A Numerical Scheme for Rate-Independent Systems

Analysis and Realization

DISSERTATION

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

(DR. RER. NAT.)

Der Fakultät für Mathematik der

TECHNISCHEN UNIVERSITÄT DORTMUND

vorgelegt von

MICHAEL SIEVERS

IM AUGUST 2020

Dissertation
A Numerical Scheme for Rate-Independent Systems – Analysis and Realization

Fakultät für Mathematik
Technische Universität Dortmund

Erstgutachter: Prof. Dr. Christian Meyer
Zweitgutachter: Prof. Dr. Dorothee Knees

Tag der mündlichen Prüfung: 11.12.2020

Danksagung

Zuallererst gilt mein besonderer Dank meinem Betreuer Prof. Dr. Christian Meyer. Nicht nur für die Möglichkeit mich diesem spannenden Thema widmen zu können, sondern insbesondere für die zahlreichen mathematischen Impulse und die ständige Unterstützung in fachlicher wie auch persönlicher Hinsicht.
In diesem Kontext möchte ich mich auch bei all jenen bedanken, die mich in der Zeit an der TU begleitet haben, speziell bei der gesamten Kollegengruppe für die Atmosphäre und die vielen anregenden Diskussionen.
Darüber hinaus gilt mein Dank Prof. Dr. Dorothee Knees sowohl für die Übernahme des Zweitgutachtens als auch für viele hilfreiche Kommentare und Antworten auf meine Fragen.
Außerhalb der Universität danke ich natürlich meinen Freunden und meiner Familie, ganz besonders meinen Eltern, ohne die all dies gar nicht erst möglich gewesen wäre. Danke für die Unterstützung (in jeglicher Hinsicht) und den Anstoß Mathematik zu studieren.
Zu guter Letzt gebührt mein Dank Ina, die zu jeder Zeit für mich da war, in allen Studienphasen, und es dabei sicher nicht immer leicht mit mir hatte.

Und falls ich nun doch jemanden vergessen habe, sei allen, die sich angesprochen fühlen, gesagt: Vielen Dank!

Zusammenfassung

Viele Materialien im Bereich der Kontinuumsmechanik lassen sich in guter Näherung als, zumindest teilweise, raten-unabhängig ansehen. Solche Systeme sind in aller Regel getrieben durch äußere Kräfte und dabei unabhängig von deren Geschwindigkeit (Rate) wohl aber abhängig von deren Bewegungsrichtung. In der vorliegenden Dissertation werden wir im Speziellen solche raten-unabhängigen Systeme betrachten, die sich mittels einer, im Allgemeinen, nichtkonvexen Energie und einer positiv homogenen Dissipation beschreiben lassen. Die Kombination beider Eigenschaften erlaubt, trotz sich gleichmäßig entwickelnder äußerer Kräfte, die Ausbildung abrupter Zustandsänderungen. Mathematisch gesehen bedeutet dies, dass sich zeitliche Unstetigkeiten (Sprünge) entwickeln können. Um solche Phänomene widerzuspiegeln, sind zunächst geeignete (schwache) Lösungsbegriffe notwendig.
Wir werden uns daher im ersten Abschnitt dieser Dissertation mit eben solchen Lösungskonzepten beschäftigen. Neben den mittlerweile vielfältig untersuchten *energetischen Lösungen* werden wir uns dabei im Besonderen den sogenannten *parametrisierten Lösungen* zuwenden, die fortan im Fokus dieser Arbeit stehen. Die wesentliche Idee dieses Lösungskonzeptes besteht grob gesagt darin, die zeitlichen Unstetigkeiten nicht als instantane Änderung von einem Zustand in den nächsten aufzufassen, sondern als (sehr schnell durchlaufene) kontinuierliche Verbindung zwischen eben diesen zwei Zuständen. Die entsprechende Verbindungslinie bildet folglich einen integralen Bestandteil dieses Lösungskonzeptes. Man kann dies auch als eine Art Parametrisierung der gesamten Lösungskurve inklusive seiner Sprünge auffassen, was nebenbei auch den Namen *parametrisierte Lösungen* begründet. Mit der Einführung dieses Lösungsbegriffs stellt sich in natürlicher Weise die Frage nach der Existenz (und Eindeutigkeit) von Lösungen dieses Typs. In der Regel gibt es hierzu zwei unterschiedliche Ansätze. Einerseits ließe sich die Ausgangsgleichung geeignet regularisieren und eine anschließende Grenzwertanalyse durchführen. Im Kontext raten-unabhängiger Systeme bezeichnet man dieses Vorgehen auch als *vanishing viscosity*. Auf der anderen Seite - und aus anwendungsspezifischer Sicht der womöglich interessantere Ansatz - steht die Diskretisierung des ursprünglichen Problems mittels eines geeigneten Zeitschritt-Verfahrens.
In Anlehnung an das in [EM06] eingeführte *local incremental minimization scheme*, werden wir im zweiten Abschnitt dieser Arbeit ein in Zeit und Ort diskretes Schema zur Lösung eines raten-unabhängigen Systems analysieren. Im Vergleich zum vorher genannten Verfahren müssen hierbei lediglich stationäre Punkte anstatt (lokaler) Minima bestimmt werden. Damit ist dieses Schema, welches wir im Weiteren *local incremental stationarity scheme* (LISS) nennen, äußerst zugänglich für numerische Verfahren. Neben dieser praktischen Anwendbarkeit, liefert unsere Konvergenzanal-

yse für LISS gleichfalls die Existenz parametrisierter Lösungen und zwar auch bei unbeschränkter Dissipation. Damit lassen sich mittels des *local incremental stationarity schemes* selbst unidirektionale Prozesse, wie sie beispielweise in Schädigungsmodellen auftreten, approximieren.

Im dritten Abschnitt dieser Arbeit werden wir uns dann mit a priori Fehlerabschätzungen für LISS beschäftigen, wobei wir hier auf eine Diskretisierung im Ort verzichten. Ein entscheidender Bestandteil dabei ist die uniforme Konvexität der Energie. Ohne diese sind Lösungen des raten-unabhängigen System im Allgemeinen weder eindeutig noch stetig, sodass Konvergenzraten in diesem Fall nicht zu erwarten sind. Unter hinreichenden Konvexitätsannahmen hingegen werden wir Konvergenzraten der Ordnung $\mathcal{O}(\sqrt{\tau})$ für ein allgemeines Setting und der Ordnung $\mathcal{O}(\tau)$ für den Fall einer semilinearen Energie beweisen. Außerdem erweitern wir letzteres Resultat auf den sogenannten lokalkonvexen Fall, bei dem sich die Lösungstrajektorie nur auf einen Bereich uniformer Konvexität der Energie beschränkt.

Der letzte Abschnitt dient schlussendlich der Darlegung einer möglichen Realisierung des eingeführten Approximationsschematas LISS, sowie der Visualisierung der numerischen Ergebnisse. Insbesondere die im dritten Abschnitt gefundenen Konvergenzordnungen werden hier durch die numerischen Bespiele gestützt.

Abschließend sei erwähnt, dass Teile dieser Arbeit bereits in Veröffentlichungen erschienen sind. Dies betrifft insbesondere den Abschnitt 3.2, welches in weiten Teilen auf den Ausarbeitungen in [MS19a] basiert, sowie Kapitel 3.3, welches in [MS20] erschienen ist.

Abstract

Many materials in the field of continuum mechanics can be considered, at least in parts, as rate-independent. Such systems are generally driven by external forces and thereby independent of their speed (rate) but still dependent on their direction. In this dissertation, we will consider those rate-independent systems that can be described by means of a, in general, nonconvex energy and a positively homogeneous dissipation. Both properties in combination allow the formation of abrupt changes in state, even if the external forces evolve smoothly. Mathematically speaking, this means that temporal discontinuities (jumps) may develop. In order to be able to reflect such phenomena, suitable (weak) notions of solutions are required.

The first section of this dissertation is therefore devoted to the presentation of precisely such solution concepts. In addition to the *energetic solutions*, which are by now widely known and analyzed, we will particularly focus on the so-called *parametrized solutions*. The essential idea of this solution concept is to resolve the possible temporal discontinuities by making the path from one state to the other an integral part of the solution. This can also be seen as a kind of parameterization of the solution curve including its jumps, which also explains the name *parametrized solutions*. Following the introduction of this solution concept, the question of the existence of solutions of this type naturally arises. There are usually two different approaches to this. On the one hand, one may regularize the initial equation appropriately and perform a limit analysis afterwards. In the context of rate-independent systems, this is commonly referred to as vanishing viscosity. On the other hand - and from an application point of view the possibly more interesting approach - one can discretize the original problem by means of a suitable time step method.

In the second section of this work, we will analyze a scheme that provides a discretization in time and space based on the *local incremental minimization scheme* introduced in [EM06]. The main difference compared to the one in [EM06] is that, instead of solving a minimization problem, only stationary points have to be determined here. This scheme, which we will subsequently denote as *local incremental stationarity scheme* (LISS), is therefore well suited for numerical methods. Beyond this practical applicability, the convergence analysis for LISS also provides the existence of parametrized solutions, even in the case of an unbounded dissipation. Therefore, the *local incremental stationarity scheme* can also be used to approximate unidirectional processes, such as those that occur in damage models for example.

In the third section of this work, we will then deal with a priori error estimates for LISS, whereby we do not incorporate the discretization in space here. A crucial assumption in this context is the uniform convexity of the energy. Without this, solutions are, in general, not unique and not

continuous, so that convergence rates are not to be expected in this case. With sufficient convexity assumptions, on the other hand, we will prove convergence rates of the order $\mathcal{O}(\sqrt{\tau})$ for a general setting and of the order $\mathcal{O}(\tau)$ in case of a semilinear energy. We also extend the latter result to the so-called locally convex case, in which the solution trajectory remains in an area of uniform convexity of the energy.
The last section, finally, is devoted to the presentation of a possible realization of the introduced approximation scheme LISS and to visualize the numerical results. In particular, we will provide several examples that illustrate the theoretical findings of the preceding section.
After all, it should be mentioned that parts of this work have already appeared in publications. This particularly applies to the Section 3.2, which is based on the elaborations in [MS19a], and the Section 3.3, which has been published in [MS20].

Contents

List of symbols

(E)	energy balance 26
(S)	stability condition 26
$(\cdot)'$	time derivative of a function from $\mathbb{R}$ to X
$\langle \cdot, \cdot \rangle_{X^*,X}$	dual pairing 17
a.e.	almost everywhere; everywhere except on a set of Lebesgue measure zero
$BV(0,T;X)$	space of functions from $[0,T]$ to X of bounded variation, Definition A.1.1
$C(z)$	continuity set of the function z 43
$C(0,T;X)$	space of continuous functions from $[0,T]$ to X with norm $\|z\|_{C(0,T;X)} = \max_{t\in[0,T]} \|z(t)\|_X$
$\overline{C}$	closure of a set C
D_z	Fréchet derivative
∂	convex subdifferential 131
∂_t	partial time derivative
$\mathrm{Diss}_{\mathcal{R}}$	dissipation w.r.t. $\mathcal{R}$ 27
$\mathrm{Diss}_{\mathfrak{p}}$	dissipation w.r.t. the distance $\Delta_{\mathfrak{p}}$ 44
dom	domain of a convex function 132
$\hookrightarrow$	embedding
$\hookrightarrow^c$	compact embedding
$\hookrightarrow^d$	dense embedding
f.a.a.	for almost all; for all elements except for a set of Lebesgue measure zero
I_C	indicator function of the set C
I_τ	indicator function of the set $\overline{B_{\mathbb{V}}(0,\tau)}$ 58
$\mathcal{I}$	energy functional 15
$J(z)$	jump set of the function z 43
$J_{\mathcal{V}}$	Riesz isomorphism $J_{\mathcal{V}} : \mathcal{V} \to \mathcal{V}^*$
$\lvert\cdot\rvert_{Lip}$	Lipschitz constant
$L^p(0,T;X)$	Bochner-Lebesgue space of functions from $[0,T]$ to X
P_h	Ritz-projection 57
$\mathfrak{p}$	vanishing viscosity contact potential 36, 44

Π_h	Orthogonal projection onto finite dimensional subspace 56
$\mathfrak{P}$	projection for parametrized solutions 43
RIS	rate-independent system, see RIS
$\mathcal{R}$	dissipation potential 16
$\mathcal{R}^*$	Fenchel conjugate functional of $\mathcal{R}$, cf. Definition A.3.2
$\mathcal{S}_{glob}(t)$	global stability set 21
$S_{\tau,h}$	artificial end time at discretization fineness τ, h 71
$\mathcal{S}_{loc}(t)$	local stability set 21
sgn	multivalued sign operator 14
$\mathbb{V}$	self-adjoint, coercive operator on $\mathcal{V}$ 15
Var_X	variation w.r.t. the norm in X 125
$W^{1,p}(0,T;X)$	Bochner-Sobolev space of functions $[0,T] \to X$
$z(t^+)$	right-hand sided limit of z, i.e., $z(t^+) = \lim_{s\to t,\, s>t} z(s)$ 125
$z(t^-)$	left-hand sided limit of z, i.e., $z(t^-) = \lim_{s\to t,\, s<t} z(s)$ 125

Chapter 1

Introduction

The effect of rate-independence occurs in various different areas of mechanics. Coulomb, for instance, already asserted around 1781 that the kinetic friction is independent of the sliding velocity[1]. About 100 years later, mechanical engineers and mathematicians started to investigate the elasto-plastic behavior of different materials as well as its mathematical description and came across a similar phenomenon. Namely, they observed that the plastic deformation does not depend on the velocity with which the external force is applied, see, e.g., [Tre64, SV70, HK09, Mis13]. Rate-independence thus describes the fact that a system is independent of the rate at which some external force is applied while it might still depend on the direction of the force. The actual term *rate-independence*, however, first came up in the middle of the 1960's in two different definitions by Pipkin and Rivlin [PR65] as well as Truesdell and Noll [TN65][2], which turned out to be equivalent soon after; cf. [OW68]. It lasted another few years until Moreau brought this field of research on a mathematical fundament using techniques known from convex analysis, see [Mor70, Mor71]. Based on this, Halphen and Nguyen introduced the concept of so-called *standard generalized materials* (see [HN75]) and therewith provided a general setting for the modeling of materials. From this point on lots of further research areas have been investigated and the notion of rate-independent systems has been established as a well suited formalism describing real life problems in the field of elastoplasticity, damage and shape-memory, to only mention a few (see, e.g., [KRZ13, FM06, Mai04, AMS08, MM09]). The interested reader is also referred to the book of Mielke and Roubíček [MR15], which provides a good starting point for further insights into the field of rate-independent systems and includes a multitude of further applications.

One main characteristic of rate-independent systems is the fact that changes in the state are solely driven by an external force. Besides, as the name already suggests, the system is independent of the rate at which the loading is applied, that is to say, whenever z is a solution to some external load ℓ, then $z \circ \alpha$ is a solution to $\ell \circ \alpha$ for every monotone increasing function α. Before we actually start with the general setting, we take a look at a simple and prototypical example for such a rate-independent system, which, in this case, is also known under the name *sweeping process*, see

[1] The original work dates back to the year 1781. In the revised version [Cou21, p. 42] from 1821 it says: " [...] il paraît que, dans tous les cas de pratique, l'on peut regarder le frottement comme étant indepéndant du degré de vitesse."

[2] To be precise, they used the term "hyper-elasticity".

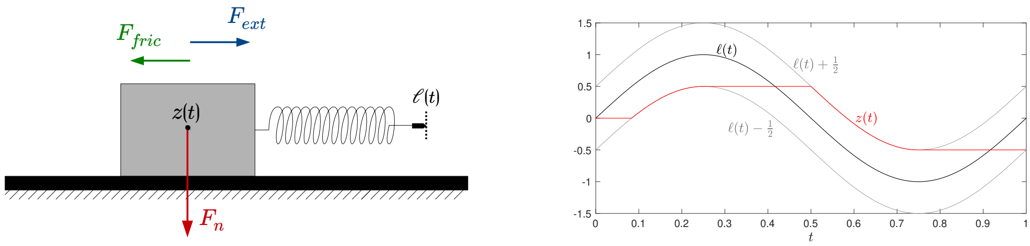


Figure 1.0.1: Left: Examplary setting of a sliding block. Right: Solution for the external load $\ell(t) = \sin(2\pi t)$ and ratio $\mu = \frac{1}{2}$.

[Mor77]. For a realization, we put a block on a solid surface and attach a spring to one of its sides, see Figure 1.0.1. The extension of the spring is prescribed by the external load $\ell(t)$. Using Hooke's law, the stored energy in the system is described by $\mathcal{I}(t, z) = \frac{\kappa}{2}(z - \ell(t))^2$, wherein κ denotes the characteristic stiffness factor of the spring. Due to the friction between the block and the surface, a certain frictional force must be overcome in order to move the block. According to Coulomb's law, this frictional force is proportional to the normal force (normal w.r.t. the surface) but independent of the velocity and opposite to its direction of motion. The behavior of the system can thus be described by the following cases:

- As long as the magnitude of the external force F_{ext} is smaller than the magnitude of the frictional force F_{fric}, i.e., $|F_{ext}| \leq |F_{fric}|$, the block stays in its position. Since the frictional force is proportional to the normal force, again by Coulomb's law, we end up with the inequality $|F_{ext}| \leq \mu|F_n|$, where μ denotes the coefficient of friction, i.e., the ratio between F_{fric} and F_n.

- As the external force reaches the frictional force, the block starts moving. However, during the movement, the forces F_{fric} and F_{ext} remain in equilibrium, i.e., $F_{ext} = -F_{fric}$. Once more, Coulomb's law implies that this frictional force is always exerted in a direction opposite to the movement. Moreover, it is independent of the velocity, so that $F_{fric} = -\mu \operatorname{sgn}(z')|F_n|$. Combining these properties we have $F_{ext} = \mu \operatorname{sgn}(z')|F_n|$ whenever $z' \neq 0$.

In sum, we have the two observations:

- if $z' = 0$, then $|F_{ext}| \leq \mu|F_n|$,

- if $z' \neq 0$, then $F_{ext} = \mu \operatorname{sgn}(z')|F_n|$.

In particular, we see that the external force never exceeds the frictional one. For simplicity, let us assume that the block is pressed onto the ground only by the gravitation and that the corresponding gravitational force is equal to -1, which implies that $F_n = -1$. Including $F_{ext} = -D_z\mathcal{I}(t, z) = \ell(t) - z$, both cases are easily brought together into one differential inclusion

$$0 \in \mu \operatorname{sgn}(z') + z - \ell(t), \quad z(0) = 0, \tag{1.0.1}$$

wherein $\operatorname{sgn} : \mathbb{R} \rightrightarrows \mathbb{R}$ denotes the multivalued sign-operator, that is,

$$\operatorname{sgn}(x) = \begin{cases} 1, & \text{if } x > 0, \\ [-1, 1], & \text{if } x = 0, \\ -1, & \text{if } x < 0. \end{cases}$$

Except for the fact that (1.0.1) is an inclusion rather than an equation this still has the form of a typical force balance law. The term $\mu \operatorname{sgn}(z')$ describes the dissipative forces, in this case the frictional force, while the term $z - \ell(t)$ corresponds to the (external) potential force coming from the energy functional.

Of course, this example is greatly simplified - e.g., we tacitly assumed that the coefficient of friction is the same for static and kinetic friction, which does not hold true in most of the cases. However, this example suffices to illustrate some important features as well as the main structure for rate-independent systems that will occur in this thesis. For the specific loading $\ell(t) = \sin(2\pi t)$ and coefficient $\mu = \frac{1}{2}$, Figure 1.0.1 shows the solution for the rate-independent system (1.0.1). We observe that it is Lipschitz continuous but does not provide further (temporal) regularity. Moreover, we see that the solution follows the external load with some delay. Both of these observations are well-known phenomenons for rate-independent systems.

Now, bringing (1.0.1) into a broader setting, we consider the differential inclusion

$$0 \in \partial\mathcal{R}(z'(t)) + D_z\mathcal{I}(t, z(t)), \qquad t \in [0, T], \quad z(0) = z_0. \tag{RIS}$$

As indicated above, one can still see this inclusion as a balance of forces, which means that the dissipative force $\partial\mathcal{R}$ and the potential force $-D_z\mathcal{I}(t, z)$ must annihilate each other. Beyond that, we note that the potential energy only depends on the time t and the state z and is thus, in particular, independent of the velocity z'. Physically, one can interpret this as a quasistatic description, which means that the load is applied slowly enough so that inertial forces can be ignored. In the above example of the moving block, this means that the spring is pulled slowly along the path so that the magnitude of the acceleration is negligible. The characteristic feature of the formulation in (RIS) is the positive 1-homogeneity of the dissipation $\mathcal{R}$. It is this property which induces that (RIS) is rate-independent. Indeed, it can be easily shown (see Lemma 2.1.1) that, in this case, the subdifferential $\partial\mathcal{R}(\cdot)$ is positively 0-homogenous, i.e., $\partial\mathcal{R}(\lambda v) = \partial\mathcal{R}(v)$ for all $\lambda > 0$. Thus, assuming that $z(t)$ is a sufficiently regular solution and $\alpha : [0, S] \to [0, T]$ is any (differentiable) monotone rescaling of the time, we obtain by the chain rule that $(z \circ \alpha)'(s) = (z' \circ \alpha(s))\, \alpha'(s)$. Since α is monotone, we have $\alpha' \geq 0$, so that the positive 0-homogeneity of $\partial\mathcal{R}(\cdot)$ implies that $\tilde{z}(s) = (z \circ \alpha)(s)$ is also a solution but to the external load $\ell \circ \alpha$. Nevertheless, at first glance, it seems slightly more restrictive to consider only dissipative forces that can be represented by a potential $\mathcal{R}$. We might more generally allow for set-valued, monotone and 0-homogeneous operators $A : \mathcal{Z} \rightrightarrows \mathcal{Z}^*$. However, it can be shown that if such a mapping A is given, which is additionally maximal monotone, there exists a proper, lower semicontinuous, convex and 1-homogeneous function $\mathcal{R} : \mathcal{Z} \to \mathbb{R} \cup \{\infty\}$ such that $A(v) = \partial\mathcal{R}(v)$, see [MR15, Prop. 3.2.1]. Thus, this requirement is natural in the context of rate-independent systems. Though, we do not want to hold back that there exist applications which do not exhibit the structure from (RIS). This applies, for example, to the model of crack growth in brittle material, cf. [DMFT05]. We will briefly elaborate on this at the end of Section 2.3 about energetic solutions. Now, before we proceed with the assumptions on the involved quantities, we want to give a further example dealing with the case of an energy that is nonconvex in z.

Example 1.0.1. [Existence of discontinuous solutions] In this example, we consider the nonconvex energy $\mathcal{I}: \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ with $\mathcal{I}(t,z) = \mathcal{E}(z) - tz$, wherein $\mathcal{E}(z) = 2|z|^3 - \frac{5}{2}z^2 + (\frac{5}{6})^3$ and the dissipation $\mathcal{R}(z) = |z|$. By direct calculations, one obtains the following solution for (RIS) (we will see in the next section, that this is in fact a so-called *BV solution*):

$$z(t) = \begin{cases} -\frac{1}{2}, & t \in [0, \frac{3}{2}), \\ -\frac{1}{6}(2 + \sqrt{10-6t}), & t \in [\frac{3}{2}, \frac{5}{3}), \\ \frac{1}{6}(2 + \sqrt{6t-2}), & t \in [\frac{5}{3}, 2]. \end{cases}$$

In particular, we observe that this solution is no longer continuous but performs a jump at $t = 5/3$ from $z^- = -\frac{1}{3}$ to $z^+ = \frac{1}{3}(1+\sqrt{2})$. This missing regularity makes it necessary to develop appropriate concepts of solutions that are able to handle these kind of discontinuities.

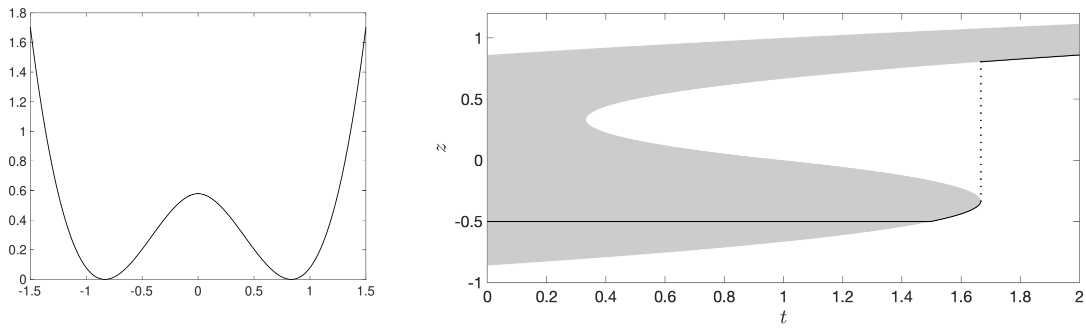


Figure 1.0.2: Left: Energy landscape for $t = 0$. Right: Solution z of the nonconvex energy $\mathcal{I}$ depending on the time t. The gray region is defined by $\{(t,z) : -D_z\mathcal{I}(t,z) \in \partial\mathcal{R}(0)\} = \cup_{t\in[0,2]}(t, \mathcal{S}_{loc}(t))$, where $\mathcal{S}_{loc}(t)$ is the so-called *set of local stability* (see (2.1.11)).

Assumptions

Let us now introduce the assumptions on the quantities in (RIS)[3]. We assume that $\mathcal{Z}$ and $\mathcal{X}$ are Banach spaces with $\mathcal{Z} \subset \mathcal{X}$. In parts of this thesis, we additionally require an intermediate Hilbert space $\mathcal{Z} \overset{c,d}{\hookrightarrow} \mathcal{V} \hookrightarrow \mathcal{X}$, where $\hookrightarrow^d$ and $\hookrightarrow^c$ refer to dense and compact embedding, respectively. We equip $\mathcal{V}$ with the norm $\|v\|_{\mathbb{V}} := \langle \mathbb{V}v, v\rangle^{1/2}_{\mathcal{V}^*,\mathcal{V}}$, where $\mathbb{V} \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$ is a self-adjoint and coercive operator, i.e., there exists a constant $\gamma > 0$ such that $\langle \mathbb{V}v, v\rangle_{\mathcal{V}^*,\mathcal{V}} \geq \gamma\|v\|^2_{\mathcal{V}}$ for all $v \in \mathcal{V}$, where $\|\cdot\|_{\mathcal{V}}$ denotes the natural norm associated with the scalar product in $\mathcal{V}$. Moreover, the energy $\mathcal{I}(t,z)$ is supposed to fulfill:

[3]The following assumptions are valid throughout the whole thesis and will be strenghtened, if necessary, at the appropriate positions.

(E1) $\mathcal{I} \in C^1([0,T] \times \mathcal{Z}; \mathbb{R})$.

(E2) *For all $t \in [0,T]$ the energy $\mathcal{I}(t,\cdot)$ is weakly lower semicontinuous and coercive on $\mathcal{Z}$ with $\mathcal{I}(t,z) \geq c_1 \|z\|_{\mathcal{Z}} - c_0$ for some constants $c_0, c_1 > 0$.*

(E3) *There exists $\beta > 0$ and $\mu \in L^1(0,T)$ with $\mu \geq 0$ such that for all $t \in [0,T]$:*

$$|\partial_t \mathcal{I}(t,z)| \leq \mu(t)(\mathcal{I}(t,z) + \beta) \quad \forall z \in \mathcal{Z}.$$

(E4) *For all sequences $t_k \to t$ and $z_k \rightharpoonup z$ in $\mathcal{Z}$ it holds:*

$$\partial_t \mathcal{I}(t_k, z_k) \to \partial_t \mathcal{I}(t,z).$$

Note that the combination of (E1)–(E2) already yields that, for all sequences $t_k \to t$ and $z_k \rightharpoonup z$ in $\mathcal{Z}$, it holds

$$\mathcal{I}(t,z) \leq \liminf_{k\to\infty} \mathcal{I}(t_k, z_k). \tag{1.0.2}$$

Regarding the dissipation $\mathcal{R} : \mathcal{X} \to [0,\infty]$, we assume that

(R1) *$\mathcal{R}$ is proper, convex and lower semicontinuous,*

(R2) *$\mathcal{R}$ is positively 1-homogeneous, i.e., $\mathcal{R}(\lambda v) = \lambda \mathcal{R}(v) \quad \forall v \in \mathcal{X}, \ \lambda > 0$,*

(R3) $\exists \rho > 0 : \ \rho \|v\|_{\mathcal{X}} \leq \mathcal{R}(v)$.

Combining the convexity and the positive 1-homogeneity of $\mathcal{R}$, it is easy to verify the following triangle inequality

$$\mathcal{R}(u-w) \leq \mathcal{R}(u-v) + \mathcal{R}(v-w) \quad \forall u,v,w \in \mathcal{Z}. \tag{1.0.3}$$

Note that, on the one hand, these conditions are formulated in such a way that they are general enough to be applied to various different examples, and, on the other hand, also sufficiently rigorous to allow for a concise description of the notions of solutions and their correlations. Depending on the actual setting and the solution concept slightly weaker forms of the above assumptions exist, see, e.g., [KZ18]. In further parts of this thesis, especially for the a priori error analysis in Section 3.3, we need to specify an additional condition on the energy, namely:

Definition 1.0.2 (κ-uniform convexity)**.** *We say that $\mathcal{I}$ is **κ-uniformly convex** if for all $t \in [0,T]$ it holds that $\mathcal{I}(t,\cdot) \in C^2(\mathcal{Z};\mathbb{R})$ and there exists a constant $\kappa > 0$ such that*

$$\langle D_z^2 \mathcal{I}(t,z)v, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \kappa \|v\|_{\mathcal{Z}}^2$$

for all $z, v \in \mathcal{Z}$.

This in particular implies that for every $\lambda \in [0,1]$ it holds

$$\mathcal{I}(t, \lambda z_1 + (1-\lambda) z_2) + \lambda(1-\lambda)\frac{\kappa}{2}\|z_1 - z_2\|_{\mathcal{Z}}^2 \leq \lambda \mathcal{I}(t,z_1) + (1-\lambda)\mathcal{I}(t,z_2) \quad \forall z_1, z_2 \in \mathcal{Z}. \tag{1.0.4}$$

In the same context we will need to assume that $\mathcal{I}(t,\cdot)$ provides slightly more regularity than $C^2(\mathcal{Z};\mathbb{R})$, precisely:

Definition 1.0.3. *We write $\mathcal{I}(t,\cdot) \in C^{2,1}_{loc}(\mathcal{Z};\mathbb{R})$ if for all $t \in [0,T]$ and all $r > 0$ there exists $C(r) \geq 0$, only depending on r, such that for all $z_1, z_2 \in B_{\mathcal{Z}}(0,r)$ it holds*

$$\langle \big[D_z^2\mathcal{I}(t,z_1) - D_z^2\mathcal{I}(t,z_2)\big]v, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq C(r)\|z_1 - z_2\|_{\mathcal{Z}}\|v\|^2_{\mathcal{Z}}. \tag{1.0.5}$$

Finally, we assume that the initial state z_0 satisfies $z_0 \in \mathcal{Z}$ and $0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(0,z_0)$.

Notation

In the subsequent chapters, we use the following notation:
Given two normed linear spaces X, Y, we denote by $\langle\cdot,\cdot\rangle_{X^*,X}$ the dual pairing and suppress the subscript if there is no risk for ambiguity. By $\|\cdot\|_X$, we denote the norm in X and by $\mathcal{L}(X,Y)$ the space of linear and bounded operators from X to Y. Furthermore, $B_X(x,r)$ is the open ball in X around $x \in X$ with radius $r > 0$. If X is embedded in Y we write $X \hookrightarrow Y$, as well as $X \overset{d}{\hookrightarrow} Y$ and $X \overset{c}{\hookrightarrow} Y$, if these embeddings are dense and compact, respectively. Given a convex functional $f : X \to \mathbb{R} \cup \{+\infty\}$, we denote the (convex) subdifferential of f at x by $\partial f(x) \subset X^*$ and its conjugate functional by $f^* : X^* \to \mathbb{R} \cup \{+\infty\}$. In addition, for a given function $f : \mathbb{R} \times X \to \mathbb{R}$ depending on time and space, we denote by $\partial_t f(t,x)$ its partial time derivative and by $D_z f(t,x)$ the Fréchet derivative with respect to the underlying space X. Moreover, we denote by z' the time derivative of a time dependent function $z : \mathbb{R} \to X$. Concerning the one-sided limit of functions, we let $t^+ := \lim_{s\to t, s>t} s$ and $z(t^+) := \lim_{s\to t, s>t} z(s)$ and t^- as well as $z(t^-)$ correspondingly. For $T > 0$ and $p \in [1,\infty]$ we furthermore write $L^p(0,T;X)$ for the Bochner space of p-integrable functions and $W^{1,p}(0,T;X)$ for the Bochner-Sobolev space. Additionally, the space of continuous functions and the space of functions of bounded variation mapping from $[0,T]$ into X are denoted by $C(0,T;X)$ and $BV(0,T;X)$, respectively. Beyond this, $|\Omega|$ stands for the Lebesgue measure of a set $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$. In the context of discretization, bold face letters always describe column vectors. Furthermore, c and C always stand for positive generic constants which may also change within one row. We will additionally write $C(\cdot)$ to highlight that the constant is dependent on some variable and use subscripts to denote specific fixed constants. Finally, we remark that, for the sake of clarity, we sometimes omit super- and subscripts for the dual pairing and/or variables if it becomes clear from the context.

Chapter 2

Solution concepts

In this chapter, we provide suitable notions of weak solutions for rate-independent systems in the form of (RIS) starting with the rather restrictive concept of differential solutions. These should be seen as a sort of strong solutions in the context of PDEs, i.e., they provide the necessary smoothness such that (RIS) holds for almost all $t \in [0, T]$. However, as indicated in Example 1.0.1, solutions might lack sufficient regularity to give (RIS) a reasonable meaning. In order to handle such cases, several (distinct) notions of solutions have been developed throughout the last 30 years, starting from the seminal paper [MT99]. Herein, the authors provide a general and widely used concept called *(global) energetic solutions*, which will be addressed in Section 2.3. In some sense opposite to this notion, we introduce the so-called *parametrized solutions* that are based on the work [EM06] and will form the fundamental concept of solutions throughout this thesis. We focus on this type of notion in the Section 2.4 and afterwards give a brief overview of two further concepts, namely *local* and *BV solutions*, in Section 2.5. Note that, in any case, while some of the conditions are supposed to hold almost everywhere in $[0, T]$, the solutions are defined for all $t \in [0, T]$. We conclude this chapter with some remarks on the relations between the presented concepts in Section 2.6.

2.1 Equivalent formulations for RIS

Before we actually enlarge upon the different concepts of solutions, we give several equivalent reformulations of the original problem

$$0 \in \partial \mathcal{R}(z'(t)) + D_z \mathcal{I}(t, z(t)), \quad z(0) = z_0. \tag{RIS}$$

On the one hand, this allows for a certain flexibility while handling (RIS) and, on the other hand, forms the basis for the definition of the different types of solutions. Note that, in all subsequent reformulations, we will tacitly assume that the involved terms are well-defined, in particular that $z'(t)$ and $-D_z \mathcal{I}(t, z(t))$ provide appropriate regularity (e.g., for the validity of the chain rule). We start by inserting the definition of the convex subdifferential, which directly leads us to the following *evolutionary variational inequality*

$$\langle D_z\mathcal{I}(t,z(t)), w - z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \mathcal{R}(w) - \mathcal{R}(z'(t)) \geq 0 \quad \forall w \in \mathcal{Z}, \tag{2.1.1}$$

which has to hold for almost all $t \in [0,T]$. Such time-dependent variational inequalities have been studied, for example, in [Kre99] or [BKS04] for some general settings as well as in [Wac11] for quasistatic plasticity. By employing standard results from convex analysis, we see that (RIS) can also be written as a *generalized gradient flow*, that is,

$$z'(t) \in \partial\mathcal{R}^*(-D_z\mathcal{I}(t,z(t))). \tag{2.1.2}$$

In this case, the choice $\mathcal{R}(v) = \frac{1}{2}\langle\mathbb{V}v, v\rangle_{\mathcal{V}}$ yields the known *(viscous) gradient flow*, cf. [RS06, AGS08, San17], which is no longer rate-independent. In fact, this reformulation is a useful access point in order to prove existence results for a viscous regularization of (RIS) such as (2.4.1), see, e.g., [MRS16, Prop. 4.17] or the argumentation in the proof of [MS19b, Thm. 3.23]. Another standard result from convex analysis, more specifically the Fenchel-Young inequality (A.3.2), additionally implies that (RIS) is equivalent to the sole inequality

$$\mathcal{R}(z'(t)) + \mathcal{R}^*(-D_z\mathcal{I}(t,z(t))) \leq \langle -D_z\mathcal{I}(t,z(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}. \tag{2.1.3}$$

Now, by exploiting a chain rule for the energy functional, precisely

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}(t,z(t)) = \langle D_z\mathcal{I}(t,z(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \partial_t\mathcal{I}(t,z(t)),$$

see also Appendix A.2.6, we find the upper energy estimate

$$\mathcal{I}(t,z(t)) + \int_0^t \mathcal{R}(z'(s)) + \mathcal{R}^*(-D_z\mathcal{I}(s,z(s)))\,\mathrm{d}s \leq \mathcal{I}(0,z_0) + \int_0^t \partial_t\mathcal{I}(s,z(s))\,\mathrm{d}s, \tag{2.1.4}$$

which is in fact an equality. Indeed, this energy inequality is already equivalent to (RIS). To see this, we exploit, again, the Fenchel-Young inequality as well as the chain rule to obtain from (2.1.4) that

$$\begin{aligned} &\langle -D_z\mathcal{I}(t,z(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq \mathcal{R}(z'(t)) + \mathcal{R}^*(-D_z\mathcal{I}(t,z(t))) \quad \text{a.e. in } [0,T] \\ \text{and} \quad &\int_0^T \mathcal{R}(z'(r)) + \mathcal{R}^*(-D_z\mathcal{I}(r,z(t)))\,\mathrm{d}t \leq \int_0^T \langle -D_z\mathcal{I}(t,z(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\,\mathrm{d}t\,, \end{aligned}$$

which immediately implies

$$\langle -D_z\mathcal{I}(t,z(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} = \mathcal{R}(z'(t)) + \mathcal{R}^*(-D_z\mathcal{I}(t,z(t))) \quad \text{a.e. in } [0,T],$$

and therefore also (RIS). This observation is useful in various existence proofs for solutions of rate-independent system and will also be taken up at some points in this thesis, see, e.g., Lemma 2.4.6.

In addition, all of the above reformulations do not rely on any special structure of $\mathcal{R}$ except convexity. To be more precise, $\mathcal{R}$ needs to be proper, lower semicontinuous and convex. However, for rate-independent systems, the dissipation provides an additional property, namely the positive 1-homogeneity. Taking into account this property, we obtain the following characteristics, which we will frequently exploit throughout the thesis.

Lemma 2.1.1. *Let $\mathcal{R} : \mathcal{Z} \to [0, \infty]$ be a positively $1-$homogeneous, convex and lower semicontinuous function. Then, for all $v \in \mathcal{Z}$ and $\xi \in \mathcal{Z}^*$,*

$$\partial\mathcal{R}(v) \subset \partial\mathcal{R}(0), \tag{2.1.5}$$

$$\partial\mathcal{R}(0) = \{\eta \in \mathcal{Z}^* \,:\, \mathcal{R}(w) \geq \langle \eta, w\rangle_{\mathcal{Z}^*,\mathcal{Z}} \quad \forall w \in \mathcal{V}\}, \tag{2.1.6}$$

$$\partial\mathcal{R}(v) = \{\eta \in \partial\mathcal{R}(0) \,:\, \langle \eta, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} = \mathcal{R}(v)\}, \tag{2.1.7}$$

$$\mathcal{R}^*(\xi) = I_{\partial\mathcal{R}(0)}(\xi) \tag{2.1.8}$$

holds. In particular, the subdifferential $\partial\mathcal{R}(\cdot)$ is 0-homogeneous.

Proof. By the definition of the subdifferential we have that

$$\xi \in \partial\mathcal{R}(v) \quad \Longleftrightarrow \quad \mathcal{R}(w) \geq \mathcal{R}(v) + \langle \xi, w - v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \qquad \forall w \in \mathcal{Z}\,. \tag{2.1.9}$$

Thus, the statement (2.1.6) directly follows from $\mathcal{R}(0) = 0$. Testing (2.1.9) with $w = 0$ and $w = 2v$ and exploiting the positive 1-homogeneity of $\mathcal{R}$ implies (2.1.7). Thereby, (2.1.5) is an easy consequence of the characterizations in (2.1.7) and (2.1.6). Finally, exploiting once more the 1-homogeneity of $\mathcal{R}$, we obtain

$$\begin{aligned}
\mathcal{R}^*(\xi) &= \sup_{v\in\mathcal{Z}}(\langle \xi, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} - \mathcal{R}(v)) \\
&= \sup_{\alpha>0}\sup_{v\in\mathcal{Z}}(\langle \xi, \alpha v\rangle_{\mathcal{Z}^*,\mathcal{Z}} - \mathcal{R}(\alpha v)) \\
&= \sup_{\alpha>0}\alpha \sup_{v\in\mathcal{Z}}(\langle \xi, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} - \mathcal{R}(v)) \\
&= \begin{cases} 0, & \text{if } \langle \xi, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} - \mathcal{R}(v) \leq 0 \quad \forall v \in \mathcal{Z} \\ +\infty, & \text{else} \end{cases} \\
&= I_{\partial\mathcal{R}(0)}(\xi),
\end{aligned}$$

which proves (2.1.8). The $0-$homogeneity of $\partial\mathcal{R}(\cdot)$ is now a direct consequence of the statement in (2.1.7) and the 1-homogeneity of $\mathcal{R}$. □

This characterization of the subdifferential of $\mathcal{R}$ allows us to reformulate (RIS) by means of the following two conditions, which have to hold almost everywhere in $[0, T]$:

$$-D_z\mathcal{I}(t, z(t)) \in \partial\mathcal{R}(0), \tag{2.1.10a}$$

$$\mathcal{R}(z'(t)) = \langle -D_z\mathcal{I}(t, z(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}. \tag{2.1.10b}$$

The inclusion (2.1.10a) is referred to as the *condition of local stability* and the corresponding state z satisfying (2.1.10a) is said to be *locally stable.* In this context we also define

$$\mathcal{S}_{loc}(t) := \{z \in \mathcal{Z} \,:\, -D_z\mathcal{I}(t,z) \in \partial\mathcal{R}(0)\}, \tag{2.1.11}$$

the so-called *set of local stability*.

Let us additionally note that by applying the chain rule from (A.2.6) in (2.1.10b) and integrating the resulting term, we find that (RIS) can be written as

$$\mathcal{R}(v) \geq \langle -D_z\mathcal{I}(t,z(t)), v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \quad \forall v \in \mathcal{Z} \quad \text{f.a.a. } t \in [0,T], \tag{2.1.12a}$$

$$\mathcal{I}(t,z(t)) + \int_0^t \mathcal{R}(z'(s))\,\mathrm{d}s = \mathcal{I}(0,z_0) + \int_0^t \partial_t\mathcal{I}(s,z(s))\,\mathrm{d}s \quad \forall t \in [0,T]. \tag{2.1.12b}$$

Again, it suffices to require "$\leq$" in (2.1.12b). This reformulation, in particular the energy inequality, is the basis for most of the solution concepts for rate-independent systems. Moreover, it forms the fundament for the convergence analysis of approximation schemes, cf. Section 3.2.4. Lastly, in view of Section 2.3, we highlight that in the case of a convex energy $\mathcal{I}(t,\cdot)$ condition (2.1.10a) is equivalent to

$$\mathcal{I}(t,z(t)) \leq \mathcal{I}(t,z) + \mathcal{R}(z - z(t)) \quad \forall z \in \mathcal{Z}. \tag{2.1.13}$$

In this context, we define the *set of global stability* as

$$\mathcal{S}_{glob}(t) := \{z \in \mathcal{Z} \,:\, \mathcal{I}(t,z) \leq \mathcal{I}(t,v) + \mathcal{R}(v-z) \quad \forall v \in \mathcal{Z}\} \tag{2.1.14}$$

and accordingly call a state $z \in \mathcal{S}_{glob}(t)$ *globally stable*. For convex energies, any locally stable point is also globally stable, that means the set of local stability and the set of global stability coincide. However, in the general case, these sets differ from each other and thus provide distinct notions of solutions in a natural way, see also [Ste09]. Indeed, we have:

Lemma 2.1.2. *If $\mathcal{I}(t,\cdot) \in C^1(\mathcal{Z};\mathbb{R})$, then $\mathcal{S}_{glob}(t) \subset \mathcal{S}_{loc}(t)$ holds. If additionally $\mathcal{I}(t,\cdot)$ is convex, then even $\mathcal{S}_{glob}(t) = \mathcal{S}_{loc}(t)$ holds.*

Proof. Let $z \in \mathcal{S}_{glob}(t)$, which means that

$$\mathcal{I}(t,z) \leq \mathcal{I}(t,v) + \mathcal{R}(v-z) \quad \forall v \in \mathcal{Z},$$

and furthermore let $\varepsilon > 0$ and $w \in \mathcal{Z}$ be arbitrary. Testing the above inequality with $v = z + \varepsilon w$, exploiting the 1-homogeneity of $\mathcal{R}$ and rearranging terms leads to $\mathcal{I}(t,z) - \mathcal{I}(t,z+\varepsilon w) \leq \varepsilon\mathcal{R}(w)$. Dividing by ε and taking the limit ε to zero, we then obtain $\langle -D_z\mathcal{I}(t,z), w\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq \mathcal{R}(w)$. Since w was arbitrary this gives $-D_z\mathcal{I}(t,z) \in \partial\mathcal{R}(0)$ by the characterization in Lemma 2.1.1, which is equivalent to saying $z \in \mathcal{S}_{loc}(t)$. If now $\mathcal{I}(t,\cdot)$ is additionally convex, then

$$\langle D_z\mathcal{I}(t,z), v - z\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq \mathcal{I}(t,v) - \mathcal{I}(t,z) \quad \forall v \in \mathcal{Z}$$

holds. Therefore, if $z \in \mathcal{S}_{loc}(t)$, then $-D_z\mathcal{I}(t,z) \in \partial\mathcal{R}(0)$ or equivalently

$$\langle -D_z\mathcal{I}(t,z), w\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq \mathcal{R}(w) \quad \forall w \in \mathcal{Z},$$

and testing with $w = v - z$ as well as inserting the inequality from above, we get $z \in \mathcal{S}_{glob}(t)$. $\square$

Overall, we visualize the different formulations in Figure 2.1.1.

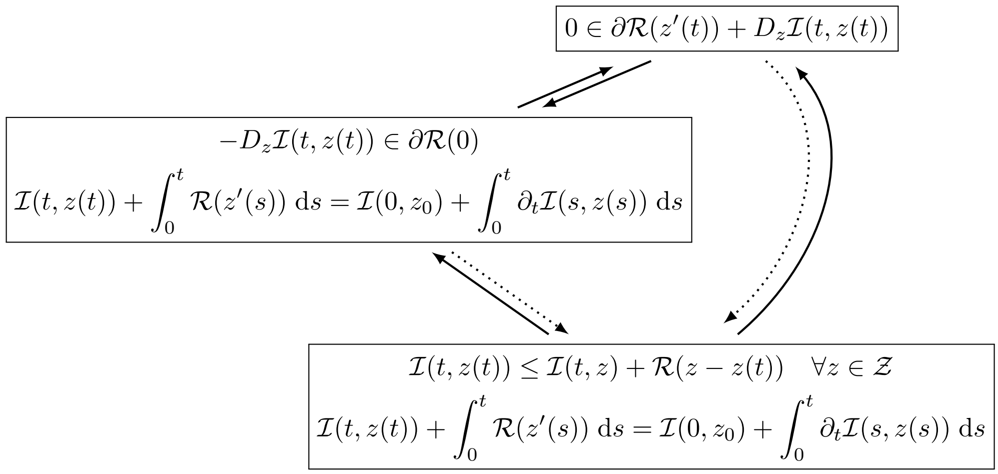


Figure 2.1.1: Overview of the essential reformulations for (RIS), that will become important in the definition of the solution concepts. Here, the dashed lines denote a direction that holds if $\mathcal{I}(t, \cdot)$ is convex.

2.2 Differential solutions

Historically, first existence results for solutions of (RIS) have been developed, e.g., in [Bré73] for quadratic and in [CV90] for general smooth and uniformly convex energies. Therein, the authors provide solutions in $W^{1,\infty}(0, T; \mathcal{Z})$ and $H^1(0, T; \mathcal{Z})$, respectively, satisfying the differential inclusion pointwise almost everywhere in $[0, T]$. However, in order to actually ensure that (RIS) is well-defined and the reformulations in Section 2.1 hold, we require that $z \in W^{1,1}(0, T; \mathcal{Z})$ as minimal regularity. In accordance with most of the literature, we will subsequently denote this type of solution by *differential solution*.

Definition 2.2.1. *We call $z : [0, T] \to \mathcal{Z}$ a **differential solution** if $z \in W^{1,1}(0, T; \mathcal{Z})$ with $z(0) = z_0$ and*

$$0 \in \partial\mathcal{R}(z'(t)) + D_z\mathcal{I}(t, z(t))$$

holds for almost all $t \in [0, T]$.

Note that this type of solution also provides solutions of the evolutionary variational inequality, that is,

$$\langle D_z\mathcal{I}(t, z(t)), w - z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \mathcal{R}(w) - \mathcal{R}(z'(t)) \geq 0 \quad \forall w \in \mathcal{V} \tag{2.2.1}$$

holds for almost all $t \in [0, T]$. Naturally, there arises the question under which assumptions we can guarantee the existence of a differential solution. As we have already seen in the Example 1.0.1, this is not the case for general nonconvex energy functionals. However, under strong convexity and smoothness assumptions, we have the following positive result.

Theorem 2.2.2. *Let $\mathcal{I}(t,\cdot) \in C^{2,1}_{loc}(\mathcal{Z};\mathbb{R})$ (see Definition 1.0.3) be κ-uniformly convex. Moreover, let there be a constant $c > 0$ such that for all $z_1, z_2 \in \mathcal{Z}$*

$$|\partial_t \mathcal{I}(t, z_1) - \partial_t \mathcal{I}(t, z_2)| \le c \|z_1 - z_2\|_{\mathcal{Z}} \tag{2.2.2}$$

holds for almost all $t \in [0,T]$. Finally, assume that one of the following properties is fulfilled:

$$\mathcal{R} \text{ is weakly continuous on } \mathcal{Z} \tag{2.2.3a}$$

or

$$\begin{aligned} &D_z\mathcal{I}(\cdot,\cdot) \text{ is (strong,weak)-weak continuous from } \mathcal{Z} \text{ to } \mathcal{Z}^*, \text{ i.e., } \forall\, t_k \to t,\ z_k \rightharpoonup z \text{ in } \mathcal{Z}: \\ &\qquad D_z\mathcal{I}(t_k, z_k) \rightharpoonup D_z\mathcal{I}(t,z) \quad \text{in } \mathcal{Z}^*. \end{aligned} \tag{2.2.3b}$$

Then there exists for every initial state $z_0 \in \mathcal{Z}$ with $0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(0, z_0)$ a unique differential solution $z \in W^{1,\infty}(0,T;\mathcal{Z})$, i.e., it holds

$$0 \in \partial\mathcal{R}(z'(t)) + D_z\mathcal{I}(t, z(t)) \quad \text{f.a.a. } t \in [0,T]. \tag{2.2.4}$$

Proof. The idea of this proof is to, first, apply Theorem 2.3.4, which guarantees the existence of an energetic solution (see Definition 2.3.1), and then show that this solution is, in fact, more regular, precisely Lipschitz continuous, by exploiting the uniform convexity of the energy functional. This will then allow us to deduce the existence of a differential solution. In the end, we provide a uniqueness result based on [MT04, Thm. 7.4]. At this point, the reader may excuse that the proof relies on a result, which is proven only at a later point.

1. *Existence of energetic solution*

In order to take advantage of Theorem 2.3.4, we need to verify condition (2.3.10) on the closedness of the global stability set. To this end, let $t_k \to t$ and $z_k \in \mathcal{S}_{glob}(t_k)$ with $z_k \rightharpoonup z$. We distinguish the two cases from (2.2.3a) and (2.2.3b).

We start with (2.2.3a). Since $z_k \in \mathcal{S}_{glob}(t_k)$ it holds $\mathcal{I}(t_k, z_k) \le \mathcal{I}(t_k, v) + \mathcal{R}(v - z_k)$ for all $v \in \mathcal{Z}$. Now let $v \in \mathcal{Z}$ be given. For the right-hand side of this inequality we exploit assumption (E1) and the weak continuity of $\mathcal{R}$ to obtain that $\lim_{k\to\infty} \mathcal{I}(t_k, v) + \mathcal{R}(v - z_k) = \mathcal{I}(t, v) + \mathcal{R}(v - z)$. By the lower semicontinuity of $\mathcal{I}$ from (1.0.2) we therefore have

$$\mathcal{I}(t,z) \le \liminf_{k\to\infty} \mathcal{I}(t_k, z_k) \le \liminf_{k\to\infty} \mathcal{I}(t_k, v) + \mathcal{R}(v - z_k) = \mathcal{I}(t,v) + \mathcal{R}(v - z). \tag{2.2.5}$$

Since this holds for every $v \in \mathcal{Z}$ we find $z \in \mathcal{S}_{glob}(t)$.

Now, assume that condition (2.2.3b) is satisfied. We then take advantage of Lemma 2.1.2, which gives $\mathcal{S}_{glob}(t_k) = \mathcal{S}_{loc}(t_k)$. Hence, $z_k \in \mathcal{S}_{glob}(t_k)$ is equivalent to $-D_z\mathcal{I}(t_k, z_k) \in \partial\mathcal{R}(0)$ and assumption (2.2.3b) as well as the weak closedness of $\partial\mathcal{R}(0)$ directly yields $-D_z\mathcal{I}(t,z) \in \partial\mathcal{R}(0)$. Thus $z \in \mathcal{S}_{loc}(t)$ and another application of Lemma 2.1.2 gives $z \in \mathcal{S}_{glob}(t)$.

In any case, we can apply Theorem 2.3.4 which yields the existence of an energetic solution.

2. Temporal regularity

The following part is based on [MR15, Thm. 7.4.4] and aims at proving that the energetic solution is in fact Lipschitz continuous. For this purpose, let $s, t \in [0,T]$ with $t > s$ be fixed. From the global stability, that is, $z(s) \in \mathcal{S}_{glob}(s)$, and the uniform convexity of $\mathcal{I}(s,\cdot)$, cf. inequality (1.0.4), we obtain

$$\begin{aligned}\mathcal{I}(s,z(s)) &\le \mathcal{I}(s,(1-\lambda)z(s)+\lambda z(t)) + \mathcal{R}((1-\lambda)z(s)+\lambda z(t)-z(s))\\ &\le (1-\lambda)\mathcal{I}(s,z(s)) + \lambda\mathcal{I}(s,z(t)) + \lambda\mathcal{R}(z(t)-z(s))\\ &\quad - \lambda(1-\lambda)\frac{\kappa}{2}\|z(t)-z(s)\|_{\mathcal{Z}}^2\end{aligned}$$

for all $\lambda \in [0,1]$. Subtracting $\mathcal{I}(s,z(s))$, dividing by λ and taking the limit λ to zero, we end up with

$$\frac{\kappa}{2}\|z(t)-z(s)\|_{\mathcal{Z}}^2 \le \mathcal{I}(s,z(t)) - \mathcal{I}(s,z(s)) + \mathcal{R}(z(t)-z(s)).$$

To proceed, we slightly rewrite the above inequality in the following form

$$\begin{aligned}\frac{\kappa}{2}\|z(t)-z(s)\|_{\mathcal{Z}}^2 &\le \mathcal{I}(s,z(t)) - \mathcal{I}(t,z(t)) + \mathcal{I}(t,z(t)) - \mathcal{I}(s,z(s)) + \mathcal{R}(z(t)-z(s))\\ &= -\int_s^t \partial_t\mathcal{I}(r,z(t))\,\mathrm{d}r + \mathcal{I}(t,z(t)) - \mathcal{I}(s,z(s)) + \mathcal{R}(z(t)-z(s)).\end{aligned}$$

Clearly, due to the additivity of the dissipation term $\mathrm{Diss}_{\mathcal{R}}$, the energy identity (E) also holds for $[s,t]$ instead of $[0,t]$. Thus, by inserting the shifted energy identity as well as the inequality $\mathcal{R}(z(t)-z(s)) \le \mathrm{Diss}_{\mathcal{R}}(z;[s,t])$, which is an easy consequence of the definition of $\mathrm{Diss}_{\mathcal{R}}$, see (2.3.1), we can further estimate

$$\begin{aligned}\frac{\kappa}{2}\|z(t)-z(s)\|_{\mathcal{Z}}^2 &\le -\int_s^t \partial_t\mathcal{I}(r,z(t))\,\mathrm{d}r + \mathcal{I}(t,z(t)) - \mathcal{I}(s,z(s)) + \mathcal{R}(z(t)-z(s))\\ &= \int_s^t \partial_t\mathcal{I}(r,z(r)) - \partial_t\mathcal{I}(r,z(t))\,\mathrm{d}r - \mathrm{Diss}_{\mathcal{R}}(z;[s,t]) + \mathcal{R}(z(t)-z(s))\\ &\le \int_s^t |\partial_t\mathcal{I}(r,z(r)) - \partial_t\mathcal{I}(r,z(t))|\,\mathrm{d}r.\end{aligned}$$

Incorporating the assumption in (2.2.2), we ultimately arrive at

$$\frac{\kappa}{2}\|z(t)-z(s)\|_{\mathcal{Z}}^2 \le \int_s^t c\,\|z(r)-z(t)\|_{\mathcal{Z}}\,\mathrm{d}r. \tag{2.2.6}$$

In order to obtain from this the Lipschitz continuity of z, we define

$$\delta(\tau) := \|z(t)-z(t-\tau)\|_{\mathcal{Z}}, \quad \tau \in [0,t].$$

First of all, we can conclude from (2.2.6) and the boundedness of energetic solutions in $\mathcal{Z}$ (see Definition 2.3.1), that z and therefore also δ are continuous. The inequality (2.2.6) thus gives

$$\delta(\tau)^2 \le C\int_{t-\tau}^t \|z(t)-z(r)\|_{\mathcal{Z}}\,\mathrm{d}r = C\int_0^\tau \|z(t)-z(t-r)\|_{\mathcal{Z}}\,\mathrm{d}r = C\int_0^\tau \delta(r)\,\mathrm{d}r.$$

An application of Lemma A.4.2 consequently yields $\delta(\tau) \leq \int_0^\tau C \, \mathrm{d}r = C\,\tau$. By setting $\tau := t - s \in [0,t]$, we ultimately find

$$\|z(t) - z(s)\|_{\mathcal{Z}} = \delta(t-s) \leq C\,|t-s|.$$

This proves the Lipschitz continuity of the solution z, which also gives $z \in W^{1,\infty}(0,T;\mathcal{Z})$.

3. Differential solution

We have already shown that z is an energetic solution and provides the extra regularity $z \in W^{1,\infty}(0,T;\mathcal{Z}) \hookrightarrow C(0,T;\mathcal{Z})$. Strictly speaking it would suffice to have $z \in W^{1,1}(0,T;\mathcal{Z})$, which also embeds into $C(0,T;\mathcal{Z})$, in order to proceed. In any case, we need to verify that z fulfills the differential inclusion (RIS) pointwise almost everywhere. For this, we first of all choose $\varepsilon > 0$ and $v \in \mathcal{Z}$ and test (S) with $z = z(t) + \varepsilon\, v$ to obtain

$$\mathcal{I}(t, z(t)) - \mathcal{I}(t, z(t) + \varepsilon\, v) \leq \mathcal{R}(\varepsilon\, v) = \varepsilon \mathcal{R}(v). \tag{2.2.7}$$

Dividing by ε and taking the limit $\varepsilon \to 0$ yields $\langle -D_z\mathcal{I}(t,z(t)), v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq \mathcal{R}(v)$. Since $v \in \mathcal{Z}$ was arbitrary, we find $-D_z\mathcal{I}(t,z(t)) \in \partial\mathcal{R}(0)$ for all $t \in [0,T]$ by the characterization in Lemma 2.1.1. In particular, we have $\mathcal{R}^*(-D_z\mathcal{I}(t,z(t))) = 0$ for all $t \in [0,T]$ due to (2.1.8). Combining this with Lemma 2.3.2, the energy identity (E) becomes

$$\mathcal{I}(T, z(T)) + \int_0^T \mathcal{R}(z'(s)) + \mathcal{R}^*(-D_z\mathcal{I}(s,z(s))\,\mathrm{d}s = \mathcal{I}(0, z_0) + \int_0^T \partial_t\mathcal{I}(s,z(s))\,\mathrm{d}s.$$

Applying the chain rule from Lemma A.2.6 and reordering terms, we finally get

$$\int_0^T \mathcal{R}(z'(s)) + \mathcal{R}^*(-D_z\mathcal{I}(s,z(s))) - \langle -D_z\mathcal{I}(s,z(s)), z'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\,\mathrm{d}s = 0.$$

Since by the Fenchel-Young inequality the integrand is nonnegative, we deduce

$$\mathcal{R}(z'(t)) + \mathcal{R}^*(-D_z\mathcal{I}(t,z(t))) = \langle -D_z\mathcal{I}(t,z(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}$$

and hence $0 \in \partial\mathcal{R}(z'(t)) + D_z\mathcal{I}(t,z(t))$ for almost all $t \in [0,T]$, by standard convex analysis results. Overall, this shows that z is a differential solution.

4. Uniqueness

Assume that there exists two differential solutions $z_1, z_2 \in W^{1,\infty}(0,T;\mathcal{Z})$. We define $\gamma(t) := \langle D_z\mathcal{I}(t,z_1(t)) - D_z\mathcal{I}(t,z_2(t)), z_1(t) - z_2(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}$. Writing the variational inequality (2.2.1) for z_1, inserting z_2 and vice versa and adding up the resulting inequalities, we obtain

$$0 \geq \langle D_z\mathcal{I}(t,z_1(t)) - D_z\mathcal{I}(t,z_2(t)), z_1'(t) - z_2'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}.$$

Exploiting the estimate from Section A.4 we therefore have

$$\begin{aligned}\gamma'(t) &\leq C\|z_1(t) - z_2(t)\|_{\mathcal{Z}}^2 + 2\langle D_z\mathcal{I}(t,z_1(t)) - D_z\mathcal{I}(t,z_2(t)), z_1'(t) - z_2'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\leq C\|z_1(t) - z_2(t)\|_{\mathcal{Z}}^2.\end{aligned} \tag{2.2.8}$$

The κ-uniform convexity of $\mathcal{I}$ implies that $\gamma(t) \geq \kappa\|z_1(t) - z_2(t)\|_{\mathcal{Z}}^2$, so that (2.2.8) gives the estimate $\gamma'(t) \leq C/\kappa\, \gamma(t)$ and we obtain the uniqueness result by applying the Gronwall inequality from Lemma A.4.1. □

Remark 2.2.3. It is easy to see that one may replace assumptions (E3) and (E4) by the following slightly weakened time regularity and still obtain a differential solution (see also assumption (E2), condition (C2) and Theorem 2.1.6 in [MR15]):

$$\begin{gathered}\exists N_{\mathcal{I}} \subset [0,T] \text{ with } \mathcal{L}^1(N_{\mathcal{I}}) = 0 \text{ so that } \forall t \in [0,T] \setminus N_{\mathcal{I}} \text{ it holds:} \\ \partial_t \mathcal{I}(t,z) \text{ exists and satisfies (E3) for all } z \in \mathcal{Z},\end{gathered} \tag{2.2.9a}$$

$$\forall t \in [0,T] \setminus N_{\mathcal{I}} \text{ and for all sequences } z_k \rightharpoonup z \text{ it holds:} \quad \partial_t \mathcal{I}(t,z_k) \to \partial_t \mathcal{I}(t,z). \tag{2.2.9b}$$

Remark 2.2.4. Due to the 1-homogeneity of $\mathcal{R}$, it holds $\partial\mathcal{R}(v) \subset \partial\mathcal{R}(0)$ for all $v \in \mathcal{V}$, see Lemma 2.1.1. Thus, since $W^{1,1}(0,T;\mathcal{Z}) \hookrightarrow C(0,T;\mathcal{Z})$ and $D_z\mathcal{I}$ is continuous by assumption (E1), a differential solution fulfills $0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(t,z(t))$ for all $t \in [0,T]$. In particular, we can reformulate $0 \in \partial\mathcal{R}(z'(t)) + D_z\mathcal{I}(t,z(t))$ as

$$\forall v \in \mathcal{Z}: \quad \mathcal{R}(v) \geq \langle -D_z\mathcal{I}(t,z(t)), v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \qquad \forall t \in [0,T], \tag{2.2.10a}$$

$$\mathcal{R}(z'(t)) = \langle -D_z\mathcal{I}(t,z(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \qquad \text{f.a.a. } t \in [0,T]. \tag{2.2.10b}$$

This type of solution will play a crucial role in our a priori error analysis in Section 3.3.

2.3 (Global) energetic solutions

The general notion of *global energetic solutions* has first been developed in the paper [MTL02] in 2002 but was used earlier in, e.g., [MTL98, MT99] for the specific case of hysteresis in elastic materials. This formulation is originally based on a mechanical extremum principle for phase transformations, that is, a transition between two states will occur as soon as it is thermodynamically possible, cf. [Lev97, p. 929] the *postulate of realizability*. However, one can also view this notion of solutions as a generalization of the form in (2.1.12b) combined with the global stability (2.1.13). In any case, it provides a mathematical fundament for a broad class of problems. The main advantage of this concept is that it does not require any differentiability of the involved quantities (except for a time derivative of the energy) and is thus amenable to handle cases without a linear structure of the underlying spaces. Nevertheless, we will formulate the actual definition in the context of our setting.

