Technische Universität München
Fakultät für Mathematik

Rate independent evolution processes on functions of bounded variation

Thomas Roche

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Herbert Egger
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Martin Brokate
	2.	Prof. Dr. Pavel Krejčí, Academy of Science, Prag, Tschechien
	3.	Univ.-Prof. Dr. Alexander Mielke, Humboldt-Universität Berlin (schriftliche Beurteilung)

Die Dissertation wurde am 20. März 2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 20. Mai 2012 angenommen.

Acknowledgements

There are plenty of people towards whom I want to express my gratitude. Thus I will take the opportunity here to do so.

First and foremost I would like to thank my advisor Prof. Martin Brokate for all his help and encouragement. I joint his group shortly before my Bachelor's degree. I am grateful for the guidance he gave me towards doing research and the possibility to follow my own interests.

My sincere gratitude goes to Prof. Pavel Krejčí. I had the pleasure to stay at his institute for three months. I am thankful for the time he took for discussions and for the interesting problems he pointed out. This time certainly developed and shaped me as a mathematician.

I moreover would like to thank Prof. Alexander Mielke for serving as a member of my examination board. Twice I had the possibility to visit with his group. I am very thankful for the warm and welcoming atmosphere at WIAS Berlin.

I gratefully acknowledge financial support through the Max Weber-Programm, the Nečas Centre for Mathematical Modeling, a doctoral scholarship by Universität Bayern e.V. and TopMath.

My warmest thanks to all people I had the chance to collaborate with during my PhD studies. Especially Pavel Krejčí, Matthias Liero, Riccarda Rossi and Ulisse Stefanelli with all of whom I had the pleasure to work on joint projects. I have learned a lot from you!

I would also like to thank all current and former memebers of the research unit M6 for the great atmosphere. Especially thanks to Stefan Neukamm and Elena Gornostaeva with whom I had the pleasure to share office. Thanks to Kathrin Ruf for fun times while organizing TUMMS. Moreover thanks, Elena and Kathrin, for proofreading parts of this thesis. Last but certainly not least thanks to Carl-Friedrich Kreiner for regular and most enjoyable lunch sessions.

I would like to thank all my fellow students in the graduate program TopMath for mathematical insights and a lot of fun. A special thanks to Max Butz for our regular coffee sessions and discussions on mathematics and everything else. Moreover I would like to thank the program managers Ralf, Andrea and Denise for for their always open ear and their continuous support.

As has been said many times, in order to write a mathematical thesis not only mathematical input is needed. Thanks to all my friends for the enjoyable time I had in Munich. Especially thanks to Andreas Bader and Matthias Vestner not only for the fun we had but also for their support in times of need.

Thanks to Mirjam for everything!

Last but not certainly not least, I would like express my most sincere gratitude to my parents for their loving and unconditional support!

All other support though not mentioned is certainly not forgotten.

Contents

1. Introduction

1.1. Overview

The objective of this thesis is to contribute to the mathematical theory of hysteresis. Loosely speaking hysteresis describes phenomena where the evolution of a system depends not only on its current state but also on its past evolution. We are going to study both hysteresis operators, which are the mathematical building brick of describing systems with hysteresis, as well as a class of evolution system, in which hysteresis occurs as a part of the problem.

The mathematical description of hysteresis phenomena at least goes back to the beginning of the 20th century. Among the first to work in this field were Prandtl [84], who described plastic deformations using hysteresis, and Preisach [85], who studied magnetic effects. The systematic mathematical study of hysteresis however was initiated only around 1970 by the group around the Russian mathematician M.A. Krasnoselskii, see e.g. [51, 50]. Their efforts resulted in their seminal monograph [52]. A key element of their studies which was to understand hysteresis as an operator acting on functions defined over a time interval $[0, T]$. We are going to use a refined definition due to A. Visintin, see e.g. [105]. Let $M(0, T; X)$ be the set of all maps mapping the time interval $[0, T]$ to a state space X. An operator $\mathcal{W} : D(\mathcal{W}) \subset M(0, T; X) \to M(0, T; Y)$ is called an **hysteresis operator** if

- $\mathcal{W}$ is *rate-independent*, i.e. for all $f \in D(\mathcal{W})$ and all monotone increasing functions $\phi : [0, T] \to [0, T]$ with $\phi([0, T]) = [0, T]$ it holds

$$\mathcal{W}(f \circ \phi) = \mathcal{W}(f) \circ \phi$$

- and $\mathcal{W}$ has the *Volterra property*, i.e. for all $g, f \in D(\mathcal{W})$ and $t \in [0, T]$ it holds

$$g|_{[0,t]} = f|_{[0,t]} \quad \Longrightarrow \quad \mathcal{W}(g)(t) = \mathcal{W}(f)(t)\,.$$

The latter property makes $\mathcal{W}$ a memory operator, as the current state is only determined by past but not by the future. Rate-independence assures that the 'speed' or rate of the input does affect the output only insofar as it changes its rate in the same way. Hysteresis operators are the basic building blocks of systems including hysteresis, of which one can find in nature many. Apart from plasticity and magnetism, hysteresis occurs also in phase transitions, the modeling of shape memory alloys or can be artificially added into a system in order to control it, as for example using thermostats. These phenomena are described by ordinary or partial differential equations in which the hysteresis operator is one element. Hysteresis operators and systems with hysteresis have been the subject of quite some research. Following the book by Krasnoselskii and Pokrovskii quite a number of monographs have been written. We would like to point out those by Mayergoyz [68], P. Krejčí [54], A. Visintin [105], and M. Brokate and J. Sprekels [19]. They cover the main analytical aspects of dealing with hysteresis, namely the description and modeling of hysteresis phenomena, the study of hysteresis operators and the analysis of evolution systems with hysteresis.

With the present thesis we would like to contribute to two of the three above topics, namely we are going to study a special type of hysteresis operator, the **quasivariational sweeping process**, and analyze evolution systems with hysteresis, the so called **doubly nonlinear differential inclusions**. The methods we use are fairly different. For the former one we utilize what one may call the classical toolbox of calculus, e.g. convex analysis, a Riemann type integral and the Banach fixed point theorem. For the latter we employ more recent concepts, such as direct methods from the calculus of variations and a fair amount of measure theory. The focus of our study are situations where either due to the given data or degeneration effects the solution are now longer continuous but allow for jumps. If solutions do have jumps the classical concepts of solution fail. They often require at least absolute continuity. It is hence that new notions of what constitutes a solution have to be developed. This is a field of active and ongoing discussion. In the following we give a short introduction to the problems we are going to address. It is more intended to give a short impression of the problem than to properly present the problem and our main results. We will postpone this to the first chapter of the corresponding parts of this thesis.

1.2. Structure of the thesis

This thesis is split into into three parts. First we introduce mathematical preliminaries needed in the sequel. The second part is devoted to the study of the quasivariational sweeping process. In the final part we are going to apply variational methods to analyze doubly nonlinear differential inclusions. Notice that Part II and III are independent from each other.

1.2.1. Part I - Mathematical preliminaries

The mathematical preliminaries address four topics. First we are going to dwell on *convex analysis*. After recalling some elementary definitions and properties of convex sets we consider the projection in Hilbert spaces. We then spend a section on convex functions and their properties. We finish the chapter with basics on maximal monotone operators. The second chapter is reserved for the theory of *classical function spaces*. We start by recalling the definitions of spaces such as functions of bounded variation and the space of regulated function and their properties. Subsequently we introduce the Kurzweil integral, a Riemann-Stieltjes type integral which allows for the integration of a BV function against a non-continuous function. The next chapter is concerned with *measure theory* especially with regard to Banach spaces. We are going to introduce vector valued measures and subsequently talk about the measurability of Banach space valued functions. There we are going to present apart from the usual Radon-Nikodym theorem on reflexive Banach spaces also a weaker version which is suited to the duals of separable spaces. Finally we are going to talk about the extension of Young measure theory to Banach spaces. In the short last chapter we are going to recall *Gronwalls Lemma* and prove two discrete analogues.

1.2.2. Part II - Quasivariational sweeping processes on functions of bounded variation

The sweeping process was introduced by J. J. Moreau [78, 79] in the 1970s. Given a time dependent convex set $K(t)$ in a Hilbert space X and an initial value $\xi_0 \in K(0)$ it seeks for a solution to the problem

$$-\dot{\xi}(t) \in N_{K(t)}\left(\xi(t)\right)\,, \quad \xi(0) = \xi_0\,.$$

Here $N_K(x)$ is the normal cone of K in the point x. To better visualize the problem, imagine the following: You put a raisin on a table, turn a cake form upside down and put it over the raisin. Where does the raisin go, when the cake form is moved? This is what the sweeping process describes. The interest in the sweeping process comes from two angles. First it is an interesting mathematical problem, being a special and extremal case of

$$-\dot{u}(t) \in A(t)u(t)$$

for a family of maximal monotone operators $A(t)$. It can hence serve as a first approach to solve the more general inclusion problem. Second it is of interest in applications, as it is a building brick to describing more complex problems, such as elastoplasticity. In the special case of $K(t) = u(t) - Z$ the sweeping process is equivalent to the play operator, one of the basic operators of the theory of hysteresis.

The quasivariational sweeping process is an extension of the sweeping process. It allows for the shape of the convex set to depend not only on time but also on the current state $\xi(t)$. In other words we are looking for solutions satisfying

$$-\dot{\xi}(t) \in N_{K(t,\xi(t))}\left(\xi(t)\right)\,, \quad \xi(0) = \xi_0\,.$$

This problem was originally posed by M. Kunze and M. D. P. Monteiro-Marquez in [64].In the above picture replace the cake form by a modern, silicone one. Also, exchange the raisin with a small stone. Then when moving the cake form around it shape will change, when it hits the stone. Already in the original paper existence of a solution has been proven. Uniqueness however is more involved and to our knowledge the first result was proven by M. Brokate, P. Krejčí and H. Schnabel in [18] six years later. Further results in this direction are to be found in [71, 95, 100].

We are going to study the problem, when the convex set $K(\cdot, \xi)$ may jump. Then one can no longer expect that the solution is smooth, but has to assume that it jumps as well. When developing a concept of solution which allows for that the essential degree of freedom one has is to decide what happens at the jump points. Below the notion of Kurzweil solutions is employed. Loosely speaking at jump points this concept determines the solution by projecting the state before the jump onto the convex set after the jump. It turns out that proving uniqueness depends strongly on the geometry of the convex set. We treat two different settings, one where K is a polyhedron with moving faces and two where K is a convex set with smooth boundaries.
The blueprint of our approach is roughly the following. Using time discretization and a limiting procedure we show a continuity result for the map $v \mapsto \eta$, where η is the solution of the sweeping process with input $K(t, v(t))$. Employing these results with Banach's fixed point theorem to deduce existence and uniqueness for the quasivariational case. The main tool for our analysis is besides limiting theorems, a good understanding of the projection.

1.2.3. Part III - Variational approach to doubly nonlinear differential inclusions

The term doubly nonlinear differential inclusion was coined by P. Colli and A. Visintin in their paper [21]. It stands for the following class of evolution problems. Let X be a Banach space and X^* its dual. Given two maximal monotone operators $A, B : X \rightrightarrows X^*$, a load $f : [0,T] \to X^*$ and an initial value $u_0 \in X$ we seek for a solution $u : [0,T] \to X$ such that

$$A\dot{u}(t) + Bu(t) \ni f(t)\,, \quad u(0) = u_0\,.$$

Existence of a solution has been proven, in case one of both operators is cyclic, i.e. the subdifferential of a convex function. If $A = \partial\Psi$ for some 1-homogeneous function Ψ, then the above problem implicitly includes a hysteresis operator (see Section 9.1). Doubly nonlinear differential inclusions are used to model phenomena such as elastoplasticity or fracture. Hence they attracted quite a bit of attention. Most generalizations consider problems of the type

$$\partial\Psi\left(\dot{u}(t)\right) + \nabla\mathcal{E}_t(u(t)) \ni 0\,, \tag{1.1}$$

where Ψ is a convex function, often called the dissipation functional, and $\mathcal{E}$ is a time dependent, possibly non-convex energy function. Especially in the case when Ψ is 1-homogeneous one can not necessarily expect the solution to be continuous. A method to solve these problems is the vanishing viscosity approach. Roughly speaking one adds to the left hand side a term of the type $\varepsilon\partial\widetilde{\Psi}\left(\dot{u}\right)$, where $\widetilde{\Psi}$ convex function satisfying sufficient growth conditions. These systems are easier to solve and one tries to obtain the solution to the original problem by taking the limit $\varepsilon \to 0$.

It is this procedure that gave rise to the problem we are going to study. The question we are going to ask is: given a family of maximal monotone operators $(\alpha_n)_{n\in\mathbb{N}}$ and an energy functional $\mathcal{E}$. Let u_n be the solutions of

$$\alpha_n\left(\dot{u}_n(t)\right) + \nabla\mathcal{E}_t\left(u_n(t)\right) \ni 0\,.$$

If $\alpha_n \to \alpha$, do the limit points u of the sequence (u_n) solve the limit problem

$$\alpha\left(\dot{u}(t)\right) + \nabla\mathcal{E}_t\left(u(t)\right) \ni 0\,.$$

in some sense? We will show below that this indeed is the case if $\alpha_n \to \alpha$ in the sense of graph convergence. In the case the space X is reflexive Banach we will be able to allow for general maximal monotone operators. If the X however is not reflexive we need to restrict ourselves to α_n which are the subdifferential of a convex function. A rather obvious generalization is to allow for a sequence of energy functionals $(\mathcal{E}_n)$ converging to some $\mathcal{E}$ and now ask the same question. It turns out that a convergence $\alpha_n \to \alpha$ and $\mathcal{E}_n \to \mathcal{E}$ alone is not enough, instead there needs to be an interplay between these two convergences. We will give a sufficient condition on how they need to interact, such that limit points of the sequence of solutions, indeed solve the limit problem.

In order to prove our results we take an approach that is maybe somewhat surprising. We reinterpret the evolution problem as a minimization problem on the space of functions $f : [0,T] \to$

X. This is possible due to the structure of the problem. We thereby enter the realm of the calculus of variations. Our question now translates to: assume there is a sequence of functionals, are limit points of their minimizers, also minimizers of the limiting functional? This problem has already been intensively studied and we will make use of the toolbox developed during these studies to solve our question.

Part I.

Mathematical preliminaries

Introductory remarks, notation and a guide through theses preliminaries

The present part of this thesis aims to provide the mathematical framework for the sequel studies. Many concepts and results we present here are well known. Others are small extensions of already established results which however are necessary for our considerations. And some are new but fit better into the context of the preliminaries than in the following parts. This is why they are located here.

Notation

In general we follow current conventions with regards to our notation. Especially we denote by $L^p(\Omega)$ the (equivalence class of) p-integrable functions on a set Ω. $W^{k,p}(\Omega)$ denotes the Sobolev space of all k-times weakly differentiable p-integrable functions whose weak derivatives up to order k are also elements of $L^p(\Omega)$.
For any Banach space $\mathcal{Y}$ we denote by $\mathcal{Y}^*$ its dual. Moreover the dual pairing is denoted by $\langle \cdot, \cdot \rangle_{\mathcal{Y}^*,\mathcal{Y}}$. The respective norms will be denoted by $|\cdot|_{\mathcal{Y}}$ and $|\cdot|_{\mathcal{Y}^*}$. If no confusion can arise the indices may be left out. If H is a Hilbert space we denote the inner product simply by $\langle \cdot, \cdot \rangle$. The norm on H is then given by $|h| = \langle h, h \rangle^{1/2}$.
For any $n \in \mathbb{N}$ we will write $[n]$ to indicate the set $\{1, 2, ...n\}$. Moreover let I be an index set. We shall at times sloppily write $(x_i)_{i \in I} \subset Z$ when we mean that $(x_i)_{i \in I}$ is a family of elements indexed by $i \in I$ and $x_i \in Z$ for all $i \in I$.
The remaining notation will be introduced in the sequel chapters.

A short guide through the preliminaries

If the reader is only interested in certain parts of this thesis and wishes to read only the necessary preliminaries we offer a short guide.

Part II. The necessary mathematical tools for the second part of this thesis will be provided in the sections 2.1-2.3 which contains some basics in convex analysis, chapter 3 which introduces 'classical' function space theory and a special Riemann-type integral. In chapter 5 some useful discrete analogous of Gronwalls inequality are proven.

Part III. For the third part we suggest to read the following: Sections 2.3-2.4 provide basic knowledge on convex functions and maximal monotone operators and chapter 4 contains a number of useful results concerning measure theory on Banach spaces.

2. Convex analysis

This chapter is dedicated to an introduction of concepts and methods of convex analysis. We do this with a certain depth which might seem a little exceptional for these preliminaries. However as these tools and methods of convex analysis will come into play throughout the whole thesis and are at points at the very core of our study we feel that this is appropriate. Convex analysis is the subject of many textbooks and most if not all the results presented here can be found in one of them. We would like to point out the monographs [92, 34, 47] and the lecture notes [17] as our main references. Many results are presented without proof and we will indicate the appropriate reference instead.
Throughout this chapter $\mathcal{Y}$ is a real Banach space and $\mathcal{Y}^*$ its dual. We indicate their dual pairing by $\langle \cdot, \cdot \rangle_{\mathcal{Y}^*,\mathcal{Y}}$. The respective norms are denoted by $|\cdot|_{\mathcal{Y}}$ and $|\cdot|_{\mathcal{Y}^*}$. Furthermore we set X to be a Hilbert space with scalar product $\langle \cdot, \cdot \rangle$ and norm $|x| = \langle x, x \rangle^{1/2}$.

2.1. Convex sets

Definition 2.1 (Convex sets)**.** A set $Z \subset \mathcal{Y}$ is called convex if

$$\forall x, y \in Z,\ \forall \lambda \in (0,1):\ \lambda x + (1-\lambda) y \in Z\,.$$

In other words, a set Z is convex if the line segment between two points of Z lies entirely in Z. Therefore the whole space $\mathcal{Y}$ as well as any affine subspace of $\mathcal{Y}$ are convex. The empty set is convex, too. Another example for convex sets are polyhedra. They are the intersection of finitely many subspaces. To be precise we shall give the definition.

Definition 2.2. A set $P \subset \mathcal{Y}$ is called a polyhedron if there exist $n \in \mathbb{N}$, $(a_i)_{i=1}^n \subset \mathcal{Y}^*$ and $\beta \in \mathbb{R}^n$ such that

$$P = \left\{ y \in \mathcal{Y} : \langle a_i, y \rangle_{\mathcal{Y}^*,\mathcal{Y}} \leq \beta_i \right\}\,. \tag{2.1}$$

Definition 2.3. Let $y \in \mathcal{Y}$ and $Z \subset \mathcal{Y}$ be a convex set. Then we define the distance between y and Z by

$$d(y, Z) := \min \left\{ |y - z| : z \in Z \right\}\,. \tag{2.2}$$

There are several ways to define the distance between two convex sets. Throughout this thesis we are going to use the Hausdorff distance.

Definition 2.4 (Hausdorff distance of convex sets)**.** Let Z_1, Z_2 be two convex sets. Then the Hausdorff distance between Z_1 and Z_2 is defined as

$$d_H(Z_1, Z_2) = \max \left\{ \sup_{z_1 \in Z_1} d(z_1, Z_2),\ \sup_{z_2 \in Z_2} d(z_2, Z_1) \right\}\,. \tag{2.3}$$

Notice that the Hausdorff distance is always well defined in $[0, +\infty]$ and symmetric. We just mention that another common notion of distance is the so called Minkowski distance. It will however not be used in the present thesis.

Definition 2.5 (Polar set)**.** To each $Z \subset \mathcal{Y}$ we associate its polar set $Z^* \subset \mathcal{Y}^*$ which is defined by

$$Z^* := \left\{ y' \in \mathcal{Y}^* : \langle y', y \rangle_{\mathcal{Y}^*, \mathcal{Y}} \leq 1 \; \forall y \in Z \right\} . \tag{2.4}$$

For any Z the set Z^* is non-empty, convex and closed with respect to the norm on $\mathcal{Y}^*$. For example the polar set of the closed ball with radius $c > 0$ and centered in the origin

$$B_c(0) := \{ y \in \mathcal{Y} : |y|_{\mathcal{Y}} \leq 1 \}$$

is the closed ball with radius c^{-1} and center 0 in $\mathcal{Y}^*$.

Definition 2.6 (Normal cone)**.** Let $\emptyset \neq Z \subset X$ be convex. Then the normal cone of Z in a point $z \in Z$ is defined by

$$N_Z(x) := \{ y \in X : \langle y, x - z \rangle \leq 0 \;\; \forall x \in Z \} . \tag{2.5}$$

As the name already indicates, the normal cone is indeed a cone, that is a set $C \subset \mathcal{Y}$ such that

$$\forall y \in C, \forall \lambda \geq 0 : \lambda y \in C .$$

From the definition one directly obtains that the normal cone of a point $x \in \text{int}(Z)$ is in fact $\{0\}$. For $c > 0$ and $|x| = c$ we have $N_{B_c(0)}(x) = \{ \lambda x : \lambda > 0 \}$. One can also explicitly calculate the normal cone of a polyhedron.

Lemma 2.7. *Let $\emptyset \neq P \subset X$ be a polyhedron. For $z \in P$ we denote by*

$$\Gamma_P(z) = \{ i \in [n] : \; \langle a_i, z \rangle = \beta_i \} \tag{2.6}$$

the set of active constraints. Then

$$N_P(z) = \left\{ w = \sum_{k \in \Gamma_P} \alpha_k n_k : \; \alpha_k \geq 0 \right\} . \tag{2.7}$$

This result is due to V. Lovicar and has been published in [55, Lemma 7.3].

Definition 2.8 (Recession cone)**.** Let $Z \subset \mathcal{Y}$ be convex. Then the recession cone of Z is the set

$$\text{rec}(Z) := \{ y \in \mathcal{Y} : [\exists z \in Z : z + \lambda y \in Z \; \forall \lambda > 0] \} .$$

As one can easily see, the recession cone is indeed a cone. One can show that $y \in \text{rec}(Z)$ if and only if $z + \lambda y \in Z$ for all $\lambda > 0$ and *all* $z \in Z$. Therefore the recession cone is the set of all directions in which Z is unbounded. Consequently the recession cone of a bounded set is the empty set.

Proposition 2.9. *Let $Z \subset \mathcal{Y}$ be convex and non-empty. Then if Z is closed then so is* $\text{rec}(Z)$*.*

2.2. Projection onto convex sets in Hilbert spaces

Definition 2.10 (Orthogonal projection)**.** Let $\emptyset \neq Z \subset X$ be a closed convex set and $u \in X$. The point $x \in Z$ which satisfies

$$\forall z \in Z: \ |u - x| \leq |u - z|$$

is called the orthogonal projection of u onto Z. We shall denote it by $Q_Z(u) := x$.

The following result is well known.

Proposition 2.11. *For any closed convex set $\emptyset \neq Z \subset X$ and any $u \in X$ the orthogonal projection $Q_Z(u)$ exists and is unique. Furthermore it satisfies the variational inequality*

$$\langle u - Q_Z(u), Q_Z(u) - z\rangle \geq 0 \quad \forall z \in Z\,. \tag{2.8}$$

We furthermore define the complementary map P_Z by

$$P_Z := \mathrm{Id}_X - Q_Z\,. \tag{2.9}$$

We are aware that it might appear unintuitive to denote the projection by Q and its complement by P. It is however consistent with the notation in the study of hysteresis and especially the play operator, see e.g. [55, 58]. Our work here is in that context. It even more relies on the aforementioned articles. We hence decided to follow this notation.
Due to Proposition 2.11 we see that $P_Z(u) \in N_Z\left(Q_Z(u)\right)$. A number of further properties of the projection and its complement are part of the upcoming Lemma.

Lemma 2.12. *Let $\emptyset \neq Z \subset X$ be closed and convex. It holds*

(i) *Q_Z is Lipschitz continuous with constant 1, that is*

$$\forall u, v \in X: \ |Q_Z(u) - Q_Z(v)| \leq |u - v|\,,$$

(ii) $\langle P_Z x - P_Z y, Q_Z x - Q_Z y\rangle \ \geq \ 0$,

(iii) $Q_Z(Q_Z x + \lambda P_Z x) \ = \ Q_Z x \quad \forall \lambda \geq 0$,

(iv) $(x \in Z\,,\ \langle y, x - z\rangle \ \geq \ 0 \ \forall z \in Z) \iff (x = Q_Z(x + y)\,,\ y = P_Z(x + y))$.

The proof of the above assertions heavily relies on the characterization (2.8). In the special case of a projection onto a linear subspace of X we obtain a few additional properties.

Proposition 2.13. *Let $X'' \subset X'$ be a linear subspaces of X. Then the following holds true:*

(i) $\forall x \in X: Q_{X''}(Q_{X'}x) = Q_{X''}(x) = Q_{X'}(Q_{X''}x)$

(ii) $\forall x \in X: |P_{X'}x|^2 = |x|^2 - |Q_{X'}x|^2$

(iii) $\forall \lambda \in \mathbb{R}, x \in X: Q_{X'}(\lambda x) = \lambda Q_{X'}x$

(iv) $\forall x \in X, v \in X': Q_{X'}(x + v) = Q_{X'}(x) + v$

(v) $\forall x \in X, z \in X' : \langle z, x\rangle = \langle z, Q_{X'}x\rangle$

(vi) $\forall x \in X, z \in X', |z| = 1 : |\langle z, x\rangle| \leq |Q_{X'}x| \leq |x|$

Proof. Since we were not able to find a reference for the above assertions, we prove them here.

ad (i): The second equality is trivial since $Q_{X''}(x) \in X'' \subset X'$. For the sake of notational simplicity we set $z := Q_{X'}x$ and $z' := Q_{X''}(Q_{X'}x)$. For any $y \in X''$ we have $y + z - z' \in X'$. Hence we have

$$\left\langle x - z, z' - y\right\rangle \geq 0 \text{ and } \left\langle z - z', z' - y\right\rangle \geq 0\,.$$

Adding both inequalities leads to $\langle x - z', z' - y\rangle \geq 0$ for all $y \in X''$ and therefore $z' = Q_{X''}x$.

ad (ii): It suffices to show that $\langle P_{X'}x, Q_{X'}x\rangle = 0$. To this end notice that X' is a linear space and therefore 0 and $-2Q_{X'}x$ are in X'. Applying (2.8) to both points implies the claim.

ad (iii): For $\lambda = 0$ we notice that $Q_{X'}(0) = 0$ and the assertion is proven. For $\lambda \neq 0$ and $y \in X'$ we have that $\lambda^{-1}y \in X'$ and due to (2.8) we obtain

$$0 \leq \lambda^2 \left\langle x - Q_{X'}(x), Q_{X'}x - \lambda^{-1}y\right\rangle = \langle \lambda x - \lambda Q_{X'}x, \lambda Q_{X'}x - y\rangle\ .$$

This directly implies that $Q_{X'}(\lambda x) = \lambda Q_{X'}x$.

ad (iv): For any $z \in X'$ also $z - v \in X'$. Hence

$$0 \leq \langle x - Q_{X'}x, Q_{X'}x - z\rangle = \langle x + v - (Q_{X'}x + v)\,, Q_{X'}x + v - z)\rangle\ \forall z \in X'\,.$$

ad (v): Notice that $x = Q_{X'}x + P_{X'}x$ and $\langle P_{X'}x, z\rangle = 0$ for all $z \in X'$.

ad (vi): The first inequality is a consequence of (v) and the second inequality of (ii). □

2.3. Convex functions

Definition 2.14 (Convex functions)**.** A function $f : \mathcal{Y} \to (-\infty, \infty]$ is called convex if

$$\forall x, y \in Z,\ \forall \lambda \in (0,1) :\ f(\lambda x + (1-\lambda)\,y) \leq \lambda f(x) + (1-\lambda)\,f(y)\,.$$

The connection between convex sets and convex functions becomes apparent in the following well known result, see e.g. [34, Proposition 2.1].

Proposition 2.15. *A function $f : \mathcal{Y} \to (-\infty, \infty]$ is convex if and only if the epigraph*

$$\text{epi}(f) := \{(x, y) \in \mathcal{Y} \times \mathbb{R} : f(x) \leq y\}$$

is a convex subset of $\mathcal{Y} \times \mathbb{R}$.

The following result can be found in [37, Proposition 4.34].

Proposition 2.16. *Let $J \subset \mathbb{R}$ be an interval and $g : J \to (-\infty, +\infty]$. Then g is convex if for every $t_0 \in J$ with $g(t_0) < +\infty$ the difference quotient*

$$t \mapsto \frac{g(t) - g(t_0)}{t - t_0}$$

is nondecreasing in $J \setminus \{t_0\}$.

A simple example for a convex function is the so called indicator function.

Example 2.17. Let $Z \subset \mathcal{Y}$ and define its indicator function I_Z by

$$I_Z(y) := \begin{cases} 0 & \text{if } y \in Y \\ +\infty & \text{else.} \end{cases}$$

Then I_Z is convex if and only if Z is convex. The function I_Z is indeed not the only function that goes under the name of indicator function. Also the function $\chi_Z(x)$, which is 1 if $x \in Z$ and 0 else, is known as an indicator function. For clarity of notation whenever we shall use one of these letters we will mean the respective function defined above.

Definition 2.18. Let $\mathcal{Y}$ be a Banach space. We say a function $f : \mathcal{Y} \to (-\infty, \infty]$ is lower semicontinuous with respect to the strong (weak / weak-*) topology on $\mathcal{Y}$, if for any sequence $(y_n)_{n \in \mathbb{N}} \subset \mathcal{Y}$ with

$$y_n \to y \quad (y_n \rightharpoonup y / y_n \overset{*}{\rightharpoonup} y)$$

it holds

$$\liminf_{n \to \infty} f(y_n) \geq f(y) \,.$$

Proposition 2.19. *A convex function $f : \mathcal{Y} \to (-\infty, +\infty]$ is lower semicontinuous if and only if* epi(f) *is closed in the strong topology of $\mathcal{Y} \times \mathbb{R}$.*

Definition 2.20. For $f : \mathcal{Y} \to (-\infty, \infty]$ we define its effective domain by

$$\mathrm{dom}(f) := \{y \in \mathcal{Y} : f(y) < +\infty\} \,.$$

A function f is called proper if $\mathrm{dom}(f) \neq \emptyset$.

For example the indicator function I_Z is proper if and only if Z is nonempty. Its effective domain is Z.

Definition 2.21. Let $0 \in Z \subset \mathcal{Y}$ be a convex set. The function $M_Z : \mathcal{Y} \to (-\infty, \infty]$, defined by

$$M_Z(y) := \inf \left\{t > 0 : t^{-1} y \in Z\right\} \tag{2.10}$$

is called the Minkowski functional of Z. By convention the infimum of the empty set is $+\infty$.

Example 2.22. The Minkowski functional of the unit ball $B_1(0) = \{y \in \mathcal{Y} : |y|_{\mathcal{Y}} \leq 1\}$ is exactly the norm on $\mathcal{Y}$, in formula

$$M_{B_1(0)}(y) = |y|_{\mathcal{Y}} \,.$$

The following results are well known.

Proposition 2.23. *For any $Z \subset \mathcal{Y}$ convex with $0 \in Z$ the Minkowski functional M_Z is well defined. It is a proper, convex function and has the following properties:*

(i) *If $M_Z(y) < 1$ then $y \in Z$. If $M_Z(y) > 1$ then $y \notin Z$.*

(ii) *If Z is closed, then $M_Z(y) \leq 1 \Leftrightarrow y \in Z$.*

(iii) *M_Z is 1-homogeneous, i.e. $\forall \lambda > 0$ and $y \in \mathcal{Y}$ it holds $M_Z(\lambda y) = \lambda M_Z(y)$.*

Definition 2.24 (Legendre-Fenchel conjugate)**.** Let $f : \mathcal{Y} \to (-\infty, \infty]$ be a proper function. The function $f^* : \mathcal{Y}^* \to (-\infty, \infty]$ defined by

$$f^*(y') := \sup\{\langle y', y\rangle_{\mathcal{Y}^*,\mathcal{Y}} - f(y) : y \in \mathcal{Y}\}$$

is called the (Legendre-Fenchel) conjugate of f.

Proposition 2.25. *Let $f : \mathcal{Y} \to (-\infty, \infty]$ be a proper function. Then the conjugate f^* is well defined, convex and lower semicontinuous.*

For a proof of this result see [17, Lemma 6.4]. Notice that due to the definition of the conjugate function we have

$$\forall y \in \mathcal{Y},\ y' \in \mathcal{Y}^* : f(y) + f^*\left(y'\right) \geq \langle y', y\rangle_{\mathcal{Y}^*,\mathcal{Y}} . \tag{2.11}$$

Example 2.26. The indicator function and the Minkowski functional are related through the concept of conjugate functions. Indeed it holds that

(i) for $0 \in Z \subset \mathcal{Y}$ convex $(M_Z)^* = I_{Z^*}$ and

(ii) for $Z \subset \mathcal{Y}$ convex $(I_Z)^* = M_{Z^*}$.

Proposition 2.27. *Let $\mathcal{Y}$ be a reflexive Banach space, and $f : \mathcal{Y} \to (-\infty, \infty]$. Let f^{**} be the conjugate function of f^*. Then the following assertions hold*

(i) *$f^{**} \leq f$*

(ii) *If f is proper, convex and lower semicontinuous then $f^{**} = f^*$.*

Corollary 2.28. *Let $\mathcal{Y}$ be a reflexive Banach space and $f : \mathcal{Y} \to (-\infty, +\infty]$ be a proper, convex and lower semicontinuous function. Then*

$$f(y) = \sup_{y' \in \mathcal{Y}^*} \left\{\langle y', y\rangle_{\mathcal{Y}^*,\mathcal{Y}} - f^*(y')\right\} .$$

Definition 2.29 (Subgradient and subdifferential)**.** Let $f : \mathcal{Y} \to (-\infty, \infty]$ be a proper convex function and let $x \in \mathcal{Y}$. A vector $y' \in \mathcal{Y}^*$ is called subgradient of f in x if

$$\forall y \in \mathcal{Y} : \langle y', y - x\rangle_{\mathcal{Y}^*,\mathcal{Y}} \leq f(y) - f(x) . \tag{2.12}$$

The set of all subgradients in x is the subdifferential of f in x. It is denoted by $\partial f(x)$. The domain of ∂f is defined as

$$\mathrm{dom}\,(\partial f) = \{x \in \mathcal{Y} : \partial f(x) \neq \emptyset\} .$$

As a direct consequence from the definition we see that for any $x \in \mathcal{Y}$ with $f(x) = +\infty$ we have $\partial f(x) = \emptyset$, hence $\mathrm{dom}(\partial f) \subset \mathrm{dom}(f)$. If f is differentiable in $y \in \mathcal{Y}$ then $\partial f(y) = \{\nabla f(y)\}$. The connection between the subdifferential and the conjugate function is part of the following theorem, see also [17, Theorem 7.6].

Theorem 2.30. *Let $f : \mathcal{Y} \to (-\infty, \infty]$ and $y \in \mathcal{Y}$. Then*

$$y' \in \partial f(y) \iff f(y) + f^*\left(y'\right) = \left\langle y', y\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} . \tag{2.13}$$

Remark 2.31. In light of (2.11) we also have that

$$y' \in \partial f(y) \iff f(y) + f^*\left(y'\right) \leq \left\langle y', y\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} . \tag{2.14}$$

Example 2.32. Let $\emptyset \neq Z \subset X$ be a convex set. Then the subdifferential of the indicator function of Z in x is identical to the normal cone of Z in x, in formula

$$\partial I_Z(x) = N_Z(x) .$$

Definition 2.33 (Recession function)**.** Let $f : \mathcal{Y} \to (-\infty, \infty]$ be a proper convex function. The recession function of f is defined by

$$f^\infty : \mathcal{Y} \to [-\infty, \infty], \quad z \mapsto \sup\left\{f(w+z) - f(w) : w \in \mathrm{dom}(f)\right\} .$$

The following theorem shows the connection between the recession function and the recession cone defined above. Indeed it holds that the recession cone of the epigraph coincides with the epigraph of the recession function.

Theorem 2.34. *Let $f : \mathcal{Y} \to (-\infty, \infty]$ be a proper convex function. Then its recession function f^∞ is a positively 1-homogeneous, proper convex function and*

$$\mathrm{epi}\left(f^\infty\right) = \mathrm{rec}\left(\mathrm{epi} f\right) .$$

Moreover if f is lower semicontinious, then so is f^∞ and

$$f^\infty(z) = \sup_{t>0} \frac{f(w+tz) - f(w)}{t} = \lim_{t\to\infty} \frac{f(w+tz) - f(w)}{t} \quad \forall w \in \mathrm{dom}(f) . \tag{2.15}$$

A proof for the case $\mathcal{Y} = \mathbb{R}^m$ can be found in [37, Theorem 4.70]. However we could not find a proof for Banach spaces, and thus shall present it below. It relies heavily on the afore mentioned proof due to Fonseca and Leoni.

Proof. Observe that f^∞ is convex as a point wise supremum of a family of convex functions. Therefore $\mathrm{epi}\left(f^\infty\right)$ is convex as well. It then holds that

$$\begin{aligned}
(z,t) \in \mathrm{rec}\left(\mathrm{epi} f\right) &\Leftrightarrow \forall (w,s) \in \mathrm{epi} f : (w,s) + (z,t) \in \mathrm{epi} f \\
&\Leftrightarrow \forall w \in \mathrm{dom}(f), s \geq f(w) : f(w+z) \leq s + t \\
&\Leftrightarrow \forall w \in \mathrm{dom} f : f(w+z) \leq f(w) + t \\
&\Leftrightarrow f^\infty(z) \leq t \\
&\Leftrightarrow (z,t) \in \mathrm{epi}\left(f^\infty\right)
\end{aligned}$$

Since $f^\infty(z) \geq f(w+z) - f(w)$ the function f^∞ does never attain the value $-\infty$. Moreover since $f^\infty(0) = 0$ it is also a proper function. For positive 1-homogeneity notice that for any $(z, f^\infty(z)) \in \text{epi}\,(f^\infty)$ and $t > 0$ we have that also $\text{epi}(f) + t\,(z, f^\infty(z)) \in \text{epi}(f)$ because of the definition of a recession cone . Therefore for any $w \in \text{dom}(f)$ we have

$$f(w+tz) \leq f(w) + tf^\infty(z) \tag{2.16}$$

and hence $f^\infty(tz) \leq tf^\infty(z)$. For the reverse inequality it suffices to replace z and t with tz and $\frac{1}{t}$ respectively.
As for the lower semicontinuity of f^∞ we first remark that due to Proposition 2.19 $\text{epi}(f)$ is closed if f is lower semicontinuous. Using Proposition 2.9 then also $\text{rec}\,(\text{epi}f)$ is closed and therefore f^∞ is lower semicontinuous.
It remains to prove (2.15).Following Proposition 2.16 it suffices to show that

$$f^\infty(z) = \sup_{t>0} \frac{f(w+tz) - f(w)}{t}$$

for all $w \in \text{dom}(f)$ and $z \in \mathcal{Y}$. To this end first notice that due to formula (2.16) we have

$$\sup_{t>0} \frac{f(w+tz) - f(w)}{t} \leq f^\infty(z)\,.$$

For the reverse inequality observe that if

$$\sup_{t>0} \frac{f(w+tz) - f(w)}{t} = +\infty$$

then there is nothing to prove. Hence we may assume that

$$\sup_{t>0} \frac{f(w+tz) - f(w)}{t} \leq s$$

for some $s \in \mathbb{R}$. Consequently for all $t > 0$ we have $f(w+tz) \leq f(w) + s \cdot t$ and we obtain

$$(w + tz, f(w) + s \cdot t) \in \text{epi}f\,.$$

Therefore $(z, s) \in \text{rec}\,(\text{epi}f)$ which readily implies $f^\infty(z) \leq s$. As the choice of s was arbitrary we have shown that

$$f^\infty(z) \leq \sup_{t>0} \frac{f(w+tz) - f(w)}{t}$$

which completes our proof. □

Corollary 2.35. *Let $f : \mathcal{Y} \to (-\infty, \infty]$ be a proper, convex and 1-homogeneous function. Then $f^\infty = f$.*

Proof. Let $(x, y) \in \text{epi}(f)$, that is $f(x) \leq y$. Due to the 1-homogeneity we have for all $\lambda > 0$: $f(\lambda x) = \lambda f(x) \leq \lambda y$. Hence $(\lambda x, \lambda y) \in \text{epi}(f)$. Therefore we obtain $\text{epi}\,(f) \subset \text{rec}\,(\text{epi}f)$. Since f is 1-homogeneous $f(0) = 0$. Hence for any $(x, y) \in \text{rec}(\text{epi}f)$ it holds $(0, 0) + (x, y) \in \text{epi}(f)$, i.e. $(x, y) \in \text{epi}(f)$. Therefore $\text{epi}\,(f) = \text{rec}\,(\text{epi}f)$ and Theorem 2.34 implies the claim. □

2.4. Maximal monotone operators and variational notions of convergence

In this short section we shall give a few basic facts about maximal monotone operators and some related notions of convergence. For a more thorough introduction of the topics we discuss below we refer to the monographs [10] and [4] and references therein. An operator $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ is a set valued map which to each $y \in \mathcal{Y}$ associates a set $\alpha(y) \subset \mathcal{Y}^*$ we denote the graph of α by the set

$$g_\alpha := \left\{(y, y') \in \mathcal{Y} \times \mathcal{Y}^* : y' \in \alpha(y)\right\} . \tag{2.17}$$

There is of course a one-to-one connection between each operator and its graph. Hence we will denote both by α and when we write $(x, y) \in \alpha$ we mean $(x, y) \in g_\alpha$. For any α we denote its domain by

$$D(\alpha) = \{y \in \mathcal{Y}^* : \alpha(y) \neq \emptyset\} . \tag{2.18}$$

Definition 2.36. A operator $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ is said to be monotone if

$$\forall (x, y) \in \alpha : \left[\langle y - y_0, x - x_0\rangle_{\mathcal{Y}^*,\mathcal{Y}} \geq 0 \;\forall (x_0, y_0) \in \alpha\right] . \tag{2.19}$$

α is said to be maximal monotone if it is monotone and there exists no monotone operator β such that

$$g_\alpha \subset g_\beta . \tag{2.20}$$

Let $f : \mathcal{Y} \to (-\infty, \infty]$ be a convex, proper and lower semicontinuous function. Then the map $\partial f : \mathcal{Y} \to \mathcal{Y}^*$ is a maximal monotone operator. The set of all maximal monotone operators of the type ∂f can be characterized.

