are to be chosen suitably. The external force is assumed to satisfy the weaker assumption $\boldsymbol{f} \in \mathrm{L}^{r^*}(0,T;W^*) + \mathrm{L}^2(0,T;V^*)$. Further, we assume $u_0 \in U$ and $v_0 \in H$. In this case, the dissipation potential $\Psi : V \to \mathbb{R}$ is given

$$\Psi(\boldsymbol{v}) = \frac{1}{2}\int_\Omega |\nabla\boldsymbol{v}(\boldsymbol{x})|^2\,\mathrm{d}\boldsymbol{x} + \int_\Omega \psi(\boldsymbol{x},\boldsymbol{v}(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}.$$

The conjugate $\Psi^* : V^* \to \mathbb{R}$ is by Lemma 2.3.5 given by the expression

$$\Psi^*(\boldsymbol{\xi}) = \min_{\boldsymbol{\eta}\in W^*}\left(\frac{1}{2}\|\boldsymbol{\xi} - \boldsymbol{\eta}\|^2_{-1,2} + \int_\Omega \psi^*(\boldsymbol{x},\boldsymbol{\eta}(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}\right),$$

where $\psi^* : \Omega \times \mathbb{R}^m \to \mathbb{R}$ is the conjugate of ψ. The energy functional $\mathcal{E} : V \to [0, +\infty]$ is given as in the previous example. The perturbation $B : [0,T] \times W \to V^*$ is consequently given by

$$\langle B(t, \boldsymbol{u}), \boldsymbol{w} \rangle_{V^* \times V} = \int_\Omega \boldsymbol{b}(\boldsymbol{x}, t, \boldsymbol{u}(\boldsymbol{x})) \cdot \boldsymbol{w}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}.$$

It is readily seen that the assumptions for the dissipation potential follow from the assumptions on ψ. From $\frac{1}{2}\|\boldsymbol{v}\|^2_{\mathrm{H}^1_0} \leq \Psi(\boldsymbol{v})$, we find $\Psi^*(\boldsymbol{\xi}) \leq \frac{1}{2}\|\boldsymbol{\xi}\|^2_{-1,2}$, and if we choose, e.g.,

$$d \geq 3,\ p \in (1, +\infty),\ q \in \begin{cases} [1, q^*) & \text{if } p \in (1,3), \\ [1, 1+p) \cap (1, 2d/(d-2)] & \text{if } p \geq 3, \end{cases}$$

where $q^* = \min\{\frac{d(p+2)}{2(d-p)}, \frac{3d+4}{d}, \frac{2d}{(d-2)}, p/2+1\}$, $d \geq 3$, $p \in (1, +\infty)$, we obtain again the estimate (7.1.3) without any restriction on r. However, from the condition $V \overset{c}{\hookrightarrow} W$, we obtain the restriction $r \in [1, 1+p) \cap (1, 2d/(d-2)]$, which is now a larger range as opposed to the previous case. Simple calculations show that the same restrictions for the exponents p, q and r hold in the dimensions $d = 1$ and $d = 2$. Again, we obtain for every initial values $u_0 \in \mathrm{dom}(\mathcal{E})$ and $v_0 \in H$, the existence of a weak solution

$$u \in \mathrm{C}_w([0,T]; U) \cap \mathrm{H}^1(0,T; V) \cap \mathrm{W}^{1,\infty}(0,T; H) \cap \mathrm{W}^{1,r}(0,T; W) \cap \mathrm{W}^{2,r^*}(0,T; U^*)$$

with $r^* = \min\{2, r^*\}$ to (P2) such that the initial conditions $u(0) = u_0$ and $u'(0) = v_0$ are satisfied, the integral equation

$$\int_0^T \Big(\langle \boldsymbol{u}'', \boldsymbol{v} \rangle_{U^* \times U} + \int_\Omega \Big(\nabla \partial_t \boldsymbol{u} : \nabla \boldsymbol{v} + \psi'(\boldsymbol{x}, \partial_t \boldsymbol{u}) \cdot \boldsymbol{v} + |\nabla \boldsymbol{u}|^{p-2} \nabla \boldsymbol{u} : \nabla \boldsymbol{v} + \boldsymbol{p} : \nabla \boldsymbol{v} \\ + (|\boldsymbol{u}|^2 - 1)\boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{b}(\boldsymbol{x}, t, \boldsymbol{u}) \Big) \mathrm{d}x + \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{U^* \times U} \Big) \mathrm{d}t \quad \text{for all } \boldsymbol{v} \in \mathrm{L}^{\max\{2,r\}}(0,T; U)$$

with $\boldsymbol{p}(\boldsymbol{x}, t) \in \mathbf{Sgn}(\nabla \boldsymbol{u}(\boldsymbol{x}, t))$ a.e. in Ω_T is fulfilled, and the energy-dissipation inequality (5.1.10) holds.

7.3 Martensitic transformation in shape-memory alloys

In this example, we consider equations which describe a solid-solid phase transition in shape-memory alloys driven by stored-energy and a dissipation mechanism. As critically discussed in Rajagopal & Roubíček[135], a commonly used model describing this phenomena is for the isothermal case given by

$$\rho \partial_{tt} \boldsymbol{u} + \nu (-1)^n \Delta^n \partial_t \boldsymbol{u} - \nabla \cdot (\boldsymbol{\sigma}(\nabla \boldsymbol{u})) + \mu (-1)^m \Delta^m \boldsymbol{u} = \boldsymbol{f}, \tag{7.3.1}$$

where $\boldsymbol{f} \in \mathrm{L}^2(0,T; \mathrm{H}^{-1,2}(\Omega)^d)$, $m, n \in \mathbb{N}$ and $\mu, \nu \geq 0$ are non-negative real values. Here, $\rho \geq 0$ denotes the density of the body, $\boldsymbol{u} : \Omega \times [0,T] \to \mathbb{R}^d$ the displacement of the body, which is related to the deformation $\boldsymbol{y}$ by $\boldsymbol{u}(\boldsymbol{x}, \cdot) = \boldsymbol{y}(\boldsymbol{x}, \cdot) - \boldsymbol{x}$ on a reference body configuration Ω, and $\boldsymbol{\sigma} : \mathbb{R}^{d \times d} \to \mathbb{R}^{d \times d}$ the Piola–Kirchhoff stress tensor depending on the gradient ∇u. The stress $\boldsymbol{\sigma}$ is, in general, not monotone and for hyperelastic materials given by the derivative of a potential $\varphi : \mathbb{R}^{d\times d} \to \mathbb{R}$ describing the specific stored energy, i.e, $\boldsymbol{\sigma} = \varphi'$ in turn is not quasiconvex[1]. The contribution of $\mu(-1)^m \varDelta^m \boldsymbol{u}$ in the equations models a capillarity-like behaviour of the solid and $\nu(-1)^n \varDelta^n \partial_t \boldsymbol{u}$ describes a higher order viscosity. In fact, the authors in [135] suggest to incorporate a correction term into the equations which describes plasticity effects of the body. More precisely, they suggest to consider the inclusion

$$\begin{aligned} \rho\partial_{tt}\boldsymbol{u} + \nu(-1)^n \varDelta^n \partial_t \boldsymbol{u} - \nabla \cdot (\boldsymbol{\sigma}_p + \boldsymbol{\sigma}(\nabla \boldsymbol{u})) + \mu(-1)^m \varDelta^m \boldsymbol{u} = \boldsymbol{f}, \\ \boldsymbol{\sigma}_p \in \operatorname{Sgn}\left(\lambda'(\nabla \boldsymbol{u}(\boldsymbol{x})) : \nabla \partial_t \boldsymbol{u}(\boldsymbol{x})\right) \lambda'(\nabla \boldsymbol{u}(\boldsymbol{x})), \end{aligned}$$

where $\operatorname{Sgn} : \mathbb{R} \rightrightarrows \mathbb{R}$ is here the multi-valued and one-dimensional sign function, $\boldsymbol{\sigma}_p : \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$ is the plastic stress, and $\lambda : \mathbb{R}^{d\times d} \to \mathbb{R}$ is a so-called phase indicator and thus indicates the phase status of $\nabla \boldsymbol{u}$. As the inclusion contains the deformation gradient $\nabla \boldsymbol{u}$ and its time derivative $\partial_t \nabla \boldsymbol{u}$, the structure of the inclusion does not allow us to apply our abstract theory. We refer the reader to Rajagopal & Roubíček [135] and Plecháč & Roubíček [133] for a detailed analysis of both evolution inclusions and to Arndt, Griebel & Roubíček [15] when the inertial forces are neglected, i.e., when $\rho = 0$. Nevertheless, we are able to cover a good scope of cases with our theory for the model which does not incorporate plasticity effects. Many cases have been studied in the literature in different situations, i.e., for different dimensions, numbers of $n, m \in \mathbb{N}$, values of $\mu, \nu \geq 0$. We refer to [74, 135] and the references cited therein for a good overview of the existing results. Most results deal with the situation $\nu > 0$ and $n = 1$ case, since higher-order viscosity implies a regularization of the solutions. The case $\nu = 0$ and $\mu = 0$ leads to the equations of classical nonlinear elastodynamics where very few results are known, see Emmrich & Puhst [72] for a nice survey on the existing results in comparison to results for the corresponding nonlinear peridynamics, a nonlocal elasticity theory. With the theory developed here, we show the existence of solutions for the cases $\nu > 0, \mu \geq 0$ and $n \geq 1, m \geq 0$, which essentially reproduce the known results from the literature. For the sake of simplicity, we supplement the equation (7.3.1) with homogeneous Dirichlet & Neumann boundary conditions. Before we specify the values for $n, m \in \mathbb{N}$ and $\nu, \mu \geq 0$, we set a system of equations and define the associated functionals, operators and spaces. We consider the initial-boundary value problem