Definition 2.3.1. *We call $z : [0,T] \to \mathcal{Z}$ an **energetic solution** of* (RIS) *if $z \in L^\infty(0,T;\mathcal{Z}) \cap BV(0,T;X)$ and for all $t \in [0,T]$ the* global stability condition (S) *and* energy balance (E) *hold, i.e.,*

$$\mathcal{I}(t,z(t)) \leq \mathcal{I}(t,z) + \mathcal{R}(z - z(t)) \quad \forall z \in \mathcal{Z}, \tag{S}$$

$$\mathcal{I}(t,z(t)) + \mathrm{Diss}_{\mathcal{R}}(z;[0,t]) = \mathcal{I}(0,z(0)) + \int_0^t \partial_t\mathcal{I}(s,z(s))\,\mathrm{d}s. \tag{E}$$

Thereby, the dissipation $\mathrm{Diss}_{\mathcal{R}}(q;[0,t])$ *is defined as*

$$\mathrm{Diss}_{\mathcal{R}}(q;[0,t]) := \sup\Bigg\{ \sum_{k=1}^{n} \mathcal{R}(q(t_k)-q(t_{k-1})) \, : \, 0=t_0<t_1<\cdots<t_{n-1}<t_n=t,\, n\in\mathbb{N}\Bigg\}. \tag{2.3.1}$$

The condition (S) is called the condition of *global stability*, since it has to hold for all $z\in\mathcal{Z}$, i.e., $z(t)$ is the global minimum of the functional $\mathcal{I}(t,\cdot)+\mathcal{R}(\cdot-z(t))$ (see also the Definition of the set of global stability). In particular, this means that the release of energy $\mathcal{I}(t,z(t))-\mathcal{I}(t,z)$ for all possible states $z\in\mathcal{Z}$ is compensated by the dissipation. The second condition (E), which is the *energy balance*, is based on the energy equality (2.1.12b), where the term $\int_0^T \mathcal{R}(z'(s))\,\mathrm{d}s$ is replaced by the dissipation $\mathrm{Diss}_{\mathcal{R}}(z;[0,t])$ due to the lack of differentiability of z. However, if z is an element of $W^{1,1}(0,T;\mathcal{X})$ both terms coincide, which is subject of the next lemma.

Lemma 2.3.2. *Let* $\mathcal{R}:\mathcal{X}\to[0,\infty]$ *be given as in (R1)-(R3). If* $z\in W^{1,1}(0,T;\mathcal{X})$*, then* $\mathrm{Diss}_{\mathcal{R}}(z;[0,t])=\int_0^t\mathcal{R}(z'(r))\,\mathrm{d}r$ *holds for all* $t\in[0,T]$*.*

Proof. First, we note that due to Lebesgue's differentiation theorem (see, e.g., [Wac11, Thm. 3.1.40]), it holds

$$z'(t)=\lim_{h\searrow 0}\frac{z(t+h)-z(t)}{h}\quad \text{f.a.a. } t\in(0,T).$$

Now, let $t\in[0,T]$ be given and $0=t_0<t_1<\cdots<t_n=t$ be any partition of $[0,t]$. By Jensen's inequality from (A.3.3) we have for all $i=1,\dots,n$

$$\mathcal{R}(z(t_i)-z(t_{i-1}))=\mathcal{R}\left(\int_{t_{i-1}}^{t_i} z'(r)\,\mathrm{d}r\right)\le\int_{t_{i-1}}^{t_i}\mathcal{R}(z'(r))\,\mathrm{d}r.$$

Summing up this estimate yields $\sum_{i=1}^n\mathcal{R}(z(t_i)-z(t_{i-1}))\le\int_0^t\mathcal{R}(z'(r))\,\mathrm{d}r$, which in turn implies

$$\mathrm{Diss}_{\mathcal{R}}(z;[0,t])\le\int_0^t\mathcal{R}(z'(r))\,\mathrm{d}r\,. \tag{2.3.2}$$

For the opposite inequality, we observe that, by the 1-homogeneity of $\mathcal{R}$, it holds for any partition $\{t_i\}_{i=0}^n$ with $h:=\max\{(t_i-t_{i-1}):i=1,\dots,n\}$ and $0=t_0<t_1<\cdots<t_n=t$ that

$$\begin{aligned}\mathrm{Diss}_{\mathcal{R}}(z;[0,t])&\ge\sum_{i=1}^n\mathcal{R}(z(t_i)-z(t_{i-1}))\\&=\sum_{i=1}^n(t_i-t_{i-1})\mathcal{R}\left(\frac{z(t_i)-z(t_{i-1})}{t_i-t_{i-1}}\right)\\&=\sum_{i=1}^n\int_{t_{i-1}}^{t_i}\mathcal{R}\left(\frac{z(t_i)-z(t_{i-1})}{t_i-t_{i-1}}\right)\mathrm{d}r.\end{aligned}\tag{2.3.3}$$

Let us define the approximation

$$\xi_h(t):=\tfrac{z(t_i)-z(t_{i-1})}{t_i-t_{i-1}},\quad t\in[t_{i-1},t_i),$$

so that (2.3.3) implies $\mathrm{Diss}_{\mathcal{R}}(z;[0,t]) \geq \int_0^t \mathcal{R}(\xi_h(r))\,\mathrm{d}r$ and, moreover, it holds

$$\xi_h(t) \to z'(t) \quad \text{in } \mathcal{X}$$

pointwise almost everywhere in $[0,T]$. Hence, by the lower semicontinuity of $\mathcal{R}$, we also have that $\liminf_{h\to 0} \mathcal{R}(\xi_h(t)) \geq \mathcal{R}(z'(t))$ for almost all $t \in [0,T]$. Since $\mathcal{R}$ is nonnegative, we can apply Fatou's lemma to obtain

$$\liminf_{h\searrow 0} \int_0^t \mathcal{R}(\xi_h(r))\,\mathrm{d}r \geq \int_0^t \mathcal{R}(z'(r))\,\mathrm{d}r$$

which proves the claim. Note that the measurability of $r \mapsto \mathcal{R}(z'(r))$ can be discussed analogous to the proof of Lemma A.3.5. □

Before we come to the main existence result, we want to draw our attention to an approximation scheme for energetic solutions. One natural approach is to use a time-discretization for (S) and rewrite it as a minimization problem. In this way, one arrives at the following *global incremental minimization problem*

$$z_k \in \arg\min\{\mathcal{I}(t_k, z) + \mathcal{R}(z - z_{k-1}) \,:\, z \in \mathcal{Z}\}. \tag{2.3.4}$$

The existence of minimizers of (2.3.4) can be easily shown using the assumptions on $\mathcal{I}$ and $\mathcal{R}$ and applying the direct method of the calculus of variations. Moreover, from the global optimality of z_k and (E1), we obtain

$$\mathcal{I}(t_k, z_k) + \mathcal{R}(z_k - z_{k-1}) \leq \mathcal{I}(t_k, z_{k-1}) = \mathcal{I}(t_{k-1}, z_{k-1}) + \int_{t_{k-1}}^{t_k} \partial_t \mathcal{I}(s, z_{k-1})\,\mathrm{d}s. \tag{2.3.5}$$

The assumptions on $\mathcal{I}$ and $\mathcal{R}$, additionally, allow us to obtain the following a priori estimate, whose proof will be kept brief at this point. A more detailed derivation can, for example, be found in [MR15, Thm. 2.1.5] or in the proofs of Lemma 3.2.5 and Lemma 3.2.6 under slightly different assumptions.

Lemma 2.3.3. *Let $\mathcal{I}$ and $\mathcal{R}$ comply with (E1)-(E4) and (R1)-(R3), respectively. Then the following a priori estimate*

$$\mathcal{I}(t_k, z_k) + \sum_{i=1}^{k} \mathcal{R}(z_i - z_{i-1}) \leq \exp\left(\int_0^{t_k} \mu(s)\,\mathrm{d}s\right)(\mathcal{I}(0, z_0) + \beta) \tag{2.3.6}$$

holds true for all $k \in \mathbb{N}$.

Proof. For the sake of brevity, we set $m(t) = \int_0^t \mu(s)\,\mathrm{d}s$ where μ denotes the L^1-function from assumption (E3). Inequality (2.3.5) in combination with the assumption (E3) now yields

$$\begin{aligned} &\mathcal{I}(t_k, z_k) + \mathcal{R}(z_k - z_{k-1}) \\ &\quad \leq \mathcal{I}(t_{k-1}, z_{k-1}) + (\mathcal{I}(t_{k-1}, z_{k-1}) + \beta)\big(\exp(m(t_k) - m(t_{k-1})) - 1\big) && (2.3.7) \\ &\quad = (\mathcal{I}(t_{k-1}, z_{k-1}) + \beta)\exp(m(t_k) - m(t_{k-1})) - \beta. && (2.3.8) \end{aligned}$$

The nonnegativity of $\mathcal{R}$ and a summation over k for (2.3.8) thus implies

$$\mathcal{I}(t_k, z_k) + \beta \leq (\mathcal{I}(0, z_0) + \beta) \exp(m(t_k)) \quad \forall k \in \mathbb{N}. \tag{2.3.9}$$

By rewriting (2.3.7) in the form

$$\mathcal{I}(t_i, z_i) + \mathcal{R}(z_i - z_{i-1}) - \mathcal{I}(t_{i-1}, z_{i-1}) \leq (\mathcal{I}(t_{i-1}, z_{i-1}) + \beta)\big(\exp(m(t_i) - m(t_{i-1})) - 1\big)$$

another summation gives

$$\begin{aligned}\mathcal{I}(t_k, z_k) + \sum_{i=1}^{k} \mathcal{R}(z_i - z_{i-1}) &\leq \mathcal{I}(0, z_0) + \sum_{i=1}^{k} (\mathcal{I}(t_{i-1}, z_{i-1}) + \beta)\big(\exp(m(t_i) - m(t_{i-1})) - 1\big) \\ &\leq \mathcal{I}(0, z_0) + \sum_{i=1}^{k} (\mathcal{I}(0, z_0) + \beta)\big(\exp(m(t_i)) - \exp(m(t_{i-1}))\big) \\ &\leq (\mathcal{I}(0, z_0) + \beta) \exp(m(t_k)),\end{aligned}$$

where we took advantage of (2.3.9). Since this holds for all $k \in \mathbb{N}$, we end up with (2.3.6). □

These a priori bounds will allow us to extract a converging subsequence and afterwards pass to the limit in a discrete energy balance in order to obtain energetic solutions. Therefore, this minimization scheme forms an essential element of the following existence result. Despite its theoretical interest, it can, however, also be effectively used to numerically approximate energetic solutions.

Theorem 2.3.4 (Existence of energetic solution)**.** *Let $\mathcal{I}$ fulfill the assumptions (E1)-(E4) and assume that*

$$\forall \{(t_k, z_k)_{k\in\mathbb{N}}\} \subset \mathbb{R} \times \mathcal{Z} \text{ with } z_k \in \mathcal{S}_{glob}(t_k) \text{ and } t_k \to t,\ z_k \rightharpoonup z \text{ in } \mathcal{Z} : \quad z \in \mathcal{S}_{glob}(t). \tag{2.3.10}$$

Moreover, let $\mathcal{R}$ comply with assumptions (R1)-(R3) and z_0 satisfy $z_0 \in \mathcal{S}_{glob}(0)$. Then there exists an energetic solution of (RIS)*.*

Sketch of proof. We will just briefly explain the main steps to prove this result in our setting. The fully, generalized version is given for example in [MR15, Thm. 2.1.6]. Note that the condition on the initial state, i.e., $z_0 \in \mathcal{S}_{glob}(0)$, is needed in order to guarantee that the global stability (S) is satisfied for all $t \in [0, T]$. For convenience, the proof is divided into five steps. The time interval $[0, T]$ is divided into N subintervals $[t_{k-1}, t_k)$, where t_k describes the discrete time points in (2.3.4) and we set $\tau = \max\{(t_k - t_{k-1}) : k \in \{1, ..., N\}\}$ as the fineness of the approximation. Furthermore, we define the (right-continuous) piecewise constant interpolant $\underline{z}_\tau(t) := z_{k-1}$ for $t \in [t_{k-1}, t_k)$, where z_{k-1} denotes the corresponding solution of the minimization problem (2.3.4) at time t_{k-1}.

1. A priori estimates

Taking into account the definition of $\underline{z}_\tau(t)$ as well as assumption (E2), we easily obtain from

Lemma 2.3.3 the following a priori estimates

$$\mathcal{I}(t, \underline{z}_\tau(t)) + \mathrm{Diss}_{\mathcal{R}}(\underline{z}_\tau; [0,T]) \leq C \quad \text{and} \quad \|\underline{z}_\tau(t)\|_{\mathcal{Z}} \leq C \quad \forall t \in [0,T]. \tag{2.3.11}$$

Moreover, the (local) discrete energy identity (2.3.5) can be written as

$$\mathcal{I}(t_k, \underline{z}_\tau(t_k)) + \mathcal{R}(z_k - z_{k-1}) \leq \mathcal{I}(t_{k-1}, \underline{z}_\tau(t_{k-1})) + \int_{t_{k-1}}^{t_k} \partial_t \mathcal{I}(s, \underline{z}_\tau(t_{k-1}))\,\mathrm{d}s \tag{2.3.12}$$

for all $k = 1, ..., N$. Now, let $n \in \{1, ..., N-1\}$ and $t \in [t_n, t_{n+1})$ be arbitrary. By summing up the above inequality from $k = 1$ to n, we find the discrete energy inequality

$$\begin{aligned}
&\mathcal{I}(t, \underline{z}_\tau(t)) + \mathrm{Diss}_{\mathcal{R}}(\underline{z}_\tau; [0,t]) \\
&\quad = \mathcal{I}(t_n, \underline{z}_\tau(t_n)) + \mathcal{I}(t, \underline{z}_\tau(t)) - \mathcal{I}(t_n, \underline{z}_\tau(t_n)) + \sum_{k=1}^{n} \mathcal{R}(z_k - z_{k-1}) \\
&\quad = \mathcal{I}(t_n, z_n) + \sum_{k=1}^{n} \mathcal{R}(z_k - z_{k-1}) + \int_{t_n}^{t} \partial_t \mathcal{I}(s, z_n)\,\mathrm{d}s \\
&\quad \leq \mathcal{I}(0, z_0) + \int_0^{t_n} \partial_t \mathcal{I}(s, \underline{z}_\tau(s))\,\mathrm{d}s + \int_{t_n}^{t} \partial_t \mathcal{I}(s, z_n)\,\mathrm{d}s \\
&\quad = \mathcal{I}(0, z_0) + \int_0^{t} \partial_t \mathcal{I}(s, \underline{z}_\tau(s))\,\mathrm{d}s.
\end{aligned}$$

2. Selection of subsequences

By the Generalized Helly selection theorem (see Lemma A.2.11), we can extract a subsequence, w.l.o.g. denoted by the same symbol, so that

$$\forall t \in [0,T]: \quad \underline{z}_\tau(t) \rightharpoonup z(t) \quad \text{and} \quad \mathrm{Diss}_{\mathcal{R}}(\underline{z}_\tau; [0,t]) \to \delta(t)$$

and $\mathrm{Diss}_{\mathcal{R}}(z; [0,t]) \leq \delta(t)$. Furthermore, (E3), (E4) and the bound from (2.3.11), allow us to apply Lebesgue's dominated convergence theorem, so that

$$\partial_t \mathcal{I}(t, \underline{z}_\tau(t)) \to \partial_t \mathcal{I}(t, z(t)) \quad \text{in } L^1(0,T). \tag{2.3.13}$$

3. Global Stability of the limit function

From the global minimality of z_k and the triangle inequality for $\mathcal{R}$ from (1.0.3) we obtain

$$\begin{aligned}
\mathcal{I}(t_k, z_k) + \mathcal{R}(z_k - z_{k-1}) &\leq \mathcal{I}(t_k, z) + \mathcal{R}(z - z_{k-1}) \\
&\leq \mathcal{I}(t_k, z) + \mathcal{R}(z - z_k) + \mathcal{R}(z_k - z_{k-1}) \quad \forall z \in \mathcal{Z}.
\end{aligned}$$

Subtracting $\mathcal{R}(z_k - z_{k-1})$ on both sides gives $\mathcal{I}(t_k, z_k) \leq \mathcal{I}(t_k, z) + \mathcal{R}(z - z_k)$ for all $z \in \mathcal{Z}$ which in turn implies $z_k \in \mathcal{S}_{glob}(t_k)$. Hence, the global stability of $z(t)$, that is, $z(t) \in \mathcal{S}_{glob}(t)$, is an immediate consequence of the assumption (2.3.10) on the closedness of the stability set.

4. Upper energy estimate

The upper energy estimate follows from the weak lower semicontinuity of $\mathcal{I}(t, \cdot)$ and the convergence of $\partial_t \mathcal{I}(t, \underline{z}_\tau(t))$ in $L^1(0,T)$. We therefore have

$$\begin{aligned}
\mathcal{I}(t, z(t)) + \mathrm{Diss}_{\mathcal{R}}(z; [0,t]) &\leq \mathcal{I}(t, z(t)) + \delta(t) \\
&\leq \liminf_{\tau \to 0} \mathcal{I}(t, \underline{z}_\tau(t)) + \mathrm{Diss}_{\mathcal{R}}(\underline{z}_\tau; [0,t]) \\
&\leq \mathcal{I}(0, z_0) + \liminf_{\tau \to 0} \int_0^t \partial_t \mathcal{I}(s, \underline{z}_\tau(s)) \,\mathrm{d}s \\
&= \mathcal{I}(0, z_0) + \int_0^t \partial_t \mathcal{I}(s, z(s)) \,\mathrm{d}s.
\end{aligned}$$

5. Lower energy estimate
In order to gain a lower estimate for the energy, we fix $t \in [0,T]$ and take a partition of $[0,t]$, that is, $0 = t_0 < t_1 < \cdots < t_n = t$. For each $t_j \in \{t_k : k = 0, \ldots, n\}$, we have $z(t_j) \in \mathcal{S}_{glob}(t_j)$ and thus

$$\mathcal{I}(t_{j-1}, z(t_{j-1})) \leq \mathcal{I}(t_{j-1}, z(t_j)) + \mathcal{R}(z(t_j) - z(t_{j-1})).$$

Adding $\mathcal{I}(t_j, z(t_j))$ on both sides and reordering terms, we end up with

$$\mathcal{I}(t_j, z(t_j)) - \mathcal{I}(t_{j-1}, z(t_j)) \leq \mathcal{I}(t_j, z(t_j)) - \mathcal{I}(t_{j-1}, z(t_{j-1})) + \mathcal{R}(z(t_j) - z(t_{j-1})).$$

Summing up this inequality over $j = 1, \ldots, n$ and exploiting the definition of $\mathrm{Diss}_{\mathcal{R}}$, we arrive at

$$\begin{aligned}
&\mathcal{I}(t, z(t)) + \mathrm{Diss}_{\mathcal{R}}(z; [0,t]) - \mathcal{I}(0, z_0) \\
&\qquad \geq \mathcal{I}(t, z(t)) + \sum_{j=1}^n \mathcal{R}(z(t_j) - z(t_{j-1})) - \mathcal{I}(0, z_0) \\
&\qquad \geq \sum_{j=1}^n \mathcal{I}(t_j, z(t_j)) - \mathcal{I}(t_{j-1}, z(t_j)) \\
&\qquad = \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \partial_t \mathcal{I}(s, z(t_j)) \,\mathrm{d}s = \int_0^t \partial_t \mathcal{I}(s, \overline{z}^n(s)) \,\mathrm{d}s,
\end{aligned}$$

where $\overline{z}^n$ is the left-continuous approximation of z, i.e., $\overline{z}^n(t) = z(t_j)$ for $t \in (t_{j-1}, t_j]$. Now, in order to conclude that z is indeed an energetic solution, we need to show that $\int_0^t \partial_t \mathcal{I}(s, z(s)) \,\mathrm{d}s$ is bounded from above using integrals of the form $\int_0^t \partial_t \mathcal{I}(s, \overline{z}^n(s)) \,\mathrm{d}s$. Indeed, one can prove such an estimate by applying a generalized version of Lusin's theorem to z (see, e.g., [Wac11, Thm. 3.1.7]) and using suitable partitions of $[0,t]$. However, the details would go beyond the scope of this sketch of proof and we refer the interested reader to [MR15, Lem. 2.1.21] at this point. Having established this property, we end up with the lower energy estimate

$$\mathcal{I}(t, z(t)) + \mathrm{Diss}_{\mathcal{R}}(z; [0,t]) \geq \mathcal{I}(0, z_0) + \int_0^t \partial_t \mathcal{I}(s, z(s)) \,\mathrm{d}s.$$

Overall, this implies the existence of an energetic solution. □

Note that the closedness of $\mathcal{S}_{glob}$ in (2.3.10) is, for instance, fulfilled if $\mathcal{R}$ maps $\mathcal{X}$ into $[0, \infty)$ (rather than $[0, \infty]$) and is weakly continuous on $\mathcal{Z}$, cf. *Step 1* in the proof of Theorem 2.2.2. For several further conditions guaranteeing the closedness of $\mathcal{S}_{glob}$ we refer to [MR15, Sec. 2.1.5]. We now illustrate the above Definition 2.3.1 by the following example.

Example 2.3.5. We consider (RIS) with the energy functional $\mathcal{I}(t,z) = 2|z|^3 - 4z^2 + (\frac{4}{3})^3 - tz$ and the dissipation potential $\mathcal{R}(z) = |z|$. One can verify by direct calculations that

$$z(t) = \begin{cases} -\frac{1}{3}(2 + \frac{1}{\sqrt{2}}\sqrt{11-3t}), & t \in [0,1], \\ \frac{1}{3}(2 + \frac{1}{\sqrt{2}}\sqrt{5+3t}), & t \in (1,3], \end{cases}$$

is an energetic solution. Its graph is depicted in Figure 2.3.1 (left) and we can observe that it performs a jump at $t = 1$ from $z(1) = -\frac{4}{3}$ to $z(1^+) = \frac{4}{3}$. The corresponding energy landscape at the jump time is shown in Figure 2.3.1 (right). Here we see that the transition takes place, although the state $z(1)$ is still locally stable. Moreover, the solution jumps over a potential barrier that exists between the two states. Both of these phenomena are traced back to the global stability condition (S). In fact, this condition forces the system to change its state once the release in energy exceeds the amount of energy that is dissipated by the transition. This is a major disadvantage of the concept of energetic solutions, namely the fact that solutions will, in general, jump as soon as possible, ignoring potential barriers in between.

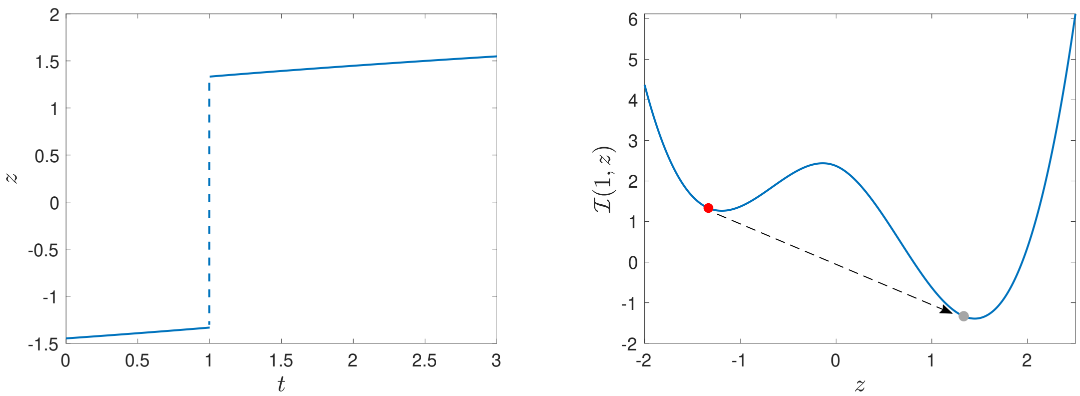


Figure 2.3.1: Left: Graph of the energetic solution z. Right: Energy landscape at the jump time $t = 1$ and the two state $z(1)$ (red) and $z(1^+)$ (gray)

At the end of this section, we briefly comment on how energetic solutions can be generalized to apply in models, which do not exhibit the structure in (RIS). As can be seen in the definition of energetic solutions, the only part where a derivative is needed, is the time component, which obviously exhibits a linear structure. One can thus also define this concept on a topological space $\mathfrak{Z}$. Instead of using the dissipation potential $\mathcal{R}$, one introduces a so-called dissipation distance $\mathcal{D} : \mathfrak{Z} \times \mathfrak{Z} \to [0,\infty]$. The actual Definition 2.3.1 however remains the same. One specific example of this is the crack growth in brittle materials, see [Mie11, Sec. 7.6] and [MR15, Sec. 4.2.4.1]. While the energy functional $\mathcal{I}(t, y, \Gamma)$ mainly consists of the elastic energy in the body plus some external loading, the dissipation distance reads

$$\mathcal{D}(\Gamma_0, \Gamma_1) = \begin{cases} \mathcal{H}^{d-1}(\Gamma_1 \setminus \Gamma_0), & \text{if } \Gamma_0 \subset \Gamma_1, \\ +\infty, & \text{else.} \end{cases}$$

Thereby $\Gamma \subset \overline{\Omega}$ describes the crack surface, $\mathcal{H}^{d-1}$ is the $(d-1)$-dimensional Hausdorff measure and the condition $\Gamma_0 \subset \Gamma_1$ represents the fact that the crack can only grow. This setting does certainly not exhibit a Banach space structure, but we can still consider $(y(t), \Gamma(t))$ as an energetic solution if it satisfies the stability condition (S) and the energy equality (E).

2.4 Parametrized solutions

The concept of parametrized solutions is a very recent notion of solutions. To motivate its definition, we want to derive the involved terms by performing (formally) the vanishing viscosity limit, i.e., $\varepsilon \searrow 0$ for

$$0 \in \partial\mathcal{R}(z'(t)) + \varepsilon \mathbb{V} z'(t) + D_z\mathcal{I}(t, z(t))\,. \tag{2.4.1}$$

In this context $\varepsilon\mathbb{V}z'$ denotes the so-called viscosity term and $\mathbb{V} \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$ is supposed to be a norm-preserving bijection. A detailed limit analysis would go beyond the scope of this work, which is why we only sketch the arguments leading to the definition of a parametrized solution. A rigorous convergence analysis can, for instance, be found in [MRS13, MRS16, MZ14, KRZ13]. The advantage in investigating (2.4.1) rather than (RIS) is that, while the original problem (RIS) does not necessarily exhibit a unique solution, (2.4.1) can be shown to be uniquely solvable providing solutions $z_\varepsilon \in W^{1,1}(0, T; \mathcal{V})$ under rather mild assumptions (see, e.g., [Col92, MRS13, MRS16, MZ14] in the case of a bounded dissipation and [KRZ13] in the case of a damage model). Beyond that, one can also interpret (2.4.1) as a regularization of the rate-independent system (RIS). Thus, it is natural to investigate (2.4.1) and perform the passage to the limit $\varepsilon \to 0$ in order to (hopefully) obtain solutions of the original problem (RIS). We illustrate this by means of the following example.

Example 2.4.1 (Vanishing viscosity limit (cf. [MR15, Ex. 1.8.3])).
In this example we consider $\mathcal{Z} = \mathcal{X} = \mathbb{R}$ and $\mathcal{I} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$\mathcal{I}(t, z) = \mathcal{E}(z) - \ell(t)\, z \;\text{ with }\; \mathcal{E}(z) = \begin{cases} \frac{1}{2}(z+4)^2, & z \le -2, \\ 4 - \frac{1}{2}z^2, & |z| < 2, \\ \frac{1}{2}(z-4)^2, & z \ge 2, \end{cases}$$

as well as $\mathcal{R}(z) = |z|$ and perform the vanishing viscosity limit. We additionaly set $z_0 = -2$ and $\ell(t) = t + 1$. For $\varepsilon > 0$ the solution of the viscous regularized problem (2.4.1) reads

$$z_\varepsilon(t) = \begin{cases} -2, & t \in [0, 2], \\ \varepsilon(\exp^{(t-2)/\varepsilon} - 1) - t, & t \in (2, t^*_\varepsilon], \\ (\varepsilon - 2 - t^*_\varepsilon) \exp^{-(t-t^*_\varepsilon)/\varepsilon} - \varepsilon + t + 4, & t \in (t^*_\varepsilon, 5], \end{cases}$$

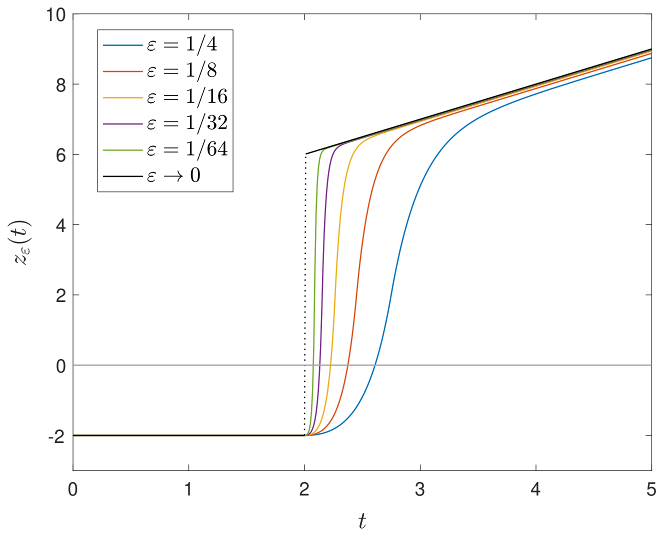


Figure 2.4.1: Viscous solutions for different viscosity parameters ε as well as the limit function z (black).

where $t_\varepsilon^* \approx 2$ is the solution of the equation $\varepsilon(\exp^{(t-2)/\varepsilon} - 1) - t = 2$. One can, for example, deduce this by concatenating solutions of ordinary differential equations. Taking ε to zero, we therefore obtain the limit

$$z(t) = \begin{cases} -2, & t \in [0, 2], \\ t + 4, & t \in (2, 5]. \end{cases}$$

However, one easily checks that

$$\mathcal{I}(t, z(t)) + \mathrm{Diss}_{\mathcal{R}}(z, [0, t]) - \mathcal{I}(0, z_0) - \int_0^t \partial_t \mathcal{I}(s, z(s))\,\mathrm{d}s = \begin{cases} 0, & t \in [0, 2], \\ -16, & t \in (2, 5], \end{cases}$$

so that the energy identity is consequently not fulfilled in this example. This follows from the fact that the release of energy during the jump is not compensated by the dissipation, that is, $\mathcal{I}(2, z(2)) - \mathcal{I}(2, z(2^+)) = 24 > 8 = \mathcal{R}(z(2^+) - z(2))$. As we will see, this transition is so fast that viscous effects must be taken into account. In fact, we will show in the subsequent analysis that this extra dissipation is a remnant of the viscous regularization term. Thus, during such a fast transition, which can be seen as a jump, the system will switch into a viscous behavior and dissipate the extra energy through this term.

Example 2.4.1 indicates that we need suitable limit equations in order to characterize such solutions, which can be obtained by the vanishing viscosity approach. This particularly refers to the resolution of (possibly) appearing jumps. Therefore, instead of performing the limit process directly, one introduces a suitable arc-length parametrization for z_ε and performs the passage to the limit afterwards. The main advantage here is that jumps do not shrink down to a single point in time. We rather obtain a whole jump curve in $\{t\} \times \mathcal{Z}$ that describes the transition between the two states. This idea was first applied in [MMMG94, MSGMM95, Bon96] for systems with dry friction and later on generalized in [EM06] and [MRS09, MZ14] for finite and infinite dimensional problems, respectively. Thus, as in [EM06], we first define the viscous regularized dissipation $\mathcal{R}_\varepsilon(v) := \mathcal{R}(v) + \frac{\varepsilon}{2}\|v\|_{\mathbb{V}}^2$. The corresponding Fenchel-conjugate is calculated exactly as in Lemma A.3.8, which gives $\mathcal{R}_\varepsilon^*(\xi) = \frac{1}{2\,\varepsilon}\,\overline{\text{dist}}_{\mathcal{V}^*}\{\xi, \partial\mathcal{R}(0)\}^2$. From the equivalence of the inclusion (2.4.1) to the energy identity (see Section 2.1, particularly (2.1.4)), we obtain here

$$\mathcal{I}(T, z_\varepsilon(T)) + \int_0^T \mathcal{R}_\varepsilon(z_\varepsilon'(s)) + \mathcal{R}_\varepsilon^*(-D_z\mathcal{I}(s, z_\varepsilon(s)))\,\mathrm{d}s = \mathcal{I}(0, z_\varepsilon(0)) + \int_0^T \partial_t\mathcal{I}(s, z_\varepsilon(s))\,\mathrm{d}s.$$

Now, for a solution z_ε of (2.4.1), we parameterize its graph by arc-length using the viscous norm $\|\cdot\|_{\mathbb{V}} = (\langle\cdot, \mathbb{V}\,\cdot\rangle_{\mathcal{V}})^{1/2}$. That is, we define $s_\varepsilon(t) := t + \int_0^t \|z_\varepsilon'(r)\|_{\mathbb{V}}\,\mathrm{d}r$, so that the function $s_\varepsilon : [0,T] \to [0, S_\varepsilon]$, where $S_\varepsilon := s_\varepsilon(T)$, is strictly monotone on $[0,T]$. Hence, it provides an inverse function $\hat{t}_\varepsilon : [0, S_\varepsilon] \to [0,T]$, by which we then introduce the rescaled function $\hat{z}_\varepsilon(s) := z_\varepsilon(\hat{t}_\varepsilon(s))$. By transforming the energy identity in terms of the "new" function $\hat{z}_\varepsilon$ and inserting the definition of $\mathcal{R}_\varepsilon$ and $\mathcal{R}_\varepsilon^*$, we obtain

$$\begin{aligned}
&\mathcal{I}(\hat{t}_\varepsilon(S_\varepsilon), \hat{z}_\varepsilon(S_\varepsilon)) \\
&\qquad + \int_0^{S_\varepsilon} \mathcal{R}(\hat{z}_\varepsilon'(s)) + \frac{\varepsilon}{2\,\hat{t}_\varepsilon'(s)}\|\hat{z}_\varepsilon'\|_{\mathbb{V}}^2 + \frac{\hat{t}_\varepsilon'(s)}{2\,\varepsilon}\,\overline{\text{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s)), \partial\mathcal{R}(0)\}^2\,\mathrm{d}s \\
&= \mathcal{I}(\hat{t}_\varepsilon(0), \hat{z}_\varepsilon(0)) + \int_0^{S_\varepsilon} \partial_t\mathcal{I}(\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s))\,\hat{t}_\varepsilon'(s)\,\mathrm{d}s.
\end{aligned} \tag{2.4.2}$$

In order to extract weakly convergent subsequences, we essentially need to bound the artificial end time S_ε uniformly in ε. The estimates on $\hat{z}_\varepsilon$ and $\hat{t}_\varepsilon$ then basically follow directly from the arc-length parametrization. However, this uniform bound does not follow readily from the analysis of the regularized problem (cf. [Mie11, Ex. 4.13]) but requires more sophisticated a priori estimates for solutions of (2.4.1), which itself need further restrictions on the energy (see Section 3.1). With this at hand, however, we can focus on the convergence of the transformed energy identity (2.4.2). Therefore, we reformulate (2.4.2) as

$$\begin{aligned}
\mathcal{I}(t_\varepsilon(S_\varepsilon), \hat{z}_\varepsilon(S_\varepsilon)) + \int_0^{S_\varepsilon} \mathcal{M}_\varepsilon(\hat{t}_\varepsilon'(s), \hat{z}_\varepsilon'(s), \overline{\text{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_\varepsilon(s), \hat{z}_\varepsilon(s)), \partial\mathcal{R}(0)\})\,\mathrm{d}s& \\
= \mathcal{I}(\hat{t}_\varepsilon(0), \hat{z}_\varepsilon(0)) + \int_0^{S_\varepsilon} \partial_t\mathcal{I}(\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s))\,\hat{t}_\varepsilon'(s)\,\mathrm{d}s&
\end{aligned}$$

with $\mathcal{M}_\varepsilon(\alpha, v, \mu) := \mathcal{R}(v) + \frac{\varepsilon}{2\,\alpha}\|v\|_{\mathbb{V}}^2 + \frac{\alpha}{2\,\varepsilon}\mu^2$. Following [KRZ13, MRS09], the term $\mathcal{M}_\varepsilon(\alpha, v, \mu)$ is Γ-convergent to

$$\mathcal{M}_0(\alpha, v, \mu) := \begin{cases} \mathcal{R}(v) + \mu\|v\|_{\mathbb{V}}, & \text{if } \alpha = 0, \\ \mathcal{R}(v) + I_0(\mu), & \text{if } \alpha > 0, \end{cases}$$

where I_0 denotes the indicator function of the singleton $\{0\}$. On the basis of this result and a weak lower semicontinuity result from [KRZ13, Lem. 6.1], [MRS12, Lem. 3.1], we define the notion of parametrized solutions as follows:

Definition 2.4.2. *Let an initial value $z_0 \in \mathcal{Z}$ be given. We call the tuple $(\hat{t}, \hat{z})$ a $\mathbb{V}$-**parametrized solution** of (RIS) if there exists an artificial end time $S \geq T$ such that the following conditions are satisfied:*

(i) Regularity:

$$\hat{t} \in W^{1,\infty}(0,S), \quad \hat{z} \in W^{1,\infty}(0,S;\mathcal{V}) \cap L^\infty(0,S;\mathcal{Z}). \tag{2.4.3}$$

(ii) Initial and end time condition:

$$\hat{t}(0) = 0, \quad \hat{z}(0) = z_0, \quad \hat{t}(S) = T. \tag{2.4.4}$$

(iii) Complementarity-like relations:

$$\hat{t}'(s) \geq 0, \quad \hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} \leq 1, \tag{2.4.5a}$$

$$\hat{t}'(s)\,\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)), \partial\mathcal{R}(0)\} = 0 \quad \text{f.a.a. } s \in (0,S), \tag{2.4.5b}$$

where $\overline{\operatorname{dist}}_{\mathcal{V}^*}\{\eta, \partial\mathcal{R}(0)\} = \inf\{\|\eta - w\|_{\mathbb{V}^{-1}} : w \in \partial\mathcal{R}(0)\}$ and $\|\eta\|^2_{\mathbb{V}^{-1}} = \langle \eta, \mathbb{V}^{-1}\eta\rangle_{\mathcal{V}^*,\mathcal{V}}$, see also Lemma A.3.8.

(iv) Energy Identity:

$$\begin{aligned} &\mathcal{I}(\hat{t}(s),\hat{z}(s)) + \int_0^s \mathcal{R}(\hat{z}'(r)) + \|\hat{z}'(r)\|_{\mathbb{V}}\,\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(r),\hat{z}(r)), \partial\mathcal{R}(0)\}\,\mathrm{d}r \\ &\qquad = \mathcal{I}(0,z_0) + \int_0^s \partial_t\mathcal{I}(\hat{t}(r),\hat{z}(r))\,\hat{t}'(r)\,\mathrm{d}r \qquad \forall\, s \in [0,S]. \end{aligned} \tag{2.4.6}$$

If, in addition to the second inequality in (2.4.5a), there exists a constant $\delta > 0$ such that $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} > \delta$ f.a.a. $s \in (0,S)$, then the solution is called nondegenerate $\mathbb{V}$-parametrized solution*, otherwise we call it* degenerate $\mathbb{V}$-parametrized solution*. In the special case where the second inequality in (2.4.5a) is fulfilled with equality, we call the solution* normalized.

Note that this definition of parametrized solutions hides its rate-independent structure in the condition $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} \leq 1$ of the arc-length parameterization. Indeed, a rescaling of the time does certainly not change the graph $\{(\hat{t}(s),\hat{z}(s)) : s \in [0,S]\} \subset [0,T] \times \mathcal{Z}$ of $(\hat{t},\hat{z})$. However, some more remarks on this definition are in order.

Remark 2.4.3. First of all, we note that the name $\mathbb{V}$-*parametrized solutions* is supposed to highlight the fact, that the parameterization is done with respect to the norm $\mathbb{V}$. Similarly, one may also consider $\mathcal{Z}$-parametrized solutions, as it is for example the case in [KRZ13]. Beyond this, there exists a further popular way of reparameterizing the viscous solutions of (2.4.1). This includes the so-called *vanishing viscosity contact potential*, which is defined as $\mathfrak{p}(v,w) := \mathcal{R}(v) + \|v\|_{\mathbb{V}}\,\overline{\operatorname{dist}}_{\mathcal{V}^*}\{w, \partial\mathcal{R}(0)\}$, see, e.g., [MRS12, Sec. 5]. One may interpret this as an "energy arc-length". The main advantage of this parameterization is the fact that a priori estimates come straight from the assumptions and even the normalization is preserved in the limit. However, in the context of (numerical) approximation, this construction is far less amenable and one loses the descriptive character of this solution concept. Moreover, since it merely holds $\mathcal{R}(v) \geq \|v\|_{\mathcal{X}}$ we no longer have control of the derivative of z_ε in the stronger $\mathcal{V}$-norm. Since $\mathcal{X}$ might not provide the Radon-Nikodým property (see, e.g., [DU77]), this energy arc-length enforces us to use spaces of absolutely continuous functions instead of the Sobolev space $W^{1,1}(0,T;\mathcal{X})$, see, e.g., [KT18]. Unfortunately, the distinction between these two parameterizations is not that strictly present in the literature so that both versions are sometimes referred to as merely *parametrized solutions*. What is more, these concepts are also denoted by *parametrized BV solutions*, which is due to the close connection with the so-called *BV solutions*, that we will take a look at in the next section. One, therefore, has to be careful when using these terms. Though, to ease the notation, we simply write parametrized solution for short in the rest of the thesis.

Remark 2.4.4. The regularity conditions in the above definition do not contain any information about the derivative of $\hat{z}$ in $\mathcal{Z}$, which might be useful with a view to the application of a chain rule, cf. Lemma 2.4.6. In fact, the regularity conditions here are chosen in such a way, that all terms contained are well-defined. Again, depending on the actual setting, particularly the choice of $\mathcal{R}$ and $\mathcal{I}$, there might exist slightly different requirements, see, e.g., [MRS16, Def. 4.2]. Nevertheless, a parametrized solution actually provides a certain weak continuity, that is,

$$\text{if } \tilde{s} \to s \quad \text{then} \quad \hat{z}(\tilde{s}) \rightharpoonup \hat{z}(s) \quad \text{in } \mathcal{Z}. \tag{2.4.7}$$

To see this, we first of all show that $\|\hat{z}(\tilde{s})\|_{\mathcal{Z}} \leq M$ for all $\tilde{s} \in [0,S]$ for some $M > 0$ independent of $\tilde{s}$. For this, we exploit the nonnegativity of the dissipation term in the energy identity (2.4.6), $\hat{t}' \leq 1$ and assumption (E3) to get

$$\begin{aligned}\mathcal{I}(\hat{t}(s),\hat{z}(s)) + c \leq \mathcal{I}(0,z_0) + \int_0^s \partial_t \mathcal{I}(\hat{t}(r),\hat{z}(r))\,\hat{t}'(r)\,\mathrm{d}r \\ \leq \mathcal{I}(0,z_0) + c + \int_0^s \mu(r)(\mathcal{I}(\hat{t}(r),\hat{z}(r)) + c)\,\mathrm{d}r.\end{aligned}$$

An application of the Gronwall lemma (Lemma A.4.1) and the lower estimate in (E2) thus implies

$$c\|\hat{z}(\tilde{s})\|_{\mathcal{Z}} \leq \mathcal{I}(\hat{t}(\tilde{s}),\hat{z}(\tilde{s})) + C \leq (\mathcal{I}(0,z_0) + c)\exp\left(\int_0^{\tilde{s}} \mu(r)\,\mathrm{d}r\right) \leq M.$$

Now, let $\tilde{s} \to s$. By the embedding $W^{1,\infty}(0,S;\mathcal{V}) \hookrightarrow C(0,S;\mathcal{V})$ we conclude that $\hat{z}(\tilde{s})$ converges to $\hat{z}(s)$ in $\mathcal{V}$. Since, moreover, $\hat{z}(\tilde{s})$ is uniformly bounded in $\mathcal{Z}$, we may extract a weakly converging subsequence with limit $z^* \in \mathcal{Z}$. Since weak and pointwise limit coincide, we conclude that $\hat{z}(s) = z^*$, which overall verifies (2.4.7).

Let us now put the parametrized solutions into some more context. We note that the additional term $\|\hat{z}'(r)\|_{\mathbb{V}}\,\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(r),\hat{z}(r)),\partial\mathcal{R}(0)\}$ which does not occur in the definition of energetic solutions (see Definition 2.3.1), can be interpreted as a viscous contribution, since it is a remnant of the vanishing viscosity approach. In particular, due to (2.4.5), this term is only present if $\hat{t}'(s) = 0$, which means that the physical time stands still and the system switches into a viscous behavior. This formulation serves as the basis for the convergence analysis of the *local iterated* *minimization scheme*, cf. Section 3.2 below. However, one could also directly pass to the limit in the viscous regularized system (2.4.1) and obtain again a subdifferential inclusion instead of the energy equality. The following Proposition shows that both approaches lead to equivalent formulations.

Proposition 2.4.5. *Let $(\hat{t}, \hat{z})$ with $\hat{z} \in W^{1,1}(0, S; \mathcal{Z})$ be given. Then the tupel $(\hat{t}, \hat{z})$ is a* nondegenerate parametrized solution *of* (RIS) *if and only if there exists a measurable function $\lambda : [0, S] \to [0, \infty)$, such that for almost all $s \in [0, S]$ it holds*

$$t(0) = 0, \quad z(0) = z_0, \quad t'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} \leq 1, \tag{2.4.8a}$$

$$0 \in \partial\mathcal{R}(\hat{z}'(s)) + \lambda(s)\mathbb{V}\hat{z}'(s) + D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \tag{2.4.8b}$$

$$\hat{t}'(s) \geq 0, \quad \lambda(s) \geq 0, \quad \lambda(s)\hat{t}'(s) = 0. \tag{2.4.8c}$$

Proof. Let $(\hat{t}, \hat{z})$ be a nondegenerate parametrized solution, i.e., there exists a $\delta > 0$ such that $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} \geq \delta$ almost everywhere in $[0, S]$. We then have

$$\begin{aligned} \mathcal{I}(\hat{t}(s), \hat{z}(s)) + \int_0^s \mathcal{R}(\hat{z}'(r)) + \|\hat{z}'(r)\|_{\mathbb{V}} \,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(r), \hat{z}(r)), \partial\mathcal{R}(0)\}\, \mathrm{d}r \\ = \mathcal{I}(0, z_0) + \int_0^s \partial_t\mathcal{I}(\hat{t}(r), \hat{z}(r))\, \hat{t}'(r)\, \mathrm{d}r \qquad \forall\, s \in [0, S]. \end{aligned} \tag{2.4.9}$$

By applying the parametrized version of the chain rule from Lemma A.2.5, we get

$$\begin{aligned} \int_0^s \mathcal{R}(\hat{z}'(r)) + \|\hat{z}'(r)\|_{\mathbb{V}} \,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(r), \hat{z}(r)), \partial\mathcal{R}(0)\}\, \mathrm{d}r \\ = \int_0^s \langle -D_z\mathcal{I}(\hat{t}(r), \hat{z}(r)), \hat{z}'(r)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\, \mathrm{d}r \qquad \forall\, s \in [0, S]. \end{aligned}$$

In fact, this is the point where the additional regularity of $\hat{z}$ is needed. To proceed, we exploit Lemma A.1.7 to obtain

$$\begin{aligned} \mathcal{R}(\hat{z}'(s)) + \|\hat{z}'(s)\|_{\mathbb{V}} \,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} \\ = \langle -D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \hat{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \end{aligned} \tag{2.4.10}$$

for almost all $s \in [0, S]$. Now, let $s \in [0, S]$ be a point where (2.4.10), (2.4.5b) and $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} \geq \delta$ hold. If $\|\hat{z}'(s)\|_{\mathbb{V}} = 0$, then $\hat{t}'(s) > 0$ by the nondegeneracy, which in turn implies that $\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} = 0$ by the complementarity in (2.4.5b) and we can set $\lambda(s) = 0$ so that (2.4.8b) and (2.4.8c) are satisfied. If otherwise $\tau := \|\hat{z}'(s)\|_{\mathbb{V}} > 0$, we set

$$\mathcal{R}_\tau(v) := \mathcal{R}(v) + I_\tau(v), \quad I_\tau(v) = \begin{cases} 0, & \text{if } \|v\|_{\mathbb{V}} \leq \tau, \\ +\infty, & \text{else,} \end{cases}$$

and Corollary A.3.9 implies

$$\begin{aligned} \langle -D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \hat{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} &= \mathcal{R}(\hat{z}'(s)) + \|\hat{z}'(s)\|_{\mathbb{V}} \,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} \\ &= \mathcal{R}_\tau(\hat{z}'(s)) + \mathcal{R}_\tau^*(-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s))). \end{aligned} \tag{2.4.11}$$

From the Fenchel duality in Lemma A.3.3, we thus get

$$-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)) \in \partial\mathcal{R}_\tau(\hat{z}'(s)). \tag{2.4.12}$$

We may now use the sum rule for convex subdifferentials (see Theorem A.3.6), which is applicable since $0 \in \mathrm{dom}(\mathcal{R}) \cap \mathrm{dom}(I_\tau)$ and I_τ is continuous in 0. Consequently, Lemma A.3.11 implies the existence of a $\lambda(s) \in \mathbb{R}$ with $\lambda(s) \geq 0$ such that

$$-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)) \in \partial\mathcal{R}(\hat{z}'(s)) + \lambda(s)\mathbb{V}\hat{z}'(s). \tag{2.4.13}$$

Since this holds for almost all $s \in [0,S]$, we conclude that there exists a function $\lambda : [0,S] \to [0,\infty)$ with

$$0 \in \partial\mathcal{R}(\hat{z}'(s)) + \lambda(s)\mathbb{V}\hat{z}'(s) + D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)) \tag{2.4.14}$$

for almost all $s \in [0,S]$. Finally, we want to convince ourselves that λ is indeed measurable. For this, we again take a point where the inclusion (2.4.14) holds. By the characterization of $\partial\mathcal{R}$ from Lemma 2.1.1 we thus have $-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)) - \lambda(s)\mathbb{V}\hat{z}'(s) \in \partial\mathcal{R}(0)$ so that $\mathcal{R}^*(-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)) - \lambda(s)\mathbb{V}\hat{z}'(s)) = 0$ by Lemma 2.1.1 and consequently

$$\begin{aligned} &\langle -D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)) - \lambda(s)\mathbb{V}\hat{z}'(s), \hat{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} = \mathcal{R}(\hat{z}'(s)) \\ \iff\quad &\langle -D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)), \hat{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} = \mathcal{R}(\hat{z}'(s)) + \lambda(s)\|\hat{z}'(s)\|_{\mathbb{V}}^2. \end{aligned}$$

Comparing this with (2.4.10) we conclude

$$\lambda(s)\|\hat{z}'(s)\|_{\mathbb{V}}^2 = \|\hat{z}'(s)\|_{\mathbb{V}}\,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)),\partial\mathcal{R}(0)\}. \tag{2.4.15}$$

We therefore set

$$\lambda(s) = \begin{cases} \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)),\partial\mathcal{R}(0)\}/\|\hat{z}'(s)\|_{\mathbb{V}}, & \text{if } \|\hat{z}'(s)\|_{\mathbb{V}} > 0, \\ 0, & \text{else.} \end{cases} \tag{2.4.16}$$

Since $\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(r),\hat{z}(r)),\partial\mathcal{R}(0)\}$ and $\|\hat{z}'(s)\|_{\mathbb{V}}$ are measurable functions, this also transfers to λ. (Note that as indicated above, if $\|\hat{z}'(s)\|_{\mathbb{V}} = 0$, then $\hat{t}'(s) > 0$ by the nondegeneracy and therefore $\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(r),\hat{z}(r)),\partial\mathcal{R}(0)\} = 0$ due to the complementarity (2.4.5). This justifies the choice $\lambda(s) = 0$ whenever $\|\hat{z}'(s)\|_{\mathbb{V}} = 0$.) The opposite dircection is proven by reversing the steps from above. That is, from (2.4.14) we find (2.4.12) with $\tau = \|\hat{z}'(s)\|_{\mathbb{V}} > 0$ and therewith also (2.4.11). Then integrating and applying the chain rule from Lemma A.2.5 gives (2.4.9) and consequently $(\hat{t},\hat{z})$ is a parametrized solution. □

The assumption on the nondegeneracy in the aforementioned lemma is essential in order to obtain a function λ that allows us to reformulate the energy identity as the differential inclusion (2.4.8b). Indeed, one easily notices that whenever $-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)) \notin \partial\mathcal{R}(0)$ but $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} = 0$, then there cannot exist a $\lambda(s) \in \mathbb{R}$ such that the inclusion (2.4.8b) is fulfilled. Nevertheless, it is always possible to retransform any parametrized solution in a way, such that the transformed version is still a parametrized solution, but nondegenerate and even normalized. The key idea here is to cut out all intervals where $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} = 0$ and to scale the artificial time appropriately. However, since the normalization plays only a minor role in this thesis, the proof of this fact is postponed to the Appendix, see Lemma A.4.3. As indicated in Section 2.1 by the upper energy estimate (2.1.12b), it suffices to require that a parametrized solution $(\hat{t}, \hat{z})$ fulfills the energy identity (2.4.6) with "$\leq$" instead of "$=$". Indeed, the following lemma, whose statement concerns exactly this equivalence, is an essential component in order to show existence of parametrized solutions. We will accordingly also take advantage of this fact in the proof of Theorem 3.2.19 in Section 3.2.4.

Lemma 2.4.6. *Let $(\hat{t}, \hat{z})$ be a pair with $\hat{t} \in W^{1,\infty}(0,S)$ and $\hat{z} \in W^{1,\infty}(0,S;\mathcal{V}) \cap W^{1,1}(0,S;\mathcal{Z})$ satisfying (2.4.4) and (2.4.5). Then $(\hat{t}, \hat{z})$ is a parametrized solution if and only if the following energy inequality is fulfilled:*

$$\begin{aligned}\mathcal{I}(\hat{t}(s), \hat{z}(s)) + \int_0^s \mathcal{R}(\hat{z}'(r)) + \|\hat{z}'(r)\|_{\mathbb{V}} \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(r), \hat{z}(r)), \partial\mathcal{R}(0)\} \,\mathrm{d}r \\ \leq \mathcal{I}(0, z_0) + \int_0^s \partial_t\mathcal{I}(\hat{t}(r), \hat{z}(r))\, \hat{t}'(r) \,\mathrm{d}r \qquad \forall\, s \in [0,S].\end{aligned} \tag{2.4.17}$$

Proof. The proof of this lemma is based on [KRZ13, Lem. 6.6]. Since every parametrized solution satisfies the above inequality with equality, the first implication is trivial. Hence, let $(\hat{t}, \hat{z})$ be given as in the assumptions, in particular $\hat{z} \in W^{1,1}(0,S;\mathcal{Z})$. This allows us to apply the parametrized chain rule from Lemma A.2.5, which gives for almost all $s \in [0,S]$:

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{I}(\hat{t}(s), \hat{z}(s)) = \langle D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \hat{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \partial_t\mathcal{I}(\hat{t}(s), \hat{z}(s))\, \hat{t}'(s). \tag{2.4.18}$$

Since $\|\hat{z}'(s)\|_{\mathbb{V}} \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\}$ is integrable, it must be finite almost everywhere. Thus, assume first that $\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} < \infty$, then there exists $\xi(s) \in \partial\mathcal{R}(0)$ such that $\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} = \|-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)) - \xi(s)\|_{\mathbb{V}^{-1}}$, cf. Lemma A.3.8. Exploiting the characterization of $\partial\mathcal{R}(0)$ from (2.1.6), we may consequently estimate

$$\begin{aligned}-\frac{\mathrm{d}}{\mathrm{d}s}&\mathcal{I}(\hat{t}(s), \hat{z}(s)) + \partial_t\mathcal{I}(\hat{t}(s), \hat{z}(s))\, \hat{t}'(s) \\ &= \langle -D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)) - \xi(s), \hat{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \langle \xi(s), \hat{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\leq \|-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)) - \xi(s)\|_{\mathbb{V}^{-1}} \|\hat{z}'(s)\|_{\mathbb{V}} + \mathcal{R}(\hat{z}'(s)) \\ &= \|\hat{z}'(s)\|_{\mathbb{V}} \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} + \mathcal{R}(\hat{z}'(s)).\end{aligned}$$

Otherwise $\|\hat{z}'(s)\|_{\mathbb{V}} = 0$ must hold and the former estimate clearly remains valid since $\mathcal{R} \geq 0$. Integration with respect to time and inserting the energy inequality, we obtain

$$\begin{aligned}\mathcal{I}(0, z_0) - \mathcal{I}(\hat{t}(s), \hat{z}(s)) + \int_0^s \partial_t\mathcal{I}(\hat{t}(r), \hat{z}(r))\hat{t}'(r) \,\mathrm{d}r \\ \leq \int_0^s \|\hat{z}'(r)\|_{\mathbb{V}} \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(r), \hat{z}(r)), \partial\mathcal{R}(0)\} + \mathcal{R}(\hat{z}'(r)) \,\mathrm{d}r \\ \leq \mathcal{I}(0, z_0) + \int_0^s \partial_t\mathcal{I}(\hat{t}(r), \hat{z}(r))\, \hat{t}'(r) \,\mathrm{d}r - \mathcal{I}(\hat{t}(s), \hat{z}(s))\end{aligned}$$

for all $s \in [0, S]$. Hence, $(\hat{t}, \hat{z})$ also satisfies the energy identity in (2.4.6) and, consequently, is a parametrized solution. $\square$

Note that we need to have an additional regularity for $\hat{z}$, namely $W^{1,1}(0, S; \mathcal{Z})$ instead of $L^\infty(0, S; \mathcal{Z})$, in order to be able to apply the chain rule from Lemma A.2.5. This regularity for example holds for the limit of the approximate parametrized solution in Chapter 3, see the enhanced convergence in (3.2.65). Strictly speaking, it certainly suffices to verify (2.4.18) with "$\geq$" for the equivalence of energy identity and energy inequality to hold. Clearly, the result also remains valid if (2.4.18) holds true with $\langle \cdot, \cdot \rangle_{\mathcal{V}^*, \mathcal{V}}$ instead of $\langle \cdot, \cdot \rangle_{\mathcal{Z}^*, \mathcal{Z}}$. In this case, the additional regularity $\hat{z} \in W^{1,1}(0, S; \mathcal{Z})$ is no longer necessary. This approach is usual, for example, if $\mathcal{R}$ is bounded from above, i.e., $\mathcal{R}(\cdot) \leq C \|\cdot\|_{\mathcal{V}}$ since then $\partial\mathcal{R}(0)$ is a bounded subset of $\mathcal{V}$, see, e.g., [MRS16, Thm. 4.4], but has also been applied to some damage model, see [KRZ19, Lem. 2.16]. Nevertheless, the energy identity (2.4.6) itself can still be formulated without the extra regularity of $\hat{z}$, which is why we chose this definition of parametrized solutions.

Remark 2.4.7. The reformulation in Lemma 2.4.5 in terms of a subdifferential inclusion allows to identify three different regimes and ascribe them a physical meaning (see [MR15] and the Figure 2.4.2 below):

- **Sticking:**
  In this case, the potential forces are too small so that $\hat{z}'(s) = 0$ and $\hat{t}'(s) = 1$ and the state does not change.

- **Rate-independent slip:**
  Here, it holds $0 < \|\hat{z}'(s)\|_{\mathbb{V}} < 1$ and $0 < \hat{t}'(s) < 1$ so that the state indeed changes, but in such a manner that the dissipation is strong enough to compensate the driving forces.

- **Viscous-jump:**
  In this case, we have $\|\hat{z}'(s)\|_{\mathbb{V}} = 1$ and $\hat{t}'(s) = 0$ which means that the system may switch into a viscous behavior. Meanwhile, the physical time stands still ($\hat{t}'(s) = 0$) so that this viscous transition is seen as a jump.

To conclude this section, we take a further look at the Example 2.4.1 from the beginning of this section.

Example 2.4.8. The considerations from above suggest that the solution z from Example 2.4.1 should be seen as a parametrized or, to be more precise, BV solution (see Section 2.5, Definition 2.5.3). In fact, we can specify the corresponding parametrized solution, namely by

$$\hat{z}(s) = \begin{cases} -2, & s \in [0, 2], \\ s - 4, & s \in (2, 10], \\ (s+2)/2, & s \in (10, 16], \end{cases} \quad \text{and} \quad \hat{t}(s) = \begin{cases} s, & s \in [0, 2], \\ 2, & s \in (2, 10], \\ (s-6)/2, & s \in (10, 16], \end{cases}$$

and the factor

$$\lambda(s) = \begin{cases} 0, & s \in [0,2], \\ 2-s, & s \in (2,6], \\ s-10, & s \in (6,10], \\ 0, & s \in (10,16]. \end{cases}$$

We note that during the viscous jump, we obtain the additional viscous dissipation $\lambda(s)\|\hat{z}'(s)\|^2$ which is missing in Example 2.4.1.

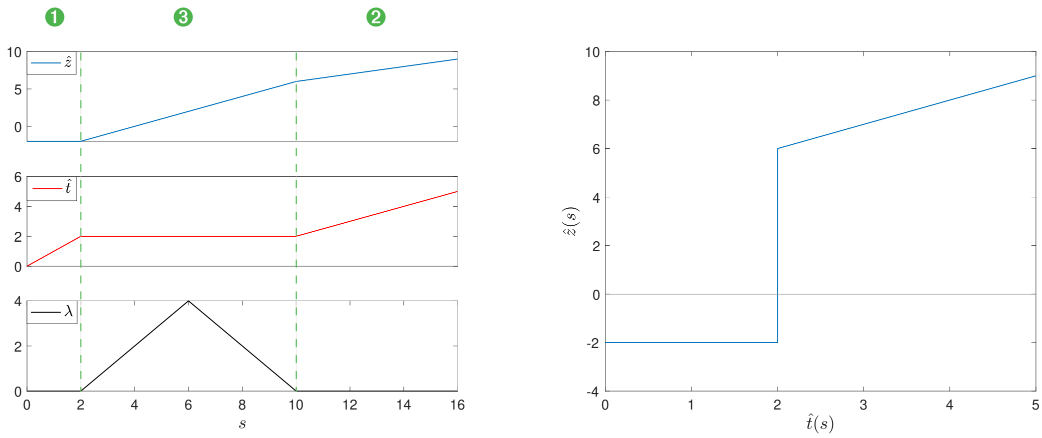


Figure 2.4.2: Left: Plot of the functions $\hat{z}$, $\hat{t}$ and λ (from top to bottom) depending on the artificial time s. The numbers indicate the different regimes *Sticking* ❶, *Rate-independent slip* ❷ and *Viscous jump* ❸. Right: Graph $\{(\hat{t}(s), \hat{z}(s)) : s \in [0,S]\} \subset [0,T] \times \mathbb{R}$ of the parametrized solution $(\hat{t}, \hat{z})$.

2.5 Further concepts in brief

There exist further notions of solutions for (RIS), for example *CD*, *local*, *semi-energetic* or *BV solutions* (see, e.g., [MR15, pp.131, p.229]). However, we restrict our presentation here to the concepts of *local* and *BV solutions*. We start with the former of them.

Definition 2.5.1. *We call $z : [0,T] \to \mathcal{Z}$ a **local solution** of* (RIS) *if*

$$0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(t,z) \quad \text{f.a.a. } t \in [0,T] \text{ and} \tag{2.5.1a}$$

$$\mathcal{I}(t,z(t)) + \mathrm{Diss}_{\mathcal{R}}(z;[0,t]) \le \mathcal{I}(0,z(0)) + \int_0^t \partial_t\mathcal{I}(s,z(s))\,\mathrm{d}s \quad \forall t \in [0,T]. \tag{2.5.1b}$$

The difference compared to the definition of energetic solutions is obviously the local stability in (2.5.1a), which replaces its global counterpart (S), and the energy inequality (2.5.1b). In fact, local solutions might exhibit a possible loss of energy, i.e., (2.5.1b) can be strict. In contrast, this is not possible for energetic or parametrized solutions, see Lemma 2.4.6 and *Step 5* in the proof of Theorem 2.3.4. This concept of solutions is a very broad one, that means, it contains all the other notions of solutions presented in this thesis. This is easy to see for the differential and energetic solutions but less so for parametrized solutions, since they are defined in the extended state space $[0,T] \times \mathcal{Z}$. Therefore, we define the following set of projections for a parametrized solution $(\hat{t}, \hat{z})$ with $\hat{t}(0) = 0$ and $\hat{t}(S) = T$ as well as $\hat{t}' \geq 0$ (see [MR15, p. 222], [Mie11, p. 89]):

$$\mathfrak{p}(\hat{t}, \hat{z}) := \{z : [0,T] \to \mathcal{Z} \,|\, \forall t \in [0,T]\, \exists s \in [0,S] : (t, z(t)) = (\hat{t}(s), \hat{z}(s))\}. \tag{2.5.2}$$

This set contains all functions z whose graph is a subset of the image of the curve $(\hat{t}, \hat{z})$. In particular, z is unique on parts where $\hat{t}$ is strictly monotone, whereas at the plateaus of $\hat{t}$, which are exactly the viscous jump parts, the function z may take any of the states in between the jump. Beyond this, for any given parametrized solution $(\hat{t}, \hat{z})$, the corresponding projections fulfill $z \in BV(0,T;\mathcal{V})$, since $\mathrm{Var}_{\mathcal{V}}(z, [0,T]) \leq \mathrm{Var}_{\mathcal{V}}(\hat{z}, [0,S]) < \infty$ by the Lipschitz continuity of $\hat{z}$. Hence, we may define the continuity and jump set as

$$C(z) := \{t \in [0,T] \,:\, z(t^-) = z(t) = z(t^+)\} \text{ and } J(z) := [0,T] \setminus C(z) \tag{2.5.3}$$

where $z(t^-) = \lim_{s \nearrow t} z(s)$ and $z(t^+) = \lim_{s \searrow t} z(s)$ is the left-hand and right-hand limit, respectively. Now, as already mentioned, there holds the following proposition.