Definition 2.37. Let $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ be a maximal monotone operator. Then α is called cyclic monotone if for all families $(y_i, y_i')_{i=0}^n$ with $(y_i, y_i') \in \alpha$ it holds

$$\sum_{i=0}^{n-1} \left\langle y_i', y_i - y_{i+1}\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} + \left\langle y_n', y_n - y_0\right\rangle_{X^*,X} \geq 0 . \tag{2.21}$$

Theorem 2.38 (Rockafeller [91])**.** *Let $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ be a maximal monotone operator. Then α is cyclically monotone if and only if there exists a convex, lower semicontinuous function $f : \mathcal{Y} \to (-\infty, \infty]$ such that*

$$\alpha = \partial f . \tag{2.22}$$

Moreover for a given maximal monotone and cyclic α the function f is uniquely determined up to a constant.

We also define the inverse of an maximal monotone operator α.

Definition 2.39. Let $\mathcal{Y}$ be a reflexive Banach space and $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ be a maximal monotone operator. Then the inverse of α is denoted by $\alpha^{-1} : \mathcal{Y}^* \rightrightarrows \mathcal{Y}$ and defined by

$$(y', y) \in \alpha^{-1} \quad \Longleftrightarrow \quad (y, y') \in \alpha \tag{2.23}$$

It is simple to see that α^{-1} is also a maximal monotone operator.

We shall finish this chapter by introducing a few notions of convergences that are important in the sequel. The first is a notion of convergence for convex sets and convex functions.

Definition 2.40 (Mosco convergence ([80]))**.** Let $(C_n)_{n\in\mathbb{N}}$ be a family of sets $C_n \subset \mathcal{Y}$. We say that C_n converges to a set $C \subset \mathcal{Y}$ in the sense of Mosco convergence, in formula $C_n \xrightarrow{M} C$ if

(i) for all families $(y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}$, $y_n \in C_n$, $y_n \rightharpoonup y$ it holds $y \in C$ and

(ii) for all $y \in C$ there exists a family $(y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}$, $y_n \in C_n$ such that $y_n \to y$.

Let $(f_n)_{n\in\mathbb{N}}$ be a family of functions $f_n : \mathcal{Y} \to (-\infty, \infty]$. We say f_n converges to $f : \mathcal{Y} \to (-\infty, \infty]$ in the sense of Mosco convergence, in formula $f_n \xrightarrow{M} f$ if the following two conditions are met

(i) **($\liminf$-inequality).** For all sequences $(y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}$, $y_n \rightharpoonup y$ it holds

$$\liminf_{n\to\infty} f_n(y_n) \geq f(y)\,. \tag{2.24}$$

(ii) **(Existence of a strong recovery sequence).** For all $y \in \mathcal{Y}$ there exists a sequence $(y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}$, $y_n \to y$ such that

$$\limsup_{n\to\infty} f_n(y_n) \leq f(y)\,. \tag{2.25}$$

The notion of Mosco convergence has been developed by Umberto Mosco in the late 60s in order to study variational inequalities. We will list a number of well known and important properties of this notion of convergence.

Proposition 2.41. *Let $(f_n)_{n\in\mathbb{N}}$ be a family of convex functions $f_n : \mathcal{Y} \to (-\infty, \infty]$. Then it holds*

$$f_n \xrightarrow{M} f \quad \Longleftrightarrow \quad \mathrm{epi}(f_n) \xrightarrow{M} \mathrm{epi}(f)\,.$$

On reflexive Banach spaces Mosco convergence of a sequence of functions carries over to its convex conjugate functions.

Theorem 2.42 ([4, Theorem III.3.18])**.** *Let $\mathcal{Y}$ be a reflexive Banach space and $(f_n)_{n\in\mathbb{N}}$ be a family of proper, convex and lower semicontinuous functions $f_n : \mathcal{Y} \to (-\infty, \infty]$. Then it holds*

$$f_n \xrightarrow{M} f \quad \Longleftrightarrow \quad f_n^* \xrightarrow{M} f^*\,.$$

A related notion of convergence for maximal monotone operators is the so-called graph convergence.

Definition 2.43 (Graph convergence)**.** Let $\mathcal{Y}$ be a reflexive Banach space. Let $(\alpha_n)_{n\in\mathbb{N}}$ be a family of maximal monotone operators $\alpha_n : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$. We say that α_n converges to a maximal monotone operator α in the sense of graph convergence, $\alpha_n \xrightarrow{g} \alpha$, if for all $(y, y') \in \alpha$ there exists a family $(y_n, y_n')_{n\in\mathbb{N}}$, $(y_n, y_n') \in \alpha_n$ such that

$$y_n \to y \quad \wedge \quad y_n' \to y'\,. \tag{2.26}$$

Notice that by its very definition if $\alpha_n \xrightarrow{g} \alpha$ then also $\alpha_n^{-1} \xrightarrow{g} \alpha^{-1}$.

Proposition 2.44 ([4, Proposition III.3.59]). *Let $\mathcal{Y}$ be a reflexive Banach space and $(\alpha_n)_{n\in\mathbb{N}}$ be a family of maximal monotone operators $\alpha_n : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$. Assume $\alpha_n \xrightarrow{g} \alpha$. Then for any sequence $(y_n, y_n')_{n\in\mathbb{N}}$, $(y_n, y_n') \in \alpha_n$ with either*

$$\begin{aligned} y_n \to y \quad &\wedge \quad y_n' \rightharpoonup y' \qquad \text{or} \\ y_n \rightharpoonup y \quad &\wedge \quad y_n' \to y' \end{aligned}$$

it holds $(y, y') \in \alpha$.

The connection from graph convergence and Mosco convergence is the following.

Theorem 2.45 ([4, Theorem III.3.66]). *Let $\mathcal{Y}$ be a reflexive Banach space and $(f_n)_{n\in\mathbb{N}}$ be a family of proper, convex and lower semicontinuous functions $f_n : \mathcal{Y} \to (-\infty, \infty]$. Then the following two are equivalent:*

(i) $f_n \xrightarrow{M} f$

(ii) $\partial f_n \xrightarrow{g} \partial f$ *and there is $(y.y')$, $y' \in \partial f(y)$ for which a family $(y_n, y_n')_{n\in\mathbb{N}}$, $(y_n, y_n') \in \partial_n$ exists such that*

$$y_n \to y \quad \wedge \quad y_n' \to y' \quad \wedge \quad f_n(y_n) \to f(y) \, .$$

3. Classical function spaces and integration theory

We are going to introduce 'classical' notions of function spaces and corresponding notions of integration. The concepts we are going to present here are used in Part II of the this thesis.

3.1. Classical function spaces: BV, G and S

In this section we are going to introduce the spaces of functions of (essentially) bounded variation BV ($\overline{\text{BV}}$), of regulated functions G and of step functions S. For a thorough investigation of these spaces and many of the proofs we omit in the sequel we refer to the book by Aumann [5]. Throughout this section we will assume that $-\infty < r < s < \infty$ and $\mathcal{Y}$ is a Banach space equipped with norm $|\cdot|$.

Definition 3.1 (One sided limits)**.** Let $f : [r, s] \to \mathcal{Y}$. For $t \in [r, s)$ assume that an $f_t^+ \in \mathcal{Y}$ exists such that for any sequence $(h_n)_{n\in\mathbb{N}}$, $h_n \to 0$ with $h_n > 0$ and $t + h_n \leq s$ for all $n \in \mathbb{N}$ it holds

$$f_t^+ = \lim_{n\to\infty} f(t + h_n) \ . \tag{3.1}$$

If such an f_t^+ exists we call it the right side limit of f at t and shortly write $f(t+)$. Analogously we define for any $t \in (r, s]$ the left side limit $f(t-)$ if it exits.

By convention we will set $f(s+) = f(s)$ and $f(r-) = f(r)$.

Definition 3.2 (Regulated functions)**.** A function $f : [r, s] \to \mathcal{Y}$ is called regulated if for all $t \in (r, s]$ the left hand side limit $f(t-)$ and for all $t \in [r, s)$ the right hand side limit $f(t+)$ exists. Furthermore we denote the space of all regulated functions on $[r, s]$ with valued in $\mathcal{Y}$ by $G(r, s; \mathcal{Y})$.

Of course, if we are to work with this function space, we will need a suitable norm on it. To this end we choose the well known supremum norm. Just to remind ourselves: For $f : [r, s] \to \mathcal{Y}$ we denote by

$$\|f\|_\infty := \sup\{|f(t)| : t \in [r, s]\} \tag{3.2}$$

the sup-norm of f. We say that f is bounded with respect to $\|\cdot\|_\infty$ if $\|f\|_\infty < \infty$. If we at times laxly state that 'f is bounded' then we imply that it is bounded with respect to this norm.

Proposition 3.3. *For any $f \in G(r, s; \mathcal{Y})$ it holds $\|f\|_\infty < \infty$. Furthermore $\|\cdot\|_\infty$ is a norm on $G(r, s; \mathcal{Y})$ and $(G(r, s; \mathcal{Y}), \|\cdot\|_\infty)$ is a Banach space.*

Definition 3.4. For $-\infty < r < s < \infty$ we denote the set of all finite partitions of $[r, s]$ by

$$\mathcal{P}_{[r,s]} = \{(t_k)_{k=0}^n : n \in \mathbb{N},\ t_0 = r,\ t_n = s\ \wedge\ (\forall k \in [n] : t_{k-1} < t_k)\} \ .$$

We will write $\mathcal{P}$ for $\mathcal{P}_{[0,T]}$.

Definition 3.5 (Step function)**.** A function $f : [r,s] \to \mathcal{Y}$ is called a step function, in formula $S(r,s;\mathcal{Y})$, if there exists a partition $(t_k)_{k=0}^n \in \mathcal{P}_{[r,s]}$ and $(f_k)_{k=1}^n$, $(\widehat{f}_k)_{k=0}^n$ such that

$$f(t) = \sum_{k=1}^{n} \chi_{(t_{k-1},t_k)}(t) f_k + \sum_{k=0}^{n} \chi_{\{t_k\}}(t) \widehat{f}_k \,. \tag{3.3}$$

Lemma 3.6. *For any regulated function $f \in G(r,s;\mathcal{Y})$ and any $\varepsilon > 0$ there exists a partition $(t_k)_{k=0}^n \subset \mathcal{P}_{[r,s]}$ such that*

$$\forall k \in [n] : \forall t,s \in (t_{k-1},t_k) : |f(t) - f(s)| \leq \varepsilon \,. \tag{3.4}$$

Especially the set of step functions $S(r,s;\mathcal{Y})$ is dense in $G(r,s;\mathcal{Y})$ with respect to $\|\cdot\|_\infty$.

A proof of this Lemma can be found in [5, p. 237f].

Definition 3.7 (Variation and Essential Variation)**.** For $f : [r,s] \to \mathcal{Y}$ its variation is defined by

$$\mathrm{Var}(f,[r,s]) := \sup \left\{ \sum_{k=0}^{n} |f(t_k) - f(t_{k-1})| : (t_k)_{k=0}^n \in \mathcal{P}_{[r,s]} \right\} \,. \tag{3.5}$$

If $f \in G(r,s;\mathcal{Y})$ then we define its essential variation by

$$\overline{\mathrm{Var}}(f,[r,s]) := \sup \left\{ \sum_{k=0}^{n} |f(t_k-) - f(t_{k-1}+)| : (t_k)_{k=0}^n \in \mathcal{P}_{[r,s]} \right\} \,. \tag{3.6}$$

Notice that whereas the variation can be defined for any $f : [r,s] \to \mathcal{Y}$ the essential variation does only make sense if $f \in G(r,s;\mathcal{Y})$. Once again if $[r,s] = [0,T]$ we shall just write $\mathrm{Var}(f)$ and $\overline{\mathrm{Var}}(f)$ respectively. A useful property of the variation is the following

Proposition 3.8. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions $f_n : [r,s] \to \mathcal{Y}$ such that $f_n \to f$ with respect to $\|\cdot\|_\infty$. It then holds that*

$$\mathrm{Var}(f,[r,s]) \leq \liminf_{n\to\infty} \mathrm{Var}(f_n,[r,s]) \,. \tag{3.7}$$

Definition 3.9 (Functions of (essentially) bounded variation)**.** We define the space of functions of bounded variation to be

$$\mathrm{BV}(r,s;\mathcal{Y}) = \{f : [r,s] \to \mathcal{Y} : \mathrm{Var}(f,[r,s]) < \infty\} \tag{3.8}$$

and the space of functions of essentially bounded variation as

$$\overline{\mathrm{BV}}(r,s;\mathcal{Y}) = \left\{f \in G(r,s;\mathcal{Y}) : \overline{\mathrm{Var}}(f,[r,s]) < \infty\right\} \,. \tag{3.9}$$

The relation between all the spaces introduced so far is content of the following theorem.

Theorem 3.10. *It holds*

$$S(r,s;\mathcal{Y}) \subset \mathrm{BV}(r,s;\mathcal{Y}) \subset \overline{\mathrm{BV}}(r,s;\mathcal{Y}) \subset G(r,s;\mathcal{Y}) \tag{3.10}$$

where each injection is dense with respect to the sup*-norm.*

The classical way to equip BV with a norm and make it a Banach space is the following.

Proposition 3.11. *For any $f \in \mathrm{BV}(r, s; \mathcal{Y})$ we define*

$$\|f\|_{\mathrm{BV}} = |f(0)| + \mathrm{Var}(f, [r, s]) . \tag{3.11}$$

Then $\| \cdot \|_{\mathrm{BV}}$ is a norm on BV and $(\mathrm{BV}(r, s; \mathcal{Y}), \| \cdot \|_{\mathrm{BV}})$ is a Banach space. Moreover there exists an $C > 0$ such that for any $f \in \mathrm{BV}(r, s; \mathcal{Y})$

$$\|f\|_\infty \leq C\|f\|_{\mathrm{BV}} . \tag{3.12}$$

We are also going to make use of a norm on BV, which is defined in terms of a '*weighted total variation*'.

Definition 3.12 (Weighted total variation)**.** Let $w : [0, T] \to \mathbb{R}_{>0}$ be monotone decreasing. We call

$$\mathrm{Var}_w(y) := \sup \left\{ \sum_{k=1}^{N} \big[|y(t_{k-1}) - y(t_k)| w(t_k) \big] : (t_k)_{k=0}^N \in \mathcal{D} \right\} \tag{3.13}$$

a weighted variation on $[0, T]$ with weight w.

One can generalize this definition to positive functions w which are bounded away from zero. However in our subsequent analysis we only employ this notion with monotone decreasing functions. Hence we restrict ourselves to this case and avoid some technical difficulties in the upcoming proofs.

Proposition 3.13. *Let $w : [0, T] \to \mathbb{R}_{>0}$ be monotone decreasing. Then*

$$|y|_w := |y(0)| + \mathrm{Var}_w(y) \tag{3.14}$$

is a norm on $BV(0, T; X)$ which is equivalent to $| \cdot |_{BV}$, i.e.

$$\min\{1, w(T)\}|y|_{BV} \leq |y|_w \leq \max\{w(0), 1\}|y|_{BV} . \tag{3.15}$$

We call $| \cdot |_w$ a weighted norm with weight w.

We omit the proof; it is straightforward. Let us point out that any space closed with respect to $| \cdot |_{BV}$ is also closed with respect to $| \cdot |_w$. We need the following (lower semi-) continuity results for the dependence of $| \cdot |_w$ on its weight.

Proposition 3.14. *Let $(w_n)_{n \in \mathbb{N}}, w$ be a (sequence of) monotone decreasing functions*

$$w_n, w : [0, T] \to \mathbb{R}_{>0} . \tag{3.16}$$

Assume $w_n \to w$ with respect to $\| \cdot \|_\infty$. Then for all $y \in BV(0, T; \mathcal{Y})$

$$\mathrm{Var}_{w_n}(y) \to \mathrm{Var}_w(y) . \tag{3.17}$$

Furthermore assume $(y_n)_{n \in \mathbb{N}} \subset BV(0, T; \mathcal{Y})$ with $\mathrm{Var}(y_n) \leq C$ for some C independent of n and $y_n \to y$ w.r.t. $\| \cdot \|_\infty$, then

$$\mathrm{Var}_w(y) \leq \liminf_{n \to \infty} \mathrm{Var}_{w_n}(y_n) . \tag{3.18}$$

Proof. Let $y \in BV(0,T;\mathcal{Y})$. Note that

$$\begin{aligned}&\text{Var}(y)\|w_n - w\|_\infty \\ &\geq \sum_{k=1}^{N} \left[|y(t_{k-1}) - y(t_k)|w_n(t_k)\right] - \sum_{k=1}^{N} \left[|y(t_{k-1}) - y(t_k)|w(t_k)\right]\end{aligned}$$

for all $(t_k) \in \mathcal{D}$. Thus

$$\text{Var}(y)\|w_n - w\|_\infty \geq \text{Var}_{w_n}(y) - \text{Var}_w(y) \tag{3.19}$$

and by the same method we derive

$$-\text{Var}(y)\|w_n - w\|_\infty \leq \text{Var}_{w_n}(y) - \text{Var}_w(y)\,. \tag{3.20}$$

Both together imply the first statement via

$$|\text{Var}_{w_n}(y) - \text{Var}_w(y)| \leq \text{Var}(y)\|w_n - w\|_\infty\,. \tag{3.21}$$

For the second claim notice that Var_w is lower semicontinuous with respect to the $\|\cdot\|_\infty$. Therefore we have

$$\liminf_{n\to\infty} \text{Var}_{w_n}(y_n) \geq \liminf_{n\to\infty} \left(\text{Var}_w(y_n) - C\|w_n - w\|_\infty\right) \geq \text{Var}_w(y)\,. \tag{3.22}$$

□

For any of the spaces introduced above we indicate its subspace of left continuous functions, i.e. functions $f : [r,s] \to \mathcal{Y}$ such that $f(t-) = f(t)$ for all $t \in (r,s]$, by the index L. A subspace which will play an important role in the sequel is the space of left continuous functions of bounded variation, whose size of discontinuities is bounded.

Definition 3.15 (BV functions with small jumps)**.** Let $c \geq 0$, $r < s$. Then we denote by

$$\text{BV}_L^c(r,s;\mathcal{Y}) := \{f \in \text{BV}_L(r,s;\mathcal{Y}) : \forall t \in [r,s) : |f(t) - f(t+)| \leq c\} \tag{3.23}$$

the space of all left continuous functions of bounded variation such that the size of every discontinuity is less than c.

If $c = 0$ then any function in BV_L^c is continuous, that is $\text{BV}_L^c = CBV$. As this space is of some importance in the sequel we shall spent some time analyzing it. For convenience we therefore shall set $[r,s] = [0,T]$. However the results we are going to show can be derived for arbitrary $[r,s]$ by a straightforward transfer of arguments.

Proposition 3.16 (Closedness of BV_L^c)**.** *Let* $c \geq 0$ *and* $(u_n)_{n\in\mathbb{N}} \subset \text{BV}_L^c(0,T;\mathcal{Y})$ *with either*

$$u_n \overset{\|\cdot\|_{BV}}{\longrightarrow} u \qquad \textit{or} \tag{3.24}$$

$$\text{Var}(u_n) \leq C \;\textit{ and }\; u_n \overset{\|\cdot\|_\infty}{\longrightarrow} u \tag{3.25}$$

then $u \in \text{BV}_L^c(0,T;\mathcal{Y})$.

Proof. If $u_n \overset{\|\cdot\|_{BV}}{\longrightarrow} u$ then $\mathrm{Var}(u_n)$ is bounded and $u_n \overset{\|\cdot\|_{\infty}}{\longrightarrow} u$. Thus it suffices to prove the second assertion. First of all $u \in \mathrm{BV}(0,T;\mathcal{Y})$ since

$$\mathrm{Var}(u) \leq \liminf_{n\to\infty} \mathrm{Var}(u_n) \leq C\,. \tag{3.26}$$

Now choose $t \in (0,T]$ arbitrary. For each $\varepsilon > 0$ there exists some $\delta_n > 0$ such that

$$\forall s \leq t : |t-s| \leq \delta_n : \quad |u_n(t) - u_n(s)| \leq \frac{\varepsilon}{3}\,. \tag{3.27}$$

Furthermore choose $N \in \mathbb{N}$ large enough such that

$$\|u_n - u\|_\infty \leq \frac{\varepsilon}{3} \ \ \forall n \geq N\,. \tag{3.28}$$

Thus for all $s < t$ with $|t-s| \leq \delta_N$ we have

$$|u(t) - u(s)| \leq \varepsilon \tag{3.29}$$

and we infer that u is left continuous. It remains to show that the jump size is bounded by c. Since $u, u_n \in BV_L$ we know that u_n and u admit a right hand side limits at every point. For $t \in [0,T)$ therefore there exist $\delta, \delta_n > 0$ such that

$$\begin{aligned} |u_n(t+) - u_n(s)| &\leq \frac{\varepsilon}{3} \ \forall s > t : |t-s| \leq \delta_n \text{ and} \\ |u(t+) - u(s)| &\leq \frac{\varepsilon}{3} \ \forall s > t : |t-s| \leq \delta\,. \end{aligned}$$

Choose N as above and set $\delta_n' := \min\{\delta_n, \delta\}$. Then for all $n \geq N$ and $s > t$ with $|s-t| \leq \delta_n'$ we have

$$|u(t+) - u_n(t+)| \leq |u(t+) - u(s)| + |u_n(t+) - u_n(s)| + |u_n(s) - u(s)| \leq \varepsilon\,. \tag{3.30}$$

Therefore $u_n(t+) \to u(t+)$ for all $t \in [0,T]$ and we obtain

$$|u(t) - u(t+)| \leq \liminf_{n\to\infty} |u_n(t) - u_n(t+)| \leq c \tag{3.31}$$

to complete the proof. □

Lemma 3.17. *For any $u \in BV_L^c(0,T;\mathcal{Y})$ and any $\varepsilon > 0$ there exists a partition $(t_n)_{n=0}^N \in \mathcal{P}$ and values $(\widehat{u}_n)_{n=0}^N \subset \mathcal{Y}$ such that for*

$$\widehat{u}(t) = \widehat{u}_0 \chi_{\{0\}}(t) + \sum_{n=1}^N \widehat{u}_n \chi_{(t_{n-1},t_n]}(t) \tag{3.32}$$

the following conditions are satisfied:

$$\|\widehat{u} - u\|_\infty \leq \varepsilon\,, \ \ \mathrm{Var}(\widehat{u}) \leq \mathrm{Var}(u) \ \ \textit{and} \tag{3.33}$$

$$\forall t \in [0,T] : \ \mathrm{Var}(u,[0,t]) + \varepsilon \geq \mathrm{Var}(\widehat{u},[0,t]) \geq \mathrm{Var}(u,[0,t]) - 2\varepsilon\,. \tag{3.34}$$

Proof. The idea of the proof is based on a classical proof by Aumann [5, p. 257f]. It relies on two facts: There exists a sequence $(s_n)_{n=0}^{N_1} \in \mathcal{P}$ such that

$$\sum_{n=1}^{N_1} |u(s_n) - u(s_{n-1})| \geq \mathrm{Var}(u) - \varepsilon \tag{3.35}$$

and for every $u \in BV_L(0,T;\mathcal{Y})$ and for every $\varepsilon > 0$ there exists a partition $(r_n)_{n=0}^{N_2} \in \mathcal{P}$, with $r_0 = 0$ and $r_{N_2} = T$ such that

$$\forall n \in [N_2] : \forall s,t \in (r_{n-1}, r_n] : |u(s) - u(t)| \leq \varepsilon\,. \tag{3.36}$$

The first fact is plainly due to the definition of the variation whereas the second fact can be derived from Lemma 3.6. The choice of the half-open interval $(r_{n-1}, r_n]$ is due to the left-continuity of u.
Let $(t_n)_{n=1}^{N} \in \mathcal{P}$ such that $\{t_n : n \in [N]\} \supset \{s_n : n \in [N_1]\} \cup \{r_n : n \in [N_2]\}$ and define

$$\widehat{u}(t) := u(0)\chi_{\{0\}}(t) + \sum_{n=1}^{N} \left(u(t_{n-1}+)\chi_{\left(t_{n-1}, \frac{t_{n-1}+t_n}{2}\right]}(t) + u(t_n)\chi_{\left(\frac{t_{n-1}+t_n}{2}, t_n\right]}(t) \right).$$

It is easy to see that $\widehat{u} \in BV_L^{c/\delta}(0,T;X)$ and $\|\widehat{u} - u\|_\infty \leq \varepsilon$. For each $n \in [N]$ choose a sequence $\tau_n^k \downarrow t_{n-1}$ as $k \to \infty$ with $t_{n-1} < \tau_n^k \leq t_n$. Then

$$\sum_{n=0}^{N} \left(|u(t_{n-1}) - u(\tau_n^k)| + |u(\tau_n^k) - u(t_n)| \right) \leq \mathrm{Var}(u) \tag{3.37}$$

and the lower semicontinuity of the norm gives (3.33). For $t = 0$ and $t = T$ (3.34) is straightforward; for the latter use the choice of s_n. For $t \in (0,T)$ there exists some $n \in [N]$ such that $t \in (t_{n-1}, t_n]$ and we have

$$\mathrm{Var}(\widehat{u}, [0,t]) = \sum_{k=1}^{n} |u(t_{k-1}+) - u(t_{k-1})| + |u(t_{k-1}+) - u(t_k)| \leq \mathrm{Var}(u, [0,s]) + \varepsilon$$

for any $s \in (t_{n-1}, t]$. Here we use $|u(t_{k-1}+) - u(t_k)| \leq \varepsilon$ and a limit argument similar as for (3.37). It remains to proof the second inequality. Once again let $t \in (0,T)$ and choose $n \in [N]$ such that $t \in (t_{n-1}, t_n]$. For $t = t_n$ we get

$$\mathrm{Var}(\widehat{u}, [0,t_n]) \geq \mathrm{Var}(u, [0,T]) - \varepsilon - \mathrm{Var}(\widehat{u}, [t_n, T])\,. \tag{3.38}$$

Using $\mathrm{Var}(u,[0,T]) = \mathrm{Var}(u,[0,t]) + \mathrm{Var}(u,[t,T])$ and the definition of $\widehat{u}$ we see that $\mathrm{Var}(\widehat{u},[t_n,T]) \leq \mathrm{Var}(u,[t_n,T])$. Therefore

$$\mathrm{Var}(\widehat{u}, [0,t_n]) \geq \mathrm{Var}(u, [0,t_n]) - \varepsilon\,. \tag{3.39}$$

For $t \in (t_{n-1}, t_n)$ use the above equation,

$$\mathrm{Var}(\widehat{u}, [0,t]) \geq \mathrm{Var}(\widehat{u}, [0,t_k]) - \varepsilon \;\text{ and }\; \mathrm{Var}(u, [0,t]) \leq \mathrm{Var}(u, [0,t_k]) \tag{3.40}$$

to obtain the desired inequality. □

Remark 3.18. As a direct consequence of (3.34) we have that for any $[r,s] \subset [0,T]$

$$\text{Var}(u,[r,s]) + 3\varepsilon \geq \text{Var}(\widehat{u},[r,s]) \geq \text{Var}(u,[r,s]) - 3\varepsilon\,. \tag{3.41}$$

Corollary 3.19 (Approximation with step functions)**.** *For any $c \geq 0$ and $u \in BV_L^c(0,T;\mathcal{Y})$ there exists either a sequence*

$$(u_n)_{n\in\mathbb{N}} \subset S(0,T;\mathcal{Y}) \cap BV_L^c(0,T;\mathcal{Y}) \quad \text{if } c > 0 \quad \text{or} \tag{3.42}$$

$$(u_n)_{n\in\mathbb{N}} \subset S(0,T;\mathcal{Y}) \cap BV_L^\delta(0,T;\mathcal{Y}) \quad \text{if } c = 0 \tag{3.43}$$

with $\delta > 0$ arbitrary such that

$$\text{Var}(u_n) \leq \text{Var}(u) \;\; \text{and} \;\; u_n \overset{\|\cdot\|_\infty}{\to} u\,. \tag{3.44}$$

In both cases $(u_n)_{n\in\mathbb{N}}$ can be chosen such that $\text{Var}(u_n,[0,t])\to\text{Var}(u,[0,t])$ *uniformly in* $[0,T]$.

Proof. For $u \in BV_L^c(0,T;X)$ and $n \in \mathbb{N}$ construct u_n as in Lemma 3.17 with $\varepsilon = \min\{1/n, c\}$ if $c > 0$ and $\varepsilon = \min\{1/n, \delta\}$ if $c = 0$. ☐

3.2. Kurzweil-Henstock Integral

The Kurzweil-Henstock integral (or Henstock / Henstock-Kurzweil / Kurzweil integral) is a Riemann(-Stieltjes) type of integral which allows the integration of non necessarily continuous functions with respect to a BV function. It has been introduced independently by the Czech mathematician Jarsolav Kurzweil in [65] and shortly afterwards by the English mathematician Ralph Henstock in [42]. Whereas for the former it was a means to an end, namely the study of parameter dependent ordinary differential equations, the latter developed a thorough theory for this notion of integration, see e.g. [43, 44, 45]. In so far as we do not show new results, our presentation here is based on the article [59] by Pavel Krejčí.

Before we can formally define the Kurzweil-Henstock integral, we shortly need to introduce some concepts and notation. First let $-\infty < a < b < +\infty$. Then we set

$$\Gamma(a,b) := \{\delta : [a,b] \to \mathbb{R} : \delta(t) > 0 \;\forall t \in [a,b]\} \tag{3.45}$$

the set of gauge functions. Furthermore for a gauge $\delta \in \Gamma(a,b)$ and $t \in [a,b]$ we write

$$I_\delta(t) := (t - \delta(t), t + \delta(t))\;. \tag{3.46}$$

Using this notions we shall define the set of δ-fine partitions. Note that there is a little ambiguity with the concept of partitions as a δ-fine partition indeed is not a partition, but in fact based on two partitions.

Definition 3.20 (δ-fine partitions)**.** Let $(t_i)_{i=0}^n \in \mathcal{P}_{[a,b]}$ be a partition and $\delta \in \Gamma(a,b)$ a gauge. We say that the family

$$D = (\tau_i, [t_{i-1}, t_i])_{i=1}^n \tag{3.47}$$

is a δ-fine partition if

$$\forall i \in [n]: \quad \tau_i \in [t_{i-1}, t_i] \text{ and} \forall i \in [n]: \quad [t_{i-1}, t_i] \in I_\delta \tau_i\,. \tag{3.48}$$

The set of all δ-fine partitions D is denoted by $\mathcal{F}(a,b,\delta)$.

For a given δ-fine partition $D = (\tau_i, [t_{i-1}, t_i])_{i=1}^n$ we define the following shorthand for the 'Kurzweil sum'

$$K_D(f, g) := \sum_{i=1}^{n} \langle f(\tau_i), g(t_i) - g(t_{i-1})\rangle_{\mathcal{Y}^*,\mathcal{Y}} \,. \tag{3.49}$$

We are now able to introduce the Kurzweil-Henstock integral.

Definition 3.21 (Kurzweil-Henstock integral)**.** Let $f : [a, b] \to \mathcal{Y}^*$ and $g : [a, b] \to \mathcal{Y}$ be given. A value $J \in \mathbb{R}$ is said to be the Kurzweil-Henstock integral of f with respect to g over $[a, b]$, in formula

$$J = (\mathrm{K}) \int_a^b \langle f(t), dg(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} \tag{3.50}$$

if for every $\varepsilon > 0$ there exits a gauge $\delta \in \Gamma(a, b)$ such that for any δ-fine partitions $D \in \mathcal{F}(a, b, \delta)$ it holds

$$|K_D(f, g) - J| \le \varepsilon \,. \tag{3.51}$$

Notice that due to Cousin's theorem ([22, p. 22], [46, p. 425]) for any gauge δ there exists a δ-fine partition. Hence the last condition indeed is sensible.

The following existence result can be found in [61, Theorem 1.9].

Theorem 3.22. *Let either $f \in G(a, b, \mathcal{Y}^*)$ and $g \in BV(a, b; \mathcal{Y})$ or $f \in BV(a, b, \mathcal{Y}^*)$ and $g \in G(a, b; \mathcal{Y})$ then the Kurzweil-Henstock integral*

$$(\mathrm{K}) \int_a^b \langle f(t), dg(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$$

exists and the satisfies the inequality

$$(\mathrm{K}) \int_a^b \langle f(t), dg(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} \le \min \left\{ \|f\|_\infty \mathrm{Var}(g, [a, b]), (|f(a)|_{\mathcal{Y}} + |f(b)|_{\mathcal{Y}} + \mathrm{Var}(f, [a, b]))\, |g|_\infty \right\} \,.$$

Unlike for the Riemann-Stieltjes integral the Kurzweil-Henstock integral allows for the integration of non-continuous functions with respect to a BV-function. This is a key feature needed in our subsequent study of sweeping processes and therefore is the reason why we use this notion of integration.
Until we say otherwise we shall denote the Kurzweil-Henstock integral by $\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$. However we point out that later on in this section we have to go back the original notation as we introduce a further notion of integration. The Kurzweil-Henstock integral enjoys the usual linearity properties [61, Proposition 1.3 & 1.4].

Proposition 3.23. *Let $f, f_1, f_2 : [a, b] \to \mathcal{Y}^*$ and $g, g_1, g_2 : [a, b] \to \mathcal{Y}$. Furthermore let $s \in (a, b)$ and $\lambda > 0$. Then the following assertions hold:*

(i) *If* $\int_a^b \langle f_1(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *and* $\int_a^b \langle f_2(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *exist then also* $\int_a^b \langle f_1(t)+f_2(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *exists and it holds*

$$\int\limits_a^b \langle f_1(t)+f_2(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \int\limits_a^b \langle f_1(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} + \int\limits_a^b \langle f_2(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} \tag{3.52}$$

(ii) *If* $\int_a^b \langle f(t),\, \mathrm{d}g_1(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *and* $\int_a^b \langle f(t),\, \mathrm{d}g_2(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *exist, then* $\int_a^b \langle f(t),\, \mathrm{d}\left(g_1+g_2\right)(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *exists as well and it holds*

$$\int\limits_a^b \langle f(t),\, \mathrm{d}\left(g_1+g_2\right)(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \int\limits_a^b \langle f(t),\, \mathrm{d}g_1(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} + \int\limits_a^b \langle f(t),\, \mathrm{d}g_2(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} \tag{3.53}$$

(iii) *If* $\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *exists, then so do* $\int_a^b \langle \lambda f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *and* $\int_a^b \langle f(t),\, \mathrm{d}\lambda g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$. *Furthermore it holds*

$$\lambda \int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \int_a^b \langle \lambda f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \int_a^b \langle f(t),\, \mathrm{d}\lambda g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} \tag{3.54}$$

(iv) *If* $\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *exists, then* $\int_a^s \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *and* $\int_s^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *do as well.*

(v) *If* $\int_a^s \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *and* $\int_s^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *exist, then also* $\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *exists. It holds*

$$\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \int_a^s \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} + \int_s^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} \tag{3.55}$$

In order to obtain some consistency in the extremal case $s = a$ and $s = b$ we set

$$\int_s^s \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} = 0 \tag{3.56}$$

for all $f : [a,b] \to \mathcal{Y}^*$ and $g : [a,b] \to \mathcal{Y}$. Furthermore the Kurzweil-Henstock integral has following continuity property [61, Proposition 1.11].

Proposition 3.24. *Let* $(f_n)_{n\in\mathbb{N}} \subset G(a,b;\mathcal{Y}^*)$, $f \in G(a,b;\mathcal{Y}^*)$, $(g_n)_{n\in\mathbb{N}} \subset BV(a,b;\mathcal{Y})$ *and* $g \in BV(a,b;\mathcal{Y})$. *If*

$$\|f_n - f\|_\infty \to 0 \;\wedge\; [\exists C > 0 : \forall n \in \mathbb{N} : \mathrm{Var}(g_n) \le C] \;\wedge\; \|g_n - g\|_\infty \to 0 \tag{3.57}$$

then

$$\lim_{n\to\infty} \int\limits_a^b \langle f_n,\, \mathrm{d}g_n(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \int\limits_a^b \langle f,\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\;. \tag{3.58}$$

Furthermore the following rules (see [61, Proposition 1.5]) help calculating the Kurzweil-Henstock integral in some concrete cases.