$$\text{(P3)} \begin{cases} \rho\partial_{tt}\boldsymbol{u} + \nu(-1)^n \varDelta^n \partial_t \boldsymbol{u} - \nabla \cdot (\boldsymbol{\sigma}(\nabla \boldsymbol{u})) + \mu(-1)^m \varDelta^m \boldsymbol{u} = \boldsymbol{f} & \text{in } \varOmega_T, \\ \boldsymbol{u}(\boldsymbol{x}, 0) \ = \boldsymbol{u}_0(\boldsymbol{x}) \quad \text{on } \varOmega, \\ \boldsymbol{u}'(\boldsymbol{x}, 0) = \boldsymbol{v}_0(\boldsymbol{x}) \quad \text{on } \varOmega, \\ \frac{\partial^k \boldsymbol{u}}{\partial \boldsymbol{\nu}^k}(\boldsymbol{x}, t) \ = 0 \qquad \text{on } \partial\varOmega \times [0, T],\ k = 0, \dots, \max\{m, n\} - 1, \end{cases}$$

where we assume $\nu, \mu \geq 0$ and $\rho : \mathbb{R}^d \to [0, \infty)$ to be a measurable function satisfying $\bar{\rho} \geq \rho(\boldsymbol{x}) \geq \underline{\rho} > 0$ for a.e. $\boldsymbol{x} \in \varOmega$. We want to show the existence of a weak solution to (P3) for any initial data $\boldsymbol{u}_0 \in \mathrm{H}_0^m(\varOmega)^d$ and $\boldsymbol{v}_0 \in \mathrm{L}^2(\varOmega)^d$ and external forces $\boldsymbol{f} \in \mathrm{L}^2(0, T; \mathrm{H}^{-\max\{m,n\}}(\varOmega)^d)$, i.e., a function $u \in \mathrm{C}_w([0, T]; \mathrm{H}_0^m(\varOmega)^d) \cap \mathrm{H}^1(0, T; \mathrm{H}_0^n(\varOmega)^d) \cap$

[1] See, e.g., Roubíček [145, Remark 6.5, p. 175].

$\mathrm{W}^{1,\infty}(0,T;\mathrm{L}^2(\Omega)^d)\cap\mathrm{H}^2(0,T;\mathrm{H}^{-\max\{m,n\}}(\Omega)^d)$ satisfying the initial conditions $\boldsymbol{u}(0)=\boldsymbol{u}_0$, $\boldsymbol{v}(0)=\boldsymbol{v}_0$, the integral equation

$$\begin{aligned}&\int_0^T\left(\langle\rho\boldsymbol{u}'',\boldsymbol{v}\rangle+\int_\Omega(\nu\nabla^n\partial_t\boldsymbol{u}:\nabla^n\boldsymbol{v}+\boldsymbol{\sigma}(\nabla\boldsymbol{u}):\nabla\boldsymbol{v}+\mu\nabla^m\boldsymbol{u}:\nabla^m\boldsymbol{v})\,\mathrm{d}x\right)\mathrm{d}t,\\&=\int_0^T\langle\boldsymbol{f},\boldsymbol{v}\rangle\,\mathrm{d}t\quad\text{for all }\boldsymbol{v}\in\mathrm{L}^2(0,T;\mathrm{H}_0^{\max\{m,n\}}(\Omega)^d),\end{aligned}\tag{7.3.2}$$

and an energy-dissipation inequality which becomes an equality if $n\geq m$, where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $\mathrm{H}_0^{\max\{m,n\}}(\Omega)^d$ and its dual space $\mathrm{H}^{-\max\{m,n\}}(\Omega)^d$, where $\mathrm{H}_0^k(\Omega)^d$ is the SOBOLEV space of all measurable functions whose weak derivative exist up to the order $k\in\mathbb{N}$ and are square-integrable, and the traces of all derivatives up to the order $k-1$ vanish on the boundary $\partial\Omega$. It is readily seen that these spaces equipped with the inner product $(\boldsymbol{v},\boldsymbol{w})_{\mathrm{H}_0^k\times\mathrm{H}^{-k}}=\int_\Omega\nabla^k\boldsymbol{v}:\nabla^k\boldsymbol{w}\,\mathrm{d}\boldsymbol{x}$ form a HILBERT space and that by a classical density argument and the POINCARÉ–FRIEDRICHS inequality, the norm induced by this inner product is equivalent to the standard norm. Now, since the stored energy φ was not supposed to satisfy any convexity assumption, we have in general two possibilities of approaching this problem. On the one hand, we can treat the stress $\boldsymbol{\sigma}$ as strongly continuous perturbation of the capillarity if $\boldsymbol{\sigma}$ has at most linear growth. On the other hand, if we assume the stress satisfies an ANDREWS–BALL type condition allowing any polynomial growth for $\boldsymbol{\sigma}$, we can treat the stored energy φ as part of the energy functional.

$\boldsymbol{\sigma}$ as perturbation: In the first case, we do not make any monotonicity assumption for $\boldsymbol{\sigma}$ and do not assume that the material is hyperelastic, i.e., $\boldsymbol{\sigma}$ has a potential. More precisely, we impose the following two conditions on $\boldsymbol{\sigma}$:

(7.3.1a) The stress $\boldsymbol{\sigma}$ is continuous.

(7.3.1b) There exists a positive constant $C_\sigma>0$ such that $|\boldsymbol{\sigma}(\boldsymbol{F})|\leq C_\sigma(1+|\boldsymbol{F}|)$ for all $\boldsymbol{F}\in\mathbb{R}^{d\times d}$.

As we do not have more structure of the perturbation at our disposal, we want to treat the stress as a strongly continuous perturbation which leads to the restriction $m\geq 2$ and $\mu>0$ if $n=1$. Finding ourselves in Case **(a)**, we choose the framework $U=\mathrm{H}_0^{\max\{m,n\}}(\Omega)^d, V=\mathrm{H}_0^{\max\{m,n\}}(\Omega)^d$, $W=\mathrm{H}_0^1(\Omega)^d$ all equipped with the seminorm and the LEBESGUE space $H:=\mathrm{L}^2(\Omega)^d$ equipped with the inner product $(\boldsymbol{w},\boldsymbol{v})=\int_\Omega\rho(\boldsymbol{x})\boldsymbol{w}(\boldsymbol{x})\cdot\boldsymbol{v}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$, which is equivalent to the standard one by the assumption for ρ. Then, the dissipation potential $\Psi:V\to\mathbb{R}$ and the energy functional $\mathcal{E}:V\to(-\infty,+\infty]$ are defined by

$$\Psi(\boldsymbol{v})=\frac{\nu}{2}\int_\Omega|\nabla^n\boldsymbol{v}(\boldsymbol{x})|^2\,\mathrm{d}\boldsymbol{x}=\frac{\nu}{2}|\boldsymbol{v}|^2_{\mathrm{H}_0^n}$$

and

$$\mathcal{E}(\boldsymbol{u})=\begin{cases}\frac{\mu}{2}\int_\Omega|\nabla^m\boldsymbol{u}(\boldsymbol{x})|^2\,\mathrm{d}\boldsymbol{x}=\frac{\mu}{2}|\boldsymbol{u}|^2_{\mathrm{H}_0^m}&\text{if }\boldsymbol{u}\in\mathrm{dom}(\mathcal{E})=\mathrm{H}_0^m(\Omega)^d\\+\infty\quad\text{otherwise.}\end{cases}$$

Consequently, the perturbation $B : W \to V^*$ is given by

$$\langle B(\boldsymbol{u}), \boldsymbol{w}\rangle_{W^*\times W} = \int_\Omega \sigma(\nabla \boldsymbol{u}(\boldsymbol{x})) : \boldsymbol{w}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

We will distinguish the cases $m > n$ and $n \geq m$ by treating the system as either a linear or a nonlinearly damped inertial system despite the dissipation potential being quadratic in both cases. In the former case, we will obtain an energy-dissipation inequality, and in the latter case, an equality instead. As in the previous example seen, it is straightforward to show that the dissipation Ψ fulfills Assumption (5.Ψ). In order to verify the conditions for the energy, we note that the energy $\mathcal{E}$ is convex, sequentially weakly lower semicontinuous on V and is time-independent so that the Assumptions (5.Ea)-(5.Ed) and (5.Ef) are verified. To verify the (sequential) weak lower semicontinuity, it is sufficient to check that $\mathcal{E}$ is lower semicontinuous, since $\mathcal{E}$ is convex. The lower semicontinuity in turn can easily be verified with Lemma 2.1.2 by showing the closedness of the sublevel sets of $\mathcal{E}$, i.e., showing that $J_\alpha = \{v \in V : \mathcal{E}(v) \leq \alpha\}$ is closed for all $\alpha \in \mathbb{R}$. If $n \geq m$, then this is clear. Otherwise, let $\alpha \in \mathbb{R}$ and $(\boldsymbol{u}_k)_{k\in\mathbb{N}} \subset J_\alpha$ such that $\boldsymbol{u}_k \to \boldsymbol{u}$ in V. Then, $(\boldsymbol{u}_k)_{k\in\mathbb{N}}$ is bounded in $\mathrm{H}_0^m(\Omega)^d$ and therefore weakly convergent (up to a subsequence) to an element $\boldsymbol{u} \in V$. Since the norm on V is weakly lower semicontinuous, we infer $\boldsymbol{u} \in J_\alpha$. To verify Condition (5.Ee), we note that for $\boldsymbol{u} \in \mathrm{dom}(\partial\mathcal{E})$, there holds $\boldsymbol{\xi} \in \partial_U \mathcal{E}(\boldsymbol{u})$ if and only if $\boldsymbol{\xi} = \mu(-1)^m \Delta^m \boldsymbol{u} \in U^*$. Now, let $\boldsymbol{u}_k \overset{*}{\rightharpoonup} \boldsymbol{u}$ in $\mathrm{L}^\infty(0,T;U) \cap \mathrm{H}^1(0,T;V)$, $\boldsymbol{u}_k \to \boldsymbol{u}$ in $\mathrm{L}^2(0,T;V)$, and $\boldsymbol{\xi}_k \rightharpoonup \boldsymbol{\xi}$ in $\mathrm{L}^2(0,T;U^*)$ with $\boldsymbol{\xi}_k(t) \in \partial_U \mathcal{E}(\boldsymbol{u}_k(t))$ in V^* for a.e. $t \in (0,T)$ and $\sup_{k\in\mathbb{N}, t\in[0,T]} \mathcal{E}(\boldsymbol{u}_k(t)) < +\infty$ such that

$$\limsup_{k\to\infty} \int_0^T \langle \boldsymbol{\xi}_k(t), \boldsymbol{u}_k(t)\rangle_{U^*\times U}\,\mathrm{d}t \leq \int_0^T \langle \boldsymbol{\xi}(t), \boldsymbol{u}(t)\rangle_{U^*\times U}\,\mathrm{d}t.$$