Proposition 2.5.2. *Let (E1)-(E4) and (R1)-(R3) hold for $\mathcal{I}$ and $\mathcal{R}$, respectively. If $(\hat{t}, \hat{z})$ is a parametrized solution in the sense of Definition 2.4.2, then every $z \in \mathfrak{p}(\hat{t}, \hat{z})$ is a local solution.*

Proof. Since $z \in BV(0,T;\mathcal{V})$ the jump set $J(z)$ is countable and thus has measure zero. In order to prove (2.5.1a) we define $G := \{s \in [0,S] \,:\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} > 0\}$. From the complementarity in (2.4.5b), we deduce that $\hat{t}'(s) = 0$ holds almost everywhere in G and thus $\hat{t}(G)$ has measure zero. Hence, the set $[0,T] \setminus \hat{t}(G)$ is dense in $[0,T]$ and we can conclude from the closedness of $\partial\mathcal{R}(0)$ and the regularity of $\mathcal{I}$ in (E1), particularly the continuity of $D_z\mathcal{I}$, that $0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(t, z(t))$ holds for every $t \in [0,T]$ where z is continuous, i.e., for every $t \in C(z)$. Since $J(z)$ has measure zero, this already proves (2.5.1a). For the second condition (2.5.1b) we observe that on account of Lemma 2.3.2 and the monotonicity of $\hat{t}$ we have $\int_0^s \mathcal{R}(\hat{z}'(r))\,\mathrm{d}r = \mathrm{Diss}_{\mathcal{R}}(\hat{z}; [0,s]) \geq \mathrm{Diss}_{\mathcal{R}}(z, [0, \hat{t}(s)])$. The change of variable formula in Appendix A.2.10, moreover, yields that $\int_0^s \partial_t\mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma))\hat{t}'(\sigma)\,\mathrm{d}\sigma = \int_0^{\hat{t}(s)} \partial_t\mathcal{I}(r, z(r))\,\mathrm{d}r$. Thus, by the nonnegativity of the term $\|\hat{z}'(s)\|_{\mathbb{V}}\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(r), \hat{z}(r)), \partial\mathcal{R}(0)\}$ in (2.4.6), we conclude that (2.5.1b) is also valid which proves that z is, indeed, a local solution. □

As indicated above, the energy inequality can be strict, which stems from an additional dissipation during a jump. Hence, in order to further restrict the set of local solutions, we need to specify this additional term. This, however, comes directly from the definition of parametrized solutions. Indeed, let $t \in J(z)$, which induces that there exists $s_1^{(t)} < s_2^{(t)}$ depending on t, such that $\hat{t}^{-1}(\{t\}) = [s_1^{(t)}, s_2^{(t)}]$. Then we have for $v(s) := \hat{z}(s_1^{(t)} + s\,(s_2^{(t)} - s_1^{(t)}))$ that

$$\begin{aligned} &\int_{s_1^{(t)}}^{s_2^{(t)}} \mathcal{R}(\hat{z}'(s)) + \|\hat{z}'(s)\|_{\mathbb{V}}\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t, \hat{z}(s)), \partial\mathcal{R}(0)\}\,\mathrm{d}s \\ &\qquad = \int_0^1 \mathfrak{p}(\dot{v}(\sigma), -D_z\mathcal{I}(t, v(\sigma)))\,\mathrm{d}\sigma, \end{aligned}$$

with the *vanishing viscosity contact potential* $\mathfrak{p}(v,\xi) := \mathcal{R}(v) + \|v\|_{\mathcal{V}}\,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{\xi, \partial\mathcal{R}(0)\}$ (see [MR15, p. 224]), which we have already introduced in the context of parametrized solutions. Moreover, exploiting (2.5.11) in combination with the chain rule in Lemma A.2.5 it is easy to obtain the lower estimate

$$\int_0^1 \mathfrak{p}(\dot v(\sigma), -D_z\mathcal{I}(t,v(\sigma)))\,\mathrm{d}\sigma \geq \mathcal{I}(t,v(0)) - \mathcal{I}(t,v(1)) = \mathcal{I}(t,\hat z(s_1^{(t)})) - \mathcal{I}(t,\hat z(s_2^{(t)})),$$

see (2.5.12). Now, taking into account that $\hat t'(s) = 0$ for almost all $s \in [s_1^{(t)}, s_2^{(t)}]$, the energy identity (2.4.6) for parametrized solutions implies

$$\mathcal{I}(t,\hat z(s_1^{(t)})) - \mathcal{I}(t,\hat z(s_2^{(t)})) = \int_{s_1^{(t)}}^{s_2^{(t)}} \mathcal{R}(\hat z'(s)) + \|\hat z'(s)\|_{\mathbb{V}}\,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t,\hat z(s)), \partial\mathcal{R}(0)\}\,\mathrm{d}s \tag{2.5.4}$$

and we observe that this dissipation is minimal compared to all possible transitions between $\hat z(s_1^{(t)})$ and $\hat z(s_2^{(t)})$. Therefore, we define

$$\Delta_{\mathfrak{p}}(t,z_1,z_2) := \inf\Big\{ \int_0^1 \mathfrak{p}(\dot v(s), -D_z\mathcal{I}(t,v(s)))\mathrm{d}s \,: v \in W^{1,1}(0,1;\mathcal{Z}), \\ v(0) = z_1,\ v(1) = z_2 \Big\}, \tag{2.5.5}$$

which somehow describes an augmented "dissipative" distance between two points z_1 and z_2. Indeed, the term $\Delta_{\mathfrak{p}}(t,z_1,z_2)$ contains as well the dissipation due to $\mathcal{R}$ as the viscous effects that may arise during the jumps and consequently fulfills $\Delta_{\mathfrak{p}}(t,z_1,z_2) \geq \mathcal{R}(z_2 - z_1)$. Summing up this augmented "dissipative" distance at every jump and adding it to the remaining dissipation, we arrive at the following

$$\begin{aligned}\mathrm{Diss}_{\mathfrak{p}}(z;[0,t]) := \mathrm{Diss}_{\mathcal{R}}(z,[0,t]) &+ \Delta_{\mathfrak{p}}(t,z(0),z(0^+)) + \Delta_{\mathfrak{p}}(t,z(t^-),z(t)) \\ &+ \sum_{s\in J(z)\cap(0,t)} \Delta_{\mathfrak{p}}(s,z(s^-),z(s)) + \Delta_{\mathfrak{p}}(s,z(s),z(s^+)) \\ &- \mathcal{R}(z(0^+) - z(0)) - \mathcal{R}(z(t) - z(t^-)) \\ &- \sum_{s\in J(z)\cap(0,t)} \mathcal{R}(z(s) - z(s^-)) + \mathcal{R}(z(s^+) - z(s)).\end{aligned} \tag{2.5.6}$$

Note that the subtraction of $\mathcal{R}(z(s)-z(s^-))$ and $\mathcal{R}(z(s^+)-z(s))$ is necessary since these dissipative parts already contribute to the terms $\Delta_{\mathfrak{p}}(s,z(s^-),z(s))$ and $\Delta_{\mathfrak{p}}(s,z(s),z(s^+))$, respectively. In comparison with the local solutions, this augmented dissipation takes into account the missing viscous dissipation during a jump. All in all, we are led to the following definition.

Definition 2.5.3. *We call $z \in BV(0,T;\mathcal{V}) \cap L^\infty(0,T;\mathcal{Z})$ a $\mathbb{V}$-**parameterizable balanced viscosity (BV) solution** if*

$$\forall t \in C(z): \quad z \in \mathcal{S}_{loc}(t) \tag{2.5.7}$$

$$\text{and } \forall t \in [0,T]: \quad \mathcal{I}(t,z(t)) + \mathrm{Diss}_{\mathfrak{p}}(z;[0,t]) = \mathcal{I}(0,z(0)) + \int_0^t \partial_t\mathcal{I}(s,z(s))\,\mathrm{d}s, \tag{2.5.8}$$

where $\mathrm{Diss}_{\mathfrak{p}}(z;[0,t])$ *is defined as in* (2.5.6)*, and, moreover,*

$$\forall t \in J(z)\, \exists v_t \in W^{1,1}(0,1;\mathcal{Z}) : v_t(0) = z^-,\, v_t(1) = z^+,\, \exists r \in [0,1] : v_t(r) = z(t) \\ \text{and } \mathcal{I}(t,z^+) - \mathcal{I}(t,z^-) = \Delta_{\mathfrak{p}}(t,z^-,z^+), \tag{2.5.9}$$

$$\sum_{t\in J(z)} \int_0^1 \|v_t'(r)\|_{\mathbb{V}}\, \mathrm{d}r < \infty. \tag{2.5.10}$$

Remark 2.5.4. Similar to the parametrized solutions, cf. Remark 2.4.3, one has to distinguish between BV and $\mathbb{V}$-parameterizable BV solutions and the reader is referred to [MRS16, Def. 3.10] for the definition of general *BV solutions*. Indeed, this distinction is heavily related to the one for parametrized solutions. Loosely speaking, the additional conditions in (2.5.9) and (2.5.10) are required since the sole dissipation $\Delta_{\mathfrak{p}}$ does not necessarily contain information about the term $\int_0^1 \|v_t'(r)\|_{\mathbb{V}}\,\mathrm{d}r$ of the jump paths v_t. This is for example the case if $\mathcal{R}(\cdot) = \|\cdot\|_{\mathcal{X}}$ and $\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t,v_t(s)), \partial\mathcal{R}(0)\} = 0$ along the jump path, cf. Example 2.6.2. Hence, it is a priori not clear if there even exist jump paths connecting z^- and z^+ and whether or not their $\mathbb{V}$-length is additionally summable. Beyond this, the notion of *connectable BV solutions*, which only requires (2.5.9), has been introduced in [Mie11, Def. 4.21]. This concept merely requires $z \in BV(0,T;\mathcal{X})$ and neglects condition (2.5.10). However, in order to show that a connectable BV solution can be turned into a $\mathbb{V}$-parametrized solution, the missing regularity $z \in BV(0,T;\mathcal{V})$ and the summability from (2.5.10) thus become an additional assumption (cf. Proposition 4.24 and condition (44) in [Mie11]), see also Proposition 2.5.5.

Note that from the sole definition of $\Delta_{\mathfrak{p}}$, there must not exist an optimal path that realizes the infimum. This condition thus becomes a crucial ingredient in the Definition 2.5.3. Indeed, it allows us to obtain the following equivalence of $\mathbb{V}$-parameterizable BV and $\mathbb{V}$-parametrized solutions. To shorten the notation, we denote $G := \cup_{t\in J(z)} \hat{t}^{-1}(\{t\}) \subset [0,S]$.

Proposition 2.5.5 (Equivalence between parametrized and BV solutions)**.** *For* $\mathcal{I}$ *and* $\mathcal{R}$ *fulfilling (E1)-(E4) and (R1)-(R3), respectively, it holds:*

(i) If $(\hat{t},\hat{z})$ *is a* $\mathbb{V}$*-parametrized solution which satisfies* $\hat{z} \in W^{1,1}(a,b;\mathcal{Z})$ *for every connected component* $[a,b] \subseteq G$*, then every* $z \in \mathfrak{P}(\hat{t},\hat{z})$ *is a* $\mathbb{V}$*-parametrizable BV solution in the sense of Definition 2.5.3.*

(ii) If conversely $z \in BV(0,T;\mathcal{V}) \cap L^\infty(0,T;\mathcal{Z})$ *is a* $\mathbb{V}$*-parameterizable BV solution, then there exists a* $\mathbb{V}$*-parametrized solution* $(\hat{t},\hat{z})$ *such that* $z \in \mathfrak{P}(\hat{t},\hat{z})$*.*

Sketch of Proof. First of all, we observe that by the Corollary A.3.9 and the Fenchel-Young inequality in (A.3.2), it holds for all $v \in \mathcal{Z}$, $\xi \in \mathcal{Z}^*$ that

$$\mathfrak{p}(v,\xi) = \mathcal{R}(v) + \|v\|_{\mathbb{V}}\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{\xi, \partial\mathcal{R}(0)\} = \mathcal{R}_{\|v\|_{\mathbb{V}}}(v) + \mathcal{R}^*_{\|v\|_{\mathbb{V}}}(\xi) \geq \langle \xi, v\rangle_{\mathcal{Z}^*,\mathcal{Z}}. \tag{2.5.11}$$

For the first assertion *(i)* let $(\hat{t},\hat{z})$ be a parametrized solution that satisifies $z \in W^{1,1}(a,b;\mathcal{Z})$ for every connected component $[a,b]$ of G. Condition (2.5.7) can be derived exactly as for Proposition 2.5.2. Moreover, $z \in BV(0,T;\mathcal{V})$ is an easy consequence of the fact that $\mathrm{Var}_{\mathcal{V}}(z,[0,T]) \leq$

$\mathrm{Var}_{\mathcal{V}}(\hat{z},[0,S]) < \infty$ by the Lipschitz continuity of $\hat{z}$. Concerning the second condition (2.5.8) we proceed as follows: By (2.5.11) and the chain rule (A.2.5) we have

$$\begin{aligned}\int_0^1 \mathfrak{p}(v'(\sigma), -D_z\mathcal{I}(t,v(\sigma)))\,\mathrm{d}\sigma &\geq \int_0^1 \langle -D_z\mathcal{I}(t,v(\sigma)), v'(\sigma)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\,\mathrm{d}\sigma \\ &= \mathcal{I}(t,z_1) - \mathcal{I}(t,z_2)\end{aligned} \tag{2.5.12}$$

for every $v \in W^{1,1}(0,1;\mathcal{Z})$ with $v(0) = z_1$ and $v(1) = z_2$. Thus, $\Delta_{\mathfrak{p}}(t,z_1,z_2) \geq \mathcal{I}(t,z_1) - \mathcal{I}(t,z_2)$ holds true. Combining the energy identity (2.4.6) with (2.5.4), we see that this estimate holds with equality for the transition defined by the parametrized solution. Thus, the infimum in (2.5.5) is attained and coincides with the left-hand side in (2.5.4), which gives (2.5.9) and, due to $\hat{z} \in W^{1,\infty}(0,S;\mathcal{V})$, also (2.5.10). This, particularly, implies that

$$\mathrm{Diss}_{\mathfrak{p}}(z;[0,t]) = \int_0^{\hat{t}(s)} \mathcal{R}(\hat{z}'(\sigma)) + \|\hat{z}'(\sigma)\|_{\mathbb{V}}\,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s),\hat{z}(\sigma)), \partial\mathcal{R}(0)\}\,\mathrm{d}\sigma,$$

where $s \in [0,S]$ is chosen such that $(\hat{t}(s),\hat{z}(s)) = (t,z(t))$, which exists by construction of z. Therewith, condition (2.5.8) is now an easy consequence of (2.4.6).

Concerning the second assertion *(ii)*, we refer the reader to [MRS16, Prop. 4.9], [Mie11, Prop. 4.24]. However, we want to remark that the proof is mainly based on the fact that (again) there exists an optimal transition curve for the jump points, which can be concatenated to the part where the BV solution is continuous. The resulting graph is then reparameterized by arc-length such that we find a parametrized solution $(\hat{t},\hat{z})$ with $z \in \mathfrak{P}(\hat{t},\hat{z})$. More precisely, we define

$$s(t) := t + \mathrm{Var}_{\mathbb{V}}(z,[0,t]) + \sum_{s\in J(z)\cap[0,t]} \left\{ \int_0^1 \|v_s'(r)\|_{\mathbb{V}}\,\mathrm{d}r - \|z(s^+) - z(s)\|_{\mathbb{V}} - \|z(s) - z(s^-)\|_{\mathbb{V}} \right\},$$

so that $s : [0,T] \to [0,S]$ with $S = s(T)$ is strictly increasing (but not necessarily continuous). Note that the sum in the above definition is clearly positive and, moreover, the jump set of s coincides with the one of z, i.e., $J(s) = J(z) =: \{t_n\}_{n\in\mathbb{N}}$. Let us denote $I_n := (s(t_n^-), s(t_n^+))$. Since s is increasing, we may apply [Leo17, Thm. 1.8] to obtain a left-inverse function $\hat{t} : [0,S] \to [0,T]$, which is constant on $I := \cup_{n\in\mathbb{N}} I_n$. In addition, it is easy to see that $|I_n| \geq \int_0^1 \|v_{t_n}'(r)\|_{\mathbb{V}}\,\mathrm{d}r$. By [AGS08, Lem. 1.1.4(b)], there exists an increasing absolutely continuous map $r_n : (0,1) \to [0,L_n]$ with $L_n = \int_0^1 \|v_{t_n}'(r)\|_{\mathbb{V}}\,\mathrm{d}r$ so that $v_{t_n} = \hat{v}_{t_n} \circ r_n$ for some Lipschitz continuous function $\hat{v}_{t_n}$ with $\|\hat{v}_{t_n}'\|_{\mathbb{V}} = 1$ almost everywhere in $[0,L_n]$. Therewith, we define

$$\hat{z}(r) := \begin{cases} z(\hat{t}(r)), & \text{if } r \in [0,S]\setminus I, \\ \hat{v}_{t_n}(\sigma_n(r)), & \text{if } r \in I_n \subset I, \end{cases}$$

whereby $\sigma_n : \overline{I_n} \to [0,L_n]$ is the unique affine and strictly increasing function mapping $\overline{I_n}$ to $[0,L_n]$. By direct calculations, one verifies that $\hat{t}$ and $\hat{z}$ are Lipschitz continuous with $0 \leq \hat{t}'(s)$ and $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} \leq 1$. Since $\hat{t}$ is constant on the jump set of z, we additionally find (2.4.5b). The initial conditions as well as the regularity are now obvious and, finally, the energy identity (2.4.6) is a consequence of (2.5.8) □

We remark that the results in the above Proposition 2.5.5 also hold true in the case of (general) BV solutions and we refer the reader to [MRS16] for more details. Note that, here, the additional regularity, i.e., $z \in W^{1,1}(a, b; \mathcal{Z})$ for every connected component of G, is necessary since we do not exclude the possibility of an unbounded dissipation as it is the case in [MRS16]. In fact, the equivalence as given in Proposition 2.5.5 relies on a chain rule argument which is valid here only if z provides a derivative in $\mathcal{Z}$. In view of this, it is possible to modify the definition of $\mathbb{V}$-parameterizable BV solutions, i.e., Definition 2.5.3, if one can guarantee a chain rule for $\mathcal{I}$ to hold in a different space (e.g. $\mathcal{V}$ instead of $\mathcal{Z}$), cf. [KRZ19] particularly Remark 5.2 therein. The existence of these various slightly different notions, which, in general, depend on the actual setting of the rate-independent system, makes it difficult to bring all these into one single definition. Nevertheless, due to the above result, one should regard the notions of $\mathbb{V}$-parameterizable BV and $\mathbb{V}$-parametrized solutions as two naturally related concepts.

2.6 Relations between different concepts

As indicated in the previous section, there exist even further concepts of solutions for (RIS), for example *local*, *semi energetic* or *approximable solutions*, see [MR15] and the references therein. Nevertheless, we will focus on those concepts presented in Sections 2.2–2.4. In this last part of the chapter, we want to highlight the connections between these different notions of solutions. An overview hereof is given at the end of this section. We start with the simple case of a uniformly convex energy, see Definition 1.0.2. In this case, all notions provide the same solution.

Lemma 2.6.1. *Let $\mathcal{I}$ and $\mathcal{R}$ satisfy assumptions (E1)-(E4) and (R1)-(R3), respectively. Moreover, let $\mathcal{I}(t, \cdot) \in C^{2,1}_{loc}(\mathcal{Z}; \mathbb{R})$ (see Definition 1.0.3) be κ-uniformly convex, satisfy the Lipschitz condition in* (2.2.2) *and either of the two conditions* (2.2.3a) *or* (2.2.3b)*. Then there exists a unique differential solution z and, denoting by $(\hat{t}, \hat{z})$ a parametrized solution, the following holds true:*

- *i) Differential and energetic solution coincide.*
- *ii) Every $z \in \mathfrak{P}(\hat{t}, \hat{z})$ is a differential solution and, conversely, the arc-length parametrization of z is a parametrized solution.*

Proof. First of all, the existence and uniqueness of a differential solution is provided by Theorem 2.2.2. In addition, the fact that, under the given assumptions, every energetic solution is also a differential one, is proven exactly as in the *Steps 2* and *3* of the proof to Theorem 2.2.2. For the opposite inclusion, we note that due to Lemma 2.1.2 the global and local stability sets coincide, which gives condition (S), see also Remark 2.2.4. Moreover, by exploiting the chain rule from Lemma A.2.5 we deduce from the characterization $\mathcal{R}(z'(t)) = \langle -D_z\mathcal{I}(t, z(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}$ in (2.2.10a) that

$$\mathcal{I}(t, z(t)) + \int_0^t \mathcal{R}(z'(s))\,\mathrm{d}s = \mathcal{I}(0, z_0) + \int_0^t \partial_t\mathcal{I}(s, z(s))\,\mathrm{d}s.$$

Thus, by Lemma 2.3.2, we ultimately have that z is also an energetic solution. Regarding (ii), we combine Lemma 2.6.4 and (i) to obtain that every $z \in \mathfrak{P}(\hat{t}, \hat{z})$ is a differential solution. For the opposite direction, we refer to Lemma 2.6.6. □

We proceed with investigating an energy functional which is merely convex. In this case, it is no longer guaranteed that there exists a differential solution. Even further, the existence of such a solution does not imply that it is unique, as the following Example shows.

Example 2.6.2 (Existence of two differential solutions). We set $\mathcal{Z} = \mathcal{X} = \mathbb{R}$ as well as

$$\mathcal{I} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad \mathcal{I}(t, z) = \mathcal{E}(z) - \ell(t)\, z \;\text{ with }\; \mathcal{E}(z) = \begin{cases} \frac{1}{2}(z-1)^2, & z \geq 1, \\ 0, & |z| < 1, \\ \frac{1}{2}(z+1)^2, & z \leq -1. \end{cases}$$

Furthermore, we take $z_0 = -2$ and

$$\ell(t) = \begin{cases} t, & t \in [0,1], \\ 1, & t \in (1,2), \\ t-1, & t \in [2,3]. \end{cases}$$

For this problem, one can easily verify by direct calculations, that

$$z_1(t) = \begin{cases} t-2, & t \in [0,1], \\ 2t-3, & t \in (1,2], \\ t-1, & t \in (2,3], \end{cases} \quad \text{and} \quad z_2(t) = \begin{cases} t-2, & t \in [0,1], \\ 4t-5, & t \in (1, \frac{3}{2}], \\ 1, & t \in (\frac{3}{2}, 2], \\ t-1, & t \in (2,3], \end{cases}$$

are two distinct differential solutions.

Nevertheless, for convex energies, we still obtain from each differential solution an energetic one. This is proven exactly as in Lemma 2.6.1. The same also holds for parametrized solutions, that is, every differential solution provides a parametrized one, see Lemma 2.6.6. The opposite direction, however, does not have to hold. In fact, even though we will show in the proof of Lemma 2.6.4 that $\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\}$ in (2.4.8b) vanishes for convex energies, this does not guarantee that the solution has no jumps, as can be seen in the following example.

Example 2.6.3 (Convex energies can have jumping solutions). We take up again the setting from Example 2.6.2. Additionally to the two differential solutions given there, we find the energetic solution

$$z_3(t) = \begin{cases} t-2, & t \in [0,1), \\ 1, & t \in [1,2), \\ t-1, & t \in [2,3], \end{cases}$$

which performs a jump at time $t = 1$ from $z(1^-) = -1$ to $z(1^+) = 1$. In particular, it does not suffice to require $\mathcal{I}$ to be merely convex in order to obtain exclusively continuous solutions.

For convex energies, however, parametrized solutions can still be transformed into energetic solutions by the same projection that is used to derive BV solutions, see (2.5.2).

Lemma 2.6.4. *Let $\mathcal{I}$ and $\mathcal{R}$ comply with (E1)-(E4) and (R1)-(R3), respectively. Moreover, let $\mathcal{I}(t,\cdot)$ be convex for all $t \in [0,T]$ and let $(\hat{t},\hat{z})$ be a parametrized solution to* (RIS) *with $\hat{z} \in W^{1,1}(0,S;\mathcal{Z})$. Then every $z \in \mathfrak{P}(\hat{t},\hat{z})$ is an energetic solution.*

Proof. Let $(\hat{t},\hat{z})$ be a parametrized solution with $\hat{z} \in W^{1,1}(0,S;\mathcal{Z})$. W.l.o.g. we may assume that $(\hat{t},\hat{z})$ is nondegenerate (otherwise we rescale $(\hat{t},\hat{z})$ according to Lemma A.4.3 and obtain a normalized solution $(\tilde{t},\tilde{z})$, whose projection remains unaltered, i.e., $z \in \mathfrak{P}(\tilde{t},\tilde{z}) = \mathfrak{P}(\hat{t},\hat{z})$). Moreover, for the sake of brevity, we define $m(s) := \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)), \partial\mathcal{R}(0)\}$. Now, due to the embedding $W^{1,1}(0,S;\mathcal{Z}) \hookrightarrow C(0,S;\mathcal{Z})$, the function $\hat{z}$ is continuous and therewith, by the continuity of $D_z\mathcal{I}(\cdot,\cdot)$, also $D_z\mathcal{I}(\hat{t}(\cdot),\hat{z}(\cdot))$. The closedness of $\partial\mathcal{R}(0)$ thus implies that the set $\mathcal{S} := \{s \in [0,S] \,:\, -D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)) \in \partial\mathcal{R}(0)\}$ is closed and consequently $\mathcal{T} := [0,S]\setminus\mathcal{S}$ is relatively open. Now let G be any connected component of $\mathcal{T}$. By the complementarity in (2.4.5b) there must hold $\hat{t}'(s) = 0$ almost everywhere in G. Given any $\tilde{s} \in G$, we set $t := \hat{t}(\tilde{s})$ and take

$$s_1 := \inf\{s \in [0,S] \,:\, \hat{t}(s) = t\} \quad \text{as well as} \quad s_2 := \sup\{s \in [0,S] \,:\, \hat{t}(s) = t\}.$$

We then have $s_1 < s_2$ and it holds $\hat{t}(s) \equiv t$ for all $s \in [s_1,s_2]$ as well as $-D_z\mathcal{I}(\hat{t}(s_1),\hat{z}(s_1)) \in \partial\mathcal{R}(0)$. Indeed, it must hold $s_1 \in \mathcal{S}$ since otherwise $s_1 \in \mathcal{T}$ and as $\mathcal{T}$ is relatively open, there exists $s < s_1$ with $s \in G \subset \mathcal{T}$. Thus, $\hat{t}(s) = t$ again due to the complementarity (2.4.5b), which holds on G, and consequently s_1 cannot be infimal. After these preparatory steps, we start with Jensen's inequality, cf. (A.3.3), to obtain

$$\mathcal{R}(\hat{z}(s_2) - \hat{z}(s_1)) = \mathcal{R}\left(\int_{s_1}^{s_2} \hat{z}'(r)\,\mathrm{d}r\right) \leq \int_{s_1}^{s_2} \mathcal{R}(\hat{z}'(r))\,\mathrm{d}r. \tag{2.6.1}$$

Exploiting the energy identity (2.4.6) and taking into account that $\hat{t}'(s) = 0$ almost everywhere in $[s_1,s_2]$, we further estimate

$$\begin{aligned}\int_{s_1}^{s_2} \mathcal{R}(\hat{z}'(r))\,\mathrm{d}r &= \mathcal{I}(\hat{t}(s_1),\hat{z}(s_1)) - \mathcal{I}(\hat{t}(s_2),\hat{z}(s_2)) - \int_{s_1}^{s_2} \|\hat{z}'(r)\|_{\mathbb{V}}\, m(r)\,\mathrm{d}r \\ &\leq \mathcal{I}(\hat{t}(s_1),\hat{z}(s_1)) - \mathcal{I}(\hat{t}(s_2),\hat{z}(s_2)),\end{aligned} \tag{2.6.2}$$

where we also used the nonnegativity of m. Since, by construction, $\hat{t}(s_1) = \hat{t}(s_2) = t$, the convexity of $\mathcal{I}(t,\cdot)$ yields

$$\begin{aligned}\mathcal{I}(\hat{t}(s_1),\hat{z}(s_1)) - \mathcal{I}(\hat{t}(s_2),\hat{z}(s_2)) &= \mathcal{I}(t,\hat{z}(s_1)) - \mathcal{I}(t,\hat{z}(s_2)) \\ &\leq \langle -D_z\mathcal{I}(t,\hat{z}(s_1)), \hat{z}(s_2) - \hat{z}(s_1)\rangle_{\mathcal{Z}^*,\mathcal{Z}}.\end{aligned} \tag{2.6.3}$$

As indicated above, we have $\partial\mathcal{R}(0) \ni -D_z\mathcal{I}(\hat{t}(s_1),\hat{z}(s_1)) = -D_z\mathcal{I}(t,\hat{z}(s_1))$. Thus, following the inequalities (2.6.1), (2.6.2) and (2.6.3) and exploiting the characterization of $\partial\mathcal{R}(0)$ from (2.1.6) with $v = \hat{z}(s_2) - \hat{z}(s_1)$, we find

$$\begin{aligned}\mathcal{R}(\hat{z}(s_2) - \hat{z}(s_1)) &\leq \mathcal{I}(\hat{t}(s_1),\hat{z}(s_1)) - \mathcal{I}(\hat{t}(s_2),\hat{z}(s_2)) - \int_{s_1}^{s_2} \|\hat{z}'(r)\|_{\mathbb{V}}\, m(r)\,\mathrm{d}r \\ &\leq \langle -D_z\mathcal{I}(t,\hat{z}(s_1)), \hat{z}(s_2) - \hat{z}(s_1)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq \mathcal{R}(\hat{z}(s_2) - \hat{z}(s_1)).\end{aligned}$$

Hence, all these inequalities hold in fact with equality which implies that $\|\hat{z}'(s)\| m(s) = 0$ almost everywhere in $[s_1, s_2]$. Exploiting the nondegeneracy of $(\hat{t}, \hat{z})$, we conclude that already $m(s) = 0$ for almost all $s \in [s_1, s_2]$. Again by the continuity of $-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s))$ and the closedness of $\partial\mathcal{R}(0)$, we find $m(s) = 0$ everywhere in $[s_1, s_2]$ and, since G was chosen arbitrary in $\mathcal{T}$, also in $[0, S]$. Therefore

$$-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)) \in \partial\mathcal{R}(0) \quad \forall s \in [0, S],$$

which means that $\hat{z}(s) \in \mathcal{S}_{loc}(\hat{t}(s))$. Thus, exploiting Lemma 2.1.2 we get $z(t) \in S_{glob}(t)$ for all $t \in [0, T]$. It remains to show that the energy identity (E) holds. More precisely, it suffices to show the lower energy estimate, that is, "$\leq$" in (E), since the opposite inequality is proven by adapting the arguments from *Step 5* in the proof of Theorem 2.3.4. To proceed, we note that given $t \in [0, T]$ there exists $s_t \in [0, S]$, depending on t, with $\hat{t}(s_t) = t$ by the construction of z. Consequently, Lemma 2.3.2 applied to $\hat{z}$ leads to

$$\mathrm{Diss}_{\mathcal{R}}(z; [0, t]) \leq \mathrm{Diss}_{\mathcal{R}}(\hat{z}; [0, s_t]) = \int_0^{s_t} \mathcal{R}(\hat{z}'(r))\,\mathrm{d}r. \tag{2.6.4}$$

Therewith, we obtain from the energy identity for parametrized solutions (2.4.6):

$$\begin{aligned}
&\mathcal{I}(t, z(t)) + \mathrm{Diss}_{\mathcal{R}}(z, [0, t]) \\
&\quad \leq \mathcal{I}(\hat{t}(s_t), \hat{z}(s_t)) + \int_0^{s_t} \mathcal{R}(\hat{z}'(r))\,\mathrm{d}r \\
&\quad \leq \mathcal{I}(\hat{t}(s_t), \hat{z}(s_t)) + \int_0^{s_t} \mathcal{R}(\hat{z}'(r)) + \|\hat{z}'(r)\|_{\mathbb{V}}\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(r), \hat{z}(r)), \partial\mathcal{R}(0)\}\,\mathrm{d}r \\
&\quad = \mathcal{I}(0, z_0) + \int_0^{s_t} \partial_t\mathcal{I}(\hat{t}(r), \hat{z}(r))\hat{t}'(r)\,\mathrm{d}r \\
&\quad = \mathcal{I}(0, z_0) + \int_0^{t} \partial_t\mathcal{I}(r, z(r))\,\mathrm{d}r,
\end{aligned}$$

where we used the change of variable formula from Appendix A.2.8 in the last line. Overall this proves that $z \in \mathfrak{P}(\hat{t}, \hat{z})$ is an energetic solution. □

Remark 2.6.5. The additional regularity $\hat{z} \in W^{1,1}(0, S; \mathcal{Z})$ in the previous lemma, is in fact only necessary in order to obtain that the set $\mathcal{S}$ is closed. Hence, we may preserve the statement of Lemma 2.6.4 if one of the following two conditions is fulfilled (see also (2.2.3a) and (2.2.3b)):

(i) $\mathcal{R}$ is weakly continuous on $\mathcal{Z}$

(ii) $D_z\mathcal{I}(\cdot, \cdot)$ is (strong,weak)-weak continuous from $\mathcal{Z}$ to $\mathcal{Z}^*$, i.e. $\forall\, t_k \to t,\ z_k \rightharpoonup z$ in $\mathcal{Z}$:

$$D_z\mathcal{I}(t_k, z_k) \rightharpoonup D_z\mathcal{I}(t, z) \quad \text{in } \mathcal{Z}^*.$$

Indeed, in the first case, we apply Lemma 2.1.2 and adapt the argumentation in *Step 1* of Theorem 2.2.2. For the second case, we combine the weak continuity from (2.4.7) and the weak closedness of $\partial\mathcal{R}(0)$.

The problem with the transformation of an energetic solution z into a parametrized one is the missing regularity. Indeed, if z could be shown to be more regular, e.g., if z is an element of

$BV(0,T;\mathcal{V})$, then we would also obtain the opposite direction in Lemma 2.6.4. The idea here is to "fill up" the jumps of z using affine interpolation and to parameterize the obtained function by arc-length. Nevertheless, for convex energies, it holds that every differential solution also leads to a parametrized one, as it is seen by the subsequent Lemma 2.6.6. Moreover, parametrized solutions are a subset of the energetic solutions in the sense of Lemma 2.6.4, that is, after projecting them onto the physical time. Note that all dependencies are also visualized at the end of this section.

Now, let us turn to the general nonconvex case. Here, the existence of a differential solution is no longer guaranteed. However, if it does exist, then it also gives rise to a parametrized solution.

Lemma 2.6.6. *Let $\mathcal{R}$ comply with (R1)-(R3) and $\mathcal{I}$ be as in (E1)-(E4), but not necessarily convex. If there exists a differential solution $z \in W^{1,1}(0,T;\mathcal{Z})$ to (RIS), then its arc-length parametrization $(\hat{t},\hat{z})$ is a parametrized solution.*

Proof. We define $s(t) := t + \int_0^t \|z'(r)\|_{\mathbb{V}}\,\mathrm{d}r$ and $S = s(T)$. Then $s : [0,T] \to [0,S]$ is strictly monotone and Lipschitz continuous and therefore provides an inverse function which we denote by $\hat{t} : [0,S] \to [0,T]$. In particular $\hat{t}$ is again Lipschitz continuous, since $|s(t_2) - s(t_1)| \geq |t_2 - t_1|$. For $\hat{z}(s) := z(\hat{t}(s))$, we have that $\hat{z}'(s) = z'(\hat{t}(s))\,\hat{t}'(s)$ almost everywhere and consequently, since z is a differential solution,

$$\partial\mathcal{R}(\hat{z}'(s)) + D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)) = \partial\mathcal{R}(z'(\hat{t}(s))) + D_z\mathcal{I}(\hat{t}(s), z(\hat{t}(s))) \ni 0. \tag{2.6.5}$$

Here we used the 0-homogeneity of $\partial\mathcal{R}(\cdot)$ and the fact that $\hat{t}'(s) > 0$. By the construction we have $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} = 1$ almost everywhere in $[0,S]$, which gives the regularity $(\hat{t},\hat{z}) \in W^{1,\infty}(0,S;\mathbb{R}) \times W^{1,\infty}(0,S;\mathcal{V})$. The additional property $\hat{z} \in L^\infty(0,S;\mathcal{Z})$ follows easily from the embedding $W^{1,1}(0,T;\mathcal{Z}) \hookrightarrow L^\infty(0,T;\mathcal{Z})$. Using the characterization in (2.1.7), we obtain from (2.6.5) that $\mathcal{R}(\hat{z}'(s)) = \langle -D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)), \hat{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}}$ almost everywhere, so that an application of the chain rule from Lemma A.2.5 and integration in time gives the energy identity (2.4.6) for parametrized solutions with $\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)), \partial\mathcal{R}(0)\} = 0$ (cf. the argumentation in the context of (2.1.12b)). Overall, we see that $(\hat{t},\hat{z})$ is indeed a parametrized solution. □

In contrast, this does, in general, not hold for the concept of energetic solutions, as can be seen in the subsequent Example. Moreover, parametrized and energetic solutions should be regarded as two opposite solution concepts for rate-independent systems. Indeed, the global character of the energetic solution via the global stability property induces that these solutions jump as soon as possible, ignoring the local behavior of the energy landscape, cf. Example 2.3.5. In contrast, the concept of parametrized solution coming from the analysis of vanishing viscosity produces solutions that try to delay jumps. This is due to the local character of this concept and reflects itself in the additional viscosity term $\lambda\mathbb{V}\hat{z}'$.

Example 2.6.7 (Differential and energetic solutions differ, cf. [MR15, Ex. 1.8.2]).
Here we use the same setting as in Example 2.4.1, that is, $\mathcal{Z} = \mathcal{X} = \mathbb{R}$ and

$$\mathcal{I} : \mathbb{R}\times\mathbb{R} \to \mathbb{R}, \quad \mathcal{I}(t,z) = \mathcal{E}(z) - \ell(t)\,z \ \text{ with } \ \mathcal{E}(z) = \begin{cases} \frac{1}{2}(z+4)^2, & z \leq -2, \\ 4 - \frac{1}{2}z^2, & |z| < 2, \\ \frac{1}{2}(z-4)^2, & z \geq 2, \end{cases}$$

but take $z_0 = -5$ and $\ell(t) = \min(t, 4-t)$. Then there exists both a differential solution z_1 and an energetic solution z_2, namely

$$z_1(t) = \begin{cases} t-5, & ,t \in [0,2], \\ -3, & t \in (2,4], \end{cases} \quad \text{and} \quad z_2(t) = \begin{cases} t-5, & t \in [0,1], \\ t+3, & t \in (1,2], \\ 5, & t \in (2,4], \end{cases}$$

which obviously do not coincide on the considered time horizon. In particular, the energetic solution cannot be brought into a differential one, since the state $z(t) = t-5$ is not locally stable for $t \in (1,2]$. We also refer to [Ste09, Sec. 3] for another example showing that the global and local stability set do not have to coincide. This naturally implies that differential solutions are not necessarily energetic ones and vice versa.

We close this chapter with Figure 2.6.1, which illustrates the connections between the different solution concepts. Loosely speaking, we have the three cases:

- $\mathcal{I}(t,\cdot)$ uniformly convex: differential sol. $=$ parametrized sol. $=$ energetic sol.
- $\mathcal{I}(t,\cdot)$ convex: differential sol. $\subset$ parametrized sol. $\subset$ energetic sol.
- $\mathcal{I}(t,\cdot)$ nonconvex: differential sol. $\subset$ parametrized sol. $\neq$ energetic sol.

In general, it is possible to obtain the opposite inclusion in the above listing if the solutions provide the necessary regularity. So, for example if a parametrized solution $(\hat{t}, \hat{z})$ fulfills $\hat{t}'(s) \geq \delta > 0$ then there exists an inverse function $\hat{t}^{-1}$ and the corresponding transformed function $\hat{z} \circ \hat{t}^{-1} =: \tilde{z}$ is a differential solution. A similar result also holds for energetic solutions, provided they have at least the regularity $W^{1,1}(0,T;\mathcal{Z})$ (see proof of Lemma 2.6.1).

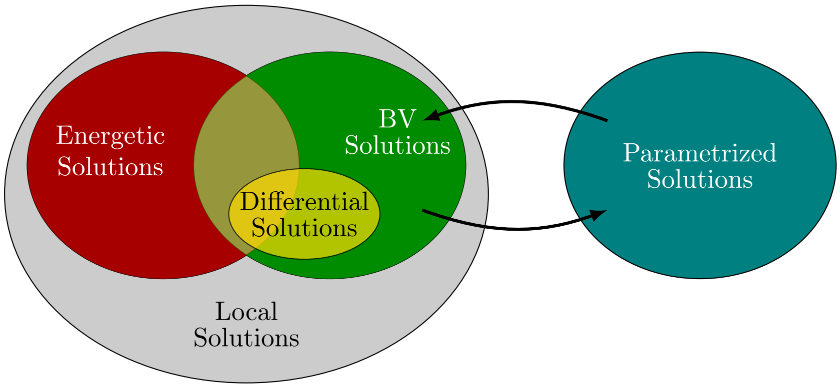


Figure 2.6.1: Overview of different solution concepts. Note that the parametrized solutions stand outside of the other notions of solutions, since they are defined on an artificial time. However, as shown in the previous section, their projections as defined in (2.5.2) are directly related to all the other concepts.

Chapter 3

Local minimization scheme for parametrized solutions

Subsequent to this introduction into the various different types of solutions, or at least the ones most commonly used, we now consider their approximability by discretization. For the class of global energetic solutions, there exists multiple papers on this topic. We only refer to [MR15, MR09] and the references therein. In contrast, less is known for parametrized solutions to the best of the author's knowledge. In [EM06], the authors introduced the following *time-incremental local minimization scheme* in order to approximate parametrized solutions:

$$z_k \in \arg\min\{\mathcal{I}(t_{k-1}, z) + \mathcal{R}(z - z_{k-1}) \,:\, z \in \mathcal{Z},\ \|z - z_{k-1}\|_{\mathbb{V}} \leq \tau\}, \tag{3.0.1a}$$

$$t_k = \min\{t_{k-1} + \tau - \|z_k - z_{k-1}\|_{\mathbb{V}}, T\}. \tag{3.0.1b}$$

The motivating background for this scheme is best explained by a comparison with the scheme (2.3.4) introduced in Section 2.3 for the approximation of global energetic solutions:

$$z_k \in \arg\min\{\mathcal{I}(t_{k-1}, z) + \mathcal{R}(z - z_{k-1}) \,:\, z \in \mathcal{Z}\}, \tag{3.0.2a}$$

$$t_k = t_{k-1} + \tau. \tag{3.0.2b}$$

Thanks to the positive homogeneity of $\mathcal{R}$, the stationarity condition of (3.0.2a) is given by

$$0 \in \partial\mathcal{R}\left(\frac{z_k - z_{k-1}}{\tau}\right) + D_z\mathcal{I}(t_{k-1}, z_k), \tag{3.0.3}$$

which, in view of (RIS), motivates the scheme in (3.0.2). However, in general, the global minimization in (3.0.2a) may induce unphysical discontinuities. Consider for instance a situation, where the energy difference between a local minimum of $\mathcal{I}(t_{k-1}, \cdot)$ in the vicinity of z_{k-1} to a global minimum of $\mathcal{I}(t_{k-1}, \cdot)$ is so large that it cannot be compensated by the dissipation. Then the iteration will jump to a global minimizer which is certainly not physically meaningful in many applications, cf. Example 2.3.5. This motivates the additional inequality constraint in (3.0.1a) which enforces the search for local minimizers in the neighborhood of the old iterate. If, however, there is no such local minimizer so that the inequality constraint in (3.0.1a) is active (i.e., fulfilled with equality), then the stationarity condition in (3.0.3) will (in general) not be fulfilled. In this case, one therefore interrupts the evolution of the physical time (see the update in (3.0.1b)) until the state z fulfills the stationarity condition (3.0.3) again. This may be seen as a discrete analogon to the viscous jump described in Remark 2.4.7. As indicated above, the literature on the approximation of parametrized solutions is rather scarce. In the aforementioned work [EM06], a convergence theory for the local minimzation scheme (3.0.1) is developed for the finite dimensional case, where $\dim(\mathcal{Z}) < \infty$. The authors prove that, for $\tau \searrow 0$, subsequences of discrete solutions (weakly) converge to a parametrized solution. However, it is not shown that the sequence $\{t_k\}$, generated by (3.0.1b), reaches the desired final time in a finite number of iterations. The same holds for a variant of (3.0.1), which is investigated in [Neg14]. Another modification of (3.0.1), which does not account for the adaptive time discretization in (3.0.1b), is considered in [MS17]. For this variant, the authors show that the final time is reached after a finite number of iterations. Recently, the original scheme in (3.0.1) was investigated in [Kne19] for a certain class of infinite dimensional problems (i.e., $\dim(\mathcal{Z}) = \infty$), providing a comprehensive convergence analysis. In particular, it is shown that the final time is reached in a finite number of steps and that subsequences of iterates (weakly) converge to a parametrized solution.

The following chapter is concerned with the analysis of the minimization scheme in (3.0.1) and relies on the papers [MS19a] and [MS20]. To be precise, we consider a scheme which is based on (3.0.1) but requires only stationary points instead of global minima, see LISS below. In addition, we treat a *full discretization in time and space* in Section 3.2. Here, the recent work [Kne19] serves as a starting point and we include the additional errors induced by the discretization of the infinite dimensional state space $\mathcal{Z}$ into the convergence analysis. Moreover, we relax the assumption on the boundedness of the dissipation and thus allow for more general settings, e.g., damage models (see Section 3.2.5). In the subsequent Section 3.3, we focus on a priori error estimates for (3.0.1) without the additional space discretization. The actual numerical realization of the scheme, as well as numerical results, are presented in Chapter 4.

3.1 General assumptions

Before we actually start with the convergence analysis of the local minimization scheme, we need to slightly strenghten the assumptions on the energy from the Introduction. We furthermore set the requirements on the spatial discretization. A prototypical example fulfilling all assumptions is given in Section 3.2.5 below. Concerning the underlying spaces, we take the same setting as in the Introduction, that is, $\mathcal{X}$ is a Banach space and $\mathcal{Z}, \mathcal{V}$ are Hilbert spaces such that $\mathcal{Z} \overset{c,d}{\hookrightarrow} \mathcal{V} \hookrightarrow \mathcal{X}$.

Energy. The energy functional $\mathcal{I}$ is supposed to have the following form:

$$\mathcal{I} : [0,T] \times \mathcal{Z} \to \mathbb{R}, \quad \mathcal{I}(t,z) = \frac{1}{2}\langle Az, z\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \mathcal{F}(z) - f(t,z)\,. \tag{$\mathcal{I}_0$}$$

Herein, $A \in \mathcal{L}(\mathcal{Z}, \mathcal{Z}^*)$ is a self-adjoint and coercive operator, i.e., there is a constant $\alpha > 0$ such that $\langle Az, z\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \alpha \|z\|^2_{\mathcal{Z}}$. The bilinear form $a : \mathcal{Z} \times \mathcal{Z} \to \mathbb{R}$ induced by $a(y,z) = \langle Ay, z\rangle_{\mathcal{Z}^*,\mathcal{Z}}$, $y, z \in \mathcal{Z}$, is thus bounded and coercive, too. Moreover, we assume that the time-dependent part

$f : [0,T] \times \mathcal{Z} \to \mathbb{R}$ fulfills for some $c_f, \mu > 0$:

$$\begin{gathered} f \in C^1([0,T] \times \mathcal{Z}; \mathbb{R}) \\ \text{with} \quad f(t,z) \leq c_f(\|z\|_{\mathcal{Z}} + 1) \quad \text{and} \quad |\partial_t f(t,z)| \leq \mu(\|z\|_{\mathcal{Z}} + 1) \end{gathered} \tag{$\mathcal{I}_{f1}$}$$

for all $t \in [0,T]$ and $z \in \mathcal{Z}$. For the derivative with respect to z we assume that there exists an intermediate space $\mathcal{Z} \subset \mathcal{W} \subset \mathcal{V}$ with $\mathcal{Z} \hookrightarrow^c \mathcal{W}$, such that

$$|\langle D_z f(t_1, z_1) - D_z f(t_2, z_2), v\rangle_{\mathcal{V}^*,\mathcal{V}}| \leq \nu(|t_1 - t_2| + \|z_1 - z_2\|_{\mathcal{W}})\|v\|_{\mathcal{V}} . \tag{$\mathcal{I}_{f2}$}$$

Finally, for the time-dependent part of the energy, we suppose that

$$\begin{gathered} f(0,\cdot) \text{ is weakly continuous and} \\ \text{for all sequences } t_k \to t \text{ and } z_k \rightharpoonup z \text{ in } \mathcal{Z} \text{ it holds:} \quad \partial_t f(t_k, z_k) \to \partial_t f(t,z). \end{gathered} \tag{$\mathcal{I}_{f3}$}$$

Writing $f(t,z_k) = f(0,z_k) + \int_0^t \partial_t f(s,z_k)\,\mathrm{d}s$ and exploiting both properties from ($\mathcal{I}_{f3}$) above, we conclude by Lebesgue's dominated convergence Theorem that, in fact, $f(t,\cdot)$ is weakly continuous for all $t \in [0,T]$.

Remark 3.1.1. The assumptions on the external force are for example fulfilled if $f(t,z)$ has a linear structure, i.e., $f(t,z) = \langle \ell(t), z\rangle_{\mathcal{V}^*,\mathcal{V}}$ for some $\ell \in C^1(0,T;\mathcal{V}^*)$. A more sophisticated example is given in Section 3.2.5

Remark 3.1.2. As described in the Introduction, a rate-independent system is solely driven by an external load. In our setting, this load is contained in the time-dependent part $f(t,z)$, i.e., $f(t,z) = \tilde{f}(\ell(t),z)$ for some $\tilde{f} : \mathcal{V} \times \mathcal{Z} \to \mathbb{R}$. However, in order to keep the notation short, we do not use this explicit dependence (see also the damage example in Section 3.2.5 below).

Regarding the nonlinearity $\mathcal{F} : \mathcal{Z} \to \mathbb{R}$ we require

$$\begin{gathered} \mathcal{F} \in C^2(\mathcal{Z}; \mathbb{R}) \text{ with } \mathcal{F} \geq 0, \\ \text{as well as} \quad D_z\mathcal{F} \in C^1(\mathcal{Z}, \mathcal{V}^*), \quad \|D_z^2\mathcal{F}(z)v\|_{\mathcal{V}^*} \leq C(1 + \|z\|_{\mathcal{Z}}^q)\|v\|_{\mathcal{Z}} \end{gathered} \tag{$\mathcal{I}_{\mathcal{F}1}$}$$

for some $q \geq 1$ so that, for every $z \in \mathcal{Z}$, $D_z\mathcal{F}(z)$ can uniquely be extended to a bounded and linear functional on $\mathcal{V}$, which we denote by the same symbol for convenience. In particular, this implies that $\mathcal{I} \in C^1([0,T] \times \mathcal{Z}; \mathbb{R})$. Moreover, $D_z\mathcal{F} : \mathcal{Z} \to \mathcal{Z}^*$ is supposed to be weak-weak continuous, i.e.,

$$\text{for all sequences } z_k \rightharpoonup z \text{ in } \mathcal{Z} \text{ it holds:} \quad D_z\mathcal{F}(z_k) \rightharpoonup D_z\mathcal{F}(z). \tag{$\mathcal{I}_{\mathcal{F}2}$}$$

Again, this property together with the fact that $D_z\mathcal{F} \in C^1(\mathcal{Z}, \mathcal{V}^*)$ and the compact embedding $\mathcal{V} \hookrightarrow \mathcal{Z}$ yields that $\mathcal{F}$ is also weakly continuous from $\mathcal{Z}$ to $\mathbb{R}$. Hence, $\mathcal{I}(t,\cdot)$ is weakly lower-semicontinuous. Furthermore, the weak continuity of $D_z\mathcal{F}$ and ($\mathcal{I}_{f2}$) in combination with the compact embedding $\mathcal{Z} \hookrightarrow^c \mathcal{W}$ implies that $D_z\mathcal{I}(\cdot,\cdot)$ is (strong,weak)-weak continuous from $\mathcal{Z}$ to $\mathcal{Z}^*$, i.e.,

$$\text{for all sequences } t_k \to t \text{ and } z_k \rightharpoonup z \text{ in } \mathcal{Z} \text{ it holds:} \quad D_z\mathcal{I}(t_k, z_k) \rightharpoonup D_z\mathcal{I}(t,z) \quad \text{in } \mathcal{Z}^*, \tag{3.1.1}$$

see also Remark 2.6.5. The above assumptions on the involved functionals allow us to obtain the following estimates on the energy. By setting $c_0 := (c_f + 1)^2/(4\alpha) + c_f$ we obtain the estimate

$$\mathcal{I}(t, z) \geq \alpha\|z\|_{\mathcal{Z}}^2 - c_f(\|z\|_{\mathcal{Z}} + 1) \geq \|z\|_{\mathcal{Z}} - c_0 \qquad \forall\, t \in [0, T],\ z \in \mathcal{Z}\,, \tag{3.1.2}$$

which in turn implies that

$$|\partial_t \mathcal{I}(t, z)| \leq \mu(\|z\|_{\mathcal{Z}} + 1) \leq \mu(\mathcal{I}(t, z) + \beta) \qquad \text{f.a.a. } t \in [0, T],\ z \in \mathcal{Z}\,,$$

with $\beta := c_0 + 1$. Gronwall's lemma thus gives for all $t, s \in [0, T]$, $z \in \mathcal{Z}$ that

$$\mathcal{I}(t, z) + \beta \leq (\mathcal{I}(s, z) + \beta) \exp(\mu|t - s|) \tag{3.1.3}$$

$$\text{and} \quad |\partial_t \mathcal{I}(t, z)| \leq \mu(\mathcal{I}(s, z) + \beta) \exp(\mu|t - s|), \tag{3.1.4}$$

Remark 3.1.3. Note that ($\mathcal{I}_0$), ($\mathcal{I}_{f1}$) and ($\mathcal{I}_{\mathcal{F}1}$) as well as the weak continuity of $\mathcal{F}$ and $f(t, \cdot)$ imply (E1) and (E2). Moreover, assumption ($\mathcal{I}_{f3}$) directly yields (E4) and finally, (E3) is guaranteed by the estimate in ($\mathcal{I}_{f1}$). Thus, the energy $\mathcal{I}$ in ($\mathcal{I}_0$) perfectly fits into the setting from the Introduction.

Dissipation. For the dissipation $\mathcal{R}$, we basically take the same assumptions as in the introduction, i.e., (R1)-(R3). We merely strengthen the continuity of $\mathcal{R}$ in the following sense:

$$\mathcal{R} \text{ is continuous on } \operatorname{dom}(\mathcal{R}). \tag{$\mathcal{R}_0$}$$

Hence, we have that $\mathcal{R} : \mathcal{Z} \to [0, \infty]$ is lower semicontinuous (l.s.c.), convex, and positively homogeneous of degree one and additionally continuous on its domain.

Remark 3.1.4. Note that from now on, we consider $\mathcal{R}$ as mapping from $\mathcal{Z}$ into $\mathbb{R}$. In fact, we will subsequently always evaluate $\mathcal{R}$ at a point in $\mathcal{Z}$. Thus, the space $\mathcal{X}$ is not used in the convergence analysis.

Initial state. The initial value z_0 is supposed to satisfy $z_0 \in \mathcal{Z}$ and $Az_0 \in \mathcal{V}^*$, which implies some additional regularity for z_0.

Spatial discretization. For the (spatial) discretization let $\mathcal{Z}_h \subset \mathcal{Z}$ be a finite dimensional subspace, where $h > 0$ indicates the fineness of the approximation, and denote by $\Pi_h : \mathcal{V} \to \mathcal{Z}_h$ the associated orthogonal projection w.r.t. the $\mathbb{V}$-norm. Then we assume that Π_h is stable w.r.t. the $\mathcal{Z}$-norm, i.e.,

$$\|\Pi_h(z)\|_{\mathcal{Z}} \leq C\|z\|_{\mathcal{Z}} \quad \forall z \in \mathcal{Z} \tag{3.1.5}$$

with a constant $C > 0$ independent of h. Note that this already implies the best approximation property of the orthogonal projection, i.e.,

$$\|z - \Pi_h(z)\|_{\mathcal{Z}} \leq \inf_{z_h \in \mathcal{Z}_h} (1 + C)\|z - z_h\|_{\mathcal{Z}}. \tag{3.1.6}$$

The stability assumption in (3.1.5) is fulfilled in prominent cases such as finite element discretizations based on shape-regular triangulations, as we will see in Section 4.1.1 below. In the following, we will frequently consider Π_h as an operator in $\mathcal{V}$ and $\mathcal{Z}$, respectively, denoted for simplicity by the same symbol.

We further introduce the Ritz-projection $P_h : \mathcal{Z} \to \mathcal{Z}_h$ as unique solution of

$$P_h(u) \in \mathcal{Z}_h\,, \quad a(P_h(u), v) = a(u, v) \qquad \forall v \in \mathcal{Z}_h\,,$$

where a is the bilinear form induced by A. For the inital value of the algorithm, we set $z_0^{\tau,h} := P_h(z_0) \in \mathcal{Z}_h$.

Furthermore, it is assumed that, for all $v \in \mathcal{V}$ and all $z \in \mathcal{Z}$, respectively, it holds

$$\Pi_h(v) \to v \quad \text{in } \mathcal{V} \quad \text{and} \quad P_h(z) \to z \quad \text{in } \mathcal{Z} \tag{3.1.7}$$

as $h \searrow 0$. Note that the stability property in (3.1.6) then automatically yields for all $z \in \mathcal{Z}$ that

$$\Pi_h(z) \to z \quad \text{in } \mathcal{Z}. \tag{3.1.8}$$

Lastly, we allow that the dissipation functional $\mathcal{R}$ is not evaluated exactly for the discrete iterate $z_k^{\tau,h}$, but merely approximated by a functional $\mathcal{R}_h : \mathcal{Z}_h \to [0, \infty]$ satisfying the following conditions:

(a) Analogously to $\mathcal{R}$, its approximation $\mathcal{R}_h$ is convex, lower semicontinuous and positively homogeneous.

(b) Furthermore, for every $v_h \in \mathcal{Z}_h$, it holds $\mathcal{R}(v_h) \le \mathcal{R}_h(v_h)$. This particularly implies that $\mathcal{R}_h \ge 0$ as well.

(c) There is a dense subset $\mathcal{U} \subset \text{dom}(\mathcal{R})$ such that, for every $v \in \mathcal{U}$, there holds $\mathcal{R}_h(\Pi_h v) \to \mathcal{R}(v)$ as $h \searrow 0$.

Note that the choice $\mathcal{R}_h = \mathcal{R}$ (i.e., no additional approximation of $\mathcal{R}$) fulfills all these assumptions. Another example that complies with all the assumptions (a)-(c) but satisfies $\mathcal{R}_h \neq \mathcal{R}$ is given in Section 4.1.1 below.

3.2 Convergence analysis

The ultimate goal of this section is to prove that the subsequent algorithm, which is based on the local incremental minimization scheme (3.0.1), provides an approximation scheme for parametrized solutions. The difference compared to (3.0.1) is that we search for stationary points of the constrained problem (alg_1) rather than global minima. Beyond that, we consider a full discretization in space and time. Thus, for a given time-discretization parameter $\tau > 0$, the fully discrete algorithm reads as follows:

Fully discrete local incremental stationarity scheme (LISS)**.**

1: *Set* $z_0^{\tau,h} = P_h(z_0)$, $t_0 = 0$, *and* $k = 1$
2: ***while*** $t_k^{\tau,h} < T$ ***do***

3: *Compute a* stationary point $z_k^{\tau,h}$, *i.e.,*

$$0 \in \partial^{\mathcal{Z}_h}(\mathcal{R}_h + I_\tau)(z_k - z_{k-1}) + D_z\mathcal{I}(t_{k-1}, z_k) \tag{alg$_1$}$$

with the indicator function I_τ *(see* (3.2.2)*), which, additionally, satisfies*

$$\mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}) + \mathcal{R}_h(z_k^{\tau,h} - z_{k-1}^{\tau,h}) \leq \mathcal{I}(t_{k-1}^{\tau,h}, z_{k-1}^{\tau,h}). \tag{alg$_2$}$$

4: *Time update:*

$$t_k^{\tau,h} = t_{k-1}^{\tau,h} + \tau - \|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathbb{V}}. \tag{alg$_3$}$$

5: *Set* $k \to k+1$.

6: ***end while***

Note that merely for technical reasons, we do not use the "min" from (3.0.1b) in the time-update. The proposed method is closely related to (3.0.1), since a local minimizer of

$$\min\{\mathcal{I}(t_{k-1}^{\tau,h}, z) + \mathcal{R}_h(z - z_{k-1}^{\tau,h}) \,:\, z \in \mathcal{Z}_h,\, \|z - z_{k-1}^{\tau,h}\|_{\mathbb{V}} \leq \tau\} \tag{3.2.1}$$

necessarily satisfies (alg$_1$) as we will see in the subsequent Lemma 3.2.1. Moreover, thanks to the assumptions on $\mathcal{I}$ and $\mathcal{R}_h$, in particular weak lower semicontinuity, the existence of a global minimum of (3.2.1) and therefore also the existence of a stationary point fulfilling (alg$_2$) is guaranteed by the direct method in the calculus of variations.

Lemma 3.2.1. *If* $z_k^{\tau,h}$ *is a local minimizer of* (3.2.1)*, then* (alg$_1$) *is satisfied. If* $z_k^{\tau,h}$ *is even a global minimizer, then* (alg$_1$) *and* (alg$_2$) *hold true.*

Proof. For the ease of readability, we suppress the superscript τ, h throughout the proof. We moreover define $I_\tau : \mathcal{V} \to [0,\infty]$ as the indicator functional associated with the constraints in (3.2.1), i.e.,

$$I_\tau(v) := \begin{cases} 0, & \text{if } \langle \mathbb{V}v, v\rangle \leq \tau^2, \\ +\infty, & \text{else.} \end{cases} \tag{3.2.2}$$

Now, let z_k be a local minimum of (3.2.1). Then it holds

$$\begin{aligned}\mathcal{I}(t_{k-1}, z_k) &+ \mathcal{R}_h(z_k - z_{k-1}) + I_\tau(z_k - z_{k-1}) \\ &\leq \mathcal{I}(t_{k-1}, z) + \mathcal{R}_h(z - z_{k-1}) + I_\tau(z - z_{k-1}) \quad \forall z \in \mathcal{Z}_h.\end{aligned}$$

For arbitrary $w \in \mathcal{Z}_h$ and $t \in (0,1]$ we test this inequality with $z = z_k + t(w - z_k)$. In combination with the convexity of I_τ and $\mathcal{R}_h$ and after rearranging terms, we thus obtain

$$\begin{aligned}(1-t)\mathcal{R}_h(z_k - z_{k-1}) &+ t\,\mathcal{R}_h(w - z_{k-1}) + (1-t)I_\tau(z_k - z_{k-1}) + t\,I_\tau(w - z_{k-1}) \\ &\geq \mathcal{I}(t_{k-1}, z_k) - \mathcal{I}(t_{k-1}, z_k + t(w - z_k)) + \mathcal{R}_h(z_k - z_{k-1}) + I_\tau(z_k - z_{k-1}) \quad \forall w \in \mathcal{Z}_h.\end{aligned}$$

Another rearrangement of terms and division by t yields

$$\begin{aligned}&\mathcal{R}_h(w - z_{k-1}) + I_\tau(w - z_{k-1})\\&\qquad\geq \frac{\mathcal{I}(t_{k-1}, z_k) - \mathcal{I}(t_{k-1}, z_k + t(w - z_k))}{t} + \mathcal{R}_h(z_k - z_{k-1}) + I_\tau(z_k - z_{k-1}).\end{aligned}$$

Substituting herein $w = v + z_{k-1}$ and passing to the limit $t \searrow 0$ we eventually end up with

$$\begin{aligned}&\mathcal{R}_h(v) + I_\tau(v)\\&\qquad\geq \mathcal{R}_h(z_k - z_{k-1}) + I_\tau(z_k - z_{k-1}) + \langle -D_z\mathcal{I}(t_{k-1}, z_k), v - (z_k - z_{k-1})\rangle_{\mathcal{Z}^*,\mathcal{Z}} \quad \forall v \in \mathcal{Z}_h,\end{aligned}$$

which is equivalent to saying

$$-D_z\mathcal{I}(t_{k-1}, z_k) \in \partial^{\mathcal{Z}_h}(\mathcal{R}_h + I_\tau)(z_k - z_{k-1}).$$

The fact that a global minimizer satisfies (alg$_2$) is obvious. □

The reason for investigating LISS instead of (3.0.1), is the fact that a numerical algorithm for solving (3.0.1a) or rather (3.2.1) naturally provides a stationary point $z_k^{\tau,h}$ that satisfies $\mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}) + \mathcal{R}_h(z_k^{\tau,h} - z_{k-1}^{\tau,h}) \leq \mathcal{I}(t_{k-1}^{\tau,h}, z_{k-1}^{\tau,h})$ but, in case of a nonconvex energy, is not guaranteed to be a global optimum of (3.0.1a) and (3.2.1), respectively. Moreover, the concept of parametrized solutions is based on a local stability condition. It is thus consistent to look for locally stable points, which are exactly the stationary points of (3.0.1a). Despite its necessity for the convergence analysis, the inequality in (alg$_2$) is also physically meaningful since it enforces the system to look for energetically preferable states, i.e., states with a lower energy cost. Concerning the exploration of this algorithm, particularly with a view to convergence, we proceed as follows: We start with characterizing properties of the stationary points. Afterwards, we turn to the essential a priori estimates that will allow a passage to the limit in the discrete version of the energy identity in (2.4.6), which is deduced in Section 3.2.3. The limit procedure itself is elaborated in the final Section 3.2.4.

3.2.1 Approximate discrete parametrized solution

The foundation for both, the a priori estimates and the discrete version of the energy identity, is given by the following Lemma 3.2.2. It provides various properties of a stationary point $z_k^{\tau,h}$ in (alg$_1$) and shows some similarities with the formulation in (2.4.8). Indeed, we will see that one can interpret the stationarity condition as a discrete version of (2.4.8).