Proposition 3.25. *For every $g : [a,b] \to \mathcal{Y}$, $a \le r \le s \le b$ and $y' \in \mathcal{Y}^*$ it holds*

$$\int_a^b \left\langle y'\chi_{\{s\}}(t),\, \mathrm{d}g(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \langle v, g\rangle_{\mathcal{Y}^*,\mathcal{Y}}(s+) - \langle v, g\rangle_{\mathcal{Y}^*,\mathcal{Y}}(s-) \text{ and} \tag{3.59}$$

$$\int_a^b \left\langle y'\chi_{(r,s)}(t),\, \mathrm{d}g(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \langle v, g\rangle_{\mathcal{Y}^*,\mathcal{Y}}(s-) - \langle v, g\rangle_{\mathcal{Y}^*,\mathcal{Y}}(r+) \tag{3.60}$$

provided the limits on the right-hand sides exist, using the convention that

$$\langle v, g\rangle_{\mathcal{Y}^*,\mathcal{Y}}(a-) = \langle v, g\rangle_{\mathcal{Y}^*,\mathcal{Y}}(a) \text{ and } \langle v, g\rangle_{\mathcal{Y}^*,\mathcal{Y}}(b+) = \langle v, g\rangle_{\mathcal{Y}^*,\mathcal{Y}}(b)\,.$$

Corollary 3.26. *Let $f \in S(a,b;\mathcal{Y}^*)$ be a step function and choose $(t_n)_{n=0}^N \in \mathcal{P}_{[a,b]}$ such that*

$$f(t) = \sum_{n=1}^N f_n\chi_{(t_{n-1},t_n)}(t) + \sum_{n=0}^N \widehat{f}_n\chi\{t_n\}(t)\,. \tag{3.61}$$

Furthermore let $g \in G(a,b;\mathcal{Y})$. Then one can evaluate the Kurzweil-Henstock integral as

$$\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \tag{3.62}$$

$$\sum_{n=1}^N \langle f_n, g(t_n-) - g(t_{n-1}+)\rangle_{\mathcal{Y}^*,\mathcal{Y}} + \sum_{n=0}^N \left\langle \widehat{f}_n, g(t_n+) - g(t_n-)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,.$$

Proof. This is a consequence of coupling Proposition 3.23 (i) with Proposition 3.25. □

The following decomposition result will be of some interest in the sequel. It has been proven by the author in [88, Appendix B].

Lemma 3.27. *Let $u \in G(0,T;\mathcal{Y}^*)$, $\xi \in BV(0,T;\mathcal{Y})$ and $[r,s] \subset [0,T]$ with $r < s$. Define*

$$\widehat{\xi} : [r,s] \to \mathcal{Y},\ t \mapsto \begin{cases} \xi(r+) & \text{if } t = r \\ \xi(t) & \text{if } t \in (r,s) \\ \xi(s-) & \text{if } t = s\,. \end{cases} \tag{3.63}$$

Then

$$\int_0^T \left\langle u(t)\chi_{(r,s)}(t), d\xi(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \int_r^s \left\langle u(t), d\widehat{\xi}(t)\right\rangle. \tag{3.64}$$

Proof. It suffices to prove the above for step functions $u \in S(0,T;\mathcal{Y}^*)$. In that case there exists a partition $(t_n)_{n=0}^N \in \mathcal{D}_{[0,T]}$, $t_0 = r$, $t_N = s$ such that

$$u(t)\chi_{(r,s)}(t) = \sum_{n\in[N]} u^n\chi_{(t_{n-1},t_n)}(t) + \sum_{n\in[N-1]} \overline{u}^n\chi_{\{t_n\}}(t) \tag{3.65}$$

for some $(u^n)_{n=1}^N, (\overline{u}^n)_{n=1}^{N-1} \subset \mathcal{Y}^*$. Then we have for the left hand side integral

$$\begin{aligned} &\int_0^T \left\langle u(t)\chi_{(r,s)}(t), d\xi(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} \qquad (3.66)\\ &= \sum_{n\in[N]} \langle u^n, \xi(t_n-) - \xi(t_{n-1}+)\rangle_{\mathcal{Y}^*,\mathcal{Y}} + \sum_{n\in[N-1]} \langle \overline{u}^n, \xi(t_n+) - \xi(t_n-)\rangle_{\mathcal{Y}^*,\mathcal{Y}} . \end{aligned}$$

For the right hand side we get

$$\begin{aligned} &\int_r^s \left\langle u(t), d\widehat{\xi}(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} \qquad (3.67)\\ &= \sum_{n\in[N]} \left\langle u^n, \widehat{\xi}(t_n-) - \widehat{\xi}(t_{n-1}+)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} + \sum_{n\in[N-1]} \left\langle \overline{u}^n, \widehat{\xi}(t_n+) - \widehat{\xi}(t_n-)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}\\ &\quad + \left\langle u(r), \widehat{\xi}(r+) - \widehat{\xi}(r)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} + \left\langle u(s), \widehat{\xi}(s) - \widehat{\xi}(s-)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} . \end{aligned}$$

Due to the definition of $\widehat{\xi}$ the last two terms are zero and the former sums agree with the ones above. Thus for all $u \in S(0,T;\mathcal{Y})$ equation (3.64) holds. Now let u be in $G(0,T;\mathcal{Y})$. Then there exists a sequence of step functions $(u_n)_{n\in\mathbb{N}}$ such that $u_n \to u$ w.r.t. $\|\cdot\|_\infty$ and the continuity of the Kurzweil integral grants the statement. □

Proposition 3.28. *Let $f : [a,b] \to \mathcal{Y}^*$ and $y \in \mathcal{Y}$. For any $s \in [a,b]$ it holds*

$$\int \left\langle f(t), \mathrm{d}\left(y\chi_{\{s\}}(t)\right)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \begin{cases} 0 & \text{if } s \in (a,b)\\ -\langle f(a), y\rangle_{\mathcal{Y}^*,\mathcal{Y}} & \text{if } s = a\\ \langle f(b), y\rangle_{\mathcal{Y}^*,\mathcal{Y}} & \text{if } s = b\,. \end{cases} \qquad (3.68)$$

Moreover for $a \le r < s \le b$ it holds

$$\int \left\langle f(t), \mathrm{d}\left(y\chi_{(r,s)}(t)\right)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \langle f(s) - f(r), y\rangle_{\mathcal{Y}^*,\mathcal{Y}} . \qquad (3.69)$$

The Kurzweil-Henstock integral in its original form however suffers one essential shortcoming: It has been shown by J. Kurzweil and P. Krejčí in [56] that it is not possible to integrate any regulated function with respect to any function of bounded variation. This however is a property which is indeed needed for our study. It for example is satisfied by the Young integral. To overcome this problem for the Kurzweil-Henstock integral P. Krejčí proposed in [59] a modification, which we shall now present.

Definition 3.29 (Negligible sets)**.** Let $-\infty < a < b < \infty$. $\mathcal{N} \subset 2^{[a,b]}$ is called a system of negligible sets of the interval $[a,b]$ if

$$\forall A \in \mathcal{N} : \quad \overline{[a,b] \setminus A} = [a,b] \qquad (3.70)$$

$$\forall A, B \in \mathcal{N} : \quad A \cup B \in \mathcal{N} . \qquad (3.71)$$

Examples for systems of negligible sets are the set of all Lebesgue measurable subsets of $[a,b]$ with measure zero or the set of all countable subsets of $[a,b]$. With the notion of negligible sets we may now introduce (δ, A)-fine partitions.

Definition 3.30 (((δ, A)-fine partitions). Let $\mathcal{N}$ be a system of negligible sets and $A \in \mathcal{N}$. A delta fine partition D is called (δ, A)-fine, in formula $D \in \mathcal{F}(a, b, \delta, A)$, if

$$\left.\begin{array}{c} \forall j \in [n-1] \setminus \{1\} : \tau_j \in (t_{j-1}, t_j) \\ \tau_j = t_{j-1} \Rightarrow j = 1 \ \wedge \ \tau_j = t_j \Rightarrow j = n \end{array}\right\} \tag{3.72}$$

$$\forall j \in [n-1] : t_j \in [a, b] \setminus A \,. \tag{3.73}$$

Now we have assembled all necessary preparations to introduce the extension of the Kurzweil-Henstock integral.

Definition 3.31 (Kurzweil-Henstock integral with exclusion of negligible sets). Let $\mathcal{N}$ be a system of negligible sets. Let furthermore $f : [a, b] \to \mathcal{Y}$ and $g : [a, b] \to \mathcal{Y}^*$ be given. A value $J \in \mathbb{R}$ is said to be the Kurzweil-Henstock integral with exclusion of negligible sets, or as a shorthand (KN)-integral of f with respect to g over $[a, b]$, in formula

$$J = (\mathrm{KN}) \int_a^b \langle f(t), dg(t) \rangle_{\mathcal{Y}^*, \mathcal{Y}} \tag{3.74}$$

if for every $\varepsilon > 0$ there exits a gauge $\delta \in \Gamma(a, b)$ and $A \in \mathcal{N}$ such that for any (δ, A)-fine partitions $D \in \mathcal{F}(a, b, \delta, A)$ it holds

$$|K_D(f, g) - J| \leq \varepsilon \,. \tag{3.75}$$

The Kurzweil-Henstock integral with exclusion of negligible sets indeed is a generalization of the usual Kurzweil-Henstock integral, as the following simple proposition shows.

Proposition 3.32. *Let $\mathcal{N}$ be a system of negligible sets and $f : [a, b] \to \mathcal{Y}$ and $g : [a, b] \to \mathcal{Y}^*$ be given. Assume that the Kurzweil-Henstock integral*

$$(\mathrm{K}) \int_a^b \langle f(t), dg(t) \rangle_{\mathcal{Y}^*, \mathcal{Y}}$$

exists. Then also the (KN)*-integral of f with respect to g exists and*

$$(\mathrm{KN}) \int_a^b \langle f(t), dg(t) \rangle_{\mathcal{Y}^*, \mathcal{Y}} = (\mathrm{K}) \int_a^b \langle f(t), dg(t) \rangle_{\mathcal{Y}^*, \mathcal{Y}} \,. \tag{3.76}$$

Proof. For any $A \in \mathcal{N}$ and $\delta \in \Gamma(a, b)$ the inclusion $\mathcal{F}(a, b, \delta, A) \subset \mathcal{F}(a, b, \delta)$ holds. By definition of the Kurzweil integral for any $\varepsilon > 0$ there exists an $\delta \in \Gamma(a, b)$ such that for any $D \in \mathcal{F}(a, b, \delta)$ we have

$$\left| K_D(f, g) - (\mathrm{K}) \int_a^b \langle f(t), dg(t) \rangle_{\mathcal{Y}^*, \mathcal{Y}} \right| \leq \varepsilon \,. \tag{3.77}$$

Consequently this inequality also holds for any $D \in \mathcal{F}(a, b, \delta, A)$ and the proof is complete. □

Due to the above proposition all results which were established for the Kurzweil-Henstock integral also hold for the (KN)-integral. Throughout this thesis we will always assume that the system of negligible sets $\mathcal{N}$ is equal to the set of all countable subsets of $[a,b]$. The following existence results justifies the introduction of this new concept of integration.

Theorem 3.33. *Let $f \in G(a,b;\mathcal{Y}^*)$ and $g \in \overline{\mathrm{BV}}(a,b,\mathcal{Y})$. Then*

$$(\mathrm{KN})\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$$

exists and it holds

$$\left|(\mathrm{KN})\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\right| \le \|f\|_\infty \overline{\mathrm{Var}}(g,[a,b])\,. \tag{3.78}$$

This result has been proven in [58, Corollary 2.6] in the special case of $\mathcal{Y}$ being a Hilbert space. The proof in the present case works essentially in the same way. For the sake of completeness we will present it below. To this end we introduce the following Lemma which contains the essential ingredient of our proof.

Lemma 3.34. *Let $(f_n), f : [a,b] \to \mathcal{Y}^*$ such that $\|f_n - f\|_\infty \to 0$ and $g \in \overline{\mathrm{BV}}(a,b;\mathcal{Y})$. If for all $n \in \mathbb{N}$ the integral $(\mathrm{KN})\int_a^b \langle f_n(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ exists then $(\mathrm{KN})\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ exists and it holds*

$$(\mathrm{KN})\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \lim_{n\to\infty}(\mathrm{KN})\int_a^b \langle f_n(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,. \tag{3.79}$$

A one-dimensional version of this Lemma has been proven in [58, Theorem 2.5].

Proof. Define

$$J_n := (\mathrm{KN})\int_a^b \langle f_n(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} \quad \forall n \in \mathbb{N}\,. \tag{3.80}$$

Our first aim is to show that $(J_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Let A be the set of all points of discontinuity of g. Since $g \in \overline{\mathrm{BV}}(a,b,\mathcal{Y})$, A is countable and hence contained in $\mathcal{N}$. Now for any $n \in \mathbb{N}$ there exists a $\delta_n \in \Gamma(a,b)$ and $A \subset A_n \in \mathcal{N}$ such that for all $D \in \mathcal{F}(a,b,\delta_n,A_n)$

$$|K_D(f_n,g) - J_n| \le \frac{1}{n}\,. \tag{3.81}$$

For any $m,n \in \mathbb{N}$ set $\delta_{m,n}(t) := \min\{\delta_m(t),\delta_n(t)\}$. Then $\delta_{m,n} \in \Gamma(a,b)$. Furthermore set $A_{m,n} := A_m \cup A_n \in \mathcal{N}$. Then any $\mathcal{F}(a,b,\delta_{m,n},A_{m,n}) \subset \mathcal{F}(a,b,\delta_n,A_n) \cap \mathcal{F}(a,b,\delta_m,A_m)$. Hence for all $D \in \mathcal{F}(a,b,\delta_{m,n},A_{m,n})$ it holds

$$|K_D(f_n,g) - J_n| \le \frac{1}{n} \quad \text{and} \quad |K_D(f_m,g) - J_m| \le \frac{1}{m}\,. \tag{3.82}$$

We then can estimate

$$\begin{aligned} |J_n - J_m| &\le |J_n - K_D(f_n,g)| + |J_m - K_D(f_m,g)| + |K_D(f_n - f_m,g)| \\ &\le \frac{1}{n} + \frac{1}{m} + \|f_n - f_m\|_\infty \overline{\mathrm{Var}}(g,[a,b])\,. \end{aligned}$$

This implies that $(J_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. We set $J := \lim_{n\to\infty} J_n$. Moreover we now can estimate for any $D \in \mathcal{F}(a,b,\delta_n,A_n)$ that

$$\begin{aligned}|K_D(f,g) - J| &\leq |K_D(f-f_n,g)| + |K_D(f_n,g) - J_n| + |J_n - J| \\ &\leq \|f_n - f\|_\infty \overline{\mathrm{Var}}(g,[a,b]) + \frac{1}{n} + |J_n - J|\,.\end{aligned}$$

Therefore for any $\varepsilon > 0$ we may choose n such that $\|f_n - f\|_\infty \overline{\mathrm{Var}}(g,[a,b]) \leq \frac{\varepsilon}{3}$, $\frac{1}{n} \leq \frac{\varepsilon}{3}$ and $|J_n - J| \leq \frac{\varepsilon}{3}$. Setting $\widehat{\delta} := \delta_n$ and $\widehat{A} := A_n$ we then obtain that for all $D \in \mathcal{F}(a,b,\widehat{\delta},\widehat{A})$ the estimate $|K_D(f,g) - J| \leq \varepsilon$ holds. Therefore $J = (\mathrm{KN}) \int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$. □

Proof of Theorem 3.33. Using Theorem 3.10 we approximate the function f uniformly with step functions $f_n \in S(a,b;\mathcal{Y}^*)$. Then due to Corollary 3.26 the integral $(\mathrm{K}) \int_a^b \langle f_n(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ exists and by virtue of Proposition so does 3.32 $(\mathrm{KN}) \int_a^b \langle f_n(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$. Applying Lemma 3.34 we then know that $(\mathrm{KN}) \int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ exists. Due to formula (3.62) we can estimate that

$$\left|(\mathrm{KN}) \int_a^b \langle f_n(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\right| \leq \|f_n\|_\infty \overline{\mathrm{Var}}(g,[a,b])\,. \tag{3.83}$$

We then obtain (3.78) by applying the convergence result (3.79). □

Lemma 3.35. *Let* $f \in \mathrm{BV}(a,b;\mathcal{Y}^*)$ *and* $g, g_n : [a,b] \to \mathcal{Y}$ *such that* $\lim \|g_n - g\|_\infty = 0$. *If the integral* $(\mathrm{KN}) \int_a^b \langle f(t),\, \mathrm{d}g_n(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *exists for all* $n \in \mathbb{N}$, *then* $(\mathrm{KN}) \int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ *exists as well and it holds*

$$(\mathrm{KN}) \int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \lim_{n\to\infty} \int_a^b \langle f(t),\, \mathrm{d}g_n(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,. \tag{3.84}$$

For real valued functions this result has been shown in the second part of [58, Theorem 2.5].

Proof. The proof is similar to the proof of Lemma 3.34. First set

$$J_n := (\mathrm{KN}) \int_a^b \langle f(t),\, \mathrm{d}g_n(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,. \tag{3.85}$$

Then for each $n \in \mathbb{N}$ there exists a gauge $\delta_n \in \Gamma(a,b)$ and a negligible set $A_n \in \mathbb{N}$ such that for any $D \in \mathcal{F}(a,b,\delta_n,A_n)$

$$|K_D(f,g_n) - J_n| < \frac{1}{n}\,.$$

For $m, n \in \mathbb{N}$ choose $\delta_{m,n}$ and $A_{m,n}$ as above. Then for any $D \in \mathcal{F}(a,b,\delta_{m,n},A_{m,n})$ it holds

$$\begin{aligned}|J_n - J_m| &\leq |J_n - K_D(f,g_n)| + |J_m - K_D(f,g_m)| + |K_D(f,g_n - g_m)| \\ &\leq \frac{1}{n} + \frac{1}{m} + (|f(a)| + |f(b)| + \mathrm{Var}(f,[a,b]))\, \|g_n - g_m\|_\infty\,.\end{aligned}$$

Thus the sequence $(J_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and we shall denote its limit by J. We then can estimate that for any $D \in \mathcal{F}(a,b,\delta_n,A_n)$

$$\begin{aligned}|K_D(f,g)-J| &\leq |K_D(f,g-g_n)|+|K_D(f,g_n)-J_n|+|J_n-J|\\ &\leq (|f(a)|+|f(b)|+\mathrm{Var}(f,[a,b]))\,\|g_n-g\|_\infty+\frac{1}{n}+|J_n-J|\end{aligned}$$

and we deduce the desired result arguing as in the last part of Lemma 3.34. □

The following theorem is a generalization of Proposition 3.24 to the (KN) integral. A version for real valued function is to be found in [58, Proposition 2.10].

Theorem 3.36. *Let $f_n, f \in G(a,b;\mathcal{Y}^*)$ and $g_n, g \in \overline{\mathrm{BV}}(a,b;\mathcal{Y})$. If*

$$\|f_n-f\|_\infty \to 0 \;\wedge\; [\exists C>0 : \forall n\in\mathbb{N} : \overline{\mathrm{Var}}(g_n,[a,b])\leq C] \;\wedge\; \|g_n-g\|_\infty \to 0$$

then

$$(\mathrm{KN})\int_a^b \langle f_n(t),\, \mathrm{d}g_n(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} \to (\mathrm{KN})\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,.$$

Proof. Choose any $w \in S(a,b;\mathcal{Y}^*)$. Then

$$\begin{aligned}&\left|(\mathrm{KN})\int_a^b \langle f_n(t),\, \mathrm{d}g_n(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} - (\mathrm{KN})\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\right|\\ &\leq \left|(\mathrm{KN})\int_a^b \langle f_n(t)-f(t),\, \mathrm{d}g_n(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\right| + \left|(\mathrm{KN})\int_a^b \langle f(t)-w(t),\, \mathrm{d}(g-g_n)(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\right|\\ &\quad + \left|(\mathrm{KN})\int_a^b \langle w(t),\, \mathrm{d}(g-g_n)(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\right|.\end{aligned}$$

Due to (3.78) we can estimate the first two terms by $C\|f_n-f\|_\infty$ and $2C\|f-w\|_\infty$ respectively. The third term can be estimated by using both Proposition 3.32 and Theorem 3.22 by $(2\|w\|_\infty+\mathrm{Var}(w,[a,b]))\,\|g_n-g\|_\infty$. Letting n go to infinity we deduce that for any $w \in S(a,b;\mathcal{Y})$ it holds

$$\lim_{n\to\infty}\left|(\mathrm{KN})\int_a^b \langle f_n(t),\, \mathrm{d}g_n(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} - (\mathrm{KN})\int_a^b \langle f(t),\, \mathrm{d}g(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\right| \leq 2C\|w-f\|_\infty\,. \quad (3.86)$$

Thus by applying Theorem 3.10 we conclude the proof. □

4. Measure theory for Banach spaces

As a standard notation for a topological (Ω, τ) we use $\mathcal{B}(\Omega)$ to denote the Borel sets of Ω w.r.t. to the topology τ.

4.1. Vector measures

A rather straightforward generalization of the concept of signed measures are measures with values in $\mathbb{R}^n$: Just take a vector of n signed measures. To define a measure on a Banach space $\mathcal{Y}$ is however more involved. We introduce measures with values in Banach spaces and present some of their properties. We aim to provide some tools which are necessary in the analysis we plan on doing in Part III of the present thesis. For a thorough introduction and deeper insight into these topics we point out the monograph [32] and references therein.

Definition 4.1 (Vector measure)**.** Let (Ω, Σ) be a measurable space and $\mathcal{Y}$ be a Banach space. A function $\mu : \Sigma \to \mathcal{Y}$ is called a (Banach space valued) vector measure, if

$$\forall \left(A_i\right)_{i\in\mathbb{N}}, A_i \in \Sigma, [i \neq j \Rightarrow A_i \cap A_j = \emptyset] \quad \text{it holds} \quad \mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i=1}^{\infty} \mu\left(A_i\right) . \tag{4.1}$$

Here the convergence of the series on the right hand side has to be understood in terms of the norm of $\mathcal{Y}$.

Notice that unlike in [32] we require a vector measure to be σ-additive instead of only finitely additive. Though the latter setting is more general it suffices for our purposes to use this more restricted setting. We shortly assemble basic properties of vector measures.

Proposition 4.2. *Let μ be a vector measure on a measure space (Ω, Σ). Then $\mu(\emptyset) = 0$ and*

$$\forall A, B \in \Sigma, A \cap B = \emptyset : \mu(A \cup B) = \mu(A) + \mu(B) . \tag{4.2}$$

Moreover for all $A, B \in \Sigma$ the identity $\mu(A) = \mu(A \setminus B) + \mu(A \cap B)$ as well as the inclusion, exclusion formula

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$$

hold.

To deduce $\mu(\emptyset) = 0$ just take $A_i = 0$ for all $i \in \mathbb{N}$ in (4.1). Finite additivity then follows from σ additivity by setting $A_i = \emptyset \in \Sigma$ for $i > 2$. The remainder assertions are a consequence of (4.2).

Definition 4.3 (Variation of a measure)**.** Let (Ω, Σ) be a measurable space, $\mathcal{Y}$ a Banach space and $\mu : \Sigma \to \mathcal{Y}$ a vector measure. Then the variation of μ, denoted by $\|\mu\| : \Sigma \to [0, \infty]$, is given by

$$\|\mu\|(A) := \sup \left\{ \sum_{i=1}^{\infty} |\mu(A_i)|_{\mathcal{Y}} : (A_i)_{i\in\mathbb{N}} \subset \Sigma \, , \, \bigcup_{i\in\mathbb{N}} A_i = A \, , \, \forall i \neq j : A_i \cap A_j = \emptyset \right\} \tag{4.3}$$

for all $A \in \Sigma$. If $\|\mu\|(\Omega) < \infty$ then we say μ is of bounded variation.

Indeed $\|\mu\|$ is a measure on (Ω, Σ), see [32, Proposition I.1.9]. We remind ourselves of the definition of absolute continuity of two measures.

Definition 4.4 (Absolute continuity and singularity of measures)**.** Let (Ω, Σ) be a measurable space and $\mu : \Sigma \to \mathcal{Y}$ be a vector measure and $\nu : \Sigma \to [0, \infty]$ be a (real valued, positive) measures. We say that μ is absolutely continuous w.r.t. to ν, in formula $\mu \ll \nu$ if

$$\forall A \in \Sigma : \Big[\nu(A) = 0 \quad \Longrightarrow \quad \mu(A) = 0 \Big] . \tag{4.4}$$

Two (real valued, positive) measures μ and ν are singular, in formulas $\mu \perp \nu$ if there exist $B_1, B_2 \in \Sigma$ with $B_1 \cup B_2 = \Omega$ and $B_1 \cap B_2 = \emptyset$ such that

$$\forall A \in \Sigma : \qquad \mu(A) = \mu\left(A \cap B_1\right) \quad \wedge \quad \nu(A) = \nu\left(A \cap B_2\right) . \tag{4.5}$$

It is easy to see that $\mu \ll \|\mu\|$. The following generalization of the Lebesgue decomposition theorem has been proven in [32, Theorem I.5.9].

Theorem 4.5 (Lebesgue decomposition theorem)**.** *Let (Ω, Σ) be a measure space, μ be a Banach space valued measure of bounded variation and λ a real valued, positive measure. Then there exists two unique vector measures μ_{ac}, μ_s on (Ω, Σ), which are of bounded variation, such that*

$$\|\mu_{ac}\| \ll \lambda\,, \ \|\mu_s\| \perp \lambda \quad \textit{and} \quad \mu = \mu_{ac} + \mu_s\,. \tag{4.6}$$

4.2. Measurability of Banach space valued functions

We introduce several concept of measurability for functions with values in Banach spaces.

Definition 4.6 (Weak and weak star measurability)**.** Let (Ω, Σ) be a measure space, $\mathcal{Y}$ a Banach space and $\mathcal{Y}^*$ its dual. A function $f : \Omega \to \mathcal{Y}$ is weakly measurable if the map

$$\Omega \ni \omega \mapsto \left\langle y', f(\omega)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} \quad \text{is measurable for all } y' \in \mathcal{Y}^*. \tag{4.7}$$

Moreover we say that a function $g : \Omega \to \mathcal{Y}^*$ is weakly star measurable if the map

$$\Omega \ni \omega \mapsto \langle g(\omega), y\rangle_{\mathcal{Y}^*,\mathcal{Y}} \quad \text{is measurable for all } y \in \mathcal{Y}. \tag{4.8}$$

In both cases measurability has to be understood as the measurability with respect to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Definition 4.7 (Simple functions and strong μ-measurability)**.** Let (Ω, Σ) be a measurable space and $\mathcal{Y}$ a Banach space. A function $f : \Omega \to \mathcal{Y}$ is called simple, if

$$\exists (A_i)_{i=1}^n, A_i \in \Sigma, (y_i)_{i=1}^n, y_i \in \mathcal{Y} : \quad f(\omega) = \sum_{i=1}^n \chi_{A_i}(\omega) y_i\,.$$

Let (Ω, Σ, μ) be a measure space. Then a function $f : \Omega \to \mathcal{Y}$ is called strongly μ-measurable if there exists a sequence $f_n : \Omega \to \mathcal{Y}$ simple such that

$$f_n(\omega) \to f(\omega) \quad \text{for } \mu - \text{almost every } \omega \in \Omega\,.$$

Definition 4.8. Let (Ω, Σ, μ) be a measure space and $f : \Omega \to \mathcal{Y}$. f is called μ-almost separable if there exists a set $\mathcal{N} \in \Sigma$ with $\mu(\mathcal{N}) = 0$ such that the range $f(\Omega \setminus \mathcal{N}) = 0$.

Now strong μ-measurability can be characterized in terms of almost separability and weak measurability.

Theorem 4.9 (Pettis theorem [82, Theorem 1.1])**.** *Let (Ω, Σ, μ) be a measure space, $\mathcal{Y}$ a Banach space and $f : \Omega \to \mathcal{Y}$. Then f is strongly μ-measurable if and only if f is μ-almost separable and weakly measurable.*

Due to the above theorem in separable Banach spaces weak measurability is equivalent to strong measurability. If the space moreover is reflexive then also weak star measurability is equivalent to strong measurability.

Notice that if f_n is a sequence of simple functions such that $f_n \to f$ μ almost everywhere. Then the function $\omega \mapsto |f_n(\omega) - f(\omega)|_{\mathcal{Y}}$ is measurable. Thus we can define the Bochner integral for Banach space valued functions. This notion of integration was introduced by Salomon Bochner in his article [12].

Definition 4.10 (Bochner integral)**.** Let (Ω, Σ, μ) be a measure space with $\mu(\Omega) < \infty$ and $\mathcal{Y}$ be a Banach space. For a simple function $f : \Omega \to \mathcal{Y}$, i.e. $f(\omega) = \sum_{i=1}^{n} \chi_{A_i}(\omega) y_i$ the Bochner integral of f w.r.t. μ is defined by

$$\int_{\Omega} f(\omega)\,\mathrm{d}\mu(\omega) := \sum_{i=1}^{n} y_i \cdot \mu(A_i)\,. \tag{4.9}$$

Let $f : \Omega \to \mathcal{Y}$ be strongly μ-measurable, such that there exists a sequence of simple functions $(f_n)_{n\in\mathbb{N}}$ with

$$f_n \to f \;\; \mu - \text{a.e.} \quad \text{and} \quad \int_{\Omega} |f_n - f|_{\mathcal{Y}}\,\mathrm{d}\mu \to 0 \tag{4.10}$$

then f is Bochner-integrable w.r.t μ and the Bochner-integral of f is defined by

$$\int_{\Omega} f(\omega)\,\mathrm{d}\mu(\omega) := \lim_{n\to\infty} \int_{\Omega} f_n(\omega)\,\mathrm{d}\mu(\omega)\,. \tag{4.11}$$

Notice that a function f is integrable if and only if it is strongly measurable and $\omega \mapsto |f(\omega)|_{\mathcal{Y}}$ is integrable. We just state a few of the properties the Bochner integral has.

Proposition 4.11. *Let (Ω, Σ, μ) be a measure space, then it holds*

$$f \mapsto \int_{\Omega} f(\omega)\,\mathrm{d}\mu(\omega) \quad \textit{is linear on the space of } \mu - \textit{Bochner-integrable functions.} \tag{4.12}$$

Moreover if f is μ-Bochner integrable then

$$\left| \int_{\Omega} f(\omega)\,\mathrm{d}\mu(\omega) \right|_{\mathcal{Y}} \leq \int_{\Omega} |f(\omega)|_{\mathcal{Y}}\,\mathrm{d}\mu(\omega)\,. \tag{4.13}$$

Let $\mathcal{X}$ be a Banach space and $T : \mathcal{Y} \to \mathcal{X}$ be a bounded linear operator. If $f : \Omega \to \mathcal{Y}$ is μ-Bochner integrable then so is $Tf : \Omega \to \mathcal{X}$ and it holds

$$\int_\Omega Tf(\omega)\,\mathrm{d}\mu(\omega) = T\int_\Omega f(\omega)\,\mathrm{d}\mu(\omega)\,. \tag{4.14}$$

Remember that the essential supremum of a function $f : \Omega \to \mathbb{R}$ with respect to a measure μ is defined via

$$\text{ess sup}_\mu(f) := \inf\left\{a \in \mathbb{R} : \mu\left(\{\omega \in \Omega : f(\omega) > a\}\right) = 0\right\}\,. \tag{4.15}$$

We now can define L^p spaces of Banach space valued functions.

Definition 4.12. Let (Ω, Σ, μ) be a finite measure space and $\mathcal{Y}$ a Banach space. For $1 \le p < \infty$ we define

$$L^p_\mu(\Omega;\mathcal{Y}) := \left\{ f : \Omega \to \mathcal{Y} : f \text{ is strongly } \mu-\text{measurable and } \int_\Omega |f|^p_{\mathcal{Y}}\,\mathrm{d}\mu < \infty \right\}\,. \tag{4.16}$$

Furthermore we define

$$L^\infty_\mu(\Omega;\mathcal{Y}) := \left\{ f : \Omega \to \mathcal{Y} : f \text{ is strongly } \mu-\text{measurable and ess sup}_\mu\left(|f|_{\mathcal{Y}}\right) < \infty \right\}\,. \tag{4.17}$$

The L^p spaces are Banach spaces with respect to the usual norms

$$\|f\|_{L^p_\mu(\Omega;\mathcal{Y})} := \begin{cases} \left(\int_\Omega |f|^p_{\mathcal{Y}}\,\mathrm{d}\mu\right)^{\frac{1}{p}} & 1 \le p < \infty\,, \\ \text{ess sup}_\mu\left(|f|_{\mathcal{Y}}\right) & p = \infty\,. \end{cases}$$

If $\Omega = (0,T)$, $\Sigma = \mathcal{B}((0,T))$ and μ is the Lebesgue measure on $(0,T)$ we simply denote the space of p-Bochner integrable functions by $L^p(0,T;\mathcal{Y})$. A further helpful result is a generalization of Lebesgue's differentiation theorem to Banach space valued functions. In order to be able to state it we remind ourselves of the following definition.

Definition 4.13. Let Ω be a topological space and (Ω, Σ, μ) a measure space. If

$$\begin{aligned} &\mathcal{B}(\Omega) \subset \Sigma\,, \quad \mu(K) < \infty \text{ for all } K \subset \Omega \text{ compact} \\ &\forall A \subset \Omega \text{ open} : \mu(A) = \sup\{\mu(K) : K \subset A \text{ compact}\} \quad \text{and} \\ &\forall A \in \Sigma : \mu(E) = \inf\{\mu(A) : A \supset E \text{ open}\} \end{aligned}$$

then μ is called a Radon measure.

Theorem 4.14 (Generalized Lebesgue differentiation theorem [35, Theorem 2.9.9])**.** *Let μ be a Radon measure on $((0,T), \sigma_{\mathcal{L}}(0,T))$ where $\sigma_{\mathcal{L}}(0,T)$ is the σ-algebra of all Lebesgue measurable subsets of $(0,T)$. Furthermore let $f \in L^1_\mu(0,T;\mathcal{Y})$. Then for μ almost every $t_0 \in (0,T)$*

$$\lim_{\varepsilon \to 0} \frac{1}{\mu(B(t_0,\varepsilon) \cap (0,T))} \int_{\mu(B(t_0,\varepsilon)\cap(0,T))} |f(t_0) - f(t)|\,\mathrm{d}t = 0\,. \tag{4.18}$$

We also introduce the concept of weak derivatives and Sobolev spaces for Banach space valued functions.

Definition 4.15.

(i) Let $\mathcal{Y}, \mathcal{X}$ be Banach spaces and $\mathcal{Y} \hookrightarrow \mathcal{X}$. Let $f \in L^1(0,T;\mathcal{Y})$ and $v \in L^1(0,T;\mathcal{X})$. Then v is called a weak derivative of f if for all $\phi \in C_0^\infty(0,T)$ the identity

$$\int_0^T f(t)\dot{\phi}(t)\,\mathrm{d}t = -\int_0^T v(t)\phi(t)\,\mathrm{d}t \quad \text{holds.}$$

Here the identity has to be understood in such a way that the left hand side is embedded into the space $\mathcal{X}$. We denote the weak derivative of f by $\dot{f}$.

(ii) For $1 \leq p \leq \infty$ we denote by $W^{1,p}(0,T;\mathcal{Y})$ the space of all p-integrable functions who possess a weak derivative, which is also p-integrable, that is

$$W^{1,p}(0,T;\mathcal{Y}) := \left\{ f \in L^p(0,T;\mathcal{Y}) : \dot{f} \text{ exists and } \dot{f} \in L^p(0,T;\mathcal{Y}) \right\} .$$

It is easy to see that if a weak derivative exists it is unique. Moreover for any $1 \leq p \leq \infty$ space $W^{1,p}(0,T;\mathcal{Y})$ is a Banach space when equipped with the norm

$$\|f\|_{W^{1,p}(0,T;\mathcal{Y})} := \begin{cases} \left(\int_\Omega |f|_\mathcal{Y}^p + |\dot{f}|_\mathcal{Y}^p\,\mathrm{d}t\right)^{\frac{1}{p}} & 1 \leq p < \infty \\ \text{ess sup}\left(|f|_\mathcal{Y} + |\dot{f}|_\mathcal{Y}\right) & p = \infty \end{cases}$$

Remember that a function $f : [0,T] \to \mathcal{Y}$ is absolutely continuous if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all sequences $[x_k, y_k]_{k=1}^\infty$ of pairwise disjoint integrals with

$$\sum_{k=1}^\infty |y_k - x_k| \quad \text{it holds} \quad \sum_{k=1}^\infty |f(y_k) - f(x_k)|_\mathcal{Y} < \varepsilon\,.$$

Unlike it is the case in finite dimensions for Banach space valued functions f absolute continuity does in general not imply $f \in W^{1,1}(0,T;\mathcal{Y})$. Banach spaces $\mathcal{Y}$ for which this assertion is true are called Banach spaces with Radon-Nikodym property.

Definition 4.16 (Radon-Nykodym property [32, Definition III.1.3])**.** A Banach space $\mathcal{Y}$ has the Radon-Nykodym property if for every finite measure space (Ω, Σ, μ) and every vector measure $\nu \ll \mu$ there exists a function $f \in L^1_\mu(\Omega, \mathcal{Y})$ such that for all $E \in \Sigma$

$$\nu(E) = \int_E f(\omega)\,\mathrm{d}\mu(\omega)\,. \tag{4.19}$$

The function f is then called the Radon-Nykodym derivative of ν with respect to μ and denoted by $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$.

Banach spaces which possess the Radon-Nykodym property are for example separable dual spaces $\mathcal{Y}^*$ (Dunford-Pettis theorem [32, Theorem III.3.1]) and all reflexive Banach spaces ([32, Corollary III.2.13]). Prominent examples for spaces which do not have the Radon-Nykodym property are $L^1(\Omega)$ and $L^\infty(\Omega)$ for open and bounded $\Omega \subset \mathbb{R}^n$. A way we use to overcome this problem are weak star derivatives. To this end we use the following result, see [32, p. 84].

Theorem 4.17 (Weak-star representation theorem)**.** *Let $\mathcal{Y}$ be a Banach space and $\mathcal{Y}^*$ be its dual. Let (Ω, Σ, μ) be a finite measure space and $T : L^1_\mu(\Omega) \to \mathcal{Y}^*$ be bounded linear operator. Then there exists a weak star measurable function $g : \Omega \to \mathcal{Y}^*$ such that for each $y \in \mathcal{Y}$ and $f \in L^1_\mu(\Omega)$ it holds*

$$\langle T(f), y\rangle_{\mathcal{Y}^*,\mathcal{Y}} := \int\limits_\Omega f(\omega) \, \langle g(\omega), y\rangle_{\mathcal{Y}^*,\mathcal{Y}} \, \mathrm{d}\mu(\omega) \, . \tag{4.20}$$

This result makes it possible to deduce a weaker version of the Radon-Nykodym theorem which holds for all dual spaces.

Theorem 4.18 (Weak star densities)**.** *Let (Ω, Σ, μ) be a finite measure space, $\mathcal{Y}$ a Banach space and $\mathcal{Y}^*$ it's dual. Then for every $\mathcal{Y}^*$ valued measure ν of bounded variation with $\nu \ll \mu$ there exists a weak star measurable function $g : \Omega \to \mathcal{Y}^*$ such that for all $y \in \mathcal{Y}$ and all $E \in \Sigma$*

$$\langle \nu(E), y\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \int\limits_E \langle g(\omega), y\rangle_{\mathcal{Y}^*,\mathcal{Y}} \, \mathrm{d}\mu(\omega) \, . \tag{4.21}$$

We call g the weak star density of ν with respect to μ. Notice that $\omega \mapsto \langle g(\omega), y\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ is in $L^\infty_\mu(\Omega)$.

This result is obtained by following the proof of [32, Theorem III.1.5] and employing Theorem 4.17.