Then, by Lemma 2.4.2 and Theorem 2.3.7 *ii)*, there holds $\boldsymbol{\xi}(t) \in \partial_U \mathcal{E}(\boldsymbol{u}(t))$ in U^* for almost every $t \in (0,T)$ and $\lim_{n\to\infty} \int_0^T \mathcal{E}(\boldsymbol{u}_n)\,\mathrm{d}t = \int_0^T \mathcal{E}(\boldsymbol{u})\,\mathrm{d}t$, i.e., norm convergence in $\mathrm{L}^2(0,T;U)$. Since $\mathrm{L}^2(0,T;U)$ is a uniformly convex space, weak convergence and norm convergence imply strong convergence, whence $\lim_{n\to\infty} \mathcal{E}(\boldsymbol{u}_n(t)) = \mathcal{E}(\boldsymbol{u}(t))$ for a.e. $t \in (0,T)$. It remains to show that the subgradient of the energy can be controlled by the energy: let $\boldsymbol{u} \in D(\partial_U \mathcal{E})$ and $\boldsymbol{\xi} \in \partial_U \mathcal{E}(\boldsymbol{u})$. Then, there holds

$$\begin{aligned}
\langle \boldsymbol{\xi}, \boldsymbol{v}\rangle_{U^*\times U} &= \mu \int_\Omega \nabla^m \boldsymbol{u} \cdot \nabla^m \boldsymbol{v}\,\mathrm{d}\boldsymbol{x} \\
&\leq \mu \left(\int_\Omega |\nabla^m \boldsymbol{u}|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \left(\int_\Omega |\nabla^m \boldsymbol{v}|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \\
&\leq \left(\mu/2 + \mu/2 \int_\Omega |\nabla^m \boldsymbol{u}|^2\,\mathrm{d}\boldsymbol{x}\right) \|\boldsymbol{v}\|_U \\
&= (\mu/2 + \mathcal{E}(\boldsymbol{u}))\|\boldsymbol{v}\|_U,
\end{aligned}$$

whence (5.Eg).

Finally, we verify the Assumptions (5.Ba) and (5.Bb) for the perturbation. The continuity Condition (5.Ba) follows immediately from the dominated convergence theorem and the Assumptions (7.3.1a) and (7.3.1b). The growth Condition (5.Bb) follows from the fact that $\Psi^*(\boldsymbol{\xi}) = \frac{1}{2\nu}\|\boldsymbol{\xi}\|_{-m,2}^2$ for all $\boldsymbol{\xi} \in V^* = \mathrm{H}^{-m}(\Omega)^d$ and the inequality

$$\begin{aligned}\langle B(\boldsymbol{u}), \boldsymbol{v}\rangle_{V^*\times V} &= \int_\Omega \sigma(\nabla \boldsymbol{u}) : \nabla \boldsymbol{v}\,\mathrm{d}\boldsymbol{x}\\ &\le \left(\int_\Omega |\sigma(\nabla \boldsymbol{u})|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \left(\int_\Omega |\nabla \boldsymbol{v}|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}}\\ &\le C\left(\int_\Omega (1+|\nabla \boldsymbol{u}|)^2\right)^{\frac{1}{2}} |\boldsymbol{v}|_{n,2}\\ &\le C\,(1+|\boldsymbol{u}|_{1,2})\,|\boldsymbol{v}|_{n,2}\\ &\le C\,(1+|\boldsymbol{u}|_{m,2})\,|\boldsymbol{v}|_{n,2}\\ &\le C\left(1+\mathcal{E}(\boldsymbol{u})^{\frac{1}{2}}\right)|\boldsymbol{v}|_{n,2} \quad \text{for all } \boldsymbol{v}\in V.\end{aligned}$$

Hence, we obtain

$$\|B(\boldsymbol{u})\|_{-n,2}^2 \le C(1+\mathcal{E}(\boldsymbol{u})) \quad \text{for all } \boldsymbol{u}\in U.$$

Then, by Theorem 5.1.4, for all initial values $u_0 \in \mathrm{dom}(\mathcal{E})$, $v_0 \in H$, there exists a weak solution

$$\boldsymbol{u} \in \mathrm{C}_w([0,T];U) \cap \mathrm{H}^1(0,T;V) \cap \mathrm{W}^{1,\infty}(0,T;H) \cap \mathrm{H}^2(0,T;U^*)$$

to (P3) such that (7.3.2) is fulfilled, and the energy-dissipation inequality

$$\begin{aligned}&\frac{1}{2}\|\rho\boldsymbol{u}'(t)\|_{\mathrm{L}^2(\Omega)^d}^2 + \frac{\mu}{2}|\boldsymbol{u}(t)|_{m,2}^2 + \int_s^t \frac{1}{2\nu}|\boldsymbol{f}(r)+\boldsymbol{\xi}(r)+\nabla\cdot\boldsymbol{\sigma}(\nabla\boldsymbol{u}(r)) - \boldsymbol{u}''(r)|_{-n,2}^2\,\mathrm{d}r\\ &+\int_s^t \frac{\nu}{2}|\boldsymbol{u}'(r)|_{n,2}^2\,\mathrm{d}r\\ &\le \frac{1}{2}\|\rho\boldsymbol{u}'(s)\|_{\mathrm{L}^2(\Omega)^d}^2 + \frac{\mu}{2}|\boldsymbol{u}(s)|_{m,2}^2 + \int_s^t \left(\langle \boldsymbol{f}(t), \boldsymbol{u}'(r)\rangle_{V^*\times V} + \boldsymbol{\sigma}(\nabla\boldsymbol{u}(t)) : \nabla\boldsymbol{u}'(r)\right)\mathrm{d}r\end{aligned} \tag{7.3.3}$$

holds for all $t \in [0,T]$ for $s = 0$ and almost every $s \in (0,t)$, where $\boldsymbol{\xi} \in \mathrm{L}^\infty(0,T;U^*)$ with $\boldsymbol{\xi}(r) = \mu(-1)^m \Delta^m \boldsymbol{u}(r)$ a.e. in $(0,T)$. The inequality (7.3.3) holds as an equality for almost every $t \in [0,T]$ and $s = 0$ if $n \ge m$. This stems from the fact that we can test with $\boldsymbol{u}'$ in the integral equation (7.3.2) or that we can treat the system of equations alternatively as a nonlinearly damped inertial system with the same choices for $\mathcal{E}, \Psi, B, U, W$ and H but with $V = \mathrm{H}_0^m(\Omega)^d$.

$\boldsymbol{\sigma}$ as energy: Now, we suppose that $\boldsymbol{\sigma}$ fulfills, aside from certain growth and continuity conditions, a potential, and that $\boldsymbol{\sigma}$ satisfies an Andrews–Ball type condition which was originally introduced by Andrews & Ball to show global existence of solutions for the one-dimensional equations in viscoelastodynamics, i.e., when $\nu > 0, n = 1$ and $\mu = 0$, see [12, 13]. The existence of weak solutions to the aforementioned case in arbitrary dimensions has already been studied in a more general abstract setting in Emmrich & Šiška [74] by making the crucial assumption that the operator $B + \lambda A$ is monotone for some $\lambda > 0$, which in practice generalizes the Andrews–Ball condition. The latter condition states that $\boldsymbol{\sigma}$ is monotone in the large, i.e., there exists a positive value $R > 0$ such that

$$\left(\boldsymbol{\sigma}(\boldsymbol{F}) - \boldsymbol{\sigma}(\tilde{\boldsymbol{F}})\right) : \left(\boldsymbol{F} - \tilde{\boldsymbol{F}}\right) > 0 \quad \text{for all } \boldsymbol{F}, \boldsymbol{F} \in \mathbb{R}^{d\times d} \text{ with } |\boldsymbol{F} - \tilde{\boldsymbol{F}}| \ge R.$$

We will impose the more general assumption of the convexity of $\varphi + \frac{\lambda}{2}$ which is in this smooth setting equivalent to the monotonicity of $\boldsymbol{\sigma} + \lambda\mathrm{id}$. However, if $m \in \mathbb{N}$ is sufficiently large so that we can again treat the stored energy φ as strongly continuous perturbation, then the previous condition is redundant. Therefore, we will not explicitly focus on this case. Having said that, the exact conditions which we impose on the stress $\boldsymbol{\sigma}$ are the following:

(7.3.2a) There exists a continuously differentiable function $\varphi : \mathbb{R}^{d\times d} \to \mathbb{R}$ such that $\boldsymbol{\sigma} = \varphi'$.

(7.3.2b) There exist positive constants $c_\sigma^1, C_\sigma^1 > 0$ and $p > 1$ such that

$$\begin{aligned} c_\sigma^1|\boldsymbol{F}|^p - C_\sigma^1 &\le \boldsymbol{\sigma}(\boldsymbol{F}) : \boldsymbol{F} \\ |\boldsymbol{\sigma}(\boldsymbol{F})| &\le C_\sigma^1(1 + |\boldsymbol{F}|^{p-1}) \\ c_{\boldsymbol{\sigma}}^1|\boldsymbol{F}|^p - C_\sigma^1 \le |\varphi(\boldsymbol{F})| &\le C_\sigma^1(1 + |\boldsymbol{F}|^p) \quad \text{for all } \boldsymbol{F} \in \mathbb{R}^{d\times d}. \end{aligned}$$

(7.3.2c) There exists a positive number $\lambda > 0$ such that $\varphi + \frac{\lambda}{2}|\cdot|^2$ is convex.