Lemma 3.2.2 (Discrete optimality System)**.** *Let $k \geq 1$ and $z_k^{\tau,h}$ be an arbitrary stationary point in the sense of* (alg$_1$) *with associated $t_{k-1}^{\tau,h}$ given by* (alg$_3$)*. Then the following properties are satisfied: There exists a Lagrange multiplier $\lambda_k^{\tau,h} \geq 0$ such that*

$$\lambda_k^{\tau,h}(\|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathbb{V}} - \tau) = 0, \tag{3.2.3a}$$

$$\tau\,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\Pi_h^* D_z\mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} = \lambda_k^{\tau,h}\|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathbb{V}}^2, \tag{3.2.3b}$$

$$\left.\begin{aligned}\mathcal{R}_h(z_k^{\tau,h} - z_{k-1}^{\tau,h}) + \tau\,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\Pi_h^* D_z\mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}), \partial(\mathcal{R}_h \circ \Pi_h)(0)\}\\ = \langle -D_z\mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}), z_k^{\tau,h} - z_{k-1}^{\tau,h}\rangle_{\mathcal{Z}^*,\mathcal{Z}}\end{aligned}\right\} \tag{3.2.3c}$$

$$\mathcal{R}_h(\Pi_h(v)) \geq -\langle \lambda_k^{\tau,h}\mathbb{V}(z_k^{\tau,h} - z_{k-1}^{\tau,h}) + \Pi_h^* D_z\mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}), v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \quad \forall v \in \mathcal{Z}. \tag{3.2.3d}$$

Herein, $\overline{\mathrm{dist}}_{\mathcal{V}^*}\{\,\cdot\,, \partial(\mathcal{R}_h \circ \Pi_h)(0)\}$ *denotes the extended distance as defined in Lemma A.3.8*

Proof. Again, to shorten the notation, we suppress the superscripts τ, h throughout the proof. Hence, let $z_k \in \mathcal{Z}_h$ be a stationary point as in (alg_1), i.e., z_k satisfies

$$0 \in \partial^{\mathcal{Z}_h}(\mathcal{R}_h + I_\tau)(z_k - z_{k-1}) + D_z\mathcal{I}(t_{k-1}, z_k) \tag{3.2.4}$$

with the indicator functional I_τ from (3.2.2). By definition this is equivalent to

$$\begin{aligned}\mathcal{R}_h(v_h) + I_\tau(v_h) \geq \mathcal{R}_h(z_k - z_{k-1}) + I_\tau(z_k - z_{k-1})\\ + \langle -D_z\mathcal{I}(t_{k-1}, z_k), v_h - (z_k - z_{k-1})\rangle_{\mathcal{Z}_h^*,\mathcal{Z}_h} \quad \forall v_h \in \mathcal{Z}_h.\end{aligned} \tag{3.2.5}$$

Note that the term $\partial^{\mathcal{Z}_h}\mathcal{R}_h$ in (3.2.5) corresponds to the subdifferential of $\mathcal{R}_h$ with respect to the space $\mathcal{Z}_h$, see Definition A.3.1. However, in view of the forthcoming convergence analysis, we transform $\partial^{\mathcal{Z}_h}\mathcal{R}_h$ into a subdifferential with respect to the space $\mathcal{Z}$ and its dual pairing. Therefore, we make use of the projection operator $\Pi_h : \mathcal{Z} \to \mathcal{Z}_h$, whose surjectivity allows us to formulate (3.2.5) as

$$\begin{aligned}\mathcal{R}_h(\Pi_h(v)) + I_\tau(\Pi_h(v)) \geq \mathcal{R}_h(\Pi_h(z_k - z_{k-1})) + I_\tau(\Pi_h(z_k - z_{k-1}))\\ + \langle -D_z\mathcal{I}(t_{k-1}, z_k), \Pi_h(v) - \Pi_h(z_k - z_{k-1})\rangle_{\mathcal{Z}_h^*,\mathcal{Z}_h} \quad \forall v \in \mathcal{Z}.\end{aligned} \tag{3.2.6}$$

Since $z_k, z_{k-1} \in \mathcal{Z}_h$, we have $\Pi_h(z_k - z_{k-1}) = z_k - z_{k-1}$ so that the nonexpansivity of the projection Π_h implies the equivalence of (3.2.6) and

$$\begin{aligned}\mathcal{R}_h(\Pi_h(v)) + I_\tau(v) \geq \mathcal{R}_h(\Pi_h(z_k - z_{k-1})) + I_\tau(z_k - z_{k-1})\\ + \langle -D_z\mathcal{I}(t_{k-1}, z_k), \Pi_h(v) - \Pi_h(z_k - z_{k-1})\rangle_{\mathcal{Z}_h^*,\mathcal{Z}_h} \quad \forall v \in \mathcal{Z}.\end{aligned} \tag{3.2.7}$$

Using the abbreviation $\mathcal{R}_{\tau,h} := \mathcal{R}_h \circ \Pi_h + I_\tau$, we have

$$\begin{aligned}(3.2.7) \quad &\Longleftrightarrow \quad \begin{cases}\mathcal{R}_{\tau,h}(v) \geq \mathcal{R}_{\tau,h}(z_k - z_{k-1})\\ \qquad + \langle -\Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k), v - (z_k - z_{k-1})\rangle_{\mathcal{Z}^*,\mathcal{Z}}\end{cases} \quad \forall v \in \mathcal{Z}\\ &\Longleftrightarrow \quad -\Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k) \in \partial\mathcal{R}_{\tau,h}(z_k - z_{k-1}) \subset \mathcal{Z}^*.\end{aligned} \tag{3.2.8}$$

Thus, an arbitrary solution z_k of (3.2.4) also satisfies the stationary condition

$$0 \in \partial\mathcal{R}_{\tau,h}(z_k - z_{k-1}) + \Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k). \tag{3.2.9}$$

This describes an equation in $\mathcal{Z}^*$ and not in $\mathcal{Z}_h^*$, which is the essential advantage in (3.2.9) compared to (3.2.4). Now, we can follow the lines of [MS19a]:

Thanks to a classical result of convex analysis (see Lemma A.3.3), (3.2.9) is equivalent to

$$\begin{aligned}\mathcal{R}_{\tau,h}(z_k - z_{k-1}) + \mathcal{R}^*_{\tau,h}(-\Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k))& \\ &= \langle -\Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k), z_k - z_{k-1}\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &= \langle -D_z\mathcal{I}(t_{k-1}, z_k), \Pi_h(z_k - z_{k-1})\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &= \langle -D_z\mathcal{I}(t_{k-1}, z_k), z_k - z_{k-1}\rangle_{\mathcal{Z}^*,\mathcal{Z}}.\end{aligned} \tag{3.2.10}$$

Again $\Pi_h(z_k - z_{k-1}) = z_k - z_{k-1}$ and the fact that $\|z_k - z_{k-1}\|_{\mathbb{V}} \leq \tau$ yield

$$\mathcal{R}_{\tau,h}(z_k - z_{k-1}) = \mathcal{R}_h(z_k - z_{k-1}). \tag{3.2.11}$$

Moreover, from Lemma A.3.8, we infer

$$\mathcal{R}^*_{\tau,h}(-\Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k)) = \tau\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k), \partial(\mathcal{R}_h \circ \Pi_h)(0)\}.$$

Inserting this together with (3.2.11) in (3.2.10) gives (3.2.3c).
To prove (3.2.3a), we consider (3.2.7) once more. Since $0 \in \mathrm{dom}(\mathcal{R}_h \circ \Pi_h) \cap \mathrm{dom}(I_\tau)$ and I_τ is continuous in 0, the sum rule for convex subdifferentials is applicable giving the existence of a $\zeta_k \in \partial I_\tau(z_k - z_{k-1})$ (note Lemma A.3.10), such that

$$0 \in \partial(\mathcal{R}_h \circ \Pi_h)(z_k - z_{k-1}) + \zeta_k + \Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k) \tag{3.2.12}$$

and thereby

$$\begin{aligned}\mathcal{R}_h(z_k - z_{k-1}) + (\mathcal{R}_h \circ \Pi_h)^*(-\zeta_k - \Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k))& \\ &= -\langle \zeta_k + \Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k), z_k - z_{k-1}\rangle_{\mathcal{V}^*,\mathcal{V}} \\ &= -\langle \zeta_k, z_k - z_{k-1}\rangle_{\mathcal{V}^*,\mathcal{V}} - \langle D_z\mathcal{I}(t_{k-1}, z_k), z_k - z_{k-1}\rangle_{\mathcal{Z}^*,\mathcal{Z}}.\end{aligned}$$

A comparison with (3.2.3c), shows that

$$\begin{aligned}&(\mathcal{R}_h \circ \Pi_h)^*(-\zeta_k - \Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k)) \\ &\quad = \tau\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} - \langle \zeta_k, z_k - z_{k-1}\rangle_{\mathcal{V}^*,\mathcal{V}}.\end{aligned} \tag{3.2.13}$$

Now, the fact that $\zeta_k \in \partial I_\tau(z_k - z_{k-1})$, and the characterization in Lemma A.3.11 imply the existence of a multiplier $\lambda_k \in \mathbb{R}$ with

$$\lambda_k \geq 0\,, \quad \zeta_k = \lambda_k \mathbb{V}(z_k - z_{k-1}), \quad \lambda_k(\|z_k - z_{k-1}\|_{\mathbb{V}} - \tau) = 0. \tag{3.2.14}$$

which is just (3.2.3a).
Next, we verify (3.2.3b). For this purpose, first observe that $\mathcal{J} := \mathcal{R}_h \circ \Pi_h$ is also convex and positively 1-homogeneous so that Lemma 2.1.1 implies $\partial(\mathcal{R}_h \circ \Pi_h)(z_k - z_{k-1}) \subset \partial(\mathcal{R}_h \circ \Pi_h)(0)$. The characterization of the conjugate functional from Lemma 2.1.1 in combination with (3.2.12)

thus yields

$$-\zeta_k - \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k) \in \partial(\mathcal{R}_h \circ \Pi_h)(z_k - z_{k-1}) \subset \partial(\mathcal{R}_h \circ \Pi_h)(0) \tag{3.2.15}$$

$$\implies \quad (\mathcal{R}_h \circ \Pi_h)^*(-\zeta_k - \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k)) = 0. \tag{3.2.16}$$

Inserting this and the second equation in (3.2.14) into (3.2.13) we arrive at (3.2.3b). Finally, (3.2.3d) is an immediate consequence of (3.2.15), i.e.,

$$\mathcal{R}_h(\Pi_h(v)) \geq -\langle \zeta_k + \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k), v \rangle_{\mathcal{Z}^*, \mathcal{Z}} \quad \forall v \in \mathcal{Z}$$

and the characterization of ζ_k in (3.2.14). □

Remark 3.2.3. In fact, since (3.2.4) and (3.2.9) are equivalent and (3.2.9) is in turn equivalent to the properties (3.2.3a)–(3.2.3d), we observe that $z_k^{\tau,h}$ is a stationary point for (alg$_1$) if and only if it satisfies (3.2.3a)–(3.2.3d). Hence, for the actual numerical realization of LISS, it might be practical to exploit the characterization via (3.2.3a)–(3.2.3d) instead of (3.2.4) in order to calculate a stationary point. Moreover, we will solely build upon this discrete optimality system (and the inequality (alg$_2$)) for the convergence analysis.

Let us take a further look at (3.2.12). Inserting the characterization of $\zeta_k^{\tau,h}$ from (3.2.14), we find

$$0 \in \partial(\mathcal{R}_h \circ \Pi_h)(z_k^{\tau,h} - z_{k-1}^{\tau,h}) + \lambda_k^{\tau,h} \mathbb{V}(z_k^{\tau,h} - z_{k-1}^{\tau,h}) + \Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h})\,. \tag{3.2.17}$$

Since $\lambda_k^{\tau,h} > 0$ only if the local stability $0 \in \partial(\mathcal{R}_h \circ \Pi_h)(z_k^{\tau,h} - z_{k-1}^{\tau,h}) + \Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h})$ is violated, we can interpret this inclusion as a discrete version of (2.4.8b). Moreover, due to the time update (alg$_3$) we have $t_k^{\tau,h} - t_{k-1}^{\tau,h} = \tau - \|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathbb{V}} \geq 0$ such that the complementarity-like conditions from (3.2.14) yield

$$\lambda_k^{\tau,h} \geq 0\,, \qquad \frac{t_k^{\tau,h} - t_{k-1}^{\tau,h}}{\tau} \geq 0, \qquad \lambda_k^{\tau,h} \left(\frac{t_k^{\tau,h} - t_{k-1}^{\tau,h}}{\tau} \right) = 0.$$

This can as well be seen as a discretization of the complementarity conditions in (2.4.8c). These observations will be taken up again in Section 3.2.3 in order to derive, similarly to the continuous case (see [MR15, MRS12]), a discrete version of the energy equality in (2.4.6). Finally, we note that combining (3.2.3a) and (3.2.3b) allows to characterize λ_k as

$$\lambda_k = \frac{1}{\tau} \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-\Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}), \partial(\mathcal{R}_h \circ \Pi_h)(0)\}. \tag{3.2.18}$$

3.2.2 A priori estimates

Based on the previous Lemma 3.2.2, we subsequently provide several a priori estimates that will allow a passage to the limit in the discrete energy identity in Section 3.2.3 and 3.2.4, respectively. Furthermore, we show that the discrete physical time $t_k^{\tau,h}$ given the time update in (alg$_3$) reaches the final time T in a finite number of iterations, see Proposition 3.2.12 below. We start with the following result, whose proof is actually the only point, where one uses that $z_k^{\tau,h}$ is energetically preferred, that means (alg_2) holds, and not only a stationary point satisfying (3.2.3a)–(3.2.3d).

Lemma 3.2.4 (Local energy inequality)**.** *For all $h, \tau > 0$ and all $k \in \mathbb{N}$, it holds*

$$\mathcal{I}(t_k^{\tau,h}, z_k^{\tau,h}) + \mathcal{R}_h(z_k^{\tau,h} - z_{k-1}^{\tau,h}) \leq \mathcal{I}(t_{k-1}^{\tau,h}, z_{k-1}^{\tau,h}) + \int_{t_{k-1}^{\tau,h}}^{t_k^{\tau,h}} \partial_t \mathcal{I}(s, z_k^{\tau,h})\, \mathrm{d}s\,. \tag{3.2.19}$$

Proof. Adding the term $\mathcal{I}(t_k^{\tau,h}, z_k^{\tau,h})$ to both sides of the energy inequality $\mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}) + \mathcal{R}(z_k^{\tau,h} - z_{k-1}^{\tau,h}) \leq \mathcal{I}(t_{k-1}^{\tau,h}, z_{k-1}^{\tau,h})$ from (alg_2) and using the continuous differentiability of $f(\cdot, z_k^{\tau,h})$, we find

$$\begin{aligned}\mathcal{I}(t_k^{\tau,h}, z_k^{\tau,h}) + \mathcal{R}_h(z_k^{\tau,h} - z_{k-1}^{\tau,h}) &\leq \mathcal{I}(t_{k-1}^{\tau,h}, z_{k-1}^{\tau,h}) + \mathcal{I}(t_k^{\tau,h}, z_k^{\tau,h}) - \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}) \\ &= \mathcal{I}(t_{k-1}^{\tau,h}, z_{k-1}^{\tau,h}) + \int_{t_{k-1}^{\tau,h}}^{t_k^{\tau,h}} \partial_t \mathcal{I}(s, z_k^{\tau,h})\, \mathrm{d}s\,,\end{aligned} \tag{3.2.20}$$

which gives the assertion. □

More or less as a direct consequence of the prior lemma and the assumptions on $\mathcal{I}$, in particular the estimate (3.1.4), we obtain the following.

Lemma 3.2.5 (Boundedness for energy and dissipation)**.** *For all $h, \tau > 0$ and all $k \in \mathbb{N}$, it holds*

$$\mathcal{I}(t_k^{\tau,h}, z_k^{\tau,h}) + \sum_{i=1}^{k} \mathcal{R}_h(z_i^{\tau,h} - z_{i-1}^{\tau,h}) \leq (\beta + \mathcal{I}(0, z_0^{\tau,h})) \exp(\mu T), \tag{3.2.21}$$

where β and μ are the constants from Section 3.1.

Proof. For the ease of clarity we suppress the superscripts τ, h, except for $z_0^{\tau,h}$ in order to avoid confusion with the initial value. We start by employing (3.1.4) into (3.2.19) to estimate

$$\begin{aligned}&\mathcal{I}(t_k, z_k) + \mathcal{R}_h(z_k - z_{k-1}) \\ &\quad\leq \mathcal{I}(t_{k-1}, z_{k-1}) + \int_{t_{k-1}}^{t_k} \mu(\mathcal{I}(t_{k-1}, z_k) + \beta) \exp(\mu\,(s - t_{k-1}))\, \mathrm{d}s \\ &\quad= \mathcal{I}(t_{k-1}, z_{k-1}) + (\mathcal{I}(t_{k-1}, z_k) + \beta)(\exp(\mu\,(t_k - t_{k-1})) - 1).\end{aligned}$$

From the nonnegativity of $\mathcal{R}_h$ by assumption (b) in combination with (alg_2) we find $\mathcal{I}(t_{k-1}, z_k) \leq \mathcal{I}(t_{k-1}, z_{k-1})$ so that

$$\begin{aligned}&\mathcal{I}(t_k, z_k) + \mathcal{R}_h(z_k - z_{k-1}) \\ &\quad\leq \mathcal{I}(t_{k-1}, z_{k-1}) + (\mathcal{I}(t_{k-1}, z_{k-1}) + \beta)(\exp(\mu\,(t_k - t_{k-1})) - 1)\end{aligned}. \tag{3.2.22}$$

holds. By exploiting once again $\mathcal{R}_h \geq 0$, this implies

$$\mathcal{I}(t_k, z_k) \leq (\mathcal{I}(t_{k-1}, z_{k-1}) + \beta) \exp(\mu\,(t_k - t_{k-1})) - \beta$$

such that induction over k already gives the desired result for the energy:

$$\begin{aligned}\mathcal{I}(t_k, z_k) &\leq (\mathcal{I}(0, z_0^{\tau,h}) + \beta) \prod_{j=1}^{k} \exp(\mu\,(t_j - t_{j-1})) - \beta \\ &\leq (\mathcal{I}(0, z_0^{\tau,h}) + \beta) \exp(\mu\, t_k) - \beta.\end{aligned} \tag{3.2.23}$$

To include the dissipation in the estimate, we sum up (3.2.22) to obtain

$$\mathcal{I}(t_k, z_k) + \sum_{j=1}^{k} \mathcal{R}_h(z_j - z_{j-1}) \leq \mathcal{I}(0, z_0^{\tau,h}) + \sum_{j=1}^{k} (\mathcal{I}(t_{j-1}, z_{j-1}) + \beta)(\exp(\mu(t_j - t_{j-1})) - 1).$$

Inserting (3.2.23) and adding β on both sides, we finally obtain

$$\begin{aligned}&\mathcal{I}(t_k, z_k) + \sum_{j=1}^{k} \mathcal{R}_h(z_j - z_{j-1}) + \beta \\ &\qquad \leq (\mathcal{I}(0, z_0^{\tau,h}) + \beta) + \sum_{j=1}^{k} (\mathcal{I}(0, z_0^{\tau,h}) + \beta) \exp(\mu t_{j-1})(\exp(\mu\,(t_j - t_{j-1})) - 1) \\ &\qquad\qquad = (\mathcal{I}(0, z_0^{\tau,h}) + \beta) \exp(\mu t_k) \leq (\mathcal{I}(0, z_0^{\tau,h}) + \beta) \exp(\mu\, T),\end{aligned}$$

which is the claimed estimate. □

We therefore see that the energy is bounded along the iterates $z_k^{\tau,h}$. Since by the estimate in (3.1.2) the energy is coercive, we infer that the iterates itself are uniformly bounded.

Lemma 3.2.6 (Uniform a priori estimate for iterates)**.** *The iterates of Algorithm LISS fulfill*

$$\sup_{h,\tau>0,\, k\in\mathbb{N}} \|z_k^{\tau,h}\|_{\mathcal{Z}} < \infty. \tag{3.2.24}$$

Proof. The lower bound on $\mathcal{I}$ from (3.1.2) implies for every $t \in [0, T]$ and every $z \in \mathcal{Z}$ that

$$\|z\|_{\mathcal{Z}} \leq \mathcal{I}(t, z) + c_0\,.$$

Combining this with (3.2.21) and using $\mathcal{R}_h \geq 0$, we arrive at

$$\begin{aligned}\|z_k^{\tau,h}\|_{\mathcal{Z}} &\leq c_0 + \mathcal{I}(t_k^{\tau,h}, z_k^{\tau,h}) \\ &\leq c_0 + (\mathcal{I}(0, z_0^{\tau,h}) + \beta) \exp(\mu\, T).\end{aligned}$$

Due to (3.1.7) and the continuity of $\mathcal{I}$ by assumption, $\mathcal{I}(0, z_0^{\tau,h})$ converges to $\mathcal{I}(0, z_0)$ and is thus bounded, which gives the assertion. □

Remark 3.2.7. As a consequence of Lemma 3.2.6 we have that $z_k^{\tau,h} \in B_{\mathcal{Z}}(0, R)$ for some $R > 0$ independent of τ and h.

The estimate (3.2.24) will, on the one hand, provide us with a uniform L^∞-bound for the linear interpolants and, on the other hand, allows us to obtain a bound for the nonlinearity $D_z\mathcal{F}$ as the following Lemma 3.2.8 reveals.

Lemma 3.2.8. *For every $r > 0$ and $\epsilon > 0$, there exists $C_{r,\epsilon} > 0$, independent of h, such that*

$$|\langle D_z\mathcal{F}(z_1^h) - D_z\mathcal{F}(z_2^h), z_1^h - z_2^h\rangle_{\mathcal{V}^*,\mathcal{V}}| \leq \epsilon\|z_1^h - z_2^h\|_{\mathcal{Z}}^2 + C_{r,\epsilon}\mathcal{R}_h(z_1^h - z_2^h)\|z_1^h - z_2^h\|_{\mathbb{V}}$$

for all $z_1^h, z_2^h \in \mathcal{Z}_h \cap B_{\mathcal{Z}}(0,r)$.

Proof. The proof is analogous to the nondiscretized case in [Kne19], one just has to employ Assumption (b) on $\mathcal{R}_h$ from Section 3.1 at the end. For convenience of the reader, we explain the arguments in detail. According to Ehrling's lemma, see Lemma A.2.1, for every $\delta > 0$, there exists a constant C_δ (obviously independent of h) such that

$$\|z\|_{\mathcal{V}} \leq \delta\|z\|_{\mathcal{Z}} + C_\delta\|z\|_{\mathcal{X}} \quad \forall\, z \in \mathcal{Z}. \tag{3.2.25}$$

Now, let $z_1^h, z_2^h \in B_{\mathcal{Z}}(0,r) \cap \mathcal{Z}_h$ be arbitrary. Using the growth condition on $D_z^2\mathcal{F}$ in ($\mathcal{I}_{\mathcal{F}1}$) and the above inequality for $\delta = \epsilon/(2C(1+r^q))$ together with Young's inequality gives

$$\begin{aligned}
&|\langle D_z\mathcal{F}(z_1^h) - D_z\mathcal{F}(z_2^h), z_1^h - z_2^h\rangle_{\mathcal{V}^*,\mathcal{V}}| \\
&\qquad\leq \|D_z\mathcal{F}(z_1^h) - D_z\mathcal{F}(z_2^h)\|_{\mathcal{V}^*}\|z_1^h - z_2^h\|_{\mathcal{V}} \\
&\qquad\leq C(1+r^q)\|z_1^h - z_2^h\|_{\mathcal{Z}}(\delta\|z_1^h - z_2^h\|_{\mathcal{Z}} + C_\delta\|z_1^h - z_2^h\|_{\mathcal{X}}) \\
&\qquad\leq \epsilon\|z_1^h - z_2^h\|_{\mathcal{Z}}^2 + \tilde{C}_{r,\epsilon}\|z_1^h - z_2^h\|_{\mathcal{X}}^2
\end{aligned}$$

with a constant $\tilde{C}_{r,\varepsilon}$ depending only on ε and r. Finally, (R3) and assumption (b) on the discretization of $\mathcal{R}$ result in

$$\|z_1^h - z_2^h\|_{\mathcal{X}} \leq \frac{1}{\rho}\mathcal{R}(z_1^h - z_2^h) \leq \frac{1}{\rho}\mathcal{R}_h(z_1^h - z_2^h), \tag{3.2.26}$$

which, together with the embedding $\mathcal{V} \hookrightarrow \mathcal{X}$, completes the proof. □

Remark 3.2.9. In fact, combining Lemma 3.2.8 and Lemma 3.2.6, precisely Remark 3.2.7, we find that

$$\begin{aligned}
&|\langle D_z\mathcal{F}(z_k^{\tau,h}) - D_z\mathcal{F}(z_{k-1}^{\tau,h}), z_k^{\tau,h} - z_{k-1}^{\tau,h}\rangle_{\mathcal{V}^*,\mathcal{V}}| \\
&\qquad\qquad\leq \delta\|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathcal{Z}}^2 + C_\delta\mathcal{R}_h(z_k^{\tau,h} - z_{k-1}^{\tau,h})\|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathbb{V}}
\end{aligned} \tag{3.2.27}$$

for all $k \in \mathbb{N}$ with C_δ only depending on the choice of δ, particularly independent of τ and h.

In the same way, we also find:

Lemma 3.2.10. *For every $\varepsilon > 0$ there exists $c_{\varepsilon,\nu} > 0$ independent of h, such that*

$$|\langle D_z f(t_1, z_1) - D_z f(t_2, z_2), v\rangle_{\mathcal{V}^*,\mathcal{V}}| \leq \nu|t_1 - t_2|\|v\|_{\mathcal{V}} + c_{\varepsilon,\nu}\mathcal{R}_h(z_1 - z_2)\|v\|_{\mathcal{V}} + \varepsilon\|z_1 - z_2\|_{\mathcal{Z}}\|v\|_{\mathcal{V}}$$

for all $t_1, t_2 \in [0,T]$, $z_1, z_2 \in \mathcal{Z}_h$ and $v \in \mathcal{V}$.

Proof. Again, by Ehrling's lemma A.2.1, we obtain for every $\delta > 0$ a constant C_δ (obviously independent of h) such that $\|z\|_{\mathcal{W}} \le \delta\|z\|_{\mathcal{Z}} + C_\delta\|z\|_{\mathcal{X}}$ for all $z \in \mathcal{Z}$. Taking herein $\delta = \frac{\varepsilon}{\nu}$ and combining the resulting estimate with ($\mathcal{I}_{f2}$) and the lower bound on $\mathcal{R}_h$ in (3.2.26), we arrive at

$$\begin{aligned}\langle D_z f(t_1, z_1) - D_z f(t_2, z_2), v\rangle_{\mathcal{V}^*,\mathcal{V}} &\le \nu(|t_1 - t_2| + \|z_1 - z_2\|_{\mathcal{W}})\|v\|_{\mathcal{V}} \\ &\le \nu(|t_1 - t_2| + C_{\alpha,\nu}\|z_1 - z_2\|_{\mathcal{X}} + \frac{\varepsilon}{\nu}\|z_1 - z_2\|_{\mathcal{Z}})\|v\|_{\mathcal{V}} \\ &\le \nu|t_1 - t_2|\|v\|_{\mathcal{V}} + c_{\alpha,\nu}\mathcal{R}_h(z_1 - z_2)\|v\|_{\mathcal{V}} + \varepsilon\|z_1 - z_2\|_{\mathcal{Z}}\|v\|_{\mathcal{V}}\end{aligned}$$

which holds for all $t_1, t_2 \in [0, T]$, $z_1, z_2 \in \mathcal{Z}_h$ and $v \in \mathcal{V}$. □

As a last preparatory lemma, before we turn to the essential a priori estimates that will allow a passage to the limit, we prove a weak convergence result for $\Pi_h^* D_z\mathcal{I}(0, z_0^{\tau,h})$ in $\mathcal{V}^*$, which then also provides us with a uniform bound of the same term.

Lemma 3.2.11. *Let $z_0 \in \mathcal{Z}$ be such that $Az_0 \in \mathcal{V}^*$. Then it holds*

$$\Pi_h^* D_z\mathcal{I}(0, z_0^{\tau,h}) \rightharpoonup D_z\mathcal{I}(0, z_0) \quad \text{in } \mathcal{V}^*, \text{ as } h \searrow 0\,.$$

Proof. First of all, we note that, by construction, we have

$$z_0^{\tau,h} = P_h(z_0) \to z_0 \quad \text{ in } \mathcal{Z}. \tag{3.2.28}$$

Furthermore, the energy functional is continuously differentiable in $\mathcal{Z}$ with

$$D_z\mathcal{I}(0, z_0^{\tau,h}) = Az_0^{\tau,h} + D_z\mathcal{F}(z_0^{\tau,h}) - D_z f(0, z_0^{\tau,h}) \in \mathcal{Z}^*.$$

We consider each term separately. For the nonlinear part, we exploit (3.2.28) and $D_z\mathcal{F} \in C(\mathcal{Z}; \mathcal{V}^*)$, cf. ($\mathcal{I}_{\mathcal{F}1}$), so that, for every $v \in \mathcal{V}$,

$$\begin{aligned}\langle \Pi_h^*(D_z\mathcal{F}(z_0^{\tau,h})), v\rangle_{\mathcal{V}^*,\mathcal{V}} &= \langle D_z\mathcal{F}(z_0^{\tau,h}), \Pi_h(v)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &= \langle D_z\mathcal{F}(z_0^{\tau,h}), \Pi_h(v)\rangle_{\mathcal{V}^*,\mathcal{V}} \to \langle D_z\mathcal{F}(z_0), v\rangle_{\mathcal{V}^*,\mathcal{V}},\end{aligned} \tag{3.2.29}$$

where we also used that $\Pi_h(v) \to v$ in $\mathcal{V}$ by (3.1.7). Moreover, the definition of the Ritz-projection and the fact that $Az_0 \in \mathcal{V}^*$ by assumption imply for every $v \in \mathcal{V}$ that

$$\begin{aligned}\langle \Pi_h^*(Az_0^{\tau,h}), v\rangle_{\mathcal{V}^*,\mathcal{V}} = \langle AP_h(z_0), \Pi_h v\rangle_{\mathcal{Z}^*,\mathcal{Z}} &= \langle Az_0, \Pi_h v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &= \langle Az_0, \Pi_h v\rangle_{\mathcal{V}^*,\mathcal{V}} \to \langle Az_0, v\rangle_{\mathcal{V}^*,\mathcal{V}}.\end{aligned}$$

Finally, for the term involving $D_z f(0, z_0^{\tau,h})$, we exploit the assumption in ($\mathcal{I}_{f2}$) and once again the convergence $\Pi_h(v) \to v$ in $\mathcal{V}$ to obtain

$$\begin{aligned}\langle \Pi_h^*(D_z f(0, z_0^{\tau,h})), v\rangle_{\mathcal{V}^*,\mathcal{V}} &= \langle D_z f(0, z_0^{\tau,h}), \Pi_h(v)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &= \langle D_z f(0, z_0^{\tau,h}), \Pi_h(v)\rangle_{\mathcal{V}^*,\mathcal{V}} \to \langle D_z f(0, z_0), v\rangle_{\mathcal{V}^*,\mathcal{V}}.\end{aligned}$$

Altogether, this yields the assertion. □

As indicated in the introduction, one major issue in the convergence analysis for parametrized solutions concerns the boundedness of the artificial time, even in the continuous setting, see also the discussion in [MR15, p. 218]. For the discrete counterpart, the artificial time reads $s_n = \sum_{k=1}^n t_k - t_{k-1} + \|z_k - z_{k-1}\|_{\mathbb{V}}$. In order to bound this term, we need to estimate $\sum_{k=1}^n \|z_k - z_{k-1}\|_{\mathbb{V}}$, which is purpose of the next proposition. Note that we will prove this boundedness even in the stronger $\mathcal{Z}$-norm, which might raise the question, whether the assumptions made are too restrictive. However, this is a well-known problem in the context of parametrized solutions, see, e.g., [KRZ13, Mie11]. Moreover, we will show that the physical end time T is reached after a finite number of iterations, which guarantees that the algorithm finishes in a finite number of steps.

Proposition 3.2.12 (Bound on artificial time)**.** *For every parameter $h, \tau > 0$ there exists an index $N(\tau, h) \in \mathbb{N}$ such that $t^{\tau,h}_{N(\tau,h)} \geq T$. Moreover, there are constants $C_1, C_2, C_3 > 0$ independent of τ, h such that, for all $h, \tau > 0$, it holds*

$$\sum_{i=1}^{N(\tau,h)} \|z_i^{\tau,h} - z_{i-1}^{\tau,h}\|_{\mathbb{V}} \leq C_1, \tag{3.2.30}$$

$$\sum_{i=1}^{N(\tau,h)} \|z_i^{\tau,h} - z_{i-1}^{\tau,h}\|_{\mathcal{Z}}^2 \leq C_2\, \tau, \tag{3.2.31}$$

and $$\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-\Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} \leq C_3 \quad \forall\, k = 0, ..., N(\tau, h). \tag{3.2.32}$$

Proof. The arguments are similar to [Kne19]. However, we additionally provide the estimate (3.2.31) and we also have to take account of the discretization at several points. Therefore, we present the arguments in detail. Let $k \in \mathbb{N}$ be arbitrary. For convenience, we again suppress the superscript τ, h throughout the proof, except for $z_0^{\tau,h}$ in order to avoid confusion with the initial data. We start by testing (3.2.3d) with $v = z_{k+1} - z_k$ to obtain

$$\begin{aligned} \mathcal{R}_h(z_{k+1} - z_k) &\geq -\langle \lambda_k \mathbb{V}(z_k - z_{k-1}) + \Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &= -\langle \lambda_k \mathbb{V}(z_k - z_{k-1}), z_{k+1} - z_k\rangle_{\mathcal{V}^*,\mathcal{V}} - \langle D_z\mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}}. \end{aligned} \tag{3.2.33}$$

Inserting (3.2.3b) into (3.2.3c) and rewriting this identity for the index $k+1$ (instead of k) gives $\mathcal{R}_h(z_{k+1} - z_k) + \lambda_{k+1}\|z_{k+1} - z_k\|_{\mathbb{V}}^2 = \langle -D_z\mathcal{I}(t_k, z_{k+1}), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}}$, so that (3.2.33) can be written as

$$\begin{aligned} 0 \geq \lambda_{k+1}\|z_{k+1} - z_k\|_{\mathbb{V}}^2 - \lambda_k\langle \mathbb{V}(z_k - z_{k-1}), z_{k+1} - z_k\rangle_{\mathcal{V}^*,\mathcal{V}} \\ + \langle D_z\mathcal{I}(t_k, z_{k+1}) - D_z\mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}}\,. \end{aligned}$$

With this inequality at hand, we can now follow the lines of [Kne19, Proposition 2.3]: On account of $D_z\mathcal{F}(\cdot) \in \mathcal{V}^*$ by assumption, inserting the definition of $\mathcal{I}$ into this inequality gives

$$\begin{aligned} &\langle D_z\mathcal{F}(z_k) - D_z\mathcal{F}(z_{k+1}), z_{k+1} - z_k\rangle_{\mathcal{V}^*,\mathcal{V}} - \langle D_z f(t_{k-1}, z_k) - D_z f(t_k, z_{k+1}), z_{k+1} - z_k\rangle_{\mathcal{V}^*,\mathcal{V}} \\ &\quad \geq \lambda_{k+1}\|z_{k+1} - z_k\|_{\mathbb{V}}^2 - \lambda_k\langle \mathbb{V}(z_k - z_{k-1}), z_{k+1} - z_k\rangle_{\mathcal{V}^*,\mathcal{V}} + \langle A(z_{k+1} - z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\quad \geq \lambda_{k+1}\|z_{k+1} - z_k\|_{\mathbb{V}}^2 - \lambda_k\|z_k - z_{k-1}\|_{\mathbb{V}}\|z_{k+1} - z_k\|_{\mathbb{V}} + \alpha\, \|z_{k+1} - z_k\|_{\mathcal{Z}}^2, \end{aligned} \tag{3.2.34}$$

where we used the coercivity of A for the last inequality. We now estimate each term on the left-hand side of (3.2.34) separately. For the first term, we apply Lemma 3.2.8, precisely (3.2.27) from Remark 3.2.9, with $\varepsilon = \alpha/4$ to get

$$\begin{aligned}\langle D_z\mathcal{F}(z_k) - D_z\mathcal{F}(z_{k+1}), z_{k+1} - z_k\rangle_{\mathcal{V}^*,\mathcal{V}} \\ \leq \frac{\alpha}{4}\|z_{k+1} - z_k\|_{\mathcal{Z}}^2 + C_\alpha\|z_{k+1} - z_k\|_{\mathbb{V}}\,\mathcal{R}_h(z_{k+1} - z_k),\end{aligned} \tag{3.2.35}$$

with a constant $C_\alpha > 0$, which is independent of τ, h, and k. Concerning the second term on the left-hand side of (3.2.34), we use Lemma 3.2.10 with $\varepsilon = \alpha/4$ to estimate

$$\begin{aligned}&\langle D_z f(t_{k-1}, z_k) - D_z f(t_k, z_{k+1}), z_{k+1} - z_k\rangle_{\mathcal{V}^*,\mathcal{V}} \\ &\quad\leq \nu((t_k - t_{k-1})\|z_{k+1} - z_k\|_{\mathcal{V}} + c_{\alpha,\nu}\mathcal{R}_h(z_{k+1} - z_k))\|z_{k+1} - z_k\|_{\mathcal{V}} \\ &\qquad\qquad + \frac{\alpha}{4}\|z_{k+1} - z_k\|_{\mathcal{Z}}\|z_{k+1} - z_k\|_{\mathcal{V}}.\end{aligned} \tag{3.2.36}$$

Hence, inserting (3.2.35) and (3.2.36) in (3.2.34) yields

$$\begin{aligned}&\lambda_{k+1}\|z_{k+1} - z_k\|_{\mathbb{V}}^2 - \lambda_k\|z_k - z_{k-1}\|_{\mathbb{V}}\|z_{k+1} - z_k\|_{\mathbb{V}} + \frac{\alpha}{2}\|z_{k+1} - z_k\|_{\mathcal{Z}}^2 \\ &\quad\leq C\|z_{k+1} - z_k\|_{\mathbb{V}}\,\mathcal{R}_h(z_{k+1} - z_k) \\ &\qquad + \nu((t_k - t_{k-1})\|z_{k+1} - z_k\|_{\mathcal{V}} + c_{\alpha,\nu}\mathcal{R}_h(z_{k+1} - z_k))\|z_{k+1} - z_k\|_{\mathcal{V}},\end{aligned} \tag{3.2.37}$$

which, thanks to the continuous embedding $\mathcal{Z} \hookrightarrow \mathcal{V}$ and the norm equivalence of $\|\cdot\|_{\mathcal{V}}$ and $\|\cdot\|_{\mathbb{V}}$, in turn implies

$$\lambda_{k+1}\|z_{k+1} - z_k\|_{\mathbb{V}} - \lambda_k\|z_k - z_{k-1}\|_{\mathbb{V}} + c\|z_{k+1} - z_k\|_{\mathcal{Z}} \leq C\left(\mathcal{R}_h(z_{k+1} - z_k) + (t_k - t_{k-1})\right).$$

Summing up this estimate with respect to k we thus have

$$\lambda_{k+1}\|z_{k+1} - z_k\|_{\mathbb{V}} + c\sum_{i=1}^{k}\|z_{i+1} - z_i\|_{\mathcal{Z}} \leq \lambda_1\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}} + C\Big(t_k + \sum_{i=1}^{k}\mathcal{R}_h(z_{i+1} - z_i)\Big). \tag{3.2.38}$$

On account of (3.2.21), this inequality already nearly gives (3.2.30), provided that the term $\lambda_1\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}$ is bounded independent of τ and h, which is shown next. To this end, we again insert (3.2.3b) into (3.2.3c) to obtain for $k = 1$:

$$\mathcal{R}_h(z_1 - z_0) + \lambda_1\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}^2 = \langle -D_z\mathcal{I}(0, z_1), z_1 - z_0^{\tau,h}\rangle_{\mathcal{Z}^*,\mathcal{Z}}\,.$$

Adding a zero, using $\mathcal{R}_h \geq 0$, and rearranging terms yields

$$\begin{aligned}&\langle D_z\mathcal{I}(0, z_1) - D_z\mathcal{I}(0, z_0^{\tau,h}), z_1 - z_0^{\tau,h}\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \lambda_1\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}^2 \\ &\quad\leq \langle -D_z\mathcal{I}(0, z_0^{\tau,h}), z_1 - z_0^{\tau,h}\rangle_{\mathcal{Z}^*,\mathcal{Z}} = \langle -\Pi_h^* D_z\mathcal{I}(0, z_0^{\tau,h}), z_1 - z_0^{\tau,h}\rangle_{\mathcal{V}^*,\mathcal{V}}\,.\end{aligned} \tag{3.2.39}$$

The first term on the left-hand side is treated completely analogous to above, that is, inserting the concrete form of $\mathcal{I}$, exploiting the coercivity of A and estimating as in (3.2.35) and (3.2.36), resulting in

$$\begin{aligned}\frac{\alpha}{2}&\|z_1 - z_0^{\tau,h}\|_{\mathcal{Z}}^2 + \lambda_1\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}^2 \\ &\leq C\,\mathcal{R}_h(z_1 - z_0^{\tau,h})\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}} + \langle -\Pi_h^* D_z\mathcal{I}(0, z_0^{\tau,h}), z_1 - z_0^{\tau,h}\rangle_{\mathcal{V}^*,\mathcal{V}}\,.\end{aligned} \tag{3.2.40}$$

Hence, using again $\mathcal{Z} \hookrightarrow \mathcal{V}$ and the norm equivalence of $\|\cdot\|_{\mathbb{V}}$ and $\|\cdot\|_{\mathcal{V}}$, we obtain

$$\lambda_1\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}} + c\|z_1 - z_0^{\tau,h}\|_{\mathcal{Z}} \leq C\left(\mathcal{R}_h(z_1 - z_0^{\tau,h}) + \|\Pi_h^* D_z\mathcal{I}(0, z_0^{\tau,h})\|_{\mathcal{V}^*}\right). \tag{3.2.41}$$

By adding (3.2.41) to (3.2.38) and applying (3.2.21), we arrive at

$$\begin{aligned}\lambda_{k+1}&\|z_{k+1} - z_k\|_{\mathbb{V}} + c\sum_{i=0}^{k}\|z_{i+1} - z_i\|_{\mathcal{Z}} \\ &\leq C\Big(t_k + \sum_{i=0}^{k}\mathcal{R}_h(z_{i+1} - z_i) + \|\Pi_h^* D_z\mathcal{I}(0, z_0^{\tau,h})\|_{\mathcal{V}^*}\Big) \\ &\leq C\Big(T + (\mathcal{I}(0, z_0^{\tau,h}) + \beta)\exp(\mu T) + \|\Pi_h^* D_z\mathcal{I}(0, z_0^{\tau,h})\|_{\mathcal{V}^*}\Big),\end{aligned} \tag{3.2.42}$$

where we used that $t_k^{\tau,h} \leq T$ by the time update in (alg$_3$) for the last estimate. On account of Lemma 3.2.11, we know that $\Pi_h^* D_z\mathcal{I}(0, z_0^{\tau,h})$ converges weakly in $\mathcal{V}^*$ and is thus bounded. Moreover, as already seen at the end of the proof of Lemma 3.2.6, $\mathcal{I}(0, z_0^{\tau,h})$ is bounded independent of h, which yields

$$T + (\mathcal{I}(0, z_0^{\tau,h}) + \beta)\exp(\mu T) + \|\Pi_h^* D_z\mathcal{I}(0, z_0^{\tau,h})\|_{\mathcal{V}^*} \leq C. \tag{3.2.43}$$

This in turn implies

$$\lambda_{k+1}\|z_{k+1} - z_k\|_{\mathbb{V}} + c\sum_{i=0}^{k}\|z_{i+1} - z_i\|_{\mathcal{Z}} \leq C, \tag{3.2.44}$$

i.e., (3.2.30) for $k \geq 0$. Note that the constant C is independent of τ, h, and k. Now, let us turn towards (3.2.31), whose proof is very similar to the steps above. To this end, we consider (3.2.37) once more. Thanks to the constraint $\|z_{k+1} - z_k\|_{\mathbb{V}} \leq \tau$ and (3.2.3a), we have

$$-\lambda_k\|z_{k+1} - z_k\|_{\mathbb{V}} \geq -\lambda_k\tau = -\lambda_k\|z_k - z_{k-1}\|_{\mathbb{V}},$$

which in turn implies

$$\lambda_{k+1}\|z_{k+1} - z_k\|_{\mathbb{V}}^2 - \lambda_k\|z_k - z_{k-1}\|_{\mathbb{V}}^2 + c\|z_{k+1} - z_k\|_{\mathcal{Z}}^2 \leq C\,\tau\,\left(\mathcal{R}_h(z_{k+1} - z_k) + (t_k - t_{k-1})\right).$$

Again, summing up this estimate with respect to k we find

$$\begin{aligned}\lambda_{k+1}&\|z_{k+1} - z_k\|_{\mathbb{V}}^2 + c\sum_{i=1}^{k}\|z_{i+1} - z_i\|_{\mathcal{Z}}^2 \\ &\leq \lambda_1\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}^2 + C\,\tau\Big(t_k + \sum_{i=1}^{k}\mathcal{R}_h(z_{i+1} - z_i)\Big).\end{aligned} \tag{3.2.45}$$

Combining this with (3.2.40) and exploiting $\|z_1 - z_0^{\tau,h}\|_{\mathcal{V}} \leq \tau$ gives exactly as in (3.2.42)

$$\begin{aligned}\lambda_{k+1}\|z_{k+1} - z_k\|_{\mathbb{V}}^2 + c\sum_{i=0}^{k}\|z_{i+1} - z_i\|_{\mathcal{Z}}^2 \\ &\leq C\,\tau\Big(t_k + \sum_{i=0}^{k}\mathcal{R}_h(z_{i+1} - z_i) + \|\Pi_h^* D_z\mathcal{I}(0, z_0^{\tau,h})\|_{\mathcal{V}^*}\Big) \\ &\leq C\,\tau\Big(T + (\mathcal{I}(0, z_0^{\tau,h}) + \beta)\exp(\mu T) + \|\Pi_h^* D_z\mathcal{I}(0, z_0^{\tau,h})\|_{\mathcal{V}^*}\Big).\end{aligned}$$

As seen in (3.2.43), the last term is bounded independent of τ and h, so that the nonnegativity of the multiplier λ_{k+1} yields

$$c\sum_{i=0}^{k}\|z_{i+1} - z_i\|_{\mathcal{Z}}^2 \leq C\,\tau \tag{3.2.46}$$

for all $k \geq 0$, which proves (3.2.31). We proceed with showing (3.2.32). For this purpose, we first note that, since $\|z_k - z_{k-1}\|_{\mathbb{V}} \leq \tau$ by the constraint in (alg_1), the identity (3.2.3b) implies

$$\overline{\text{dist}}_{\mathcal{V}^*}\{-\Pi_h^* D_z\mathcal{I}(t_{k-1}, z_k), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} \leq \lambda_k\|z_k - z_{k-1}\|_{\mathbb{V}}\,. \tag{3.2.47}$$

The estimate in (3.2.32) is thus an easy consequence of (3.2.44). Finally, we show that the final time T is reached after a finite number of steps. For this, we observe that by the embedding $\mathcal{Z} \hookrightarrow \mathcal{V}$ and the norm equivalence of $\|\cdot\|_{\mathcal{V}}$ and $\|\cdot\|_{\mathbb{V}}$ estimate (3.2.44) implies that $\sum_{k=1}^{\infty}\|z_k - z_{k-1}\|_{\mathbb{V}}$ is convergent, thus bounded. Summing up (alg_3) from $k = 1$ to n and exploiting (3.2.30) we therefore obtain

$$t_n = t_0 + n\tau - \sum_{k=1}^{n}\|z_k - z_{k-1}\|_{\mathbb{V}} \geq t_0 + n\tau - C \to \infty \quad \text{for } n \to \infty.$$

Hence, there must exist a finite index $N(\tau, h)$, possibly depending on τ and h, so that $t_{N(\tau,h)} \geq T$. Lastly, since (3.2.44) and (3.2.46) hold for every k, we obtain (3.2.30) and (3.2.31), respectively. □

In what follows we will abbreviate the index $N(\tau, h)$ simply by N having in mind that the number N of time steps always depends on τ and h. We close this section with a remark on the regularity assumption for the initial state z_0.

Remark 3.2.13. Revisiting the above proof, we may alternatively assume that the discrete initial state $z_0^{\tau,h}$ fulfills

$$-D_z\mathcal{I}(0, z_0^{\tau,h}) \in \partial\mathcal{R}_h(0) \quad \text{and} \quad z_0^{\tau,h} \to z_0 \quad (h \searrow 0),$$

and the dissipation satisfies

$$\mathcal{R}_h(v) \leq c\|v\|_{\mathcal{V}} \quad \forall v \in \text{dom}(\mathcal{R}_h) \quad \text{for all } h \geq 0$$

for some $c > 0$ independent of h. Moreover, in this case, we may also relax the assumption in ($\mathcal{I}_{f2}$) to

$$|\langle D_z f(t_1, z_1) - D_z f(t_2, z_2), v\rangle_{\mathcal{Z}^*,\mathcal{Z}}| \leq \nu(|t_1 - t_2| + \|z_1 - z_2\|_{\mathcal{W}})\|v\|_{\mathcal{Z}}\,. \tag{$\mathcal{I}_{f2}'$}$$

Indeed, in this case, the above assumptions guarantee that

$$\langle -D_z\mathcal{I}(0,z_0^{\tau,h}), z_1 - z_0^{\tau,h}\rangle_{\mathcal{Z}^*,\mathcal{Z}} \le \mathcal{R}_h(z_1 - z_0^{\tau,h}) \le c\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}.$$

Hence, from (3.2.39) and (3.2.40) it follows

$$\frac{\alpha}{2}\|z_1 - z_0^{\tau,h}\|_{\mathcal{Z}}^2 + \lambda_1\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}^2 \le C\,\mathcal{R}_h(z_1 - z_0^{\tau,h})\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}} + c\|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}$$

so that (3.2.44) and therewith Lemma 3.2.12 remain valid. This is in particular noteworthy in terms of generalizations for the structure of the energy functional, e.g., quasilinear instead of semilinear, see also the note after the proof of Theorem 3.2.19.

3.2.3 Discrete energy identity

In the following section, we aim at deriving a discrete analogon to the energy identity (2.4.6). To this end, we introduce the piecewise affine as well as the left- and right-continuous piecewise constant interpolants associated with the iterates $z_k^{\tau,h}$. As depicted in the context of parametrized solutions in Section 2.4, potential discontinuities of the solution are resolved by introducing an artificial time. The physical time is accordingly interpreted as a function of the very same and jumps are characterized by the plateaus of this function. This is also reflected by the time-incremental stationarity scheme (LISS), where, loosely speaking, the artificial time is divided into equidistant subintervals with step size τ and the approximation of the parametrized solution is implicitly defined through the optimization in (LISS). To be more precise, we set $s_k^{\tau,h} := k\tau$, so that

$$\begin{aligned} s_N^{\tau,h} = N\tau &= \sum_{i=1}^{N}(t_i^{\tau,h} - t_{i-1}^{\tau,h} + \|z_i^{\tau,h} - z_{i-1}^{\tau,h}\|_{\mathbb{V}}) \\ &= t_N^{\tau,h} + \sum_{i=1}^{N}\|z_i^{\tau,h} - z_{i-1}^{\tau,h}\|_{\mathbb{V}} \le T + \tau + \sum_{i=1}^{N}\|z_i^{\tau,h} - z_{i-1}^{\tau,h}\|_{\mathbb{V}} \le C_S \end{aligned} \tag{3.2.48}$$

by Proposition 3.2.12 with a constant $C_S > 0$ which is neither depending on τ nor h so that the artificial time interval is indeed bounded independent of the discretization. Hence, we can proceed with the construction of the interpolants. For $s \in [s_{k-1}^{\tau,h}, s_k^{\tau,h}) \subset [0, s_N^{\tau,h})$, the continuous and piecewise affine interpolants are defined through

$$\begin{aligned} \hat{z}_{\tau,h}(s) &:= z_{k-1}^{\tau,h} + \frac{(s - s_{k-1}^{\tau,h})}{\tau}(z_k^{\tau,h} - z_{k-1}^{\tau,h}), \\ \hat{t}_{\tau,h}(s) &:= t_{k-1}^{\tau,h} + \frac{(s - s_{k-1}^{\tau,h})}{\tau}(t_k^{\tau,h} - t_{k-1}^{\tau,h}), \end{aligned} \tag{3.2.49}$$

while the piecewise constant interpolants are given by

$$\overline{z}_{\tau,h}(s) := z_k^{\tau,h}, \qquad \overline{t}_{\tau,h}(s) := t_k^{\tau,h}, \qquad \underline{z}_{\tau,h}(s) := z_{k-1}^{\tau,h}, \qquad \underline{t}_{\tau,h}(s) := t_{k-1}^{\tau,h}. \tag{3.2.50}$$

Moreover, we define the artificial end time $S_{\tau,h}$ as that point where $\hat{t}$ reaches the end time T, i.e., it holds (see also Figure 3.2.1)

$$\hat{t}_{\tau,h}(S_{\tau,h}) = T, \quad s_{N-1}^{\tau,h} < S_{\tau,h} \le s_N^{\tau,h} \quad \text{and} \quad S_{\tau,h} \le C_S, \tag{3.2.51}$$

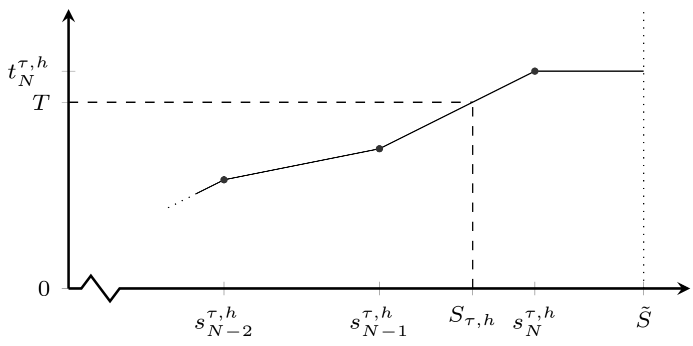


Figure 3.2.1: Qualitative illustration of the affine interpolant $\hat{t}$, the choice of the artificial end time $S_{\tau,h}$ via the equality $\hat{t}(S_{\tau,h}) = T$ and the upper bound $\tilde{S}$.

whereby the boundedness follows directly from (3.2.48). Since the artificial end time $S_{\tau,h}$ depends on the chosen discretization level, we extend all interpolants constantly onto $[0, \tilde{S}]$ with $\tilde{S} := \sup_{\tau,h} S_{\tau,h}$ where this is necessary, i.e., where $s_N^{\tau,h} < \tilde{S}$. Hence, we let

$$\left.\begin{aligned} \overline{z}_{\tau,h}(s) = \underline{z}_{\tau,h}(s) = \hat{z}_{\tau,h}(s) := z_N^{\tau,h} \\ \text{and} \quad \overline{t}_{\tau,h}(s) = \underline{t}_{\tau,h}(s) = \hat{t}_{\tau,h}(s) := T \end{aligned}\right\} \quad \forall\, s \in [s_N^{\tau,h}, \tilde{S}]\,. \tag{3.2.52}$$

Observe that still $\tilde{S} \leq C_S$ by (3.2.51). Moreover, due to the time update in (alg$_3$), we clearly have that $(\hat{t}_{\tau,h}, \hat{z}_{\tau,h}) \in W^{1,\infty}(0, \tilde{S}; \mathbb{R}) \times W^{1,\infty}(0, \tilde{S}; \mathcal{V})$, but we even obtain the following pointwise properties.

Lemma 3.2.14 (Properties of affine interpolants)**.** *For almost all $s \in [0, S_{\tau,h}]$, the affine interpolants from* (3.2.49) *fulfill*

$$\hat{t}'_{\tau,h}(s) \geq 0, \qquad \hat{t}'_{\tau,h}(s) + \|\hat{z}'_{\tau,h}(s)\|_{\mathbb{V}} = 1, \tag{3.2.53}$$

$$\hat{t}'_{\tau,h}(s)\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\Pi_h^* D_z \mathcal{I}(\underline{t}_{\tau,h}(s), \overline{z}_{\tau,h}(s)), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} = 0. \tag{3.2.54}$$

Proof. The first statement in (3.2.53) is a direct consequence of the constraint in (alg$_1$) and the time update in (alg$_3$), which immediately implies $t_k^{\tau,h} - t_{k-1}^{\tau,h} \geq 0$. To prove the second one, we again exploit (alg$_3$) to obtain for every $s \in [s_{k-1}^{\tau,h}, s_k^{\tau,h})$ that

$$\hat{t}'_{\tau,h}(s) + \|\hat{z}'_{\tau,h}(s)\| = \frac{(t_k^{\tau,h} - t_{k-1}^{\tau,h})}{\tau} + \frac{\|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathbb{V}}}{\tau} = 1,$$

which gives (3.2.53). Finally the complementarity in (3.2.54) is a direct consequence of (3.2.3a), since

$$0 = \lambda_k^{\tau,h}(\tau - \|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathbb{V}}) = \tau\, \lambda_k^{\tau,h}(1 - \|\hat{z}'_{\tau,h}(s)\|_{\mathbb{V}})$$

for $s \in [s_{k-1}^{\tau,h}, s_k^{\tau,h})$. Thus, inserting (3.2.18) and exploiting the identity in (3.2.53) yields (3.2.54). □

Once more, we note the similarity between the continuous case in (2.4.5a) and (2.4.5b) and its discrete version in Lemma 3.2.14. In the subsequent, last preparatory lemma, we collect the main a priori bounds of our interpolants, which will be essential to pass to the limit in the discrete energy identity, which is elaborated afterwards.

Lemma 3.2.15. *There exists $C > 0$, independent of τ and h, so that*

$$\|\hat{t}_{\tau,h}\|_{W^{1,\infty}(0,\tilde{S})}, \|\hat{z}_{\tau,h}\|_{W^{1,\infty}(0,\tilde{S};\mathcal{V})}, \|\hat{z}_{\tau,h}\|_{L^\infty(0,\tilde{S};\mathcal{Z})}, \|\hat{z}_{\tau,h}\|_{H^1(0,\tilde{S};\mathcal{Z})} \leq C.$$

Proof. While the first three bounds are an immediate consequence of the results in Lemma 3.2.14 and Lemma 3.2.6, the last one requires some slighlty more explanation. Due to the bound in $L^\infty(0,\tilde{S};\mathcal{Z})$, it suffices to estimate the $L^2(0,\tilde{S};\mathcal{Z})$-norm of the time-derivative $\hat{z}'_{\tau,h}$. Hence, inserting the definition of $\hat{z}$ from (3.2.52) and keeping in mind that $S_{\tau,h} \leq s_N^{\tau,h}$, we have

$$\begin{aligned}\|\hat{z}'_{\tau,h}\|_{L^2(0,\tilde{S};\mathcal{Z})} = \int_0^{\tilde{S}} \|\hat{z}'_{\tau,h}(r)\|^2_{\mathcal{Z}}\,\mathrm{d}r &= \int_0^{S_{\tau,h}} \|\hat{z}'_{\tau,h}(r)\|^2_{\mathcal{Z}}\,\mathrm{d}r \\ &\leq \sum_{k=1}^N \int_{s_{k-1}^{\tau,h}}^{s_k^{\tau,h}} \left\|\frac{z_k^{\tau,h} - z_{k-1}^{\tau,h}}{\tau}\right\|^2_{\mathcal{Z}} \mathrm{d}r = \frac{1}{\tau}\sum_{k=1}^N \|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|^2_{\mathcal{Z}}.\end{aligned}$$

Lemma 3.2.12, precisely (3.2.31), thus implies that this term is bounded independent of τ and h, which proves the desired $H^1(0,\tilde{S};\mathcal{Z})$ estimate. □

Eventually, we are now in the position to show a discrete version of the energy equality. Its proof is based on Lemma 3.2.14, the a priori estimates derived in Section 3.2.2 and assumption (a) on the discretization of $\mathcal{R}$, which essentially ensures that $\mathcal{R}_h$ has the same properties as $\mathcal{R}$.