Proof. The idea is to show that a Radon measure generates a bounded linear operator acting on $L^1_\mu(\Omega)$ with values in $\mathcal{Y}^*$. To this end remember that the variation $\|\nu\|$ is a positive measure on (Ω, Σ). Moreover $\|\nu\|$ vanishes on μ null sets and therefore we obtain that $\|\nu\| \ll \mu$. Due to the Radon-Nykodym theorem for real valued measures (see e.g. [15, Theorem 2]) there exists a function $h \in L^1_\mu(\Omega)$ such that

$$\|\nu\|(E) = \int\limits_E h(\omega) \, \mathrm{d}\mu(\Omega) \ \text{ for all } E \in \Sigma \, .$$

Define for $n \in \mathbb{N}$ the set $E_n := \{\omega \in \Omega : n-1 \leq h(\omega) < n\} \in \Sigma$. Since both $\|\nu\|$ and μ are positive measures it holds $h \geq 0$ μ-almost everywhere. Moreover $E_n \cap E_m = \emptyset$ for $n \neq m$ and $\bigcup\limits_{n\in\mathbb{N}} E_n = \Omega$. For any n and any simple function $f = \sum\limits_{i=1}^k a_i \chi_{A_i}$, $A_i \in \Sigma$, $A_i \cap A_j = \emptyset$ if $i \neq j$ we define

$$T_n(f) = \sum_{i=1}^k \alpha_i \nu(E_n \cap A_i) \, . \tag{4.22}$$

We then can estimate that

$$|T_n(f)|_{\mathcal{Y}} \leq \sum_{i=1}^{k} |\alpha_i| \, |\nu(E_n \cap A_i)|_{\mathcal{Y}^*} \leq \sum_{k=1}^{n} |\alpha_i| \|\nu\|(E_n \cap A_i) \leq n \|f\|_{L^1_\mu(\Omega)} \,. \tag{4.23}$$

Hence T_n can be extended to a bounded linear operator from $L^1_\mu(\Omega)$ to $\mathcal{Y}^*$. Due to Theorem 4.17 there exists a weak-star measurable function $g_n : \Omega \to \mathcal{Y}^*$ such that for all $y \in \mathcal{Y}$ and $f \in L^1_\mu(\Omega)$

$$\langle T_n(f), y \rangle_{\mathcal{Y}^*,\mathcal{Y}} = \int_\Omega f(\omega) \, \langle g_n(\omega), y \rangle_{\mathcal{Y}^*,\mathcal{Y}} \, \mathrm{d}\mu(\omega) \tag{4.24}$$

and especially for all $E \in \Sigma$

$$\langle \nu(E \cap E_n), y \rangle_{\mathcal{Y}^*,\mathcal{Y}} = \int_E \langle g_n(\omega), y \rangle_{\mathcal{Y}^*,\mathcal{Y}} \, \mathrm{d}\mu(\omega) \,. \tag{4.25}$$

We define the function $g : \Omega \to \mathcal{Y}^*$ by $g(\omega) = g_n(\omega)$ for all $\omega \in E_n$. Then for any $y \in \mathcal{Y}$ we have

$$\int_{\bigcup_{n=1}^{m} E_n} | \langle g(\omega, y) \rangle_{\mathcal{Y}^*,\mathcal{Y}} | \, \mathrm{d}\mu(\omega) \leq \|\nu(\Omega\| |y|_{\mathcal{Y}} \tag{4.26}$$

and hence $\omega \mapsto \langle g(\omega, y) \rangle_{\mathcal{Y}^*,\mathcal{Y}}$ is in $L^1_\mu(\Omega)$. Moreover since ν is σ-additive we have for all $E \in \Sigma$

$$\begin{aligned}
\langle \nu(E), y \rangle_{\mathcal{Y}^*,\mathcal{Y}} &= \lim_{m\to\infty} \left\langle \nu \left(E \bigcap \left(\bigcup_{n=1}^{m} E_n \right) \right), y \right\rangle_{\mathcal{Y}^*,\mathcal{Y}} \\
&= \lim_{m\to\infty} \int_{E \bigcap (\bigcup_{n=1}^{m} E_n)} \langle g(\omega, y) \rangle_{\mathcal{Y}^*,\mathcal{Y}} \, \mathrm{d}\mu(\omega) = \int_E \langle g(\omega, y) \rangle_{\mathcal{Y}^*,\mathcal{Y}} \, \mathrm{d}\mu(\omega)
\end{aligned}$$

where the last equality is due to the dominated convergence theorem. □

4.3. Young-measure theory in Banach spaces

The concept of Young measures was introduced by L. C. Young [106, 107, 108] in order to provide existence results in optimal control theory. Subsequently it was successfully applied to a vast number of problems in the calculus of variations, such as the analysis of micro structures in continuum mechanics, see [81] and references therein.

Given a set Ω and a family of measurable functions $u_k : \Omega \to \mathbb{R}^d$ the Young measure ν generated by the sequence (u_k) is a family $\nu = (\nu_x)_{x\in\Omega}$ of (sub-)probability measures which give insight in the limiting behavior of u_k as $k \to \infty$. Roughly speaking for some $x_0 \in \Omega$ and $A \subset \mathbb{R}^d$ measurable $\nu_{x_0}(A)$ is the probability of $\lim u_k(x) \in A$ for 'x close to x_0' ([97, Section 5.6]). This is formulated in a precise way in the fundamental theorem of Young measures.

Theorem 4.19 (Fundamental theorem of Young measures [81, Theorem 3.1]). *Let $\Omega \subset \mathbb{R}^n$ be a measurable set with finite measure and $(u_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions $u_n : \Omega \to \mathbb{R}^d$. Then there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ and a weak-*-measurable map $\nu : \Omega \to \mathcal{M}(\mathbb{R}^d)$ such that*

(i) $\nu_x \geq 0$ *and* $\nu_x(\mathbb{R}^d) \leq 1$ *for almost every* $x \in \Omega$,

(ii) *for all* $f \in C_0(\mathbb{R}^d)$ *it holds*

$$f(u_{n_k}) \overset{*}{\rightharpoonup} \overline{f} \ \text{ in } L^\infty(\Omega)$$

where

$$\overline{f}(x) = \int_{\mathbb{R}^d} f(z)\, \mathrm{d}\nu_x(z)\,, \tag{4.27}$$

(iii) *if* $K \subset \mathbb{R}^d$ *is compact, then*

$$\operatorname{dist}(u_{n_k}, K) \to 0 \text{ in measure } \Rightarrow \operatorname{supp}\{\nu_x\} \subset K\,, \tag{4.28}$$

(iv) $\nu_x(\mathbb{R}^d) = 1$ *for almost all* $x \in \Omega$ *if and only if*

$$\lim_{M\to\infty} \sup_{k\in\mathbb{N}} |\{x : \|u_{n_k}(x)\| > M\}| = 0\,, \tag{4.29}$$

(v) *if* (4.29) *holds, if* $A \subset \Omega$ *is measurable, if* $f \in C(\mathbb{R}^d)$ *and if*

$$f(u_{n_k}) \text{ is relatively weakly compact in } L^1(A)\,, \tag{4.30}$$

then

$$f(u_{n_k}) \rightharpoonup \overline{f} \tag{4.31}$$

where $\overline{f}$ *is defined as above,*

(vi) *if* (4.29) *holds then* (iii) *is in fact an equivalence.*

Condition (4.29) is for example satisfied if u_{n_k} is bounded in $L^p(\Omega, \mathbb{R}^d)$ for some $p \geq 1$. From a mathematical point of view especially properties (ii) and (v) are very interesting as they characterize weak (-star) limits of *nonlinear* functions of u_{n_k}. Remember that even if a weak(-star) limit of a sequence (u_{n_k}) exists it provides no information of the weak (-star) limit of the sequence $f(u_{n_k})$ if f is nonlinear.

It is apparent that such a tool would also be helpful in the study of evolution equations. A typical framework for these problems would contain a family functions (u_n) mapping a (finite) time interval $[0,T]$ into a Banach space $\mathcal{Y}$. The first result obtained in this direction has been obtained by E. Balder [8], who proved a fundamental theorem of Young measure in a very general setting. We also point out the lecture notes [104, 9] for a comprehensive introduction to this topic. However, these results have been proven for spaces endowed with a metrizable topology and can therefore not directly be applied to Banach spaces endowed with their weak topologies. This problem has been overcome by R. Rossi and G. Savaré in [94, Theorem 3.2] who proved a fundamental theorem of Young measures for weak topologies in separable Hilbert spaces. We shall go a small step further and provide this result for the topology induced by the weak-star convergence in the dual of a separable normed space. Before we start, a short definition is needed.

Definition 4.20. A $\mathcal{L} \times \mathcal{B}(\mathcal{Y}^*)$-measurable function $h : (0,T) \times \mathcal{Y}^* \to (-\infty, \infty]$ is called a weakly-star normal integrand if the map

$$\mathcal{Y}^* \ni y' \to h(t, y') \text{ is weakly-star lower semicontinuous for a.e. } t \in [0,T].$$

A sequence $(y'_n)_{n\in\mathbb{N}} \subset L^1(0,T;\mathcal{Y})$ is called weakly-star tight, if there exists a non-negative normal integrand $h : (0,T) \times \mathcal{Y}^* \to [0,\infty)$ such that

$$\lim_{|y'|\to\infty} h(t,y') = +\infty \quad \text{for a.e. } t \in [0,T] \quad \text{and}$$

$$\sup_{n\in\mathbb{N}} \int_0^T h(t, u_n(t))\,\mathrm{d}t < \infty$$

Theorem 4.21 (Fundamental theorem for weak-star topologies)**.** *Let $\mathcal{Y}$ be separable normed space and $\mathcal{Y}^*$ be its dual. Then for any weakly-star tight sequence $(y'_n)_{n\in\mathbb{N}}$ in $L^1(0,T;\mathcal{Y}^*)$ there exists a subsequence $\left(y'_{n_k}\right)_{k\in\mathbb{N}}$ and a family of parametrized measures $\boldsymbol{\nu} = (\nu_t)_{t\in(0,T)}$, $\nu_t \in \mathcal{M}(\mathcal{Y}^*)$ such that for almost all $t \in (0,T)$ it holds*

$$\nu_t(\mathcal{Y}^*) = 1\,, \quad \limsup_{k\uparrow\infty} \left|y'_{n_k}(t)\right|_{\mathcal{Y}^*} < \infty \quad \textit{and} \quad \operatorname{supp}(\nu_t) \subset \bigcap_{j=1}^{\infty} \overline{\left\{y'_{n_k}(t) : k \geq j\right\}}^{w*}. \tag{4.32}$$

Moreover for every weakly-star normal integrand $h : [0,T] \times \mathcal{Y}^* \to (-\infty,\infty]$ such that $h^-\left(\cdot, y'_{n_k}(\cdot)\right)$ is uniformly integrable it holds

$$\liminf_{k\to\infty} \int_0^T h\left(t, y'_{n_k}(t)\right)\,\mathrm{d}t \geq \int_0^T \int_{\mathcal{Y}}^{*} h\left(t, y'\right)\,\mathrm{d}\nu_t(y')\,\mathrm{d}t\,. \tag{4.33}$$

The proof follows the lines of [94, Theorem 3.2].

Proof. Since $\mathcal{Y}$ is separable there exists a sequence $\boldsymbol{y} = (y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}$ with $|y_n| = 1$ for all $n \in \mathbb{N}$ such that

$$\mathcal{Y} = \overline{\operatorname{span}\{y_n : n \in \mathbb{N}\}}\,. \tag{4.34}$$

We now define the $\|\cdot\|_{\boldsymbol{y}}$ via

$$\|y'\|_{\boldsymbol{y}} := \sum_{n=1}^{\infty} 2^{-n} \left|\langle y', y_n\rangle_{\mathcal{Y}^*,\mathcal{Y}}\right|. \tag{4.35}$$

Then $\|\cdot\|_{\boldsymbol{y}}$ is a norm with $\|y'\|_{\boldsymbol{y}} \leq |y'|_{\mathcal{Y}^*}$ and for any sequence $(y'_n)_{n\in\mathbb{N}}$ bounded with respect to $|\cdot|_{\mathcal{Y}^*}$ it holds $y'_n \overset{*}{\rightharpoonup} y'$ if and only if $\|y'_n - y'\|_{\boldsymbol{y}} \to 0$ ([13, Theorem 1]). Therefore any bounded ball $\{y' \in \mathcal{Y}^* : |y|_{\mathcal{Y}^*} \leq c\} \subset \mathcal{Y}^*$ is compact with respect to $\|\cdot\|_{\boldsymbol{y}}$. Because of [33, Theorem I.6.15] any bounded ball is therefore separable with respect to $\|\cdot\|_{\boldsymbol{y}}$. [1] Now define the space

$$E := \left\{(y', w) \in \mathcal{Y}^* \times \mathbb{R} : |y'|_{\mathcal{Y}^*} \leq w\right\} \subset \mathcal{Y}^* \times \mathbb{R} \tag{4.36}$$

[1] Note that on bounded subsets the weak star topology of $\mathcal{Y}^*$ is the same as the topology induced by $\|\cdot\|_{\boldsymbol{y}}$. We would like to point out a small peculiarity: The separability of $\mathcal{Y}^*$ w.r.t. the weak-star topology does not guarantee the separability of a bounded ball in $\mathcal{Y}^*$ w.r.t. the same topology, see [27].

and the function

$$d : E \times E \to [0,\infty)\,, \quad \big((y_1', w_1), (y_2', w_2)\big) \mapsto \left\|y_1' - y_2'\right\|_{\boldsymbol{y}} + |w_1 - w_2|\,. \tag{4.37}$$

Following our above considerations d is a metric on E and it holds that

$$\forall (y_n', w_n)_{n\in\mathbb{N}} \subset E : \Big[(y_n', w_n) \stackrel{d}{\longrightarrow} (y', w) \Leftrightarrow y_n' \stackrel{*}{\rightharpoonup} y' \;\wedge\; w_n \to w\Big]\,. \tag{4.38}$$

Therefore E is complete and separable with respect to the metric d and bounded weakly-star closed subsets of E are compact with respect to d. Moreover for any closed ball $B \subset \mathcal{Y}^* \times \mathbb{R}$ it holds $B \cap E$ is a Borel set of E. This implies that

$$B \subset \mathcal{B}\left(\mathcal{Y}^* \times \mathbb{R}\right) \Rightarrow B \cap E \in \mathcal{B}(E) \tag{4.39}$$

and hence any Borel measure on E can be trivially extended to a Borel measure on $\mathcal{Y}^* \times \mathbb{R}$.

We now may apply Balder's theorem [8, Theorem 3] to the sequence $(\boldsymbol{u}_n)_n = (y_n', |y_n'|_{\mathcal{Y}^*})_n \subset E$. It grants the existence of a subsequence $(\boldsymbol{u}_{n_k})_k$ and a family of parameterized measure $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]}$ such that for almost all $t \in [0,T]$

$$\operatorname{supp}(\mu_t) \subset \bigcap_{j=1}^{\infty} \overline{\{u_{n_k}(t) : k \geq j\}}^{d}\,. \tag{4.40}$$

Moreover for any E-normal integrand g for which $g^-\left(\cdot, u_{n_k}(\cdot), |u_{n_k}(\cdot)|\right)$ is uniformly integrable the inequality

$$\liminf_{k\to 0} \int_0^T g\left(t, u_{n_k}(t), |u_{n_k}(t)|_{\mathcal{Y}^*}\right)\,\mathrm{d}t \geq \int_0^T \left(\int_E g(t, y', w)\,\mathrm{d}\mu_t(y', w)\right)\,\mathrm{d}t \tag{4.41}$$

holds. Defining the parametrized measure $\boldsymbol{\nu} = (\nu_t)_{t\in[0,T]}$ via

$$\nu_t(B) = \mu_t\left(B \times [0,\infty)\right) \;\; \forall B \in \mathcal{B}(\mathcal{Y}^*) \tag{4.42}$$

we see that $\boldsymbol{\nu}$ fulfills (4.32) and (4.33). □

In applications it might also be of interest to generalize a result due to U. Stefanelli [101, Theorem 4.3] who provided a $\liminf$-inequality in terms of Young measures for the $\Gamma - \liminf$ of a sequence of functionals.

Theorem 4.22 ($\Gamma - \liminf$ result in weak star topologies)**.** *Let $\mathcal{Y}$ be separable normed space and $\mathcal{Y}^*$ be its dual. Furthermore let $(y_n')_{n\in\mathbb{N}}$ be a weakly-star tight sequence in $L^1(0,T;\mathcal{Y}^*)$ and $g_n, g : [0,T] \times \mathcal{Y}^* \to (-\infty,\infty]$ be weakly star normal integrands such that for all $y' \in \mathcal{Y}^*$ and almost all $t \in [0,T]$*

$$g(t, y') \leq \inf\left\{\liminf_{n\to\infty} g_n(t, y_n') : y_n' \stackrel{*}{\rightharpoonup} y' \text{ in } \mathcal{Y}^*\right\}\,. \tag{4.43}$$

Then there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ and a family of parametrized measures $\boldsymbol{\nu} = (\nu_t)_{t\in(0,T)}$, $\nu_t \in \mathcal{M}(\mathcal{Y}^*)$ such that for almost all $t \in (0,T)$

$$\nu_t(\mathcal{Y}^*) = 1\,, \quad \limsup_{k\uparrow\infty} \left|y'_{n_k}(t)\right|_{\mathcal{Y}^*} < \infty \quad \textit{and} \quad \mathrm{supp}(\nu_t) \subset \bigcap_{j=1}^{\infty} \overline{\left\{y'_{n_k}(t) : k \geq j\right\}}^{w*}\,. \tag{4.44}$$

and if the sequence $g^-_{n_k}\left(\cdot, y'_{n_k}(\cdot)\right)$ is uniformly integrable it holds

$$\liminf_{k\to\infty} \int_0^T g_{n_k}\left(t, y'_{n_k}(t)\right)\,\mathrm{d}t \geq \int_0^T \int_{\mathcal{Y}}^* g\left(t, y'\right)\,\mathrm{d}\nu_t(y')\,\mathrm{d}t\,. \tag{4.45}$$

The proof is just a straightforward adaption of the proof of [101, Theorem 4.3], just make use of the norm $\|\cdot\|_{\boldsymbol{y}}$, and shall therefore be omitted here. Notice that condition (4.43) says that $g(t,\cdot)$ is for almost every $t \in [0,T]$ less or equal than the $\Gamma - \liminf$ of the sequence $g_n(t,\cdot)$.

The results stated here contain the previous mentioned results [94, Theorem 3.2] and [101, Theorem 4.3]. These have been proven in the setting of weak topologies of either separable Hilbert spaces or separable and reflexive Banach spaces. To see this just remember that the weak star topology of a reflexive Banach space is just the same as the weak topology and that if a reflexive Banach space is separable so is its dual.

5. Gronwall's inequality

In the vast toolbox for the study of evolution equation Gronwall's inequality certainly takes a prominent role. Proven in [39] more than 90 years ago it is nowadays still an indispensable and widely used instrument. Today quite a big number of variants of this inequality are in common usage. We shall state it in its original form here.

Lemma 5.1 (Gronwall's Lemma)**.** *When, for $x_0 \le x \le x_0+h$, the continuous function $z = z(x)$ satisfies the inequalities*

$$0 \le z \le \int_{x_0}^{x} Mz + A \,\mathrm{d}x$$

where the constants M and A are positive or zero, then

$$0 \le z \le Ahe^{Mh}, \quad x_0 \le x \le x_0 + h\,.$$

The proof of this Lemma is well known and shall be omitted here. Instead we are going to present two discrete analoga of this results, which are needed in the sequel. The first one has been proven in a joint work with P. Krejčí [62, Lemma A.1].

Lemma 5.2. *Let $N \in \mathbb{N}$, $(\delta_k)_{k=0}^N, (c_k)_{k=1}^N, (a_k)_{k=1}^N \subset \mathbb{R}^+$ be given sequences. Assume that*

$$\delta_k - \delta_{k-1} \le c_k + a_k\delta_{k-1} \ \ \forall k \in [N]\,. \tag{5.1}$$

Then for all $n \in [N]$ it holds

$$\delta_n \le \exp\left(\sum_{k=1}^n a_k\right)\left(\delta_0 + \sum_{k=1}^n c_k\right) \tag{5.2}$$

Proof. We have by hypothesis

$$\frac{\delta_k}{\prod_{i=1}^k(1+a_i)} - \frac{\delta_{k-1}}{\prod_{i=1}^{k-1}(1+a_i)} \le \frac{c_k}{\prod_{i=1}^k(1+a_i)} \le c_k \quad \forall k \in [N]\,. \tag{5.3}$$

Summing up over $k \in [n]$ we obtain

$$\delta_n \le \prod_{i=1}^n(1+a_i)\left(\delta_0 + \sum_{k=1}^n c_k\right)\,,$$

and the inequality $1 + a_i \le \exp(a_i)$ completes the proof. □

The second discrete version of Gronwall's inequality has been proven by the author in [90, Lemma A.1]. It is adapted to the framework, in which it shall be applied in the sequel, and might, at first glance, appear to have little in common with the original inequality. However it is little more than yet another variation of Lemma 5.1.

Lemma 5.3. *Let $N \in \mathbb{N}$, $(\xi_k)_{k=0}^N, (\eta_k)_{k=0}^N \subset X$, $(b_k)_{k=0}^N, (a_k)_{k=0}^N \subset \mathbb{R}_{\geq 0}$ with*

$$|\Delta_k \xi| + \Delta_k b \leq a_k \left(|\xi_{k-1}| + |\eta_{k-1}|\right) + \delta |\Delta_k \eta| \;\; \forall k \in [N] \tag{5.4}$$

for some $0 \leq \delta < 1$. Here $\Delta_k f$ indicates the difference $f_k - f_{k-1}$ for all $k \in [N]$ and $f \in \{\xi, \eta, b, a\}$. Assume furthermore that $b_0 = 0$ and $\xi_0 = \eta_0 = 0$. Then there exists $\varepsilon \geq 0$ and $0 \leq \rho < 1$ such that for

$$w_k = \exp\{-\frac{1}{\varepsilon}\sum_{i=1}^k a_k\} \;\; \forall k \in [N] \tag{5.5}$$

it holds

$$\sum_{k=1}^N |\Delta_k \xi| w_k \leq \rho \sum_{k=1}^N |\Delta_k \eta| w_k \,. \tag{5.6}$$

Proof. Choose any $\varepsilon < (1-\delta)/2$ and set

$$\rho := \frac{\delta + \varepsilon}{1 - \varepsilon} < 1 \,. \tag{5.7}$$

We multiply both sides of (5.4) by w_k. Since w_k is a decreasing sequence and $b_0 = 0$ we have

$$\sum_{k=1}^N (\Delta_k b)\, w_k = b_N w_N + \sum_{k=1}^{N-1} b_k (w_k - w_{k+1}) - b_0 w_1 \geq 0 \,. \tag{5.8}$$

Therefore we obtain

$$\sum_{k=1}^N |\Delta_k \xi| w_k \leq \sum_{k=1}^N a_k w_k \left(|\xi_{k-1}| + |\eta_{k-1}|\right) + \delta \sum_{k=1}^N |\Delta_k \eta| w_k \,. \tag{5.9}$$

We estimate the first term on the right hand side via

$$\begin{aligned} \sum_{k=1}^N a_k w_k |\xi_{k-1}| &\overset{\xi_0 = 0}{\leq} \varepsilon \sum_{k=2}^N \frac{1}{\varepsilon} a_k w_k \left(\sum_{j=1}^{k-1} |\Delta_j \xi| \right) \\ &= \varepsilon \sum_{j=1}^{N-1} |\Delta_j \xi| \left(\sum_{k=j+1}^N \frac{1}{\varepsilon} a_k w_k \right) . \end{aligned}$$

Furthermore we can calculate

$$\sum_{k=j+1}^N \frac{1}{\varepsilon} a_k w_k \leq \int_{\frac{1}{\varepsilon}\sum_{k=1}^j a_k}^{\frac{1}{\varepsilon}\sum_{k=1}^N a_k} \exp(-x)\,\mathrm{d}x \leq \exp\left\{-\frac{1}{\varepsilon}\sum_{k=1}^j a_k\right\} = w_j \tag{5.10}$$

by interpreting the left hand side as a Riemann sum. Proceeding in exactly the same way for $\sum a_k w_k |\eta_{k-1}|$ we obtain

$$\sum_{k=1}^N |\Delta_k \xi| w_k \leq \varepsilon \sum_{k=1}^N \left(|\Delta_k \xi| + |\Delta_k \eta|\right) w_k + \delta \sum_{k=1}^N |\Delta_k \eta| w_k \,. \tag{5.11}$$

This completes the proof. □

Part II.

Quasivariational sweeping processes on functions of bounded variation

6. Introduction and main results

We establish existence and uniqueness results of quasivariational (or implicit) sweeping processes on functions of bounded variation on a separable Hilbert space X. To extend the sweeping process to BV we make use of the so-called Kurzweil formulation. The results we are able to prove depend on the shape of the involved convex sets. If the convex sets is a polyhedron, we prove - under suitable conditions - existence and uniqueness on the whole space of functions of essentially bounded variation. If the convex sets are smooth we are able to establish this result for functions of bounded variation with small jumps. Indeed in the latter case we provide a counterexample to uniqueness in case the jumps are too big.

6.1. The sweeping process

The sweeping process has been introduced by J. J. Moreau in two seminal articles [78, 79] in the 1970s. He proposed the following evolution problem:
Given a moving convex set $K(t)$ and a point $\xi(t)$ we assume that the evolution of the point $\xi(t)$ is governed by the evolution of $K(t)$ through two rules. First $\xi(t)$ has to stay within $K(t)$. And second if $\xi(t)$ moves then it moves only in the opposite direction of the normal cone of $K(t)$ at point $\xi(t)$. Especially if $\xi(t)$ is in the interior of $K(t)$ then it does not move. The question we are going to ask is: Does there exist an unique solution to this problem? Or to write it down more precisely:

Problem 6.1 (Sweeping process)**.** *Given a time dependent convex set $K(t) \subset X$ and an initial value $\xi_0 \in K(0)$, does there exist an absolutely continuous function ξ such that $\xi(0) = \xi_0$ and*

$$-\dot{\xi}(t) \in \partial I_{K(t)}(\xi) \ \text{ a.e. in } [0,T] \tag{6.1}$$

hold?

Moreau provided an existence and uniqueness result for this problem in [79] under suitable conditions on the convex set. Laxly formulated the condition is that the convex set $K(t)$ does evolve with absolutely continuous speed. One might wonder whether solutions of (6.1) are differentiable. In general this not the case. To illustrate this fact consider a simple example.

Example 6.2. Let $X = \mathbb{R}$, $K(t) = [-1+t, 1+t]$ and $\xi_0 = 0$. Then the solution to the sweeping process is given by

$$\xi(t) = \begin{cases} 0 & \text{if } t \leq 1 \\ t-1 & \text{if } t > 1 \end{cases}$$

which is clearly not differentiable in $t = 1$.

As in [18] we shall slightly rewrite the problem: Let $Z(r)$ be a family of convex sets indexed by elements r of an reflexive Banach space $\mathcal{R}$, $u \in W^{1,1}(0,T;X)$ and $r \in W^{1,1}(0,T;\mathcal{R})$. We look for absolutely continuous functions ξ such that $\xi(0) = \xi_0$ and

$$\begin{aligned} u(t) - \xi(t) &\in Z(r(t)) \qquad \forall t \in [0,T] \qquad \text{and} \\ \left\langle \dot{\xi}(t), u(t) - \xi(t) - y \right\rangle &\geq 0 \quad \forall y \in Z(r(t)) \text{ for a.e. } t \in [0,T]\,. \end{aligned} \tag{6.2}$$

This is nothing but a reformulation of (6.1) where $K(t)$ takes the form $K(t) = u(t) - Z\,(r(t))$. Note that every problem formulated in terms of u, r and Z can be formulated in terms of $K(t)$ and vice versa. One direction we have just shown and for the other direction set $\mathcal{R} = \mathbb{R}$, $Z(t) = -K(t)$, $u = 0$ and $r(t) = t$. The use of (6.2) is motivated by applications where it often appears in a natural way. In the special case of $Z(r) = Z$ the map $u \mapsto \xi$ which maps the input u to the solution of the sweeping process ξ is exactly the play operator from the theory of hysteresis. This theory has been developed from the late 1960s by the Russian school around M. A. Krasnosel'skiĭ (see e.g. [51, 50]). Their efforts culminated in the seminal monograph [52] which opened the path to a rigorous mathematical treatment of hysteresis. Indeed we will make use of many methods developed for the study of multidimensional play operators. For an extensive study of these we refer to [55] and references therein. For the study of one dimensional hysteresis operators and especially the one dimensional play, which exhibits a few additional interesting properties, we refer to the book [19].

6.2. Extending the sweeping process to BV

When extending the sweeping process to functions of (essentially) bounded variation, i.e. allowing u and r and consequently also the solution ξ to be in this space, one encounters a difficulty: The formulation of (6.1) or (6.2) can no longer be applied as ξ might no longer have a weak derivative. Therefore one must come up with a different formulation. It is then natural to ask for two conditions to be satisfied:

1. If the input functions u and r are absolutely continuous, then the solution ξ of the new formulation must also solve the original problem.

2. The formulation may only contain concepts which have a meaning on BV.

However this does not uniquely determine a way to extend the sweeping process to functions of bounded variation. The degree of freedom one essentially has, is to decide what happens at points where the functions jump. This has been solved in different ways. The first concept is already due to J. J. Moreau [79] who proposed to use the weak derivative with respect to the underlying Radon measure of the function ξ.

For the study of play operators two other approaches have been developed. P. Krejčí and Ph. Laurençot proposed to use an integral formulation of the variational inequality (6.2). This approach was later on extended to sweeping processes in [61]. A different proposal is due to V. Recupero [86, 87, 88]. Here the evolution at the jump points is determined by, roughly speaking, 'filling in' the jump with a function that connects the starting point with the end point of the jump. The solution is then determined by letting the evolution run on the 'filled in' function. We however will follow the concept of P. Krejčí and Ph. Laurençot, which is nowadays known as the *Kurzweil formulation*. It is named after the Czech mathematician Jaroslav Kurzweil, who developed the integral which plays a crucial role in this formulation (see Chapter 3.2). We will now shortly introduce the idea behind this approach. For simplicity we restrict ourselves to the play operator. Afterwards we state the sweeping process in its Kurzweil formulation and prove the existence of a unique solution.

Before we proceed a short remark on our wording is in order: We will talk about a *'sweeping process on BV'* when we want to say that the input and output functions are allowed to be in BV. One might argue this to be an odd wording, as it might appear more natural to talk about a *'sweeping process on X'* as X is the space in which the convex set moves. However, as we have now stated what we mean by our choice of words, we hope that it is acceptable.

By definition ξ is the output of the play operator if $u(t) - \xi(t) \in Z$ and

$$\left\langle \dot{\xi}(t), u(t) - \xi(t) - y \right\rangle \geq 0 \quad \forall y \in Z, \text{ for a.e. } t \in [0,T] \,.$$

In [55] it has been shown that this condition is equivalent to an integral inequality, namely

$$\int_0^T \left\langle u(t) - \xi(t) - y(t), \dot{\xi}(t) \right\rangle \geq 0 \quad \forall y \in C(0,T;Z) \,.$$

The idea now is to substitute the above integral by an integral where the right side of the inner product is a generalized derivative of a BV function. Probably the first integral that comes into ones mind is the well known Riemann-Stieltjes integral. However it is not the right choice here as it asks for the left side to be continuous, which in general will not be the case. Still one can stay in the realm of 'classical theory' and does not have use integrals based on measures - this of course is also a possible path. Our choice is to employ the Kurzweil-Henstock integral (Chapter 3.2). Additionally we extend the space of admissible test functions to all regulated functions, that is functions which allow left and right side limits at every point, with values in Z. One now might be tempted to ask for functions $\xi \in \overline{\mathrm{BV}}(0,T;X)$ satisfying $u(t) - \xi(t) \in Z$ and

$$\int_0^T \langle u(t) - \xi(t) - y(t), \, \mathrm{d}\xi(t) \rangle \geq 0 \quad \forall y \in G(0,T;Z) \,. \tag{6.3}$$

This however is a bad choice. Indeed in general there will be no function ξ which can satisfy this inequality, as the following example shows.

Example 6.3. Choose $Z = [-1,1]$, $x_0 = 0$ and define

$$u(t) = \begin{cases} 0 & t = 0 \\ 2 & t > 0 \end{cases} \,.$$

Assume ξ is satisfies (6.3). Then it holds that $\xi(0) = 0$. We choose the test function y to be defined as

$$y(t) = \begin{cases} z & t = 0 \\ u(t) - \xi(t) & t > 0 \end{cases}$$

with $z \in [-1,1]$. By plugging this into (6.3) we obtain $\langle -z, \xi(0+) - \xi(0)\rangle \geq 0$ for every $z \in [-1,1]$, which can only be satisfied if $\xi(0+) = \xi(0) = 0$. However, since ξ is a solution we especially have $\xi(t) - u(t) \in Z$. By choice of u this means $\xi(t) \geq 1$ for all $t > 0$, which is a contradiction.

As this is an unsatisfying state, it is necessary to tweak (6.3) in order to obtain a better suited condition. Indeed the correct condition reads

$$\int_0^T \langle u(t+) - \xi(t+) - y(t),\, \mathrm{d}\xi(t)\rangle \geq 0 \quad \forall y \in G(0,T;Z)\,. \tag{6.4}$$

Assume that ξ has a jump at any point $s \in [0,T]$. We choose the testfunction y as $y(t) = u(t) - \xi(t)$ if $t \neq s$ and $y(s) = z \in Z$. Then the integral inequality reduces to

$$\langle u(s+) - \xi(s+) - z, \xi(s+) - \xi(s-)\rangle \geq 0\,.$$

This inequality can be fulfilled namely by setting $\xi(s+) = Q_Z\left(u(s+) + \xi(s-)\right)$. Indeed this formulation can be generalized to sweeping processes and it then reads as follows:

Problem 6.4 (Sweeping process)**.** *For given input functions $u \in \overline{\mathrm{BV}}(0,T;X)$, $r \in \overline{\mathrm{BV}}(0,T;\mathcal{R})$ and initial condition $x_0 \in Z(r(0))$, we look for a function $\xi \in \overline{\mathrm{BV}}(0,T;X)$ such that*

$$x(t) := u(t) - \xi(t) \in Z(r(t)) \qquad \forall t \in [0,T]\,, \tag{6.5}$$

$$x(0) = x_0\,, \tag{6.6}$$

$$\int_0^s \langle x(t+) - y(t),\, \mathrm{d}\xi(t)\rangle \geq 0 \tag{6.7}$$

for all $s \in [0,T]$ and every $y \in G(0,s;X)$ such that $y(t) \in Z(r(t+))$ for every $t \in [0,s]$.

Of course the question arises whether and under which condition a solution to this problem exists. We settle this question with, at least for our purposes, sufficient generality in the following theorem.

Theorem 6.5. *Let $Z(r)$ be a family of closed convex sets indexed by a parameter $r \in \mathcal{R}$. Assume that*

$$d_H\left(Z(r), Z(s)\right) \leq L\|r - s\|_{\mathcal{R}}\,.$$

Then for every $u \in \overline{\mathrm{BV}}(0,T;X)$, $r \in \overline{\mathrm{BV}}(0,T;\mathcal{R})$ and $x_0 \in Z(r(0))$ there exists a unique solution $\xi \in \overline{\mathrm{BV}}(0,T;X)$ to Problem 6.4 and

$$\overline{\mathrm{Var}}(\xi) \leq \overline{\mathrm{Var}}(u) + L\overline{\mathrm{Var}}(r)\,. \tag{6.8}$$

Furthermore the sweeping process is locally $\frac{1}{2}$-Hölder continuous with respect to $\|\cdot\|_\infty$.

Up to our knowledge this result has not been proven or at least published under these precise assumptions. However there are a number of related results: For the play operator existence and uniqueness of a solution has been shown in [60]. Unique solvability of a slightly more general problem was shown in [61], however only for left-continuous functions of bounded variation. A closely related result has been shown in the lecture notes [58] for right continuous functions. Our proof relies on the methods developed in these articles and is written down mainly for the sake of completeness.

Proof. We proceed in three steps. We first establish a continuity result which will also imply the uniqueness of a solution. We then explicitly calculate solutions for stepfunctions and thus establish the existence. Finally we expand the result to the whole space $\overline{\mathrm{BV}}$.

Step 1 - Hölder continuity: Let $u, v \in \overline{\mathrm{BV}}(0,T;X)$, $r, s \in \overline{\mathrm{BV}}(0,T;\mathcal{R})$, $x_0 \in Z(r(0))$ and $y_0 \in Z(s(0))$. Assume that $\xi \in \overline{\mathrm{BV}}(0,T;X)$ and $\eta \in \overline{\mathrm{BV}}(0,T;X)$ are both solutions to the sweeping process with data u, r, x_0 and v, s, y_0 respectively. Let $t_0 \in [0,T]$ and define

$$\begin{aligned} z(t) \; &:= \; \begin{cases} Q_{Z(r(t+)}(v(t+) - \eta(t+)) & t < t_0 \\ Q_{Z(r(t_0)}(v(t_0) - \eta(t_0)) & t = t_0 \end{cases} \quad \text{and} \\ \widehat{z}(t) \; &:= \; \begin{cases} Q_{Z(s(t+)}(u(t+) - \xi(t+)) & t < t_0 \\ Q_{Z(s(t_0)}(u(t_0) - \eta(t_0)) & t = t_0 \end{cases} \end{aligned}$$

Notice that z and $\widehat{z}$ are admissible test functions for (6.7) for ξ and η respectively. Indeed $z(t) \in Z(r(t+))$ for all $t \in [0, t_0]$ by construction. Furthermore notice that for any sequences $z^n \to z$, $r^n \to r$ it holds $Q_{Z(r^n)}(z^n) \to Q_{Z(r)}(z)$. Hence z is a regulated function. Analogous arguments can be applied to $\widehat{z}$. By testing (6.7) with these two functions we obtain

$$\int_0^{t_0} \langle u(t+) - \xi(t+) - \widehat{z}(t),\, \mathrm{d}\xi(t)\rangle \geq 0 \;\text{ and }\; \int_0^{t_0} \langle v(t+) - \eta(t+) - z(t),\, \mathrm{d}\eta(t)\rangle \geq 0\,.$$

Therefore we can deduce

$$\begin{aligned} &\int_0^{t_0} \langle \xi(t+) - \eta(t+),\, \mathrm{d}(\xi - \eta)(t)\rangle \\ &\leq \int_0^{t_0} \langle u(t+) - v(t+),\, \mathrm{d}\xi(t)\rangle + \int_0^{t_0} \langle v(t+) - \eta(t+) - \widehat{z}(t),\, \mathrm{d}\xi(t)\rangle \\ &\qquad + \int_0^{t_0} \langle v(t+) - u(t+),\, \mathrm{d}\eta(t)\rangle + \int_0^{t_0} \langle z(t) - (u(t+) - \xi(t+))\,,\, \mathrm{d}\eta(t)\rangle \\ &\leq \left(\overline{\mathrm{Var}}(\xi) + \overline{\mathrm{Var}}(\eta)\right) \left(\|u - v\|_\infty + \sup_{t\in[0,t_0]} d_H\left(Z(r(t), Z(s(t))\right) \right) \\ &\leq \left(\overline{\mathrm{Var}}(\xi) + \overline{\mathrm{Var}}(\eta)\right) \left(\|u - v\|_\infty + L\,\|r - s\|_\infty\right) \end{aligned}$$

We hence have

$$\frac{1}{2}\|\xi - \eta\|_\infty^2 \leq \left(\overline{\mathrm{Var}}(\xi) + \overline{\mathrm{Var}}(\eta)\right) \left(\|u - v\|_\infty + L\,\|r - s\|_\infty\right) + |x_0 - y_0| + |u(0) - v(0)|\,.$$

This indeed proves local Hölder continuity of the sweeping process - and thereby uniqueness - should a solution exist.