Condition (7.3.2c) is in fact equivalent to the following Andrews–Ball type condition:

$$\Big(\boldsymbol{\sigma}(\boldsymbol{F}) - \boldsymbol{\sigma}(\tilde{\boldsymbol{F}})\Big) : \Big(\boldsymbol{F} - \tilde{\boldsymbol{F}}\Big) \ge -\lambda|\boldsymbol{F} - \tilde{\boldsymbol{F}}|^2 \quad \text{for all } \boldsymbol{F}, \boldsymbol{F} \in \mathbb{R}^{d\times d}. \tag{7.3.4}$$

This follows from the convexity and Gâteaux differentiability of $\varphi + \frac{\lambda}{2}|\cdot|^2$, the parallelogram identity of $|\cdot|$ and Lemma 2.2.2 and 2.2.6. The Andrews–Ball condition in turn necessitates (7.3.4) if $\boldsymbol{\sigma}$ is in addition locally Lipschitz continuous, see [74].

The obvious choice of the spaces are $U = \mathrm{H}_0^{\max\{n,m\}}(\varOmega)^d \cap \mathrm{W}_0^{1,p}(\varOmega)^d, V = W = \mathrm{H}_0^n(\varOmega)^d$ and H as before. Further, we assume again $\boldsymbol{f} \in \mathrm{L}^2(0,T;\mathrm{H}^{-\max\{m,n\}}(\varOmega)^d)$. Then, the dissipation potential $\varPsi$ and the energy functional $\mathcal{E}$ are given by

$$\varPsi(\boldsymbol{v}) = \frac{\nu}{2}\int_\varOmega |\nabla^n \boldsymbol{v}(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x}, \quad \text{and} \quad \begin{aligned} \mathcal{E}(\boldsymbol{u}) &= \int_\varOmega \Big(\varphi(\nabla\boldsymbol{u}(\boldsymbol{x})) + \frac{\mu}{2}|\nabla^m \boldsymbol{u}(\boldsymbol{x})|^2\Big)\,\mathrm{d}\boldsymbol{x} \\ &= \mathcal{E}_1(\boldsymbol{u}) + \mathcal{E}_2(\boldsymbol{u}), \end{aligned}$$

respectively, and therefore, $B \equiv 0$. There are essentially two cases to be discussed here: the first case where the weak solution does not satisfy the energy-dissipation inequality, and the second case where it does. In both cases, we assume $n \ge 1, \nu > 0$ and $m \ge 1$. The first case then includes all $\mu \ge 0$, whereas the second case is limited to $\mu \ge \frac{\lambda}{C_{\mathrm{H}_0^m,\mathrm{H}_0^1}}$, where $C_{H_0^m,H_0^1} > 0$ denotes the constant of the embedding $\mathrm{H}_0^m(\varOmega)^d \hookrightarrow \mathrm{H}_0^1(\varOmega)^d$. The lower bound on μ ensures the convexity of $\mathcal{E}$ and thus implies that $\mathcal{E}$ is weakly lower semicontinuity which is necessary to show the energy-dissipation inequality/balance. The weak lower semicontinuity of $\mathcal{E}_1$ is, in general, not given. However, if $\mathcal{E}_1$ is weakly lower semicontinuous, then we allow $\mu \ge 0$. We focus first on the case where the weak solution fulfills the energy-dissipation inequality (balance), i.e., when $\mathcal{E}$ is weakly lower semicontinuous. Then, the assumptions on the dissipation potential $\varPsi$ are for both cases easily verified. Hence, we need to verify the Conditions (5.Ea)-(5.Eg) for $\mathcal{E}$. With the remarks made above and the fact that $\mathcal{E}$ is time-independent, Conditions (5.Ea)-(5.Ed) and (5.Ef) are fulfilled. From Assumption (7.3.2c) it follows that $\mathcal{E}$ is λ-convex, and Assumption (7.3.2a) and (7.3.2b) imply that the energy functional $\mathcal{E}$ is Gâteaux differentiable on U and that the derivative is given by

$$\langle \mathrm{D}_G\mathcal{E}(\boldsymbol{u}), \boldsymbol{v}\rangle_{U^*\times U} = \int_\Omega \boldsymbol{\sigma}(\nabla \boldsymbol{u}(\boldsymbol{x})) : \nabla \boldsymbol{v}(\boldsymbol{x}) + \mu \Delta \boldsymbol{u}(\boldsymbol{x}) \cdot \Delta \boldsymbol{v}(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}.$$

Consequently, by the subdifferential calculus, the subdifferential of the energy with respect to U is single-valued with $\partial_U \mathcal{E}(\boldsymbol{u}) = \{\mathrm{D}\mathcal{E}(\boldsymbol{u})\}$ and hence $\boldsymbol{\xi} \in \partial_U \mathcal{E}(\boldsymbol{u})$ if and only if $\boldsymbol{\xi} = -\nabla \cdot \boldsymbol{\sigma}(\nabla \boldsymbol{u}) + \mu(-1)^m \Delta^m \boldsymbol{u} \in U^*$. Before we proceed with showing the remaining conditions, we note that since $\lambda \neq 0$, there holds $V_\lambda = U \cap V$. Then, Condition (5.Eg) follows from the following estimate:

$$\begin{aligned}
\langle \mathrm{D}_G\mathcal{E}(\boldsymbol{u}), \boldsymbol{v}\rangle_{(U^*+V^*)\times(U\cap V)} &= \int_\Omega \boldsymbol{\sigma}(\nabla \boldsymbol{u}(\boldsymbol{x})) : \nabla \boldsymbol{v}(\boldsymbol{x}) + \mu \nabla^m \boldsymbol{u}(\boldsymbol{x}) \cdot \nabla^m \boldsymbol{v}(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \\
&\leq \left(\int_\Omega |\boldsymbol{\sigma}(\nabla \boldsymbol{u}(\boldsymbol{x}))|^{p/(p-1)} \,\mathrm{d}x\right)^{(p-1)/p} \left(\int_\Omega |\nabla \boldsymbol{v}(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{p}} \\
&\quad + \mu \left(\int_\Omega |\nabla^m \boldsymbol{u}(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \left(\int_\Omega |\nabla^m \boldsymbol{v}(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \\
&\leq \left(C_\sigma 2^{p/(p-1)} \int_\Omega (1 + |\nabla \boldsymbol{u}(\boldsymbol{x})|^p) \,\mathrm{d}\boldsymbol{x}\right)^{(p-1)/p} \|\boldsymbol{v}\|_{U\cap V} \\
&\quad + \mu \left(\int_\Omega |\nabla^m \boldsymbol{u}(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \|\boldsymbol{v}\|_{U\cap V} \\
&\leq C\left(1 + \int_\Omega |\nabla \boldsymbol{u}(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x} + \frac{\mu}{2}\int_\Omega |\Delta \boldsymbol{u}(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x}\right) \|\boldsymbol{v}\|_{U\cap V} \\
&\leq C\left(1 + \int_\Omega \varphi(\nabla \boldsymbol{u}(\boldsymbol{x})) \,\mathrm{d}\boldsymbol{x} + \frac{\mu}{2}\int_\Omega |\Delta \boldsymbol{u}(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x}\right) \|\boldsymbol{v}\|_{U\cap V} \\
&\leq C(1 + \mathcal{E}(\boldsymbol{u}))\|\boldsymbol{v}\|_{U\cap V},
\end{aligned}$$

where we made use of Hölder's and Young's inequality as well as the growth condition (7.3.2b). Finally, we verify the closedness condition for $\partial_{U\cap V}\mathcal{E}$, i.e. (5.Ee): let $\boldsymbol{u}_k \overset{*}{\rightharpoonup} \boldsymbol{u}$ in $\mathrm{L}^\infty(0,T;U) \cap \mathrm{H}^1(0,T;V)$, $\boldsymbol{u}_k \to \boldsymbol{u}$ in $\mathrm{L}^2(0,T;V)$, and $\mathrm{D}_G\mathcal{E}(\boldsymbol{u}_k) \rightharpoonup \boldsymbol{\xi}$ in $\mathrm{L}^2(0,T;U^*)$ such that $\sup_{k\in\mathbb{N},t\in[0,T]} \mathcal{E}(\boldsymbol{u}_k(t)) < +\infty$ and

$$\limsup_{k\to\infty} \int_0^T \langle \mathrm{D}_G\mathcal{E}(\boldsymbol{u}_k(t)), \boldsymbol{u}_k(t)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}t \leq \int_0^T \langle \boldsymbol{\xi}(t), \boldsymbol{u}(t)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}t.$$