Lemma 3.2.16 (Discrete energy equality)**.** *For all $s \in [0, S_{\tau,h}]$, it holds*

$$\begin{aligned}&\mathcal{I}(\hat{t}_{\tau,h}(s), \hat{z}_{\tau,h}(s)) \\ &\quad + \int_0^s \mathcal{R}_h(\hat{z}'_{\tau,h}(\sigma)) + \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-\Pi_h^* D_z\mathcal{I}(\underline{t}_{\tau,h}(\sigma), \overline{z}_{\tau,h}(\sigma)), \partial(\mathcal{R}_h \circ \Pi_h)(0)\}\,\mathrm{d}\sigma \\ &= \mathcal{I}(\hat{t}_{\tau,h}(0), \hat{z}_{\tau,h}(0)) \\ &\quad + \int_0^s \partial_t\mathcal{I}(\hat{t}_{\tau,h}(\sigma), \hat{z}_{\tau,h}(\sigma))\,\hat{t}'_{\tau,h}(\sigma)\,\mathrm{d}\sigma + \int_0^s r_{\tau,h}(\sigma)\,\mathrm{d}\sigma\,,\end{aligned} \tag{3.2.55}$$

where

$$r_{\tau,h}(s) := \langle D_z\mathcal{I}(\hat{t}_{\tau,h}(s), \hat{z}_{\tau,h}(s)) - D_z\mathcal{I}(\underline{t}_{\tau,h}(s), \overline{z}_{\tau,h}(s)), \hat{z}'_{\tau,h}(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\,. \tag{3.2.56}$$

Moreover, the complementarity condition

$$\hat{t}'_{\tau,h}(s)\,\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-\Pi_h^* D_z\mathcal{I}(\underline{t}_{\tau,h}(s), \overline{z}_{\tau,h}(s)), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} = 0 \tag{3.2.57}$$

is fulfilled f.a.a. $s \in (0, S_{\tau,h})$, and there exists a constant $C > 0$ such that the remainder $r_{\tau,h}$ satisfies for all $h, \tau > 0$ and all $s \in [0, S_{\tau,h}]$

$$\int_0^s r_{\tau,h}(\sigma)\,\mathrm{d}\sigma \leq C\tau. \tag{3.2.58}$$

Proof. The complementarity in (3.2.57) has already been proven in Lemma 3.2.14. Hence, we turn to the discrete energy identity. Since the affine interpolants in (3.2.49) are by construction elements of $W^{1,\infty}(0,S_{\tau,h})$ and $W^{1,\infty}(0,S_{\tau,h};\mathcal{Z})$, respectively, and due to $\mathcal{I}\in C^1([0,T]\times\mathcal{Z})$ by assumption, the chain rule is applicable and gives for $s\in(s_{k-1}^{\tau,h},s_k^{\tau,h})$ that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{I}(\hat{t}_{\tau,h}(s),\hat{z}_{\tau,h}(s))\\
&\quad=\partial_t\mathcal{I}(\hat{t}_{\tau,h}(s),\hat{z}_{\tau,h}(s))\,\hat{t}'_{\tau,h}(s)+\langle D_z\mathcal{I}(\hat{t}_{\tau,h}(s),\hat{z}_{\tau,h}(s)),\hat{z}'_{\tau,h}(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\\
&\quad=\partial_t\mathcal{I}(\hat{t}_{\tau,h}(s),\hat{z}_{\tau,h}(s))\,\hat{t}'_{\tau,h}(s)+\frac{1}{\tau}\langle D_z\mathcal{I}(\underline{t}_{\tau,h}(s),\overline{z}_{\tau,h}(s)),z_k^{\tau,h}-z_{k-1}^{\tau,h}\rangle_{\mathcal{Z}^*,\mathcal{Z}}\\
&\qquad+\langle D_z\mathcal{I}(\hat{t}_{\tau,h}(s),\hat{z}_{\tau,h}(s))-D_z\mathcal{I}(\underline{t}_{\tau,h}(s),\overline{z}_{\tau,h}(s)),\hat{z}'_{\tau,h}(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\,.
\end{aligned}$$

From (3.2.3c), we have in combination with the 1-homogeneity of $\mathcal{R}_h$ that

$$\begin{aligned}
&-\frac{1}{\tau}\langle D_z\mathcal{I}(\underline{t}_{\tau,h}(s),\overline{z}_{\tau,h}(s)),z_k^{\tau,h}-z_{k-1}^{\tau,h}\rangle_{\mathcal{Z}^*,\mathcal{Z}}\\
&\quad=\frac{1}{\tau}\Big(\mathcal{R}(z_k^{\tau,h}-z_{k-1}^{\tau,h})+\tau\,\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-\Pi_h^*D_z\mathcal{I}(t_{k-1}^{\tau,h},z_k^{\tau,h}),\partial(\mathcal{R}_h\circ\Pi_h)(0)\}\Big)\\
&\quad=\mathcal{R}_h(\hat{z}'_{\tau,h})+\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-\Pi_h^*D_z\mathcal{I}(t_{k-1}^{\tau,h},z_k^{\tau,h}),\partial(\mathcal{R}_h\circ\Pi_h)(0)\}).
\end{aligned}$$

By taking into account the definition of $r_{\tau,h}$ in (3.2.56), integration over (σ_1,σ_2) then yields (3.2.55). It remains to estimate $r_{\tau,h}$. To this end, first observe that the definition of the affine and constant interpolants in (3.2.49) and (3.2.50) implies for every $k\in\{1,...,N\}$ and every $s\in[s_{k-1}^{\tau,h},s_k^{\tau,h})$ that

$$\hat{z}_{\tau,h}(s)-\overline{z}_{\tau,h}(s)=(s-s_k^{\tau,h})\hat{z}'_{\tau,h}(s)\ \text{ and }\ \hat{t}_{\tau,h}(s)-\underline{t}_{\tau,h}(s)=(s-s_{k-1}^{\tau,h})\hat{t}'_{\tau,h}(s),$$

which is frequently used in the following estimates. Now, let $k\in\{1,...,N\}$ and $s\in[s_{k-1}^{\tau,h},s_k^{\tau,h})$ be arbitrary. Then, by inserting the concrete form of $\mathcal{I}$ into the definition of $r_{\tau,h}$ in (3.2.56) and employing the coercivity of A together with $(s-s_k^{\tau,h})<0$, we arrive at

$$\begin{aligned}
r_{\tau,h}(s)&=(s-s_k^{\tau,h})\langle A(\hat{z}'_{\tau,h}(s)),\hat{z}'_{\tau,h}(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\\
&\quad+\langle D_z\mathcal{F}(\hat{z}_{\tau,h}(s))-D_z\mathcal{F}(\overline{z}_{\tau,h}(s)),\frac{\hat{z}_{\tau,h}(s)-\overline{z}_{\tau,h}(s)}{(s-s_k^{\tau,h})}\rangle_{\mathcal{V}^*,\mathcal{V}}\\
&\qquad-\langle D_zf(\hat{t}_{\tau,h}(s),\hat{z}_{\tau,h}(s))-D_zf(\underline{t}_{\tau,h}(s),\overline{z}_{\tau,h}(s)),\hat{z}'_{\tau,h}(s)\rangle_{\mathcal{V}^*,\mathcal{V}}\\
&\le\alpha(s-s_k^{\tau,h})\|\hat{z}'_{\tau,h}(s)\|_{\mathcal{Z}}^2\\
&\quad+\frac{1}{|s-s_k^{\tau,h}|}\,|\langle D_z\mathcal{F}(\overline{z}_{\tau,h}(s))-D_z\mathcal{F}(\hat{z}_{\tau,h}(s)),\overline{z}_{\tau,h}(s)-\hat{z}_{\tau,h}(s)\rangle_{\mathcal{V}^*,\mathcal{V}}|\\
&\qquad+|\langle D_zf(\hat{t}_{\tau,h}(s),\hat{z}_{\tau,h}(s))-D_zf(\underline{t}_{\tau,h}(s),\overline{z}_{\tau,h}(s)),\hat{z}'_{\tau,h}(s)\rangle_{\mathcal{V}^*,\mathcal{V}}|.
\end{aligned}\tag{3.2.59}$$

We apply Lemma 3.2.8 with $\varepsilon=\alpha/4$ to the second term on the right-hand side to obtain

$$\begin{aligned}
&\frac{1}{|s-s_k^{\tau,h}|}\,|\langle D_z\mathcal{F}(\overline{z}_{\tau,h}(s))-D_z\mathcal{F}(\hat{z}_{\tau,h}(s)),\overline{z}_{\tau,h}(s)-\hat{z}_{\tau,h}(s)\rangle_{\mathcal{V}^*,\mathcal{V}}|\\
&\qquad\le\frac{\alpha}{4}\,|s-s_k^{\tau,h}|\,\|\hat{z}'_{\tau,h}(s)\|_{\mathcal{Z}}^2+C_\alpha\,|s-s_k^{\tau,h}|\,\mathcal{R}_h(\hat{z}'_{\tau,h}(s))\,\|\hat{z}'_{\tau,h}(s)\|_{\mathbb{V}},
\end{aligned}$$

where we also used the positive homogeneity of $\mathcal{R}_h$. Likewise, using Lemma 3.2.10 with $\varepsilon = \alpha/4$, the third term is estimated by

$$\begin{aligned}
&|\langle D_z f(\hat{t}_{\tau,h}(s), \hat{z}_{\tau,h}(s)) - D_z f(\underline{t}_{\tau,h}(s), \overline{z}_{\tau,h}(s)), \hat{z}'_{\tau,h}(s)\rangle_{\mathcal{V}^*,\mathcal{V}}| \\
&\quad \le \nu(|\hat{t}_{\tau,h}(s) - \underline{t}_{\tau,h}(s)| + \|\hat{z}_{\tau,h}(s) - \overline{z}_{\tau,h}(s)\|_{\mathcal{W}})\|\hat{z}'_{\tau,h}(s)\|_{\mathcal{V}} \\
&\quad \le \nu|s - s_{k-1}^{\tau,h}|\hat{t}'_{\tau,h}(s)\|\hat{z}'_{\tau,h}(s)\|_{\mathbb{V}} + \frac{\alpha}{4}\,|s - s_k^{\tau,h}|\,\|\hat{z}'_{\tau,h}(s)\|_{\mathcal{Z}}^2 + C_\alpha\,|s - s_k^{\tau,h}|\,\mathcal{R}_h(\hat{z}'_{\tau,h}(s))\,\|\hat{z}'_{\tau,h}(s)\|_{\mathbb{V}}.
\end{aligned}$$

By inserting both estimates in (3.2.59) and using again that $(s - s_k^{\tau,h}) < 0$ as well as $\|\hat{z}'_{\tau,h}(s)\|_{\mathbb{V}} \le 1$, one deduces

$$\begin{aligned}
r_{\tau,h}(s) &\le C\Big(\mathcal{R}_h(\hat{z}'_{\tau,h}(s))\,(s_k^{\tau,h} - s) + \hat{t}'_{\tau,h}(s)\,(s - s_{k-1}^{\tau,h})\Big) \\
&\le C\,\tau\,\Big(\mathcal{R}_h(\hat{z}'_{\tau,h}(s)) + \hat{t}'_{\tau,h}(s)\Big).
\end{aligned}$$

Integrating and exploiting the definition of $\hat{z}_{\tau,h}$ and $\hat{t}_{\tau,h}$, respectively, then yields

$$\begin{aligned}
\int_0^s r_{\tau,h}(\sigma)\,\mathrm{d}\sigma &\le \sum_{i=1}^N \int_{s_{i-1}^{\tau,h}}^{s_i^{\tau,h}} C\,\tau\Big(\mathcal{R}_h(\hat{z}'_{\tau,h}(s)) + \hat{t}'_{\tau,h}(s)\Big)\,\mathrm{d}s \\
&= C\,\tau\sum_{i=1}^N \left\{\mathcal{R}_h\left(\frac{z_i^{\tau,h} - z_{i-1}^{\tau,h}}{s_i^{\tau,h} - s_{i-1}^{\tau,h}}\right) + \frac{t_i^{\tau,h} - t_{i-1}^{\tau,h}}{s_i^{\tau,h} - s_{i-1}^{\tau,h}}\right\}(s_i^{\tau,h} - s_{i-1}^{\tau,h}) \\
&\le C\,\tau\Big(T + \sum_{i=1}^N \mathcal{R}_h(z_i^{\tau,h} - z_{i-1}^{\tau,h})\Big).
\end{aligned}$$

Thanks to Lemma 3.2.5, the bracket on the right-hand side is bounded independent of τ and h so that (3.2.58) is proven, too. □

Remark 3.2.17. A comparison of the discrete energy identity in (3.2.55) and the continuous one in (2.4.6) shows that the coefficient $\|\hat{z}'_{\tau,h}\|$ is missing in front of the distance. It would be possible to reformulate the optimality conditions in Lemma 3.2.2 in a way such that this coefficient would arise in (3.2.55). This, however, would complicate the passage to the limit in the next section. As we will see at the end of the proof of Theorem 3.2.19, (3.2.55) is sufficient to obtain the desired energy identity in (2.4.6).

3.2.4 Convergence theorem

Before we come to the main result, i.e., the passage to the limit in the discrete energy identity and therewith ultimately the existence of parametrized solutions, we need one last preparatory result, which guarantees the weak lower semicontinuity of the distance term in (3.2.55).

Lemma 3.2.18. *Let $\xi_h \in \mathcal{Z}^*$ with $\xi_h \rightharpoonup \xi$ in $\mathcal{Z}^*$ for $h \to 0$. Suppose, moreover, that the distance is uniformly bounded, i.e., $\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-\xi_h, \partial(\mathcal{R}_h \circ \Pi_h)(0)\} \le C$ with C independent of h. Then the following weak lower semicontinuity result holds true:*

$$\liminf_{h\to 0}\ \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-\xi_h, \partial(\mathcal{R}_h \circ \Pi_h)(0)\} \ge \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-\xi, \partial\mathcal{R}(0)\}. \tag{3.2.60}$$

Proof. First of all, thanks to Lemma A.3.8, we know that the minimum in the definition of the distance is attained so that there exists $\mu_h \in \partial\mathcal{R}_h(0)$ with

$$\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\xi_h, \partial(\mathcal{R}_h \circ \Pi_h)(0)\} = \|\mu_h + \xi_h\|_{\mathbb{V}^{-1}}. \tag{3.2.61}$$

Therewith, we define $\eta_h := \mu_h + \xi_h$ and infer $\|\eta_h\|_{\mathcal{V}^*} \le C$ by assumption. Hence, we may extract a weakly convergent subsequence $\eta_{h_n} \rightharpoonup \eta$ in $\mathcal{V}^*$ for $n \to \infty$. In particular, due to the lower semicontinuity of the norm $\|\cdot\|_{\mathbb{V}^{-1}}$, it holds

$$\|\eta\|_{\mathbb{V}^{-1}} \le \liminf_{n\to\infty} \|\eta_{h_n}\|_{\mathbb{V}^{-1}} = \liminf_{n\to\infty} \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\xi_{h_n}, \partial(\mathcal{R}_{h_n} \circ \Pi_{h_n})(0)\}. \tag{3.2.62}$$

We proceed with showing that $\eta = \mu + \xi$ for some $\mu \in \partial\mathcal{R}(0)$. To this end, we first note that by $\mathcal{V}^* \subset \mathcal{Z}^*$ and the weak convergence of ξ_{h_n} it holds $\mu_{h_n} = \xi_{h_n} - \eta_{h_n} \rightharpoonup \xi - \eta$ in $\mathcal{Z}^*$ and we define $\mu = \xi - \eta$. Now, $\mu_{h_n} \in \partial(\mathcal{R}_{h_n} \circ \Pi_{h_n})(0)$ is equivalent to

$$\mathcal{R}_{h_n}(\Pi_{h_n} z) \ge \langle \mu_{h_n}, z\rangle_{\mathcal{Z}^*,\mathcal{Z}} \quad \forall\, z \in \mathcal{Z}.$$

By weak convergence, the right-hand side converges to $\langle \mu, z\rangle_{\mathcal{Z}^*,\mathcal{Z}}$. The left-hand side converges to $\mathcal{R}(z)$ on a dense subset $\mathcal{U}$ by assumption (c) on the approximation of the dissipation potential. By density of $\mathcal{U} \subset \mathrm{dom}(\mathcal{R})$ and continuity of $\mathcal{R}$ on $\mathrm{dom}(\mathcal{R})$, we thus obtain $\mathcal{R}(z) \ge \langle \mu, z\rangle_{\mathcal{Z}^*,\mathcal{Z}}$ for all $z \in \mathrm{dom}(\mathcal{R})$ and therefore clearly for all $z \in \mathcal{Z}$, so that $\mu \in \partial\mathcal{R}(0)$. Hence, we conclude from (3.2.62) that

$$\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\xi, \partial\mathcal{R}(0)\} \le \|\mu + \xi\|_{\mathbb{V}^{-1}} = \|\eta\|_{\mathbb{V}^{-1}} \le \liminf_{n\to\infty} \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\xi_{h_n}, \partial(\mathcal{R}_{h_n} \circ \Pi_{h_n})(0)\}.$$

Since this holds for all subsequence of η_h, we ultimately arrive at the desired lower semicontinuity in (3.2.60). □

We now have everything at hand to prove our main convergence result.

Theorem 3.2.19 (Convergence towards parametrized solutions)**.** *There exists a sequence of parameters $\{\tau_n, h_n\}_{n\in\mathbb{N}} \subset \mathbb{R}_+ \times \mathbb{R}_+$ converging to zero so that the affine interpolants generated by the fully discrete local stationarity scheme (LISS) and the artificial end time defined in (3.2.51) satisfy*

$$S_{\tau_n,h_n} \to S, \tag{3.2.63}$$

$$\hat{t}_{\tau_n,h_n} \overset{*}{\rightharpoonup} \hat{t} \quad \text{in } W^{1,\infty}(0,S;\mathbb{R}), \tag{3.2.64}$$

$$\hat{z}_{\tau_n,h_n} \overset{*}{\rightharpoonup} \hat{z} \quad \text{in } W^{1,\infty}(0,S;\mathcal{V}) \cap H^1(0,S;\mathcal{Z}), \tag{3.2.65}$$

$$\hat{z}_{\tau_n,h_n}(s) \rightharpoonup \hat{z}(s) \quad \text{in } \mathcal{Z} \text{ for every } s \in [0,S], \tag{3.2.66}$$

and the limit $(\hat{t}, \hat{z})$ is a parametrized solution in the sense of Definition 2.4.2.

Moreover, every accumulation point $(\hat{t}, \hat{z})$ of sequences in the sense of (3.2.63)–(3.2.66) is a parametrized solution.

Proof. The arguments are similar to the semi-discrete case without a spatial discretization and

with bounded dissipation, discussed in [Kne19]. However, we have to include the passage to limit $h \searrow 0$ and take care of the unboundedness of $\partial\mathcal{R}(0)$.
The existence of a (sub-)sequence satisfying (3.2.63)–(3.2.65) is an immediate consequence of the uniform estimates in Lemma 3.2.6, Lemma 3.2.15, and (3.2.51). By the Aubin-Lions lemma $W^{1,\infty}(0,S;\mathcal{V}) \cap L^\infty(0,S;\mathcal{Z})$ compactly embeds in $C(0,S;\mathcal{V})$ so that $\hat{z}_{\tau_n,h_n}$ uniformly converges in $\mathcal{V}$ to $\hat{z}$. However, Lemma 3.2.6 tells us that, for every $s \in [0,S]$, $\{\hat{z}_{\tau_n,h_n}(s)\}_{n\in\mathbb{N}}$ is bounded in $\mathcal{Z}$ and therefore there is a weakly convergent subsequence. Due to the uniform convergence in $\mathcal{V}$, the pointwise limit equals $\hat{z}(s)$, which implies (3.2.66).
It remains to show that every (weak) limit is a parametrized solution. For this purpose, let $\{\tau_n, h_n\}$ be an arbitrary null sequence and assume that the convergences in (3.2.63)–(3.2.66) hold. In order to simplify the notation, we indicate by $\{\cdot\}_n$ the sequence of $\{\cdot\}_{\tau,h}$ corresponding to $\{\tau_n, h_n\}$. Analogously, we abbreviate the index h_n simply by n. We proceed in several steps and start with the following:

Convergence of piecewise constant interpolants. First, we show that the piecewise constant interpolants converge pointwise to the same limit. We exemplarily consider $\overline{z}_{\tau,h}$. Because of (3.2.24), there is a subsequence, for convenience also denoted by $\overline{z}_n$, converging in every $s \in [0,S]$ weakly in $\mathcal{Z}$ to some $\tilde{z}(s)$. Hence, the compact embedding of $\mathcal{Z}$ in $\mathcal{V}$ implies $\overline{z}_n(s) \to \tilde{z}(s)$ in $\mathcal{V}$ for all $s \in [0,S]$. Moreover, by (3.2.49) and (3.2.50), we have for all $k \in \{1,...,N\}$ and all $s \in [s^n_{k-1}, s^n_k)$ that

$$\|\hat{z}_n(s) - \overline{z}_n(s)\|_{\mathbb{V}} = |s - s^n_k|\, \|\hat{z}'_n(s)\|_{\mathbb{V}} \le \tau \to 0,$$

where we used (3.2.53) and (3.2.51) for the last estimate. Hence, we obtain $\tilde{z}(s) = \hat{z}(s)$ for all $s \in [0,S]$ and the uniqueness of the weak limit implies the weak convergence of the whole sequence $\{\overline{z}_n\}$. For the other piecewise constant interpolants, one argues completely analogously so that

$$\underline{t}_n(s), \overline{t}_n(s) \to \hat{t}(s), \qquad \underline{z}_n(s), \overline{z}_n(s) \rightharpoonup \hat{z}(s) \ \text{ in } \mathcal{Z} \quad \forall\, s \in [0,S] \tag{3.2.67}$$

is obtained, as desired.

Initial and end time conditions. Since the Ritz projection trivially fulfills

$$\hat{z}_n(0) = z_0^{\tau,h} = P_n(z_0) \to z_0 \quad \text{in } \mathcal{Z}, \tag{3.2.68}$$

the pointwise convergence in (3.2.66) implies $\hat{z}(0) = z_0$ as desired. Moreover, thanks to (3.2.64), $\hat{t}_n$ converges uniformly to $\hat{t}$ so that

$$0 = \hat{t}_n(0) \to \hat{t}(0) \quad \text{and} \quad T = \hat{t}_n(S_n) \to \hat{t}(S),$$

where we also used (3.2.63).

Complementarity relations. We continue with the complementarity-like relations in (2.4.5). First, the set

$$\{(\tau, \zeta) \in L^2(0,S) \times L^2(0,S;\mathcal{V}) : \tau(s) \ge 0,\ \tau(s) + \|\zeta(s)\|_{\mathbb{V}} \le 1 \text{ f.a.a. } s \in (0,S)\}$$

is clearly convex and closed, thus weakly closed and consequently, we obtain that the weak limit

$(\hat{t}, \hat{z})$ satisfies the inequalities in (2.4.5a). Next, we turn to (2.4.5b), whose derivation is by far more involved. On account of the weak continuity assumptions for $D_z f$ and $D_z \mathcal{F}$, it follows from (3.2.67) that

$$D_z\mathcal{I}(\underline{t}_n(s), \overline{z}_n(s)) \rightharpoonup D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)) \;\text{ in } \mathcal{Z}^* \quad \forall\, s \in [0, S].$$

Thanks to (3.1.8), i.e., $\Pi_n(z) \to z$ in $\mathcal{Z}$ for every $z \in \mathcal{Z}$, this also gives

$$\begin{aligned}\langle \Pi_n^* D_z\mathcal{I}(\underline{t}_n(s), \overline{z}_n(s)), z\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ = \langle D_z\mathcal{I}(\underline{t}_n(s), \overline{z}_n(s)), \Pi_n z\rangle_{\mathcal{Z}^*,\mathcal{Z}} \to \langle D_z\mathcal{I}(\hat{t}_n(s), \hat{z}(s)), z\rangle_{\mathcal{Z}^*,\mathcal{Z}} \quad \forall\, z \in \mathcal{Z},\end{aligned} \tag{3.2.69}$$

i.e., weak convergence of $\Pi_n^* D_z\mathcal{I}(\underline{t}_n(s), \overline{z}_n(s))$ to $D_z\mathcal{I}(\hat{t}(s), \hat{z}(s))$ in $\mathcal{Z}^*$. Now we can take a closer look at the distance in (2.4.5b). The weak convergence of $\Pi_n^* D_z\mathcal{I}(\underline{t}_n(s), \overline{z}_n(s))$ and the uniform boundedness of the distance from (3.2.32) allow us to apply Lemma 3.2.18, which gives

$$\begin{aligned}\liminf_{n\to\infty} \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-\Pi_n^* D_z\mathcal{I}(\underline{t}_n(s), \overline{z}_n(s)), \partial(\mathcal{R}_n \circ \Pi_n)(0)\} \\ \geq \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\}.\end{aligned} \tag{3.2.70}$$

To show (2.4.5b), let us abbreviate

$$\begin{aligned}\xi_n(s) &:= \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-\Pi_n^* D_z\mathcal{I}(\underline{t}_n(s), \overline{z}_n(s)), \partial(\mathcal{R}_n \circ \Pi_n)(0)\}, \\ \xi(s) &:= \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\},\end{aligned}$$

so that (3.2.70) reads

$$\liminf_{n\to\infty} \xi_n(s) \geq \xi(s) \geq 0 \quad \forall\, s \in [0, S]. \tag{3.2.71}$$

We next address the measurability of ξ. The embedding $H^1(0, S; \mathcal{Z}) \hookrightarrow C(0, T; \mathcal{Z})$ and the continuity of $D_z\mathcal{I}$ imply that $s \mapsto -D_z\mathcal{I}(\hat{t}(s), \hat{z}(s))$ is continuous. Exploiting Lemma 3.2.18, we can conclude that ξ is lower semicontinuous and therefore, indeed, measurable.
Now, consider an arbitrary $\kappa \geq 0$ and define $\xi_{n,\kappa}(s) := \min\{\xi_n(s), \xi(s), \kappa\}$ such that $\xi_{n,\kappa}(s)$ converges to $\xi_\kappa(s) := \min\{\xi(s), \kappa\}$ almost everywhere in $(0, S)$. Since ξ_κ is measurable (as ξ is so) and $\kappa \geq \xi_{n,\kappa}(s)$, Lebesgue's dominated convergence theorem gives

$$\xi_{n,\kappa} \to \xi_\kappa \text{ in } L^1(0, S).$$

Thus, thanks to $\xi_n(s) \geq \xi_{n,\kappa}(s)$ and the weak* convergence of $\hat{t}'$, we obtain from (3.2.57) that

$$0 = \liminf_{n\to\infty} \int_0^S \hat{t}_n'(s)\, \xi_n(s)\, \mathrm{d}s \geq \liminf_{n\to\infty} \int_0^S \hat{t}_n'(s)\, \xi_{n,\kappa}(s)\, \mathrm{d}s = \int_0^S \hat{t}'(s)\, \xi_\kappa(s)\, \mathrm{d}s.$$

Since $\kappa \geq 0$ was arbitrary, this inequality holds for every κ so that Fatou's lemma yields

$$0 \geq \liminf_{\kappa\to\infty} \int_0^S \hat{t}'(s)\, \xi_\kappa(s)\, \mathrm{d}s \geq \int_0^S \hat{t}'(s)\, \xi(s)\, \mathrm{d}s \geq 0.$$

Because of $\xi \geq 0$ and $\hat{t}' \geq 0$ a.e. in $(0, S)$, cf. (2.4.5a), this gives (2.4.5b).

Energy identity. Let $s \in [0, S]$ be arbitrary. Thanks to the weak lower semicontinuity of

$\mathcal{I}(t,\cdot)$ and assumption ($\mathcal{I}_{f3}$), (3.2.64) and (3.2.66) yield

$$\mathcal{I}(\hat{t}(s),\hat{z}(s) \leq \liminf_{n\to\infty} \mathcal{I}(\hat{t}_n(s),\hat{z}_n(s)).$$

Moreover, the lower semicontinuity from Lemma A.3.5 and assumption (b) on the discretization of $\mathcal{R}$ imply, in view of (3.2.65), that

$$\int_0^s \mathcal{R}(\hat{z}'(\sigma))\,\mathrm{d}\sigma \leq \liminf_{n\to\infty} \int_0^s \mathcal{R}(\hat{z}_n'(\sigma))\,\mathrm{d}\sigma \leq \liminf_{n\to\infty} \int_0^s \mathcal{R}_n(\hat{z}_n'(\sigma))\,\mathrm{d}\sigma.$$

Lastly, we obtain from $\|\hat{z}'(s)\|_{\mathbb{V}} \leq 1$, (3.2.70) and Fatou's lemma

$$\begin{aligned}
&\int_0^s \|\hat{z}'(\sigma)\|_{\mathbb{V}}\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(\sigma),\hat{z}(\sigma)),\partial\mathcal{R}(0)\}\,\mathrm{d}\sigma\\
&\quad\leq \int_0^s \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(\sigma),\hat{z}(\sigma)),\partial\mathcal{R}(0)\}\,\mathrm{d}\sigma\\
&\quad\leq \liminf_{n\to\infty} \int_0^s \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\Pi_n^* D_z\mathcal{I}(\underline{t}_n(s),\overline{z}_n(s)),\partial(\mathcal{R}_n\circ\Pi_n)(0)\}\,\mathrm{d}\sigma
\end{aligned}$$

so that, altogether, Lemma 3.2.16 yields

$$\begin{aligned}
&\mathcal{I}(\hat{t}(s),\hat{z}(s)) + \int_0^s \mathcal{R}(\hat{z}'(\sigma)) + \|\hat{z}'(\sigma)\|_{\mathbb{V}}\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(\sigma),\hat{z}(\sigma)),\partial\mathcal{R}(0)\}\,\mathrm{d}\sigma\\
&\leq \liminf_{n\to\infty}\Big(\mathcal{I}(\hat{t}_n(s),\hat{z}_n(s))\\
&\qquad + \int_0^s \mathcal{R}_n(\hat{z}_n'(\sigma)) + \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\Pi_n^* D_z\mathcal{I}(\underline{t}_n(s),\overline{z}_n(s)),\partial(\mathcal{R}_n\circ\Pi_n)(0)\}\,\mathrm{d}\sigma\Big)\\
&= \liminf_{n\to\infty}\Big(\mathcal{I}(\hat{t}_n(0),\hat{z}_n(0)) + \int_0^s \partial_t\mathcal{I}(\hat{t}_n(\sigma),\hat{z}_n(\sigma))\,\hat{t}_n'(\sigma)\,\mathrm{d}\sigma + \int_0^s r_n(\sigma)\,\mathrm{d}\sigma\Big).
\end{aligned}$$

In order to eventually arrive at the energy inequality (2.4.17), we investigate the three terms on the right-hand side separately. Due to the continuity of $\mathcal{I}$ and the strong convergence of $\hat{z}_n(0)$ to z_0 in $\mathcal{Z}$, we have $\mathcal{I}(\hat{t}_n(0),\hat{z}_n(0)) = \mathcal{I}(0,\hat{z}_n(0)) \to \mathcal{I}(0,z_0)$. In addition, exploiting (3.2.58) from Lemma 3.2.16, it holds

$$\limsup_{n\to\infty} \int_0^s r_n(\sigma)\,\mathrm{d}\sigma \leq 0.$$

Finally, concerning the second term, we make use of assumption ($\mathcal{I}_{f3}$) and the convergences (3.2.64) and (3.2.66), which guarantees that for almost all $s\in[0,S]$ it holds

$$\partial_t\mathcal{I}(\hat{t}_n(s),\hat{z}_n(s)) = \partial_t f(\hat{t}_n(s),\hat{z}_n(s)) \to \partial_t f(\hat{t}(s),\hat{z}(s)).$$

On account of assumption ($\mathcal{I}_{f1}$), i.e., $|\partial_t f(\hat{t}_n(s),\hat{z}_n(s))| \leq \mu(\|\hat{z}_n(s)\|_{\mathcal{Z}}+1)$, and the $L^\infty(0,S;\mathcal{Z})$-bound for $\hat{z}_n$, we may therefore apply Lebesgue's dominated convergence theorem to obtain

$$\partial_t\mathcal{I}(\hat{t}_n(s),\hat{z}_n(s)) \to \partial_t\mathcal{I}(\hat{t}(s),\hat{z}(s)) \quad \text{in } L^1(0,S).$$

Overall, exploiting the weak-$*$ convergence of $\hat{t}_n'$ in $L^\infty(0,S)$, we end up with

$$\mathcal{I}(\hat{t}(s),\hat{z}(s)) + \int_0^s \mathcal{R}(\hat{z}'(\sigma)) + \|\hat{z}'(\sigma)\|_{\mathbb{V}}\,\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(\sigma),\hat{z}(\sigma)),\partial\mathcal{R}(0)\}\,\mathrm{d}\sigma \\ \leq \mathcal{I}(0,z_0) + \int_0^s \partial_t\mathcal{I}(\hat{t}(\sigma),\hat{z}(\sigma))\,\hat{t}'(\sigma)\,\mathrm{d}\sigma, \quad (3.2.72)$$

which is the desired energy inequality. Taking into account that $\hat{z} \in H^1(0,S;\mathcal{Z})$, it follows from Lemma 2.4.6 that the sole inequality (3.2.72) is already equivalent to the energy identity (2.4.6), which completes the proof. □

Unfortunately, we do not obtain the nondegeneracy let alone normalization of the limit $(\hat{t},\hat{z})$ here. The main problem is the fact that the weak convergence of $\hat{z}_n$ in $H^1(0,S;\mathcal{Z})$ from (3.2.65) is not sufficient in order to pass to the limit in (3.2.53), that is, $\hat{t}_n'(s) + \|\hat{z}_n'(s)\| = 1$, and still obtain equality in the end. In [MZ14, EM06], the authors therefore provide sufficient conditions, which guarantee the nondegeneracy of the limit function. Moreover, in [EM06], a condition is given, which also preserve the normalization. Nevertheless, as shown in Lemma A.4.3, it is always possible to reparameterize a parametrized solution and obtain in order to normalize it. Regardless of this fact, we note that the above Theorem, while dedicated to the convergence analysis of the fully discrete local stationarity scheme, also provides an existence result for parametrized solutions in case of an unbounded dissipation $\mathcal{R}$ (choose $\mathcal{Z}_h = \mathcal{Z}$). Moreover, it is to be expected that relaxations on the form of the energy $\mathcal{I}$ and the regularity of the time-dependent part (e.g., $D_z f(t,z) \in \mathcal{Z}^*$ instead of $\mathcal{V}^*$) are possible, if the energy still fulfills some Gårding-like inequality, i.e.,

$$\langle D_z\mathcal{I}(t,z_1) - D_z\mathcal{I}(t,z_2), z_1 - z_2\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \alpha\|z_1 - z_2\|_{\mathcal{Z}}^2 - \lambda\|z_1 - z_2\|_{\mathcal{V}}^2.$$

Indeed, the specific form of $\mathcal{I}$ and the additional regularity of its components $D_z f$ and $D_z\mathcal{F}$ are mainly used in order to get this Gårding-like inequality (see, e.g., (3.2.34)-(3.2.35) to find (3.2.37)) and in order to obtain the weak-convergence of the spatial approximations in Lemma 3.2.11. However, this is purpose of further research.

3.2.5 Two examples

The purpose of this last section within Chapter 3 is to present two different examples fitting into the setting given in Section 3.1. The first of these possesses the classical semilinear structure for the energy, which has also been investigated in various papers on parametrized solutions, see, e.g., [Mie11, MZ14, KT18, Kne19]. More specifically:

Semilinear example. We consider a bounded domain $\Omega \subset \mathbb{R}^d$ with $d = 2,3$ and choose the spaces $\mathcal{Z} = H_0^1(\Omega)$, $\mathcal{V} = L^2(\Omega)$, and $\mathcal{X} = L^1(\Omega)$. For the operator $\mathbb{V} : L^2(\Omega) \to L^2(\Omega)^*$ corresponding to the viscosity part we use the Riesz isomorphism and, moreover, we let

$$A = -\Delta : H_0^1(\Omega) \to H^{-1}(\Omega),$$

the usual Laplace operator, such that the coercivity constant α equals Poincaré's constant. For the nonlinear part $\mathcal{F}$ we choose a scaled version of the well-known double well potential, i.e.,

$$\mathcal{F}(z) := 48 \int_\Omega \big(1 - z(x)^2\big)^2 \,\mathrm{d}x,$$

which is nonnegative and twice Fréchet-differentiable as a functional in $L^4(\Omega)$ and, via Sobolev embeddings, also in $H_0^1(\Omega)$ with

$$D_z\mathcal{F}(z)h = -192 \int_\Omega (1 - z(x)^2) z(x)\, h(x) \,\mathrm{d}x.$$

Due to $H_0^1(\Omega) \hookrightarrow L^6(\Omega)$, the derivative can be extended to a linear functional on $L^2(\Omega)$ and is continuous in these spaces. Moreover, the second derivative is given by

$$D_z^2\mathcal{F}(z)[h, v] = -192 \int_\Omega (1 - 3z(x)^2) v(x) h(x) \,\mathrm{d}x.$$

Using once again Sobolev embeddings and Hölder's inequality, one thus finds

$$|D_z^2\mathcal{F}(z)[v, h]| \le C\big(1 + \|z\|_{H^1(\Omega)}^2\big)\|v\|_{H^1(\Omega)}\, \|h\|_{L^2(\Omega)} \quad \forall\, z, v \in H_0^1(\Omega), h \in L^2(\Omega),$$

which is ($\mathcal{I}_{\mathcal{F}1}$) with $q = 2$. Lastly, we let the time-dependent part $f(t, z)$ itself only depend on some external load $\ell(t, x)$ with $\ell \in C^1(0, T; L^2(\Omega))$, so that

$$f(t, z) = \langle \ell(t, \cdot), z\rangle_{\mathcal{V}} = \int_\Omega \ell(t, x) z(x) \,\mathrm{d}x, \quad t \in [0, T].$$

It is easy to see that $f(t, z)$ satisfies the regularity requirements in ($\mathcal{I}_{f1}$)–($\mathcal{I}_{f3}$). Overall, we have

$$\mathcal{I} : [0, T] \times H^1(\Omega) \to \mathbb{R}, \quad \mathcal{I}(t, z) = \int_\Omega |\nabla z(x)|^2 + 48(1 - z(x)^2)^2 - \ell(t, x) z(x) \,\mathrm{d}x.$$

To complete this setting, we set the dissipation functional to the L^1-norm, i.e.,

$$\mathcal{R}(v) = \|v\|_{L^1(\Omega)},$$

so that (R3) is fulfilled with $\underline{\rho} = 1$. Hence, all the assumptions from Section 3.1 are satisfied and we will come back to this example in Section 4.1 about the actual realization of the local stationarity scheme. As a second example we consider a rate-independent damage model as described in [MS19c, KRZ13].

Damage Model. In this case, we let $\Omega \subset \mathbb{R}^2$ be a domain that corresponds to an elastic body. Within the time interval $[0, T]$, time dependent boundary conditions u_D as well as external boundary and volume forces ℓ are applied, which lead to a certain displacement u and possibly even to a damage, represented by the variable z, of the body. Usually, z is supposed to take values in $[0, 1]$ whereby $z(t, x) = 0$ means the body is completely sound and, correspondingly, $z(t, x) = 1$ means the body is comletely damaged. With view to the energy functional, we define

$\mathcal{U} = \{v \in H^1(\Omega, \mathbb{R}^2) : v|_{\Gamma_D} = 0\}$, $\mathcal{Z} = H^1(\Omega)$, $\mathcal{V} = L^2(\Omega)$ and set $\mathcal{E} : [0, T] \times V \times \mathcal{Z} \to \mathbb{R}$ as

$$\begin{aligned}\mathcal{E}(t, u, z) &= \frac{1}{2}\int_\Omega |\nabla z|^2 \,\mathrm{d}x + \int_\Omega f(z) \,\mathrm{d}x \\ &\quad + \frac{1}{2}\int_\Omega g(z)\, \mathbb{C}\varepsilon(u + u_D(t)) : \varepsilon(u + u_D(t)) \,\mathrm{d}x - \langle \ell(t), u\rangle_{\mathcal{U}} \\ &= \mathcal{I}_1(z) + \mathcal{E}_2(t, u, z).\end{aligned}$$

where $\mathbb{C}$ is the usual elasticity tensor and $\varepsilon(u) = \frac{1}{2}(\nabla u + \nabla u^\top)$ the linearized strain tensor. Moreover, the dissipation $\mathcal{R} : L^1(\Omega) \to [0, \infty]$ is given by

$$\mathcal{R}(v) = \begin{cases} \kappa \int_\Omega |v(x)| \,\mathrm{d}x, & \text{if } v \geq 0 \text{ a.e. in } \Omega, \\ +\infty, & \text{else,} \end{cases}$$

with the so-called fracture toughness $\kappa > 0$. The aim at this point is not to give a detailed description of the model, let alone prove the subsequent statements, but rather provide an application-oriented example that fits into the general setting from Section 3.1. The interested reader is therefore referred to the elaborations in [MS19c, KRZ13]. However, let us mention that the function g, which somehow represents the preservation of the elasticity of the material depending on the state of damage, is supposed to fulfill:

$$g \in C^2(\mathbb{R}), \text{ with } g', g'' \in L^\infty(\mathbb{R}), \text{ and } \exists \gamma_1, \gamma_2 > 0 : \forall z \in \mathbb{R} : \gamma_1 \leq g(z) \leq \gamma_2.$$

In particular, the lower bound $g \geq \gamma_1 > 0$ is to be noted here. It implies that even if the material is completely damaged, it does not lose all its rigidity. This is often referred to as *partial damage model*. Now, in order to bring this model into the form in ($\mathcal{I}_0$), it is convenient to reduce the system to the damage variable. This means that we require the displacement $u(t)$ to minimize the energy $\mathcal{E}(t, \cdot, z(t))$ at every time point $t \in [0, T]$, i.e.,

$$u(t) \in \arg\min\{\mathcal{E}(t, v, z(t)) : v \in \mathcal{U}\}. \tag{3.2.73}$$

Hence, we define $\mathcal{I}_2 : [0, T] \times \mathcal{Z} \to \mathbb{R}$ by $\mathcal{I}_2(t, z) = \inf_{v \in \mathcal{U}} \mathcal{E}_2(t, v, z)$ and let $\mathcal{I}(t, z) = \mathcal{I}_1(z) + \mathcal{I}_2(t, z)$. Taking for example $f(z) = (1 - z)^2$, which is a common choice in this context, we observe that $\mathcal{I}_1$ is exactly of the form $\frac{1}{2}\langle Az, z\rangle + \mathcal{F}(z)$. Concerning the assumptions ($\mathcal{I}_{f1}$)-($\mathcal{I}_{f3}$) on the time-dependent part, we rely on the following result.

Lemma 3.2.20 ([KRZ13, Lem. 2.4, Lem. 2.6, Lem. 2.8 and Cor. 2.9])**.** *Let the assumptions from [KRZ13, Lem. 2.8] hold and assume that $p > 4$ in [KRZ13, Eq. (2.15)]. Then there exist constants $C_1, C_2, c_3 > 0$ such that*

$$\mathcal{I}_2(t, z) \geq -C_1 \quad \textit{and} \quad |\partial_t \mathcal{I}_2(t, z)| \leq C_2 \tag{3.2.74}$$

as well as

$$\langle D_z\mathcal{I}_2(t_1, z_1) - D_z\mathcal{I}_2(t_2, z_2), v\rangle_{L^2(\Omega)} \leq c_3(|t_2 - t_1| + \|z_1 - z_2\|_{L^r(\Omega)})\|v\|_{L^2(\Omega)}.$$

for every $r \in [\frac{6p}{p-4}, \infty)$. *Moreover, for any sequences* $t_k \to t$ *and* $z_k \rightharpoonup z$ *in* $\mathcal{Z}$, *it holds*

$$D_z\mathcal{I}_2(t_k, z_k) \rightharpoonup D_z\mathcal{I}_2(t, z) \quad \textit{in } \mathcal{Z}^*, \tag{3.2.75}$$

$$\mathcal{I}_2(t_k, z_k) \to \mathcal{I}_2(t, z) \quad \textit{and} \quad \partial_t\mathcal{I}_2(t_k, z_k) \to \partial_t\mathcal{I}_2(t, z). \tag{3.2.76}$$

This guarantees that the time-dependent part $f(t, z) = -\mathcal{I}_2(t, z)$ indeed complies with the assumptions in Section 3.1. Unfortunately, the assumption on the exponent p in the above lemma, which relates to the integrability of the minimizer in (3.2.73), is rather restrictive and results in this direction have only been obtained under additional conditions, see also [KRZ13, Sec. 2.4]. Nevertheless, relying on [KRZ13, Cor. 2.10], the energy $\mathcal{I}$ in this example fulfills the already mentioned Gårding-like inequality (see end of Section 3.2.4)

$$\langle D_z\mathcal{I}(t, z_1) - D_z\mathcal{I}(t, z_2), z_1 - z_2\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \alpha\|z_1 - z_2\|^2_{\mathcal{Z}} - \lambda\|z_1 - z_2\|^2_{\mathcal{V}}$$

for any $p > 2$. On that basis, we expect that a convergence result in the sense of Theorem 3.2.19 can still be obtained. However, since it was not the main purpose of this thesis to incorporate the model above into the setting from Section 3.1, this gives rise to further research.

3.3 A priori error analysis

Theorem 3.2.19 above shows that the local incremental stationarity scheme LISS provides approximate parametrized solutions whose (weak) limits converge to exactly this type of solution as $\tau \searrow 0$. In this sense, the described algorithm is thus consistent. Nevertheless, let us mention that there exist other discretization methods to approximate parametrized solutions, such as relaxed local minimization schemes as proposed in [ACFS17] or alternating minimization schemes, if a second variable enters the energy functional, as for example in the damage model from Section 3.2.5, see also, e.g., [Rou15, KN17, Alm20]. Moreover, time discretization and viscous regularization can be coupled to approximate a parametrized solution, see [KS13, MRS16]. For a detailed overview, the reader is referred to [Kne19]. However, when it comes to rates of convergence for discretizations using either (3.0.1) or LISS, the literature becomes rather scarce. Since in the case of nonconvex energies the (parametrized) solutions of (RIS) are, in general, not unique (not even locally) as there might be a whole continuum of solutions, one can, in general, hardly expect any a priori estimates. The situation changes if one turns to *uniformly convex energies*. In this case, however, there is no need for a localized scheme as in (3.0.1) so that one can drop the additional constraint in (3.0.1a) and simply use the a time update of the form $t_k = t_{k-1} + \tau$. The method arising in this way is the *global incremental minimization scheme* from (3.0.2) and can be shown to converge to the *global energetic solution*, which is unique in case of a uniformly convex energy, see the proof of Theorem 2.2.2. Additionally, in [MT04, MPPS10], the authors show that the error between the discrete solution of this scheme and the global energetic solution is of order $\mathcal{O}(\sqrt{\tau})$. This result has been improved in [Mie06] and, slightly more general, in [Bar14] to rates of order $\mathcal{O}(\tau)$ for the case of a quadratic and coercive energy. An energy functional with these properties arises for instance in case of quasistatic elastoplasticity with linear kinematic hardening, where several convergence results have been obtained by various authors, see, e.g., [HR99, AC00] and the references therein. Recently, in [RSS17], the authors provide an a priori error estimate for the global minimization scheme in case of a semilinear and uniformly convex energy including a spatial discretization.

In contrast, to the best of the author's knowledge, there exist no such convergence results for the local incremental stationarity scheme in (3.0.1), not even in the case of a uniformly convex energy. The ultimate goal of the present Section 3.3 is therefore exactly this, to derive an a priori estimate for the approximations based on LISS. Moreover, we provide an a priori estimate if the energy functional is only *locally uniformly convex* along a given solution trajectory. At this point, the local incremental stationarity scheme turns out to be superior to the global minimization one, since the latter, in general, does not satisfy such an a priori estimate as we will demonstrate by means of a counterexample. Note that, however, we do not incorporate a spatial discretization in the a priori analysis.

The plan of this section looks as follows: First of all, we will convince ourselves, by means of a simple one-dimensional example, that, indeed, one cannot expect any convergence result for the whole sequence of discrete solutions without any further restrictions on the energy such as (local) uniform convexity. Hence, we will subsequently state some additional assumptions on the involved quantities and provide some basic estimates that are frequently used throughout the convergence analysis. The following Sections 3.3.3 and 3.3.4 are then devoted to the derivation of the a priori estimates. We start by considering a semilinear energy functional with the additional assumption that the driving force is Lipschitz continuous with a sufficiently small Lipschitz constant. Here, we derive a priori estimates of optimal order, first in the case that the energy is (globally) uniformly convex and afterwards also for the case of local uniform convexity. In Section 3.3.4, we then consider slightly more general energies in the form of ($\mathcal{I}_0$) and also drop the smallness assumption on the Lipschitz constant. Despite of a global uniform convexity condition, we do not obtain the optimal order of convergence in this case, see Remark 3.3.18 below. The actual numerical experiments of our theoretical findings are illustrated in Chapter 4, particularly Section 4.1.3 and Section 4.2. Lastly, it is to be noted that several of the subsequent results have already been published in [MS20], particularly the ones in Sections 3.3.1–3.3.3.

3.3.1 A counterexample in the case of a nonconvex energy

Before we actually start our error analysis, let us take a look at a first numerical example for the local minimization algorithm, which illustrates that one cannot expect any convergence result going beyond Theorem 3.2.19 without further assumptions. For this example, we set $\mathcal{Z} = \mathcal{V} = \mathcal{X} = \mathbb{R}$ as well as

$$\mathcal{R}(v) = |v| \quad \text{and} \quad \mathcal{I}(t, z) = \frac{1}{2} z^2 + \mathcal{F}(z) - \ell(t) z \tag{3.3.1}$$

with

$$\mathcal{F}(z) = 2|z|^3 - 5/2\, z^2 + 1 \quad \text{and} \quad \ell(t) = -24(t - 1/4)^2 + 5/3.$$

We already considered this type of energy functional several times in Chapter 2, particularly in order to demonstrate that solutions might be discontinuous in the case of nonconvex energies. Nevertheless, the lack of convexity also implies that solutions are, in general, not unique, see, e.g., Example 2.6.2. However, it is not clear beforehand, whether or not the discrete approximations may not actually converge to some particular parametrized solution (potentially even with some rate) or not. The following example demonstrates that this is, in general, *not* the case. For $z_0 = -1/3$, straightforward calculations show that

$$z_1(t) \equiv -1/3 \quad \text{and} \quad z_2(t) = \begin{cases} -1/3, & t \in [0, 1/4), \\ 1/3(1+\sqrt{2}), & t \in [1/4, 1/2], \end{cases}$$

are BV-solutions of the rate-independent system (3.3.1). The two numerical results depicted in Figure 3.3.1, which can be obtained by the algorithm LISS, show that, although z_1 is continuous, the discrete solution approximates either z_1 or z_2 depending on the choice of the parameter τ. As indicated above, it is therefore not clear if any of the solutions is preferred by the algorithm, without any further restrictions on the energy functional, particularly some kind of (uniform) convexity. In addition, an a priori error estimate can hardly be expected. As a consequence of this example, we will impose additional assumptions on the energy to derive a priori error estimates. On the one hand, we will assume that the energy is uniformly convex (Sections 3.3.3 & 3.3.4) and later on generalize our results for the case of locally uniformly convex energies (Section 3.3.3).

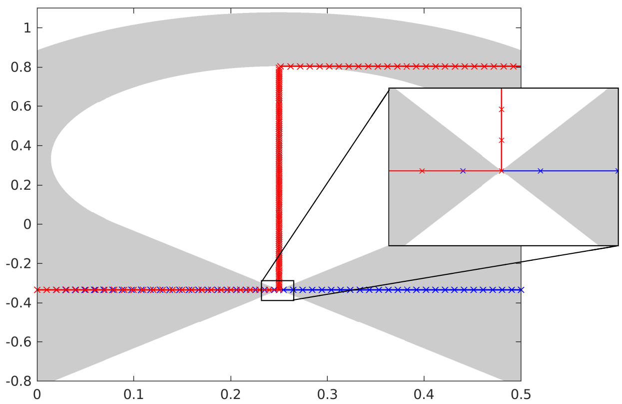


Figure 3.3.1: Approximations of two different parametrized solutions. Depending on the choice of τ either of two solutions is approximated. The set of local stability, i.e., $\cup_{t\in[0,0.5]}\mathcal{S}(t)$, is depicted in gray.

3.3.2 Additional assumptions

As already mentioned above, we need to sharpen the assumptions on the quantities in Section 3.1. This means that, in addition to $(\mathcal{I}_0)$,$(\mathcal{I}_{f1})$-$(\mathcal{I}_{f3})$ and $(\mathcal{I}_{\mathcal{F}1})$-$(\mathcal{I}_{\mathcal{F}2})$, we require:

Spaces. Concerning the underlying spaces, we stick with the Banach space $\mathcal{X}$ and the Hilbert spaces $\mathcal{Z}, \mathcal{V}$ whereby $\mathcal{Z} \overset{c,d}{\hookrightarrow} \mathcal{V} \hookrightarrow \mathcal{X}$. However, for convenience, we will assume w.l.o.g. that the embedding constant $c_\mathcal{Z}$ of $\mathcal{Z} \to \mathcal{V}$ fulfills $c_\mathcal{Z} = 1$. Otherwise only the constants in the corresponding estimates will change. For the same reason, we will use the natural norm in $\mathcal{V}$ rather than $\mathbb{V}$ as carried out in Section 3.2. The Riesz isomorphism associated with $\mathcal{V}$ is denoted by $J_\mathcal{V} : \mathcal{V} \to \mathcal{V}^*$.

Energy. For the energy functional we also maintain the semilinear form:

$$\mathcal{I} : [0,T] \times \mathcal{Z} \to \mathbb{R}, \quad \mathcal{I}(t,z) = \frac{1}{2}\langle Az, z\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \mathcal{F}(z) - f(t,z)\,.$$

Still, $A \in \mathcal{L}(\mathcal{Z}, \mathcal{Z}^*)$ is a self-adjoint and coercive operator but concerning f and $\mathcal{F}$ we additionally assume that for all $r > 0$ there exists $C_\mathcal{F}(r) \geq 0$ such that for all $z_1, z_2 \in \overline{B_\mathcal{Z}(0,r)}$ it holds

$$\langle \big[D_z^2\mathcal{F}(z_1) - D_z^2\mathcal{F}(z_2)\big]v, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq C(r)\|z_1 - z_2\|_\mathcal{Z}\|v\|_\mathcal{Z}^2. \tag{3.3.2}$$

Moreover, we require the same property to hold for f, i.e., $f(t,\cdot) \in C^2(\mathcal{Z};\mathbb{R})$ and

$$\langle \big[D_z^2 f(t,z_1) - D_z^2 f(t,z_2)\big]v, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq C_f(r)\|z_1 - z_2\|_\mathcal{Z}\|v\|_\mathcal{Z}^2 \quad \forall t \in [0,T],\ \forall z_1, z_2 \in \overline{B_\mathcal{Z}(0,r)} \tag{3.3.3}$$

for some constant $C_f(r) \geq 0$ depending only on the radius r. Additionally, we suppose that

$$|\partial_t f(t,z_1) - \partial_t f(t,z_2)| \leq c\|z_1 - z_2\|_\mathcal{Z} \quad \forall z_1, z_2 \in \mathcal{Z}, \tag{3.3.4}$$

for all $t \in [0,T]$ and some constant $c \geq 0$. Note that these assumptions imply that $\mathcal{I}(t,\cdot) \in C^{2,1}_{loc}(\mathcal{Z};\mathbb{R})$, see Definition 1.0.3. Lastly, we require $\mathcal{I}$ to be (at least locally) uniformly convex, see the Assumption $\mathbf{GC}_\kappa$ and Assumption $\mathbf{LC}_\kappa$ below. This will be indicated at the appropriate places, though.

Initial data. Finally, we assume that the initial state z_0 satisfies $z_0 \in \mathcal{Z}$ and $0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(0,z_0)$, i.e., z_0 is locally stable.

Remark 3.3.1. Due to the convexity of $\mathcal{I}(t,\cdot)$ and the assumption on the initial state z_0, i.e., $0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(0,z_0)$, there holds $\mathcal{I}(0,z_0) \leq \mathcal{I}(0,z) + \mathcal{R}(z - z_0)$ for all $z \in \mathcal{Z}$ so that $z_1 = z_0$ is the unique minimizer of (3.0.1a), and consequently the first iterate of the local stationarity algorithm always equals the initial state. This also entails $t_1 = t_1 - t_0 = \tau$. We will use this fact at some places throughout the thesis. Note that the uniform convexity of $\mathcal{I}(0,\cdot)$ on $B_\mathcal{Z}(z_0,\tau)$ is perfectly sufficient for the above argument, which will become important in Section 3.3.3 below.

The convexity of $\mathcal{I}$ moreover implies that any iterate z_k of LISS is in fact a minimizer of the (local) minimization in (3.0.1). The scheme in LISS thus turns into the local minimization scheme (3.0.1), which is why we will subsequently use the terms *local minimizer* and *locally minimal* in order to denote a point satisfying the stationarity conditions in (3.3.6a)-(3.3.6d).

Before we start with our error analysis, let us recall three essential results from the previous Section 3.2 that are frequently used in what follows (without the additional spatial discretization; i.e., we have $\mathcal{R}_h = \mathcal{R}$, $\mathcal{Z}_h = \mathcal{Z}$ and $\Pi_h = \mathrm{Id}$):

- The iterates of Algorithm 3.0.1 fulfill (see Lemma 3.2.6):

$$\sup_{\tau>0,\, k\in\mathbb{N}} \|z_k\|_{\mathcal{Z}} < \infty. \tag{3.3.5}$$

  This particularly implies that $z_k \in B_{\mathcal{Z}}(0, r_0)$ for all $k \in \mathbb{N}$ for some r_0 independent of τ.

- Let $k \geq 1$ and z_k be an arbitrary solution of (3.0.1a) with associated t_k given by (3.0.1b). Then the following optimality properties are satisfied, see also Lemma 3.2.2: There exists a Lagrange multiplier $\lambda_k \geq 0$ such that

$$\lambda_k(\|z_k - z_{k-1}\|_{\mathcal{V}} - \tau) = 0, \tag{3.3.6a}$$

$$\tau\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_{k-1}, z_k), \partial\mathcal{R}(0)\} = \lambda_k\|z_k - z_{k-1}\|_{\mathcal{V}}^2, \tag{3.3.6b}$$

$$\left\{\begin{aligned} &\mathcal{R}(z_k - z_{k-1}) + \tau\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_{k-1}, z_k), \partial\mathcal{R}(0)\} \\ &\qquad\qquad = \langle -D_z\mathcal{I}(t_{k-1}, z_k), z_k - z_{k-1}\rangle_{\mathcal{Z}^*,\mathcal{Z}}, \end{aligned}\right. \tag{3.3.6c}$$

$$\mathcal{R}(v) \geq -\langle \lambda_k J_{\mathcal{V}}(z_k - z_{k-1}) + D_z\mathcal{I}(t_{k-1}, z_k), v\rangle_{\mathcal{V}^*,\mathcal{V}} \quad \forall v \in \mathcal{V}. \tag{3.3.6d}$$

  In particular

$$0 \in \partial\mathcal{R}(z_k - z_{k-1}) + \lambda_k J_{\mathcal{V}}(z_k - z_{k-1}) + D_z\mathcal{I}(t_{k-1}, z_k). \tag{3.3.7}$$

  In addition, (3.3.6a) and (3.3.6b) give

$$\lambda_k = \frac{1}{\tau}\, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_{k-1}, z_k), \partial\mathcal{R}(0)\}. \tag{3.3.8}$$

- From Proposition 3.2.12, particularly (3.2.44), we also have: For every $\tau > 0$, there exists an index $N(\tau) \in \mathbb{N}$ such that $t_{N(\tau)} \geq T$. Moreover, it holds

$$\sum_{k=1}^{N(\tau)} \|z_k - z_{k-1}\|_{\mathcal{Z}} \leq C_\Sigma \tag{3.3.9}$$

  for some $C_\Sigma = C_\Sigma(\alpha, \mathcal{F}, |\ell|_{Lip}, z_0, T) > 0$ independent of τ.

As a consequence thereof, we henceforth denote by $N(\tau)$ the number of necessary iterates to reach the final time T at fineness τ.

Remark 3.3.2. In order to keep the following arguments concise, we will proceed with the iteration for $t_{N(\tau)}$ until we find $z_{N(\tau)+n} \in \mathcal{Z}$, which is locally stable again, i.e., $0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(t_{N(\tau)}, z_{N(\tau)+n})$. Since, $t_{N(\tau)}$ might be strictly larger than T, we extend $f(\cdot, z)$ onto $[0, \infty)$ by constant continuation, i.e., $f(t, z) = f(T, z)$ for all $t \geq T$ and $z \in \mathcal{Z}$. In Lemma 3.3.8 below, we will see that, under suitable assumptions, a locally stable point is found after a finite number of steps, which is bounded independent of τ. Eventually, we denote $\hat{N}(\tau) := N(\tau) + n$.

As a preliminary point, before we start with the basic estimates used in this section, let us briefly recall the definition of the interpolants for the approximation of parametrized solutions as given in (3.2.49):

For $s \in [s_{k-1}, s_k)$, the continuous and piecewise affine interpolants are defined through

$$\hat{z}_\tau(s) := z_{k-1} + \frac{(s - s_{k-1})}{\tau}(z_k - z_{k-1}), \quad \hat{t}_\tau(s) := t_{k-1} + \frac{(s - s_{k-1})}{\tau}(t_k - t_{k-1}). \tag{3.3.10}$$

We moreover define, again, the artificial end time S_τ as that point where $\hat{t}$ reaches the end time T, i.e., it holds

$$\hat{t}(S_\tau) = T, \quad s_{N(\tau)-1} < S_\tau \leq s_{N(\tau)} \quad \text{and} \quad S_\tau \leq C_S, \tag{3.3.11}$$

whereby the boundedness of S_τ is easily obtained following the estimates in (3.2.48) and exploiting (3.3.9), see also Figure 3.2.1.

Basic estimates

As mentioned above, large parts of the subsequent error analysis are based on the following

Assumption GC$_\kappa$ (κ-uniform convexity)**.** *The energy functional $\mathcal{I}$ is* κ-uniformly convex*, that is, there exists a $\kappa > 0$ such that, for all $t \in [0,T]$ and all $z, v \in \mathcal{Z}$, it holds $\langle D_z^2 \mathcal{I}(t,z)v, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \kappa \|v\|_{\mathcal{Z}}^2$.*

This property especially implies that

$$\langle D_z\mathcal{I}(t, z_2) - D_z\mathcal{I}(t, z_1), z_2 - z_1\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \kappa \|z_2 - z_1\|_{\mathcal{Z}}^2 \qquad \forall z_1, z_2 \in \mathcal{Z}.$$

Later on, in Section 3.3.3, we will relax this assumption and turn to *locally* uniformly convex energies, see Assumption **LC$_\kappa$** below.

We now state some basic estimates, which will be used at several points in the subsequent analysis. The first of these already occurred in a very similar form in Lemma 3.2.8 and does not require any kind of convexity.

Lemma 3.3.3. *There exists $C_{\mathcal{F},r_0} > 0$, such that for all $z_1, z_2 \in \overline{B_{\mathcal{Z}}(0, r_0)}$ it holds that*

$$|\langle D_z\mathcal{F}(z_1) - D_z\mathcal{F}(z_2), v - w\rangle_{\mathcal{V}^*,\mathcal{V}}| \leq C_{\mathcal{F},r_0} \, \|z_1 - z_2\|_{\mathcal{Z}} \, \|v - w\|_{\mathcal{V}}$$

for all $v, w \in \mathcal{Z}$.

Proof. The result is a direct consequence of the growth-condition on $D_z^2\mathcal{F}$. Let $v, w \in \mathcal{Z}$ be arbitrary. Using the aforementioned growth condition in ($\mathcal{I}_{\mathcal{F}1}$) together with the embedding $\mathcal{Z} \hookrightarrow \mathcal{V}$ yields

$$|\langle D_z\mathcal{F}(z_1) - D_z\mathcal{F}(z_2), v - w\rangle_{\mathcal{V}^*,\mathcal{V}}| \leq C(1 + r_0^q) \, \|z_1 - z_2\|_{\mathcal{Z}} \, \|v - w\|_{\mathcal{V}}.$$

Taking $C_{\mathcal{F},r_0} = C(1 + r_0^q)$ we end up with the assertion. □

Remark 3.3.4. In what follows, Lemma 3.3.3 is commonly used with z_1, z_2 being iterates of the local stationarity scheme. Hence, Lemma 3.2.6, i.e., the uniform boundedness of the iterates (see (3.3.5)), allows to specialize Lemma 3.3.3 as follows: There exists a constant $C > 0$ such that, for all iterates $z_k, z_j \in \mathcal{Z}$ it holds

$$|\langle D_z\mathcal{F}(z_k) - D_z\mathcal{F}(z_j), v - w\rangle_{\mathcal{V}^*,\mathcal{V}}| \leq C\,\|z_k - z_j\|_{\mathcal{Z}}\,\|v - w\|_{\mathcal{Z}}.$$

We proceed with the following result, whose proof is more or less a direct consequence of the κ-uniform convexity of $\mathcal{I}$.

Lemma 3.3.5. *Under the Assumption* $\mathbf{GC}_\kappa$ *we have for all iterates* $k \in \mathbb{N}$, $k \leq N(\tau)$*:*

$$0 \geq \kappa\|z_{k+1} - z_k\|_{\mathcal{Z}}^2 - \nu(t_k - t_{k-1})\|z_{k+1} - z_k\|_{\mathcal{V}} + (\lambda_{k+1} - \lambda_k)\tau^2. \tag{3.3.12}$$

Proof. First, we observe that, due to the complementarity condition in (3.3.6a), it holds that $\lambda_k\|z_k - z_{k-1}\|_{\mathcal{V}}^2 = \lambda_k\tau^2$. Now, by inserting (3.3.6b) in (3.3.6c) and writing the resulting equation for the iteration $k+1$, we obtain

$$\mathcal{R}(z_{k+1} - z_k) = \langle -D_z\mathcal{I}(t_k, z_{k+1}), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} - \lambda_{k+1}\tau^2. \tag{3.3.13}$$

Testing the inequality (3.3.6d) with $v = z_{k+1} - z_k$ yields

$$\begin{aligned}\mathcal{R}(z_{k+1} - z_k) &\geq \langle -D_z\mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} - \lambda_k\|z_k - z_{k-1}\|_{\mathcal{V}}\|z_{k+1} - z_k\|_{\mathcal{V}} \\ &\geq \langle -D_z\mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} - \lambda_k\tau^2.\end{aligned}$$

Subtracting hereof the terms in (3.3.13), exploiting the κ-uniform convexity of $\mathcal{I}(t,\cdot)$ and estimate ($\mathcal{I}_{f2}$), we obtain

$$\begin{aligned}0 &\geq \langle D_z\mathcal{I}(t_k, z_{k+1}) - D_z\mathcal{I}(t_k, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\quad + \langle D_z\mathcal{I}(t_k, z_k) - D_z\mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} + (\lambda_{k+1} - \lambda_k)\tau^2 \\ &\geq \kappa\|z_{k+1} - z_k\|_{\mathcal{Z}}^2 - \nu(t_k - t_{k-1})\|z_{k+1} - z_k\|_{\mathcal{V}} + (\lambda_{k+1} - \lambda_k)\tau^2,\end{aligned} \tag{3.3.14}$$

which was claimed. □

Remark 3.3.6. Revisiting the proof of Lemma 3.3.5, we observe that the κ-uniform convexity is only necessary in order to obtain the last estimate in (3.3.14). Since we will take advantage of this fact in Section 3.3.3, it is useful to state this estimate explicitly here for reference purposes: For all $k \in \mathbb{N}$, $k \leq N(\tau)$, it holds (without assuming that $\mathcal{I}$ is uniformly convex) that

$$\begin{aligned}0 &\geq \langle D_z\mathcal{I}(t_k, z_{k+1}) - D_z\mathcal{I}(t_k, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\quad + \langle D_z\mathcal{I}(t_k, z_k) - D_z\mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} + (\lambda_{k+1} - \lambda_k)\tau^2.\end{aligned} \tag{3.3.15}$$

Finally, the uniform convexity of the energy functional allows us to estimate the difference of two consecutive iterates with respect to the difference in the time points.

Lemma 3.3.7. *Under Assumption* $\mathbf{GC}_\kappa$ *it holds for any* $k \in \mathbb{N}$ *with* $k \leq N(\tau)$ *that*

$$0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(t_{k-1}, z_k) \quad\Longrightarrow\quad \|z_{k+1} - z_k\|_{\mathcal{Z}} \leq \frac{\nu}{\kappa}\,(t_k - t_{k-1}).$$

Proof. Let $0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(t_{k-1}, z_k)$, which directly implies that $\lambda_k = 0$, due to (3.3.6a) and (3.3.6b). Thanks to Lemma 3.3.5 and the nonnegativity of the multiplier λ_{k+1}, we thus arrive at $\kappa\|z_{k+1} - z_k\|_{\mathcal{Z}}^2 - \nu(t_k - t_{k-1})\|z_{k+1} - z_k\|_{\mathcal{Z}} \leq 0$, where we used the embedding $\mathcal{Z} \hookrightarrow \mathcal{V}$ with constant $c_{\mathcal{Z}} = 1$. □

3.3.3 Linear time dependence in energy functional

For the first part of this a priori analysis we let

$$f : [0, T] \times \mathcal{Z} \to \mathbb{R}, \quad f(t, z) = \langle \ell(t), z\rangle_{\mathcal{V}^*,\mathcal{V}}$$

for some $\ell \in W^{1,\infty}(0, T; \mathcal{V}^*)$ whose Lipschitz constant is denoted by $|\ell|_{Lip}$. By the structure of f, it is easy to see that (3.3.3) is valid with $C_f(r) = 0$. In combination with assumption (3.3.2) this implies that the energy functional $\mathcal{I}$ indeed fulfills for all $z_1, z_2 \in B_{\mathcal{Z}}(0, r)$

$$\langle \big[D_z^2\mathcal{I}(t, z_1) - D_z^2\mathcal{I}(t, z_2)\big] v, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq C(r)\|z_1 - z_2\|_{\mathcal{Z}}\|v\|_{\mathcal{Z}}^2 \tag{3.3.16}$$

with a constant $C(r) > 0$ only depending on the radius $r > 0$ and particularly not on the time t. Note that, again, this induces $\mathcal{I}(t, \cdot) \in C^{2,1}_{loc}(\mathcal{Z}; \mathbb{R})$ which is the necessary regularity in order to obtain existence and uniqueness results for differential solutions, see also Definition 1.0.3 and Theorem 2.2.2. Moreover, the special structure of f also allows us to refine the estimate ($\mathcal{I}_{f2}$) as follows:

$$|\langle D_z f(t_1, z_1) - D_z f(t_2, z_2), v\rangle_{\mathcal{V}^*,\mathcal{V}}| = \langle \ell(t_1) - \ell(t_2), v\rangle_{\mathcal{V}^*,\mathcal{V}} \leq |\ell|_{Lip}\,|t_1 - t_2|\,\|v\|_{\mathcal{V}}. \tag{3.3.17}$$

Hence, $f(t, z) = \langle \ell(t), z\rangle_{\mathcal{V}^*,\mathcal{V}}$ fulfills ($\mathcal{I}_{f2}$) with $\nu = |\ell|_{Lip}$. Clearly, assumptions ($\mathcal{I}_{f1}$) and ($\mathcal{I}_{f3}$) as well as (3.3.4) also hold true.

Globally uniformly convex energy functional

We are now in the position to actually start our error analysis. We begin with the case of a uniformly convex energy, see Assumption $\mathbf{GC}_\kappa$. Beside this, we additionally assume:

Assumption $\mathbf{A}_{Lip}$ (Bound on the Lipschitz constant of the driving force)**.**
There exists $\delta \in (0, \kappa]$ so that $|\ell|_{Lip} \leq \kappa - \delta$.

We will drop this condition in Section 3.3.4 for the price of losing the optimal rate of convergence, see Theorem 3.3.17 below.