Step 2 - Existence of a solution for stepfunctions: Let $u \in S(0,T;X)$ and $r \in S(0,T;X)$ be stepfunctions with respect to the same partition $(t_i))_{i=0}^n$, i.e. there exist $(u_i)_{i=1}^n \subset X$,

$(\widehat{u}_i)_{i=0}^n \subset X$ and $(r_i)_{i=1}^n \subset \mathcal{R}$, $(\widehat{r}_i)_{i=0}^n \subset \mathcal{R}$ respectively such that

$$\begin{aligned} u(t) &= \sum_{i=0}^{N} \widehat{u}_i \chi_{\{t_i\}}(t) + \sum_{i=1}^{N} u_i \chi_{(t_{i-1},t_i)}(t) \quad \text{and} \\ r(t) &= \sum_{i=0}^{N} \widehat{r}_i \chi_{\{t_i\}}(t) + \sum_{i=1}^{N} r_i \chi_{(t_{i-1},t_i)}(t)\,. \end{aligned}$$

Then the unique solution to the sweeping process is also of the form

$$\xi(t) = \sum_{i=0}^{N} \widehat{\xi}_i \chi_{\{t_i\}}(t) + \sum_{i=1}^{N} \xi_i \chi_{(t_{i-1},t_i)}(t)$$

where $\widehat{\xi_0} = \widehat{u}_0 - x_0$ and

$$\left.\begin{array}{rl} & u_1 - \xi_1 = Q_{Z(r_1)}\left(u_1 - \widehat{\xi}_0\right) \;\wedge\; \xi_1 - \widehat{\xi}_0 = P_{Z(r_1)}\left(u_1 - \widehat{\xi}_0\right) \\ \forall 1 < k \le n: & u_k - \xi_k = Q_{Z(r_k)}\left(u_k - \xi_{k-1}\right) \;\wedge\; \xi_k - \xi_{k-1} = P_{Z(r_k)}\left(u_k - \xi_{k-1}\right) \\ \forall 1 \le k \le n: & \widehat{u}_k - \widehat{\xi}_k = Q_{Z(\widehat{r}_k)}\left(\widehat{u}_k - \xi_k\right) \;\wedge\; \widehat{\xi}_k - \xi_k = P_{Z(\widehat{r}_k)}\left(\widehat{u}_k - \xi_k\right) \end{array}\right\} \tag{6.9}$$

Indeed let $s \in (0,T]$ and $y \in G(0,T;X)$ such that $y(t) \in Z(r(t+))$. Remember that by convention for $t = s$ this implies $y(s) \in Z(r(s))$. By applying the calculus of the Kurzweil integral we obtain that

$$\begin{aligned} &\int_0^s \langle u(t+) - \xi(t+) - y(t), d\xi(t)\rangle \\ &\quad = \left\langle u_1 - \xi_1 - y(0+), \xi_1 - \widehat{\xi}_0 \right\rangle + \sum_{k\in[n]:t_k<t} \langle u_{k+1} - \xi_{k+1} - y(t_k+), \xi_{k+1} - \xi_k\rangle \\ &\qquad + \langle u(s) - \xi(s) - y(s), \xi(s) - \xi(s-)\rangle \end{aligned}$$

The first two terms are greater or equal zero due to (6.9). If $s \neq t_k$ for all $k \in [n]$, then the last term is zero, since $\xi(s) = \xi(s-)$. Otherwise if there exists an $k \in [n]$ such that $s = t_k$ then we obtain

$$\langle u(s) - \xi(s) - y(s), \xi(s) - \xi(s-)\rangle = \left\langle \widehat{u}_k - \widehat{\xi}_k - y(t_k), \widehat{\xi}_k - \xi_k \right\rangle$$

which is also non negative. Therefore ξ is the, by the first step, unique solution to the sweeping process with input u and r. Now one can easily calculate that

$$|\xi_k - \xi_{k-1}| \le |u_k - u_{k-1}| + d_H\left(Z(r_k), Z(r_{k-1})\right) \le 2\,|u_k - u_{k-1}| + L\,\|r_k - r_{k-1}\|_{\mathcal{R}} \tag{6.10}$$

We therefore can estimate

$$\overline{\mathrm{Var}}(\xi) = \sum_{k=2}^{n} |\xi_k - \xi_{k-1}| \le \overline{\mathrm{Var}}(u) + L\overline{\mathrm{Var}}(r)\,. \tag{6.11}$$

Step 3 - Existence on $\overline{\mathrm{BV}}$: Let $u \in \overline{\mathrm{BV}}(0,T;X)$, $r \in \overline{\mathrm{BV}}(0,T;\mathcal{R})$ and $x_0 \in Z(r(0))$. Then due to Theorem 3.10 there exist sequence $(u^n)_{n\in\mathbb{N}} \subset S(0,T;X)$ and $(r^n)_{n\in\mathbb{N}} \subset S(0,T;\mathcal{R})$ such that

$$\overline{\mathrm{Var}}(f^n) \le \overline{\mathrm{Var}}(f) \;\text{ and }\; \|f^n - f\|_\infty \to 0\,,$$

for $f \in \{r, s\}$. Without loss of generality we may assume that $r^n(0) = r(0)$ and $u^n(0) = u(0)$. Let ξ^n be the solution of the sweeping process with data u^n, r^n, x_0. Then we know that

$$\overline{\mathrm{Var}}(\xi^n) \leq \overline{\mathrm{Var}}(u^n) + C\overline{\mathrm{Var}}(r^n) \leq \overline{\mathrm{Var}}(u) + C\overline{\mathrm{Var}}(r)\,. \tag{6.12}$$

Since $(r^n)_{n\in\mathbb{N}}$ and $(u^n)_{n\in\mathbb{N}}$ are Cauchy sequences with respect to $\|\cdot\|_\infty$. Due to the first step so is $(\xi_n)_{n\in\mathbb{N}}$. Hence there exists a function ξ such that $\|\xi^n - \xi\|_{n\in\mathbb{N}} \to 0$. In view of (6.12) we obtain that $\xi \in \overline{\mathrm{BV}}(0, T; X)$ and ξ satisfies (6.11). Notice that $u^n(t) - \xi^n(t) \in Z(r^n(t))$ and the strong convergence of $\xi^n(t)$, $u^n(t)$ and $r^n(t)$ implies $u(t) - \xi(t) \in Z(r(t))$. It remains to show that the variational inequality (6.7) is satisfied. Therefore let $s \in [0, T]$ and choose $y \in G(0, s; X)$ such that for all $t \in [0, s)$ it holds $y(t) \in Z(r(t+))$ and $y(s) \in Z(r(s))$. Define test functions y^n by

$$y^n(t) := \begin{cases} Q_{Z(r^n(t+))}(y(t)) & t < s \\ Q_{Z(r^n(s))}(y(s)) & t = s\,. \end{cases}$$

For all $t \in [0, s)$ we can estimate

$$|y^n(t) - y(t)| \leq d_H\left(Z\left(r^n(t+)\right), Z\left(r(t+)\right)\right) \leq C\|r^n - r\|_\infty\,.$$

A similar estimate can be made for $t = s$. We therefore obtain $\|y^n - y\|_\infty \to 0$. Arguing as in Step 1 we furthermore obtain $y^n \in G(0, T; X)$. Hence y^n is an admissible testfunction and (6.7) holds with u^n, ξ^n and r^n. Due to the continuity properties of the Kurzweil integral, see Theorem 3.36 , we may pass to the limit and obtain that (6.7) also holds true for y, u, ξ and r. Hence ξ is an admissible solution to Problem 6.4. □

Remark 6.6. In case we allow for u, r and ξ to be left (or right) continuous only condition (6.7) can be simplified. It is then equivalent to ask for the inequality to hold for $s = T$.

6.3. Quasivariational sweeping processes

Quasivariational sweeping processes extend the sweeping processes in the following way. The shape of the convex set K does now no longer depend only on time t but also on the current state $\xi(t)$. Figuratively spoken: Put a small stone on a table, turn a cake form upside down and put it over the stone. Then if the cake form is moved, the stone is moving as well. This is the sweeping process. Now for a quasivariational sweeping process the cake form is no longer a classical one but made out of silicon. Then if the stone is heavy enough the shape of the silicon cake form depends on where the stone is. Formally the problem can be written as follows

Problem 6.7 (Quasivariational sweeping process). *Let $K : (t, \xi) \mapsto K(t, \xi) \subset X$ where $K(t, \xi)$ is convex for all $(t, \xi) \in [0, T] \times X$. For an initial value $\xi_0 \in K(0, \xi_0)$, does there exist an absolutely continuous function ξ such that $\xi(0) = \xi_0$ and*

$$-\dot{\xi}(t) \in \partial I_{K(t,\xi(t)))}(\xi) \ \text{ a.e. in } [0, T] \qquad \text{holds?}$$

This problem was first proposed by M. Kunze and M. D. P. Monteiro Marquez in [64]. Therein they also provided a first prove of existence of a solution under fairly general assumptions. Uniqueness of the solution however is more involved. Here the first results are due to M. Brokate, P. Krejčí and H. Schnabel in [18] in the case of smooth and bounded convex sets. Here smooth has to be understood in such a way that roughly speaking the outer normals depend Lipschitz continuous on both the position and the current state. This result was later on generalized to smooth and unbounded convex sets by A. Mielke and R. Rossi [71]. Also due to R. Rossi and U. Stefanelli are uniqueness results which do not make use of the smoothness of a solution. Instead in [95, 100] a certain order principle for the convex sets was assumed. In the former of the two papers up to the authors knowledge the term *'quasivariational sweeping process'* was used for the first time. It is probably derived from the seminal book [7] where the term *'quasivariational inequality'* was coined to describe problems of the type

$$\text{find } u \text{ such that } \langle u, v\rangle \leq f(v) \text{ for all } v \in K(u)\,.$$

As in the case of the sweeping process there are several possibilities to extend the quasivariational sweeping process to functions of (essentially) bounded variation. We are going to employ the Kurzweil formulation. To make this precise let us shortly formulate the problem in the setting we want to analyze it.

Problem 6.8 (Quasivariational sweeping process on $\overline{\mathrm{BV}}$)**.** *Consider a family $Z(r) \subset X$ of closed convex sets parameterized by elements r of a reflexive Banach space $\mathcal{R}$. Assume that $u \in \overline{\mathrm{BV}}(0,T;X)$, $g \in \overline{\mathrm{BV}}(0,T;C^1(X\times X;\mathcal{R}))$ and $x_0 \in Z(g(0,u(0),x(0)-u(0)))$ are given. We look for a function $\xi \in \overline{\mathrm{BV}}(0,T;X)$ such that*

$$x(t) \;:=\; u(t)-\xi(t) \in Z(g(t,u(t),\xi(t))) \qquad \forall t\in[0,T]\,, \tag{6.13}$$

$$x(0) = x_0\,, \tag{6.14}$$

$$\int_0^T \langle x(t+)-y(t),\, \mathrm{d}\xi(t)\rangle \;\geq\; 0 \; \textit{for every } y \in \mathcal{T}(\xi)\,, \tag{6.15}$$

where

$$\mathcal{T}(\xi) := \{y \in G(0,T;X) : y(t) \in Z(g(t+,u(t+),\xi(t+)))\ \forall t\in[0,T]\}$$

is the set of all admissible testfunctions.

Our aim is to prove existence and uniqueness of a solution to the above problem. However we are not able to treat any family $Z(r)$ of convex sets. Therefore we have to restrict ourselves to two cases. In the upcoming Chapter 7 we are going to study the case where the convex set $Z(r)$ is a polyhedron. Chapter 8 is dedicated to the analysis of Problem 6.8 where $Z(r)$ is assumed to be smooth. The precise conditions are going to be stated below.

In both cases we are able to prove existence and uniqueness of a solution (under some restrictions). Indeed the rough strategy is the same, we make use of Banach's contraction principle. However the techniques employed to deduce the desired estimates are different and depend very much on the shape of the convex sets.

7. Quasivariational sweeping processes with polyhedral characteristics

In this chapter we consider quasivariational sweeping processes where the convex sets $Z(r)$ are polyhedra. Remember that a polyhedron is the intersection of finitely many subspaces. In other words P is a polyhedron, if there exist $m \in \mathbb{N}$, $(a_i)_{i=1}^m \subset X$ and $\beta \in \mathbb{R}^m$ such that

$$P = \{x \in X : \langle a_i, x\rangle \leq \beta_i\} \ .$$

One reason to study quasivariational sweeping processes with polyhedral characteristics is that the play operator with polyhedral characteristics enjoys remarkably good continuity properties. It is globally Lipschitz continuous with respect to both the norm on $W^{1,1}$ and the $\| \cdot \|_\infty$ (see [83, 55, 31]). Also the Skokhorod process, a generalization of the sweeping process, has been studied for polyhedra and global Lipschitz continuity again with respect to both the above norms has been obtained [57, 53]. In this case however the dependence of the convex set on the parameter r is of a very special type. The faces of the convex sets are only allowed to move along their normal direction but may not be rotated. We also will stick to this assumption and provide existence of a unique solution to the quasivariational sweeping process. However we will not use the methods introduced in the latter paper. Instead we will use the approach due to I. Picek [83] for the play operator and generalize it in order to show global Lipschitz continuity of the sweeping process. We shall then use this result to provide existence and uniqueness of a solution by a contraction argument.

This chapter is organized as follows. We are first going to fix notation and assumptions and introduce the main results in the upcoming section. Section 7.2 will be dedicated to proving the global Lipschitz continuity of the sweeping process. Thereafter we are going to prove existence and uniqueness of the quasivariational sweeping process in Section 7.3. We shall also demonstrate the sharpness of our conditions in dimensions less or equal then two.

7.1. Main result

Let us start by fixing our assumptions on the convex set.

Hypothesis 7.1. *There exist $m \in \mathbb{N}$, $(a_i)_{i=1}^m \subset X$ with $|a_i| = 1$ for all $i \in [m]$ and a map $\beta : \mathcal{R} \to \mathbb{R}^m$ such that for all $r \in \mathcal{R}$*

$$Z(r) = \{x \in X : \langle a_i, x\rangle \leq \beta_i(r)\} \ . \tag{7.1}$$

Furthermore we assume that $X = \text{span}\{a_i : i \in [m]\}$.

As indicated before this assumption implies that the faces of the polyhedron can move only along their normal directions but are not allowed to rotate. The last assumption, namely that we require that the outer normals are a generating system of X, seems to be very restrictive. It appears as if we were only able to handle finite dimensional Hilbert spaces. However this is not the case. If $X \neq \operatorname{span}\{a_i : i \in [m]\}$ we decompose $X = Y \oplus Y^{\perp}$, where $Y = \operatorname{span}\{a_i : i \in [m]\}$. Then for $x \in X$ there exist unique $\widehat{x}, \widehat{x}^{\perp} \in X$ such that

$$\widehat{x} \in Y, \widehat{x}^{\perp} \in Y^{\perp} : \; x = \widehat{x} + \widehat{x}^{\perp}.$$

It is then immediate that if ξ is a solution to the sweeping process $\left(\widehat{\xi}(t)\right)^{\perp} \equiv \widehat{\xi}_0^{\perp}$. Assuming that ξ solves the quasivariational sweeping process. By definition it solves the sweeping process with $r(t) = g(t, u(t), \xi(t))$. Hence by rewriting

$$g\big(t, u(t), \xi(t)\big) = \widetilde{g}\left(t, \widehat{u}(t), \widehat{u}^{\perp}(t), \widehat{\xi}(t), \widehat{\xi}^{\perp}(t)\right) = \widehat{g}\left(t, \widehat{u}(t), \widehat{\xi}(t)\right)$$

we reduce the problem to solving the quasivariational sweeping process for $\widehat{\xi}$ with input data $\widehat{g}$, $\widehat{u}$ and $\widehat{\xi}_0$ on Y. In general, if conditions are imposed on $\widehat{g}$, we subsequently need to check whether any conditions on $\widehat{u}^{\perp}$ and eventually on u need to be made. As we will see below, in the present case, we do not need to do so. Before we are able to start we need a further definition.

Definition 7.2. Assume Hypothesis 7.1 holds. Set $N = \dim(X)$ and $\mathcal{A}_0 = \{\{0\}\} \subset 2^X$. For $k \in [N]$ we define $\mathcal{A}_k \subset 2^X$ by

$$\mathcal{A}_k := \big\{X' = \operatorname{span}\{a_{i_j} : j \in [n]\} : \; n \in \mathbb{N}, \; \forall j \in [r] : i_j \in [m], \; \dim(X') = k\big\} \tag{7.2}$$

Here m is the number of normal vectors defining $Z(r)$, see Hypothesis 7.1. $\mathcal{A}_k$ in other words is the set of all k dimensional planes in X, which can be generated from the family of normal vectors $(a_i)_{i=1}^m$. Furthermore set $M_0 := 0$ and define for all $k \in [N]$

$$\varepsilon_k := \max\{|Q_{X'} a_j| : \; X' \in \mathcal{A}_k, \; a_j \notin X'\} \tag{7.3}$$

and

$$M_k := \left(\frac{1}{1 - \varepsilon_{k-1}^2}\left(1 + M_{k-1}^2 + 2\varepsilon_{k-1} M_{k-1}\right)\right)^{1/2}. \tag{7.4}$$

Now everything is in place to state the main result of this chapter.

Theorem 7.3 (Existence and uniqueness)**.** *Let Hypothesis 7.1 hold and let M_N be defined as in Definition 7.2. Assume that $\beta : \mathcal{R} \to \mathbb{R}^m$ is Lipschitz continuous with Lipschitz constant C. Let $g \in \overline{\mathrm{BV}}(0, T; C^{0,1}_{\mu,\lambda}(X \times X; \mathbb{R}^m_{\geq 0}))$, $u \in \overline{\mathrm{BV}}(0, T; X)$ and $x_0 \in Z(g(0, u(0), u(0) - x_0))$ with*

$$C M_N \lambda =: \delta < 1\,. \tag{7.5}$$

Then there exists a unique solution ξ of Problem 6.8.

We will show that (7.5) is sharp for $dim(X) \leq 2$. The main tools of the proof are Banach's contraction principle and the global Lipschitz continuity of the sweeping process.

Theorem 7.4 (Global Lipschitz continuity)**.** *Let Hypothesis 7.1 hold and let M_N be defined as in Definition 7.2. Assume that $\beta : \mathcal{R} \to \mathbb{R}^m$ is Lipschitz continuous in the sense that*

$$|\beta(r) - \beta(s)|_\infty \le C\|r - s\|_{\mathcal{R}} \,. \tag{7.6}$$

Then for all $r \in \overline{\mathrm{BV}}(0,T;\mathcal{R})$, $u \in \overline{\mathrm{BV}}(0,T;X)$ and $x_0 \in Z(r(0))$ there exists a unique solution $\xi \in \overline{\mathrm{BV}}(0,T;X)$ to the sweeping process. Furthermore let $r, s \in \overline{\mathrm{BV}}(0,T;\mathcal{R})$, $u, v \in \overline{\mathrm{BV}}(0,T;X)$, $x_0 \in Z(r(0))$ and $y_0 \in Z(s(0))$. Let $\xi, \eta \in \overline{\mathrm{BV}}(0,T;X)$ be solutions to the sweeping process with input u, r, x_0 and v, s, y_0 respectively. Then

$$\|\xi - \eta\|_\infty \le M_N \left(\|u - v\|_\infty + C\|r - s\|_\infty + |x_0 - y_0|\right) \,. \tag{7.7}$$

The fact that the sweeping process is globally Lipschitz continuous is not new. It has already been shown for an even more general problem in [57, 53]. However there the Lipschitz constant could not be explicitly calculated which we can do here by using Definition 7.2.

7.2. Global Lipschitz continuity of the polyhedral sweeping process

Let $Z(r)$ satisfy Hypothesis 7.1. Let us just shortly remember that as in Lemma 2.7 we define the set of all active constraints at some point $x \in Z(r)$ by

$$\Gamma_{Z(r)}(x) := \left\{ i \in [m] : \langle a_i, x\rangle_{X^*,X} = \beta_i(r) \right\} \,.$$

We start our path to proving Theorem 7.4 by making a simple observation.

Proposition 7.5. *Let the assumption of Theorem 7.4 hold. For all $r, s \in \mathcal{R}$, $x = u - \xi \in Z(r)$ and $y = v - \eta \in Z(s)$ it holds that*

$$\forall i \in \Gamma_{Z(r)}(x) : \ \langle a_i, \xi - \eta\rangle \le |u - v| + C\|r - s\|_{\mathcal{R}} \tag{7.8}$$

Proof. Notice that $y \in Z(s)$ directly implies $\langle y - (\beta_i(r) - \beta_i(s))a_i, a_i\rangle \le \beta_i(r)$. Therefore

$$\langle x - (y - (\beta_i(r) - \beta_i(s))a_i)\,, a_i\rangle \ge 0 \Rightarrow \langle \xi - \eta, a_i\rangle \le \langle u - v, a_i\rangle + \beta_i(r) - \beta_i(s) \,.$$

The proof is completed by remembering that $|a_i| = 1$ and β is Lipschitz continuous with constant C. □

Now let u, v, r, s be step functions on the same partition $(t_k)_{k=0}^n \in \mathcal{D}_{[0,T]}$. In other words we assume that for there exist $(u_k)_{k=1}^n\,, (v_k)_{k=1}^n \subset X$, $(r_k)_{k=1}^n\,, (s_k)_{k=1}^n \subset \mathcal{R}$ and $(\widehat{u}_k)_{k=0}^n\,, (\widehat{v}_k)_{k=0}^n \subset X$, $(\widehat{r}_k)_{k=0}^n\,, (\widehat{s}_k)_{k=0}^n \subset \mathcal{R}$ such that

$$f(t) = \sum_{k=1}^n f_k \chi_{(t_{k-1},t_k)}(t) + \sum_{k=0}^n \widehat{f}_k \chi_{\{t_k\}}(t) \,, \tag{7.9}$$

where f stands for u, v, r, s. Let ξ, η be solutions to the sweeping process with input u, r, x_0 and v, s, y_0 respectively. Then also ξ and η are step functions with respect to the same partition $(t_k)_{k=0}^n$ and can be denoted in the fashion of (7.9). Furthermore it holds that $\widehat{\xi}_0 = \widehat{u}_0 - x_0$ and

$$\left.\begin{array}{rl} & u_1 - \xi_1 = Q_{Z(r_1)}\left(u_1 - \widehat{\xi}_0\right) \ \wedge\ \xi_1 - \widehat{\xi}_0 = P_{Z(r_1)}\left(u_1 - \widehat{\xi}_0\right) \\ \forall 1 < k \leq n: & u_k - \xi_k = Q_{Z(r_k)}\left(u_k - \xi_{k-1}\right) \ \wedge\ \xi_k - \xi_{k-1} = P_{Z(r_k)}\left(u_k - \xi_{k-1}\right) \\ \forall 1 \leq k \leq n: & \widehat{u}_k - \widehat{\xi}_k = Q_{Z(\widehat{r}_k)}\left(\widehat{u}_k - \xi_k\right) \ \wedge\ \widehat{\xi}_k - \xi_k = P_{Z(\widehat{r}_k)}\left(\widehat{u}_k - \xi_k\right) \end{array}\right\} \tag{7.10}$$

Respective equations hold for η. We now introduce the key element of our proof.

Definition 7.6. For any $c > \|u - v\|_\infty + C\|r - s\|_\infty + |x_0 - y_0|$ we define

$$V_c(g) := \max\left\{M_k^2 c^2 + |P_{X'}(g)| : X' \in \mathcal{A}_k, k \in [N] \cup \{0\}\right\} \tag{7.11}$$

where M_k and $\mathcal{A}_k$ are defined in Definition 7.2.

Notice that due to Proposition 2.13, we have $V_c(-g) = V_c(g)$. In a certain sense the function V_c shall play the role of a Lyapunov function for the difference $\xi - \eta$. However we cannot assure that V_c is monotonic decreasing at all points but need to use a more subtle notion. We will make this precise in the following Lemma, whose proof is the main effort in providing the Lipschitz continuity result.

Lemma 7.7. *For any $c > \|u-v\|_\infty + C\|r-s\|_\infty + |x_0 - y_0|$ the function $V_c(\xi - \eta)$ is monotonic decreasing in the sense that for $g = \xi - \eta$ it holds that $V_c(g_1) \leq V_c(\widehat{g}_0)$ and*

$$\forall i \in [n] : V_c(g_i) \leq V_c(g_{i-1}) \ \wedge\ V_c(\widehat{g}_i) \leq V_c(g_i)\ .$$

Proof. The proof heavily relies on the properties of the projection onto linear subspaces, which were assembled in Proposition 2.13. Our strategy is a proof by contradiction. If the assertion is wrong then there exists $c > \|u - v\|_\infty + C\|r - s\|_\infty + |x_0 - y_0|$ such that $V_c(g_1) > V_c(\widehat{g}_0)$ or

$$\exists i \in [n] : V_c(g_i) > V_c(g_{i-1}) \ \vee\ V_c(\widehat{g}_i) > V_c(g_i)\ .$$

We will assume that

$$V_c(g_i) > V_c(g_{i-1})\ . \tag{7.12}$$

For the other cases the proof works analogously. To simplify the notation we will drop the index c and just write V. Furthermore we shall denote by Δf the term $f_i - f_{i-1}$ whenever feasible. Now remember that $\mathcal{A}_N = \{X\}$ and $P_X = 0$. Hence $V(g_{i-1}) \geq M_N^2 c^2$ and $V(g_i) = M_k^2 c^2 + |P_{X''} g_i|$ for some $k \in [N-1] \cup \{0\}$ and $X'' \in \mathcal{A}_k$. Especially for g_i the maximum in (7.11) is not attained for $k = N$. Furthermore $|P_{X''}(g_i)| > |P_{X''}(g_{i-1})|$. Our aim now is to show that

$$M_k^2 c^2 + |P_{X''}(g_i)| < M_0^2 c^2 + \left|P_{\{0\}}(g_i)\right|^2 \quad \text{holds.} \tag{7.13}$$

This constitutes a contradiction. Indeed it is plainly impossible for $k = 0$. However if $k \neq 0$ then $V(g_i) < M_0^2 c^2 + |P_0(g_i)|^2$, which cannot be due to (7.11). To start we observe that

$$\langle \Delta g, g_i \rangle \geq |g_i|^2 - |g_i|\,|g_{i-1}| \geq \frac{1}{2}\left(|g_i|^2 - |g_{i-1}|^2\right)\ . \tag{7.14}$$

Furthermore remember that X'' is a linear subspace and therefore also $-Q_{X''}(g_i), -Q_{X''}(g_{i-1})$ and their linear combinations are elements of X''. Therefore by definition of the projection we obtain

$$\begin{aligned} 0 &\geq \langle P_{X''}(g_i), Q_{X''}(g_i) - 2Q_{X''}(g_i)\rangle \quad \text{and} \\ 0 &\geq \langle P_{X''}(g_{i-1}), Q_{X''}(g_{i-1}) - (Q_{X''}(g_{i-1}) + Q_{X''}(g_i))\rangle . \end{aligned}$$

Adding both inequalities we have

$$\langle P_{X''}(g_i) - P_{X''}(g_{i-1}), -Q_{X''}(g_i)\rangle \geq 0 .$$

Using $P_{X''} = \mathrm{Id} - Q_{X''}$ we conclude

$$\langle \Delta g, -Q_{X''}(g_i)\rangle \geq \frac{1}{2}\left(|Q_{X''}(g_i)|^2 - |Q_{X''}(g_{i-1})|^2\right) . \tag{7.15}$$

Combining this equation with (7.14) we obtain, by using Proposition 2.13 (ii), that

$$\langle \Delta g, P_{X''}(g_i)\rangle \geq \frac{1}{2}\left(|P_{X''}(g_i)|^2 - |P_{X''}(g_{i-1})|^2\right) > 0 . \tag{7.16}$$

We now claim that either $\langle \Delta\xi, P_{X''}(g_i)\rangle > 0$ or $\langle \Delta\eta, P_{X''}(-g_i)\rangle > 0$. Assume that the second assertion is wrong that is $0 \geq \langle \Delta\eta, P_{X''}(-g_i)\rangle = \langle -\Delta\eta, P_{X''}(g_i)\rangle$. Applying this to (7.16) we obtain that $\langle \Delta\xi, P_{X''}(g_i)\rangle > 0$.

Without loss of generality we shall assume that $\langle \Delta\xi, P_{X''}(g_i)\rangle > 0$. For the other case it suffices to interchange ξ and η. Remember that $\Delta\xi = P_{Z(r_i)}(\xi_{i-1} + u_i)$ and consequently $\Delta\xi \in \partial I_{Z(r_i)}(x_i)$, where x once again denotes the term $u - \xi$. Due to Lemma 2.7 there exists for every $l \in \Gamma_{Z(r_i)}(x)$ some $\lambda_l \geq 0$ such that

$$\Delta\xi = \sum_{l \in \Gamma_{Z(r_i)}(x)} \lambda_l a_l .$$

Therefore we have that

$$\sum_{l \in \Gamma_{Z(r_i)}(x)} \lambda_l \langle a_l, P_{X''}(g_i)\rangle > 0$$

and thus there exists at least one $j \in \Gamma_{Z(r_i)}(x)$ such that $\langle a_j, P_{X''}(g_i)\rangle > 0$.

Notice that $a_j \notin X''$ since if $a_j \in X''$ then also $Q_{X''}(g_i) \pm n_j \in X''$. Testing the variational inequality defining the projection with both leads to $\langle a_j, P_{X''}(g_i)\rangle = 0$, a contradiction. We define $X' := X'' \oplus \mathrm{span}\{a_j\}$. There exist $\rho, \eta \in \mathbb{R}$ and $v \in X''$ such that $|v| = 1$ and

$$Q_{X'}(g_i) = \rho v + \nu a_j . \tag{7.17}$$

Employing Proposition 2.13 (i) we have

$$|Q_{X''}(g_i)| = |Q_{X''}(Q_{X'}(g_i))| \geq |\langle Q_{X'}(g_i), v\rangle| = |a\varepsilon + b| , \tag{7.18}$$

where $\varepsilon = \langle v, n_j\rangle$. Due to (7.3) we furthermore know that $|\varepsilon| \leq \varepsilon_k$. Moreover due to Proposition 7.5 we have

$$c > \langle a_j, g_i\rangle = \langle a_j, Q_{X'}(g_i)\rangle = \rho + \nu\varepsilon . \tag{7.19}$$

Here the first equality is due to the fact that $a_j \in X'$ and Proposition 2.13 (v), whereas the second is owed to (7.17). On the other hand by assumption we have

$$\langle a_j, g_i \rangle > \langle a_j, Q_{X''}(g_i) \rangle = \langle a_j, Q_{X''}(Q_{X'}(g_i)) \rangle = \rho \langle a_j, Q_{X''}(a_j) \rangle + \nu\varepsilon, \tag{7.20}$$

which by combination with the previous equation implies that $\rho \langle a_j, a_j - Q_{X''}(a_j) \rangle > 0$. Since $\langle a_j, a_j - Q_{X''}(a_j) \rangle \geq \left(|a_j|^2 - |Q_{X''}(a_j)|^2 \right)/2 > 0$ we derive that $\rho > 0$. We now employ (7.18) and (7.19) to obtain

$$\rho - \rho\varepsilon^2 < c - \nu\varepsilon - \rho\varepsilon^2 \leq c + |\varepsilon||b + a\varepsilon| \leq c + |\varepsilon| \left| Q_{X''}(g_m) \right| . \tag{7.21}$$

By assumption and using Proposition 2.13 (ii) we deduce

$$M_{k+1}^2 c^2 - |Q_{X'}(g_i)| \leq M_k^2 c^2 - |Q_{X''}(g_i)| \, .$$

Furthermore we can estimate that

$$\begin{aligned}
|Q_{X'}(g_i)| &= \rho^2 + 2\rho\nu\varepsilon + \nu^2 \\
&= (\rho\varepsilon + \nu)^2 + \rho^2 \left(1 - \varepsilon^2\right) \\
&\leq |Q_{X''}(g_i)| + a\left(c + |\varepsilon| \, |Q_{X''}(g_m)|\right) \\
&\leq |Q_{X''}(g_i)| + \frac{1}{1-\varepsilon}\left(c + |\varepsilon| \, |Q_{X''}(g_m)|\right)^2 \\
&= \frac{1}{1-\varepsilon}\left(c^2 + 2c|\varepsilon| \, |Q_{X''}(g_m)| + |Q_{X''}(g_m)|^2\right) \\
&\leq \frac{1}{1-\varepsilon_k}\left(c^2 + 2c\varepsilon_k \, |Q_{X''}(g_m)| + |Q_{X''}(g_m)|^2\right) .
\end{aligned}$$

By putting both above inequalities together we obtain

$$M_{k+1}^2 c^2 - M_k^2 c^2 + |Q_{X''}(g_i)| \leq \frac{1}{1-\varepsilon_k}\left(c^2 + 2c\varepsilon_k \, |Q_{X''}(g_m)| + |Q_{X''}(g_m)|^2\right) . \tag{7.22}$$

Applying the definition of M_{k+1} we see that

$$\varepsilon_k^2 M_k^2 c^2 + 2\varepsilon_k M_k c^2 - \varepsilon_k^2 \, |Q_{X''}(g_i)|^2 \leq 2c\varepsilon_k \, |Q_{X''}(g_i)| \, .$$

By regrouping the terms and multiplying with ε_k^{-1} we then derive

$$\varepsilon_k \left(M_k^2 c^2 - |Q_{X''}(g_i)|^2 \right) + c\left(M_k r - |Q_{X''}(g_i)|\right) < 0 \, . \tag{7.23}$$

Since both ε_k and c are positive this can only be satisfied if $M_k c < |Q_{X''}(g_i)|$ and consequently $\left(M_k^2 c^2 - |Q_{X''}(g_i)|^2 \right) < 0$. Employing Proposition 2.13 (ii) for a last time gives us

$$\begin{aligned}
0 &> M_k^2 c^2 - |Q_{X''}(g_i)|^2 \\
&= M_k^2 c^2 + |P_{X''}(g_i)|^2 - \left(M_0^2 c^2 - \left| P_{\{0\}}(g_i) \right|^2 \right) ,
\end{aligned}$$

where last equality is due to the fact that $P_{\{0\}} = \mathrm{Id}$ and $M_0 = 0$. As explained at the beginning this is a contradiction. Hence our assumption is wrong and consequently $V(g_i) \leq V(g_{i-1})$. □

We are now able to show our main result of this section.

Proof of Theorem 7.4. The proof is divided into three parts. We first use the Lemma 7.7 to establish global Lipschitz continuity for step functions. We then show existence and uniqueness of a solution on the whole space $\overline{\mathrm{BV}}$ by applying Theorem 6.5. Finally we extend the Lipschitz continuity result to the whole space of functions of essentially bounded variation by a density argument.

Step 1 - Lipschitz estimate for step functions: Let $u, v \in S(0,T;X)$, $r, s \in S(0,T;\mathcal{R})$, $x_0 \in Z(r(0))$ and $y_0 \in Z(s(0))$. Furthermore let $\xi \in S(0,T;X)$ be the solution of the sweeping process with input u, r, x_0 and η the solution with input v, s, y_0. Without loss of generality we can assume that there exists a partition $(t_n)_{n=0}^{N}$ such that all functions can be written in the fashion of (7.9) by using the union of all partitions of the input functions. Now due to Lemma 7.7 we know that for every $t \in [0,T]$ and every $c > \|u-v\|_\infty + C\|r-s\|_\infty + |x_0 - y_0|$ we have

$$V_r\left(\xi(t) - \eta(t)\right) \leq V_c\left(\xi(0) - \eta(0)\right) .$$

From (7.11) we derive that $|\xi(t) - \eta(t)|^2 \leq V_c\left(\xi(t) - \eta(t)\right)$; set $k = 0$. On the other hand since $|\xi(0) - \eta(0)| \leq |u(0) - v(0)| + |x_0 - y_0| < c$ we have that for all $k \in [N-1]$ and $X' \in \mathcal{A}_k$

$$M_k^2 c^2 + |P_{X'}\left(\xi(0) - \eta(0)\right)| \leq \left(M_k^2 + 1\right) c^2 < M_{k+1}^2 c^2 .$$

Hence $V_c(\xi(0) - \eta(0)) = M_N^2 c^2$ and we therefore obtain

$$|\xi(t) - \eta(t)| \leq M_N c .$$

Since this holds for all $t \in [0,T]$ and $c > \|u-v\|_\infty + C\|r-s\|_\infty + |x_0 - y_0|$ can be chosen arbitrarily Theorem 7.4 is proven for all step functions.

Step 2 - Existence and uniqueness on $\overline{\mathrm{BV}}$: We apply Theorem 6.5 to show existence and uniqueness of a solution. In order to do so, we need to ensure that (6.5) holds. Indeed for $r, s \in \mathcal{R}$ we have

$$d_H\left(Z(r), Z(s)\right) \leq M_N C \|r-s\|_{\mathcal{R}} . \tag{7.24}$$

One easy way to see this, is using the result from the first step: Let $r, s \in \mathcal{R}$ and $x_0 \in Z(r)$. Define on the interval $[0,1]$ the functions $u \equiv 0$, $r_1 \equiv r$ and $r_2(t) = r\chi_{[0,1/2]}(t) + s\chi_{(1/2,1]}(t)$. Let ξ_1 and ξ_2 be the solutions associated to u, r_1, x_0 and u, r_2, x_0 respectively. From the first step we deduce that $\|\xi_1 - \xi_2\| \leq M_N C \|r-s\|_{\mathcal{R}}$. Remembering (7.10) this inequality and $u \equiv 0$ lead to $\operatorname{dist}(x_0, Z(s)) = \|Q_{Z(s)}(x_0) - x_0\| \leq M_N C \|r-s\|_{\mathcal{R}}$. The same argument can also be applied to $y_0 \in Z(s)$. Taking the supremum over all $x_0 \in Z(r)$ and $y_0 \in Z(s)$ implies (7.24) .

Step 3 - Lipschitz estimate on $\overline{\mathrm{BV}}$: Now let $u, v \in \overline{\mathrm{BV}}(0,T;X)$, $r, s \in \overline{\mathrm{BV}}(0,T;\mathcal{R})$, $x_0 \in Z(r(0))$ and $y_0 \in Z(s(0))$. Then there exist sequences of step functions $(u^n)_{n\in\mathbb{N}}, (v^n)_{n\in\mathbb{N}} \subset S(0,T;X)$ and $(r^n)_{n\in\mathbb{N}}, (s^n)_{n\in\mathbb{N}} \subset S(0,T;\mathcal{R})$ such that

$$[\forall n \in \mathbb{N} : f^n(0) = f(0) \ \wedge \ \mathrm{Var}(f^n) \leq \mathrm{Var}(f)] \ \wedge \ \|f^n - f\|_\infty \to 0 \tag{7.25}$$

where f is subsequently replaced by u, v, r, s. Let ξ^n, η^n and ξ,η be the associated solutions. Due to Theorem 6.5 we furthermore have $\|\xi^n - \xi\|_\infty \to 0$ and $\|\eta^n - \eta\|_\infty \to 0$. Due to our above argument we know that for every $n \in \mathbb{N}$

$$\|\xi^n - \eta^n\|_\infty \le M_N \left(\|u^n - v^n\|_\infty + C\|r^n - s^n\|_\infty + |x_0 - y_0|\right)$$

holds. Passing to the limit $n \to \infty$ completes the proof. □

7.3. Uniqueness and non-uniqueness of the quasivariational sweeping process

Due to our preparations in the previous section proving Theorem 7.3 is only little work.