This implies

$$\begin{aligned}
&\limsup_{k\to\infty} \int_0^T \langle \mathrm{D}_G\mathcal{E}(\boldsymbol{u}_k(t))(t), \boldsymbol{u}_k(t)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}t \\
&= \limsup_{k\to\infty} \int_0^T \langle \mathrm{D}_G\mathcal{E}(\boldsymbol{u}_k(t)), \boldsymbol{u}_k(t)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}t + \frac{\lambda}{2}\int_0^T |\nabla \boldsymbol{u}(t)|^2 \,\mathrm{d}\boldsymbol{x} \\
&\quad - \frac{\lambda}{2}\int_0^T |\nabla \boldsymbol{u}(t)|^2 \,\mathrm{d}\boldsymbol{x} \\
&= \limsup_{k\to\infty} \left(\int_0^T \langle \mathrm{D}_G\mathcal{E}(\boldsymbol{u}_k(t)), \boldsymbol{u}_k(t)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}t + \frac{\lambda}{2}\int_0^T |\nabla \boldsymbol{u}(t)|^2 \,\mathrm{d}\boldsymbol{x}\right)
\end{aligned}$$

$$\begin{aligned}&- \frac{\lambda}{2}\int_0^T |\nabla \boldsymbol{u}(t)|^2 \,\mathrm{d}\boldsymbol{x}\\ &\le \int_0^T \langle \boldsymbol{\xi}(t), \boldsymbol{u}(t)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}t,\end{aligned}$$

whence

$$\begin{aligned}&\limsup_{k\to\infty}\int_0^T \langle \mathrm{D}_G\mathcal{E}(\boldsymbol{u}_k(t)) + \frac{\lambda}{2}\boldsymbol{u}_k(t), \boldsymbol{u}_k(t)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}t\\ &\le \int_0^T \langle \boldsymbol{\xi}(t) + \frac{\lambda}{2}\boldsymbol{u}(t), \boldsymbol{u}(t)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}t.\end{aligned}$$

The convexity of $\mathcal{E} + \frac{\lambda}{2}|\cdot|^2$ together with Lemma 2.4.2 and Theorem 2.3.7 implies that $\boldsymbol{\xi}(t) = \mathrm{D}_G\mathcal{E}(\boldsymbol{u}(t)) = -\nabla\cdot\boldsymbol{\sigma}(\nabla\boldsymbol{u}(t)) + \mu(-1)^m\Delta^m\boldsymbol{u}(t)$ a.e. in $(0,T)$ and

$$\begin{aligned}&\lim_{k\to\infty}\int_0^T \langle \mathrm{D}_G\mathcal{E}(\boldsymbol{u}_k(t)), \boldsymbol{u}_k(t)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}t\\ &= \lim_{k\to\infty}\int_0^T\int_\Omega \Big(\boldsymbol{\sigma}(\nabla\boldsymbol{u}_k(t)) : \nabla\boldsymbol{u}_k(t) + \mu|\nabla^m\boldsymbol{u}_k(t)|^2\Big)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\ &= \int_0^T\int_\Omega \Big(\sigma(\nabla\boldsymbol{u}(t)) : \nabla\boldsymbol{u}(t) + \mu|\nabla\boldsymbol{u}(t)|^2\Big)\,d\boldsymbol{x}\,\mathrm{d}t\\ &= \int_0^T \langle \boldsymbol{\xi}(t), \boldsymbol{u}(t)\rangle_{(U^*+V^*)\times(U\cap V)} \,\mathrm{d}t.\end{aligned}$$

Then, the strong convergence $\boldsymbol{u}_k \to \boldsymbol{u}$ in $\mathrm{L}^2(0,T;V)$ implies $\nabla\boldsymbol{u}_k \to \nabla\boldsymbol{u}$ a.e. in Ω_T as $k\to\infty$. Together with Assumption (7.3.2b), this yields $\boldsymbol{u}_n \to \boldsymbol{u}$ in $\mathrm{L}^p(0,T;\mathrm{W}_0^{1,p}(\Omega))$ and with the continuity of φ finally $\mathcal{E}(\boldsymbol{u}_n(t)) \to \mathcal{E}(\boldsymbol{u}(t))$ a.e. in $(0,T)$ as $n\to\infty$. Therefore, by Theorem 5.1.4, for every $\boldsymbol{u}_0 \in U$ and $\boldsymbol{v}_0 \in H$, there exists a weak solution $\boldsymbol{u} \in \mathrm{C}_w([0,T];\mathrm{H}_0^m(\Omega)^d) \cap \mathrm{H}^1(0,T;\mathrm{H}_0^n(\Omega)^d) \cap \mathrm{W}^{1,\infty}(0,T;\mathrm{L}^2(\Omega)^d) \cap \mathrm{H}^2(0,T;\mathrm{H}^{-\max\{m,n\}}(\Omega)^d)$ satisfying the integral equation (7.3.2) and the energy dissipation inequality

$$\begin{aligned}&\frac{1}{2}\|\rho\boldsymbol{u}'(t)\|^2_{\mathrm{L}^2(\Omega)^d} + \frac{\mu}{2}|\boldsymbol{u}(t)|^2_{m,2} + \int_\Omega \varphi(\nabla\boldsymbol{u}(t))\,\mathrm{d}x + \int_s^t \frac{\nu}{2}|\boldsymbol{u}'(r)|^2_{n,2}\,\mathrm{d}r\\ &\quad + \frac{1}{2\nu}\int_s^t |\boldsymbol{f}(r) - \nabla\cdot\boldsymbol{\sigma}(\nabla\boldsymbol{u}(r)) + \mu(-1)^m\Delta^m\boldsymbol{u}(r) - \boldsymbol{u}''(r)|^2_{-n,2}\,\mathrm{d}r\\ &\le \frac{1}{2}\|\rho\boldsymbol{u}'(s)\|^2_{\mathrm{L}^2(\Omega)^d} + \frac{\mu}{2}|\boldsymbol{u}(s)|^2_{m,2} + \int_\Omega \varphi(\nabla\boldsymbol{u}(s))\,\mathrm{d}x + \int_s^t \langle \boldsymbol{f}(t), \boldsymbol{u}'(r)\rangle_{V^*\times V}\,\mathrm{d}r,\end{aligned}$$

which holds for all $t \in [0,T]$ for $s = 0$ and almost every $s \in (0,t)$. The inequality (7.3.3) becomes again an equality and holds for almost every $t \in [0,T]$ and $s = 0$ if $n \ge m$, which stems from the fact that the system can be treated as a nonlinearly damped inertial system.

If $\mu = 0$, by Remark 5.5.1, we still obtain a solution satisfying the differential inclusion (5.1.9) if there exists $r_0 > 0$ such that $v \mapsto \frac{1}{r_0}\Psi(\boldsymbol{v}) + \mathcal{E}(\boldsymbol{v})$ is sequentially weakly lower semicontinuous. But since we assumed $\mathcal{E} + \frac{\lambda}{2}|\cdot|^2_{1,2}$ to be convex, which implies that $\mathcal{E} + \frac{1}{2}\lambda C_{\mathrm{H}_0^n,\mathrm{H}_0^1}|\cdot|^2_{1,n}$ is convex with $C_{\mathrm{H}_0^n,\mathrm{H}_0^1} > 0$ being the constant of the embedding $\mathrm{H}_0^n(\Omega)^d \hookrightarrow \mathrm{H}_0^1(\Omega)^d$, this particularly implies that $\mathcal{E} + \frac{\lambda C_{\mathrm{H}_0^n,\mathrm{H}_0^1}}{2}|\cdot|^2_{1,n}$ is sequentially weakly lower semicontinuous. Therefore, by the procedure carried out above, all conditions are fulfilled so that we obtain the existence of a weak solution in the sense of (7.3.2), which, in general, does not fulfill the energy-dissipation inequality.

7.4 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

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

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.

We make the following assumptions on the functions ψ 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.

(7.4.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 $y \in \mathbb{R}$.

(7.4.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 $z \in \mathbb{R}^m, y \in \mathbb{R}, |y| \leq R$.

(7.4.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

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

Accordingly, the function spaces are given by $V = \mathrm{W}_0^{1,q}(\Omega)$, $U = \mathrm{H}_0^1(\Omega), \widetilde{W} = \mathrm{L}^{\max\{p,2\}}(\Omega)$ and $H = \mathrm{L}^2(\Omega)$. Then, we identify the dissipation potential $\Psi : V \to \mathbb{R}$ and the energy functional $\mathcal{E} : U \to [0,+\infty)$ as

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

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

$$\langle B(u), w\rangle_{\widetilde{W}^*\times\widetilde{W}} = \int_\Omega b(u(x))w(x)\,\mathrm{d}x.$$

We note that the conjugate functional Ψ_u^* can, in general, not be expressed as an integral functional over Ω, since it is defined on $\mathrm{W}^{-1,q^*}(\Omega)$.

Obviously, $\mathcal{E}$ satisfies all Conditions 6.1. In view of the compact embedding $\mathrm{H}_0^1(\Omega) \overset{c}{\hookrightarrow} \mathrm{C}(\overline{\Omega})$ and FATOU's lemma, it is readily that Ψ_u satisfies Conditions (6.Ψa) and (6.Ψb). In order to verify Condition (6.Ψc), we show that for every sequence $u_n \rightharpoonup u$ in U with $\sup_{n\in\mathbb{N}} \mathcal{E}(u_n) < +\infty$, there holds $\Psi_{u_n} \xrightarrow{\mathrm{M}} \Psi_u$ as $n \to \infty$. As we mentioned in Remark 2.4.5, the MOSCO-convergence $\Psi_{u_n} \xrightarrow{\mathrm{M}} \Psi_u$ implies the MOSCO-convergence of the related integral functionals defined by (2.3.6) that in turn implies Condition (6.Ψc). The liminf estimate in the MOSCO-convergence follows from IOFFE [95, 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 $f \in \mathrm{L}^2(0, T; H)$, it is easy to check in the same way as in the previous examples that Conditions (6.Ba), (6.Bb), and (6.Bb) are also fulfilled. Therefore, Theorem 6.1.4 ensures that for every initial values $v_0 \in H$ and $u_0 \in U$, the existence of a solution $u \in \mathrm{C}_w([0, T]; U) \cap \mathrm{W}^{1,\infty}(0, T; H) \cap \mathrm{W}^{2,q^*}(0, T; U^* + V^*)$ with $u - u_0 \in \mathrm{W}^{1,q}(0, T; V)$ to (P4) satisfying the integral equation

$$\int_0^T \left(\langle u'' v \rangle_{(U^*+V^*)\times(U\cap V)} + \int_\Omega \mathbf{p} \cdot \nabla v + b(u) v \,\mathrm{d}x \right) \mathrm{d}t = \int_0^T \int_\Omega f v \,\mathrm{d}x \,\mathrm{d}t$$

for all $v \in \mathrm{L}^{\min\{2,q^*\}}(0, T; U^* + V^*)$ with $\mathbf{p}(x, t) \in \partial_v \psi(x, u(x, t), \nabla \partial_t u(x, t))$ a.e. in Ω_T, and the energy-dissipation balance

$$\begin{aligned} &\frac{1}{2}\|u'(t)\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2}\|u(t)\|^2_{H_0^1(\Omega)} + \int_0^t \Big(\Psi_{u(t)}(u'(r)) + \Psi^*_{u(t)}(f(r) - u''(r) - \Delta u(r)) \Big) \,\mathrm{d}r \\ &= \frac{1}{2}\|v_0\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2}\|u_0\|^2_{H_0^1(\Omega)} + \int_0^t \langle f(r), u'(r) \rangle_{\mathrm{L}^2(\Omega)\times\mathrm{L}^2(\Omega)} \,\mathrm{d}r \end{aligned}$$