The basic idea of our convergence proof is to first transform the affine interpolant back into the physical time and then to compare it with the unique differential solution of the rate-independent system (RIS), whose existence is ensured by Theorem 2.2.2. In order to guarantee that the back-transformation is well-defined and fulfills some upper bounds, we need the subsequent lemma.

Lemma 3.3.8. *Let Assumption* $\mathbf{GC}_\kappa$ *and Assumption* $\mathbf{A}_{Lip}$ *be fulfilled. Then it holds that*

$$\|z_{k+1} - z_k\|_{\mathcal{Z}} \leq \frac{\kappa - \delta}{\kappa}\,(t_k - t_{k-1}) \quad \forall\, 1 \leq k \leq N(\tau), \tag{3.3.18}$$

and $(1 - \frac{\kappa-\delta}{\kappa}) = \frac{\delta}{\kappa} \leq \hat{t}'_\tau(s) \leq 1$ *for almost all* $s \in [0, S_\tau]$*. Moreover, it holds* $\hat{N}(\tau) = N(\tau) + 1$*.*

Proof. We argue by induction. Since $z_1 = z_0$ holds by Remark 3.3.1, we have that $\partial\mathcal{R}(z_1 - z_0) + D_z\mathcal{I}(t_0, z_1) = \partial\mathcal{R}(0) + D_z\mathcal{I}(t_0, z_0) \ni 0$ so that Lemma 3.3.7 and Assumption $\mathbf{A}_{Lip}$ imply

$$\|z_2 - z_1\|_{\mathcal{Z}} \leq \frac{|\ell|_{Lip}}{\kappa}\,(t_1 - t_0) \leq \frac{\kappa - \delta}{\kappa}\,(t_1 - t_0),$$

which is (3.3.18) for $k = 1$. Now, let $k \geq 2$ be arbitrary and assume that (3.3.18) holds for $k - 1$, i.e., $\|z_k - z_{k-1}\|_{\mathcal{Z}} \leq \frac{\kappa-\delta}{\kappa}\,(t_{k-1} - t_{k-2}) < \tau$. Consequently, the complementarity conditions in (3.3.6a) and (3.3.7) imply

$$0 \in \partial\mathcal{R}(z_k - z_{k-1}) + D_z\mathcal{I}(t_{k-1}, z_k) \subset \partial\mathcal{R}(0) + D_z\mathcal{I}(t_{k-1}, z_k).$$

Thus, by applying again Lemma 3.3.7 and Assumption $\mathbf{A}_{Lip}$, we obtain (3.3.18) for the next iteration.
For $s \in (0, \tau)$, the lower bound on $\hat{t}'(s)$ follows immediately from $t_1 - t_0 = \tau$, see Remark 3.3.1. For $s > \tau$, it is a direct consequence of (3.3.18), the embedding $\mathcal{Z} \hookrightarrow \mathcal{V}$, and the time update formula (alg$_3$). Finally, by (3.3.18) and the complementarity condition (3.3.6a), we have $\lambda_{N(\tau)+1} = 0$, so that indeed $\hat{N}(\tau) = N(\tau) + 1$ thanks to (3.3.6b). □

We are now in the position to prove our main result on the convergence rate for parametrized solutions. By the lemma above, there exists a unique inverse function $\hat{s}_\tau(t) : [0, T] \mapsto [0, \hat{S}_\tau]$ of $\hat{t}_\tau$. We will then denote by $z_\tau(t) := \hat{z}_\tau(s_\tau(t))$ the retransformed discrete parametrized solution (see also end of the proof of Theorem 3.3.9).

Theorem 3.3.9. *Let Assumption* $\mathbf{GC}_\kappa$ *and Assumption* $\mathbf{A}_{Lip}$ *hold. Moreover, assume that* $\ell \in W^{1,\infty}(0, T; \mathcal{V}^*)$ *with* $\ell' \in BV(0, T; \mathcal{V}^*)$*. Then the sequence* $\{z_\tau\}_{\tau>0}$ *of retransformed discrete parametrized solutions satisfies the a priori error estimate*

$$\|z_\tau(t) - z(t)\|_{\mathcal{Z}} \leq K\,\tau \qquad \forall t \in [0, T], \tag{3.3.19}$$

where $K = K(\alpha, \kappa, \ell, z_0, T, \mathcal{F}, \|A\|_{\mathcal{L}(\mathcal{Z},\mathcal{Z}^*)}) > 0$ *is independent of* τ*.*

Proof. For convenience of the reader we split the rather lengthy proof into eight parts, which are as follows:

0. First, we will see that, due to the uniform convexity of the energy, (RIS) even admits a unique differential solution and not only a parametrized one.

1. Based on Lemma 3.3.8, we can transform the piecewise affine interpolants introduced above into the physical time. This allows us to compare the discrete solution with the exact (differential) solution, which, of course, also "lives" in the physical time. The error analysis,

however, uses a slightly different piecewise affine interpolant, denoted by $\tilde{z}_\tau$ providing a certain shift in the time steps.

2. In analogy to [MT04], we introduce a quantity $\gamma(t)$, which dominates the pointwise error $\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}}$. This error measure enables us to deal with uniformly convex energy functionals instead of just quadratic and coercive ones.

3. The error measure is essentially estimated by two contributions, denoted by $E(t)$ and $R(t)$. Both contributions depend only differences of $D_z\mathcal{I}$ evaluated at different time points and different discrete solutions.

4./5. $E(t)$ and $R(t)$ are estimated by using the smoothness properties of $\mathcal{F}$ and the load ℓ. In addition, the uniform convexity of $\mathcal{I}$ plays an essential role for the estimate of R. In this way, one obtains an estimate of $\mathcal{O}(\tau^2)$ for the L^1-norms of E and R.

6. Together with Gronwall's lemma, this estimate yields a bound of $\mathcal{O}(\tau)$ for the error indicator γ and thus also for the error $\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}}$.

7. Finally, we relate $\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}}$ with the auxiliary interpolant $\tilde{z}_\tau$ to the "true error" containing the "correct" interpolant $z_\tau = \hat{z}_\tau \circ s_\tau$ as introduced above.

Step 0: Differential solution

First of all, we want to exploit Theorem 2.2.2 in order to obtain a unique differential solution of the rate-independent system (see also Remark 2.2.3). To this end, we need to verify assumptions (2.2.2) and (2.2.3b). Clearly, (2.2.2) is satisfied due to the structure of $\mathcal{I}$ and the assumptions on ℓ, see also (3.3.17). Moreover, since $\mathcal{I}$ complies with the requirements in Section 3.1, the (strong,weak)-weak convergence of $D_z\mathcal{I}$, i.e., (2.2.3b), is also valid in this case; cf. (3.1.1). Hence, there exists a unique (differential) solution $z \in W^{1,\infty}(0,T;\mathcal{Z})$ of the rate-independent system. In particular, it holds for almost all $t \in [0,T]$ that $0 \in \partial\mathcal{R}(z'(t)) + D_z\mathcal{I}(t, z(t))$, which can be reformulated as (see (2.2.10)):

$$\forall v \in \mathcal{Z}: \quad \mathcal{R}(v) \geq \langle -D_z\mathcal{I}(t, z(t)), v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \qquad \forall t \in [0,T], \tag{3.3.20a}$$

$$\mathcal{R}(z'(t)) = \langle -D_z\mathcal{I}(t, z(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \qquad \text{f.a.a. } t \in [0,T]. \tag{3.3.20b}$$

Since $z \in W^{1,\infty}(0,T;\mathcal{Z})$, it additionally holds that

$$\|z'(t)\|_{\mathcal{Z}} \leq C \quad \text{f.a.a. } t \in [0,T]. \tag{3.3.21}$$

Step 1: Construction of interpolants in the physical time

Given $t \in [t_{k-1}, t_k)$ with $k \leq N(\tau)$, we define the following affine interpolant

$$\tilde{z}_\tau(t) = z_k + \frac{t - t_{k-1}}{t_k - t_{k-1}}(z_{k+1} - z_k). \tag{3.3.22}$$

Note that $[t_{k-1}, t_k)$ is nonempty and that $\lambda_k = 0$ due to Lemma 3.3.8. Thus, from the stationarity condition, i.e., (3.3.7), we know that $0 \in \partial\mathcal{R}(\tilde{z}_\tau'(t)) + D_z\mathcal{I}(t_k, z_{k+1})$. Analogously to *Step 0*, this can be reformulated as

$$\forall v \in \mathcal{Z}: \qquad \mathcal{R}(v) \geq \langle -D_z\mathcal{I}(t_k, z_{k+1}), v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \qquad \forall k \in \{0,\dots,N(\tau)\}, \tag{3.3.23a}$$

$$\mathcal{R}(\tilde{z}_\tau'(t)) = \langle -D_z\mathcal{I}(t_k, z_{k+1}), \tilde{z}_\tau'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \qquad \text{f.a.a. } t \in [0,T]. \tag{3.3.23b}$$

Exploiting Lemma 3.3.8, we additionally have

$$\|\tilde{z}_\tau'(t)\|_{\mathcal{Z}} \leq C \quad \text{f.a.a. } t \in [0,T]. \tag{3.3.24}$$

Step 2: Introduction of an error measure

We now basically follow along the lines of [MT04, Thm 7.4], but we have to adapt the underlying analysis at some points. Therefore we present the arguments in detail. Let us define

$$\gamma(t) := \langle D_z\mathcal{I}(t, \tilde{z}_\tau(t)) - D_z\mathcal{I}(t, z(t)), \tilde{z}_\tau(t) - z(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\,. \tag{3.3.25}$$

Due to the κ-uniform convexity of $\mathcal{I}(t,\cdot)$, we have

$$\gamma(t) \geq \kappa\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}}^2, \tag{3.3.26}$$

so that γ measures the discretization error. In full analogy to [MT04, Thm 7.4], we can estimate (see Appendix A.4)

$$\dot{\gamma}(t) \leq C\,\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}}^2 + 2\langle D_z\mathcal{I}(t, \tilde{z}_\tau(t)) - D_z\mathcal{I}(t, z(t)), \tilde{z}_\tau'(t) - z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \tag{3.3.27}$$

for almost all $t \in [0,T]$. We split the second term into two parts, namely

$$\begin{aligned} &e_1(t) := 2\,\langle D_z\mathcal{I}(t, z(t)) - D_z\mathcal{I}(t, \tilde{z}_\tau(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ \text{and} \quad &e_2(t) := 2\,\langle D_z\mathcal{I}(t, \tilde{z}_\tau(t)) - D_z\mathcal{I}(t, z(t)), \tilde{z}_\tau'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}. \end{aligned}$$

Step 3: Estimates for the error e_i

Let again $k \leq N(\tau)$ and $t \in [t_{k-1}, t_k)$ be arbitrary. First, observe that due to the convexity of $\partial\mathcal{R}(0)$, it holds for

$$\theta(t) = \frac{t - t_{k-1}}{t_k - t_{k-1}}$$

that

$$-(1-\theta(t))\,\xi_{k-1} - \theta(t)\,\xi_k \in \partial\mathcal{R}(0)$$

with $\xi_{k-1} := D_z\mathcal{I}(t_{k-1}, z_k)$ and $\xi_k := D_z\mathcal{I}(t_k, z_{k+1})$. From the characterization of $\partial\mathcal{R}(0)$, we infer $\mathcal{R}(v) \geq -\langle(1-\theta(t))\,\xi_{k-1} + \theta(t)\,\xi_k, v\rangle_{\mathcal{Z}^*,\mathcal{Z}}$ for all $v \in \mathcal{Z}$. Inserting herein $v = z'(t)$ and subtracting (3.3.20b), we can estimate

$$\begin{aligned} e_1(t) &= 2\,\langle D_z\mathcal{I}(t, z(t)) - (1-\theta(t))\,\xi_{k-1} - \theta(t)\,\xi_k, z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\qquad + 2\langle(1-\theta(t))\,\xi_{k-1} + \theta(t)\,\xi_k - D_z\mathcal{I}(t, \tilde{z}_\tau(t)), z'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\leq 2\,\|(1-\theta(t))\,\xi_{k-1} + \theta(t)\,\xi_k - D_z\mathcal{I}(t, \tilde{z}_\tau(t))\|_{\mathcal{Z}^*}\|z'(t)\|_{\mathcal{Z}} \end{aligned} \tag{3.3.28}$$

for almost all $t \in [t_{k-1}, t_k)$.
Next, we turn to the term e_2. Similarly, we take $v = \tilde{z}_\tau'(t)$ in (3.3.20a) and subtract (3.3.23b) to obtain $0 \geq \langle D_z\mathcal{I}(t_k, z_{k+1}) - D_z\mathcal{I}(t, z(t)), \tilde{z}_\tau'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}$, from which we deduce

$$\begin{aligned}
e_2(t) &\leq 2\,\langle D_z\mathcal{I}(t, \tilde{z}_\tau(t)) - D_z\mathcal{I}(t_k, z_{k+1}), \tilde{z}_\tau'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&\leq 2\langle D_z\mathcal{I}(t, \tilde{z}_\tau(t)) - (1-\theta(t))\,\xi_{k-1} - \theta(t)\,\xi_k, \tilde{z}_\tau'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&\qquad + 2\langle (1-\theta(t))\,\xi_{k-1} + \theta(t)\,\xi_k - D_z\mathcal{I}(t_k, z_{k+1}), \tilde{z}_\tau'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&\leq 2\,\|(1-\theta(t))\,\xi_{k-1} + \theta(t)\,\xi_k - D_z\mathcal{I}(t, \tilde{z}_\tau(t))\|_{\mathcal{Z}^*}\|\tilde{z}_\tau'(t)\|_{\mathcal{Z}} \\
&\qquad + 2(1-\theta(t))\langle D_z\mathcal{I}(t_{k-1}, z_k) - D_z\mathcal{I}(t_k, z_{k+1}), \tilde{z}_\tau'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}.
\end{aligned} \tag{3.3.29}$$

Next, let us define

$$E(t) := \|(1-\theta(t))\,\xi_{k-1} + \theta(t)\,\xi_k - D_z\mathcal{I}(t, \tilde{z}_\tau(t))\|_{\mathcal{Z}^*} \tag{3.3.30}$$

$$\text{and} \quad R(t) := 2(1-\theta(t))\langle D_z\mathcal{I}(t_{k-1}, z_k) - D_z\mathcal{I}(t_k, z_{k+1}), \tilde{z}_\tau'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}. \tag{3.3.31}$$

Then we insert (3.3.30) and (3.3.31) into (3.3.28) and (3.3.29). The resulting estimates for e_1 and e_2 are, in turn, inserted in (3.3.27), which, together with the boundedness of $\|z'(t)\|_{\mathcal{Z}}$ and $\|\tilde{z}_\tau'(t)\|_{\mathcal{Z}}$ by (3.3.21) and (3.3.24), yields

$$\gamma'(t) \leq C(\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}}^2 + E(t) + R(t)). \tag{3.3.32}$$

The aim now is to provide an estimate for the additional error terms E and R, which is done next.

Step 4: Estimate for $E(t)$
The particular structure of $\mathcal{I}$, together with the linearity of A and the definition of $\tilde{z}_\tau$ gives

$$\begin{aligned}
E(t) &\leq \|(1-\theta(t))D_z\mathcal{F}(z_k) + \theta(t)D_z\mathcal{F}(z_{k+1}) - D_z\mathcal{F}\left((1-\theta(t))z_k + \theta(t)z_{k+1}\right)\|_{\mathcal{Z}^*} \\
&\qquad + \|(1-\theta(t))\ell(t_{k-1}) + \theta(t)\ell(t_k) - \ell(t)\|_{\mathcal{Z}^*} \\
&=: I_1(t) + I_2(t).
\end{aligned}$$

Exploiting the regularity of $\mathcal{F}$, in particular (3.3.2), we can estimate

$$\begin{aligned}
I_1(t) &= \Big\|\theta(t)(D_z\mathcal{F}(z_{k+1}) - D_z\mathcal{F}(z_k)) \\
&\qquad\qquad - \theta(t)\int_0^1 D_z^2\mathcal{F}(z_k + s\theta(t)(z_{k+1} - z_k))[z_{k+1} - z_k]\,\mathrm{d}s\Big\|_{\mathcal{Z}^*} \\
&\leq \theta(t)\|z_{k+1} - z_k\|_{\mathcal{Z}} \\
&\qquad \int_0^1 \big\|D_z^2\mathcal{F}(z_k + s(z_{k+1} - z_k)) - D_z^2\mathcal{F}(z_k + s\theta(t)(z_{k+1} - z_k))\big\|_{\mathcal{L}(\mathcal{Z},\mathcal{L}(\mathcal{Z},\mathcal{Z}^*))}\,\mathrm{d}s \\
&\leq C\|z_{k+1} - z_k\|_{\mathcal{Z}}^2,
\end{aligned}$$

where we also used $\theta(t) \in [0,1]$ and the boundedness of the iterates z_k in $\mathcal{Z}$ independent of τ from

(3.3.5). For I_2, we proceed similarly by exploiting the regularity of ℓ:

$$I_2(t) \le \int_{t_{k-1}}^{t} \left\| \frac{\ell(t_k) - \ell(t_{k-1})}{t_k - t_{k-1}} - \ell'(s) \right\|_{\mathcal{V}^*} \mathrm{d}s \le \tau \|\ell'\|_{BV(t_{k-1},t_k;\mathcal{V}^*)}.$$

Since $\|z_{k+1} - z_k\|_{\mathcal{Z}} \le C\tau$ by Lemma 3.3.8, the above estimates for $I_1(t)$ and $I_2(t)$ imply for all $t \in [t_{k-1}, t_k)$ that $E(t) \le C\tau^2 + \tau\|\ell'\|_{BV(t_{k-1},t_k;\mathcal{V}^*)}$. Now integrating E yields

$$\int_0^T E(t)\,\mathrm{d}t \le C\tau^2 + \tau^2\|\ell'\|_{BV(0,T;\mathcal{V}^*)} \le C\tau^2. \tag{3.3.33}$$

Step 5: Estimate for $R(t)$

First, we abbreviate $\mathcal{E}(z) := \langle Az, z\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \mathcal{F}(z)$ so that $\mathcal{I}(t,z) = \mathcal{E}(z) - \langle \ell(t), z\rangle_{\mathcal{V}^*,\mathcal{V}}$. Moreover, we set

$$\Delta t_k := t_k - t_{k-1}, \qquad \mathrm{d}_\tau \ell_k := \frac{\ell(t_k) - \ell(t_{k-1})}{\Delta t_k}, \qquad k = 1, ..., N(\tau),$$
$$\mathrm{d}_\tau z_{k+1} := \frac{z_{k+1} - z_k}{\Delta t_k}, \quad \mathrm{d}_\tau D_z\mathcal{E}_{k+1} := \frac{D_z\mathcal{E}(z_{k+1}) - D_z\mathcal{E}(z_k)}{\Delta t_k}, \quad k = 1, ..., N(\tau),$$

as well as $\mathrm{d}_\tau \ell_0 = 0$, $\mathrm{d}_\tau z_1 = 0$, and $\mathrm{d}_\tau D_z\mathcal{E}_1 = 0$. Note that by Remark 3.3.2 and Lemma 3.3.8, the quantities $\mathrm{d}_\tau z_{N(\tau)+1}$ and $\mathrm{d}_\tau \mathcal{E}_{N(\tau)+1}$ are well-defined. By Lemma 3.3.8, we have

$$\|\mathrm{d}_\tau z_k\|_{\mathcal{Z}} \le C \quad \forall k = 1, ..., N(\tau) + 1. \tag{3.3.34}$$

Now, on account of $-D_z\mathcal{I}(t_{k-1}, z_k) \in \partial\mathcal{R}(z_k - z_{k-1})$, we deduce from (3.3.23a) tested with $z_k - z_{k-1}$ that $0 \ge \langle D_z\mathcal{I}(t_{k-1}, z_k) - D_z\mathcal{I}(t_k, z_{k+1}), z_k - z_{k-1}\rangle_{\mathcal{Z}^*,\mathcal{Z}}$. Inserting the definitions of $\tilde{z}$ and $\theta(t)$, we thus obtain for $t \in [t_{k-1}, t_k)$ that

$$\begin{aligned}
R(t) &= 2(1 - \theta(t))\langle D_z\mathcal{I}(t_{k-1}, z_k) - D_z\mathcal{I}(t_k, z_{k+1}), \tilde{z}_\tau'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&= 2(t_k - t)\langle (\Delta t_k)^{-1}[D_z\mathcal{I}(t_{k-1}, z_k) - D_z\mathcal{I}(t_k, z_{k+1})], \mathrm{d}_\tau z_{k+1} - \mathrm{d}_\tau z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&\qquad + 2(t_k - t)\langle (\Delta t_k)^{-1}[D_z\mathcal{I}(t_{k-1}, z_k) - D_z\mathcal{I}(t_k, z_{k+1})], \mathrm{d}_\tau z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&\le 2(t_k - t)\langle (\Delta t_k)^{-1}[D_z\mathcal{I}(t_{k-1}, z_k) - D_z\mathcal{I}(t_k, z_{k+1})], \mathrm{d}_\tau z_{k+1} - \mathrm{d}_\tau z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&= 2(t_k - t)\langle -\mathrm{d}_\tau D_z\mathcal{E}_{k+1} + \mathrm{d}_\tau \ell_k, \mathrm{d}_\tau z_{k+1} - \mathrm{d}_\tau z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}}.
\end{aligned}$$

Integrating then gives

$$\begin{aligned}
\int_0^T R(t)\,\mathrm{d}t &\le \sum_{k=1}^{N(\tau)} (\Delta t_k)^2 \langle -\mathrm{d}_\tau D_z\mathcal{E}_{k+1} + \mathrm{d}_\tau \ell_k, \mathrm{d}_\tau z_{k+1} - \mathrm{d}_\tau z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&\le \tau^2 \sum_{k=1}^{N(\tau)} \langle -\mathrm{d}_\tau D_z\mathcal{E}_{k+1}, \mathrm{d}_\tau z_{k+1} - \mathrm{d}_\tau z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \langle \mathrm{d}_\tau \ell_k, \mathrm{d}_\tau z_{k+1} - \mathrm{d}_\tau z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}}.
\end{aligned} \tag{3.3.35}$$

For the terms involving ℓ we have

$$\begin{aligned}
&\sum_{k=1}^{N(\tau)} \langle \mathrm{d}_\tau \ell_k, \mathrm{d}_\tau z_{k+1} - \mathrm{d}_\tau z_k \rangle_{\mathcal{V}^*,\mathcal{V}} \\
&\quad = \sum_{k=1}^{N(\tau)} \langle \mathrm{d}_\tau \ell_k, \mathrm{d}_\tau z_{k+1} \rangle_{\mathcal{V}^*,\mathcal{V}} - \langle \mathrm{d}_\tau \ell_k - \mathrm{d}_\tau \ell_{k-1}, \mathrm{d}_\tau z_k \rangle_{\mathcal{V}^*,\mathcal{V}} - \langle \mathrm{d}_\tau \ell_{k-1}, \mathrm{d}_\tau z_k \rangle_{\mathcal{V}^*,\mathcal{V}},
\end{aligned}$$

where we used $\mathrm{d}_\tau \ell_0 = 0$. The second term is estimated analogously to I_2, exploiting the regularity of ℓ as well as the boundedness of $\|\mathrm{d}_\tau z_k\|_{\mathcal{V}}$ from (3.3.34), which yields

$$\begin{aligned}
&|\langle \mathrm{d}_\tau \ell_k - \mathrm{d}_\tau \ell_{k-1}, \mathrm{d}_\tau z_k \rangle_{\mathcal{V}^*,\mathcal{V}}| \\
&\quad = \Big| \int_0^1 \langle \ell'(t_{k-1} + s(t_k - t_{k-1})) - \ell'(t_{k-2} + s(t_{k-1} - t_{k-2}))\,\mathrm{d}s, \mathrm{d}_\tau z_k \rangle_{\mathcal{V}^*,\mathcal{V}} \Big| \\
&\quad \le \|\ell'\|_{BV(t_{k-2},t_k;\mathcal{V}^*)} \|\mathrm{d}_\tau z_k\|_{\mathcal{V}} \\
&\quad \le C \|\ell'\|_{BV(t_{k-2},t_k;\mathcal{V}^*)}.
\end{aligned}$$

Hence, thanks to $\mathrm{d}_\tau \ell_0 = 0$ and (3.3.34), we have

$$\begin{aligned}
&\sum_{k=1}^{N(\tau)} \langle \mathrm{d}_\tau \ell_k, d_t z_{k+1} - d_t z_k \rangle_{\mathcal{V}^*,\mathcal{V}} \\
&\quad \le \sum_{k=1}^{N(\tau)} \langle \mathrm{d}_\tau \ell_k, \mathrm{d}_\tau z_{k+1} \rangle_{\mathcal{V}^*,\mathcal{V}} - \langle \mathrm{d}_\tau \ell_{k-1}, \mathrm{d}_\tau z_k \rangle_{\mathcal{V}^*,\mathcal{V}} + C \|\ell'\|_{BV(t_{k-2},t_k;\mathcal{V}^*)} \\
&\quad \le \langle \mathrm{d}_\tau \ell_{N(\tau)}, \mathrm{d}_\tau z_{N(\tau)+1} \rangle_{\mathcal{V}^*,\mathcal{V}} + 2C \|\ell'\|_{BV(0,T;\mathcal{V}^*)} \qquad (3.3.36) \\
&\quad \le C(|\ell|_{Lip} + \|\ell'\|_{BV(0,T;\mathcal{V}^*)}). \qquad (3.3.37)
\end{aligned}$$

Now, for the terms involving $D_z\mathcal{E}$, we first calculate

$$\begin{aligned}
\langle \mathrm{d}_\tau D_z \mathcal{E}_{k+1}, \mathrm{d}_\tau z_k \rangle_{\mathcal{Z}^*,\mathcal{Z}} &= \Big\langle \frac{D_z\mathcal{E}(z_{k+1}) - D_z\mathcal{E}(z_k)}{t_k - t_{k-1}}, \mathrm{d}_\tau z_k \Big\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&= \int_0^1 \langle D_z^2 \mathcal{E}(z_k + s(z_{k+1} - z_k))[\mathrm{d}_\tau z_{k+1}], \mathrm{d}_\tau z_k \rangle_{\mathcal{Z}^*,\mathcal{Z}}\,\mathrm{d}s.
\end{aligned}$$

Since $D_z^2\mathcal{E}$ is symmetric, we obtain

$$\begin{aligned}
&2\int_0^1 \langle D_z^2 \mathcal{E}(z_k + s(z_{k+1} - z_k))[\mathrm{d}_\tau z_{k+1}], \mathrm{d}_\tau z_k \rangle_{\mathcal{Z}^*,\mathcal{Z}}\,\mathrm{d}s \\
&= -\int_0^1 \langle D_z^2 \mathcal{E}(z_k + s(z_{k+1} - z_k))[\mathrm{d}_\tau z_{k+1} - \mathrm{d}_\tau z_k], \mathrm{d}_\tau z_{k+1} - \mathrm{d}_\tau z_k \rangle_{\mathcal{Z}^*,\mathcal{Z}}\,\mathrm{d}s \\
&\quad + \int_0^1 \langle D_z^2 \mathcal{E}(z_k + s(z_{k+1} - z_k))[\mathrm{d}_\tau z_{k+1}], \mathrm{d}_\tau z_{k+1} \rangle_{\mathcal{Z}^*,\mathcal{Z}}\,\mathrm{d}s \\
&\quad + \int_0^1 \langle (D_z^2 \mathcal{E}(z_k + s(z_{k+1} - z_k)) - D_z^2\mathcal{E}(z_{k-1} + s(z_k - z_{k-1})))[\mathrm{d}_\tau z_k], \mathrm{d}_\tau z_k \rangle_{\mathcal{Z}^*,\mathcal{Z}}\,\mathrm{d}s \\
&\quad + \int_0^1 \langle D_z^2 \mathcal{E}(z_{k-1} + s(z_k - z_{k-1}))[\mathrm{d}_\tau z_k], \mathrm{d}_\tau z_k \rangle_{\mathcal{Z}^*,\mathcal{Z}}\,\mathrm{d}s.
\end{aligned}$$

Thus, thanks to the convexity and regularity of $\mathcal{E}$, we have

$$\begin{aligned}\langle \mathrm{d}_\tau D_z\mathcal{E}_{k+1}, \mathrm{d}_\tau z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq \frac{1}{2}\langle \mathrm{d}_\tau D_z\mathcal{E}_k, \mathrm{d}_\tau z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \frac{1}{2}\langle \mathrm{d}_\tau D_z\mathcal{E}_{k+1}, \mathrm{d}_\tau z_{k+1}\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ + \frac{1}{2}C\|\mathrm{d}_\tau z_k\|_{\mathcal{Z}}^2(\|z_{k+1}-z_k\|_{\mathcal{Z}} + \|z_k - z_{k-1}\|_{\mathcal{Z}}).\end{aligned}$$

Combining (3.3.9) with Lemma 3.3.8 for $k = N(\tau)$, we clearly have $\sum_{k=1}^{N(\tau)}\|z_{k+1}-z_k\|_{\mathcal{Z}} \leq C_\Sigma + \tau$. Thus, exploiting (3.3.34), we eventually end up with

$$\begin{aligned}&\sum_{k=1}^{N(\tau)}\langle \mathrm{d}_\tau D_z\mathcal{E}_{k+1}, \mathrm{d}_\tau z_k\rangle - \langle \mathrm{d}_\tau D_z\mathcal{E}_{k+1}, \mathrm{d}_\tau z_{k+1}\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\quad\leq \frac{1}{2}\sum_{k=1}^{N(\tau)}\{\langle \mathrm{d}_\tau D_z\mathcal{E}_k, \mathrm{d}_\tau z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} - \langle \mathrm{d}_\tau D_z\mathcal{E}_{k+1}, \mathrm{d}_\tau z_{k+1}\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\qquad + C\|\mathrm{d}_\tau z_k\|_{\mathcal{Z}}^2(\|z_{k+1}-z_k\|_{\mathcal{Z}} + \|z_k - z_{k-1}\|_{\mathcal{Z}})\} \\ &\quad\leq \frac{1}{2}C(2C_\Sigma + \tau) + \frac{1}{2}\langle \mathrm{d}_\tau D_z\mathcal{E}_1, \mathrm{d}_\tau z_1\rangle_{\mathcal{Z}^*,\mathcal{Z}} - \frac{1}{2}\langle \mathrm{d}_\tau D_z\mathcal{E}_{N(\tau)+1}, \mathrm{d}_\tau z_{N(\tau)+1}\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq C,\end{aligned}$$

wherein the last estimate is due to Remark 3.3.1, i.e., $\langle \mathrm{d}_\tau D_z\mathcal{E}_1, \mathrm{d}_\tau z_1\rangle = 0$, and the convexity of $\mathcal{E}$, that is, $\langle \mathrm{d}_\tau D_z\mathcal{E}_{N(\tau)+1}, \mathrm{d}_\tau z_{N(\tau)+1}\rangle \geq 0$. Inserting this together with (3.3.37) into (3.3.35) and combining the resulting estimate with (3.3.33), overall we have shown that

$$\int_0^T E(t)\,\mathrm{d}t + \int_0^T R(t)\,\mathrm{d}t \leq C\tau^2. \tag{3.3.38}$$

With this inequality at hand, we return to the estimate in (3.3.32) in order to, eventually, obtain a bound for the error measure γ.

Step 6: Obtain convergence rate by Gronwall's lemma
Exploiting the fact that $\gamma(t)/\kappa \geq \|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}}^2$ in (3.3.32), one obtains

$$\gamma'(t) \leq C(\gamma(t) + E(t) + R(t))\,.$$

Integrating this and using Gronwall's inequality from Lemma A.4.1 as well as the estimates (3.3.38) on E and R yields

$$\gamma(t) \leq (\gamma(0) + C\tau^2)\exp^{Ct} \leq C(\gamma(0) + \tau^2)\,.$$

Due to $\tilde{z}_\tau(0) = z(0) = z_0$, we have $\gamma(0) = 0$. Using another time the κ-uniform convexity of $\mathcal{I}$, we therefore finally obtain

$$\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}}^2 \leq \gamma(t)/\kappa \leq C\tau^2. \tag{3.3.39}$$

Step 7: Comparing interpolants
By $\hat{z}_\tau$ we denote the affine interpolation of the discrete approximations with step size τ in the artificial time, see (3.3.10). From Lemma 3.3.8, we conclude that $\hat{t}_\tau(s)$ is monotonically increasing and $\hat{t}'_\tau(s) \geq 1 - \frac{\kappa-\delta}{\kappa}$ holds almost everywhere in $[0, S_\tau]$ (for the definition of S_τ see (3.3.11)). Thus, there exists a unique inverse function $s_\tau : [0,T] \to [0, S_\tau]$ with $1 \leq s'_\tau(t) \leq \frac{1}{1-\frac{\kappa-\delta}{\kappa}}$ almost everywhere in $[0,T]$. Given this inverse, one can define z_τ as the retransformed affine interpolant,

i.e.,

$$z_\tau(t) := \hat{z}_\tau(s_\tau(t)).$$

By elementary calculations, one verifies that the explicit formula for z_τ reads as follows:

$$z_\tau(t) = z_{k-1} + \frac{t - t_{k-1}}{t_k - t_{k-1}}(z_k - z_{k-1}), \quad t \in [t_{k-1}, t_k),$$

i.e., z_τ is just the affine interpolant in the physical time. Comparing z_τ with $\tilde{z}_\tau$ from (3.3.22) results in

$$\begin{aligned}\|z_\tau(t) - \tilde{z}_\tau(t)\|_{\mathcal{Z}} &= \|z_{k-1} + \theta(t)(z_k - z_{k-1}) - z_k - \theta(t)(z_{k+1} - z_k)\|_{\mathcal{Z}} \\ &\leq (1 - \theta(t))\|z_{k-1} - z_k\|_{\mathcal{Z}} + \theta(t)\|z_k - z_{k+1}\|_{\mathcal{Z}} \leq \tau,\end{aligned}$$

where we exploited (3.3.18) once more. Since $k \leq N(\tau)$ was arbitrary, we have $\|z_\tau(t) - \tilde{z}_\tau(t)\|_{\mathcal{Z}} \leq \tau$ for all $t \in [0, T]$. In combination with (3.3.39), this finally gives

$$\|z_\tau(t) - z(t)\|_{\mathcal{Z}} \leq K\tau,$$

which is the desired result. A careful analysis of the used estimates and the corresponding constants yields that K provides the claimed dependencies. □

Some remarks and comments concerning the assertion of Theorem 3.3.9 and its proof are in order.

Remark 3.3.10. In preparation for the following section, we note that the uniform convexity of the energy is only needed at three places in the above analysis: first for the estimate in (3.3.18); second for the lower bound on γ in (3.3.26); and third for the inequality

$$\int_0^1 \langle D_z^2\mathcal{E}(z_k + s(z_{k+1} - z_k))[\mathrm{d}_\tau z_{k+1} - \mathrm{d}_\tau z_k], \mathrm{d}_\tau z_{k+1} - \mathrm{d}_\tau z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}}\, \mathrm{d}s \geq 0. \tag{3.3.40}$$

However, estimates (3.3.18) and (3.3.40) remain valid if $\mathcal{I}(t_k, \cdot)$ is only κ-uniformly convex on a ball $B_{\mathcal{Z}}(z(t_k), \Delta)$ with radius $\Delta > \tau > 0$ and $z_k, z_{k+1} \in B_{\mathcal{Z}}(z(t_k), \Delta)$. To see this, note that (3.3.18) follows from estimate (3.3.12), see proof of Lemma 3.3.8, which itself is a consequence of $\langle D_z\mathcal{I}(t_k, z_{k+1}) - D_z\mathcal{I}(t_k, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \kappa\|z_{k+1} - z_k\|_{\mathcal{Z}}^2$. This inequality, just as inequality (3.3.40), only requires that z_k and z_{k+1} lay in a region of uniform convexity of $\mathcal{I}$. The estimate on γ finally necessitates that $\tilde{z}_\tau(t) \in B_{\mathcal{Z}}(z(t), \Delta)$ and that $\mathcal{I}$ is uniformly convex on $B_{\mathcal{Z}}(z(t), \Delta)$ for all $t \in [0, T]$; cf. the definition of γ in (3.3.25).

Remark 3.3.11. In view of the regularity of the differential solution, i.e., $z \in W^{1,\infty}(0, T; \mathcal{Z})$, the convergence rate of $\mathcal{O}(\tau)$ in Theorem 3.3.9 can be regarded as optimal, since the piecewise affine interpolation of the solution does not yield a better convergence rate.

Remark 3.3.12. It is to be expected that a spatial discretization can also be included in the above a priori estimates, following, e.g., the lines of [MPPS10].

Locally uniformly convex energy functional

As already mentioned in the introduction, the local incremental minimization algorithm is actually not necessary if the energy is globally uniformly convex. In this case, one could also use the global incremental minimization scheme, which is easier to implement, since the additional inequality constraint in (3.0.1a) is omitted. The situation changes, however, if the energy is no longer globally uniformly convex, but only locally around a given evolution z. Then the local incremental minimization scheme still approximates the (local) solution with optimal order (provided that $|\ell|_{Lip}$ is not too large), while the global scheme might fail to converge to this solution, as we will demonstrate by means of a numerical example in Section 4.2.2. Our precise notion of local uniform convexity is as follows:

Assumption LC$_\kappa$ (Local κ-uniform convexity)**.** *We call $\mathcal{I}$* locally κ-uniformly convex *around $z : [0,T] \to \mathcal{Z}$ if there exist $\kappa, \Delta > 0$, independent of t, such that $\mathcal{I}(t,\cdot)$ is κ-uniformly convex on $\overline{B_\mathcal{Z}(z(t),\Delta)}$ for all $t \in [0,T]$, i.e.,*

$$\langle D_z^2\mathcal{I}(t,\tilde{z})v, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \kappa\|v\|_\mathcal{Z}^2 \qquad \forall \tilde{z} \in \overline{B_\mathcal{Z}(z(t),\Delta)},\ v \in \mathcal{Z}. \tag{3.3.41}$$

Note that local uniform convexity is always referred to an evolution z. The Assumption **LC$_\kappa$** especially implies that

$$\langle D_z\mathcal{I}(t,z_2) - D_z\mathcal{I}(t,z_1), z_2 - z_1\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \kappa\|z_2 - z_1\|_\mathcal{Z}^2 \qquad \forall z_1, z_2 \in \overline{B_\mathcal{Z}(z(t),\Delta)} \tag{3.3.42}$$

holds. Indeed, using (3.3.41), we obtain

$$\begin{aligned}&\langle D_z\mathcal{I}(t,z_2) - D_z\mathcal{I}(t,z_1), z_2 - z_1\rangle_{\mathcal{Z}^*,\mathcal{Z}}\\ &\quad= \int_0^1 \langle D_z^2\mathcal{I}(t, z_1 + s(z_2 - z_1))[z_2 - z_1], z_2 - z_1\rangle_{\mathcal{Z}^*,\mathcal{Z}}\mathrm{d}s \geq \kappa\|z_2 - z_1\|_\mathcal{Z}^2,\end{aligned}$$

where we used the fact that $z_1 + s(z_2 - z_1) \in \overline{B_\mathcal{Z}(z(t),\Delta)}$ for all $s \in [0,1]$. Now, in order to prove a convergence rate in the local uniform convex case, we again have to estimate the difference of iterates in the $\mathcal{Z}$-norm. Since it is not a priori clear that the iterate remains in the neighborhood of convexity of $\mathcal{I}$, we need to alter the proof of Lemma 3.3.7.

Lemma 3.3.13. *Let $0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(t_{k-1}, z_k)$ for some $k \in \mathbb{N}$. Then $\|z_{k+1} - z_k\|_\mathcal{Z} \leq C_{loc}\,\tau$ for some constant $C_{loc} = C_{loc}(\mathcal{F}, \alpha, |\ell|_{Lip}) > 0$.*

Proof. Let $k \in \mathbb{N}$ be given. From (3.3.15) we know that

$$\begin{aligned}0 \geq\ &\langle D_z\mathcal{I}(t_k, z_{k+1}) - D_z\mathcal{I}(t_k, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}}\\ &+ \langle D_z\mathcal{I}(t_k, z_k) - D_z\mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} + (\lambda_{k+1} - \lambda_k)\tau^2.\end{aligned}$$

Since $0 \in \partial\mathcal{R}(0) + D_z\mathcal{I}(t_{k-1}, z_k)$ holds by assumption, (3.3.8) implies $\lambda_k = 0$. Inserting the definition of $\mathcal{I}$ and exploiting Remark 3.3.4, we can thus further estimate

$$
\begin{aligned}
0 &\geq \langle A(z_{k+1} - z_k), z_{k+1} - z_k \rangle_{\mathcal{Z}^*,\mathcal{Z}} + \langle D_z\mathcal{F}(z_{k+1}) - D_z\mathcal{F}(z_k), z_{k+1} - z_k \rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&\quad + \langle \ell(t_{k-1}) - \ell(t_k), z_{k+1} - z_k \rangle_{\mathcal{V}^*,\mathcal{V}} + \lambda_{k+1}\tau^2 \\
&\geq \alpha \|z_{k+1} - z_k\|_{\mathcal{Z}}^2 - C_{\mathcal{F}} \|z_{k+1} - z_k\|_{\mathcal{Z}} \|z_{k+1} - z_k\|_{\mathcal{V}} - |\ell|_{Lip}(t_k - t_{k-1}) \|z_{k+1} - z_k\|_{\mathcal{V}}.
\end{aligned}
$$

Therefore, by applying the generalized Young-inequality, it follows from the constraint in (3.0.1a) that

$$
\begin{aligned}
0 &\geq \alpha \|z_{k+1} - z_k\|_{\mathcal{Z}}^2 - \frac{\alpha}{2} \|z_{k+1} - z_k\|_{\mathcal{Z}}^2 - C_{\mathcal{F},\alpha} \|z_{k+1} - z_k\|_{\mathcal{V}}^2 - |\ell|_{Lip}\, \tau^2 \\
&\geq \frac{\alpha}{2} \|z_{k+1} - z_k\|_{\mathcal{Z}}^2 - C_{\mathcal{F},\alpha}\tau^2 - |\ell|_{Lip}\, \tau^2,
\end{aligned}
$$

so that, indeed, $C_{loc}^2 \tau^2 \geq \|z_{k+1} - z_k\|_{\mathcal{Z}}^2$ with $C_{loc}^2 = \frac{2}{\alpha}(C_{\mathcal{F},\alpha} + |\ell|_{Lip})$. □

With this at hand, we can now show an a priori estimate in the case of an energy functional, which is only locally uniformly convex around a differential solution.

Theorem 3.3.14. *Let $z \in W^{1,\infty}(0,T;\mathcal{Z})$ be a (differential) solution. Furthermore, let $\mathcal{I}$ be locally κ-uniformly convex around z with radius $\Delta > 0$, and assume that $\ell \in W^{1,\infty}(0,T;\mathcal{V}^*)$ with $|\ell|_{Lip} \leq \kappa - \delta$ (see Assumption $\mathbf{A}_{Lip}$) and $\ell' \in BV(0,T;\mathcal{V}^*)$. Then there exists a constant $K_{loc} > 0$, independent of τ, such that for the back-transformed parametrized solution $z_\tau : [0,T] \to \mathcal{Z}$ and all $\tau \leq \bar{\tau}$ with $\bar{\tau}$ sufficiently small, it holds that*

$$
\|z_\tau(t) - z(t)\|_{\mathcal{Z}} \leq K_{loc}\, \tau \quad \forall t \in [0,T]. \tag{3.3.43}
$$

Proof. The proof basically follows the steps in the proof of Theorem 3.3.9, though we need to ensure that the iterates remain in the region of uniform convexity of $\mathcal{I}$; see Remark 3.3.10. Therefore, we will show by means of induction, that $z_k, z_{k+1} \in \overline{B_{\mathcal{Z}}(z(t),\Delta)}$ for $t \in [t_{k-1}, t_k]$. As an easy consequence, the affine interpolant $\tilde{z}_\tau$, defined in (3.3.45) below, fulfills $\tilde{z}_\tau(t) \in \overline{B_{\mathcal{Z}}(z(t),\Delta)}$ for $t \in [t_{k-1}, t_k]$, which yields that the estimates in Remark 3.3.10 also hold in the local convex case and we can proceed as in the proof of Theorem 3.3.9.

Step 0: Preparation
We start by choosing

$$
\tau \leq \min\left(\frac{\Delta}{3\, C_{loc}}, \frac{\Delta}{3\, K'}, \frac{\Delta}{3\, |z|_{Lip}}, \frac{\Delta}{3} \right) =: \bar{\tau}, \tag{3.3.44}
$$

where C_{loc} denotes the constant from Lemma 3.3.13, and K' is the constant from Theorem 3.3.9. To be precise here, assume that $\mathcal{I}$ is globally κ-uniform convex. Then, by Theorem 3.3.9, there would exist a constant K' such that the a priori estimate (3.3.19) would hold on $[0,T]$. This is the constant we refer to here. In order to prove (3.3.43), we will now successively show that the affine interpolant defined by

$$
\tilde{z}_\tau(t) := z_k + \frac{t - t_{k-1}}{t_k - t_{k-1}}(z_{k+1} - z_k), \qquad t \in [t_{k-1}, t_k), \tag{3.3.45}
$$

fulfills (3.3.43) on every interval $[t_{k-1}, t_k]$. Since we might have $[t_{k-1}, t_k) = \emptyset$, this definition is, at first, only formal. However, we will successively show by means of induction w.r.t k, that $t_k - t_{k-1} \geq \varepsilon\tau$ for some fixed $\varepsilon \in (0,1)$ independent of τ.

Step 1: Initialization
We show (3.3.43) for $t \in [t_0, t_1]$. To do so, we observe that due to the choice of τ, we have $\overline{B_{\mathcal{Z}}(z_0,\tau)} \subset B_{\mathcal{Z}}(z_0,\Delta)$. Hence, $\mathcal{I}(0,\cdot)$ is convex on $\overline{B_{\mathcal{Z}}(z_0,\tau)}$ and consequently, we can argue exactly as in Remark 3.3.1 to obtain $z_1 = z_0 \in \overline{B_{\mathcal{Z}}(z(0),\Delta)}$ and $t_1 - t_0 = \tau$. The choice of τ, again, and the Lipschitz continuity of z imply that $z_0 = z_1 \in \overline{B_{\mathcal{Z}}(z(t_1),\Delta/3)}$. This together with Lemma 3.3.13 gives

$$\|z_2 - z(t_1)\|_{\mathcal{Z}} \leq \|z_2 - z_1\|_{\mathcal{Z}} + \|z_1 - z(t_1)\|_{\mathcal{Z}} \leq C_{loc}\tau + |z|_{Lip}\tau \leq \Delta/3 + \Delta/3 < \Delta.$$

Hence, $z_2, z_1 \in \overline{B_{\mathcal{Z}}(z(t_1),\Delta)}$ and the κ-uniform convexity of $\mathcal{I}(t_1,\cdot)$ on $\overline{B_{\mathcal{Z}}(z(t_1),\Delta)}$ imply that the estimates (3.3.18) and (3.3.40) hold for $k = 1$ (see Remark 3.3.10), in particular

$$\|z_2 - z_1\|_{\mathcal{V}} \leq \|z_2 - z_1\|_{\mathcal{Z}} \leq \frac{\kappa - \delta}{\kappa}(t_1 - t_0) \leq \frac{\kappa - \delta}{\kappa}\tau,$$

so that $\tilde{z}_\tau$ is well-defined on $[t_0, t_1)$ and satisfies $\|\tilde{z}_\tau'(t)\|_{\mathcal{Z}} \leq C$ for all $t \in [t_0, t_1)$. Moreover, due to the Lipschitz continuity of z and the choice of τ, we have

$$\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}} \leq \|z_1 - z(t_1)\|_{\mathcal{Z}} + \|z(t_1) - z(t)\|_{\mathcal{Z}} + \frac{t - t_0}{t_1 - t_0}\|z_2 - z_1\|_{\mathcal{Z}} \leq \Delta/3 + |z|_{Lip}\tau + \tau \leq \Delta,$$

and thus $\tilde{z}_\tau(t) \in \overline{B_{\mathcal{Z}}(z(t),\Delta)}$ for all $t \in [t_0, t_1]$. Therefore, we can exploit the convexity of $\mathcal{I}(t,\cdot)$ on $\overline{B_{\mathcal{Z}}(z(t),\Delta)}$, giving that (3.3.26) holds for $t \in [t_0, t_1]$, too. Then, as illustrated in Remark 3.3.10, we can argue analogous to the proof of Theorem 3.3.9 (*Steps 2–6*) to obtain $\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}} \leq K'\,\tau$ for all $t \in [t_0, t_1]$.

Step 2: Induction
Let $k \geq 2$ be given with

$$\|z_k - z_{k-1}\|_{\mathcal{V}} \leq \frac{\kappa - \delta}{\kappa}\tau, \tag{3.3.46}$$

$$\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}} \leq K'\,\tau \quad \forall\, t \in [t_0, t_{k-1}]. \tag{3.3.47}$$

In the first step of the proof, we have seen that these conditions are fulfilled for $k = 2$ and we will now show that these estimates can be extended to $[t_0, t_k]$. For this, we observe that, since $\tau \leq \frac{\Delta}{3\,K'}$, the inequality (3.3.47) gives $z_k = \tilde{z}_\tau(t_{k-1}) \in \overline{B_{\mathcal{Z}}(z(t_{k-1}),\Delta/3)}$. Thus, by exploiting the Lipschitz continuity of z and the choice of τ from (3.3.44), it follows that $\|z_k - z(t_k)\|_{\mathcal{Z}} \leq \frac{2\Delta}{3}$. Combining this with Lemma 3.3.13 and exploiting again (3.3.44), we find $\|z_{k+1} - z(t_k)\|_{\mathcal{Z}} \leq \|z_{k+1} - z_k\|_{\mathcal{Z}} + \|z_k - z(t_k)\|_{\mathcal{Z}} \leq \Delta$, so that $z_{k+1}, z_k \in \overline{B_{\mathcal{Z}}(z(t_k),\Delta)}$. Hence, the estimates (3.3.18) and (3.3.40) hold true (see Remark 3.3.10). It remains to show that $\tilde{z}_\tau(t) \in \overline{B_{\mathcal{Z}}(z(t),\Delta)}$ for all $t \in [t_{k-1}, t_k]$ so that we have (3.3.26) on the next time interval, see again Remark 3.3.10. By (3.3.46) and (3.3.6a) $\lambda_k = 0$ holds such that the inequality (3.3.15), in combination with $\lambda_{k+1} \geq 0$, reduces to

$$0 \geq \langle D_z\mathcal{I}(t_k, z_{k+1}) - D_z\mathcal{I}(t_k, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ + \langle D_z\mathcal{I}(t_k, z_k) - D_z\mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k\rangle_{\mathcal{Z}^*,\mathcal{Z}}.$$

The κ-uniform convexity of $\mathcal{I}(t_k, \cdot)$ on $\overline{B_{\mathcal{Z}}(z(t_k), \Delta)}$ together with $z_k, z_{k+1} \in \overline{B_{\mathcal{Z}}(z(t_k), \Delta)}$ thus gives $0 \geq \kappa\|z_{k+1} - z_k\|_{\mathcal{Z}}^2 - |\ell|_{Lip}(t_k - t_{k-1})\|z_{k+1} - z_k\|_{\mathcal{V}}$, which implies

$$\|z_{k+1} - z_k\|_{\mathcal{V}} \leq \|z_{k+1} - z_k\|_{\mathcal{Z}} \leq |\ell|_{Lip}/\kappa\,\tau \leq \frac{\kappa - \delta}{\kappa}\tau \tag{3.3.48}$$

by the assumption on $|\ell|_{Lip}$. By the time update (alg$_3$) and (3.3.46), we have

$$t_k - t_{k-1} \geq \delta/\kappa\,\tau, \tag{3.3.49}$$

which consequently gives the well-posedness of our interpolant and the boundedness of its derivative in $\mathcal{Z}$ due to (3.3.48). From (3.3.48) and, again, the choice of τ, we moreover conclude for $t \in [t_{k-1}, t_k]$ that

$$\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}} \leq \|z_k - z(t_{k-1})\|_{\mathcal{Z}} + \|z(t_{k-1}) - z(t)\|_{\mathcal{Z}} + \frac{t - t_{k-1}}{t_k - t_{k-1}}\|z_{k+1} - z_k\|_{\mathcal{Z}} \\ \leq K'\,\tau + |z|_{Lip}(t_k - t_{k-1}) + \tau \leq \Delta/3 + \Delta/3 + \Delta/3 = \Delta\,.$$

Hence, $\tilde{z}_\tau(t) \in \overline{B_{\mathcal{Z}}(z(t), \Delta)}$ for all $t \in [t_0, t_k]$ so that the uniform convexity of $\mathcal{I}(t, \cdot)$ on $\overline{B_{\mathcal{Z}}(z(t), \Delta)}$ implies that (3.3.26) holds on $[t_0, t_k]$. Thus, we can again argue as in the proof of Theorem 3.3.9 (*Steps 2–6*) to show (3.3.47) on the extended time interval $[t_0, t_k]$. In summary, we have shown that (3.3.46)–(3.3.47) holds with k instead of $k-1$, which completes the induction step. Hence, we find $\|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}} \leq K'\,\tau$ on the whole time interval $[0, T]$.

Step 3: Comparing interpolants

We again define the affine interpolant $\hat{t}_\tau$ as in (3.3.10). From (3.3.49), it follows that $\hat{t}_\tau' \geq \delta/\kappa$ for almost all $s \in [0, S_\tau]$. Thus, there exists a unique inverse function $s_\tau : [0, T] \to [0, S_\tau]$ with $1 \leq s_\tau'(t) \leq \frac{1}{1 - \frac{\kappa-\delta}{\kappa}}$ almost everywhere in $[0, T]$. In full analogy to the proof of Theorem 3.3.9 (*Step 7*), we obtain $\|z_\tau(t) - \tilde{z}_\tau(t)\|_{\mathcal{Z}} \leq \tau$, where again z_τ is the retransformed affine interpolation, i.e., $z_\tau(t) := \hat{z}_\tau(s_\tau(t))$. Thus, we finally get

$$\|z_\tau(t) - z(t)\|_{\mathcal{Z}} \leq \|z_\tau(t) - \tilde{z}_\tau(t)\|_{\mathcal{Z}} + \|\tilde{z}_\tau(t) - z(t)\|_{\mathcal{Z}} \leq K_{loc}\,\tau,$$

which was claimed. □

3.3.4 General time dependence in energy functional

The second part of the a priori analysis is concerned with the case of a general time-dependent part $f(t, z)$. While we do not impose further requirements on f, we suppose again that $\mathcal{I}$ complies with the Assumption **GC**$_\kappa$, i.e., $\mathcal{I}$ is κ-uniformly convex. Recall that this means that there exists a $\kappa > 0$ such that, for all $t \in [0, T]$ and all $z, v \in \mathcal{Z}$, it holds $\langle D_z^2\mathcal{I}(t, z)v, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \kappa\|v\|_{\mathcal{Z}}^2$, which particularly implies

$$\langle D_z\mathcal{I}(t,z_2) - D_z\mathcal{I}(t,z_1), z_2 - z_1\rangle_{\mathcal{Z}^*,\mathcal{Z}} \geq \kappa\|z_2 - z_1\|_{\mathcal{Z}}^2 \qquad \forall z_1, z_2 \in \mathcal{Z}.$$

Now, in the case, where the Lipschitz constant, or more general the constant ν does not necessarily fulfill $\nu < \kappa$, we can no longer guarantee that the algorithm always makes progress w.r.t. time, which implies that the back-transformation onto the physical time need not be continuous. In order to handle these cases, we will transform the differential solution into the artificial time interval of the discrete parametrized solution. By this means, both solutions "live" on the same time horizon and we can adapt the arguments from Theorem 3.3.9. However, we still have to guarantee that the multipliers λ_k are uniformly bounded, which is part of the following lemma.

Lemma 3.3.15. *Let Assumption* **GC**$_\kappa$ *hold. Then the Lagrange multipliers* λ_k *are bounded, i.e.,* $\lambda_k \leq \nu$ *for all* $k \in \mathbb{N}$, $k \leq N(\tau)$ *with* ν *from* ($\mathcal{I}_{f2}$).

Proof. W.l.o.g. let k be the last iterate with $\lambda_k = 0$. By Remark 3.3.1 we have $t_1 - t_0 > 0$ and therefore $\lambda_0 = 0$, so that there always exists such an index $k \leq N(\tau)$. We will first show that λ_{k+1} is bounded by the constant ν of f. Afterwards, we will show that the sequence $\{\lambda_{k+l}\}_{l\geq 1}$ is monotonically decreasing by some constant value. Since all multipliers are nonnegative, this will give some index $m \geq 1$ so that $\lambda_{k+m} = 0$. Consequently there exists, again, an index $\tilde{k} > k$ with $\lambda_{\tilde{k}} = 0$ and we may repeat the same steps showing that, indeed, all λ_k are bounded.

Step 1: Boundedness of λ_{k+1}
Since $\lambda_k = 0$, Lemma 3.3.5 gives

$$\begin{aligned}0 &\geq \kappa\|z_{k+1} - z_k\|_{\mathcal{Z}}^2 - \nu(t_k - t_{k-1})\|z_{k+1} - z_k\|_{\mathcal{V}} + \lambda_{k+1}\tau^2 \\ &\geq -\nu(t_k - t_{k-1})\|z_{k+1} - z_k\|_{\mathcal{V}} + \lambda_{k+1}\tau^2 \geq -\nu\tau^2 + \lambda_{k+1}\tau^2,\end{aligned}$$

so that, indeed, $\lambda_{k+1} \leq \nu$.

Step 2: Monotonicity of $\{\lambda_{k+l}\}_{l\geq 1}$
To proceed, let $l \geq 2$ iterations be given with $\lambda_{k+l} > 0$. Due to the time update (3.0.1b) and the complementarity in (3.3.6a) this implies

$$t_{k+l} = t_{k+l-1} = \dots = t_k \tag{3.3.50}$$

$$\text{and} \quad \|z_{k+l} - z_{k+l-1}\|_{\mathcal{V}} = \|z_{k+l-1} - z_{k+l-2}\|_{\mathcal{V}} = \dots = \tau. \tag{3.3.51}$$

We will now show that the sequence $\{\lambda_{k+l}\}_{l\geq 1}$ is monotonically decreasing by some constant value. Together with (3.3.12) for the index $k + l - 1$, (3.3.50) implies

$$0 \geq \kappa\|z_{k+l} - z_{k+l-1}\|_{\mathcal{Z}}^2 + \lambda_{k+l}\tau^2 - \lambda_{k+l-1}\tau^2.$$

Using the embedding $\mathcal{Z} \hookrightarrow \mathcal{V}$ and inserting (3.3.51), we obtain $0 \geq \kappa\tau^2 + \lambda_{k+l}\tau^2 - \lambda_{k+l-1}\tau^2$. Combining this with the bound on λ_{k+1} from above and rearranging terms then yields

$$\lambda_{k+l} \leq \lambda_{k+l-1} - \kappa \quad \Longrightarrow \quad \lambda_{k+l} \leq \lambda_{k+1} - (l-1)\kappa \leq \nu - (l-1)\kappa,$$

which finally gives that $\lambda_{k+m} = 0$ for $m = \lceil \nu/\kappa \rceil + 1$ due to the nonnegativity of the multipliers. Hence, we are in the situation of Step 1, which ultimately proves the claim. □

Finally, we need an estimate for the iterates in the stronger $\mathcal{Z}$-norm, in order to get a uniform bound for the derivative of the linear interpolants.

Lemma 3.3.16. *Let Assumption* $\mathbf{GC}_\kappa$ *be satisfied. Then there exists a constant* $C = C(\nu, \kappa) > 0$ *such that* $\|z_k - z_{k-1}\|_{\mathcal{Z}} \leq C\,\tau$ *for all iterations* $k \leq \hat{N}(\tau)$.

Proof. For $k = 1$ this easily follows from Remark 3.3.1. Hence, let $k \geq 2$. In the proof of Lemma 3.3.15, we have seen that the multipliers λ_k are bounded by ν for all $k \leq N(\tau)$. Another application of Lemma 3.3.5 thus gives

$$\kappa\|z_k - z_{k-1}\|_{\mathcal{Z}}^2 \leq \nu(t_{k-1} - t_{k-2})\|z_k - z_{k-1}\|_{\mathcal{V}} - (\lambda_k - \lambda_{k-1})\tau^2 \leq \nu\,\tau^2 + \lambda_{k-1}\,\tau^2 \leq 2\,\nu\,\tau^2,$$

where we exploited the positivity of the multiplier λ_k. □

Before we actually state the a priori estimate in the general case, let us once more stress the basic idea of its proof. First, the differential solution z, whose existence and uniqueness is guaranteed by Theorem 2.2.2, is transformed into the artificial time of the approximate parametrized solution. Here, techniques from Section 3.3.3 can be used to derive a pointwise estimate similar to (3.3.19) but in the artificial rather than the physical time. However, the fact that we obtain a pointwise characterization also allows us to provide an estimate between a suitably chosen interpolant in the physical time and the "original" differential solution afterwards.

Theorem 3.3.17. *Let Assumption* $\mathbf{GC}_\kappa$ *hold. Moreover, let* $\hat{t}_\tau$ *be defined as in* (3.3.10). *Then the sequence of approximate parametrized solutions* $\{\hat{z}_\tau\}_{\tau>0}$ *generated by the scheme in LISS satisfies the a priori error estimate*

$$\|\hat{z}_\tau(s) - z(\hat{t}_\tau(s))\|_{\mathcal{Z}} \leq C\,\sqrt{\tau} \qquad \forall s \in [0, S_\tau], \tag{3.3.52}$$

with $C > 0$ *independent of* τ.

Proof. For convenience of the reader we split the rather lengthy proof into eight parts, which are as follows:

0. First, we will see that $z(\hat{t}_\tau(s))$ is also a differential solution but in the artificial time horizon $[0, S_\tau]$.
1. For the error analysis we use, again, a slightly different piecewise affine interpolant, denoted by $\tilde{z}_\tau$ providing a certain shift in the time steps.
2. Similar to [MT04], we introduce a quantity $\gamma(s)$, which dominates the pointwise error $\|\tilde{z}_\tau(s) - z(\hat{t}_\tau(s))\|_{\mathcal{Z}}$.
3. The error measure is essentially estimated by two contributions, denoted by $E(s)$, $R(s)$ and $r(s)$. Compared to the proof of Theorem 3.3.17, a "new" term $r(s)$ occures here. This quantity contains additional errors that arise from the multipliers λ_k.

4./5. $E(s)$ and $R(s)$ are estimated by using the smoothness properties of $\mathcal{F}$ and f, while the bound for r is based on Lemma 3.3.15.

6. Together with Gronwall's lemma, these estimates yield a bound of $\mathcal{O}(\tau)$ for the error indicator γ and thus also for the error $\|\tilde{z}_\tau(s) - z(\hat{t}_\tau(s))\|_{\mathcal{Z}}^2$.

7. Finally, we relate the auxiliary interpolant $\tilde{z}_\tau$ to the "true error" containing the "correct" interpolant $\hat{z}_\tau$.

Note that we will denote by $\tilde{z}$ the differential solution which is transformed into the artificial time $[0, S_\tau]$ and by $\tilde{z}_\tau$ the auxiliary interpolant which also "lives" in $[0, S_\tau]$.

Step 0: Differential solution
First of all, due to Theorem 2.2.2, there exists a unique (differential) solution $z \in W^{1,\infty}(0,T;\mathcal{Z})$ of the rate-independent system. Applying Lemma A.2.7, we find $\tilde{z} := z \circ \hat{t}_\tau \in W^{1,\infty}(0, S_\tau; \mathcal{Z})$ so that, by the 0-homogeneity of $\partial\mathcal{R}$, it holds

$$0 \in \partial\mathcal{R}(\tilde{z}'(s)) + D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}(s))$$

for almost all $s \in [0, S_\tau]$. This can be reformulated as (see (2.2.10)):

$$\begin{aligned} \forall v \in \mathcal{Z}: \qquad \mathcal{R}(v) &\geq \langle -D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}(s)), v\rangle_{\mathcal{Z}^*,\mathcal{Z}} && \forall s \in [0, S_\tau], && \text{(3.3.53a)}\\ \mathcal{R}(\tilde{z}'(s)) &= \langle -D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}(s)), \tilde{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} && \text{f.a.a. } s \in [0, S_\tau]. && \text{(3.3.53b)} \end{aligned}$$

Since $\tilde{z} \in W^{1,\infty}(0, S_\tau; \mathcal{Z})$, it additionally holds

$$\|\tilde{z}'(s)\|_{\mathcal{Z}} \leq C \quad \text{f.a.a. } s \in [0, S_\tau]. \tag{3.3.54}$$

Step 1: Construction of auxiliary interpolants in the artificial time
Given $s \in [s_{k-1}, s_k)$ with $k \leq N(\tau)$, we define the following affine interpolant, which, in comparison to the ones in (3.3.10), incorporates a shift in the iterates

$$\tilde{z}_\tau(s) = z_k + \frac{s - s_{k-1}}{s_k - s_{k-1}}(z_{k+1} - z_k). \tag{3.3.55}$$

Moreover, we define $z_{N(\tau)+1} = z_{N(\tau)}$ and extend the affine interpolant by constant continuation onto $[0, s_{N(\tau)} + \tau]$, i.e., $\tilde{z}_\tau(s) = z_{N(\tau)}$ for $s \in [s_{N(\tau)}, s_{N(\tau)} + \tau]$. Note that $s_k - s_{k-1} = \tau$ for all $k \leq N(\tau)$ so that $\tilde{z}_\tau'(s) = (z_{k+1} - z_k)/\tau$ and we can reformulate (3.3.6b)-(3.3.6d) as

$$\forall k \in \{0, \ldots, N(\tau)\}\, \forall v \in \mathcal{Z}: \quad \begin{cases} \mathcal{R}(v) \geq \langle -D_z\mathcal{I}(t_k, z_{k+1}) \\ \qquad\qquad - \lambda_{k+1}\, J_{\mathcal{V}}(z_{k+1} - z_k), v\rangle_{\mathcal{Z}^*,\mathcal{Z}}, \end{cases} \tag{3.3.56a}$$

$$\text{f.a.a. } t \in [0, S_\tau]: \quad \mathcal{R}(\tilde{z}_\tau'(s)) = \langle -D_z\mathcal{I}(t_k, z_{k+1}) - \tau\, \lambda_{k+1}\, J_{\mathcal{V}}(\tilde{z}_\tau'(s)), \tilde{z}_\tau'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}}. \tag{3.3.56b}$$

Exploiting Lemma 3.3.16, we additionally have

$$\|\tilde{z}_\tau'(s)\|_{\mathcal{Z}} \leq C \quad \text{f.a.a. } s \in [0, S_\tau]. \tag{3.3.57}$$

Step 2: Introduction of an error measure
We now basically follow the lines of [MT04, Thm 7.4], but have to adapt the underlying analysis at some points. Therefore we present the arguments in detail. Let us define

$$\gamma(s) := \langle D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}_\tau(s)) - D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}(s)), \tilde{z}_\tau(s) - \tilde{z}(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \,. \tag{3.3.58}$$

Due to the κ-uniform convexity of $\mathcal{I}(t, \cdot)$, we have

$$\gamma(s) \geq \kappa\|\tilde{z}_\tau(s) - \tilde{z}(s)\|_{\mathcal{Z}}^2, \tag{3.3.59}$$

so that γ measures the discretization error. Analogous to Appendix A.4 we find

$$\gamma'(s) \leq C\,\|\tilde{z}_\tau(s) - \tilde{z}(s)\|_{\mathcal{Z}}^2 + 2\langle D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}_\tau(s)) - D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}(s)), \tilde{z}_\tau'(s) - \tilde{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\,, \tag{3.3.60}$$

for almost all $s \in [0, S_\tau]$. Again, we split the second term into two parts, namely

$$\begin{aligned} & e_1(s) := 2\,\langle D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}(s)) - D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}_\tau(s)), \tilde{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ \text{and}\quad & e_2(s) := 2\,\langle D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}_\tau(s)) - D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}(s)), \tilde{z}_\tau'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\,. \end{aligned}$$

Both terms can be estimated analogous to *Step 3* in the proof of Theorem 3.3.9. For convenience, we briefly repeat the arguments.