Proof of Theorem 7.3. Let $u \in \overline{\mathrm{BV}}(0,T;X)$, $g \in \overline{\mathrm{BV}}\left(0,T;C^{0,1}_{\mu,\lambda}(X \times X;\mathcal{R})\right)$ and $x_0 \in Z(g(0,u(0),x_0 - u(0)))$. Define the set $\Omega \subset \overline{\mathrm{BV}}(0,T;X)$ by

$$\Omega := \left\{ \eta \in \overline{\mathrm{BV}}(0,T;X) : \begin{array}{l} \eta(0) = x_0 - u(0) \wedge \\ \overline{\mathrm{Var}}(\eta) \le \frac{M_N C}{1-\delta}\left(\overline{\mathrm{Var}}(g) + \left(\frac{1}{M_N C} + \mu\right)\overline{\mathrm{Var}}(u)\right) \end{array} \right\}.$$

Notice that due to Ω is closed with respect to the topology induced by the $\|\cdot\|_\infty$-norm. Define $\mathcal{S} : \Omega \to \overline{\mathrm{BV}}(0,T;X)$ as the map $\mathcal{S} : \eta \mapsto \xi$ where ξ is the solution to the sweeping process with inputs u, x_0 and $r(t) = g(t,u(t),\eta(t))$. We claim that $\mathcal{S}$ is a contraction on Ω. We first show that $\mathcal{S}(\Omega) \subset \Omega$. To this end let r be defined as above. First notice that since ξ is a solution we have $\xi(0) = u(0) - x_0$. Furthermore it is easy to see that

$$\overline{\mathrm{Var}}(r) \le \overline{\mathrm{Var}}(g) + \mu\overline{\mathrm{Var}}(u) + \lambda\overline{\mathrm{Var}}(\eta)\,. \tag{7.26}$$

Applying this to (6.8) with $L = M_N C$ (due to step 2 of the above proof) we obtain that

$$\begin{aligned} \overline{\mathrm{Var}}(\xi) &\le \overline{\mathrm{Var}}(u) + M_N C\left(\overline{\mathrm{Var}}(g) + \mu\overline{\mathrm{Var}}(u) + \lambda\overline{\mathrm{Var}}(\eta)\right) \\ &\le M_N C\left(\frac{\delta}{1-\delta} + 1\right)\overline{\mathrm{Var}}(g) \\ &\quad + (1 + M_N C\mu)\,\overline{\mathrm{Var}}(u) + M_N C\frac{\delta}{1-\delta}\left(\frac{1}{M_N C} + \mu\right)\overline{\mathrm{Var}}(u) \\ &= \frac{M_N C}{1-\delta}\left(\overline{\mathrm{Var}}(g) + (1+\mu)\overline{\mathrm{Var}}(u)\right)\,. \end{aligned}$$

Indeed if ξ is a solution to the implicit sweeping process then we can simply calculate

$$\overline{\mathrm{Var}}(\xi) \le \overline{\mathrm{Var}}(u) + M_N C\left(\overline{\mathrm{Var}}(g) + \mu\overline{\mathrm{Var}}(u) + \lambda\overline{\mathrm{Var}}(\xi)\right)$$

and obtain that $\xi \in \Omega$. For $\eta^1, \eta^2 \in \Omega$ we define $r^i(t) = g(t,u(t),\eta^i(t))$ for $i \in [2]$. Then $\|r^1 - r^1\|_\infty \le \lambda\|\eta^1 - \eta^2\|_\infty$ and by applying Theorem 7.4 we get

$$\left\|\mathcal{S}\left(\eta^1\right) - \mathcal{S}\left(\eta^2\right)\right\|_\infty \le \delta\left\|\eta^1 - \eta^2\right\|_\infty\,.$$

This is the contraction property and Banach's fixed point theorem yields the existence of a unique $\xi \in \Omega$ such that $\mathcal{S}(\xi) = \xi$. Thereby ξ is a solution to Problem 6.8. Furthermore, since any solution needs to be in Ω, it is also unique. □

We will now show that condition (7.5) is sharp in dimension less or equal two. This implies that the Lipschitz constant in Theorem 7.4 is sharp with respect to $\|r-s\|_\infty$ in these dimensions. We thereby complement results for the play operator (see [55]), which imply that the Lipschitz constant is also sharp with respect to $\|u-v\|_\infty$ in dimension less or equal to two. There are no results for dimensions greater than two. Indeed even for the play operator it is not known whether the Lipschitz bound is sharp for dimension three or higher.

We treat each dimension in a separate example. As can be expected the case $\dim(X) = 1$ is the simplest one and therefore we shall start with it.

Example 7.8. For $N = \dim(X) = 1$ we can without loss of generality assume that $X = \mathbb{R}$. For simplicity also assume that $\mathcal{R} = \mathbb{R}$ and define $Z(r)$ by

$$Z(r) = \{x \in \mathbb{R} : x \leq r\} \ .$$

This indeed is a polyhedron and we can denote it in terms of (7.1) by writing $x = x \cdot 1$ and $\beta(r) = r$. We hence obtain $C = 1$ and by definition $M_N = M_1 = 1$. We choose

$$g_\lambda(t, u, \xi) = \lambda\,(u - \xi)\,,\ u(t) = t \text{ and } x_0 = 0\,.$$

For $\lambda < 1$ Theorem 7.3 grants the existence of a unique solution, namely $\xi^\lambda(t) = t$. In contrast for $\lambda = 1$ uniqueness is lost. Indeed both

$$\xi_1(t) = t \text{ and } \xi_2 \equiv 0$$

are solutions to Problem (6.8). Since both functions are differentiable it suffices to show that $u(t) - \xi_i(t) \in Z(g_1(t, u, \xi_i(t))$ and $\dot{\xi}_i(t) \in \partial I_{Z(g_1(t,u(t),\xi_i(t)))}(u(t) - \xi_i(t))$ for both $i \in [2]$. The first condition is simple; it is readily granted by the definition of $Z(r)$ and g_1. For ξ_2 the second condition is trivial as well. For any closed convex set K and any $x \in K$, it holds $0 \in \partial I_K(x)$. For ξ_1 notice that $\dot{\xi}_1 \equiv 1$, $\partial I_{Z(r)}(x) = 0$ for $x \neq r$ and $\partial I_{Z(r)}(x) = (-\infty, 0]$ for $x = r$.

Remark 7.9. In the above example one can even show that for $\lambda = 1$ any non decreasing function ξ with $\xi(t) \leq t$ is a solution to the implicit sweeping process.

For $\dim(X) = 2$ the example is slightly more elaborate. The bigger Lipschitz constant is owed to the possibility of the interplay of two faces, which we will make use of in our example. Let us choose $X = \mathbb{R}^2 = \mathcal{R}$. Note that we have to choose a norm for $\mathcal{R}$, which can be more or less arbitrary, as long at it is a norm. X, which has to be a Hilbert space, has to be equipped with an Euclidean norm. We decide in both cases for the standard 2-norm. Let $a_1, a_2 \in \mathbb{R}^2$ such that $|\langle a_1, a_2\rangle| \neq 1$, or in other words choosing a_1, a_2 linearly independent, we define the set $Z(r)$ by

$$Z(r) := \left\{x \in \mathbb{R}^2 : \langle a_i, x\rangle \leq r_i\right\} \ . \tag{7.27}$$

In other words we once again we set $\beta : r \mapsto r$. We can explicitly calculate the value of M_2 and obtain

$$M_2 = \sqrt{\frac{2}{1 - |\langle a_1, a_2\rangle|}}\,. \tag{7.28}$$

Notice that the above term is finite by assumption. The way we need to exploit the interplay of the two faces depends on the sign of $\langle a_1, a_2\rangle$. Roughly speaking, if it is positive then the faces should move in the same direction, if it is negative they shall move in opposite direction. We shall treat the cases in two separate examples.

Example 7.10. We set $\mathcal{R} = X = \mathbb{R}^2$. We equip $\mathcal{R}$ with the $\|\cdot\|_\infty$ norm whereas we equip X as an Hilbert space with the usual inner product. Due to our choice of norm we have $C = 1$. Assume that $Z(r)$ is of form (7.27) and $\langle a_1, a_2\rangle \leq 0$. Define the vector

$$v := \frac{a_1 + a_2}{|a_1 + a_2|} .$$

For the initial value we choose $x_0 = 0 \in \mathbb{R}^2$. We choose the function g once again independent of t as

$$g(u,\xi) := \frac{1}{M_2} \langle u - \xi, v\rangle \begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

Therefore to make notation simpler we drop the dependence of g on time as it obviously does not depend on time. It is now a straightforward computation to see that

$$|g(u,\xi) - g(w,\eta)|_\infty \leq \frac{1}{M_2} \left(|u - w| - |\xi - \eta|\right) .$$

Thus we are once again in the critical case $\delta = 1$. Finally we choose the function u to be

$$u(t) = t \cdot v .$$

It is easy to see that $\xi_1 \equiv u$ is a solution of the quasivariational sweeping process. It is clear that $\xi_1 - u \equiv 0$ is an element of $Z(g(u(t), \xi_1(t)))$. Also $\dot{\xi}_1(t) \equiv v \in \partial I_{Z(g(u(t),\xi_1(t)))}(0)$ due to Lemma 2.7. On the other hand also $\xi_2 \equiv 0$ is a solution to the problem. Indeed on the one hand we have that

$$g(u(t), \xi_2(t)) = \frac{t}{M_2\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = t\sqrt{\frac{1 + \langle a_1, a_2\rangle}{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and on the other hand we can calculate

$$\langle a_1, u_2(t) - \xi_2(t)\rangle = t \left\langle a_1, \frac{a_1 + a_2}{|a_1 + a_2|} \right\rangle = t\sqrt{\frac{1 + \langle a_1, a_2\rangle}{2}} = \langle a_2, u_2(t) - \xi_2(t)\rangle .$$

Therefore $u(t) - \xi_2(t) \in Z\left(g\left(u(t), \xi_2(t)\right)\right)$. Since $\dot{\xi}_2(t) \equiv 0$ the differential inclusion is also satisfied.

In the case $\langle a_1, a_2\rangle \geq 0$ set

$$v := \frac{a_1 - a_2}{|a_1 - a_2|} \tag{7.29}$$

and define g as above. Define $u(t) = t \cdot a_1$ and choose $x_0 = 0 \in \mathbb{R}^2$. Then on the one hand $\xi_1 \equiv u$ is a solution, since both $\xi_1(t) - u(t) = 0 \in Z\left(g(u(t), \xi_1(t))\right)$ and also $a_1 \in \partial I_{Z(g(u(t),\xi_1(t)))}(0)$. Moreover

$$\xi_2(t) := t \cdot a_2 \tag{7.30}$$

is a solution. It is an easy computation to check that

$$u(t) - \xi_2(t) = t(a_1 - a_2) \in Z(g(t \cdot a_1, t \cdot a_2)) . \tag{7.31}$$

Indeed it turns out that $\Gamma_{Z(g(t\cdot a_1, t\cdot a_2))}\big(t\,(a_1 - a_2)\big) = \{1, 2\}$ and hence also

$$\dot{\xi}_2(t) \in \partial I_{Z(g(u(t),\xi_2(t)))}(u(t) - \xi_2(t)) .$$

8. Quasivariational sweeping processes with smooth characteristics

The aim of this chapter is to prove existence and uniqueness of a quasivariational sweeping process for smooth characteristics on (left continuous) functions of bounded variations. This generalizes a previous result due to M. Brokate, P. Krejčí and H. Schnabel [18] who studied this problem for absolutely continuous functions. In the previous chapter the conditions which were needed to establish the existence of a unique solution were the same both for functions of bounded variation and for absolutely continuous functions. Here however things are different. We cannot straightforwardly transfer the arguments of [18]. It turns out that we pay for allowing BV functions by the need to employ additional structural assumptions. This is due to the fact that at every point the solution jumps it has to satisfy a static quasivariational inequality, that is a variational problem of the form

$$\text{find } \xi \text{ such that } \langle u-\xi, \xi-y\rangle \geq 0 \ \ \forall y \in Z(\xi)\,.$$

We will show that the assumption of [18] do not suffice to guarantee the existence of an unique solution for this problem. Therefore throughout the main part of our exposition we will limit the size of the jumps. At the end we however will give examples for stricter conditions on the convex sets that in exchange allow for arbitrary large jumps.

The idea of our proof here is the following: The BV functions are decomposed into intervals in which only very small jumps occur and a finite number of larger jumps. We solve each part by a contraction argument. For the first part a method similar to [18] is used to show the existence of a solution. This part is indeed quite technical as we have to use differences instead of derivatives which allows a shorter and more elegant proof. For the larger jumps it suffices to analyze the static quasivariational inequality. Here existence and uniqueness can only be guaranteed if the additional structural condition is satisfied, e.g. when the size of the jumps is not too big. The solution to the static problem gives the starting point for the evolution problem on the new time interval. Since we are in BV, only a finite number of restarts have to be considered. The framework of the Kurzweil solution is very suitable for this procedure as it easily allows to "glue" the solutions of the different parts together.

Let us give a short overview over this chapter: In the following Section 8.1 we are going to accurately introduce our main result. Section 8.2 is devoted to the analysis of the static quasivariational inequality. We would especially like to point out Proposition 8.9, which studies the distance of the projections of a point onto two different convex sets under fairly general assumptions. We believe this result to be new and it might be of interest also outside the context of this study. Afterwards, we are going to establish existence and uniqueness of a solution under the assumption that the involved functions have only very small jumps. There, we will slightly deviate from our main course and prove local Lipschitz continuity of the sweeping process on BV. This might be of interest in applications. The proof of the main theorem is the content of Section 8.4. In the final Section 8.5 we are first going to show that a condition on the jump size is needed. We will then give examples how this can be overcome by additional structural assumption on the convex sets.

The results we present here have been obtained and published in a joint work with P. Krejčí [62] and the article [90].

8.1. Introduction of the main result

We start by precisely formulating the assumptions on the convex sets imposed throughout this chapter. Throughout this chapter we will write

$$M(r,x) := M_{Z(r)}(x)$$

where $M_Z(x)$ is the Minkowski functional of Z evaluated in point x, see Definition 2.21.

Hypothesis 8.1. *There exists $C > 0$ such that $0 \in Z(r) \subset B_C(0)$ for all $r \in \mathcal{R}$. Furthermore the partial Fréchet derivatives $\partial_r M(r,x) \in \mathcal{R}^*$ and $\partial_x M(r,x) \in X$ exist for every $r \in \mathcal{R}$ and every $x \in X \setminus \{0\}$. We denote $B(r,x) = \frac{1}{2}M^2(r,x)$. The maps*

$$\begin{aligned} J(r,x) &= \partial_x B(r,x) = M(r,x)\partial_x M(r,x) : X \times \mathcal{R} \to X \,, \\ K(r,x) &= \partial_r B(r,x) = M(r,x)\partial_r M(r,x) : X \times \mathcal{R} \to \mathcal{R}^* \end{aligned}$$

allow continuous extensions to $x = 0$. Furthermore, there exist constants K_0, C_J, C_K such that for all $x, y \in B_C(0)$, $r, s \in \mathcal{R}$ it holds

$$\begin{aligned} \|K(r,x)\|_{\mathcal{R}^*} &\leq K_0 \,, \\ |J(r,x) - J(s,y)| &\leq C_J \left(|x-y| + \|r-s\|_{\mathcal{R}}\right) \,, \\ \|K(r,x) - K(s,y)\|_{\mathcal{R}^*} &\leq C_K \left(|x-y| + \|r-s\|_{\mathcal{R}}\right) \,. \end{aligned}$$

In [18] it was additionally assumed that all sets $Z(r)$ contain a ball centered at 0 with some radius $c > 0$. However this is already a consequence of the other assumptions:

Lemma 8.2. *If Hypothesis 8.1 holds, then $C_J C^2 > 1$. Furthermore for $c := C_J^{-1/2}$ we have $B_c(0) \in Z(r)$ for all $r \in \mathcal{R}$.*

Proof. For all $x \in X$ and $r \in \mathcal{R}$ we have $M^2(r,x) = \langle J(r,x), x\rangle$ and $J(r,0) = 0$. Hypothesis 8.1 then yields $M^2(r,x) \leq C_J|x|^2$. Using the implications $|x| > C \Rightarrow M(r,x) > 1$ and $|x| < C_J^{-1/2} \Rightarrow M(r,x) < 1$, we obtain the assertion. □

The assumptions we impose might be hard to understand in a geometrical context at first glance. What we essentially ask for is that at every point $x \in \partial Z(r)$ there exists a unique outer normal vector $n(r,x)$. Furthermore we want this vector to depend globally Lipschitz continuous on both x and r. We shall present a prototypical convex set that satisfies these assumptions below in Example 8.5.

We denote by $C^1_{\omega,\gamma}(X \times X; \mathcal{R})$ the space of functions $f \in C^1(X \times X; \mathcal{R})$, $(u,\xi) \mapsto f(u,\xi)$ such that $\|\partial_u f\|_\infty \leq \omega$ and $\|\partial_\xi f\|_\infty \leq \gamma$. The main result of this chapter reads:

Theorem 8.3 (Existence and uniqueness)**.** *Let $u \in BV_L^{c_u}(0,T;X)$, $g \in BV_L^{c_g}(0,T;C^1_{\omega,\gamma}(X \times X;\mathcal{R}))$ and $x_0 \in Z(g(0,u(0),u(0)-x_0))$. Assume that Hypothesis 8.1,*

$$\delta := CK_0\gamma \;<\; 1 \qquad \textit{and} \tag{8.1}$$

$$CK_0c_g + (1+CK_0\omega)c_u \;<\; \frac{(1-\delta)^2}{C_JC(1+\delta)} \tag{8.2}$$

hold. Then there exists a unique solution to Problem 6.8.

Remember that BV_L^c is the set of left-continuous functions of bounded variation whose jumps are smaller than c. For a more thorough discussion of these functions we refer to Definition 3.15 and the analysis thereafter.

Remark 8.4. When allowing for absolutely continuous functions only, existence and uniqueness has been shown in [18] if (8.1) holds. Our result covers this case: Absolutely continuous functions have no jumps, i.e. discontinuities, and hence (8.2) is always satisfied. It turns out that (8.2) is owed to the vectorial nature of our problem. For $X = \mathbb{R}$ it is superfluous, as we shall see in Section 8.5.

To verify that a given problem does indeed satisfy the above assumptions amounts in general to tedious calculations. It has been done for the Gurson model with absolutely continuous functions in [98]. That model, introduced in [40], describes the nucleation of voids in elastic and ideally plastic materials. Here the convex sets are ellipsoids in the space of stress tensors, that is symmetric tensors on $\mathbb{R}^3$, namely

$$Z(r) = \left\{\sigma \in \mathbb{R}^{3\times3}_{\text{symm}} : a\|d(\sigma)\|^2 + r\left(2\cosh\big(b\cdot \text{tr}(\sigma)\big) - 1\right) \le (1-r)^2\right\}\,,$$

where $\text{tr}(\sigma)$ denotes the trace of σ and $d(\sigma)$ the trace free deviator $d(\sigma) = \sigma - \frac{1}{3}\text{tr}(\sigma)\text{Id}$.

In order to develop a feeling for what this theorem is able to do and to understand its limitations we shall give an example of a problem we are able to handle with this result. The probably most simple examples which do not reduce to a problem on the real line appear to be ellipsoids in $\mathbb{R}^2$. This will also be the form of the convex set here.

Example 8.5. Choose $X = \mathbb{R}^2$ and $\mathcal{R} = \mathbb{R}$. Furthermore let the set $Z(r)$ be of the form

$$Z(r) = \left\{x \in \mathbb{R}^2 : x_1^2 + 2x_2^2 \le f(r)^2\right\}$$

with $f : \mathbb{R} \to (0,\infty)$. First we need to ensure that the conditions of Hypothesis 8.1 are satisfied. First we compute $M(x,r) = \left(x_1^2 + 2x_2^2\right)^{+1/2}/f(r)$. To give a concrete example let us assume that

$$f(r) = \frac{3}{4} + \frac{1}{4}\sin(r)\,.$$

Then it is easy to see that $Z(r) \subset B_1(0)$. Furthermore

$$J(x,r) \;=\; \left(\frac{3}{4} + \frac{1}{4}\sin(r)\right)^{-1}\begin{pmatrix} x_1 \\ 2x_2 \end{pmatrix} \quad \text{and}$$

$$K(x,r) \;=\; \frac{x_1^2 + 2x_2^2}{4\left(\frac{3}{4} + \frac{1}{4}\sin(r)\right)^3}\cos(r)\,.$$

It is possible to choose $K_0 = 4$, $C_J = 4$ and $C_K = 28$. Here the choice of C_K is rough, however as it does not appear in (8.1) and (8.2), its mere existence suffices. Due to the above estimates we need $\gamma < \frac{1}{4}$ in order to be able to apply our theorem. Assuming that g does only depend on ξ, e.g.

$$g(t, u, \xi) = \gamma \xi_1(t)\,,$$

the jump size of u needs to be bounded by $\frac{(1-4\gamma)^2}{4(1+\gamma)}$.

8.2. The static quasivariational inequality

When trying to analyze a time-dependent problem, it is natural to consider the underlying static problem first - in our case the quasivariational inequality. Apart from that being folklore wisdom we have even more reason to do so: Let ξ be a solution of Problem 6.8 that jumps at some time $\tau \in [0, T)$. Choosing the testfunction z in (6.15) by $z(t) = u(t+) - \xi(t+)$ for $t \neq \tau$ and $z(\tau) = y$ for some $y \in Z(g(\tau+, u(\tau+), \xi(\tau+)))$ leads to

$$\langle u(\tau+) - \xi(\tau+) - y, \xi(\tau+) - \xi(\tau)\rangle \geq 0\,. \tag{8.3}$$

Therefore at any time τ where ξ jumps it has to satisfy a quasivariational inequality. A short study of it will be the subject of this section. Indeed this analysis is crucial in understanding the jump size condition. Let us start by precisely formulating the problem we want to solve.

Problem 8.6. *Let $g_i \in Lip_{\omega,\gamma}(X \times X; \mathcal{R})$, $u_i \in X$ $i \in \{0, 1\}$ and $\xi_0 \in X$ such that $x_0 = u_0 - \xi_0 \in Z(g_0(x_0, u_0))$. Find ξ_1 such that*

$$\forall y \in Z(g_1(u_1, \xi_1)) : \;\; \langle u_1 - \xi_1 - y, \xi_1 - \xi_0\rangle \geq 0\,. \tag{8.4}$$

The above problem is called a quasivariational inequality. The set of admissible vectors to test against depends on the state that is tested. These problems have been intensively studied by C. Baiocchi and A. Capelo in their seminal monograph [7]. They established existence and uniqueness of solutions by employing order methods. However these results are not applicable to our case. We proceed in a different way and try to establish existence and uniqueness of a solution to Problem 8.6 by applying Banach's contraction principle. To this end we rewrite the (8.4) as find $x_1 \in X$ such that

$$u_1 - \xi_1 = Q_{Z(g(u_1,\xi_1))}(x_0 + \Delta u)\,, \tag{8.5}$$

where $\Delta f = f_1 - f_0$ for any $f \in \{u, g, \xi, x\}$. Here and for the remainder of this chapter we denote by $x_{(i)}$ the term $\xi_{(i)} - u_{(i)}$. We point out that Q_Z is the projection onto Z as defined in Section 2.2. Our aim is to show that the map

$$\begin{cases} X & \to & X \\ \eta & \mapsto & u_1 - Q_{Z(g_1(u_1,\eta))}\,(x_0 + \Delta u) \end{cases}$$

is a contraction. Hence our task is clear: Estimate the difference of the two projections

$$Q_{Z(g_1(u_1,\eta_1))}\,(x_0 + \Delta u) - Q_{Z(g_1(u_1,\eta_2))}\,(x_0 + \Delta u)\;.$$

It is perhaps surprising, that the difference does not only depends on the distance of the convex sets and therefore on $\eta_1 - \eta_2$ but also on the length of the projection which means in the end it depends on Δg and Δu.

8.2.1. The difference of projections onto two convex sets

We first establish an estimate on the difference of two projections (Proposition 8.9) under rather general assumptions on the convex sets. We believe this result to be of interest in its own right. It has been originally obtained in joint work with Pavel Krejčí [62]. We will then use the assumptions of Hypothesis 8.1 to derive the result adapted to our purposes in Corollary 8.13.

Hypothesis 8.7. *Assume $Z : \mathcal{R} \to 2^X$ satisfies the following conditions: $Z(r) \subset X$ is nonempty, closed and convex for all $r \in \mathcal{R}$ and there exist functions $j : \mathcal{R} \times \mathcal{R} \to \mathbb{R}$, $\psi : [0, \infty) \to [0, \infty)$ such that*

$$|n_r - n_s| \le j(r, s) + \psi(|x - y|)$$

for all $x \in \partial Z(r)$, $y \in \partial Z(s)$ and $n_r \in \partial I_{Z(r)}(x)$, $n_s \in \partial I_{Z(s)}(y)$, $|n_r| = |n_s| = 1$.

Notice that the outer unit normal does not need to be unique, i.e. $\psi(0) = 0$ is not required. However ψ can be interpreted as a (generalized) modulus of continuity whereas j does not necessarily need to be one.

Remark 8.8. As a shorthand we shall from now on write $d_H(r, s)$ instead of $d_H(Z(r), Z(s))$ and P_r for $P_{Z(r)}$.

Our aim here is to prove the following result.

Proposition 8.9. *Let Hypothesis 8.7 hold and let $u \in X$ be given. Then*

$$|Q_r(u) - Q_s(u)| \le d_H(r, s) + |P_r u|(j(r, s) + \psi(d_H(r, s)))$$

To prepare the proof of this statement we establish some helpful results.

Lemma 8.10. *Let Z be a convex set and Q_Z, P_Z be as in Section 2.2. Let $x, e \in X$, $|e| = 1$ be fixed and $\Lambda \subset \mathbb{R}$ be the set*

$$\Lambda = \{\lambda \in \mathbb{R} : x + \lambda e \notin Z\} .$$

For $\lambda \in \Lambda$ define

$$b_\lambda := P_Z(x + \lambda e), \quad f(\lambda) = \left\langle e, \frac{P_Z(x + \lambda e)}{|P_Z(x + \lambda e)|} \right\rangle .$$

Then for all $\mu, \lambda \in \Lambda$ it holds that

$$0 \;\le\; \frac{|b_\lambda| + |b_\mu|}{2} \left| \frac{b_\lambda}{|b_\lambda|} - \frac{b_\mu}{|b_\mu|} \right|^2 \;\le\; (\lambda - \mu)(f(\lambda) - f(\mu)) . \tag{8.6}$$

Especially, the function f is nondecreasing in Λ.

Proof. By Lemma 2.12, we have for every $\lambda, \mu \in \Lambda$ that

$$\left\langle \frac{P_Z(x + \lambda e)}{|P_Z(x + \lambda e)|}, Q_Z(x + \lambda e) - Q_Z(x + \mu e) \right\rangle \;\ge\; 0 ,$$

$$\left\langle \frac{P_Z(x + \mu e)}{|P_Z(x + \mu e)|}, Q_Z(x + \mu e) - Q_Z(x + \lambda e) \right\rangle \;\ge\; 0 .$$

Summing up the two inequalities we obtain

$$\left\langle \frac{b_\lambda}{|b_\lambda|} - \frac{b_\mu}{|b_\mu|}, b_\lambda - b_\mu \right\rangle \;\le\; (\lambda - \mu) \left\langle e, \frac{b_\lambda}{|b_\lambda|} - \frac{b_\mu}{|b_\mu|} \right\rangle ,$$

which is precisely (8.6). □

Lemma 8.11. *Under the assumptions of Proposition 8.9, we have*

$$\big||P_r(u)| - |P_s(u)|\big| \leq d_H(r,s)\,.$$

Proof. We have

$$|P_r(u)| \leq |u - Q_r(Q_s(u))| \leq |u - Q_s(u)| + |P_r(Q_s(u))| \leq |P_s(u)| + d_H(r,s)$$

Interchanging the roles of r and s, we obtain the assertion. □

Now everything is in place to move to the main proof.

Proof of Proposition 8.9. If $u \in Z(r) \cup Z(s)$ or $Q_r(u) = Q_s(u)$ the assertion is straightforward. Thus set $x_r = Q_r(u)$, $x_s = Q_s(u)$ and assume that

$$x_r \neq x_s\,,\ u \notin Z(r) \cup Z(s)\,,\ |P_r(u)| \leq |P_s(u)|\,. \tag{8.7}$$

Set

$$n_r = \frac{P_r(u)}{|P_r(u)|}, \quad n_s = \frac{P_s(u)}{|P_s(u)|}. \tag{8.8}$$

Then n_r, n_s are unit vectors belonging to $\partial I_{Z(r)}(x_r), \partial I_{Z(s)}(x_s)$, respectively, and we have

$$\begin{aligned} |x_r - x_s| &= |P_r(u) - P_s(u)| = ||P_r(u)|n_r - |P_s(u)|n_s| \\ &\leq ||P_r(u)| - |P_s(u)||\,|n_s| + |P_r(u)|\,(|n_r - n_s|) \\ &\leq d_H(r,s) + |P_r(u)|\,(|n_r - n_s|)\ . \end{aligned} \tag{8.9}$$

In the last step we employed the result of Lemma 8.11. Put $e := n_r$ and

$$x_s(\lambda) := Q_s(x_r + \lambda e)\,,\ x_r(\lambda) = Q_r(x_r + \lambda e)$$

for $\lambda \geq 0$. We have

$$|x_s(0) - x_r(0)| = |Q_s(x_r) - x_r| \leq d_H(r,s)\,.$$

Furthermore for all $\lambda \geq 0$ we have by Lemma 2.12 (iii) that

$$P_r(x_r + \lambda e) = \lambda e\,.$$

Assume that for some $\lambda \in [0, |P_r(u)|]$ we have $x_r + \lambda e \in Z(s)$. Then

$$\langle P_s(u), x_s - x_r - \lambda e\rangle \geq 0\,,$$

hence

$$\langle P_s(u), P_r(u) - P_s(u)\rangle \geq \lambda \left\langle P_s(u), \frac{P_r(u)}{|P_r(u)|}\right\rangle\,,$$

that is,

$$|P_s(u)|^2 \leq \langle P_s(u), P_r(u)\rangle \left(1 - \frac{\lambda}{|P_r(u)|}\right)\,. \tag{8.10}$$

By (8.7), we have $|P_r(u)| \leq |P_s(u)|$. This is compatible with (8.10) only if $\lambda = 0$ and $x_r - x_s = P_s(u) - P_r(u) = 0$, which contradicts (8.7). Hence we have that

$$[0, |P_r(u)|] \subset \Lambda_s = \{\lambda \in \mathbb{R} : x_r + \lambda e \notin Z(s)\}\,.$$

For $\lambda \in [0, |P_r(u)|]$ set

$$n_r(\lambda) = \frac{P_r(x_r + \lambda e)}{|P_r(x_r + \lambda e)|} = e\,, \quad n_s(\lambda) = \frac{P_s(x_r + \lambda e)}{|P_s(x_r + \lambda e)|}\,. \tag{8.11}$$

We have for all $\lambda \in [0, |P_r(u)|]$ that

$$n_r(\lambda) \in \partial I_{Z(r)}(x_r(\lambda))\,, \quad n_s(\lambda) \in \partial I_{Z(s)}(x_s(\lambda)), \tag{8.12}$$

and

$$|n_r(\lambda) - n_s(\lambda)|^2 = 2\left(1 - \left\langle e, \frac{P_s(x_r + \lambda e)}{|P_s(x_r + \lambda e)|}\right\rangle\right).$$

By Lemma 8.10 we have for $\lambda > \mu$ that

$$\begin{aligned} &|n_r(\lambda) - n_s(\lambda)|^2 - |n_r(\mu) - n_s(\mu)|^2 \\ \le\ & -\frac{|P_s(x_r + \mu e)| + |P_s(x_r + \lambda e)|}{\lambda - \mu}\left|\frac{P_s(x_r + \mu e)}{|P_s(x_r + \mu e)|} - \frac{P_s(x_r + \lambda e)}{|P_s(x_r + \lambda e)|}\right|^2. \end{aligned} \tag{8.13}$$

Note that in agreement with (8.8), for $\lambda = |P_r(u)|$ we have $n_s(\lambda) = n_s$, and for $\mu \to 0$ we get $|n_r(\mu) - n_s(\mu)| \to |e - n_s(0)|$. Using $|P_r(u)| \le |P_s(u)|$ we get

$$|n_r - n_s|^2 \le |e - n_s(0)|^2 - |n_s - n_s(0))|^2\,.$$

Especially we have

$$|n_r - n_s| \le |e - n_s(0)| \le j(r,s) + \psi(d_H(r,s))\,,$$

where the latter inequality is due to (8.2.1), (8.12), and Hypothesis 8.7. To complete the proof, it suffices to refer to (8.9). □

In order to apply Proposition 8.9 to our problem we need some results, which have been proven in [18, Section 3].

Proposition 8.12. *Let Hypothesis 8.1 hold, let c be as in Lemma 8.2. Then*

$$\frac{|x|}{C} \le M(r,x) \le \frac{|x|}{c} \quad \forall (r,x) \in \mathcal{R} \times X\,, \tag{8.14}$$

$$\frac{1}{C} \le |J(r,x)| \le \frac{1}{c} \quad \forall r \in \mathcal{R}\ \forall x \in \partial Z(r)\,, \tag{8.15}$$

$$n(r,x) = \frac{J(r,x)}{|J(r,x)|} \quad \forall r \in \mathcal{R}\ \forall x \in \partial Z(r)\,. \tag{8.16}$$

We now can derive the following assertions.

Corollary 8.13. *Let Hypothesis 8.1 hold. Then*

$$d_H(r,s) \le CK_0\|r - s\|_{\mathcal{R}} \quad \forall r,s \in \mathcal{R}\,, \tag{8.17}$$

$$\begin{aligned} &|n(r,x) - n(s,y)| \le C_J C(|x - y| + \|r - s\|_{\mathcal{R}}) \\ &\qquad \forall r,s \in \mathcal{R}\ \forall x \in \partial Z(r)\ \forall y \in \partial Z(s)\,, \end{aligned} \tag{8.18}$$

$$\begin{aligned} &|Q_r(u) - Q_s(u)| \le (CK_0 + C_J C(1 + CK_0)|P_r(u)|)\,\|r - s\|_{\mathcal{R}} \\ &\qquad \forall r,s \in \mathcal{R}\ \forall u \in X\,. \end{aligned} \tag{8.19}$$

Proof. Let $x \in Z(r)$ be arbitrary and assume that $x \notin Z(s)$. Then $M(s,x) \geq 1$, $M(r,x) \leq 1$, and

$$\begin{aligned} \operatorname{dist}(x, Z(s)) &\leq \left| x - \frac{x}{M(s,x)} \right| = \frac{|x|}{M(s,x)(1+M(s,x))}(M^2(s,x)-1) \\ &\leq C\left(\frac{1}{2}M^2(s,x) - \frac{1}{2}M^2(r,x)\right) \leq CK_0\|r-s\|_{\mathcal{R}} . \end{aligned}$$

Interchanging the roles of r and s, we obtain (8.17). Using Proposition 8.12, it is easy to check that (8.18) holds. Indeed, for all $r, s \in \mathcal{R}$, $x \in \partial Z(r)$, and $y \in \partial Z(s)$, we have

$$|n(r,x) - n(s,y)| \leq \frac{|J(r,x) - J(s,y)|}{\sqrt{|J(r,x)|\,|J(s,y)|}} \leq C|J(r,x) - J(s,y)| ,$$

and the assertion follows from Hypothesis 8.1. Hypothesis 8.7 is thus fulfilled with $j(r,s) = C_J C\|r-s\|_{\mathcal{R}}$, $\psi(p) = C_J C p$, and (8.19) follows from Proposition 8.9. □

8.2.2. Existence and uniqueness for the static quasivariational inequality

Let us start with a simple proposition.

Proposition 8.14. *Assume $x_0 \in Z(s)$ and $\xi_1 - u_1 = Q_{Z(r)}(x_0 + \Delta u)$, then*

$$|\Delta\xi| \leq |\Delta u| + d_H(r,s) .$$

Proof. First note that $\Delta\xi = x_0 + \Delta u - x_1$. We therefore estimate

$$|\Delta\xi| \leq |P_{Z(r)}(x_0 + \Delta u)| \leq |P_{Z(r)}(x_0 + \Delta u) - P_{Z(r)}(x_0)| + |P_{Z(r)}(x_0)| .$$

Due to Proposition 2.12 (ii) the first term is less than $|\Delta u|$ and the second is by definition less than $d_H(r,s)$. □

Using the above calculation we can estimate the 'jump size' $|\Delta\xi|$ of the solution.

Corollary 8.15. *If Hypothesis 8.1 holds and ξ_1 is a solution to Problem 8.6, then*

$$|\Delta\xi| \leq |\Delta u| + CK_0(|\Delta g| + \gamma|\Delta\xi| + \omega|\Delta u|) \tag{8.20}$$

Proof. Using 8.14 with $s = g_0(u_0, \xi_0)$ and $r = g_1(u_1, \xi_1)$ leads to

$$|\Delta\xi| \leq |\Delta u| + CK_0\|r-s\|_{\mathcal{R}} .$$

It remains to estimate $\|r-s\|_{\mathcal{R}}$, which is done in the usual way

$$\|g_1(u_1,\xi_1) - g_0(u_0,\xi_0)\|_{\mathcal{R}} \leq \|g_1(u_1,\xi_1) - g_1(u_0,\xi_0)\|_{\mathcal{R}} + \|g_1(u_0,\xi_0) - g_0(u_0,\xi_0)\|_{\mathcal{R}}$$

and the proof is complete. □

As in [18] we assume that $CK_0\gamma =: \delta < 1$. No matter how small the size of the jumps is this condition is indeed necessary to obtain uniqueness as the following example demonstrates.

Example 8.16. Let $X = \mathcal{R} = \mathbb{R}$ and $0 < c < 1 < \infty$. Choose

$$Z(r) = [-1, \max\{\min\{1, r\}, c\}]$$

and $g_i(u, \xi) = u - \xi$ for $i \in [2]$. Then $C = K_0 = \gamma = 1$. For $x_0 = \xi_0 = c$ and $u_0 = 1$. Choose $u_1 = \varepsilon \geq 0$. Then any $\xi_1 \in [\max\{0, u_1 - 1\}, \min\{\varepsilon, u_1 - c\}]$ is a solution to the quasivariational inequality.

Proposition 8.17. *Let Hypothesis 8.1 and $CK_0\gamma =: \delta < 1$ hold. If ξ_1 is a solution to Problem 8.6, then*

$$|\Delta\xi| \leq \frac{1}{1-\delta}(CK_0|\Delta g| + (1 + CK_0\omega)|\Delta u|) =: S(|\Delta g|, |\Delta u|)\,. \tag{8.21}$$

Furthermore

$$\forall \eta \in \overline{B}_{S(|\Delta g|,|\Delta u|)}(x_0) : \; Q_{Z(g(\eta,u_1))}(x_0 + \Delta u) \in \overline{B}_{S(|\Delta g|,|\Delta u|)}(x_0)\,.$$

Obtaining the assertion is just a simple calculation. Finally we are now able to prove the following

Proposition 8.18. *Assume Hypothesis 8.1 holds and $CK_0\gamma =: \delta < 1$. Furthermore let $|\Delta g|$ and $|\Delta u|$ be chosen small enough such that*

$$CK_0|\Delta g| + (1 + CK_0\omega)|\Delta u| \leq \frac{(1-\delta)^2}{C_J C(1+\delta)}\,. \tag{8.22}$$

Then there exists a unique solution ξ_1 to Problem 8.6 with $x_1 = u_1 - \xi_1 \in \overline{B}_{S(|\Delta g|,|\Delta u|)}$.

Proof. By Corollary 8.13 we have for all $\eta, \xi \in \overline{B}_{S(|\Delta g|,|\Delta u|)}(x_0)$ that

$$\begin{aligned}
&|Q_{Z(g(\xi,u_1))}(x_0 + \Delta u) - Q_{Z(g(\eta,u_1))}(x_0 + \Delta u)| \\
&\quad \leq \;\; (CK_0\gamma + S(|\Delta g|, |\Delta u|)C_J C(1 + CK_0\gamma))\,|\xi - \eta| \\
&\quad \leq \;\; (\delta + S(|\Delta g|, |\Delta u|)C_J C(1 + \delta))\,|\xi - \eta|\,.
\end{aligned}$$

Then (8.22) implies that

$$\delta + S(|\Delta g|, |\Delta u|)C_J C(1 + \delta) =: \delta' < 1$$

and we can apply the Banach's contraction principle. □

Note that the condition imposed in Proposition 8.18 correspond to those imposed in Theorem 8.3. If they hold, then we denote the solution operator of the quasivariational inequality by

$$\xi_1 =: \mathcal{E}(\xi_0, u_0, g_0, u_1, g_1)\,. \tag{8.23}$$

8.3. Existence and uniqueness of a solution for very small jumps

In this section we establish the existence and uniqueness of a solution to Problem 6.8 if the size of the jumps is very small. The method of the proof is similar to the method used in [18]. However we cannot use weak derivatives, but approximate the BV functions by step functions. This turns out to contain some quite technical and lengthy computations. The main result we want to prove here reads as follows.