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

7.5 Differential inclusion II

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

$$\text{(P5)} \begin{cases} \partial_{tt} u + |\partial_t u|^{q-2}\, \partial_t u + p - \Delta u = f & \text{in } \Omega_T, \\ p(\boldsymbol{x}, t) \in \operatorname{Sgn}\left(\partial_t u(\boldsymbol{x}, t)\right) & \text{a.e. in } \Omega_T, \\ u(\boldsymbol{x}, 0) \ = u_0(\boldsymbol{x}) & \text{on } \Omega, \\ u'(\boldsymbol{x}, 0) = v_0(\boldsymbol{x}) & \text{on } \Omega, \\ u(\boldsymbol{x}, t) \ = 0 & \text{on } \partial\Omega \times [0, T], \end{cases}$$

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

$$\Psi_u(v) = \int_\Omega \left(\frac{1}{q}|v(\boldsymbol{x})|^q + |v(\boldsymbol{x})|\right) \mathrm{d}\boldsymbol{x} \quad \text{and} \quad \mathcal{E}(u) = \frac{1}{2}\int_\Omega |\nabla u(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x},$$

respectively. Consequently, $B = 0$ and $\mathcal{E}_t^2 = 0$. Again, all the assumptions are easily verified, so that Theorem 6.1.4 ensures for any initial values $v_0 \in H$ and $u_0 \in U$ the existence of a solution $u \in \mathrm{C}_w([0,T];U) \cap \mathrm{W}^{1,\infty}(0,T;H) \cap \mathrm{W}^{2,q^*}(0,T;U^* + V^*)$ with $u - u_0 \in \mathrm{W}^{1,q}(0,T;V)$ to (P5) fulfilling the integral equation

$$\int_0^T \left(\langle u'' v \rangle_{(U^*+V^*)\times(U\cap V)} + \int_\Omega |\partial_t u|^{q-2}\partial_t u v + pv + \nabla u \cdot \nabla v \, \mathrm{d}\boldsymbol{x} \right) \mathrm{d}t = \int_0^T \int_\Omega f v \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}t$$

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

$$\begin{aligned}
&\frac{1}{2}\|u'(t)\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2}\|u(t)\|^2_{H^1_0(\Omega)} + \int_0^t \left(\Psi(u'(r)) + \Psi^*(f(r) - u''(r) - \Delta u(r))\right) \mathrm{d}r \\
&= \frac{1}{2}\|v_0\|^2_{\mathrm{L}^2(\Omega)} + \frac{1}{2}\|u_0\|^2_{H^1_0(\Omega)} + \int_0^t \langle f(r), u'(r)\rangle_{\mathrm{L}^2(\Omega)\times\mathrm{L}^2(\Omega)} \, \mathrm{d}r
\end{aligned}$$

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

Appendix

A.1 The Gronwall lemma

In this section, we provide two versions of the Gronwall lemma, the discrete and the classical version. The Growall lemma is indispensable for obtaining a priori estimates or to show stability or uniqueness results.

Lemma A.1.1 (Gronwall) *Let $T \in (0, +\infty]$, $s \in [0, T)$, $a, b \in \mathrm{L}^\infty(s, T)$, $\lambda \in \mathrm{L}^1(s, T)$ with $\lambda(t) \geq 0$ almost everywhere in (s, T) such that*

$$a(t) \leq b(t) + \int_s^t \lambda(r) a(r) \,\mathrm{d}r \quad \text{a.e. in } (s, T),$$

then, there holds

$$a(t) \leq b(t) + \int_s^t e^{\Lambda(t) - \Lambda(r)} \lambda(r) b(r) \,\mathrm{d}r \quad \text{a.e. in } (s, T),$$

where $\Lambda(t) = \int_r^t \lambda(r) \,\mathrm{d}r$, $t \in [s, T]$.

Proof. A proof can be found in Emmrich [71, Lemma 7.3.1, pp.180]. □

Lemma A.1.2 (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

$$a_n \leq B e^{\beta \sum_{k=1}^{n-1} \tau_k} \quad \text{for all } n \in \mathbb{N}.$$

Proof. A proof can be found in Ambrosio et al. [10, 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–Dubinskiĭ lemma and deals with the case of piecewise constant functions in time, which avoids the construction of weakly time differentiable functions.

Lemma A.2.1 (Lions–Aubin–Dubinskiĭ) *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 $(u_{\tau_n})_{n\in\mathbb{N}}$ 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*

$$\tau_n^{-1}\|\sigma_{\tau_n}u_{\tau_n} - u_{\tau_n}\|_{\mathrm{L}^r(0,T-\tau_n;Y)} + \|u_{\tau_n}\|_{\mathrm{L}^p(0,T;X)} \leq C \quad \text{for all } n \in \mathbb{N}, \tag{A.2.1}$$

where $\sigma_{\tau_n}u\colon = u(\cdot + \tau_n)$ and $C > 0$ is a constant which is independent of τ. If $p < \infty$, then $(u_{\tau_n})_{n\in\mathbb{N}}$ is relatively compact in $\mathrm{L}^p(0,T;B)$ and if $p = \infty$, there exists a subsequence of $(u_{\tau_n})_{n\in\mathbb{N}}$ converging in $\mathrm{L}^q(0,T;B)$ for all $1 \leq q < \infty$ to a limit function belonging to $\mathrm{C}([0,T];B)$.

Proof. A proof can be found in Dreher & Jüngel [64, Theorem 1]. □

References

[1] M. Agueh, N. Ghoussoub, and X. Kang, *Geometric inequalities via a general comparison principle for interacting gases*, Geom. Funct. Anal., 14 (2004), pp. 215–244.

[2] S. Aizicovici and V.-M. Hokkanen, *Doubly nonlinear equations with unbounded operators*, Nonlinear Anal., 58 (2004), pp. 591–607.

[3] S. Aizicovici and V.-M. Hokkanen, *Doubly nonlinear periodic problems with unbounded operators*, J. Math. Anal. Appl., 292 (2004), pp. 540–557.

[4] G. Akagi, *Doubly nonlinear evolution equations with non-monotone perturbations in reflexive Banach spaces*, J. Evol. Equ., 11 (2011), pp. 1–41.

[5] G. Akagi and S. Melchionna, *Elliptic-regularization of nonpotential perturbations of doubly-nonlinear flows of nonconvex energies: a variational approach*, J. Convex Anal., 25 (2018), pp. 861–898.

[6] G. Akagi and U. Stefanelli, *Periodic solutions for doubly nonlinear evolution equations*, J. Differential Equations, 251 (2011), pp. 1790–1812.

[7] H. Amann and J. Escher, *Analysis III*, Birkhäuser, Basel, 2001.

[8] L. Ambrosio, *Metric space valued functions of bounded variation*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 17 (1990), pp. 439–478.

[9] L. Ambrosio, *Minimizing movements*, Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5), 19 (1995), pp. 191–246.

[10] 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 ed., 2008.

[11] J. An and L. Ying, *On the gradient flow structure of the isotropic landau equation*, (2020).

[12] G. Andrews, *On the existence of solutions to the equation* $u_{tt} = u_{xxt} + \sigma(u_x)_x$, J. Differential Equations, 35 (1980), pp. 200 – 231.

[13] G. Andrews and J. Ball, *Asymptotic behaviour and changes of phase in one-dimensional nonlinear viscoelasticity*, J. Differential Equations, 44 (1982), pp. 306 – 341.

[14] T. Arai, *On the existence of the solution for $\partial\varphi(u'(t)) + \partial\psi(u(t)) \ni f(t)$*, J. Fac. Sci. Univ. Tokyo Sect. IA Math., 26 (1979), pp. 75–96.

[15] M. Arndt, M. Griebel, and T. Roubíček, *Modelling and numerical simulation of martensitic transformation in shape memory alloys*, Cont. Mech. Thermodyn., 15 (2003), pp. 463–485.

[16] E. Asplund, *Averaged norms*, Israel J. Math., 5 (1967), pp. 227–233.

[17] H. Attouch, *Variational Convergence for Functions and Operators*, Applicable Mathematics Series, Pitman, Boston, 1984.

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

[19] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston, 1990.

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

[21] A. Bacho, E. Emmrich, and A. Mielke, *An existence result and evolutionary Γ-convergence for perturbed gradient systems*, J. Evol. Equ., 19 (2019), pp. 479–522.

[22] J. M. Ball, *A version of the fundamental theorem for Young measures*, in PDEs and continuum models of phase transitions (Nice, 1988), Springer, Berlin, 1989, pp. 207–215.

[23] A. Bamberger, *Étude d'une équation doublement non linéaire*, J. Functional Analysis, 24 (1977), pp. 148–155.

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

[25] V. Barbu, *Existence for nonlinear Volterra equations in Hilbert spaces*, SIAM J. Math. Anal., 10 (1979), pp. 552–569.

[26] V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*, Springer-Verlag, New York, 2010.