Step 3: Estimates for the error e_i
Let again $k \leq N(\tau)$ and $s \in [s_{k-1}, s_k)$ be arbitrary. First observe that, due to the convexity of $\partial\mathcal{R}(0)$, it holds for

$$\theta(s) = \frac{s - s_{k-1}}{s_k - s_{k-1}}$$

that

$$-(1-\theta(s))\,\xi_{k-1} - \theta(s)\,\xi_k \in \partial\mathcal{R}(0)$$

with $\xi_{k-1} := D_z\mathcal{I}(t_{k-1}, z_k) + \lambda_k\, J_\mathcal{V}(z_k - z_{k-1})$ and $\xi_k := D_z\mathcal{I}(t_k, z_{k+1}) + \lambda_{k+1}\, J_\mathcal{V}(z_{k+1} - z_k)$. From the characterization of $\partial\mathcal{R}(0)$, we infer $\mathcal{R}(v) \geq -\langle(1-\theta(s))\,\xi_{k-1} + \theta(s)\,\xi_k, v\rangle_{\mathcal{Z}^*,\mathcal{Z}}$ for all $v \in \mathcal{Z}$. Inserting herein $v = \tilde{z}'(s)$ and subtracting (3.3.53b), we can estimate

$$\begin{aligned} e_1(s) &= 2\,\langle D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}(s)) - (1-\theta(s))\,\xi_{k-1} - \theta(s)\,\xi_k, \tilde{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\qquad + 2\,\langle(1-\theta(s))\,\xi_{k-1} + \theta(s)\,\xi_k - D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}_\tau(s)), \tilde{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\leq 2\,\langle(1-\theta(s))\,\xi_{k-1} + \theta(s)\,\xi_k - D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}_\tau(s)), \tilde{z}'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\leq 2\,\|(1-\theta(s))\,D_z\mathcal{I}(t_{k-1}, z_k) + \theta(s)D_z\mathcal{I}(t_k, z_{k+1}) - D_z\mathcal{I}(\hat{t}_\tau(s), \tilde{z}_\tau(s))\|_{\mathcal{Z}^*}\|\tilde{z}'(s)\|_\mathcal{Z} \\ &\qquad + 2\,(1-\theta(s))\lambda_k\langle J_\mathcal{V}(z_k - z_{k-1}), \tilde{z}'(s)\rangle_{\mathcal{V}^*,\mathcal{V}} + 2\,\theta(s)\lambda_{k+1}\langle J_\mathcal{V}(z_{k+1} - z_k), \tilde{z}'(s)\rangle_{\mathcal{V}^*,\mathcal{V}} \end{aligned} \tag{3.3.61}$$

for almost all $s \in [s_{k-1}, s_k)$. By the complementarity (3.3.6a) and the time update (3.0.1b) we moreover have $0 = \lambda_k(\tau - \|z_k - z_{k-1}\|_\mathcal{V}) = \lambda_k(t_k - t_{k-1})$. The chain rule from Lemma A.2.7 thus implies that $\lambda_k\tilde{z}'(s) = \lambda_k\hat{t}_\tau'(s)z'(\hat{t}_\tau(s)) = 0$ almost everywhere in $[s_{k-1}, s_k)$. Therefore, employing

$z_{k+1} - z_k = \tau \tilde{z}_\tau'(s+\tau)$, it holds

$$\begin{aligned} e_1(s) &\le 2\,\|(1-\theta(s))\,D_z\mathcal{I}(t_{k-1},z_k) + \theta(s) D_z\mathcal{I}(t_k,z_{k+1}) - D_z\mathcal{I}(\hat{t}_\tau(s),\tilde{z}_\tau(s))\|_{\mathcal{Z}^*}\|\tilde{z}'(s)\|_{\mathcal{Z}} \\ &\quad + 2\,\theta(s)\,\lambda_{k+1}\,\tau\,\langle J_{\mathcal{V}}(\tilde{z}_\tau'(s+\tau)),\tilde{z}'(s)\rangle_{\mathcal{V}^*,\mathcal{V}}. \end{aligned}$$

Next, we turn to the term e_2. Similarly, we take $v = \tilde{z}_\tau'(s)$ in (3.3.53a) and subtract (3.3.53b) to obtain

$$\begin{aligned} 0 &\ge \langle D_z\mathcal{I}(t_k,z_{k+1}) - D_z\mathcal{I}(\hat{t}_\tau(s),\tilde{z}(s)),\tilde{z}_\tau'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \lambda_{k+1}\tau\|\tilde{z}_\tau'(s)\|_{\mathcal{V}}^2 \\ &\ge \langle D_z\mathcal{I}(t_k,z_{k+1}) - D_z\mathcal{I}(\hat{t}_\tau(s),\tilde{z}(s)),\tilde{z}_\tau'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}}, \end{aligned}$$

from which we deduce

$$\begin{aligned} e_2(s) &\le 2\,\langle D_z\mathcal{I}(\hat{t}_\tau(s),\tilde{z}_\tau(t)) - D_z\mathcal{I}(t_k,z_{k+1}),\tilde{z}_\tau'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\le 2\,\langle D_z\mathcal{I}(\hat{t}_\tau(s),\tilde{z}_\tau(s)) - (1-\theta(s))\,D_z\mathcal{I}(t_{k-1},z_k) - \theta(s)\,D_z\mathcal{I}(t_k,z_{k+1}),\tilde{z}_\tau'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\quad + 2\,\langle (1-\theta(s))\,D_z\mathcal{I}(t_{k-1},z_k) + \theta(s)\,D_z\mathcal{I}(t_k,z_{k+1}) - D_z\mathcal{I}(t_k,z_{k+1}),\tilde{z}_\tau'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \\ &\le 2\,\|(1-\theta(s))\,D_z\mathcal{I}(t_{k-1},z_k) + \theta(s) D_z\mathcal{I}(t_k,z_{k+1}) - D_z\mathcal{I}(\hat{t}_\tau(s),\tilde{z}_\tau(s))\|_{\mathcal{Z}^*}\|\tilde{z}_\tau'(s)\|_{\mathcal{Z}} \\ &\quad + 2\,(1-\theta(s))\langle D_z\mathcal{I}(t_{k-1},z_k) - D_z\mathcal{I}(t_k,z_{k+1}),\tilde{z}_\tau'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}}. \end{aligned} \tag{3.3.62}$$

Next, we define

$$E(s) := \|(1-\theta(s))\,D_z\mathcal{I}(t_{k-1},z_k) + \theta(s)\,D_z\mathcal{I}(t_k,z_{k+1}) - D_z\mathcal{I}(\hat{t}_\tau(s),\tilde{z}_\tau(s))\|_{\mathcal{Z}^*}, \tag{3.3.63}$$

$$R(s) := 2(1-\theta(s))\langle D_z\mathcal{I}(t_{k-1},z_k) - D_z\mathcal{I}(t_k,z_{k+1}),\tilde{z}_\tau'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \tag{3.3.64}$$

$$\text{and}\quad r(s) := 2\theta(s)\lambda_{k+1}\tau\langle J_{\mathcal{V}}(\tilde{z}_\tau'(s+\tau)),\tilde{z}'(s)\rangle_{\mathcal{V}^*,\mathcal{V}}. \tag{3.3.65}$$

Then we insert (3.3.63), (3.3.64) and (3.3.65) in (3.3.61) and (3.3.62), respectively. The resulting estimates for e_1 and e_2 are in turn inserted in (3.3.60), which, together with the boundedness of $\|\tilde{z}'(s)\|_{\mathcal{Z}}$ and $\|\tilde{z}_\tau'(s)\|_{\mathcal{Z}}$ by (3.3.54) and (3.3.57), yields

$$\gamma'(s) \le C(\|\tilde{z}_\tau(s) - \tilde{z}(s)\|_{\mathcal{Z}}^2 + E(s) + R(s) + r(s)). \tag{3.3.66}$$

Step 4: Estimate for $E(s)$

The particular structure of $\mathcal{I}$ together with the linearity of A and the definition of $\tilde{z}_\tau$ gives

$$\begin{aligned} E(s) &\le \|(1-\theta(s))D_z\mathcal{F}(z_k) + \theta(s)D_z\mathcal{F}(z_{k+1}) - D_z\mathcal{F}\left((1-\theta(s))z_k - \theta(s)z_{k+1}\right)\|_{\mathcal{Z}^*} \\ &\quad + \|(1-\theta(s))D_zf(t_{k-1},z_k) + \theta(s)D_zf(t_k,z_{k+1}) - D_zf(\hat{t}_\tau(s),\tilde{z}_\tau(s))\|_{\mathcal{Z}^*} \\ &=: I_1(s) + I_2(s). \end{aligned}$$

The term I_1 can be estimated exactly as in Theorem 3.3.9. For I_2, we exploit assumption $(\mathcal{I}_{f2})$, which, in combination with the embedding $\mathcal{Z} \hookrightarrow \mathcal{W}$, implies

$$\|D_zf(t_{k-1},z_k) - D_zf(t_k,z_{k+1})\|_{\mathcal{Z}^*} \le c\,(|t_{k-1} - t_k| + \|z_{k+1} - z_k\|_{\mathcal{Z}})$$

and therefore

$$
\begin{aligned}
I_2(s) \le \theta(s) \|(D_z f(t_k, z_{k+1}) - D_z f(t_{k-1}, z_k))\|_{\mathcal{Z}^*} \\
+ \|D_z f(t_{k-1}, z_k) - D_z f(\hat{t}_\tau(s), \tilde{z}_\tau(s))\|_{\mathcal{Z}^*} \\
&\le \theta(s)\, c\, (|t_{k-1} - t_k| + \|z_{k+1} - z_k\|_{\mathcal{Z}}) + c\, (|t_{k-1} - \hat{t}_\tau(s)| + \|z_k - \tilde{z}_\tau(s)\|_{\mathcal{Z}}) \\
&\le c\,\tau + C\|z_{k+1} - z_k\|_{\mathcal{Z}}.
\end{aligned}
$$

Since $\|z_{k+1} - z_k\|_{\mathcal{Z}} \le C\tau$ by Lemma 3.3.8, the above estimates for $I_1(s)$ and $I_2(s)$ imply for all $s \in [s_{k-1}, s_k)$ that $E(s) \le C\tau$. Now, integrating E and exploiting the boundedness of S_τ from (3.3.11) yields

$$
\int_0^{S_\tau} E(s)\, \mathrm{d}s \le C\tau. \tag{3.3.67}
$$

Step 5: Estimate for $R(s)$
First of all, we observe that

$$
\begin{aligned}
R(s) &= 2(1-\theta(s)) \langle D_z \mathcal{I}(t_{k-1}, z_k) - D_z \mathcal{I}(t_k, z_{k+1}), \hat{z}_\tau'(s) \rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&= 2(1-\theta(s))\, \frac{1}{\tau}\, \langle D_z \mathcal{I}(t_{k-1}, z_k) - D_z \mathcal{I}(t_{k-1}, z_{k+1}), z_{k+1} - z_k \rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&\quad + 2(1-\theta(s))\, \langle D_z \mathcal{I}(t_{k-1}, z_{k+1}) - D_z \mathcal{I}(t_k, z_{k+1}), \hat{z}_\tau'(s) \rangle_{\mathcal{Z}^*,\mathcal{Z}}.
\end{aligned}
$$

Using the κ-uniform convexity for $\mathcal{I}(t_{k-1}, \cdot)$, we find

$$
\begin{aligned}
2(1-\theta(s))\, \frac{1}{\tau}\, \langle D_z \mathcal{I}(t_{k-1}, z_k) - D_z \mathcal{I}(t_{k-1}, z_{k+1}), z_{k+1} - z_k \rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
\le -2\kappa(1-\theta(s))\, \frac{1}{\tau}\, \|z_{k+1} - z_k\|_{\mathcal{Z}}^2 \le 0.
\end{aligned} \tag{3.3.68}
$$

Thus, by exploiting, once more, the assumption ($\mathcal{I}_{f2}$), the embedding $\mathcal{Z} \hookrightarrow \mathcal{W}$ and the boundedness of $\|\tilde{z}_\tau'(s)\|_{\mathcal{V}}$, we obtain

$$
\begin{aligned}
R(s) &\le 2(1-\theta(s))\, \langle D_z \mathcal{I}(t_{k-1}, z_{k+1}) - D_z \mathcal{I}(t_k, z_{k+1}), \hat{z}_\tau'(s) \rangle_{\mathcal{Z}^*,\mathcal{Z}} \\
&= 2(1-\theta(s))\, \langle D_z f(t_{k-1}, z_{k+1}) - D_z f(t_k, z_{k+1}), \hat{z}_\tau'(s) \rangle_{\mathcal{V}^*,\mathcal{V}} \\
&\le 2\, c\, (|t_{k-1} - t_k|) \|\tilde{z}_\tau'(s)\|_{\mathcal{V}} \le C\tau.
\end{aligned}
$$

Hence, we have

$$
\int_0^{S_\tau} R(s)\, \mathrm{d}s \le C\tau. \tag{3.3.69}
$$

again by the bound on S_τ.

Step 6: Estimate for $r(s)$
Here, we take advantage of (3.3.57), (3.3.54) and the boundedness of the multiplier λ_{k+1} by (3.3.15) which gives

$$
r(s) \le 2\, \lambda_{k+1}\, \tau\, \|\tilde{z}_\tau'(s+\tau)\|_{\mathcal{V}} \|\tilde{z}'(s)\|_{\mathcal{V}} \le C\, \tau. \tag{3.3.70}
$$

Integrating, exploiting again the boundedness of S_τ and combining the resulting estimate with

(3.3.67) and (3.3.69), we have overall shown that

$$\int_0^{S_\tau} E(s)\,\mathrm{d}s + \int_0^{S_\tau} R(s)\,\mathrm{d}s + \int_0^{S_\tau} r(s)\,\mathrm{d}s \leq C\tau. \tag{3.3.71}$$

Step 7: Obtain convergence rate by Gronwall's lemma
Exploiting that $\gamma(s)/\kappa \geq \|\tilde{z}_\tau(s) - \tilde{z}(s)\|_{\mathcal{Z}}^2$ in (3.3.66), one obtains

$$\gamma'(s) \leq C(\gamma(s) + E(s) + R(s) + r(s)).$$

Integrating this and using Gronwall's lemma as well as the estimates (3.3.71) and (3.3.11) on E, R, r and S_τ yields

$$\gamma(s) \leq (\gamma(0) + C\tau)\exp^{Cs} \leq C(\gamma(0) + \tau).$$

Due to $\tilde{z}_\tau(0) = z_1 = z_0$, see Remark 3.3.1, we have $\gamma(0) = 0$. Using another time the κ-uniform convexity of $\mathcal{I}$, we therefore finally obtain

$$\|\tilde{z}_\tau(s) - \tilde{z}(s)\|_{\mathcal{Z}}^2 \leq \gamma(s)/\kappa \leq C\tau. \tag{3.3.72}$$

Taking into account the definition of $\tilde{z}_\tau(s)$, we arrive at (3.3.52) by repeating the *Step 7* in Theorem 3.3.9. □

Some remarks and comments concerning the assertion of Theorem 3.3.17 and its proof are in order.

Remark 3.3.18. In contrast to Theorem 3.3.9, we do not obtain the optimal rate of convergence here. The reason for this is twofold. On the one hand, we rely on less restrictive assumptions on the time-dependent part, which entails that the estimates on E and R are merely of order $\mathcal{O}(\tau)$ instead of $\mathcal{O}(\tau^2)$. The crucial part of the proof, though, is the estimate of r. Unfortunately, this term does not contain any parts of the energy, so that we cannot exploit their regularity. Hence, a first potential way out could be to find sufficient conditions such that $\lambda_k = 0$ for all $k \leq N(\tau)$. In fact, this corresponds to the Assumption $\mathbf{A}_{Lip}$ from Section 3.3.3. A second potential way out could be to replace $(\hat{t}_\tau, \tilde{z}_\tau)$ by a more sophisticated interpolant. Indeed, one major issue is the fact that the complementarity does not carry over to the last term in (3.3.61). More precisely, we see that the term $\lambda_k \hat{t}'(s)$ still vanishes in $[s_{k-1}, s_k)$ (cf. the explanation after (3.3.61)) but this does unfortunately not hold for $\lambda_{k+1}\hat{t}'(s)$. In this context, it might therefore also be useful to directly construct an interpolant in the physical time as it is done in [MS20]. Nevertheless, it is to be noted that, due to the 1−homogeneity of the dissipation, it is always possible to achieve $|\ell|_{Lip} < \kappa$ by rescaling the time accordingly. Then Theorem 3.3.9 is applicable giving the optimal order in the rescaled time scale. Of course, depending on the Lipschitz constant of ℓ, the rescaled time scale might become rather small so that a large number of iterations is necessary, but this rescaling argument indicates that it should be possible to achieve the optimal order in the case of large $|\ell|_{Lip}$, too. This, however, gives rise to future research.

As a direct consequence of the pointwise estimate in Theorem 3.3.17, we also obtain a pointwise estimate in the physical time for the following affine interpolant, which will only include the iterates

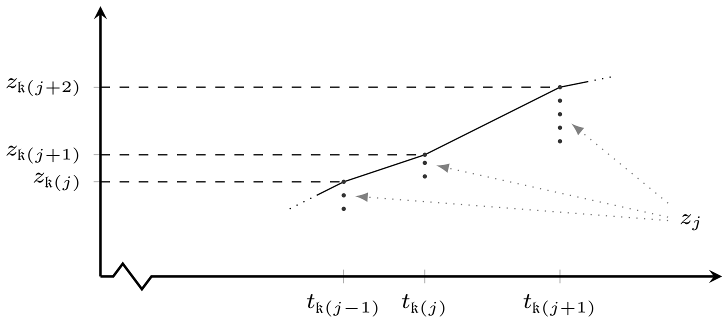


Figure 3.3.2: Qualitative illustration of the affine interpolant $\hat{t}$, the choice of the artificial end time S_τ via the equality $\hat{t}(S_\tau) = T$ and the upper bound $\tilde{S}$.

for which the time update proceeds. Thus, we set (cf. [MS20])

- $N(\tau)$ = number of iterations to reach the end time T (with step size τ),
- $\mathcal{N}(\tau) := \{k \in \{1, \dots, N(\tau)\} : t_k - t_{k-1} > 0\}$.

The iterations in $\mathcal{N}(\tau) \cup \{0\}$ are then numbered from 0 to $|\mathcal{N}(\tau)|$ and the corresponding mapping is denoted by $\Bbbk$, i.e.,

$$\Bbbk : \{0, 1, \dots, |\mathcal{N}(\tau)|\} \to \mathcal{N}(\tau) \cup \{0\} \quad \text{so that} \quad (\mathcal{N}(\tau) \cup \{0\}) = \{\Bbbk(0), \Bbbk(1), \dots, \Bbbk(|\mathcal{N}(\tau)|)\}.$$

Therewith, we define for $t \in [t_{\Bbbk(j-1)}, t_{\Bbbk(j)})$, $j = 1, \dots, |\mathcal{N}(\tau)| - 1$,

$$z_\tau(t) = z_{\Bbbk(j)} + \frac{t - t_{\Bbbk(j-1)}}{t_{\Bbbk(j)} - t_{\Bbbk(j-1)}} \left(z_{\Bbbk(j+1)} - z_{\Bbbk(j)}\right) \tag{3.3.73}$$

as well as $z_\tau(t) = z_{N(\tau)}$ for $t \in [t_{\Bbbk(|\mathcal{N}(\tau)|)}, T]$. It then holds:

Corollary 3.3.19. *Let $\mathcal{I}(t, \cdot) \in C^{2,1}_{loc}(\mathcal{Z}; \mathbb{R})$ (see (1.0.5)) as well as Assumption* $\mathbf{GC}_\kappa$ *hold. Then $\{z_\tau\}_{\tau>0}$ as defined in (3.3.73) satisfies the a priori error estimate*

$$\|z_\tau(t) - z(t)\|_{\mathcal{Z}} \le C\sqrt{\tau} \qquad \forall t \in [0, T], \tag{3.3.74}$$

for some $C > 0$ independent of τ.

Proof. By *Step 6* in the proof of Theorem 3.3.17, particularly the estimate (3.3.72), we know that $\|\tilde{z}_\tau(s) - z(\hat{t}_\tau(s))\|_{\mathcal{Z}} \le C\sqrt{\tau}$ holds for all $s \in [0, S_\tau]$ where $\tilde{z}_\tau$ denotes the interpolant given in (3.3.55). Since this estimate is valid pointwise, we conclude that

$$\|z_k - z(t_{k-1})\|_{\mathcal{Z}} = \|\tilde{z}_\tau(s_{k-1}) - z(\hat{t}_\tau(s_{k-1}))\|_{\mathcal{Z}} \le C\sqrt{\tau}$$

is valid for all $k \le N(\tau)$. To proceed, observe that, by construction, we have $t_k = t_{\Bbbk(j-1)}$ for all $k \in \{\Bbbk(j-1), \Bbbk(j-1)+1, \dots, \Bbbk(j)-1\}$, so that

$$\begin{aligned}\|z_\tau(t_{\Bbbk(j-1)}) - z(t_{\Bbbk(j-1)})\|_{\mathcal{Z}} &= \|z_{\Bbbk(j)} - z(t_{\Bbbk(j-1)})\|_{\mathcal{Z}} \\ &= \|z_{\Bbbk(j)} - z(t_{\Bbbk(j)-1})\|_{\mathcal{Z}} \leq C\sqrt{\tau} \quad \forall j \in \{0, 1, \dots, |\mathcal{N}(\tau)|\}.\end{aligned}$$

Combining this with the Lipschitz continuity of z, we finally arrive at (3.3.74). □

It is to be expected that, in the locally uniformly convex case, the corresponding results from Theorem 3.3.17 can be obtained by carrying out the same arguments as in the proof of Theorem 3.3.14.

Chapter 4

Numerical results

This chapter is devoted to the presentation of numerical results that can be obtained using the discrete local stationarity scheme LISS. Section 4.1 is concerned with the actual realization of the scheme in the case of the semilinear example from Section 3.2.5. This particularly includes the finite element discretization and the approximation $\mathcal{R}_h$ of the dissipation potential, which is performed by a mass lumping scheme. This mass lumping in fact induces that the subdifferential of $\mathcal{R}_h$ admits a component-wise characterization, so that the stationarity conditions (3.3.6a)-(3.3.6d) can be written as a system of nonsmooth equations. This system is then solved by means of a semismooth Newton method. A numerical test is presented in Section 4.1.3. In the ensuing Section 4.2 we provide three numerical examples in order to illustrate the theoretical findings of the Section 3.3.

4.1 Numerical realization

To test the fully discrete local stationarity scheme numerically, we choose the semilinear setting from Section 3.2.5, that is:

- $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, is a bounded domain.
- The spaces are chosen to be $\mathcal{Z} = H_0^1(\Omega)$, $\mathcal{V} = L^2(\Omega)$, and $\mathcal{X} = L^1(\Omega)$.
- For the operator $\mathbb{V} : L^2(\Omega) \to L^2(\Omega)^*$, we just choose the Riesz isomorphism.
- The operator A within the energy functional is set to $A = -\Delta : H_0^1(\Omega) \to H^{-1}(\Omega)$.
- The nonlinearity $\mathcal{F}$ in the energy is defined as the well-known double well potential

$$\mathcal{F}(z) := 48 \int_\Omega \big(1 - z(x)^2\big)^2 \mathrm{d}x.$$

- The external loads are only depending on t and given by

$$\ell(t, x) = \ell(t) := -48 \sin(2\pi t), \quad (t, x) \in [0, T] \times \Omega.$$

- The dissipation functional is given by the L^1-norm, i.e., $\mathcal{R}(v) = \|v\|_{L^1(\Omega)}$.

4.1.1 Finite Element discretization

We employ classical linear finite elements (FE) to discretize the energy and the dissipation functional. For this purpose, assume that a family $\{\mathcal{T}_h\}_{h>0}$ of shape-regular triangulations of the domain Ω be given. Herein, h denotes the mesh size defined by $h := \max_{T\in\mathcal{T}_h} \operatorname{diam}(T)$. To keep the discussion concise, we assume that Ω is a polygon and polyhedron, respectively, and that the triangulations exactly fit the boundary. For the discrete space, we choose the space of piecewise linear and continuous test functions, i.e.,

$$\mathcal{Z}_h := \{v \in C(\bar{\Omega}) \cap H_0^1(\Omega) : v|_T \in \mathcal{P}_1 \;\forall\, T \in \mathcal{T}_h\}.$$

By classical results on Lagrange and quasi-interpolation, respectively, the best approximation properties of the orthogonal and the Ritz projection show that the approximation assumptions in (3.1.7) are fulfilled. Moreover, as shown in [GHS16], the shape-regularity of the triangulation guarantees that the stability assumption in (3.1.5) is satisfied.

The discretization of the dissipation potential in form of the L^1-norm is performed by a mass lumping scheme, which turns out to be advantageous for the numerical calculation of a stationarity point of (alg$_1$), as we will see in Section 4.1.2. Let us denote the nodes of the triangulation $\mathcal{T}_h$ and the associated nodal basis by x_i and φ_i, $i = 1, ..., N_h$. Moreover, given a function $z_h \in \mathcal{Z}_h$, we denote the coefficient vector of z_h w.r.t. the nodal basis by $\boldsymbol{z} = (z_1, ..., z_{N_h}) \in \mathbb{R}^{N_h}$, i.e., $z_h(x) = \sum_{i=1}^{N_h} z_i\, \varphi_i(x)$. Then the discrete dissipation potential $\mathcal{R}_h : \mathcal{Z}_h \to \mathbb{R}$ is defined by

$$\mathcal{R}_h(z_h) := \int_\Omega \sum_{i=1}^{N_h} |z_i| \varphi_i(x)\, \mathrm{d}x. \tag{4.1.1}$$

It remains to verify the assumptions in (a)–(c) on $\mathcal{R}_h$, which is done next.

Proposition 4.1.1. *The discrete dissipation potential defined in* (4.1.1) *satisfies the conditions (a)–(c) from Section 3.1.*

Proof. Due to the positivity of the nodal basis, $\mathcal{R}_h$ is only a scaled version of the $|\cdot|_1$-norm on $\mathbb{R}^N$ and consequently, it fulfills assumption (a). Moreover, the nonnegativity of the nodal basis directly implies for every $v_h \in \mathcal{Z}_h$ that

$$\mathcal{R}(v_h) = \int_\Omega \Big| \sum_{i=1}^{N_h} v_i \varphi_i(x) \Big| \,\mathrm{d}x \le \int_\Omega \sum_{i=1}^{N_h} |v_i| \varphi_i(x)\, \mathrm{d}x = \mathcal{R}_h(v_h),$$

which is the second assumption (b). Before we proceed with showing (c), we make the following observation: Taking as $F_T : \hat{T} \to T$ the affine transformation to the reference element $\hat{T} = \operatorname{conv}((0,0),(1,0),(0,1))$, we obtain

$$\mathcal{R}_h(z_h) = \sum_{T\in\mathcal{T}_h} \int_{\hat{T}} \sum_{x_i \in \overline{T}} |z_i|\, (\varphi_i \circ F_T)(\hat{x})\, |\det(DF_T(\hat{x}))|\, \mathrm{d}\hat{x}.$$

Let us denote the transformed basis functions by $\hat{\varphi}_j$, $j = 1, ..., d$. Due to the nonnegativity of the nodal basis, each of the mappings

$$\mathbb{R}^d \ni (v_j)_{j=1}^d \mapsto \int_{\hat{T}} \sum_{j=1}^{d} |v_j| \hat{\varphi}_j(\hat{x}) \, \mathrm{d}\hat{x} \quad \text{and} \quad \mathbb{R}^d \ni (v_j)_{j=1}^d \mapsto \int_{\hat{T}} \Big| \sum_{j=1}^{d} v_j \hat{\varphi}_j(\hat{x}) \Big| \, \mathrm{d}\hat{x}$$

forms a norm on $\mathbb{R}^d$. Thus, by the norm-equivalence in finite dimensions, there exists a constant $c > 0$, only dependening on $d = \dim(\Omega)$, such that

$$\mathcal{R}_h(z_h) \le \sum_{T \in \mathcal{T}_h} c \int_{\hat{T}} \Big| \sum_{x_i \in \overline{T}} z_i \, (\varphi_i \circ F_T)(\hat{x}) \, |\det(DF_T(\hat{x}))| \Big| \, \mathrm{d}\hat{x} = c \, \|z_h\|_{L^1(\Omega)} \, . \tag{4.1.2}$$

Now, concerning the convergence in (c), we set $\mathcal{U} = C_c^\infty(\Omega)$ and estimate for all $u \in \mathcal{U}$:

$$\begin{aligned} &|\mathcal{R}_h(\Pi_h u) - \mathcal{R}(u)| \\ &\quad \le |\mathcal{R}_h(\Pi_h(u)) - \mathcal{R}_h(I_h(u))| + |\mathcal{R}_h(I_h(u)) - \mathcal{R}(I_h(u))| + |\mathcal{R}(I_h u) - \mathcal{R}(u)| \\ &\quad =: e_1 + e_2 + e_3, \end{aligned}$$

where $I_h : C(\bar{\Omega}) \to \mathcal{Z}_h$ denotes the Lagrange interpolation operator. Using the reverse triangle inequality in combination with (4.1.2), the first difference can be estimated by

$$\begin{aligned} e_1 &= \Big| \int_\Omega \sum_{i=1}^{N_h} \big(|(\Pi_h u)(x_i)| - |(I_h u)(x_i)|\big) \varphi_i(x) \, \mathrm{d}x \Big| \\ &\le \mathcal{R}_h(\Pi_h(u) - I_h(u)) \\ &\le c \, \|\Pi_h(u) - I_h(u)\|_{L^1(\Omega)} \le C \, \|\Pi_h\|_{\mathcal{L}(L^2(\Omega), L^1(\Omega))} \, \|u - I_h(u)\|_{L^2(\Omega)} \to 0, \quad \text{as } h \searrow 0. \end{aligned}$$

Thanks to the partition-of-unity property of the nodal basis, we obtain for the second difference by applying the reverse triangle inequality once again

$$\begin{aligned} e_2 &\le \sum_{T \in \mathcal{T}_h} \int_T \sum_{i=1}^{N_h} |(I_h u)(x_i) - (I_h u)(x)| \, \varphi_i(x) \, \mathrm{d}x \\ &\le \sum_{T \in \mathcal{T}_h} h \, \|I_h u\|_{W^{1,\infty}(T)} \int_T \sum_{i=1}^{N_h} \varphi_i(x) \, \mathrm{d}x \le C \, h \, \|u\|_{W^{1,\infty}(\Omega)} \to 0, \quad \text{as } h \searrow 0. \end{aligned}$$

Using the reverse triangle inequality a third time, we estimate the last difference by

$$e_3 \le \mathcal{R}(I_h u - u) \le C \, \|I_h u - u\|_{L^2(\Omega)} \to 0, \quad \text{as } h \searrow 0.$$

Overall, we thus obtain assumption (c) with $\mathcal{U} = C_c^\infty(\Omega)$, which is clearly dense in $L^2(\Omega)$. □

4.1.2 Numerical solution of the stationarity system

The essential advantage of the discretization of $\mathcal{R} = \|\cdot\|_{L^1(\Omega)}$ in (4.1.1) is that its subdifferential admits a *component-wise* characterization. This allows to rewrite the first-order optimality conditions associated with (alg$_1$) as a system of nonsmooth equations, which is amenable to semismooth Newton methods. To see this, note that the discrete dissipation potential can equivalently be rewritten as

$$\mathcal{R}_h(z_h) = \boldsymbol{R}(\boldsymbol{z}) := \boldsymbol{m}^\top|\boldsymbol{z}| \quad \text{with} \quad \boldsymbol{m} = (m_1, ..., m_{N_h}) := M\mathbb{1}, \tag{4.1.3}$$

where $M_{ij} = \int_\Omega \varphi_i \varphi_j \,\mathrm{d}x \in \mathbb{R}^{N_h \times N_h}$ is the mass matrix, $\mathbb{1} = (1, ..., 1) \in \mathbb{R}^{N_h}$, and $|\boldsymbol{z}| = (|z_1|, \dots, |z_{N_h}|)$. Therefore, the convex subdifferential of $\mathcal{R}_h$ can be expressed as follows:

$$\boldsymbol{q} = (q_1, ..., q_{N_h}) \in \partial \boldsymbol{R}(z_h) \quad \Longleftrightarrow \quad |q_i| \leq m_i,\ q_i\, z_i = m_i|z_i| \quad \forall\, i = 1, ..., N_h,$$

which can equivalently be formulated as

$$\max\{|q_i| - m_i, m_i|z_i| - q_i\, z_i\} = 0 \quad \forall\, i = 1, ..., N_h\,. \tag{4.1.4}$$

To reformulate the optimality conditions of (alg$_1$) as nonsmooth equation, let us abbreviate the coefficient vector associated with $z_k^{\tau,h}$ by $\boldsymbol{z}^k$. Moreover, we denote the energy functional considered as mapping acting on the coefficient vector by $\boldsymbol{I} : \mathbb{R}^{N_h} \to \mathbb{R}$. Then *Step 3* in LISS corresponds to the calculation of a stationary point of the following minimization problem for the coefficient vector $\boldsymbol{z}^k$:

$$\begin{cases} \min\limits_{\boldsymbol{z}\in\mathbb{R}^{N_h}} & \boldsymbol{I}(t_{k-1}, \boldsymbol{z}) + \boldsymbol{R}(\boldsymbol{z} - \boldsymbol{z}^{k-1}) \\ \quad \text{s.t.} & G(\boldsymbol{z}) \leq 0, \end{cases} \tag{4.1.5}$$

with $G(\boldsymbol{z}) = \frac{1}{2}((\boldsymbol{z} - \boldsymbol{z}^{k-1})^\top M(\boldsymbol{z} - \boldsymbol{z}^{k-1}) - \tau^2)$. Here and for the rest of this section, we abbreviate $t_{k-1}^{\tau,h}$ simply by t_{k-1}. Based on the above description of the convex subdifferential of $\boldsymbol{R}$, we find the following conditions for a stationary point of this finite dimensional problem:

Lemma 4.1.2. *If $\boldsymbol{z}^k \in \mathbb{R}^{N_h}$ is a stationary point of (4.1.5) in the sense of (alg$_1$), then there exists multipliers $\boldsymbol{q} \in \mathbb{R}^{N_h}$ and $\lambda \geq 0$ such that*

$$D_z\boldsymbol{I}(t_{k-1}, \boldsymbol{z}^k) + \lambda\, G'(\boldsymbol{z}^k) + \boldsymbol{q} = 0, \tag{4.1.6a}$$

$$\max\{|q_i| - m_i, m_i|z_i^k - z_i^{k-1}| - q_i(z_i^k - z_i^{k-1})\} = 0 \quad \forall\, i = 1, ..., N_h, \tag{4.1.6b}$$

$$\max\{-\lambda, G(\boldsymbol{z}^k)\} = 0. \tag{4.1.6c}$$

Proof. Since the arguments are quite standard, we will be brief at this point. First, we define

$$I^-(r) := \begin{cases} 0, & \text{if } r \leq 0, \\ +\infty, & \text{else.} \end{cases}$$

By the sum rule for convex functions, which is clearly applicable here, we find that (alg$_1$) can be written as $0 \in \partial\boldsymbol{R}(\boldsymbol{z}^k - \boldsymbol{z}^{k-1}) + \partial(I^- \circ G)(\boldsymbol{z}^k) + D_z\boldsymbol{I}(t_{k-1}, \boldsymbol{z}^k)$. Hence, there exists $\boldsymbol{q} \in \partial\boldsymbol{R}(\boldsymbol{z}^k - \boldsymbol{z}^{k-1})$ such that

$$D_z\boldsymbol{I}(t_{k-1}, \boldsymbol{z}^k) + \partial(I^- \circ G)(\boldsymbol{z}^k) + \boldsymbol{q} \ni 0.$$

It is easy to see that the chain rule for convex subdifferentials (see, e.g., [SW11, Lem. 3.4]) can be used here to obtain

$$\partial(I^- \circ G)(\boldsymbol{z}^k) = G'(\boldsymbol{z}^k)^* \partial I^-(G(\boldsymbol{z}^k)).$$

This gives the existence of $\lambda \in \partial I^-(G(\boldsymbol{z}^k))$ so that

$$D_z \boldsymbol{I}(t_{k-1}, \boldsymbol{z}^k) + G'(\boldsymbol{z}^k)^* \lambda + \boldsymbol{q} = 0.$$

Since this equation is precisely an equation in the dual space of $\mathbb{R}^{N_h}$ we may write it, with a little abuse of notation (particularly identifying the dual space of $\mathbb{R}^{N_h}$ with itself), as

$$D_z \boldsymbol{I}(t_{k-1}, \boldsymbol{z}^k) + \lambda\, G'(\boldsymbol{z}^k) + \boldsymbol{q} = 0.$$

Finally, $\lambda \in \partial I^-(G(\boldsymbol{z}^k))$ is equivalent to the conditions

$$\lambda \geq 0, \quad \lambda\, G(\boldsymbol{z}^k) = 0, \quad G(\boldsymbol{z}^k) \leq 0,$$

so that, by exploiting the characterization of $\boldsymbol{q}$ from (4.1.4), we overall end up with the system in (4.1.6). □

Remark 4.1.3. Clearly, Lemma 4.1.2 provides an analogous result compared to Lemma 3.2.2. Indeed, equation (4.1.6b), which just says that q is an element of $\partial\mathcal{R}(\boldsymbol{z}^k - \boldsymbol{z}^{k-1})$, in combination with (4.1.6a) corresponds to the properties (3.2.3c) and (3.2.3d). Moreover, (4.1.6c) describes a discrete version of the complementarity in (3.2.3a).

Remark 4.1.4. Denoting (with a little abuse of notation) the stiffness matrix associated with the FE discretization of the Laplacian by $A \in \mathbb{R}^{N_h \times N_h}$, (4.1.6a) is equivalent to

$$A\boldsymbol{z}^k - 192\Big(\int_\Omega (1 - z_k^{\tau,h}(x)^2) z_k^{\tau,h}(x)\, \varphi_i(x)\, \mathrm{d}x\Big)_{i=1}^{N_h} + \ell(t_{k-1})\, \boldsymbol{m} + \lambda\, M(\boldsymbol{z}^k - \boldsymbol{z}^{k-1}) + \boldsymbol{q} = 0.$$

Thus, by employing an appropriate quadrature rule, equation (4.1.6a) can be evaluated without any additional discretization error.

The optimality system in (4.1.6) is solved numerically by a semismooth Newton algorithm, see, e.g., [HPUU08]. To describe this in detail, let us denote the left-hand side of (4.1.6) by $F : \mathbb{R}^{2N_h+1} \to \mathbb{R}^{2N_h+1}$ so that (4.1.6) becomes $F(\boldsymbol{z}, \boldsymbol{q}, \lambda) = 0$. Of course, F depends on time discretization level k, but we suppress this dependency for the time being to shorten the notation. Now, given an iterate $\boldsymbol{x}^n = (\boldsymbol{z}^n, \boldsymbol{q}^n, \lambda^n)$, we compute the next one by solving the following semismooth Newton equation

$$H_n\,(\boldsymbol{x}^{n+1} - \boldsymbol{x}^n) = -F(\boldsymbol{x}^n) \quad \text{with} \quad H_n \in \partial^N F(\boldsymbol{x}^n),$$

where $\partial^N F$ denotes the Newton-derivative according to [IK08]. As a composition of Newton-differentiable functions, F itself is Newton-differentiable, see [HPUU08, Thm. 2.10]. For our implementation, we choose

$$H_n := \begin{pmatrix} D^2_{zz}\boldsymbol{I}(t_{k-1}, \boldsymbol{z}^n) + \lambda_n M & \mathrm{Id}_{N_h \times N_h} & M(\boldsymbol{z}^n - \boldsymbol{z}^{k-1}) \\ \mathrm{diag}(\boldsymbol{\alpha}^n) & \mathrm{diag}(\boldsymbol{\beta}^n) & 0_{N_h} \\ \chi_n(\boldsymbol{z}^n - \boldsymbol{z}^{k-1})^\top M & 0_{N_h}^\top & -1 + \chi_n \end{pmatrix} \tag{4.1.7}$$

with

$$\boldsymbol{\alpha}_i^n := \begin{cases} 0, & m_i|z_i^n - z_i^{k-1}| - q_i^n(z_i^n - z_i^{k-1}) - |q_i^n| + m_i \leq 0, \\ m_i \,\widetilde{\mathrm{sgn}}(z_i^n - z_i^{k-1}) - q_i^n, & m_i|z_i^n - z_i^{k-1}| - q_i^n(z_i^n - z_i^{k-1}) - |q_i^n| + m_i > 0, \end{cases}$$

$$\boldsymbol{\beta}_i^n := \begin{cases} \widetilde{\mathrm{sgn}}(q_i^n), & m_i|z_i^n - z_i^{k-1}| - q_i^n(z_i^n - z_i^{k-1}) - |q_i^n| + m_i \leq 0, \\ z_i^{k-1} - z_i^n, & m_i|z_i^n - z_i^{k-1}| - q_i^n(z_i^n - z_i^{k-1}) - |q_i^n| + m_i > 0, \end{cases}$$

as well as

$$\chi_n := \begin{cases} 1, & G(\boldsymbol{z}^n) > -\lambda^n, \\ 0, & G(\boldsymbol{z}^n) \leq -\lambda^n, \end{cases} \qquad \text{and} \qquad \widetilde{\mathrm{sgn}}(x) := \begin{cases} 1, & x \geq 0, \\ -1, & x < 0. \end{cases}$$

According to [IK08], $\widetilde{\mathrm{sgn}}$ constitutes an element of the Newton-derivative of the absolute value function. We choose this particular element instead of the sgn-function satisfying $\mathrm{sgn}(0) = 0$ in order to avoid the appearance of zero rows in H_n. With this choice, all matrices H_n appearing in the numerical test have shown to be invertible and the semismooth Newton method performed well with respect to both, robustness and efficiency. In particular, no globalization efforts are needed to ensure convergence of the method.

4.1.3 Numerical results

For the numerical test, we choose the unit square $\Omega = (0,1)^2$ as computational domain. Moreover, the initial state is set to $z_0 \equiv 0$ and the final time is $T = 1.0$. Note that the initial state thus satisfies $Az_0 \equiv 0 \in L^2(\Omega)$ as required by the standing assumptions in Section 3.1. The domain is discretized by a Friedrich-Keller triangulation with mesh size $h = \sqrt{2}/50$. For the time step size, we choose $\tau = 0.01$. The numerical computations are performed with MATLAB© and the linear systems of equations arising in each semismooth Newton step are solved by MATLAB's inbuilt direct solver based on UMFPACK.

We compare the local minimization algorithm with the global minimization scheme from (3.0.2), which is discretized in the same way as (alg$_1$) by using piecewise linear finite elements. The minimization problem in (3.0.2a) is also solved by means of the semismooth Newton method. In order to ensure the convergence to global minimizers, we choose the two global minimizers of the nonlinear function $\mathcal{F}$ (i.e., $z \equiv 1$ and $z \equiv -1$) as starting points for the semismooth Newton method.

Let us first comment on the results of the local minimization iteration from algorithm LISS. Since $\ell(0) \equiv 0$, the initial state $z_0 = P_h z_0$ is locally stable, meaning $-D_z\mathcal{I}(0, z_0) \in \partial\mathcal{R}_h(0)$. Consequently, the state does not change in the first iteration and, thanks to the time update in (alg$_3$), the physical time proceeds by $\tau = 0.01$. However, $z_0 \equiv 0$ is only a local maximum of the nonlinearity $\mathcal{F}$ and therefore, the external load enforces the state to jump immediately after the first time step into a local minimum in the subsequent iterations. In case of the local minimization algorithm, this jump evolves as a viscous transition, while the physical time stagnates, see Figure 4.1.2. The state after this viscous transition is shown in Figure 4.1.1b. Afterwards the system evolves in a time continuous manner until $t \approx 0.6724$. At this time, a second jump occurs and the system switches into a viscous behavior, which can be observed in Figures 4.1.1d-4.1.1h, finally yielding the state in Figure 4.1.1i. Meanwhile, the physical time again stands still (see Figure 4.1.2), so that the solution in fact changes in a jump-like fashion. The end state is shown in Figure 4.1.1j.

Let us now turn to the results of the global minimization scheme from (3.0.2). Just as in case of the algorithm LISS, the state jumps to a global minimum immediately after the first time step and evolves continuously afterwards. However, as the time evolves, both solutions show a quite different behavior. While the second discontinuity of the parametrized solution shows up at $t \approx 0.6724$, as depicted above, the global energetic solution already jumps at $t \approx 0.51$. In view of the global minimization in (3.0.2a), it is intuitively expected that the global energetic solution jumps as soon as possible (cf. also [MRS12, Mie03] and the references therein). The difference between the global energetic and the parametrized solution can even be further enhanced. For instance, by choosing $\ell(t) = -32\sin(2\pi t)$, the global energetic solution still provides a discontinuity at $t \approx 0.51$, while the parametrized one remains continuous until the end time is reached.

4.2 A priori error estimates

This last section of Chapter 4 is devoted to presentation of several numerical examples under the aspect of a priori error estimates. In particular, the results from Section 3.3 are visualized here.

4.2.1 Quadratic case

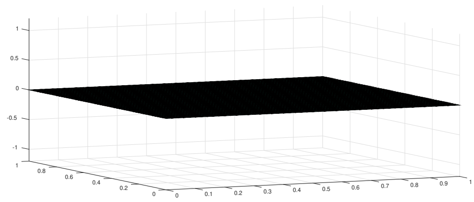
We start with an infinite-dimensional example. For that, we let $\Omega = [0,1]^2$ and choose

$$\mathcal{I}(t,z) = \frac{1}{2}\langle Az, z\rangle_{\mathcal{Z}^*,\mathcal{Z}} - \langle \ell(t), z\rangle_{\mathcal{V}}$$

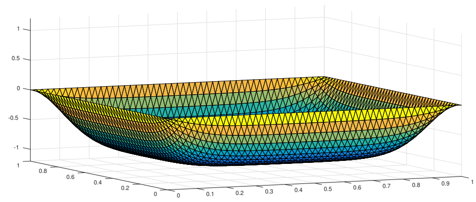
with $A = -\Delta : H_0^1(\Omega) \mapsto H^{-1}(\Omega)$ and $\ell(t,x) = \mathbb{1}_\Omega - \frac{1}{\pi}\cos(\pi\, t/2) f(x)$, wherein $f(x) = 2(x_1(1 - x_1) + x_2(1 - x_2))$. Moreover, the dissipation functional is given by the L^1-norm, i.e., $\mathcal{R}(v) = \|v\|_{L^1(\Omega)}$. Consequently, the underlying spaces are $\mathcal{Z} = H_0^1(\Omega)$, $\mathcal{V} = L^2(\Omega)$, and $\mathcal{X} = L^1(\Omega)$. In this setting, the unique (differential) solution to (RIS) reads

$$z(t,x) = \begin{cases} 0 & , t \in [0,1), \\ -\frac{1}{\pi}\cos(\frac{\pi}{2}t)\, v(x) & , t \in [1,2), \\ -\frac{1}{\pi}\, v(x) & , t \in [2,3], \end{cases} \tag{4.2.1}$$

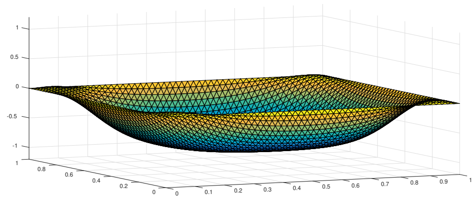
with $v(x) = x_1 x_2 (1 - x_1)(1 - x_2)$. For the spatial discretization of this system, we choose again linear finite elements on a Friedrich–Keller triangulation with mesh size $h = \sqrt{2}/100$ and the mass-



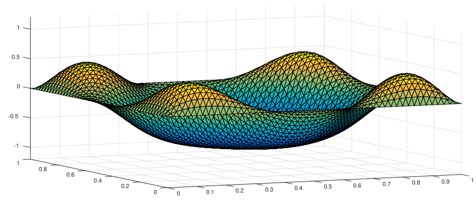
(a) t = 0.00



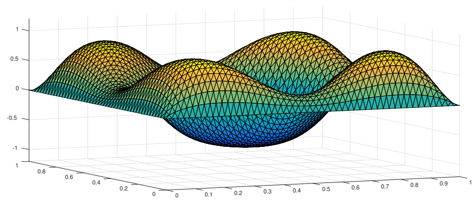
(b) t = 0.01



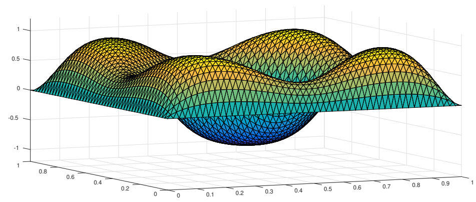
(c) t = 0.6724



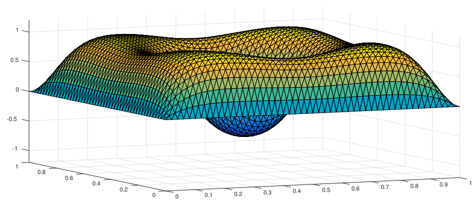
(d) t = 0.6724



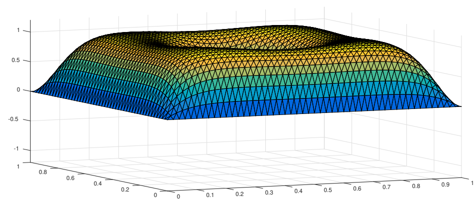
(e) t = 0.6724



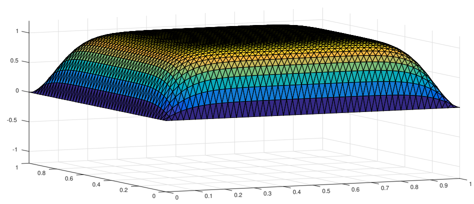
(f) t = 0.6724



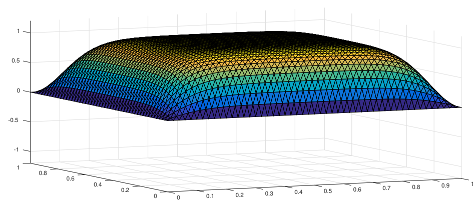
(g) t = 0.6724



(h) t = 0.6724



(i) t = 0.6729



(j) t = 1.00

Figure 4.1.1: Computed parametrized solution to the problem described in Section 4.1.2. Figures 4.1.1c–4.1.1i show the viscous transition corresponding to the discontinuity at time $t \approx 0.6724$.

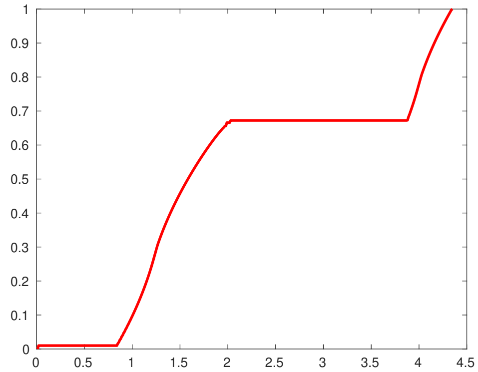


Figure 4.1.2: Evolution of the physical time as function of the artificial time. The physical time stands still during the viscous transitions at time $t \approx 0.01$ and $t \approx 0.6724$

lumping scheme for the discretization of $\mathcal{R}$ as described in Section 4.1.2. The resulting errors are shown in Figure 4.2.1. It can be seen that the error decreases in a linear fashion (w.r.t. the time parameter τ) until the error of the spatial-discretization is dominant.

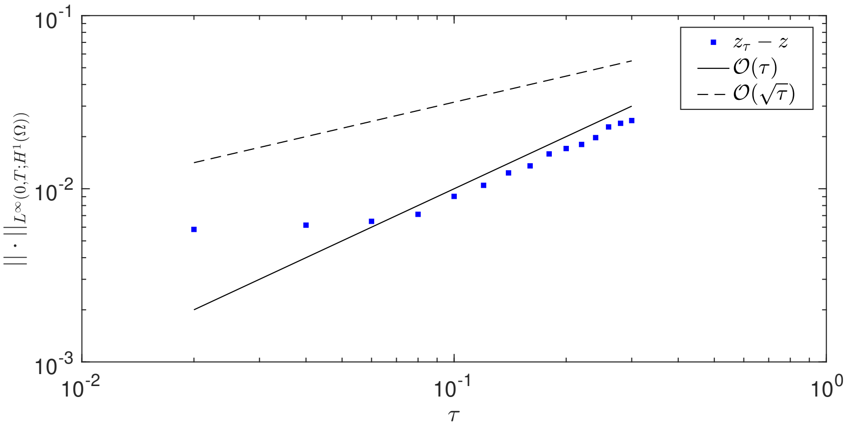


Figure 4.2.1: Errors for the approximation of the parametrized solution (4.2.1) using the local minimization scheme.

4.2.2 Local case

1D-Example

We next give a one-dimensional example, in which the energy is not globally uniformly convex. In particular, the energetic solution will no longer be continuous in time, which is seen in Figure 4.2.2. However, the parametrized solution is still Lipschitz continuous and, moreover, remains in a region, where the energy is uniformly convex, see Figure 4.2.2. For this example, we set $\mathcal{Z} = \mathcal{V} = \mathcal{X} = \mathbb{R}$

as well as

$$\mathcal{R}(v) = |v| \quad \text{and} \quad \mathcal{I}(t,z) = \frac{1}{2}z^2 + \mathcal{F}(z) - \ell(t)z \tag{4.2.2}$$

with

$$\mathcal{F}(z) = 2|z|^3 - 5/2\, z^2 + 1 \quad \text{and} \quad \ell(t) = -1/2(t-3/2)^2 + 3/2.$$

For $z_0 = -2/3$, a (differential) solution to (RIS) with (4.2.2) reads

$$z(t) = \begin{cases} -2/3 & , t \in [0,1/2), \\ -\frac{1}{3}(1 + 1/2\sqrt{1+3(t-3/2)^2}) & , t \in [1/2,2), \\ -1/2 & , t \in [2,3]. \end{cases} \tag{4.2.3}$$

By direct calculations, one verifies that z, indeed, stays in a region where $\mathcal{I}$ is uniformly convex. Thus, from the analysis in Section 3.3.3, we expect the error in the approximation to be of order $\mathcal{O}(\tau)$, which can be nicely observed in Figure 4.2.2. In contrast, due to the time discontinuity, an L^∞-error estimate in the form of (3.3.43) cannot hold for the global minimization scheme (see Figure 4.2.2, right). Recall that the iterates of the global scheme are defined by

$$z_k \in \arg\min\{\mathcal{I}(t_{k-1}, z) + \mathcal{R}(z - z_{k-1}) \,:\, z \in \mathcal{Z}\}, \quad t_k = t_{k-1} + \tau. \tag{4.2.4}$$

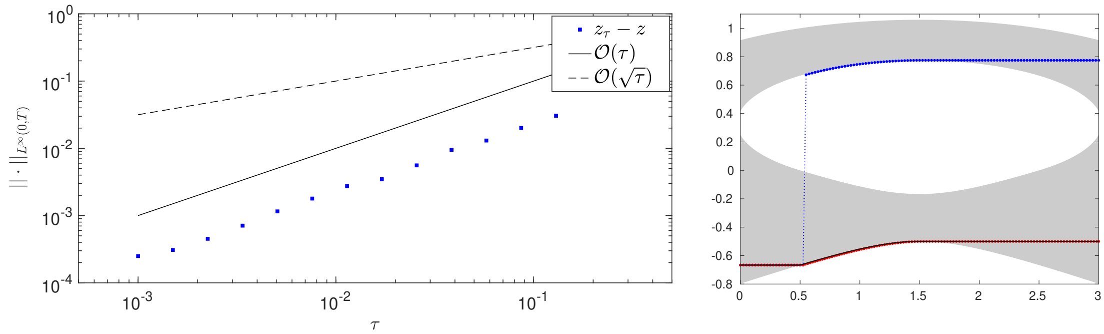


Figure 4.2.2: Left: Errors for the approximation of a parametrized solution using the local minimization scheme depending on the step size τ. Right: Corresponding differential solution (black) as well as the numerical approximations using the global (blue) and the local iterated minimization scheme (red) as functions of the time t.

2D-Example

In view of the previous example, one may wonder if it is possible to obtain error estimates in an L^p-norm, $1 \le p < \infty$, for the global minimization scheme, provided that the energy is locally uniformly convex. The following two-dimensional example demonstrates that this is not the case for any $p \ge 1$

at least not for the whole sequence of approximations. To this end, we set $\mathcal{Z} = \mathcal{V} = \mathcal{X} = \mathbb{R}^2$ as well as

$$\mathcal{R}(v) = |v|_1 \quad \text{and} \quad \mathcal{I}(t,z) = \frac{1}{2}\|z\|^2 + \mathcal{F}(z) - \ell(t)z \tag{4.2.5}$$

with

$$\mathcal{F}(z) = 2\|z\|^4 + 12z_1^2 z_2^2 - 9/2\|z\|^2 + 3 \quad \text{and} \quad \ell(t) = \frac{-1}{128}(l(t),\, l(t))^\top,$$

where $l(t) = t^3 - 27t^2 + 179t - 25$. For $z_0 = (1,0)^\top$, a (differential) solution to (RIS) with (4.2.5) reads

$$z(t) = \begin{cases} (1,\, 0)^\top, & t \in [0,1), \\ (1 + (1-t)/16,\, (1-t)/16)^\top, & t \in [1,2]. \end{cases} \tag{4.2.6}$$

Again, z stays in a region where $\mathcal{I}$ is uniformly convex, and thus the approximation error of the local minimization scheme is of order $\mathcal{O}(\tau)$, see Figure 4.2.3. In contrast to this, the global minimization scheme does not, in general, converge in any L^p-norm, $p \geq 1$, because of the ambiguity of the global minimizers. To be more precise, while the global minimization problem in (4.2.4) admits a unique global minimizer in $z_0^* = z_0$ for $t < 1$, it exhibits three different global minima at $t = 1$, namely $z_0^* = z_0$, $z_1^* = (-1,0)^\top$, and $z_2^* = (0,-1)^\top$ (see Figure 4.2.3, right). For $t > 1$, the global minimum z_0^* vanishes, but z_1^* and z_2^* both remain globally minimal. Thus, when the algorithm reaches $t = 1$, the iterates either jump to z_1^* or z_2^*, depending on the concrete algorithmic realization (e.g., choice of the optimization algorithm, initial value, etc). Therefore, either of two different energetic solutions is approximated by the global minimization scheme illustrating that an L^p-error estimate, $1 \leq p \leq \infty$, cannot, in general, be expected for this discretization scheme, at least not for the whole sequence of discrete solutions.

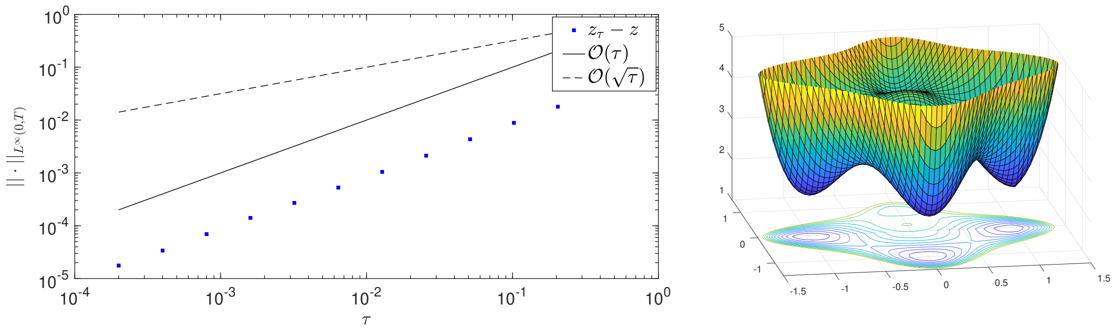


Figure 4.2.3: Left: Errors for the approximation of a parametrized solution using the local minimization scheme depending on the step size τ. Right: Surface and contour plot of $\mathcal{I}(t,z)+\mathcal{R}(z-z_0)$ at time $t = 1$ with $\mathcal{I}$ and $\mathcal{R}$ from (4.2.5).

Chapter 5

Conclusion and outlook

As we have seen at several points in this thesis, rate-independent systems naturally inhere a certain nonsmoothness which also transfers to possible solutions of the very same. In the first part of this work, we therefore concerned ourselves with some concepts of solutions being capable of handling time discontinuities. Hereby, we primarily focused on the notions of *differential*, *energetic* and *parametrized solutions*. While each of these provides their own characteristics, we brought out several interrelations depending on the properties of the underlying energy functional, particularly convexity.

With all the necessary foundations at hand, we subsequently presented a full space and time discretization scheme (LISS) based on the local incremental minimization scheme introduced in [EM06]. The major differences between these two are, on the one hand, that we allow for unbounded dissipation functionals here and, on the other hand, that we merely require the iterates to be stationary points of the underlying minimization problem. The latter property, especially, has the essential advantage that it is a natural outcome in the context of optimization algorithms (limit points are, in general, stationary). By adapting the analysis of the recent contribution [Kne19] and using arguments from [KRZ13], we proved that (weak) accumulation points of the sequence of discrete solutions for mesh and time step size tending to zero exist and are parametrized solutions of the original rate-independent system. While this is at first glance a result that verifies the consistency of the *local incremental stationarity scheme*, we, moreover, gain an existence result for parametrized solutions in the case of a nonconvex energy and unbounded dissipation.

We subsequently turned our attention to the derivation of a priori estimates for the *local incremental stationarity scheme*. Here we proved that for the semilinear energy functional, i.e., $\mathcal{I}(t,z) = \frac{1}{2}\langle Az, z\rangle + \mathcal{F}(z) - \langle \ell(t), z\rangle$, the optimal convergence rate of $\mathcal{O}(\tau)$ can be obtained if $\mathcal{I}$ is additionally (globally) uniformly convex or if $\mathcal{I}$ is only locally uniformly convex along a given solution trajectory and the time step τ is sufficiently small. Unfortunately, both cases additionally require some smallness assumption on the external load ℓ. In the general case, i.e., $\mathcal{I}(t,z) = \frac{1}{2}\langle Az, z\rangle + \mathcal{F}(z) - f(t,z)$, this convergence rate reduces to $\mathcal{O}(\sqrt{\tau})$. In summary, the overall picture concerning the a priori estimates (at least in the semilinear case) for the local incremental stationarity scheme now looks as follows:

- For an arbitrary nonconvex energy, there exists a subsequence of discrete solutions that converges (weakly) to a parametrized solution as $\tau \searrow 0$.

- If the energy is locally uniformly convex along a solution trajectory, then the discrete solution converges with optimal rate to this solution, provided that the time step size is sufficiently small. At this point, the local incremental minimization scheme turns out to be superior to the global one, since the latter does, in general, not satisfy such an a priori estimate.
- If the energy is uniformly convex, one obtains the same convergence rates as for the global incremental minimization scheme.

Finally, we turned to the actual realization of the *local incremental stationarity scheme*, where we employed standard piecewise linear finite elements for the spatial discretization. For the discretization of the dissipation functional, we made use of a mass lumping scheme which, on the one hand, turned out to be advantageous for the numerical realization of the algorithm and, on the other hand, can be incorporated into the (abstract) convergence analysis. This mass lumping allows a reformulation of the discrete optimality system arising in each step of the local stationarity scheme as nonsmooth equation, which is amenable for semismooth Newton methods. For the case of a double-well potential, we compared the local minimization scheme with another time discretization scheme known to converge to global energetic solutions. We observe that both schemes, indeed, provide different solutions which jump at different points in time.

After all, there remain several unanswered questions:

- This concerns for example the generalization of the convergence theorem for more general energy functionals, in particular quasilinear instead of semilinear. As already indicated, the main ingredient of the convergence analysis is a Gårding-like inequality, so that an improvement in this direction might be possible.
- Moreover, the gap between the semilinear and the general setting remains as an open problem here. However, under suitable smoothness assumptions on f and applying arguments from the proof in the semilinear case, it should be possible to extend the results. Still, this does not affect the rather unsatisfactory smallness assumption on ℓ and some more investigations into this direction are needed.
- Finally, as we did not include the spatial discretization into the a priori analysis, one might take a deeper look into a priori estimates combining both, discretization in space and time.