Theorem 8.19 (Existence and uniqueness for very small jumps)**.** *Assume that Hypothesis 8.1 holds. Furthermore let*

$$\delta := CK_0\gamma < 1 \tag{8.24}$$

and $u \in BV_L^{c_u}(0,T;X)$, $g \in BV_L^{c_g}(0,T;C^1_{\omega,\gamma}(X\times X;\mathcal{R}))$ *and* $x_0 \in Z(g(0,u(0),x_0-u(0)))$. *Then there exists some* $\nu > 0$ *such that there exits a unique solution to Problem 6.8 if* $c_g, c_u \le \nu$.

We start by studying the sweeping process on step functions. What we essentially do is trying to transfer the results of [18, Section 5] to our case by replacing the derivatives by differences. This is however not always possible and additional terms occur which mainly involve jump sizes. These are the main reason, why we can only handle very small jumps with this approach. On the way to the main prove we make a small deviation and prove that the sweeping process is locally Lipschitz continuous with respect to the BV-norm (Section 8.3.2). A reader not interested in this result and its proof may skip the Section. However we recommend it as it is quite short and instructive for the main proof which we finish in Section 8.3.3.

8.3.1. One step estimates

Assume that u, v, r, s are step functions on the same division $(t_k)_{k=0}^n \in \mathcal{P}$. More specifically, assume that elements $u_k, v_k \in X$, $r_k, s_k \in \mathcal{R}$ for $k = 0, \dots, n$ are given, and that

$$f(t) = f_0\chi_{\{0\}}(t) + \sum_{k=1}^{n} f_k\chi_{(t_{k-1},t_k]}(t)\,, \tag{8.25}$$

where f stands for u, v, r, s. The solutions ξ, x corresponding to u, r and initial condition x_0, and η, y corresponding to v, s and initial condition y_0, are given by formula (8.25) as well, with f replaced successively by ξ, η, x, y, where $\xi_0 = u_0 - x_0$, $\eta_0 = v_0 - y_0$, and

$$x_k = Q_{r_k}(x_{k-1} + \Delta_k u)\,, \quad \xi_k = \xi_{k-1} + P_{r_k}(u_k - \xi_{k-1})\,, \tag{8.26}$$

and similarly

$$y_k = Q_{s_k}(y_{k-1} + \Delta_k v)\,, \quad \eta_k = \eta_{k-1} + P_{s_k}(v_k - \eta_{k-1})\,, \tag{8.27}$$

for $k \in [n]$, where we denote $\Delta_k f := f_k - f_{k-1}$ for all $k \in [n]$.

Lemma 8.20. *For* $k \in [n]$ *set*

$$A_k[r,u] = \langle J(r_k, x_k), \Delta_k\xi\rangle\ .$$

Then $A_k[r,u] \ge 0$ *for all* $k \in [n]$. *If moreover* $\Delta_k\xi \ne 0$, *then*

$$\Delta_k\xi = \frac{A_k[r,u]}{|J(r_k,x_k)|^2} J(r_k,x_k)\,. \tag{8.28}$$

Proof. If $\Delta_k\xi = 0$, the inequality is trivially fulfilled. Let $\Delta_k\xi \neq 0$. It follows from (8.8), (8.16), and (8.26) that

$$\Delta_k\xi = n(r_k, x_k)|\Delta_k\xi| = \frac{|\Delta_k\xi|}{|J(r_k, x_k)|} J(r_k, x_k)\,,$$

hence

$$A_k[r, u] = |\Delta_k\xi|\,|J(r_k, x_k)| \geq 0\,,$$

and formula (8.28) holds. □

We now define $B(r_k, x_k) := \frac{1}{2}M(r_k, x_k)^2$, and similarly $B(s_k, y_k)$ for s, y.

Lemma 8.21. *Let Hypothesis 8.1 hold. Then for every $k \in [n]$ we have*

$$|\Delta_k(\xi - \eta)| \leq C|\Delta_k\xi||J(r_k, x_k) - J(s_k, y_k)| + C|A_k[r, u] - A_k[s, v]|\,, \tag{8.29}$$

and

$$\begin{aligned} &\Delta_k|B(r, x) - B(s, y)| + |A_k[r, u] - A_k[s, v]| \\ \leq\ & (C_J\,|\Delta_k u| + C_K\,\|\Delta_k r\|_{\mathcal{R}} + 2C_J\,|\Delta_k\xi|) \max_{i\in\{k,k-1\}} \|r_i - s_i\|_{\mathcal{R}} \\ & + (C_K\,\|\Delta_k r\|_{\mathcal{R}} + 2C_J\,(|\Delta_k u| + \|\Delta_k s\|_{\mathcal{R}} + |\Delta_k y| + |\Delta_k\xi|)) \max_{i\in\{k,k-1\}} |x_i - y_i| \\ & +2C_J\,(\|\Delta_k s\|_{\mathcal{R}} + |\Delta_k y|) \max_{i\in\{k,k-1\}} |u_i - v_i| \\ & +K_0\,\|\Delta_k(r - s)\|_{\mathcal{R}} + \frac{1}{c}\,|\Delta_k(u - v)|\,. \end{aligned} \tag{8.30}$$

Proof. At the heart of this proof lies a careful case analysis.
Case 1 - $\Delta_k\xi = 0 = \Delta_k\eta$: The first inequality is trivial. Furthermore we have $A_k[r, u] = 0 = A_k[s, v]$. By the mean value theorem, there exists some $\lambda \in [0, 1]$ such that for

$$\begin{pmatrix} x_\lambda \\ r_\lambda \\ y_\lambda \\ s_\lambda \end{pmatrix} = (1 - \lambda) \begin{pmatrix} x_k \\ r_k \\ y_k \\ s_k \end{pmatrix} + \lambda \begin{pmatrix} x_{k-1} \\ r_{k-1} \\ y_{k-1} \\ s_{k-1} \end{pmatrix}$$

it holds (note that $\Delta_k x = \Delta_k u$, $\Delta_k y = \Delta_k v$)

$$\begin{aligned} &\big(B(r_k, x_k) - B(s_k, y_k)\big) - \big(B(r_{k-1}, x_{k-1}) - B(s_{k-1}, y_{k-1})\big) \\ =\ & \langle J\left(r_\lambda, x_\lambda\right), \Delta_k u\rangle + \langle K\left(r_\lambda, x_\lambda\right), \Delta_k r\rangle_{\mathcal{R}',\mathcal{R}} - \big(\langle J\left(s_\lambda, y_\lambda\right), \Delta_k v\rangle \\ & + \langle K\left(s_\lambda, y_\lambda\right), \Delta_k s\rangle_{\mathcal{R}',\mathcal{R}}\big)\,. \end{aligned} \tag{8.31}$$

Taking the norm on both sides and using the notation introduced above and estimates from Proposition 8.12 we derive

$$
\begin{aligned}
&|B(r_k,x_k)-B(s_k,y_k)| \\
&\le \; |\langle J(r_\lambda,x_\lambda)-J(s_\lambda,y_\lambda),\Delta_k u\rangle| + |\langle J(s_\lambda,y_\lambda),\Delta_k u-\Delta_k v\rangle| \\
&\quad +|\langle K(r_\lambda,x_\lambda)-K(s_\lambda,y_\lambda),\Delta_k r\rangle_{\mathcal{R}',\mathcal{R}}| + |\langle K(s_\lambda,y_\lambda),\Delta_k r-\Delta_k s\rangle_{\mathcal{R}',\mathcal{R}}| \\
&\quad +|B(r_{k-1},x_{k-1})-B(s_{k-1},y_{k-1})| \\
&\le \; C_J|\Delta_k u|\,(\|r_\lambda-s_\lambda\|_{\mathcal{R}}+|x_\lambda-y_\lambda|)+|J(s_\lambda,y_\lambda)|\,|\Delta_k(u-v)| \\
&\quad + C_K\|\Delta_k r\|_{\mathcal{R}}\,(\|r_\lambda-s_\lambda\|_{\mathcal{R}}+|x_\lambda-y_\lambda|)+\|K(s_\lambda,y_\lambda)\|_{\mathcal{R}^*}\,\|\Delta_k(r-s)\|_{\mathcal{R}} \\
&\quad +|B(r_{k-1},x_{k-1})-B(s_{k-1},y_{k-1})| \\
&\le \; C_J|\Delta_k u|\,(\|r_\lambda-s_\lambda\|_{\mathcal{R}}+|x_\lambda-y_\lambda|)+\frac{1}{c}|\Delta_k(u-v)| \\
&\quad + C_K\|\Delta_k r\|_{\mathcal{R}}\,(\|r_\lambda-s_\lambda\|_{\mathcal{R}}+|x_\lambda-y_\lambda|)+K_0\,\|\Delta_k(r-s)\|_{\mathcal{R}} \\
&\quad +|B(r_{k-1},x_{k-1})-B(s_{k-1},y_{k-1})|\;.
\end{aligned}
\tag{8.32}
$$

Then by definition of x_λ, r_λ, y_λ, s_λ we have

$$
\begin{aligned}
&|B(r_k,x_k)-B(s_k,y_k)| \\
&\le \; (C_J\,|\Delta_k u|+\;C_K\,\|\Delta_k r\|_{\mathcal{R}})\left(\max_{i\in\{k,k-1\}}\|r_i-s_i\|_{\mathcal{R}}+\max_{i\in\{k,k-1\}}|x_i-y_i|\right) \\
&\quad + K_0\,\|\Delta_k(r-s)\|_{\mathcal{R}}+\frac{1}{c}\,|\Delta_k(u-v)| \\
&\quad +|B(r_{k-1},x_{k-1})-B(s_{k-1},y_{k-1})|\;.
\end{aligned}
\tag{8.33}
$$

This implies the desired inequality (8.30).

Case 2 - $\Delta_k\xi\neq 0=\Delta_k\eta$*:* Then $x_k\in\partial Z(r_k)$. By (8.28) and the inequality $|J(r_k,x_k)|\ge\frac{1}{C}$ we have

$$
|\Delta_k\xi|=\frac{|A_k[r,u]|}{|J(r_k,x_k)|}\le C|A_k[r,u]|
$$

and the first inequality is proven. The second inequality is slightly more involved. First we notice that $B(r_k,x_k)=\frac{1}{2}\ge B(s_k,y_k)$. As above we can write

$$
\begin{aligned}
&\big(B(r_k,x_k)-B(s_k,y_k)\big)-\big(B(r_{k-1},x_{k-1})-B(s_{k-1},y_{k-1})\big) \\
&= \; \langle J(r_\lambda,x_\lambda),\Delta_k x\rangle+\langle K(r_\lambda,x_\lambda),\Delta_k r\rangle_{\mathcal{R}',\mathcal{R}}-\big(\langle J(s_\lambda,y_\lambda),\Delta_k y\rangle \\
&\quad +\langle K(s_\lambda,y_\lambda),\Delta_k s\rangle_{\mathcal{R}',\mathcal{R}}\big) \\
&= \; \big(\langle J(r_\lambda,x_\lambda),\Delta_k u\rangle-\langle J(s_\lambda,y_\lambda),\Delta_k v\rangle\big)+\big(\langle K(r_\lambda,x_\lambda),\Delta_k r\rangle_{\mathcal{R}',\mathcal{R}} \\
&\quad -\langle K(s_\lambda,y_\lambda),\Delta_k s\rangle_{\mathcal{R}',\mathcal{R}}\big)-\big(\langle J(r_\lambda,x_\lambda),\Delta_k\xi\rangle-\langle J(s_\lambda,y_\lambda),\Delta_k\eta\rangle\big)
\end{aligned}
\tag{8.34}
$$

We add $A_k[r,u]-A_k[s,v]$ on both sides of this identity. Apart from the same term that we already estimated on the right hand side of (8.31), it remains to estimate on the right hand side of (8.34) the term

$$
(A_k[r,u]-A_k[s,v])-(\langle J(r_\lambda,x_\lambda),\Delta_k\xi\rangle-\langle J(s_\lambda,y_\lambda),\Delta_k\eta\rangle)\;. \tag{8.35}
$$

Though $A_k[s,v] = 0 = \langle J(s_\lambda, y_\lambda), \Delta_k \eta \rangle$, we decide to work with this term as well to have the estimate at hand for the general case.

First we rewrite (8.35) as

$$\begin{aligned}&\langle J(r_k, x_k) - J(s_k, y_k) - J(r_\lambda, x_\lambda) + J(s_\lambda, y_\lambda), \Delta_k \xi \rangle \\ &\quad + \langle J(s_k, y_k) - J(s_\lambda, y_\lambda), \Delta_k(\xi - \eta) \rangle \ .\end{aligned} \tag{8.36}$$

From the triangle inequality it follows

$$\begin{aligned}&|\langle J(r_k, x_k) - J(s_k, y_k) - J(r_\lambda, x_\lambda) + J(s_\lambda, y_\lambda), \Delta_k \xi \rangle| \\ &\leq \ 2C_J \left(\max_{i \in \{k,k-1\}} \|r_i - s_i\|_{\mathcal{R}} + \max_{i \in \{k,k-1\}} |x_i - y_i| \right) |\Delta_k \xi| \ ,\end{aligned}$$

and

$$|\langle J(s_k, y_k) - J(s_\lambda, y_\lambda), \Delta_k(\xi - \eta) \rangle| \leq C_J \left(\|\Delta_k s\|_{\mathcal{R}} + |\Delta_k y| \right) |\Delta_k(\xi - \eta)| \ .$$

Furthermore we make use of the following simple inequality

$$|\Delta_k(\xi - \eta)| \leq 2 \left(\max_{i \in \{k,k-1\}} |x_i - y_i| + \max_{i \in \{k,k-1\}} |u_i - v_i| \right)$$

Thus by putting the identity

$$\begin{aligned}&|B(r_k, x_k) - B(s_k, y_k) + A_k[r,u] - A_k[s,v]| = B(r_k, x_k) - B(s_k, y_k) + A_k[r,u] \\ &\quad = |B(r_k, x_k) - B(s_k, y_k)| + |A_k[r,u] - A_k[s,v]|\end{aligned}$$

in what we calculated above we have

$$\begin{aligned}&|B(r_k, x_k) - B(s_k, y_k)| + |A_k[r,u] - A_k[s,v]| \\ &\leq \ (C_J |\Delta_k u| + C_K \|\Delta_k r\|_{\mathcal{R}} + 2C_J |\Delta_k \xi|) \max_{i \in \{k,k-1\}} \|r_i - s_i\|_{\mathcal{R}} \\ &\quad + (C_K \|\Delta_k r\|_{\mathcal{R}} + 2C_J (|\Delta_k u| + \|\Delta_k s\|_{\mathcal{R}} + |\Delta_k y| + |\Delta_k \xi|)) \max_{i \in \{k,k-1\}} |x_i - y_i| \\ &\quad + 2C_J (\|\Delta_k s\|_{\mathcal{R}} + |\Delta_k y|) \max_{i \in \{k,k-1\}} |u_i - v_i| \\ &\quad + K_0 \|\Delta_k (r - s)\|_{\mathcal{R}} + \frac{1}{c} |\Delta_k (u - v)| \\ &\quad + |B(r_{k-1}, u_{k-1}) - B(s_{k-1}, v_{k-1})| \ ,\end{aligned}$$

which is precisely (8.30).

Case 3 - $\Delta_k \xi = 0 \neq \Delta_k \eta$*:* This case is analogous to case 2.

Case 4 - $\Delta_k \xi \neq 0 \neq \Delta_k \eta$*:* For the first inequality we proceed as in the proof of [18, Lemma 5.3]:

$$\begin{aligned}|\Delta_k(\xi - \eta)| &= \left| \frac{A_k[r,u]}{|J(r_k, x_k)|^2} J(r_k, x_k) - \frac{A_k[s,v]}{|J(s_k, y_k)|^2} J(s_k, y_k) \right| \\ &\leq \ A_k[r,u] \left| \frac{J(r_k, x_k)}{|J(r_k, x_k)|^2} - \frac{J(s_k, y_k)}{|J(s_k, y_k)|^2} \right| + \frac{1}{|J(r_k, x_k)|} |A_k[r,u] - A_k[s,v]| \\ &\leq \ \frac{|\Delta_k \xi|}{|J(s_k, y_k)|} |J(r_k, x_k) - J(s_k, y_k)| + \frac{1}{|J(r_k, x_k)|} |A_k[r,u] - A_k[s,v]| \\ &\leq \ C |\Delta_k \xi| |J(r_k, x_k) - J(s_k, y_k)| + C |A_k[r,u] - A_k[s,v]| \ .\end{aligned}$$

We derive the second inequality exactly as in the cases before, having in mind that $B(r_k, x_k) = \frac{1}{2} = B(s_k, y_k)$. Lemma 8.21 is proved. □

As a direct consequence of Lemma 8.21, we get the following result. For simplicity, we use in the formulas below suitable positive constants $\alpha_1, \alpha_2, \dots$ depending only on C, C_K, C_J, K_0, and denote

$$V_k := \|\Delta_k r\|_{\mathcal{R}} + \|\Delta_k s\|_{\mathcal{R}} + |\Delta_k u| + |\Delta_k v| \; . \tag{8.37}$$

Corollary 8.22. *Under the conditions of Lemma 8.21 it holds for every $k \in [n]$*

$$\begin{aligned}
&|\Delta_k(\xi - \eta)| + C\Delta_k|B(r,x) - B(s,y)| \\
&\leq \; CK_0 \, \|\Delta_k(r-s)\|_{\mathcal{R}} + \frac{C}{c} \, |\Delta_k(u-v)| + \alpha_1 V_k \, |x_{k-1} - y_{k-1}| \\
&\quad + \alpha_2 V_k \left(\max_{i \in \{k,k-1\}} |u_i - v_i| + (1 + V_k) \max_{i \in \{k,k-1\}} \|r_i - s_i\|_{\mathcal{R}} \right) .
\end{aligned}$$

Proof. By multiplying the second inequality of Lemma 8.21 with C and adding it to the first inequality we derive the estimate

$$\begin{aligned}
&|\Delta_k(\xi - \eta)| + C\Delta_k|B(r,x) - B(s,y)| \\
&\leq \; \alpha_3 V_k \left(\max_{i \in \{k,k-1\}} \|r_i - s_i\|_{\mathcal{R}} + \max_{i \in \{k,k-1\}} |x_i - y_i| + \max_{i \in \{k,k-1\}} |u_i - v_i| \right) \\
&\quad + CK_0 \, \|\Delta_k(r-s)\|_{\mathcal{R}} + \frac{C}{c} \, |\Delta_k(u-v)| \; .
\end{aligned}$$

Furthermore, due to Proposition 8.13, we have

$$\begin{aligned}
&|x_k - y_k| = |Q_{r_k}(x_{k-1} + \Delta_k u) - Q_{s_k}(y_{k-1} + \Delta_k v)| \\
&\leq \; |x_{k-1} - y_{k-1}| + |\Delta_k(u-v)| + \alpha_4(1 + |\Delta_k \xi|)\|r_k - s_k\|_{\mathcal{R}} \\
&\leq \; |x_{k-1} - y_{k-1}| + 2 \max_{i \in \{k,k-1\}} |u_i - v_i| + \alpha_4(1 + |\Delta_k \xi|)\|r_k - s_k\|_{\mathcal{R}} \, ,
\end{aligned}$$

and we thus obtain the desired inequality. □

8.3.2. Excursion: Local Lipschitz continuity of the sweeping process

It is at this stage only a small step to show that the sweeping process depends locally Lipschitz continuous on the input data. We hence do not want to deprive the reader from the proof. First because this result might be of interest of itself in applications. And second because of the method of the proof. To show the result we first establish the assertion for stepfunctions by employing a variant of Gronwall's Lemma. We then extend it to the whole space of functions of bounded variation by a density argument. In a certain sense this is a technically much more simple prototype of the proof of Theorem 8.19 and might therefore be instructive to read. However, as it is not needed in the sequel, it may also be skipped.

Theorem 8.23. *Let Hypothesis 8.1 hold. Then there exist constants $\alpha, \beta, \gamma > 0$ depending only on C, C_J, C_K, K_0 such that for all $u, v \in BV_L(0,T;X)$, $r, s \in BV_L(0,T;\mathcal{R})$, $x_0 \in Z(r(0))$,*

$y_0 \in Z(s(0))$, the solutions ξ, η corresponding to (u, r, x_0), (v, s, y_0), respectively, satisfy the inequality

$$\begin{aligned}
&\mathrm{Var}(\xi - \eta) + C|B(r(T), x(T)) - B(s(T), y(T))| \qquad (8.38)\\
&\quad \leq \; \alpha \, \exp(\beta V) \left(\mathrm{Var}(r - s) + \mathrm{Var}(u - v)\right)\\
&\qquad + \gamma \, \exp(\beta V)(1 + V) \left(|x_0 - y_0| + \|u - v\|_\infty + (1 + W)\|r - s\|_\infty\right) ,
\end{aligned}$$

where $\|\cdot\|_\infty$ denotes the sup-norm, and

$$\begin{aligned}
V = V(r, s, u, v) &:= \mathrm{Var}(r) + \mathrm{Var}(s) + \mathrm{Var}(u) + \mathrm{Var}(v) \, ,\\
W = W(r, s, u, v) &:= \|r\|_\infty + \|s\|_\infty + \|u\|_\infty + \|v\|_\infty \, .
\end{aligned}$$

Proof. Corollary 8.22 yields

$$\begin{aligned}
&|\Delta_k(x - y)| + C\Delta_k|B(r, x) - B(s, y)|\\
&\quad \leq \; CK_0 \, \|\Delta_k(r - s)\|_{\mathcal{R}} + \left(1 + \frac{C}{c}\right) |\Delta_k(u - v)| + \alpha_1 V_k \, |x_{k-1} - y_{k-1}|\\
&\qquad + \alpha_2 V_k \left(\max_{i \in \{k, k-1\}} |u_i - v_i| + (1 + V_k) \max_{i \in \{k, k-1\}} \|r_i - s_i\|_{\mathcal{R}}\right) .
\end{aligned}$$

Using the inequality $|x_{k-1} - y_{k-1}| \leq |x_0 - y_0| + \sum_{i=1}^{k-1} |\Delta_i(x - y)|$, we are in the situation of Lemma 5.2 with

$$\begin{aligned}
a_k &= \alpha_1 V_k \, ,\\
\delta_k &= \sum_{i=1}^{k} |\Delta_i(x - y)| + C|B(r_k, x_k) - B(s_k, y_k)| \, ,\\
c_k &= CK_0 \, \|\Delta_k(r - s)\|_{\mathcal{R}} + \left(1 + \frac{C}{c}\right) |\Delta_k(u - v)| + \alpha_1 V_k |x_0 - y_0|\\
&\quad + \alpha_2 V_k \left(\max_{i \in \{k, k-1\}} |u_i - v_i| + (1 + V_k) \max_{i \in \{k, k-1\}} \|r_i - s_i\|_{\mathcal{R}}\right) .
\end{aligned}$$

Therefore we have

$$\begin{aligned}
\sum_{k=1}^{n} a_k &= \alpha_1 V \, ,\\
\delta_0 &= C|B(r_0, x_0) - B(s_0, y_0)| \leq \alpha_5(\|r(0) - s(0)\|_{\mathcal{R}} + |x_0 - y_0|) \, ,\\
\sum_{k=1}^{n} c_k &\leq CK_0 \mathrm{Var}(r - s) + \left(1 + \frac{C}{c}\right) \mathrm{Var}(u - v)\\
&\quad + \alpha_6 V \left(\|u - v\|_\infty + (1 + W) \, \|r_i - s_i\|_\infty + |x_0 - y_0|\right) .
\end{aligned}$$

Inequality (8.38) thus holds for all step functions u, v, r, s.

Now let $u, v \in BV_L(0,T;X)$, $r, s \in BV_L(0,T;\mathcal{R})$ be arbitrary. Choose $(u^n)_{n=1}^\infty, (v^n)_{n=1}^\infty \subset S_L(0,T;X)$, $(r^n)_{n=1}^\infty, (s^n)_{n=1}^\infty \subset S_L(0,T;\mathcal{R})$ such that $\{f^n(t) : t \in [0,T]\} \subset \{f(t) : t \in [0,T]\}$, $f^n(0) = f(0)$, and

$$\left(\mathrm{Var}(f^n) \leq \mathrm{Var}(f)\right) \ \wedge \ \left(\|f^n - f\|_\infty \overset{n\to\infty}{\to} 0\right),$$

where f stands successively for u, v, r, s. Moreover following the construction of [5, p. 237f] it is possible to choose $(u^n, v^n, r^n, s^n)_{n\in\mathbb{N}}$ such that $\mathrm{Var}(u^n - v^n) \leq \mathrm{Var}(u-v)$ and $\mathrm{Var}(r^n - s^n) \leq \mathrm{Var}(r-s)$. Let ξ^n, η^n be the respective solutions associated with (u^n, r^n, x_0) and (v^n, s^n, y_0). Then, according to [61], $\mathrm{Var}(\xi^n), \mathrm{Var}(\eta^n)$ are bounded independently of n, and

$$\|\xi^n - \xi\|_\infty \overset{n\to\infty}{\to} 0, \quad \|\eta^n - \eta\|_\infty \overset{n\to\infty}{\to} 0,$$

where ξ, η are the solutions corresponding to the data (u, r, x_0), (v, s, y_0), respectively. Inequality (8.38) holds for all elements of the sequence. Now notice on the one hand that Var(.) is lower semicontinuous with respect to the norm $\|\cdot\|_\infty$; on the other hand, $\|u^n - v^n\|_\infty \to \|u - v\|_\infty$, $\|r^n - s^n\|_\infty \to \|r - s\|_\infty$ and $\mathrm{Var}(u^n - v^n) \leq \mathrm{Var}(u - v)$, $\mathrm{Var}(r^n - s^n) \leq \mathrm{Var}(r - s)$. Hence, we may pass to the limit in (8.38) as $n \to \infty$ to finish the proof. □

8.3.3. Proof of Theorem 8.19

Assume that there exist stepfunctions $u^i, \eta^i \in S(0,T;X)$, $i \in [2]$ and $g \in S(0,T;C^1_{\omega,\gamma}(X \times X;\mathcal{R}))$ are stepfunctions with respect to the same subdivision such that they can be written in the form of (8.25). Let $u^1 = u^2 := u$ and $x_0^1 = x_0^2 := x_0$ and set

$$r^i : [0,T] \to \mathcal{R}, \ t \mapsto g(t, u(t), \eta^i(t)). \tag{8.39}$$

Then both r^i's are step functions which also can be denoted in the fashion of (8.25). We set ξ^i to be the solutions of the sweeping process associated with inputs u, r^i and x_0^i.

Corollary 8.24. *Assume that Hypothesis 8.1 holds, let c be as derived in Proposition 8.2 and denote*

$$V_k := \|\Delta_k g\|_{\mathcal{R}} + \left|\Delta_k \eta^1\right| + \left|\Delta_k \eta^2\right| + |\Delta_k u|. \tag{8.40}$$

Then there exist constants β_1, β_2 depending only on C, C_K, C_J, K_0, and γ such that for every $k \in [n]$

$$\begin{aligned} &\left|\Delta_k(\xi^1 - \xi^2)\right| + C\Delta_k|B(r^1, x^1) - B(r^2, x^2)| \\ \leq \ & CK_0\gamma\left|\Delta_k(\eta^1 - \eta^2)\right| + \beta_1 V_k \left|\xi^1_{k-1} - \xi^2_{k-1}\right| \\ & + \beta_2 V_k (1 + V_k)\left(\left|\eta^1_{i-1} - \eta^2_{i-1}\right| + \left|\Delta_k(\eta^1 - \eta^2)\right|\right). \end{aligned}$$

Proof. Note that since $u^1 = u^2$ all terms concerning u cancel and especially $x_k^1 - x_k^2 = \xi_k^2 - \xi_k^1$. Note furthermore that

$$\max_{i\in\{k,k-1\}} \|r_i^1 - r_i^2\|_{\mathcal{R}} \leq \left\|r^1_{i-1} - r^2_{i-1}\right\|_{\mathcal{R}} + \left\|\Delta_i(r^1 - r^2)\right\|_{\mathcal{R}}.$$

Applying the definition of r^i we can simply estimate the first term of the right hand side by

$$\left\|r^1_{i-1} - r^2_{i-1}\right\|_{\mathcal{R}} \leq \left\|g_{i-1}(u_{i-1}, \eta^1_{i-1}) - g_{i-1}(u_{i-1}, \eta^2_{i-1})\right\|_{\mathcal{R}} \leq \gamma\left|\eta^1_{i-1} - \eta^2_{i-1}\right|.$$

It remains to derive an estimate for $\left\|\Delta_i(r^1 - r^2)\right\|_{\mathcal{R}}$. This is easily done by

$$\begin{aligned}
&\left\|\Delta_i(r^1 - r^2)\right\|_{\mathcal{R}} \\
&= \left\|g_i(u_i, \eta_i^1) - g_i(u_i, \eta_i^2) - g_{i-1}(u_{i-1}, \eta_{i-1}^1) + g_{i-1}(u_{i-1}, \eta_{i-1}^2)\right\|_{\mathcal{R}} \\
&\leq \gamma\left(\left|\eta_i^1 - \eta_i^2\right| + \left|\eta_{i-1}^1 - \eta_{i-1}^2\right|\right) \\
&\leq \gamma\left|\Delta_i(\eta^1 - \eta^2)\right| + 2\gamma\left|\eta_{i-1}^1 - \eta_{i-1}^2\right| .
\end{aligned}$$

Now plugging all of this into Corollary 8.22 we obtain the estimate. □

Now everything is at hand to proceed to the main proof.

Proof of Theorem 8.19. If $\xi \in BV_L(0,T;X)$ is a solution of the quasivariational sweeping process then

$$\forall [r,s] \subset [0,T] : \mathrm{Var}(\xi,[r,s]) \leq S\big(\mathrm{Var}(g,[r,s]), \mathrm{Var}(u,[r,s])\big) . \tag{8.41}$$

Here S is the function defined in (8.21). Therefore let us define

$$\Omega := \{\xi \in BV_L^c(0,T;X) : \xi(0) = u(0) - x_0 \text{ and (8.41) holds}\} .$$

It is yet another straightforward proof to show that Ω is closed with respect to the BV-norm. Furthermore for any $\eta \in \Omega$ and $t \in [0,T]$ it holds that

$$|\eta(t+) - \eta(t)| \leq \frac{1}{1-\delta}(CK_0\|g(t+) - g(t)\| + (1 + CK_0\omega)|u(t+) - u(t)|) .$$

We define the operator $\mathcal{A} : \Omega \to BV_L$, $\mathcal{A} : \eta \mapsto \xi$, which maps any η to the solution of the sweeping process with input u, $g(\cdot, u(\cdot), \eta(\cdot))$ and x_0. Our aim is to prove that $\mathcal{A}$ is a contraction on Ω with respect to a weighted BV-norm. First note that $\mathcal{A}(\Omega) \subset \Omega$. Now choose $\eta^1, \eta^2 \in \Omega$ and $\varepsilon > 0$. Let $(t_n^f)_{n=0}^{N_f}$ be the corresponding approximating partition as constructed in Lemma 3.17 of $f \in \{u, g\}$. Furthermore choose a partition of $(t^i)_{n=1}^{N_i}$ be a partition of $[0,T]$ such that $t_0 = 0$, $t_n = N$ and

$$\forall n \in [N_i] : \forall s,t \in (t_{n-1}^i, t_n^i] : |\eta^i(s) - \eta^i(t)| \leq \varepsilon .$$

Let $(\widehat{t}_n)_{n=0}^{\widehat{N}} \in \mathcal{P}$ contain all above partitions, that is

$$\{\widehat{t}_n : n \in [\widehat{N}] \cup \{0\}\} \supset \{t_n^r : r \in \{f, g, 1, 2\} \wedge n \in [N_r] \cup \{0\}\} .$$

We now construct the approximating step functions $\widehat{f}$ for $f \in \{u, g, \eta^1, \eta^2\}$. To not overload the notation we drop the dependence on ε. We choose the sequence $(t_n)_{n=1}^N$ such that $N := 2\widehat{N}$ and

$$t_n = \widehat{t}_{n/2} \text{ for } n \text{ even and } t_n = \frac{1}{2}\left(\widehat{t}_{(n-1)/2} + \widehat{t}_{(n+1)/2}\right) \text{ for } n \text{ odd.}$$

Now define $\widehat{f}$ by

$$\widehat{f}_k = \begin{cases} f(\widehat{t}_{k/2}) & \text{for } k \text{ even} \\ f(\widehat{t}_{(k-1)/2}+) & \text{for } k \text{ odd} . \end{cases}$$

Let $\widehat{\xi^i}$ be the corresponding solutions of the sweeping process with input functions $\widehat{g}$, $\widehat{u}$, $\widehat{\eta}^i$. Due to Corollary 8.24 we have

$$\begin{aligned}
&\left|\Delta_k(\widehat{\xi}^1 - \widehat{\xi}^2)\right| + C\Delta_k|B(\widehat{r}^1, \widehat{x}^1) - B(\widehat{r}^2, \widehat{x}^2)| \\
&\leq \quad (CK_0\gamma + \beta_2 V_k(1+V_k))\left|\Delta_k(\widehat{\eta}^1 + \widehat{\eta}^2)\right| \\
&\qquad + \beta_1 V_k\left|\widehat{\xi}^1_{k-1} - \widehat{\xi}^2_{k-1}\right| + \beta_2 V_k(1+V_k)\left|\widehat{\eta}^1_{i-1} - \widehat{\eta}^2_{i-1}\right| .
\end{aligned}$$

where $V_k = |\Delta_k\widehat{u}| + |\Delta_k\widehat{g}| + |\Delta_k\widehat{\eta}^1| + |\Delta_k\widehat{\eta}^2|$. For $k \geq 1$ odd we get

$$\begin{aligned}
|\Delta_k\widehat{\eta}^i| &= |\eta^i(\widehat{t}_{(k-1)/2}) - \eta^i(\widehat{t}_{(k-1)/2}+)| \\
&\leq \quad S\left(\left|g(\widehat{t}_{(k-1)/2}+) - g(\widehat{t}_{(k-1)/2})\right|, \left|u(\widehat{t}_{(k-1)/2}+) - u(\widehat{t}_{(k-1)/2})\right|\right) \\
&\leq \quad S(c_g, c_u) .
\end{aligned}$$

If k is even we have

$$\begin{aligned}
|\Delta_k\widehat{\eta}^i| &= |\eta^i(\widehat{t}_{k/2}) - \eta^i(\widehat{t}_{(k-1)/2}+)| \\
&\leq \quad S\big(\mathrm{Var}(g, (t_{(k-1)/2}, t_{k/2}]), \mathrm{Var}(u, (t_{(k-1)/2}, t_{k/2}]))\big) .
\end{aligned}$$

Thus we estimate $V_k \leq \widetilde{V}_k$ with

$$\widetilde{V}_k = |\nabla_k\widehat{g}| + |\nabla_k\widehat{u}| + \begin{cases} S\big(|g(\widehat{t}_{(k-1)/2}+) - g(\widehat{t}_{(k-1)/2})|, \\ \quad |u(\widehat{t}_{(k-1)/2}+) - u(\widehat{t}_{(k-1)/2})|\big) & \text{for } k \text{ odd} \\ \\ S\big(\mathrm{Var}(g, (t_{(k-1)/2}, t_{k/2}]), \\ \quad \mathrm{Var}(u, (t_{(k-1)/2}, t_{k/2}]))\big) & \text{for } k \text{ even} . \end{cases}$$

Assume that $c_g, c_u \leq \nu$ and choose $\varepsilon \leq \frac{1}{4}\nu$. Then by means of (3.41) we have

$$\mathrm{Var}(g, (t_{(k-1)/2}, t_{k/2}]) \leq \nu \text{ and } \mathrm{Var}(u, (t_{(k-1)/2}, t_{k/2}])) \leq \nu .$$

Now we choose ν such that

$$\delta + \beta_2 2\,(\nu + S(\nu,\nu))\,(1 + 2\,(\nu + S(\nu,\nu))) =: \delta' < 1 . \tag{8.42}$$

Then $\left(CK_0\gamma + \beta_2\widetilde{V}_k\left(1 + \widetilde{V}_k\right)\right) \leq \delta'$ and applying Lemma 5.3 we get

$$\sum_{k=1}^{N} |\Delta_k(\widehat{\xi}^1 - \widehat{\xi}^2)|w_k \leq \rho \sum_{k=1}^{N} |\Delta_k(\widehat{\eta}^1 - \widehat{\eta}^2)|w_k . \tag{8.43}$$

for some $\rho < 1$ and

$$w_k = \exp\left\{-\frac{1}{\vartheta}\sum_{i=1}^{k}\widetilde{V}_i\right\} .$$

with ϑ small enough. For the given ε and $(t_k)_{k=1}^N$ define $V_f^\varepsilon : [0,T] \to \mathbb{R}$ by

$$V_f^\varepsilon : t \mapsto \begin{cases} \mathrm{Var}\,(f, [0, t_{k-1}]) + |f(t_{k-1}) - f(t_{k-1}+)| & \text{for } t \in (t_{k-1}, t_k], k \text{ odd} \\ \mathrm{Var}\,(f, [0, t_k]) & \text{for } t \in (t_{k-1}, t_k], k \text{ even} . \end{cases}$$

Then (8.43) can be rewritten into

$$\mathrm{Var}_{\widehat{w}}(\widehat{\xi}^1 - \widehat{\xi}^2) \leq \rho \mathrm{Var}_{\widehat{w}}(\widehat{\eta}^1 - \widehat{\eta}^2)$$

where $\widehat{w} := \exp\left(-(1/\delta)\widetilde{V}(t)\right)$ and $\widetilde{V} : [0,T] \to \mathbb{R}^+_{>0}$ is defined by

$$\widetilde{V}(t) = \mathrm{Var}(\widehat{g}, [0,t]) + \mathrm{Var}(\widehat{u}, [0,t]) + S\left(V^\varepsilon_g(t), V^\varepsilon_u(t)\right) .$$

Letting $\varepsilon \to 0$ we have

$$\widehat{\eta}^i \overset{\|\cdot\|_\infty}{\to} \eta^i, \widehat{g} \overset{\|\cdot\|_\infty}{\to} g \text{ and } \widehat{u} \overset{\|\cdot\|_\infty}{\to} u$$

and consequently also $\widehat{\xi}^i \overset{\|\cdot\|_\infty}{\to} \xi$. We now prove that

$$\widetilde{V} \overset{\|\cdot\|_\infty}{\to} V \tag{8.44}$$

where V is defined by

$$V(t) = \mathrm{Var}(g, [0,t]) + \mathrm{Var}(u, [0,t]) + S\big(\mathrm{Var}(g, [0,t]), \mathrm{Var}(u, [0,t])\big) .$$

First remember that $\mathrm{Var}(\widehat{f}, [0,\cdot]) \overset{\|\cdot\|_\infty}{\to} \mathrm{Var}(f, [0,\cdot])$ for $f \in \{g, u\}$ by the choice of the approximating sequence. Due to the construction of $\widehat{f}$ we have for $t \in (t_{k-1}, t_k)$ that $\mathrm{Var}(\widehat{f}, [0,t]) = \mathrm{Var}(\widehat{f}, [0,t_k])$ if k is odd and $\mathrm{Var}(\widehat{f}, [0,t]) = \mathrm{Var}(\widehat{f}, [0,t_{k-1}]) + |f(t_{k-1}+) - f(t_{k-1})|$ for k even. Hence we can estimate

$$\|\mathrm{Var}(\widehat{f}, [0,t]) - V^\varepsilon_f(t)\|_\infty \leq 2\varepsilon$$

and deduce $V^\varepsilon_f \overset{\|\cdot\|_\infty}{\to} \mathrm{Var}(f, [0,\cdot))$. Since $S(\cdot,\cdot)$ is nothing but a sum of two linear terms this implies (8.44). Therefore we have

$$\widehat{w} \overset{\|\cdot\|_\infty}{\to} w = \exp\{-(1/\delta)V(t)\}$$

and due to Proposition 3.14 we obtain

$$\mathrm{Var}_w(\xi^1 - \xi^2) \leq \liminf_\varepsilon \mathrm{Var}_{\widehat{w}}(\widehat{\xi}^1 - \widehat{\xi}^2) .$$

To estimate the right hand side we use

$$\begin{aligned}
&\limsup_\varepsilon (\mathrm{Var}_{\widehat{w}}(\widehat{\eta}^1 - \widehat{\eta}^2) - \mathrm{Var}_w(\eta^1 - \eta^2)) \\
&\leq \quad \limsup_\varepsilon (\mathrm{Var}_w(\widehat{\eta}^1 - \widehat{\eta}^2) - \mathrm{Var}_w(\eta^1 - \eta^2)) \\
&\qquad + \limsup_\varepsilon (\mathrm{Var}_{\widehat{w}}(\widehat{\eta}^1 - \widehat{\eta}^2) - \mathrm{Var}_w(\widehat{\eta}^1 - \widehat{\eta}^2)) .
\end{aligned}$$

The second term tends to zero as a consequence of (3.21). Due to the choice of $\widehat{\eta}^i$ we can calculate

$$\mathrm{Var}_w(\widehat{\eta}^1 - \widehat{\eta}^2) = \sum_{k=1}^N w(t_{k-1}+) \left|\widehat{\eta}^1(t_k) - \widehat{\eta}^2(t_k) - \left(\widehat{\eta}^1(t_{k-1}) - \widehat{\eta}^2(t_{k-1})\right)\right| .$$

For the sake of readability let us denote $\eta^1 - \eta^2$ by $\delta\eta$. If k is odd we have

$$\left|\widehat{\eta}^1(t_k) - \widehat{\eta}^2(t_k) - \left(\widehat{\eta}^1(t_{k-1}) - \widehat{\eta}^2(t_{k-1})\right)\right| = \left|\delta\eta(t_{k-1}+) - \delta\eta^2(t_{k-1})\right|$$

and if k is even we can evaluate that term to

$$\left|\widehat{\eta}^1(t_k) - \widehat{\eta}^2(t_k) - \left(\widehat{\eta}^1(t_{k-1}) - \widehat{\eta}^2(t_{k-1})\right)\right| = \left|\delta\eta(t_k) - \delta\eta(t_{k-2}+)\right| \ .$$

Furthermore we can bound the weighted total variation of $\xi^1 - \xi^2$ from below by

$$\begin{aligned}
&\mathrm{Var}_w(\xi^1 - \xi^2) \\
&\geq \sum_{\{k\in[N]:k=2r-1,r\in\mathbb{N}\}} w(t_{k-1}+)\left|\delta\eta(t_{k-1}+) - \delta\eta(t_{k-1})\right| \\
&\quad + \sum_{\{k\in[N]:k=2r,r\in\mathbb{N}\}} w(t_{k-1}+)\left|\delta\eta(t_{k-2}+) - \delta\eta(t_{k-1}+)\right| \\
&\quad + \sum_{\{k\in[N]:k=2r,r\in\mathbb{N}\}} w(t_k)\left|\delta\eta(t_{k-1}+) - \delta\eta(t_k)\right| .
\end{aligned}$$

Therefore we can estimate

$$\begin{aligned}
&\mathrm{Var}_w(\widehat{\eta}^1 - \widehat{\eta}^2) - \mathrm{Var}_w(\eta^1 - \eta^2) \\
&\leq \sum_{\{k\in[N]:k=2r,r\in\mathbb{N}\}} \left(w(t_{k-1}+) - w(t_k)\right)\left|\delta\eta(t_{k-1}+) - \delta\eta(t_k)\right| \\
&\leq \varepsilon\left(w(T) - w(0)\right) \ .
\end{aligned}$$

and finally obtain

$$\mathrm{Var}_w(\xi^1 - \xi^2) \leq \rho \mathrm{Var}_w(\eta^1 - \eta^2) \, .$$

This indeed is the contraction property and the Banach fixed point theorem grants the existence of a unique solution to Problem 6.8 within the set Ω. □

By simple translation one can proof Theorem 8.19 on any time interval $[r, s]$ with $-\infty < r < s < \infty$. Under the above conditions we denote the corresponding solution operator $\mathcal{S}$, that is

$$\mathcal{S}(x_0 - u(r), u, g, [r, s]) = \xi \tag{8.45}$$

where ξ is the unique solution of the quasivariational sweeping process with initial value x_0 and input u, g on the time interval $[r, s]$.