[27] V. Barbu and T. Precupanu, *Convexity and Optimization in Banach Spaces*, Springer-Verlag, Dordrecht, 4th ed., 1986.

[28] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer-Verlag, New York, 2011.

[29] C. Bauzet and G. Vallet, *On abstract Barenblatt equations*, Differ. Equ. Appl., 3 (2011), pp. 487–502.

[30] A. Braides, *An introduction to homogenization and Gamma-convergence.* School on Homogenization, ICTP, Trieste, 1993.

[31] A. Braides, *Γ-Convergence for Beginners*, vol. 22 of Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, Oxford, 2002.

[32] H. Brézis, *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, North-Holland Publishing Co., Amsterdam, 1973.

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

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

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

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

[37] M. Bulíček, P. Kaplický, and M. Steinhauer, *On existence of a classical solution to a generalized Kelvin-Voigt model*, Pacific J. Math., 262 (2013), pp. 11–33.

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

[39] W. L. Bynum, *Characterizations of uniform convexity*, Pacific J. Math., 38 (1971), pp. 577–581.

[40] W. L. Bynum, *Weak parallelogram laws for Banach spaces*, Canad. Math. Bull., 19 (1976), pp. 269–275.

[41] J. A. Carrillo, M. G. Delgadino, L. Desvillettes, and J. Wu, *The landau equation as a gradient flow*, arXiv:2007.08591, (2020).

[42] C. Castaing, P. Raynaud de Fitte, and M. Valadier, *Young measures on topological spaces*, vol. 571 of Mathematics and its Applications, Kluwer Academic Publishers, Dordrecht, 2004. With applications in control theory and probability theory.

[43] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics, Vol. 580, Springer-Verlag, Berlin–New York, 1977.

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

[45] P. Colli and A. Visintin, *On a class of doubly nonlinear evolution equations*, Comm. Partial Differential Equations, 15 (1990), pp. 737–756.

[46] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger, *A Riemannian interpolation inequality à la Borell, Brascamp and Lieb*, Invent. Math., 146 (2001), pp. 219–257.

[47] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger, *Prékopa-Leindler type inequalities on Riemannian manifolds, Jacobi fields, and optimal transport*, Ann. Fac. Sci. Toulouse Math. (6), 15 (2006), pp. 613–635.

[48] D. Cordero-Erausquin, B. Nazaret, and C. Villani, *A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities*, Adv. Math., 182 (2004), pp. 307–332.

[49] M. G. Crandall and T. M. Liggett, *Generation of semi-groups of nonlinear transformations on general Banach spaces*, Amer. J. Math., 93 (1971), pp. 265–298.

[50] G. Dal Maso, *An Introduction to Γ-Convergence*, Birkhäuser, Boston, 1993.

[51] E. De Giorgi, *Sulla convergenza di alcune successioni d'integrali del tipo dell'area*, Rend. Mat. (6), 8 (1975), pp. 277–294.

[52] E. De Giorgi, *Γ-convergenza e G-convergenza.*, Boll. Unione Mat. Ital., V. Ser., A 14 (1977), pp. 213–220.

[53] E. De Giorgi and T. Franzoni, *Su un tipo di convergenza variazionale*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8), 58 (1975), pp. 842–850.

[54] E. De Giorgi, A. Marino, and M. Tosques, *Problems of evolution in metric spaces and maximal decreasing curve*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8), 68 (1980), pp. 180–187.

[55] E. De Giorgi and S. Spagnolo, *Sulla convergenza degli integrali dell'energia per operatori ellittici del secondo ordine*, Boll. Un. Mat. Ital. (4), 8 (1973), pp. 391–411.

[56] E. DiBenedetto and R. E. Showalter, *Implicit degenerate evolution equations and applications*, SIAM J. Math. Anal., 12 (1981), pp. 731–751.

[57] J. Diestel, *Geometry of Banach Spaces-Selected Topics*, Lecture Notes in Mathematics, Vol. 485, Springer-Verlag, Berlin-New York, 1975.

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

[59] P.-A. Dionne, *Sur les problèmes de Cauchy hyperboliques bien posés*, J. Analyse Math., 10 (1962/63), pp. 1–90.

[60] P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford, at the Clarendon Press, 1947.

[61] J. Dixmier, *General Topology*, Springer-Verlag, New York, 1984.

[62] M. Dobrowolski, *Angewandte Funktionalanalysis*, Springer-Verlag, Heidelberg, 2nd ed., 2010.

[63] J. R. Dorroh, *A nonlinear Hille-Yosida-Phillips theorem*, J. Functional Analysis, 3 (1969), pp. 345–353.

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

[65] N. Dunford and J. T. Schwartz, *Linear Operators. Part I: General Theory*, Interscience Publ., New York, 1958.

[66] A. Eikmeier and E. Emmrich, *On a multivalued differential equation with nonlocality in time*, Vietnam J. Math., 48 (2020), pp. 703–718.

[67] A. Eikmeier, E. Emmrich, and H.-C. Kreusler, *Nonlinear evolution equations with exponentially decaying memory: existence via time discretisation, uniqueness, and stability*, Comput. Methods Appl. Math., 20 (2020), pp. 89–108.

[68] I. Ekeland, *On the variational principle*, J. Math. Anal. Appl., 47 (1974), pp. 324–353.

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

[70] J. Elstrodt, *Maß- und Integrationstheorie*, Springer-Verlag, Berlin, 4th ed., 2005.

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

[72] E. Emmrich and D. Puhst, *Survey of existence results in nonlinear peridynamics in comparison with local elastodynamics*, Comput. Methods Appl. Math., 15 (2015), pp. 483–496.

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

[74] 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 (2013), pp. 3719–3746.

[75] 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 (2015), pp. 1653–1701.

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

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

[78] E. Emmrich and G. Vallet, *Existence via time discretization for a class of doubly nonlinear operator-differential equations of Barenblatt-type*, J. Differential Equations, 254 (2013), pp. 2499–2514.

[79] E. Emmrich, D. Šiška, and A. Wróblewska-Kamińska, *Equations of second order in time with quasilinear damping: existence in Orlicz spaces via convergence of a full discretisation*, Math. Methods Appl. Sci., 39 (2016), pp. 2449–2460.

[80] M. Erbar, *A gradient flow approach to the boltzmann equation*, arxiv:1603.0540, (2019).

[81] L. C. Evans, *Weak convergence methods for nonlinear partial differential equations*, vol. 74 of CBMS Regional Conference Series in Mathematics, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1990.

[82] H. O. Fattorini, *The Cauchy problem*, vol. 18 of Encyclopedia of Mathematics and its Applications, Addison-Wesley Publishing Co., Reading, Mass., 1983.

[83] A. Friedman and J. Nečas, *Systems of nonlinear wave equations with nonlinear viscosity*, Pacific J. Math., 135 (1988), pp. 29–55.

[84] H. Gajewski, K. Gröger, and K. Zacharias, *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, Akademie-Verlag, Berlin, 1974.

[85] G. Gilardi and U. Stefanelli, *Time-discretization and global solution for a doubly nonlinear Volterra equation*, J. Differential Equations, 228 (2006), pp. 707–736.

[86] G. Gilardi and U. Stefanelli, *Existence for a doubly nonlinear Volterra equation*, J. Math. Anal. Appl., 333 (2007), pp. 839–862.

[87] O. Grange and F. Mignot, *Sur la résolution d'une équation et d'une inéquation paraboliques non linéaires*, J. Functional Analysis, 11 (1972), pp. 77–92.

[88] H. Grassmann, *Extension Theory*, vol. 19, American Mathematical Society, Providence, RI; London Mathematical Society, London, 2000.

[89] M. Grmela and H. C. Öttinger, *Dynamics and thermodynamics of complex fluids. I. Development of a general formalism*, Phys. Rev. E (3), 56 (1997), pp. 6620–6632.

[90] J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, Dover Publications, New York, 1953.

[91] M. Hazewinkel, ed., *Encyclopaedia of mathematics. Supplement. Vol. II*, Kluwer Academic Publishers, Dordrecht, 2000.

[92] E. Hille, *On the generation of semi-groups and the theory of conjugate functions*, Kungl. Fysiografiska Sällskapets i Lund Förhandlingar [Proc. Roy. Physiog. Soc. Lund], 21 (1952), p. 13.

[93] E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups*, American Mathematical Society Colloquium Publications, vol. 31, American Mathematical Society, Providence, R. I., 1957.

[94] V.-M. Hokkanen, *Existence of a periodic solution for implicit nonlinear equations*, Differential Integral Equations, 9 (1996), pp. 745–760.

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

[96] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal., 29 (1998), pp. 1–17.

[97] T. Kato, *Integration of the equation of evolution in a Banach space*, J. Math. Soc. Japan, 5 (1953), pp. 208–234.

[98] T. Kato and H. Tanabe, *On the analyticity of solution of evolution equations*, Osaka Math. J., 4 (1967), pp. 1–4.

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

[100] N. Kenmochi and M. Kubo, *Periodic solutions to a class of nonlinear variational inequalities with time-dependent constraints*, Funkcial. Ekvac., 30 (1987), pp. 333–349.

[101] N. Kenmochi and M. Kubo, *Periodic behavior of solutions to parabolic-elliptic free boundary problems*, J. Math. Soc. Japan, 41 (1989), pp. 625–640.

[102] B. T. Kien, *The normalized duality mapping and two related characteristic properties of a uniformly convex Banach space*, Acta Math. Vietnam., 27 (2002), pp. 53–67.

[103] M. Koike, M. Ôtani, and S. Uchida, *Periodic problem for doubly nonlinear evolution equation*, arXiv:2010.11570, (2020).

[104] A. Kraft and S. V. Gurevich, *Time-delayed feedback control of spatio-temporal self-organized patterns in dissipative systems*, in Control of Self-Organizing Nonlinear Systems, E. Schöll, S. Klapp, and P. Hövel, eds., Springer, 2016, pp. 413–430.