Clearly, this short list is by far not complete. Despite the problems arising in this thesis, it is for example interesting to investigate dissipation functionals which are also depending on the state z, i.e., $\mathcal{R} = \mathcal{R}(z, z')$. Nevertheless, it hopefully gives some impulses for problems that can be addressed in future research on this topic.

Appendix

Throughout this appendix, X denotes a normed vector space over $\mathbb{R}$ if not otherwise specified. Moreover, we adopt the setting from the introduction, that is, $\mathcal{Z}, \mathcal{V}$ are Hilbert spaces and $\mathcal{X}$ is a Banach space with $\mathcal{Z} \hookrightarrow^{c,d} \mathcal{V} \hookrightarrow \mathcal{X}$.

A.1 Spaces

We use this section in order to briefly recall some of the function spaces used in this thesis as well as some minor results. Preliminary, we define the *set of partitions of* $[a, b]$ as

$$\mathcal{P}(a, b) := \{\{t_0, t_1, \dots, t_n\} \,:\, a = t_0 < t_1 < \cdots < t_n = b\} \tag{A.1.1}$$

and call $\{t_0, t_1, \dots, t_n\} \in \mathcal{P}(a, b)$ a *partition.*

Definition A.1.1. *Let $[a, b] \subset \mathbb{R}$ and $z : [a, b] \to X$. We define the **variation** of z as*

$$\mathrm{Var}_X(z, [a, b]) := \sup_{\mathcal{P}(a,b)} \left\{ \sum_{i=1}^{n} \|z(t_i) - z(t_{i-1})\|_X \right\}.$$

Therewith, we denote

$$BV(a, b; X) := \{z : [a, b] \to X \,:\, \mathrm{Var}_X(z, [a, b]) < \infty\},$$

*the **space of functions of (pointwise) bounded variation**.*

If the space X in the above definition is a Banach space, then $BV(a, b; X)$ is a Banach space endowed with the norm $\|z\|_{BV(a,b;X)} := \|z(a)\| + \mathrm{Var}_X(z, [a, b])$. Moreover, we simply write $BV(a, b)$ if $X = \mathbb{R}$. The next result is a well-known property of functions with bounded variation and can for example be found in [Leo17, Thm. 2.17].

Lemma A.1.2. *Let X be a Banach space and $z \in BV(a, b; X)$. Then z is continuous at all but countably many points and there exist*

$$z(t^+) = \lim_{s \to t,\, s > t} z(s) \quad \text{and} \quad z(t^-) = \lim_{s \to t, s < t} z(s)$$

for all $t \in [a, b]$ with obvious modifications if $t = a, b$.

Additionally, a function in BV allows the following estimate

Lemma A.1.3. *Let $z \in BV(a,b;X)$ and $h \in (0, b-a)$. Then*

$$\frac{1}{h}\int_{a+h}^{b} \|z(t) - z(t-h)\|_X \,\mathrm{d}s \le \mathrm{Var}_X(z, [a,b]). \tag{A.1.2}$$

Proof. see [Kre99, Thm. 8.12]. □

We next turn to the Bochner Sobolev spaces. For these, we do not explictly repeat the definition but merely refer the reader to [GGZ74]. In this context [Wac11] also provides a well chosen composition of various results. Nevertheless, we recall some properties. For the rest of this section, we let X be a Banach space. Then $W^{1,p}(0,T;X)$ is also a Banach space and, moreover, reflexive provided $p \in (1,\infty)$ and X is also reflexive.

Lemma A.1.4. *Let $z : [a,b] \to X$ be Bochner integrable, then*

$$\left\| \int_a^b z(t)\,\mathrm{d}t \right\|_X \le \int_a^b \|z(t)\|_X \,\mathrm{d}t.$$

If furthermore Y is a second Banach space and $A : X \to Y$ is a linear continuous operator, then

$$A\left(\int_a^b z(t)\,\mathrm{d}t\right) = \int_a^b A(z(t))\,\mathrm{d}t.$$

As in the real-valued case, we also obtain the Lebesgue's dominated convergence theorem:

Theorem A.1.5 (Lebesgue dominated convergence)**.** *Let $p \in [1,\infty)$ and $z_n \in L^p(a,b;X)$ as well as $m_n \in L^p(a,b)$ be two sequences with*

$$\begin{aligned} m_n &\to m \in L^p(a,b), \\ z_n(t) &\to z(t) \quad \text{f.a.a. } t \in [a,b], \\ \forall n \in \mathbb{N} \quad \|z_n(t)\|_X &\le m_n(t) \quad \text{f.a.a. } t \in [a,b]. \end{aligned}$$

Then z_n converges to z in $L^p(a,b;X)$.

Moreover, we have the following properties:

Lemma A.1.6. *There holds:*

- *Let $p \in [1,\infty)$ then $C^\infty(a,b;X)$ is dense in $W^{1,p}(a,b;X)$.*
- *If $v \in L^1(a,b;X)$ and $z(t) = z_0 + \int_a^b v(s)\,\mathrm{d}s$, then $z \in C(a,b;X) \cap W^{1,1}(a,b;X)$ with $z'(t) = v(t)$ almost everywhere in $[a,b]$. Moreover, z is absolutely continuous.*
- *If, in contrast, $z \in W^{1,1}(a,b;X)$ then $z(t) = z(a) + \int_a^b z'(s)\,\mathrm{d}s$ holds almost everywhere in $[a,b]$.*
- *Furthermore, if $z \in W^{1,1}(a,b;X)$ with $z'(t) = 0$ almost everywhere in $[a,b]$, then z is constant almost everywhere.*

- *There holds the following continuous embedding:* $W^{1,1}(a,b;X) \hookrightarrow C(a,b;X)$.
- *If additionally* $p \in (1,\infty]$ *and* Y *is another Banach space with* $Y \hookrightarrow^d X$*, then it holds* $W^{1,p}(a,b;Y) \hookrightarrow^c C(a,b;X)$.

In particular, we have:

Lemma A.1.7. *Let* $f \in L^1(0,T;X)$ *and* $\int_s^t f(r)\mathrm{d}r = 0$ *for all* $0 \leq s < t \leq T$ *hold, then* $f = 0$ *almost everywhere in* $[0,T]$.

Proof. We define $w(t) := \int_0^t f(s)\mathrm{d}s$. Then $w \in W^{1,1}(0,T)$ by construction and, moreover, $w \equiv 0$. Hence, since the weak derivative is unique, we obtain $0 = w' = f$ almost everywhere in $[0,T]$. □

Finally, we recall the following version of the Lebesgue differentiation theorem for Bochner-Sobolev functions:

Theorem A.1.8. *Let* $z \in W^{1,1}(a,b;X)$ *be given. Then*

$$z'(t) = \lim_{h \searrow 0} \frac{z(t+h) - z(t)}{h} \tag{A.1.3}$$

holds for almost all $t \in (a,b)$.

A.2 Elements of functional analysis

Ehrling lemma

The following estimate on the norm of the intermediate space $\mathcal{Z} \subset \mathcal{V} \subset \mathcal{X}$ is an essential ingredient in the convergence analysis of the local stationarity scheme LISS. Although there exist slightly more general versions as for example in [Alt12, p. 365] or [Rou13, Lem. 7.6], we stick with this one:

Lemma A.2.1. *Let* $\mathcal{X}, \mathcal{V}$ *and* $\mathcal{Z}$ *be Banach spaces with* $\mathcal{V} \hookrightarrow \mathcal{X}$ *and* $\mathcal{Z} \hookrightarrow^c \mathcal{V}$*. Then for every* $\delta > 0$ *there exists* $C_\delta > 0$ *such that*

$$\|v\|_{\mathcal{V}} \leq \delta \|v\|_{\mathcal{Z}} + C_\delta \|v\|_{\mathcal{X}} \tag{A.2.1}$$

for all $v \in \mathcal{Z}$.

Proof. by contradiction. Thus, let $\varepsilon > 0$ be given. Then there exists $v_k \in \mathcal{Z}$ with

$$\|v_k\|_{\mathcal{V}} > \varepsilon \|v_k\|_{\mathcal{Z}} + k \|v_k\|_{\mathcal{X}}.$$

Due to the embedding $\mathcal{V} \hookrightarrow \mathcal{X}$, particularly $\|v_k\|_{\mathcal{X}} \leq c \|v_k\|_{\mathcal{V}}$, we may w.l.o.g. assume that $\|v_k\|_{\mathcal{Z}} > 0$. Hence, taking $\tilde{v}_k := v_k / \|v_k\|_{\mathcal{Z}}$ it holds

$$\|\tilde{v}_k\|_{\mathcal{V}} > \varepsilon + k \|\tilde{v}_k\|_{\mathcal{X}}.$$

Since $\tilde{v}_k$ is bounded in $\mathcal{Z}$ and $\mathcal{Z} \hookrightarrow^c \mathcal{V}$, there exists a subsequence (again denoted by $\tilde{v}_k$) with $\tilde{v}_k \to \tilde{v}$ for some $\tilde{v} \in \mathcal{V}$. In particular, $\tilde{v}_k$ is bounded and also converges to $\tilde{v}$ in $\mathcal{X}$, so that we can conclude

$$0 \longleftarrow \frac{1}{k}\|\tilde{v}_k\|_{\mathcal{V}} \geq \|\tilde{v}_k\|_{\mathcal{X}} \to \|\tilde{v}_k\|_{\mathcal{X}}.$$

This implies that $\tilde{v} = 0$ and therefore $0 < \varepsilon < \|\tilde{v}_k\|_{\mathcal{V}} \to 0$, which proves the claim. □

Chain rules and change of variables formula

We now turn to the different versions of the chain rule, occurring at several points in this thesis. In preparation for this, however, we need to specify the following definition, wherein $\mathcal{H}^1$ denotes the onedimensional Hausdorff measure.

Definition A.2.2 (Lusin (N) property ([Leo17, Def. 3.34])). *Let $I \subset \mathbb{R}$ and (X, d) be a metric space. Moreover, let $u : I \to X$. We say that u satisfies the* ***Lusin (N) property*** *if*

$$\mathcal{H}^1(u(J)) = 0$$

for every set $J \subseteq I$ with Lebesgue-measure zero, i.e., $\mathcal{L}^1(J) = 0$.

Now, we can state the following very general result, which can be found in [Leo17]:

Theorem A.2.3 (General chain rule). *Let $I, J \subset \mathbb{R}$ be two intervals, let X be a normed space, and let $f : J \to X$ and $u : I \to J$ be such that f, u, and $f \circ u$ are differentiable almost everywhere in their respective domains. If f satisfies the Lusin (N) property, then for almost every $x \in I$,*

$$(f \circ u)'(x) = f'(u(x))u'(x), \tag{A.2.2}$$

where $f'(u(x))u'(x)$ is interpreted to be zero whenever $u'(x) = 0$ (even if f is not differentiable at $u(x)$).

Proof. see [Leo17, Thm 3.59] □

From this, we easily deduce:

Corollary A.2.4. *Let f, u be as in Theorem A.2.3. Let additionally f be Lipschitz continuous. Then (A.2.2) holds true, i.e.,*

$$(f \circ u)'(x) = f'(u(x))u'(x), \tag{A.2.3}$$

where $f'(u(x))u'(x)$ is interpreted to be zero whenever $u'(x) = 0$ (even if f is not differentiable at $u(x)$).

Proof. Since Lipschitz continuous functions provide the Lusin (N) property, this is an immediate consequence of Theorem A.2.3. □

Now, we turn to more concrete versions of the chain rule.

Lemma A.2.5. *Let $\mathcal{I} \in C^1([0,S] \times \mathcal{Z}, \mathbb{R})$ and $(t,z) \in W^{1,\infty}(0,S) \times W^{1,1}(0,S;\mathcal{Z})$. Then the following chain rule*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}(t(s), z(s)) = \partial_t \mathcal{I}(t(s), z(s))t'(s) + \langle D_z \mathcal{I}(t(s), z(s), z'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \quad \textit{a.e. in } [0,S]$$

holds. In particular, it holds

$$\mathcal{I}(t(s_2), z(s_2)) - \mathcal{I}(t(s_1), z(s_1)) \\ = \int_{s_1}^{s_2} \partial_t \mathcal{I}(t(s), z(s))\, t'(s) + \langle D_z \mathcal{I}(t(s), z(s), z'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \,\mathrm{d}s \tag{A.2.4}$$

for all $0 \le s_1 < s_2 \le S$.

Proof. From the theory for vector-valued Sobolev spaces we have $W^{1,1}(0,S;\mathcal{Z}) \hookrightarrow C(0,S;\mathcal{Z})$. Now let $t \in W^{1,1}(0,S)$ and $z \in W^{1,1}(0,S;\mathcal{Z})$ be arbitrary. Then, by density, there exist $t_n \in C^\infty(0,S)$ and $z_n \in C^\infty(0,S;\mathcal{Z})$ converging strongly to t in $W^{1,1}(0,S)$ and z in $W^{1,1}(0,S;\mathcal{Z})$, respectively. From the above embedding we infer that this convergence is even pointwise, so that for all $s \in [0,S]$

$$D_z \mathcal{I}(t_n(s), z_n(s)) \to D_z \mathcal{I}(t(s), z(s)) \quad \text{in } \mathcal{Z}^*, \\ \partial_t \mathcal{I}(t_n(s), z_n(s)) \to \partial_t \mathcal{I}(t(s), z(s)),$$

since $\mathcal{I} \in C^1([0,T] \times \mathcal{Z}; \mathbb{R})$. Since both terms are also uniformly bounded, we have weak* convergences of $D_z\mathcal{I}(t_n(s), z_n(s))$ to $D_z\mathcal{I}(s, z(s))$ and $\partial_t \mathcal{I}(t_n(s), z_n(s))$ to $\partial_t \mathcal{I}(t(s), z(s))$ in $L^\infty(0,S;\mathcal{Z}^*)$ and $L^\infty(0,S)$, respectively. In combination with the strong convergence of z_n' to z' in $L^1(0,S;\mathcal{Z})$ we obtain

$$\langle D_z \mathcal{I}(s, z_n(s)), z_n'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \to \langle D_z \mathcal{I}(s, z(s)), z'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \qquad \text{in } L^1(0,S)$$

and correspondingly

$$\partial_t \mathcal{I}(t_n(s), z_n(s))\, t_n'(s) \to \partial_t \mathcal{I}(t(s), z(s))\, \hat{t}'(s) \qquad \text{in } L^1(0,S)\,.$$

For every t_n, z_n we can therefor apply the chain rule so that in the end

$$\begin{aligned} &\mathcal{I}(t_n(s_2), z_n(s_2)) - \mathcal{I}(t_n(s_1), z_n(s_1)) \\ &= \int_{s_1}^{s_2} \frac{\mathrm{d}}{\mathrm{d}s} \mathcal{I}(t_n(s), z_n(s)) \,\mathrm{d}s \\ &= \int_{s_1}^{s_2} \partial_t \mathcal{I}(t_n(s), z_n(s))\, t_n'(s) + \langle D_z \mathcal{I}(t_n(s), z_n(s)), z_n'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \,\mathrm{d}s \\ &\to \int_{s_1}^{s_2} \partial_t \mathcal{I}(t(s), z(s))\, t'(s) + \langle D_z \mathcal{I}(t(s), z(s)), z'(s)\rangle_{\mathcal{Z}^*,\mathcal{Z}} \,\mathrm{d}s \end{aligned}$$

Finally, since $\mathcal{I}$ is continuous, it holds

$$\mathcal{I}(t_n(s_2), z_n(s_2)) - \mathcal{I}(t_n(s_1), z_n(s_1)) \to \mathcal{I}(t(s_2), z(s_2)) - \mathcal{I}(t(s_1), z(s_1)),$$

which proves the claim. □

Thereof, we can easily conclude the following variant.

Corollary A.2.6. *Let $\mathcal{I} \in C^1([0,T] \times \mathcal{Z}, \mathbb{R})$ and $z \in W^{1,1}(0,T;\mathcal{Z})$. Then the following chain rule holds*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}(t, z(t)) = \partial_t \mathcal{I}(t, z(t)) + \langle D_z \mathcal{I}(t, z(t), z'(t) \rangle_{\mathcal{Z}^*, \mathcal{Z}} \quad \textit{a.e. in } [0,T]. \tag{A.2.5}$$

Lastly, we state a chain rule for the composition of a $W^{1,p}(0,T;\mathcal{Z})$-function with a time-transformation in $W^{1,\infty}(0,S)$.

Lemma A.2.7. *Let $\mathcal{Z}$ be a reflexive Banach space. Moreover, suppose that $z \in W^{1,p}(0,T;\mathcal{Z})$ and $\hat{t} \in W^{1,\infty}(0,S)$ with $\hat{t}(0) = 0$, $\hat{t}(S) = T$ and $\hat{t}'(s) \geq 0$ for almost all $s \in [0,S]$. Then $z \circ \hat{t} \in W^{1,p}(0,T;\mathcal{Z})$ with $(z \circ \hat{t})'(s) = z'(\hat{t}(s))\hat{t}'(s)$ for almost all $s \in [0,S]$.*

Proof. We let $\delta \in (0,1]$ be fixed but arbitrary and extend z to $\mathbb{R}$ by constant continuation (for simplicity also denoted by z). Moreover, we set $\hat{t}_\delta(s) := \hat{t}(s) + \delta\, s$ so that $\hat{t}' \geq \delta > 0$ by the assumption on $\hat{t}$. Hence, an application of [Leo17, Cor. 3.64] implies that $(z \circ \hat{t}_\delta)$ is differentiable almost everywhere and it holds

$$(z \circ \hat{t}_\delta)'(s) = z'(\hat{t}_\delta(s))\hat{t}_\delta'(s) = z'(\hat{t}_\delta(s))(\hat{t}'(s) + \delta) \quad \text{f.a.a. } s \in [0,S]. \tag{A.2.6}$$

This leads to

$$\|(z \circ \hat{t}_\delta)'\|_{L^p(0,S;\mathcal{Z})} \leq \big(\|\hat{t}'\|_{L^\infty(0,S)} + \delta\big)\|z'\|_{L^p(0,T;\mathcal{Z})} \leq C < \infty.$$

so that $(z \circ \hat{t}_\delta) \in W^{1,p}(0, S + \delta S; \mathcal{Z})$. Now, consider a sequence $\delta \searrow 0$. Then $\{z \circ \hat{t}_\delta\}$ is bounded in $W^{1,p}(0,S;\mathcal{Z})$ and consequently, there is a subsequence converging weakly(*) in $W^{1,p}(0,S;\mathcal{Z})$. On the other hand, $\hat{t}_\delta$ converges uniformly to $\hat{t}$ on $[0,S]$ so that the continuity of z (due to the embedding of $W^{1,p}(0,T+S;\mathcal{Z}) \hookrightarrow C(0,T+S;\mathcal{Z})$) gives the pointwise convergence of $z \circ \hat{t}_\delta$ to $z \circ \hat{t}$ on $[0,S]$. Since the weak and the pointwise limit coincide, this gives $z \circ \hat{t} \in W^{1,p}(0,T;\mathcal{Z})$. Finally, the identity $(z \circ \hat{t})'(s) = z'(\hat{t}(s))\hat{t}'(s)$ follows from Theorem A.2.3 by exploiting that z is absolutely continuous and therefore satisfies the Lusin (N) property; cf. [Leo17, Thm. 8.42]. □

We close this section with the following two versions of the change of variable formula. The first of these can be found in [EG92] in a slightly more general form.

Theorem A.2.8 (Change of variables)**.** *Let $t : [0,S] \to [0,T]$ with $S \geq T$ be Lipschitz continuous and monotone increasing. Then for every $p \in L^1(0,T)$ it holds*

$$\int_0^S p(s)t'(s)\,\mathrm{d}s = \int_0^T \left[\sum_{x \in t^{-1}(y)} p(x)\right] \mathrm{d}y. \tag{A.2.7}$$

Proof. see [EG92, p. 99]. □

If the function t in the prior theorem is not necessarily Lipschitz continuous, we may still obtain the subsequent result, which is the only point where the space of absolutely continuous functions $AC(a,b)$ is mentioned, so that we refer to [Leo17, Def. 3.1] for a definition.

Theorem A.2.9. *Let $t : [a,b] \to [c,d]$ be differentiable almost everywhere and $p \in L^1(c,d)$. Then $(p \circ t)\, t'$ is integrable and it holds*

$$\int_{t(\alpha)}^{t(\beta)} p(s)\,\mathrm{d}s = \int_{\alpha}^{\beta} p(t(\sigma))\,t'(\sigma)\,\mathrm{d}\sigma \tag{A.2.8}$$

for all $\alpha, \beta \in [a,b]$ if and only if the function $f \circ t$ belongs to $AC(a,b)$, where

$$f(r) = \int_c^r p(x)\,\mathrm{d}x.$$

Proof. see [Leo17, Thm. 3.75]. □

Corollary A.2.10. *Let $t : [a,b] \to [c,d]$ be monotone increasing and absolutely continuous and $p \in L^1(c,d)$. Then $(p \circ t)\, t'$ is integrable and the change of variables formula (A.2.8) holds.*

General Helly selection theorem

The following result is an essential ingredient in the existence proof of energetic solutions.

Lemma A.2.11 (General Helly selection theorem [MR15, Thm. B.5.13])**.** *Let $\mathcal{R}$ comply with assumptions (R1)-(R3) and $\mathrm{Diss}_{\mathcal{R}}$ be defined as in (2.3.1). Moreover, let K be a weakly sequentually compact subset of $\mathcal{Z}$ and $z_k : [0,T] \to \mathcal{Z}$ be a sequence with*

$$z_k(t) \in K \quad \forall t \in [0,T] \ \ \textit{and}\ k \in \mathbb{N}, \tag{A.2.9}$$

$$\sup_{k\in\mathbb{N}} \mathrm{Diss}_{\mathcal{R}}(z_k, [0,T]) < \infty. \tag{A.2.10}$$

Then there exists a subsequence $\{z_{k_l}\}_{l\in\mathbb{N}}$, a limit function $z : [0,T] \to \mathcal{Z}$ and a nondecreasing function $\delta : [0,T] \to [0,\infty)$ so that for all $t \in [0,T]$:

$$\delta(t) = \lim_{l\to\infty} \mathrm{Diss}_{\mathcal{R}}(z_{k_l}, [0,t]), \tag{A.2.11}$$

$$z_{k_l}(t) \rightharpoonup z(t) \ \ \textit{and}\ \ z(t) \in K, \tag{A.2.12}$$

$$\mathrm{Diss}_{\mathcal{R}}(z, [s,t]) \le \delta(t) - \delta(s) \quad \forall 0 \le s < t. \tag{A.2.13}$$

Proof. see [MR15, Thm. B.5.13]. □

A.3 Elements of convex analysis

A crucial ingredient of the thesis at hand is the convex subdifferential. In order to clarify the notation, we give a Definition of the very same.

Definition A.3.1. *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a convex function. Then we denote by*

$$\partial f(x) := \{g \in X^* : f(v) \ge f(x) + \langle g, v - x\rangle\ \forall v \in X\} \subset X^*$$

the ***convex subdifferential*** *of f at $x \in X$. Sometimes, we will also write $\partial^X f(x)$ in order to stress the underlying space.*

Moreover, we denote the domain of a functional f by $\mathrm{dom}(f)$, i.e.,

$$\mathrm{dom}(f) := \{x \in X \,:\, f(x) < +\infty\}. \tag{A.3.1}$$

Finally, we introduce the concept of conjugate and biconjugate functionals and refer to [BV10] for more details. Since the degenerate cases are rather uninteresting, we restrict our definition to *proper* functionals f, i.e., $f(x) > -\infty$ for all $x \in X$ and there exists at least one $x \in X$ with $f(x) < \infty$.

Definition A.3.2. *Let $f : X \to (-\infty, \infty]$ be a proper functional. We denote by $f^* : X^* \to (-\infty, \infty]$ the* ***conjugate functional***

$$f^*(x^*) := \sup_{x \in X} \left(\langle x^*, x \rangle_{X^*,X} - f(x) \right).$$

Moreover, we define $f^{**} : X \to (-\infty, \infty]$ as

$$f^{**}(x) := \sup_{x^* \in X^*} \left(\langle x^*, x \rangle_{X^*,X} - f^*(x^*) \right)$$

and call it the ***biconjugate functional****.*

An easy consequence of the definition of the conjugate functional is the following *Fenchel-Young inequality*:

Lemma A.3.3 (Fenchel-Young inequality)**.** *Let $f : X \to \mathbb{R}$ be a proper functional. Let moreover denote f^* the conjugate of f. Then*

$$f(x) + f^*(\xi) \geq \langle \xi, x \rangle_{X^*,X} \quad \forall v \in X,\, \xi \in X^* \tag{A.3.2}$$

and equality holds if and only if $\xi \in \partial f(x)$.

Lemma A.3.4 (Jensen inequality)**.** *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex and lower semicontinuous functional. Then, for every integrable $z : [0, T] \to X$, it holds*

$$f\left(\frac{1}{(t-s)} \int_s^t z(r)\, \mathrm{d}r \right) \leq \frac{1}{(t-s)} \int_s^t f(z(r))\, \mathrm{d}r \tag{A.3.3}$$

for all $0 \leq s < t \leq T$. Note that we define the right-hand side to equal ∞ if $f \circ z$ is not integrable.

Proof. Let $s, t \in [0, T]$ with $s < t$ be given. We define $z_0 := \frac{1}{(t-s)} \int_s^t z(r)\, \mathrm{d}r \in X$. It is easy to see that, if $z_0 \notin \mathrm{dom}(f)$ then there exists $E \subset (s, t)$ with nonzero measure, such that $z(r) \notin \mathrm{dom}(f)$ for almost all $r \in (s, t)$. In this case, (A.3.3) is trivial. Thus let $z_0 \in \mathrm{dom}(f)$. Since f is proper, convex and lower semicontinuous, a well-known result from convex analysis gives $f^{**} = f$, that is, for all $x \in X$ it holds $f(x) = \sup_{x^* \in X^*} \left(\langle x^*, x \rangle_{X^*,X} - f^*(x^*) \right)$. Now, let $x^* \in X^*$ be arbitrary.

Then $f(x) \geq (\langle x^*, x\rangle_{X^*,X} - f^*(x^*))$ and therefore

$$\begin{aligned}\int_s^t f(z(r))\;\mathrm{d}r &\geq \int_s^t (\langle x^*, z(r)\rangle_{X^*,X} - f^*(x^*))\;\mathrm{d}r\\ &= \langle x^*, \int_s^t z(r)\;\mathrm{d}r\rangle_{X^*,X}\;\mathrm{d}r - (t-s)f^*(x^*)\\ &= (t-s)\left(\langle x^*, z_0\rangle_{X^*,X} - f^*(x^*)\right).\end{aligned}$$

Taking the supremum, we find

$$\frac{1}{(t-s)}\int_s^t f(z(r))\;\mathrm{d}r \geq \sup_{x^*\in X^*} (\langle x^*, z_0\rangle_{X^*,X} - f^*(x^*)) = f(z_0) = f\left(\frac{1}{(t-s)}\int_s^t z(r)\;\mathrm{d}r\right).$$

This gives (A.3.3). □

The integration of the superposition with a convex functional as in the Jensen inequality (A.3.3) occurs at various points throughout this thesis, which is why we take a look at its properties now.

Lemma A.3.5. *Let $f : X \to \mathbb{R}\cup\{+\infty\}$ be a proper, convex and lower semicontinuous functional and $[a,b] \subset \mathbb{R}$. If, moreover, f is nonnegative, then the functional*

$$F : L^1(a,b;X) \to \mathbb{R}\cup\{+\infty\}, \quad v \mapsto \begin{cases} \int_a^b f(v(t))\;\mathrm{d}t & , \textit{if } f\circ v \in L^1(a,b)\\ \infty & , \textit{else}\end{cases}$$

is proper, convex and lower semicontinuous.

Proof. Obviously, since f is proper, there exists a $v_0 \in X$ with $f(v_0) < \infty$ so that for $v \equiv v_0$ we have $F(v) < \infty$ and thus F is proper. For the convexity of F we argue as follows: Let $v, w \in L^1(a,b;X)$ and $\lambda \in [0,1]$ be given. W.l.o.g. we may assume that $v, w \in \mathrm{dom}(F)$. Clearly, $\lambda v + (1-\lambda)w$ is an element of $L^1(a,b;X)$ and thus Bochner–measurable. Since f is lower semicontinuous, it is Borel–measurable, so that $f\circ(\lambda v + (1-\lambda)w)$ is also measurable. By the convexity of f, we additionally have

$$0 \leq f(\lambda v(t) + (1-\lambda)w(t)) \leq \lambda f(v(t)) + (1-\lambda)f(w(t))$$

for almost all $t \in [a,b]$. Since $v, w \in \mathrm{dom}(F)$ it therefore holds

$$\begin{aligned}F(\lambda v + (1-\lambda)w) &= \int_a^b f(\lambda v(t) + (1-\lambda)w(t))\;\mathrm{d}t\\ &\leq \int_a^b \lambda f(v(x)) + (1-\lambda)f(w(x))\;\mathrm{d}t = \lambda F(v) + (1-\lambda)F(w),\end{aligned}$$

which proves the convexity of F. We finally turn to the lower semicontinuity property. For this, it suffices to show that the epigraph (see, e.g., [BV10]) is closed. Hence, let $(v_n, r_n) \in \mathrm{epi}(F) \subset X\times\mathbb{R}$ with $v_n \to v$ in $L^1(a,b;X)$ and $r_n \to r$. Then there exists a subsequence (denoted by the same symbol) such that $v_n(t) \to v(t)$ in X for almost all $t \in [a,b]$. The nonnegativity of f allows us to apply Fatou's lemma, which, together with the lower semicontinuity of f, gives

$$F(v) \;=\; \int_a^b f(v(t))\,\mathrm{d}t \;\leq\; \int_a^b \liminf_{n\to\infty} f(v_n(t))\,\mathrm{d}t \;\leq\; \liminf_{n\to\infty} \int_a^b f(v_n(t))\,\mathrm{d}t \;\leq\; \limsup_{n\to\infty} t_n \;=\; t.$$

The measurability of $f \circ v$ follows again by the measurability of v and f as above. Thus, we have $(v,t) \in \operatorname{epi}(F)$ which overall proves that F is indeed proper, convex and lower semicontinuous. □

With a view to calculus rules for the convex subdifferential, we state the following sum and chain rule which may be found in [ET99, Prop. I.5.6 and I.5.7].

Theorem A.3.6. *Let $f_1, f_2 : X \to \mathbb{R} \cup \{+\infty\}$ be convex. Moreover, let there exist $\hat{x} \in X$ with $x \in (\operatorname{dom}(f_1) \cap \operatorname{dom}(f_2))$ and either f_1 or f_2 is continuous in $\hat{x}$. Then it holds*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x) \quad \forall x \in X. \tag{A.3.4}$$

Theorem A.3.7. *Let Y be another normed vector space, $A \in \mathcal{L}(X,Y)$ and $f : Y \to \mathbb{R} \cup \{+\infty\}$ proper and convex. Moreover, let there be an $\hat{x} \in X$ so that f is finite and continuous in $A\hat{x}$. Then it holds*

$$\partial(f \circ A)(x) = A^* \partial f(Ax) \quad \forall x \in X,$$

where A^* denotes the adjoint operator of A.

After these more or less general results, we turn to more concrete statements concerning the dissipation $\mathcal{R}$ and $\mathcal{R}_h$, respectively. As in the proof of Lemma 3.2.2 we abbreviate $\mathcal{R}_{\tau,h} = \mathcal{R}_h \circ \Pi_h + I_\tau$, where I_τ is as defined in (3.2.2). Analogously, we set $\mathcal{R}_\tau^h := \mathcal{R}_h + I_\tau(v)$.

Lemma A.3.8. *For every $\eta \in \mathcal{Z}^*$, there holds*

$$(\mathcal{R}_{\tau,h})^*(\eta) = \tau\, \overline{\operatorname{dist}}_{\mathcal{V}^*}\{\eta, \partial(\mathcal{R}_h \circ \Pi_h)(0)\}, \tag{A.3.5}$$

where $\overline{\operatorname{dist}}_{\mathcal{V}^*}\{\eta, \partial(\mathcal{R}_h \circ \Pi_h)(0)\} = \min\{\|\eta - w\|_{\mathbb{V}^{-1}} \,:\, w \in \partial(\mathcal{R}_h \circ \Pi_h)(0)\}$ and $\|\eta\|^2_{\mathbb{V}^{-1}} = \langle \eta, \mathbb{V}^{-1}\eta\rangle$. Note that $\overline{\operatorname{dist}}_{\mathcal{V}^*}\{\eta, \partial(\mathcal{R}_h \circ \Pi_h)(0)\} = +\infty$ if there exists no $w \in \partial(\mathcal{R}_h \circ \Pi_h)(0)$ such that $\eta - w \in \mathcal{V}^*$.

Proof. We use the inf-convolution formula (see [Att84, Prop. 3.4]), which is applicable, since both functions are proper, convex and closed and we have $\operatorname{dom}(I_\tau) = B_{\mathbb{V}}(0,\tau)$. This gives

$$(\mathcal{R}_h \circ \Pi_h + I_\tau)^*(\eta) = \inf_{w \in \mathcal{Z}^*} \left((\mathcal{R}_h \circ \Pi_h)^*(w) + I_\tau^*(\eta - w) \right). \tag{A.3.6}$$

For I_τ^*, direct calculation leads to

$$I_\tau^*(\mu) = \begin{cases} \tau \|\mu\|_{\mathbb{V}^{-1}}, & \text{if } \mu \in \mathcal{V}^*, \\ +\infty, & \text{if } \mu \in \mathcal{Z}^* \setminus \mathcal{V}^*. \end{cases} \tag{A.3.7}$$

To calculate the conjugate functional of $(\mathcal{R}_h \circ \Pi_h)^*$, note that by the linearity of Π_h the composition $\mathcal{R}_h \circ \Pi_h$ is again convex and positively 1−homogeneous. Therefore, Lemma 2.1.1 implies that

$(\mathcal{R}_h \circ \Pi_h)^*(\eta) = I_{\partial(\mathcal{R}_h \circ \Pi_h)(0)}(\eta)$. Inserting this together with (A.3.7) in (A.3.6) finally yields

$$(\mathcal{R}_h \circ \Pi_h + I_\tau)^*(\eta) = \inf_{w \in \partial(\mathcal{R}_h \circ \Pi_h)(0)} \{\tau \|\eta - w\|_{\mathbb{V}^{-1}}\} = \tau\, \overline{\operatorname{dist}}_{\mathcal{V}^*}\{\eta, \partial(\mathcal{R}_h \circ \Pi_h)(0)\},$$

which is (A.3.5). The fact that the infimum, if it is finite, is attained, follows from the weak-closedness of $\partial(\mathcal{R}_h \circ \Pi_h)(0)$ and the weak lowersemicontinuity of $\|\cdot\|_{\mathbb{V}^{-1}}$. □

From the previous lemma it is easy to conclude the following corollary, where we abbreviate $\mathcal{R}_\tau = \mathcal{R} + I_\tau$ and again I_τ denotes the indicator function as defined in (3.2.2).

Corollary A.3.9. *For every $\eta \in \mathcal{Z}^*$, there holds*

$$(\mathcal{R}_\tau)^*(\eta) = \tau\, \overline{\operatorname{dist}}_{\mathcal{V}^*}\{\eta, \partial\mathcal{R}(0)\}, \tag{A.3.8}$$

where $\overline{\operatorname{dist}}_{\mathcal{V}^*}\{\eta, \partial\mathcal{R}(0)\} = \min\{\|\eta - w\|_{\mathbb{V}^{-1}} : w \in \partial\mathcal{R}(0)\}$ and $\|\eta\|^2_{\mathbb{V}^{-1}} = \langle \eta, \mathbb{V}^{-1}\eta\rangle$. Note, again, that $\overline{\operatorname{dist}}_{\mathcal{V}^*}\{\eta, \partial\mathcal{R}(0)\} = +\infty$ if there exists no $w \in \partial\mathcal{R}(0)$ such that $\eta - w \in \mathcal{V}^*$.

Lemma A.3.10. *Let $v \in \mathcal{Z}$ be arbitrary. Then $\partial^{\mathcal{Z}} I_\tau(v) = \partial^{\mathcal{V}} I_\tau(v)$.*

Proof. If $v \notin B_{\mathbb{V}}(0,\tau)$, then $\partial^{\mathcal{V}} I_\tau(v) = \emptyset = \partial^{\mathcal{Z}} I_\tau(v)$. Thus, let $v \in B_{\mathbb{V}}(0,\tau)$ and $\xi \in \partial^{\mathcal{Z}} I_\tau(v) \subset \mathcal{Z}^*$, which, by definition, means

$$\begin{aligned} I_\tau(w) \geq I_\tau(v) + \langle \xi, w - v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \quad \forall w \in \mathcal{Z} \\ \iff 0 \geq \langle \xi, w - v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \quad \forall w \in B_{\mathbb{V}}(0,\tau) \cap \mathcal{Z}. \end{aligned} \tag{A.3.9}$$

Since $v \in \mathcal{Z}$, we can estimate $\langle \xi, w\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq \langle \xi, v\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq \|\xi\|_{\mathcal{Z}^*}\|v\|_{\mathcal{Z}}$ for all $w \in B_{\mathbb{V}}(0,\tau) \cap \mathcal{Z}$. Testing with $w = \tau\tilde{w}/\|\tilde{w}\|_{\mathbb{V}}$, we find $\langle \xi, \tilde{w}\rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq \tau\|\xi\|_{\mathcal{Z}^*}\|v\|_{\mathcal{Z}}\|\tilde{w}\|_{\mathbb{V}} =: C\|\tilde{w}\|_{\mathbb{V}}$ for all $\tilde{w} \in \mathcal{Z}$. Hence, the density of $\mathcal{Z}$ in $\mathcal{V}$ allows us to extend $\xi \in \mathcal{Z}^*$ in a unique way to a subgradient in $\mathcal{V}^*$, i.e., $\xi \in \partial^{\mathcal{V}} I_\tau(v)$. In this sense, it holds $\partial^{\mathcal{Z}} I_\tau(v) \subset \partial^{\mathcal{V}} I_\tau(v)$. Note that the opposite inclusion is a direct consequence of $\mathcal{Z} \subset \mathcal{V}$. □

The lemma above allows us to simply use $\partial I_\tau(v)$ and neglect the underlying space $\mathcal{V}$ and $\mathcal{Z}$, respectively. Now, we come to the actual characterization of ∂I_τ.

Lemma A.3.11. *Let $v \in \mathcal{Z}$ be arbitrary. Then $\xi \in \partial I_\tau(v)$ if and only if there exists a multiplier $\lambda \in \mathbb{R}$ such that $\xi = \lambda \mathbb{V} v$ and*

$$\|v\|_{\mathbb{V}} \leq \tau, \quad \lambda(\|v\|_{\mathbb{V}} - \tau) = 0, \quad \lambda \geq 0.$$

Proof. According to a classical result of convex analysis in combination with (A.3.7), it holds

$$\xi \in \partial I_\tau(v) \quad \iff \quad I_\tau(v) + I_\tau^*(\xi) = \langle \xi, v\rangle \quad \iff \quad \begin{cases} \|v\|_{\mathbb{V}} \leq \tau, \\ \tau\|\xi\|_{\mathbb{V}^{-1}} = \langle \xi, v\rangle. \end{cases} \tag{A.3.10}$$

Now, the Cauchy-Schwarz-Inequality implies $\langle \xi, v\rangle = \langle \mathbb{V}(\mathbb{V}^{-1}\xi), v\rangle \leq \|\xi\|_{\mathbb{V}^{-1}}\|v\|_{\mathbb{V}} \leq \tau\|\xi\|_{\mathbb{V}^{-1}}$ so that the equivalence in (A.3.10) can only hold if $\mathbb{V}^{-1}\xi = \lambda v$ for some $\lambda \in \mathbb{R}$. Inserting this into

(A.3.10), we conclude that $\lambda \geq 0$. Moreover, if $\|v\|_{\mathbb{V}} < \tau$, then $\xi = 0$ so that λ fulfills also $\lambda(\|v\|_{\mathbb{V}} - \tau) = 0$ as claimed. □

A.4 Auxiliary results

Some Gronwall inequalities

Lemma A.4.1 (Standard Gronwall inequality)**.** *Let $m \in L^1(a,b)$ with $m \geq 0$ almost everywhere in $[a,b]$ and $\alpha \geq 0$ a positive constant. If $u : [a,b] \to \mathbb{R}$ is a continuous function satisfying*

$$u(t) \leq \alpha + \int_a^b m(s)u(s)\,\mathrm{d}s \quad \forall t \in [a,b],$$

then $u(t) \leq \alpha \exp\left(\int_a^t m(s)\,\mathrm{d}s\right)$ for all $t \in [a,b]$.

Proof. See, e.g., [Bré73, Lem. A.4]. □

Lemma A.4.2. *Let $m \in L^1(a,b)$ with $m \geq 0$ almost everywhere in $[a,b]$, $\alpha \geq 0$ a positive constant and let $u : [a,b] \to \mathbb{R}$ be a continuous function satisfying*

$$u(t)^2 \leq \alpha^2 + \int_a^t m(s)u(s)\,\mathrm{d}s \quad \forall t \in [a,b].$$

Then $|u(t)| \leq \alpha + \int_a^t m(s)\,\mathrm{d}s$.

Proof. See [Bré73, Lem. A.5]. □

Normalization of parametrized solutions

Lemma A.4.3. *Any $\mathbb{V}$-parametrized solutions $(\hat{t}, \hat{z})$ can be reparameterized such that its reparameterization $(\tilde{t}, \tilde{z})$ is normalized, i.e., it fulfills $\tilde{t}'(s) + \|\tilde{z}'\|_{\mathbb{V}} = 1$ for almost all $s \in [0, \tilde{S}]$, and $(\tilde{t}, \tilde{z})$ is still a $\mathbb{V}$-parametrized solution.*

Proof. The proof is based on [Mie11, Lem. 4.12] and [KRZ13, Rem. 6.5]. Let $(\hat{t}, \hat{z})$ by a non-normalized parametrized solution with artificial end time S. We define

$$m(s) := \int_0^s \hat{t}'(r) + \|\hat{z}'(r)\|_{\mathbb{V}}\,\mathrm{d}r, \tag{A.4.1}$$

so that $m'(s) = \hat{t}'(r) + \|\hat{z}'(r)\|_{\mathbb{V}}$ for almost all $s \in [0,S]$ and $m \in W^{1,\infty}(0,S)$. We moreover let $r(\mu) := \inf\{\sigma \geq 0 : m(\sigma) = \mu\}$ as well as $\tilde{t}(\mu) = \hat{t}(r(\mu))$ and $\tilde{z}(\mu) = \hat{z}(r(\mu))$. Note that r is monotone increasing, so that Lebesgue's differentiation theorem (see, e.g., [Leo17, Thm. 1.18]) implies that r is differentiable almost everywhere with $r' \geq 0$. By construction, we additionally have $m(r(\mu)) = \mu$ for all $\mu \in [0,R]$, where $R = r(m(S))$. To proceed, we observe that $\tilde{t}$ and $\tilde{z}$ are Lipschitz continuous. Indeed, it holds for all $0 \leq \mu_1 < \mu_2 \leq S$ that

$$
\begin{aligned}
|\tilde{t}(\mu_2) - \tilde{t}(\mu_1)| + \|\tilde{z}(\mu_2) - \tilde{z}(\mu_1)\|_{\mathbb{V}} \\
&= |\hat{t}(r(\mu_2)) - \hat{t}(r(\mu_1))| + \|\hat{z}(r(\mu_2)) - \hat{z}(r(\mu_1))\|_{\mathbb{V}} \\
&= \left| \int_{r(\mu_1)}^{r(\mu_2)} \hat{t}'(s) \, \mathrm{d}s \right| + \left\| \int_{r(\mu_1)}^{r(\mu_2)} \hat{z}'(s) \, \mathrm{d}s \right\|_{\mathbb{V}} \\
&\leq \int_{r(\mu_1)}^{r(\mu_2)} |\hat{t}'(s)| + \|\hat{z}'(s)\|_{\mathbb{V}} \, \mathrm{d}s = m(r(\mu_2)) - m(r(\mu_1)) = \mu_2 - \mu_1.
\end{aligned}
$$

Briefly summarized, we have m, $\tilde{t}$ and $\tilde{z}$ differentiable almost everywhere since they are Lipschitz-continuous and so is r due to Lebesgue's differentiation theorem. The chain rule from Theorem A.2.3 thus gives

$$
\tilde{t}'(\mu) = \hat{t}'(r(\mu))r'(\mu) \quad \text{and} \quad \tilde{z}'(\mu) = \hat{z}'(r(\mu))r'(\mu)
$$

for almost all $\mu \in [0, R]$. Combining all this, we end up with

$$
\begin{aligned}
\mu = m(r(\mu)) &= \int_0^\mu (m \circ r)'(s) \, \mathrm{d}s \\
&= \int_0^\mu m'(r(s))r'(s) \, \mathrm{d}s \\
&= \int_0^\mu (\hat{t}'(r(s)) + \|\hat{z}'(r(s))\|_{\mathbb{V}} \, r'(s) \, \mathrm{d}s \\
&= \int_0^\mu (\hat{t}'(r(s)) \, r'(s) + \|\hat{z}'(r(s))r'(s)\|_{\mathbb{V}} \, \mathrm{d}s \\
&= \int_0^\mu (\hat{t} \circ r)'(s) + \|(\hat{z} \circ r)'(s))\|_{\mathbb{V}} \, \mathrm{d}s = \int_0^\mu \tilde{t}'(s) + \|\tilde{z}'(s)\| \, \mathrm{d}s.
\end{aligned}
$$

for all $\mu \in [0, R]$, which eventually proves that $(\tilde{t}, \tilde{z})$ is normalized. The fact that $(\tilde{t}, \tilde{z})$ is still a parametrized solution is now easy to check by exploiting the change of variable formula in Theorem A.2.9. □

Estimate on $\gamma(t)$

Proof. Let $z_1, z_2 \in W^{1,\infty}(0, T; \mathcal{Z})$ and, again,

$$
\gamma(t) := \langle D_z\mathcal{I}(t, z_1(t)) - D_z\mathcal{I}(t, z_2(t)), z_1(t) - z_2(t) \rangle_{\mathcal{Z}^*, \mathcal{Z}}.
$$

First of all, using the symmetry of $D_z^2\mathcal{I}$, we calculate

$$
\begin{aligned}
\gamma'(t) = \langle \partial_t D_z\mathcal{I}(t, z_1(t)) - \partial_t D_z\mathcal{I}(t, z_2(t)), z_1(t) - z_2(t) \rangle_{\mathcal{Z}^*, \mathcal{Z}} \\
+ \langle D_z^2\mathcal{I}(t, z_1(t))[z_1(t) - z_2(t)], z_1'(t) \rangle_{\mathcal{Z}^*, \mathcal{Z}} \\
- \langle D_z^2\mathcal{I}(t, z_2(t))[z_1(t) - z_2(t)], z_2'(t) \rangle_{\mathcal{Z}^*, \mathcal{Z}} \\
+ \langle D_z\mathcal{I}(t, z_1(t)) - D_z\mathcal{I}(t, z_2(t)), z_1'(t) - z_2'(t) \rangle_{\mathcal{Z}^*, \mathcal{Z}}.
\end{aligned}
$$

Rearranging terms, we arrive at

$$\begin{aligned}\gamma'(t) &= \langle \partial_t D_z\mathcal{I}(t,z_1(t)) - \partial_t D_z\mathcal{I}(t,z_2(t)), z_1(t)-z_2(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\\ &\quad + \langle D_z^2\mathcal{I}(t,z_1(t))[z_1(t)-z_2(t)] + D_z\mathcal{I}(t,z_2(t)) - D_z\mathcal{I}(t,z_1(t)), z_1'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\\ &\quad - \langle D_z^2\mathcal{I}(t,z_2(t))[z_1(t)-z_2(t)] + D_z\mathcal{I}(t,z_1(t)) - D_z\mathcal{I}(t,z_2(t)), z_2'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\\ &\quad + 2\langle D_z\mathcal{I}(t,z_1(t)) - D_z\mathcal{I}(t,z_2(t)), z_1'(t)-z_2'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}.\end{aligned}$$

Now, due to $z_1, z_2 \in W^{1,\infty}(0,T;\mathcal{Z})$ and the regularity on $\mathcal{I}(t,\cdot)$ (see (1.0.5)), we find that

$$\begin{aligned}\gamma'(t) &\leq c\|z_1(t)-z_2(t)\|_{\mathcal{Z}}^2\\ &\quad + C\|z_1(t)-z_2(t)\|_{\mathcal{Z}}^2\|z_1'(t)\|_{\mathcal{Z}} + C\|z_1(t)-z_2(t)\|_{\mathcal{Z}}^2\|z_2'(t)\|_{\mathcal{Z}}\\ &\quad + 2\langle D_z\mathcal{I}(t,z_1(t)) - D_z\mathcal{I}(t,z_2(t)), z_1'(t)-z_2'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}}\\ &\leq C\|z_1(t)-z_2(t)\|_{\mathcal{Z}}^2 + 2\langle D_z\mathcal{I}(t,z_1(t)) - D_z\mathcal{I}(t,z_2(t)), z_1'(t)-z_2'(t)\rangle_{\mathcal{Z}^*,\mathcal{Z}},\end{aligned}$$

which is the desired estimate. □

Index

P

R

S

U

V

Bibliography

[AC00] Alberty, J. ; Carstensen, C.: Numerical Analysis of Time-Depending Primal Elastoplasticity with Hardening. In: *SIAM Journal on Numerical Analysis* 37 (2000), Nr. 4, S. 1271–1294

[ACFS17] Artina, M. ; Cagnetti, F. ; Fornasier, M. ; Solombrino, F.: Linearly constrained evolutions of critical points and an application to cohesive fractures. In: *Math. Models Methods Appl. Sci.* 27 (2017), Nr. 2, S. 231–290. http://dx.doi.org/10.1142/S0218202517500014. – DOI 10.1142/S0218202517500014

[AGS08] Ambrosio, L. ; Gigli, N. ; Savaré, G.: *Gradient flows: in metric spaces and in the space of probability measures.* Springer Science & Business Media, 2008

[Alm20] Almi, S.: Irreversibility and alternate minimization in phase field fracture: a viscosity approach. In: *Zeitschrift für angewandte Mathematik und Physik* 71 (2020), Nr. 4, S. 1–21

[Alt12] Alt, H.W.: *Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung.* Springer Berlin Heidelberg, 2012 (Masterclass)

[AMS08] Auricchio, F. ; Mielke, A. ; Stefanelli, U.: A rate-independent model for the isothermal quasi-static evolution of shape-memory materials. In: *Mathematical Models and Methods in Applied Sciences* 18 (2008), Nr. 01, S. 125–164

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

[Bar14] Bartels, S.: Quasi-optimal Error Estimates for Implicit Discretizations of Rate-Independent Evolutions. In: *SIAM Journal on Numerical Analysis* 52 (2014), Nr. 2, S. 708–716

[BKS04] Brokate, M. ; Krejčí, P. ; Schnabel, H.: On uniqueness in evolution quasivariational inequalities. 11 (2004), Nr. 1, S. 111–130

[Bon96] Bonfanti, G.: A vanishing viscosity approach to a two degree-of-freedom contact problem in linear elasticity with friction. In: *Annali dell'Universita di Ferrara* 42 (1996), Nr. 1, S. 127–154

[Bré73] Brézis, H.: *North-Holland Mathematics Studies*. Bd. 5: *Opeérateurs maximaux monotones*. North-Holland, 1973

[BV10] Borwein, J.M. ; Vanderwerff, J.D.: *Convex functions: constructions, characterizations and counterexamples*. Bd. 109. Cambridge University Press Cambridge, 2010

[Col92] Colli, P.: On some doubly nonlinear evolution equations in Banach spaces. In: *Japan Journal of Industrial and Applied Mathematics* 9 (1992), Nr. 2, S. 181

[Cou21] Coulomb, C.A.: Theorie des machines simple (Theory of simple machines). In: *Bachelier, Paris* (1821)

[CV90] Colli, P. ; Visintin, A.: On a class of doubly nonlinear evolution equations. In: *Communications in partial differential equations* 15 (1990), Nr. 5, S. 737–756

[DMFT05] Dal Maso, G. ; Francfort, G. ; Toader, R.: Quasistatic crack growth in nonlinear elasticity. In: *Archive for Rational Mechanics and Analysis* 176 (2005), Nr. 2, S. 165–225

[DU77] Diestel, J. ; Uhl, J.J.: *Vector Measures*. American Mathematical Society, 1977 (Mathematical surveys and monographs)

[EG92] Evans, L.C. ; Gariepy, R.F.: *Measure Theory and Fine Properties of Functions*. CRC Press, 1992

[EM06] Efendiev, M. A. ; Mielke, A.: On the rate-independent limit of systems with dry friction and small viscosity. In: *Journal of Convex Analysis* 13 (2006), Nr. 1, S. 151–167

[ET99] Ekeland, I. ; Temam, R.: *Convex analysis and variational problems*. SIAM, 1999

[FM06] Francfort, G. ; Mielke, A.: Existence results for a class of rate-independent material models with nonconvex elastic energies. In: *Journal für die reine und angewandte Mathematik (Crelles Journal)* 2006 (2006), Nr. 595, S. 55–91

[GGZ74] Gajewski, H. ; Gröger, K. ; Zacharias, K.: Nichtlineare Operatorgleichungen. In: *Operatordifferentialgleichungen, Berlin* (1974)

[GHS16] Gaspoz, F.D. ; Heine, C.-J. ; Siebert, K.G.: Optimal Grading of the Newest Vertex Bisection and H^1-Stability of the L^2-Projection. In: *IMA Journal of Numerical Analysis* 36 (2016), Nr. 3, S. 1217–1241

[HK09] Haar, A. ; Kármán, Th. v.: Zur Theorie der Spannungszustände in plastischen und sandartigen Medien. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1909 (1909), S. 204–218

[HN75] Halphen, B. ; Nguyen, Q.S.: Generalized standard materials. In: *Journal de mécanique* 14 (1975), Nr. 1, S. 39–63

[HPUU08] Hinze, M. ; Pinnau, R. ; Ulbrich, M. ; Ulbrich, S.: *Optimization with PDE constraints*. Bd. 23. Springer Science & Business Media, 2008

[HR99] Han, W. ; Reddy, B.D.: *Plasticity*. New York : Springer, 1999

[IK08] Ito, K. ; Kunisch, K.: *Lagrange multiplier approach to variational problems and applications*. SIAM, 2008

[KN17] Knees, D. ; Negri, M.: Convergence of alternate minimization schemes for phase-field fracture and damage. In: *Mathematical Models and Methods in Applied Sciences* 27 (2017), Nr. 09, S. 1743–1794

[Kne19] Knees, D.: Convergence analysis of time-discretisation schemes for rate-independent systems. In: *ESAIM: Control, Optimisation and Calculus of Variations* 25 (2019), S. 65

[Kre99] Krejčí, P.: Evolution variational inequalities and multidimensional hysteresis operators. In: *Nonlinear differential equations (Chvalatice, 1998)* Bd. 404. Chapman & Hall/CRC, Boca Raton, FL, 1999, S. 47–110

[KRZ13] Knees, D. ; Rossi, R. ; Zanini, C.: A vanishing viscosity approach to a rate-independent damage model. In: *Mathematical Models and Methods in Applied Sciences* 23 (2013), Nr. 04, S. 565–616

[KRZ19] Knees, D. ; Rossi, R. ; Zanini, C.: Balanced viscosity solutions to a rate-independent system for damage. In: *European Journal of Applied Mathematics* 30 (2019), Nr. 1, S. 117–175

[KS13] Knees, D. ; Schröder, A.: Computational aspects of quasi-static crack propagation. In: *Discrete and Continuous Dynamical Systems. Series S* 6 (2013), Nr. 1, S. 63–99. http://dx.doi.org/10.3934/dcdss.2013.6.63. – DOI 10.3934/dcdss.2013.6.63

[KT18] Knees, D. ; Thomas, S.: Optimal control of a rate-independent system constrained to parametrized balanced viscosity solutions. In: *arXiv preprint arXiv:1810.12572* (2018)

[KZ18] Knees, D. ; Zanini, C.: *Existence of parameterized BV-solutions for rate-independent systems with discontinuous loads*. 2018

[Leo17] Leoni, G.: *A First Course in Sobolev Spaces*. Second. American Mathematical Society, 2017 (Graduate studies in mathematics)

[Lev97] Levitas, V.I.: Phase transitions in elastoplastic materials: continuum thermomechanical theory and examples of control—part I. In: *Journal of the Mechanics and Physics of Solids* 45 (1997), Nr. 6, S. 923–947

[Mai04] MAINIK, A.: *A rate-independent model for phase transformations in shape-memory alloys*, Universität Stuttgart, Diss., 2004. http://dx.doi.org/10.18419/opus-4749. – DOI 10.18419/opus–4749

[Mie03] MIELKE, A.: Energetic formulation of multiplicative elasto-plasticity using dissipation distances. In: *Continuum Mechanics and Thermodynamics* 15 (2003), Nr. 4, S. 351–382

[Mie06] MIELKE, A.: Chapter 6: Evolution Of Rate-Independent Systems. In: *Handbook of Differential Equations, Evolutionary Equations* 2 (2006), 12

[Mie11] MIELKE, A.: Differential, energetic, and metric formulations for rate-independent processes. In: *Nonlinear PDE's and Applications* 2028 (2011), S. 87–167

[Mis13] MISES, R. v.: Mechanik der festen Körper im plastisch-deformablen Zustand. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1913 (1913), S. 582–592

[MM09] MAINIK, A. ; MIELKE, A.: Global existence for rate-independent gradient plasticity at finite strain. In: *Journal of Nonlinear Science* 19 (2009), Nr. 3, S. 221–248

[MMMG94] MARTINS, JA C. ; MONTEIRO MARQUES, MD P. ; GASTALDI, Fabio: On an example of non-existence of solution to a quasistatic frictional contact problem. In: *European Journal of Mechanics - A./Solids* 13 (1994), Nr. 1, S. 113–133

[Mor70] MOREAU, J.J.: Sur les lois de frottement, de plasticité et de viscosité. In: *C.R. Acad. Sci., Paris* 271 (1970), S. 608–611

[Mor71] MOREAU, J.J.: Fonctions de résistance et fonctions de dissipation, Séminaire d'analyse convexe. In: *Montpellier, exposé* (1971), Nr. 6

[Mor77] MOREAU, J.J.: Evolution problem associated with a moving convex set in a Hilbert space. In: *Journal of Differential Equations* 26 (1977), Nr. 3, S. 347–374

[MPPS10] MIELKE, A. ; PAOLI, L. ; PETROV, A. ; STEFANELLI, U.: Error estimates for space-time discretizations of a rate-independent variational inequality. In: *SIAM Journal on Numerical Analysis* 48 (2010), Nr. 5, S. 1625–1646

[MR09] MIELKE, A. ; ROUBÍČEK, T.: Numerical approaches to rate-independent processes and applications in inelasticity. In: *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique* 43 (2009), Nr. 3, S. 399–428

[MR15] MIELKE, A. ; ROUBÍČEK, T.: *Rate-Independent Systems: Theory and Application.* New York : Springer, 2015

[MRS09] MIELKE, A. ; ROSSI, R. ; SAVARÉ, G.: Modeling solutions with jumps for rate-independent systems on metric spaces. In: *Discrete & Continuous Dynamical Systems-A* 25 (2009), Nr. 2, S. 585–615

[MRS12] Mielke, A. ; Rossi, R. ; Savaré, G.: BV solutions and viscosity approximations of rate-independent systems. In: *ESAIM: Control, Optimisation and Calculus of Variations* 18 (2012), Nr. 1, S. 36–80

[MRS13] Mielke, A. ; Rossi, R. ; Savaré, G.: Nonsmooth analysis of doubly nonlinear evolution equations. In: *Calculus of Variations and Partial Differential Equations* 46 (2013), Nr. 1-2, S. 253–310

[MRS16] Mielke, A. ; Rossi, R. ; Savaré, G.: Balanced viscosity (BV) solutions to infinite-dimensional rate-independent systems. In: *J. Eur. Math. Soc. (JEMS)* 18 (2016), Nr. 9, S. 2107–2165

[MS17] Matteo, N. ; Scala, R.: A quasi-static evolution generated by local energy minimizers for an elastic material with a cohesive interface. In: *Nonlinear Analysis: Real World Applications* 38 (2017), S. 271–305

[MS19a] Meyer, C. ; Sievers, M.: Finite Element Discretization of Local Minimization Schemes for Rate-Independent Evolutions. In: *Calcolo* 56 (2019), Nr. 6. http://dx.doi.org/10.1007/s10092-018-0301-4. – DOI 10.1007/s10092–018–0301–4

[MS19b] Meyer, C. ; Susu, L.M.: Analysis of a Viscous Two-Field Gradient Damage Model I: Existence and Uniqueness. In: *Zeitschrift fuer Analysis und Ihre Anwendungen* 38 (2019), Nr. 3, S. 249–287

[MS19c] Meyer, C. ; Susu, L.M.: Analysis of a Viscous Two-Field Gradient Damage Model II: Penalization Limit. In: *Zeitschrift für Analysis und ihre Anwendungen* 38 (2019), Nr. 4, S. 439–474

[MS20] Meyer, C. ; Sievers, M.: A Priori Error Analysis of Local Incremental Minimization Schemes for Rate-Independent Evolutions. In: *SIAM Journal on Numerical Analysis* 58 (2020), Nr. 4, S. 2376–2403

[MSGMM95] Martins, J.A.C. ; Simões, F.M.F. ; Gastaldi, F. ; Monteiro Marques, M.D.P.: Dissipative graph solutions for a 2 degree-of-freedom quasistatic frictional contact problem. In: *International journal of engineering science* 33 (1995), Nr. 13, S. 1959–1986

[MT99] Mielke, A. ; Theil, F.: A mathematical model for rate-independent phase transformations with hysteresis. In: *Proceedings of the Workshop on "Models of Continuum Mechanics in Analysis and Engineering*, 1999, S. 117–129

[MT04] Mielke, A. ; Theil, F.: On rate-independent hysteresis models. In: *NoDEA: Nonlinear Differential Equations and Applications* 11 (2004), Nr. 2, S. 151–189

[MTL98] Mielke, A. ; Theil, F. ; Levitas, V.I.: Mathematical formulation of quasistatic phase transformations with friction using an extremum principle. In: *Preprint A8, Hannover* (1998)

[MTL02] MIELKE, A. ; THEIL, F. ; LEVITAS, V.I.: A Variational Formulation of Rate-Independent Phase Transformations Using an Extremum Principle. In: *Archive for Rational Mechanics and Analysis* 162 (2002), Nr. 2, S. 137–177

[MZ14] MIELKE, A. ; ZELIK, S.: On the vanishing-viscosity limit in parabolic systems with rate-independent dissipation terms. In: *Ann. Sc. Norm. Super. Pisa, Cl. Sci.* 13 (2014), Nr. 1, S. 67–135

[Neg14] NEGRI, M.: Quasi-static rate-independent evolutions: characterization, existence, approximation and application to fracture mechanics. In: *ESAIM Control Optim. Calc. Var.* 20 (2014), Nr. 4, S. 983–1008. http://dx.doi.org/10.1051/cocv/2014004. – DOI 10.1051/cocv/2014004

[OW68] OWEN, D.R. ; WILLIAMS, W.O.: On the concept of rate-independence. In: *Quarterly of Applied Mathematics* 26 (1968), Nr. 3, S. 321–329

[PR65] PIPKIN, A.C. ; RIVLIN, R.S.: Mechanics of rate-independent materials. In: *Zeitschrift für angewandte Mathematik und Physik ZAMP* 16 (1965), May, Nr. 3, S. 313–327

[Rou13] ROUBÍČEK, T.: *Nonlinear partial differential equations with applications*. Bd. 153. Springer Science & Business Media, 2013

[Rou15] ROUBÍČEK, T.: Maximally-dissipative local solutions to rate-independent systems and application to damage and delamination problems. In: *Nonlinear Analysis: Theory, Methods & Applications* 113 (2015), S. 33 – 50. http://dx.doi.org/https://doi.org/10.1016/j.na.2014.09.020. – DOI https://doi.org/10.1016/j.na.2014.09.020

[RS06] ROSSI, R. ; SAVARÉ, G.: Gradient flows of non convex functionals in Hilbert spaces and applications. In: *ESAIM: Control, Optimisation and Calculus of Variations* 12 (2006), Nr. 3, S. 564–614

[RSS17] RINDLER, F. ; SCHWARZACHER, S. ; SÜLI, E.: Regularity and approximation of strong solutions to rate-independent systems. In: *Mathematical Models and Methods in Applied Sciences* 27 (2017), Nr. 13, S. 2511–2556. http://dx.doi.org/10.1142/S0218202517500518. – DOI 10.1142/S0218202517500518

[San17] SANTAMBROGIO, Filippo: {Euclidean, metric, and Wasserstein} gradient flows: an overview. In: *Bulletin of Mathematical Sciences* 7 (2017), Nr. 1, S. 87–154

[Ste09] STEFANELLI, U.: A variational characterization of rate-independent evolution. In: *Mathematische Nachrichten* 282 (2009), Nr. 11, S. 1492–1512

[SV70] SAINT-VENANT, B. de: Sur l'établissement des équations des mouvements intérieurs opérés dans les corps solides ductiles au delà des limites où l'élasticité pourrait les ramener á leur premier état. In: *Comptes Rendus de l'Ac. des Sciences* 70 (1870), S. 473–480

[SW11] Schiela, A. ; Wollner, W.: Barrier methods for optimal control problems with convex nonlinear gradient state constraints. In: *SIAM Journal on Optimization* 21 (2011), Nr. 1, S. 269–286

[TN65] Truesdell, C. ; Noll, W.: *The Nonlinear Field theories of Mechanics, in in S. Flügge, editor*. Bd. 3. 1965. – 3 S.

[Tre64] Tresca, H.E.: *Sur l'ecoulement des corps solides soumis a de fortes pressions.* Imprimerie de Gauthier-Villars, successeur de Mallet-Bachelier, rue de Seine-Saint-Germain, 1864

[Wac11] Wachsmuth, G.: *Optimal control of quasistatic plasticity – An MPCC in function space*, Chemnitz University of Technology, Diss., 2011