[105] J. Leray, *Hyperbolic differential equations*, The Institute for Advanced Study, Princeton, N. J., 1953.

[106] J.-L. Lions, *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires*, Dunod; Gauthier-Villars, Paris, 1969.

[107] J.-L. Lions and E. Magenes, *Problèmes aux Limites Non Homogènes et Applications. Vol. 1*, Travaux et Recherches Mathématiques, No. 17, Dunod, Paris, 1968.

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

[109] P.-L. Lions, *Solutions généralisées des équations de Hamilton-Jacobi du premier ordre*, C. R. Acad. Sci. Paris Sér. I Math., 292 (1981), pp. 953–956.

[110] G. Lumer and R. S. Phillips, *Dissipative operators in a Banach space*, Pacific J. Math., 11 (1961), pp. 679–698.

[111] E. Maitre and P. Witomski, *A pseudo-monotonicity adapted to doubly nonlinear elliptic-parabolic equations*, Nonlinear Anal., 50 (2002), pp. 223–250.

[112] J. Málek, J. Nečas, M. Rokyta, and M. Růžička, *Weak and measure-valued solutions to evolutionary PDEs*, vol. 13 of Applied Mathematics and Mathematical Computation, Chapman & Hall, London, 1996.

[113] A. Matas and J. Merker, *On doubly nonlinear evolution equations with non-potential or dynamic relation between the state variables*, J. Evol. Equ., 17 (2017), pp. 869–881.

[114] J. M. Mazón, J. D. Rossi, and J. Toledo, *Fractional p-Laplacian evolution equations*, J. Math. Pures Appl. (9), 105 (2016), pp. 810–844.

[115] A. Mielke, *Evolution of rate–independent inelasticity with microstructure using relaxation and Young measures*, in IUTAM Symposium on Computational Mechanics of Solid Materials at Large Strains (Stuttgart Aug. 2001), C. Miehe, ed., Kluwer, 2003, pp. 33–44.

[116] A. Mielke, *Evolution in rate-independent systems (Ch. 6)*, in Handbook of Differential Equations, Evolutionary Equations, vol. 2, C. Dafermos and E. Feireisl, eds., Elsevier B.V., Amsterdam, 2005, pp. 461–559.

[117] A. Mielke, *Deriving amplitude equations via evolutionary Γ-convergence*, Discr. Cont. Dynam. Systems Ser. A, 35 (2015), pp. 2679–2700.

[118] A. Mielke, *Deriving effective models for multiscale systems via evolutionary Γ-convergence*, in Control of Self-Organizing Nonlinear Systems, E. Schöll, S. Klapp, and P. Hövel, eds., Springer-Verlag, 2016, pp. 235–251.

[119] A. Mielke, *On evolutionary Γ-convergence for gradient systems*, in Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity, vol. 3 of Lect. Notes Appl. Math. Mech., Springer-Verlag, 2016, pp. 187–249.

[120] A. Mielke, R. Rossi, and G. Savaré, *Modeling solutions with jumps for rate-independent systems on metric spaces*, Discr. Cont. Dynam. Systems Ser. A, 25 (2009), pp. 585–615.

[121] A. Mielke, R. Rossi, and G. Savaré, *BV solutions and viscosity approximations of rate-independent systems*, ESAIM Control Optim. Calc. Var., 18 (2012), pp. 36–80.

[122] A. Mielke, R. Rossi, and G. Savaré, *Nonsmooth analysis of doubly nonlinear evolution equations*, Calc. Var. Part. Diff. Eqns., 46 (2013), pp. 253–310.

[123] A. Mielke, R. Rossi, and G. Savaré, *Balanced-viscosity (BV) solutions to infinite-dimensional rate-independent systems*, J. Europ. Math. Soc., 18 (2016), pp. 2107–2165.

[124] A. Mielke and T. Roubíček, *Rate-Independent Systems: Theory and Application*, Springer-Verlag, New York, 2015.

[125] U. Mosco, *Approximation of the solutions of some variational inequalities*, Ann. Scuola Norm. Sup. Pisa, 21 (1967), pp. 373–394. Erratum, ibid. (3) 765.

[126] M. Ôtani, *Nonmonotone perturbations for nonlinear parabolic equations associated with subdifferential operators, Cauchy problems*, J. Differential Equations, 46 (1982), pp. 268–299.

[127] M. Ôtani, *Nonmonotone perturbations for nonlinear parabolic equations associated with subdifferential operators, periodic problems*, J. Differential Equations, 54 (1984), pp. 248–273.

[128] F. Otto, *Dynamics of labyrinthine pattern formation in magnetic fluids: a mean-field theory*, Arch. Rational Mech. Anal., 141 (1998), pp. 63–103.

[129] F. Otto, *Evolution of microstructure in unstable porous media flow: a relaxational approach*, Comm. Pure Appl. Math., 52 (1999), pp. 873–915.

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

[131] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal., 173 (2000), pp. 361–400.

[132] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, vol. 44 of Applied Mathematical Sciences, Springer-Verlag, New York, 1983.

[133] P. Plecháč and T. Roubíček, *Visco-elasto-plastic model for martensitic phase transformation in shape-memory alloys*, Math. Methods Appl. Sci., 25 (2002), pp. 1281–1298.

[134] D. Puhst, *On the evolutionary fractional p-Laplacian*, Appl. Math. Res. Express. AMRX, (2015), pp. 253–273.

[135] K. R. Rajagopal and T. Roubíček, *On the effect of dissipation in shape-memory alloys*, Nonlinear Anal. Real World Appl., 4 (2003), pp. 581–597.

[136] M. M. Rao and Z. D. Ren, *Theory of Orlicz Spaces*, vol. 146 of Monographs and Textbooks in Pure and Applied Mathematics, Marcel Dekker, Inc., New York, 1991.

[137] R. T. Rockafellar, *Characterization of the subdifferentials of convex functions*, Pacific J. Math., 17 (1966), pp. 497–510.

[138] R. T. Rockafellar, *Integrals which are convex functionals*, Pacific J. Math., 24 (1968), pp. 525–539.

[139] R. T. Rockafellar, *Convex Analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970.

[140] R. T. Rockafellar, *Convex integral functionals and duality*, in Contributions to nonlinear functional analysis (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1971), 1971, pp. 215–236.

[141] R. T. Rockafellar, *Integrals which are convex functionals, II*, Pacific J. Math., 39 (1971), pp. 439–469.

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

[143] R. Rossi and G. Savaré, *Gradient flows of non convex functionals in Hilbert spaces and applications*, ESAIM Control Optim. Calc. Var., 12 (2006), pp. 564–614.

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

[145] T. Roubíček, *Nonlinear Partial Differential Equations with Applications*, vol. 153 of International Series of Numerical Mathematics, Birkhäuser, Basel, 2nd ed., 2013.

[146] A. M. Ruf, *Convergence of a full discretization for a second-order nonlinear elastodynamic equation in isotropic and anisotropic Orlicz spaces*, Z. Angew. Math. Phys., 68 (2017), pp. Paper No. 118, 24.

[147] E. Sandier and S. Serfaty, *Gamma-convergence of gradient flows with applications to Ginzburg-Landau*, Comm. Pure Appl. Math., LVII (2004), pp. 1627–1672.

[148] G. Schimperna, A. Segatti, and U. Stefanelli, *Well-posedness and long-time behavior for a class of doubly nonlinear equations*, Discr. Cont. Dynam. Systems Ser. A, 18 (2008), pp. 15–38.

[149] E. Schrödinger, *Quantisierung als eigenwertproblem*, Annalen der Physik, 384 (1926), pp. 361–376.

[150] L. Schwartz, *Les équations d'évolution liées au produit de composition*, Ann. Inst. Fourier (Grenoble), 2 (1950), pp. 19–49 (1951).

[151] T. Senba, *On some nonlinear evolution equation*, Funkcial. Ekvac., 29 (1986), pp. 243–257.

[152] S. Serfaty, *Gamma-convergence of gradient flows on Hilbert spaces and metric spaces and applications*, Discr. Cont. Dynam. Systems Ser. A, 31 (2011), pp. 1427–1451.

[153] M. S. Skaff, *Vector valued Orlicz spaces generalized N-functions. I*, Pacific J. Math., 28 (1969), pp. 193–206.

[154] M. S. Skaff, *Vector valued Orlicz spaces. II*, Pacific J. Math., 28 (1969), pp. 413–430.

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

[156] J. Swift and P. Hohenberg, *Hydrodynamic fluctuations at the convective instability*, Physical Review A, 15 (1977), pp. 319–328.

[157] R. Temam, *Navier-Stokes equations. Theory and Numerical Analysis*, vol. 2 of Studies in Mathematics and its Applications, North-Holland Publishing Co., Amsterdam-New York, revised ed., 1979.

[158] C. Villani, *Optimal Transport. Old and New*, vol. 338 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 2009.

[159] J. W. von Goethe, *Faust. Eine Tragödie von Goethe*, J. G. Cotta, Tübigen, 1808.

[160] J. Wloka, *Partielle Differentialgleichungen*, B. G. Teubner, Leipzig, 1982.

[161] Y. Yamada, *On evolution equations generated by subdifferential operators*, J. Fac. Sci. Univ. Tokyo Sect. IA Math., 23 (1976), pp. 491–515.

[162] K. Yosida, *On the differentiability and the representation of one-parameter semi-group of linear operators*, J. Math. Soc. Japan, 1 (1948), pp. 15–21.

[163] L. Young, *Generalized curves and the existence of an attained absolute minimum in the calculus of variations*, C. R. Soc. Sci. Varsovie, Cl. III, 30 (1937), pp. 212–234.

[164] E. Zeidler, *Nonlinear Functional Analysis and its Applications II/B*, Springer-Verlag, New York, 1990.

[165] M. Zemek, *Strong monotonicity and Lipschitz-continuity of the duality mapping*, Acta Univ. Carolin. Math. Phys., 32 (1991), pp. 61–64.