8.4. Proof of Theorem 8.3

We finally prove the main theorem (Theorem 8.3) of this chapter. We divide this task into the classical two parts. First we show the existence of a solution (Proposition 8.26) and afterwards prove its uniqueness (Proposition 8.28). We start by remembering the following simple property.

Proposition 8.25. *Let* $u \in BV_L(0, T; X)$ *then for each* $\varepsilon > 0$ *there exists an* $N = N(\varepsilon, \mathrm{Var}(u)) \in \mathbb{N}$ *such that*

$$\#\{t \in [0, T) : |u(t) - u(t+)| \geq \varepsilon\} \leq N \, .$$

For proving the existence of a solution to Problem 6.8 we are now going to explicitly construct it.

Proposition 8.26. *Assume that Hypothesis 8.1,* (8.1) *and* (8.2) *are satisfied. Then there exists a solution to the quasivariational sweeping process.*

Proof. Choose ν such that (8.42) is satisfied. By Proposition 8.25 we can choose $(t_n)_{n=0}^N \in \mathcal{D}_{[0,T]}$ such that $t_0 = 0$, $t_N = T$ and

$$\forall n \in [N]: \ \forall t \in (t_{n-1}, t_n): \ |u(t) - u(t+)| \le \nu \ \wedge \ |g(t) - g(t+)| \le \nu\,. \tag{8.46}$$

For any $f \in G(0,T;X)$ and each $n \in [N]$ we denote by

$$\widehat{f}^n : \begin{cases} [t_{n-1}, t_n] & \to & \mathcal{Y} \\ t & \mapsto & \begin{cases} f(t) & \text{for } t \in (t_{n-1}, t_n] \\ f(t_{n-1}+) & \text{for } t = t_{n-1} \end{cases} \end{cases}$$

a restriction of f to $[t_{n-1}, t_n]$ with no jump at its initial time. We set $\xi_0 = u(0) - x_0$ and $k = 1$. While $k \le N$ do

1. $\overline{\xi}_{k-1} := \mathcal{E}(\xi_{k-1}, u(t_{k-1}), g(t_{k-1}), u(t_{k-1}+), g(t_{k-1}+))$
2. $\widehat{\xi}^k := \mathcal{S}(\overline{\xi}_{k-1}, \widehat{u}^k, \widehat{g}^k, [t_{k-1}, t_k])$
3. $\xi_k := \widehat{\xi}^k(t_k)$, $k := k + 1$

We define $\xi : [0,T] \to X$ by

$$\xi(t) := \begin{cases} \xi_0 & \text{if } t = 0 \\ \widehat{\xi}^n(t) & \text{if } t \in (t_{n-1}, t_n] \end{cases}.$$

It remains to prove that ξ is a solution to Problem 6.8. First we note that $\xi \in BV_L(0,T;X)$. This is due to

$$\operatorname{Var}(\xi) \le \sum_{k=1}^{N} \left(\operatorname{Var}\left(\widehat{\xi}^k, [t_{k-1}, t_k]\right) + |\xi_{k-1} - \overline{\xi}_{k-1}|\right)$$

and the left-continuity of $\widehat{\xi}^k$. Due to (8.20) we have for all $n \in [N]$

$$\xi(t_{n-1}+) = \widehat{\xi}^n(t_{n-1}+) = \widehat{\xi}^n(t_{n-1}) = \overline{\xi}_{n-1}\,. \tag{8.47}$$

For $z \in \mathcal{T}(\xi)$ we compute

$$\begin{aligned} &\int_0^T \langle u(t+) - \xi(t+) - z(t), d\xi(t)\rangle \\ &= \sum_{n=1}^{N} \Bigg(\int_0^T \big\langle (u(t+) - \xi(t+) - z(t))\, \chi_{\{t_{n-1}\}}(t), d\xi(t) \big\rangle \\ &\qquad + \int_0^T \big\langle (u(t+) - \xi(t+) - z(t))\, \chi_{(t_{n-1}, t_n)}(t), d\xi(t) \big\rangle \Bigg)\,. \end{aligned}$$

The first term within can be evaluated by

$$\begin{aligned}
&\int_0^T \left\langle \left(u(t+) - \xi(t+) - z(t)\right) \chi_{\{t_{n-1}\}}(t), d\xi(t)\right\rangle \\
&\quad = \langle u(t_{n-1}+) - \xi(t_{n-1}+) - z(t_{n-1}), \xi(t_{n-1}+) - \xi(t_{n-1})\rangle \\
&\quad \geq 0\,.
\end{aligned}$$

Remember that $z(t_{n-1}) \in Z(g(t_{n-1}+, u(t_{n-1}+), \xi(t_{n-1}+)))$. The inequality is due to (8.47) and the definition of $\overline{\xi}_{n-1}$. For the second term we make use of Lemma 3.27 to see that

$$\begin{aligned}
&\int_0^T \left\langle \left(u(t+) - \xi(t+) - z(t)\right) \chi_{(t_{n-1},t_n)}(t), d\xi(t)\right\rangle \\
&\quad = \int_{t_{n-1}}^{t_n} \left\langle \left(\widehat{u}^n(t+) - \widehat{\xi}^n(t+) - z(t)\right) \chi_{(t_{n-1},t_n)}(t), d\widehat{\xi}^n(t)\right\rangle \\
&\quad \geq 0\,.
\end{aligned}$$

Here the inequality is a consequence of the fact that $\widehat{\xi}^n$ is, by definition, a solution of the quasi-variational sweeping process on the interval $[t_{n-1}, t_n]$. Thus we have for any $z \in \mathcal{T}(\xi)$ that

$$\int_0^T \langle u(t+) - \xi(t+) - z(t), d\xi(t)\rangle \geq 0$$

and the proof is complete. □

We now show that the solution constructed above is indeed a unique solution. Before we can do that we need another tool.

Lemma 8.27. *Let $\xi \in BV_L(0, T; X)$ be a solution of the quasivariational sweeping process on $[0, T]$. Choose any $[r, s] \subset [0, T]$ and define for $f \in \{g, u, \xi\}$ $\widehat{f} : [r, s] \to \mathcal{Y}$ (where $\mathcal{Y}$ is the appropriate space) by*

$$\widehat{f} : t \mapsto \begin{cases} f(r+) & \text{if } t = r \\ f(t) & \text{else} \end{cases}$$

Then $\widehat{\xi}$ solves the quasivariational sweeping process on $[r, s]$ with input functions $\widehat{g}$, $\widehat{u}$ and initial value $\xi(r+)$.

Proof. Obviously $\widehat{\xi}(t)$ has the desired initial value. Choose any $\widehat{z} \in G(r, s; X)$ such that $\widehat{z}(t) \in Z(g(t+, u(t+), \xi(t+)))$. Define

$$z(t) := \left(u(t+) - \xi(t+)\right) \chi_{[0,T]\setminus(r,s)}(t) + \widehat{z}(t)\chi_{(r,s)}(t)$$

Then due to Lemma 3.27 - note that ξ is left continuous - we have

$$0 \leq \int_0^T \langle u(t+) - \xi(t+) - z(t), d\xi(t)\rangle = \int_r^s \left\langle u(t+) - \widehat{\xi}(t+) - \widehat{z}(t), d\widehat{z}(t)\right\rangle.$$

Thus $\widehat{\xi}$ satisfies the quasi-variational inequality and the proof is complete. □

Proposition 8.28. *Assume that Hypothesis 8.1 holds. Furthermore let* (8.1) *and* (8.2) *be fulfilled. Then the solution to Problem 6.8 is unique.*

Proof. Let $(t_n)_{n=0}^N$ be the partition of $[0,T]$ corresponding to (8.46). Assume that $\xi, \eta \in BV_L(0,T;X)$ solve Problem 6.8 with $\eta \neq \xi$, that is there exists some $t \in [0,T]$ such that $\xi(t) \neq \eta(t)$. Let

$$\tau := \inf\{t \in [0,T] : \xi(t) \neq \eta(t)\}\,.$$

We make the convention that $\tau = T$ if $\xi \equiv \eta$. This is motivated by the fact that $\xi(\tau) = \eta(\tau)$. If $\tau = 0$ it is immediate, otherwise it is a consequence of the left continuity. We first show that if $\tau \in (t_{n-1}, t_n]$, then $\tau = t_n$. First notice $\xi(t) = \eta(t)$ for all $t \in (t_{n-1}, \tau]$ and especially $\xi(t_{n-1}+) = \eta(t_{n-1}+)$. Due to Lemma 8.27 both $\widehat{\xi^n}$ and $\widehat{\eta}^n$ solve the quasivariational evolution equation with initial value $\xi(t_{n-1}+)$. By Theorem 8.19 we know the solution to be unique, in other words

$$\widehat{\xi}^n = \mathcal{S}(\xi(t_{n-1}+), \widehat{u}^n, \widehat{g}^n, [t_{n-1}, t_n]) = \widehat{\eta^n}\,.$$

Therefore $\xi(t) = \eta(t)$ for all $t \in (t_{n-1}, t_n]$ and we can safely assume that $\tau = t_n$ for some $n \in [N] \cup 0$. If $n = N$ we already have $\xi \equiv \eta$. Otherwise by testing inequality (6.15) with

$$z(t) = (u(t+) - \xi(t+))\chi_{[0,T]\setminus\{t_n\}}(t) + y\chi_{\{t_n\}}(t)$$

for any $y \in Z(g(t_n+, u(t_n+), g(t_n+)))$ and the respective function for η we derive that

$$\xi(t_n+) = \mathcal{E}(\xi(t_n), u(t_n), g(t_n), u(t_n+), g(t_n+)) = \eta(t_n+)\,.$$

Due to Lemma 8.27 we know that $\widehat{\xi}^{n+1}$ and $\widehat{\eta}^{n+1}$ solve the quasivariational evolution equation on $[t_n, t_{n+1}]$ with initial value $\xi(t_n+)$. Since its solution is unique we know that $\xi(t) = \eta(t)$ for all $t \leq t_{n+1}$, hence $\tau \geq t_{n+1}$ - a contradiction. Therefore $\tau = t_N = T$ and $\xi \equiv \eta$. ☐

8.5. On the jump condition

In this section we discuss condition (8.1) restricting the jump size in Theorem 8.3 and whether it is necessary. We start by showing that (8.1) alone is not sufficient to guarantee uniqueness. However in some cases (8.2) can be replaced by different structural conditions for which we will give two examples. As we see from the above proof it suffices to consider Problem 8.6 which is the critical part. If we can guarantee a unique solution for it, we can guarantee a unique solution for Problem 6.8 as well.

We give an example of non uniqueness of the quasivariational inequality for $CK_0\gamma$ arbitrarily small. Thus condition (8.1) is not sufficient and an additional one has to be chosen.

Example 8.29. Choose $f \in C^\infty(\mathbb{R})$ such $f(x) = x$ in $[-1/2, 1/2]$, $\|f\|_\infty \leq 1$ and $\|f'\|_\infty \leq 1$. We furthermore choose the spaces $X = \mathbb{R}^2$ with $\mathcal{R} = \mathbb{R}$. Consider polyhedra of the following type

$$Z(r) = \left\{x \in \mathbb{R}^2 : A(r)x \leq w(r)\right\}$$

with

$$A(r) = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & -1 \\ f(r) & l \end{pmatrix}\,,\; w(r) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 - f(r)^2 \end{pmatrix}\,.$$

The real number $l > 1$ remains to be chosen. Note that for $l > 1$ we have $0 \in Z(r)$ and we can calculate the Minkowski functional of $Z(r)$ by

$$M(r, x) = \max\left\{\frac{a_i \cdot x}{w_i}, i \in [4]\right\}.$$

Here a_i denotes the i-th row of A. Therefore it is easy to compute that

$$\partial_r M(r, x) = \frac{\left\langle \begin{pmatrix} f'(r) \\ 0 \end{pmatrix}, x \right\rangle}{(l - f(r)^2)} - 2\frac{\left\langle \begin{pmatrix} f(r) \\ l \end{pmatrix}, x \right\rangle f(r)f'(r)}{(l - f(r)^2)^2}$$

if $4 \in \operatorname{argmax}\left\{\frac{a_i \cdot x}{w_i}, i \in [4]\right\}$ and 0 otherwise. For $l \geq 2$ we have

$$1/4|a_4(r)|^2 = 1/4f(r)^2 + 1/4l^2 \leq 1/2l^2 \leq \left(l - f(r)^2\right)^2$$

and therefore $B_{1/2}(0) \subset K(r)$. Moreover we see that $B_2(0) \supset K(r)$. Thus we can set $C = 2$. Using Lemma 3.1 of [18] we have $M(r, x) \leq 2|x|$ for any $x \in \mathbb{R}^2$ and therefore $M(r, x) \leq 4$ for all $x \in B_C(0)$. Therefore and by our above calculations we can estimate

$$|K(r, x)| \leq c\frac{1}{l - 1}$$

where c is a suitable constant. Setting $g : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $(u, \xi) \mapsto u - \xi$ we have $\gamma = 1$. Thus we finally obtain

$$\delta = CK_0\gamma \leq c\frac{1}{l - 1}.$$

Choosing l arbitrary large we can make δ arbitrary small. However setting $x^0 = (0, 1)$, $u^0 = (0, 0)$, $u^1 = (0, l)$ we have that any $x^1 = (z, 1)$ with $z \in [-1/2, 1/2]$ is a solution to

$$x_1 = Q_{Z(g(u^1, u^1 - x^1))}(x^0 + \Delta u).$$

Note that the set $Z(r)$ a polyhedron and thus does not satisfy Hypothesis 8.1. However by smoothing the edges (let's say for $|x_1| > 3/4$) there is no essential change in the behavior and the aforementioned conditions are satisfied.

As the above example shows a condition on the jump size is in general necessary. This however may in certain applications be unpleasant. Therefore we will present two examples where this condition can be replaced or dismissed if stronger assumptions on the convex sets $Z(r)$ are made. The perhaps the simplest possible example is the restriction

$$Z(r) = B_{|r|}(0).$$

where $B_{|r|}(0) = \{x \in X : |x| \leq |r|\}$ is the closed ball with radius $|r|$. Furthermore we set

$$g : X \to \mathbb{R}^+ : |g(x) - g(y)| < |x - y| \ \ \forall x, y \in X, x \neq y.$$

Proposition 8.30. *For any $y \in X$ there exists a unique solution to the problem: Find ξ such that*

$$\xi = Q_{Z(g(\xi))}(y). \tag{8.48}$$

Proof. For uniqueness assume that both ξ,η solve the above problem. Without loss of generality let $|\xi| \le |\eta|$. It is straightforward to see that

$$\xi = \frac{y}{|y|}|\xi| \text{ and } \eta = \frac{y}{|y|}|\eta|$$

with $|\xi|, |\eta| \le |y|$. Therefore if $\xi \ne \eta$, then $|\xi| < |\eta| \le |y|$. In this case $Q_{K(g(\xi))}(y) \ne y$ and we immediately derive $|\xi| = g(\xi)$. Furthermore by definition $|\eta| \le g(\eta)$. We then derive

$$|\xi - \eta| \le |\eta| - |\xi| \le g(\eta) - g(\xi) < |\eta - \xi|\,,$$

a contradiction. To prove existence, denote by $Y = \{\lambda y : 1 \ge \lambda \ge 0\}$ the line between 0 and y. Let $\mathcal{A} : X \to Y$ be the continuous operator which maps $z \mapsto Q_{Z(g(z))}(y)$. Since Y is compact by the Schauder fixed point theorem $\mathcal{A}$ has a fixed point $\xi \in Y$, which solves (8.48). □

Things work out so smoothly since the problem reduces to a problem on the real line. In fact by a similar method one can show that the implicit sweeping process on the real line always has a unique solution for BV input functions provided that $CK_0\gamma < 1$. This should cause to no surprise since every closed convex set on the real line is an interval and hence a polyhedron. Therefore this case in fact is already covered by the results of Chapter 7.

However the above condition on $Z(r)$ clearly is a toy example for which little application comes in mind. We are going to present a condition, which allows for a greater variety of convex sets. We assume that for all $r \in \mathcal{R}$ there exists some $\vartheta > 0$ such that

$$|n(r,x) - n(r,y)| \ge \vartheta |x - y| \;\; \forall x, y \in \partial Z(r)\,. \tag{8.49}$$

In other words what we are asking for is that the curvature of the boundary is uniformly bounded from below for all r. With this we can sharpen the condition (8.1) such that a restriction on the jump size is no longer necessary. It in a certain sense shows that the Example 8.29 is indeed prototypical for all situations where the jump size plays a role as we made use of the fact that we project onto a straight line segment of the convex set.

Proposition 8.31. *Let Hypothesis 8.1 and* (8.49) *hold. If*

$$\left(CK_0 + \vartheta^{-1} CC_J (CK_0 + 2)\right)\gamma < 1$$

then there exists a unique solution to Problem 8.6.

Proof. The proof relies on a careful examination of the proof of Proposition 8.9. For $r, s \in \mathbb{R}$ let $Q_{r/s}$ denote the orthogonal projection onto $Z(r)$ and $Z(s)$ respectively. Our aim is to prove that

$$|Q_r(x) - Q_s(x)| \le \left(CK_0 + \vartheta^{-1} CC_J (CK_0 + 2)\right) \|r - s\|_{\mathcal{R}}\,. \tag{8.50}$$

If $x \in Z(s)$, the left hand side can be estimated from above by the Hausdorff distance $d_H(r, s)$. Lemma 8.13 then grants the claim. The same holds for $x \in Z(r)$. If $x \notin Z(s) \cup Z(s)$ we use equation (8.13), where it was shown that

$$|n\left(r, Q_r\left(x\right)\right) - n\left(s, Q_s\left(x\right)\right)| \le |n\left(r, Q_r\left(x\right)\right) - n\left(s, Q_s\left(Q_r\left(x\right)\right)\right)|\,.$$

Due to Proposition 8.13 we can estimate the right hand side from above by

$$|n\left(r, Q_r\left(x\right)\right) - n\left(s, Q_s\left(Q_r\left(x\right)\right)\right)| \le CC_J(d_H(r,s) + \|r - s\|_{\mathcal{R}})\,.$$

On the other hand the left hand side can be estimated from below by

$$\begin{aligned}&|n\left(r,Q_r\left(x\right)\right)-n\left(s,Q_s\left(x\right)\right)|\\ \geq\ & \left|n(r,Q_r\left(x\right))-\frac{J(r,Q_s\left(x\right))}{|J(r,Q_s\left(x\right))|}\right|-\left|\frac{J(r,Q_s\left(x\right))}{|J(r,Q_s\left(x\right))|}-\frac{J(s,Q_s\left(x\right))}{|J(s,Q_s\left(x\right))|}\right| .\end{aligned}$$

Note that the outer normal is only defined at the boundary, but the term $\frac{J(r,\cdot)}{|J(r,\cdot)|}$ is defined everywhere in X, coincides with the outer normal on the boundary and is 0-homogeneous. For $\nu = 1/M(r, Q_s(x))$ we have $\nu Q_s(x) \in \partial Z(r)$. Therefore we can estimate

$$\begin{aligned}&\left|n(r,Q_r\left(x\right))-\frac{J(r,Q_s\left(x\right))}{|J(r,Q_s\left(x\right))|}\right|\\ &= \ |n(r,Q_r\left(x\right))-n(r,\nu Q_s\left(x\right))|\\ &\geq \ \vartheta\,|Q_r(x)-Q_s(x)|-\vartheta\frac{|Q_s(x)|}{M(r,Q_s(x))}|M(r,Q_s(x))-1|\end{aligned}$$

Since $x \notin Z(s)$ we have $Q_s(x) \in \partial Z(s)$. Using [18, Lemma 3.1] we get $M(r, \frac{Q_s(x)}{|Q_s(x)|}) \geq C^{-1}$. If $M(r, Q_s(x)) \geq 1$ then

$$M^2(r,Q_s(x))-M(s,Q_s(x))^2 \geq M(r,Q_s(x))-1 \geq 0\,,$$

if $M(r, Q_s(x)) < 1$ then

$$M^2(r,Q_s(x))-M^2(s,Q_s(x)) \leq M(r,Q_s(x))-1 \leq 0\,.$$

Thus we have

$$|M(r,Q_s(x))-1| \leq |M^2(r,Q_s(x))-M(s,Q_s(x))^2| \leq K_0\|r-s\|_{\mathcal{R}}$$

and we finally obtain

$$\begin{aligned}&|n\left(r,Q_r\left(x\right)\right)-n\left(s,Q_s\left(x\right)\right)|\\ &\geq \ \vartheta\,|Q_r(x)-Q_s(x)|-\vartheta CK_0\|r-s\|_{\mathcal{R}}-C_JC\|r-s\|_{\mathcal{R}}\,.\end{aligned}$$

Putting all together we derive (8.50). An application of Banach's fixed point theorem now grants a unique solution to Problem 8.6. □

Going through the computations we see that under the conditions of Proposition 8.31 there exists a unique solution to Problem 6.8 for all $u \in BV_L(0, T; X)$, $g \in BV_L(0, , T; C^1_{\omega,\gamma}(X \times X; \mathcal{R}))$ and $x_0 \in Z(g(0, u(0), u(0) - x_0))$. However we refrain from doing so here as this would only result in more computations but at least in our opinion no new insight. Concluding this section we see that, if we allow for functions of bounded variation instead of only absolutely continuous functions, we pay by imposing a structural assumption in addition to the conditions of [18]. These additional assumptions can be either on the involved functions, as in the main part of our exposition, or on the convex sets, as we showed here.

Part III.

Variational methods for doubly nonlinear evolution equations

9. Introduction

Let $\mathcal{Y}$ be a Banach space, $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ a maximal monotone operator and $\mathcal{E} : [0,T] \times \mathcal{Y} \to (-\infty, \infty]$ a lower semicontinuous energy functional. The topic of the present part is the study of doubly nonlinear differential inclusions

$$\alpha\left(\dot{u}(t)\right) + \partial_u \mathcal{E}_t\left(u(t)\right) \ni 0 \quad \text{a.e. in } [0,T]\,. \tag{9.1}$$

In fact we are interested in approximation or as one might say stability results for the above evolution system. The question we are going to ask is what happens if the dissipation operator α is approximated (in a sensible way) by a family (α_n). If (u_n) is a family of corresponding solutions, do the limit points solve the (9.1) in some sense?
We will give positive answers to these questions in two settings.

- If $\mathcal{Y}$ is a reflexive Banach space and (α_n) a sequence of general maximal monotone operators converging the graph sense to α. Then all limit points of the sequence of corresponding solutions u_n solves the limit equation.
- If $\mathcal{Y}$ is not reflexive we need to assume that (α_n) is a sequence of cyclic maximal monotone operators. This means that for each α_n there exists a convex function Ψ_n such that $\alpha_n = \partial\Psi_n$. If Ψ_n converges to Ψ in the sense of Mosco convergence then we are able to prove similar results. As $\mathcal{Y}$ is not necessarily reflexive quite a few useful properties are missing. Therefore a number of technical points need to be addressed.

In the following we shall go into more detail about doubly nonlinear differential inclusions and how they can be approached by variational methods. The precise statement of the results will be part of the following two chapters. Part of the results presented here has been obtained in joint work with Riccarda Rossi and Ulisse Stefanelli.

We would like to remark that the notation used here differs from the notation of the previous part. The differences and reasons therefore are described in Section 9.4. For a reader only looking at this part and used to the notation employed by e.g., Mielke, Rossi, Stefanelli, Visintin and others in the context of rate-independent systems the notation will appear familiar.

9.1. Doubly nonlinear differential inclusions

The term doubly nonlinear differential inclusions goes back to the seminal article [21] by P. Colli and A. Visintin. They studied equations of the form

$$A\dot{u}(t) + Bu(t) \ni f(t) \quad \text{a.e. in } [0,T]\,, \tag{9.2}$$

where both A and B are maximal monotone operators and at least one of them is cyclic, i.e. the subdifferential of a convex function. Existence of a solution was shown provided both operators satisfied growth conditions of the form

$$\forall (y,y') \in B : \quad \langle y', y \rangle_{\mathcal{Y}^*,\mathcal{Y}} \geq C|y|^2 - c \quad \text{and}$$
$$\forall (y,y') \in A : \quad \langle y', y \rangle_{\mathcal{Y}^*,\mathcal{Y}} \geq c_1|y|^p + c_2|y'|^q - c_3$$

for some $p, q > 1$. A number of papers later on studied generalizations of the above equation. The main attention focused on equations of the type

$$\partial\Psi\left(\dot{u}(t)\right) + \partial_u \mathcal{E}_t(u(t)) \ni 0\,, \quad \text{for a.a. } t \in [0,T] \tag{9.3}$$

where Ψ is a convex, lower semicontinuous functional and $\mathcal{E}$ a possibly non convex energy function. The operator ∂_u is meant to be a generalized, possibly set valued derivative with respect to state of the time dependent energy function $\mathcal{E}$. This focus was mainly driven by applications, such as elastoplasticity [48, 89, 41, 25], fracture propagation [26, 49, 14], phase transitions in shape memory alloys [77, 75, 6] and many others. On an abstract level it has recently been proven by A. Mielke, R. Rossi and G. Savare [69] that if the potential Ψ has superlinear growth, i.e.

$$\lim_{|x|\to\infty} \frac{\Psi(x)}{|x|} = +\infty\,, \tag{9.4}$$

the above differential inclusion has a solution under very general assumptions on the energy functional $\mathcal{E}$. Also the same authors studied an extension of doubly nonlinear inclusions to metric spaces [93]. To define a rate independent evolution in a metric space they generalized the framework of extending gradient flows to those spaces, see e.g. the recent monograph [3]. From an applications point of view, the limiting case of convex, 1-homogeneous Ψ, i.e. potentials satisfying

$$\forall \lambda \geq 0 : \quad \Psi(\lambda x) = \lambda \Psi(x)\,,$$

is of strong interest. It corresponds to the case where the evolution is rate independent. Indeed let u be sufficiently differentiable and $\varphi \in C^1(0,T)$ be a monotone increasing map with $\varphi([0,T]) = [0,T]$. Then assume $u(t)$ satisfies (9.3) then also $u \circ \varphi$ satisfies the doubly nonlinear inclusion with energy $\mathcal{E}_{\varphi(t)}$. To see this remember that if Ψ is 1-homogeneous, then its subdifferential $\partial\Psi$ is 0-homogeneous, meaning that for all $\lambda \geq 0$ it holds $\partial\Psi\left(\lambda x\right) \supset \partial\Psi(x)$. Hence for almost all $t \in [0,T]$

$$\partial\Psi\left(\frac{\mathrm{d}}{\mathrm{d}t} u\left(\varphi(t)\right)\right) = \partial\Psi\left(\dot{u}\left(\varphi(t)\right)\dot{\varphi}(t)\right) \ni \partial\mathcal{E}(u(\cdot)) \circ \varphi(t)\,.$$

Notice moreover that

$$\partial\mathcal{E}_{\varphi(t)}\left(u\left(\varphi(t)\right)\right) = \partial_u\mathcal{E}_{(\cdot)}(u(\cdot)) \circ \varphi(t)\,.$$

Hence if (9.3) is satisfied it as well holds that

$$\partial\Psi\left(\frac{\mathrm{d}}{\mathrm{d}t} u\left(\varphi(t)\right)\right) + \partial_u\mathcal{E}_{\varphi(t)}\left(u\left(\varphi(t)\right)\right) \ni 0\,, \quad \text{for a.a. } t \in [0,T]$$

A different interpretation is given by the following observation. Remember that

$$\xi \in \partial\Psi(u) \quad \Longleftrightarrow \quad u \in \partial\Psi^*(\xi)\,.$$

Moreover if Ψ is lower semicontinuous, bounded from below and 1-homogeneous, then there exists a closed convex set $K \in X$ such that $\Psi(x) = M_K(x)$. Here M_K is the Minkowski functional of K. In other words equation (9.3) is equivalent to

$$\dot{u}(t) \in I_{K^*}\left(-\partial_u \mathcal{E}_t(u(t))\right) . \tag{9.5}$$

This resembles the sweeping process we discussed in the section before. It is indeed nothing but the sweeping process in the case $\mathcal{E}_t(u) = \frac{1}{2}u^2 - f(t)u$.

The 1-homogeneous case is however mathematically challenging. The reason is that in lack of condition (9.4) the solution u may contain jumps, see e.g. [102]. Hence as (absolutely) continuous solutions can no longer be expected, a new concept of solution is necessary to allow for jumps. The first such notion we are aware of was proposed by A. Mielke and F. Theil in their seminal article [76]. They proposed the so called energetic solution, which reads as follows. A function $u : [0, T] \to X$ is called an energetic solution to (9.3) if and only if it satisfies

$$\forall t \in [0, T]: \qquad \mathcal{E}_t(u(t)) \leq \mathcal{E}_t(z) + \Psi(z - u(t)) \tag{9.6}$$

$$\forall t \in [0, T]: \qquad \mathcal{E}_t(u(t)) + \mathrm{Diss}_\Psi(u, [0, t)) = \mathcal{E}_0(u(0)) + \int_0^t \partial_s \mathcal{E}_s(u(s))\, \mathrm{d}s\, . \tag{9.7}$$

Here Diss_Ψ is essentially the variation of u with respect to Ψ. The first condition is called the stability condition. It states that at the the energy of the state $u(t)$ has to be less than the energy of any other state plus the amount of energy that would be dissipated by spontaneously jumping to this state. The second condition is an energy equality. The energy of the system at time t plus the energy dissipated until that point is equal to the initial energy plus the applied work. This formulation does not include any time derivatives of u and thus allows for solutions with jumps. It has been applied to a number of problems, see e.g. [63, 66, 96, 103]. However the energetic formulation has some drawbacks. Assuming once more that $\Psi = M_K$ the stability condition (9.6) implies that $-\partial_u \mathcal{E}_t(u(t)) \in K^*$ but the converse is not true - a counterexample can be found in [102]. This enforces the system to jump early. Other notion of solutions have been proposed, e.g. [72, 2, 73]. We start from a notion of solution for absolutely continuous functions. In the passage to the limit we arrive at a BV function and conditions this function has to satisfy. One may take these conditions to define a solution on BV.

9.2. Representative functions

The main theme of the calculus of variations is broadly speaking the study the existence and properties of (local) minimizers of functionals $\mathcal{I} : \mathcal{Y} \to (-\infty, \infty]$. The interest dates back to Bernoulli . Since then a vast toolbox has been developed. We are interested in the so called direct methods, which allow the direct study of the minimization problem, without the use of e.g. Euler-Lagrange equations. Key elements of this theory are lower semicontinuity, different types of convexity and coercivity. For a thorough introduction we refer to the monographs [23, 37].

In order to apply methods of the calculus of variations to the evolution problem (9.1) we need to rewrite it as a minimization problem. To this end let us shortly assume that α is a cyclic operator. Then there exists a convex function $\psi : \mathcal{Y} \to (-\infty, \infty]$ such that $\alpha = \partial\psi$. Moreover let ψ^* be the usual convex conjugate. The following equivalence is well known (see Remark 2.31)

$$y \in \partial\psi(x) \Longleftrightarrow \psi(x) + \psi^*(y) \leq \langle y, x\rangle_{\mathcal{Y}^*,\mathcal{Y}} \,. \tag{9.8}$$

Moreover for any $(x, y) \in \mathcal{Y} \times \mathcal{Y}^*$ it is true that $\psi(x) + \psi^*(y) \geq \langle y, x\rangle_{\mathcal{Y}^*,\mathcal{Y}}$. Therefore for any given x the problem to find $y \in \mathcal{Y}^*$ such that $y \in \alpha(x)$ is equivalent to find $y \in \mathcal{Y}^*$ such that

$$\mathcal{I}(y) := \psi(x) + \psi^*(y) - \langle y, x\rangle \leq \mathcal{I}(z) \quad \forall z \in \mathcal{Y}^* \quad \wedge \quad \mathcal{I}(y) = 0\,. \tag{9.9}$$

In fact the differential inclusion $y \in \partial\psi(x)$ is not equivalent to a pure minimization problem. We also require that the minimizer of $\mathcal{I}$ takes value is 0. The concept of Fitzpatrick functions provides a possibility to generalize the equivalence (9.8) to general maximal monotone operator. In fact what we are looking for is a function $f : \mathcal{Y} \times \mathcal{Y}^* \to (-\infty, \infty]$ such that

$$\left.\begin{array}{l} f : \mathcal{Y} \times \mathcal{Y}^* \to (-\infty, \infty] \text{is convex and lower semicontinuous,} \\ \forall (x, y) \in \mathcal{Y} \times \mathcal{Y}^* : f(x, y) \geq \langle y, x\rangle_{\mathcal{Y}^*,\mathcal{Y}} \quad \text{and} \\ y \in \alpha(x) \Longleftrightarrow f(x, y) = \langle y, x\rangle_{\mathcal{Y}^*,\mathcal{Y}} \,. \end{array}\right\} \tag{9.10}$$

Definition 9.1. Let $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ be a maximal monotone operator with $D(\alpha) \neq \emptyset$. A function $f : \mathcal{Y} \times \mathcal{Y}^* \to (-\infty, \infty]$ is called a representative function (of the operator α) if the conditions (9.10) are met.

It was S. Fitzpatrick in his work [36] who - to the authors knowledge - first developed a function which satisfied the above conditions. Its definition is as follows

Definition 9.2 (Fitzpatrick function)**.** Let $\mathcal{Y}$ be a reflexive Banach space and $\mathcal{Y}^*$ its dual. Let $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ be maximal monotone. The Fitzpatrick function $f_\alpha : \mathcal{Y} \times \mathcal{Y}^* \to (-\infty, \infty]$ of α is defined by

$$f(x, y) := \sup\Big\{\langle y, x_0\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \langle y_0, x_0 - x\rangle_{\mathcal{Y}^*,\mathcal{Y}} : x_0 \in D(\alpha), y_0 \in \alpha(x_0)\Big\} \,. \tag{9.11}$$

Let us first remark that f_α indeed satisfies the property (9.10). We shall in the following gather basic properties of Fitzpatrick functions needed in this work.

Notice that in general the Fitzpatrick f_α of a cyclic maximal monotone operator $\alpha = \partial\psi$ does not coincide with $\psi(\xi) + \psi^*(\xi^*)$. For example consider $\mathcal{Y} = \mathbb{R} = \mathcal{Y}^*$ and set α to be the subdifferential of $\psi(x) = \frac{1}{2}x^2$. It is easy to see that $\psi^*(y) = \frac{1}{2}(y)^2$. However the Fitzpatrick is

$$f_{\partial\psi}(x, y) = \frac{1}{4}x^2 + \frac{1}{4}y^2 + \frac{1}{2}xy\,. \tag{9.12}$$

This clearly shows that for a given maximal monotone operator α the function f satisfying (9.10) is not unique. Indeed there are a number of functions which satisfy these properties. One prominent and important example are the so called Penot functions.

Definition 9.3 (Penot function)**.** Let $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ be maximal monotone. Then the Penot function of α is defined by $\rho_\alpha : \mathcal{Y} \times \mathcal{Y}^* \mapsto (-\infty; \infty]$ as

$$\rho_\alpha := \Big(\langle\cdot,\cdot\rangle_{\mathcal{Y}^*,\mathcal{Y}} + I_\alpha(\cdot,\cdot)\Big)^{**} \tag{9.13}$$

Here I_α is the indicator function of the graph of α.

Remark 9.4. Among the representative functions the Fitzpatrick and the Penot are extremal, in the sense that for any given α and any representative function f it holds

$$f_\alpha \le f \le \rho_\alpha\,. \tag{9.14}$$

Moreover both functions are dual in the sense that

$$f_\alpha^* = \rho_{\alpha^{-1}} \quad \wedge \quad \rho_\alpha^* = f_{\alpha^{-1}}\,. \tag{9.15}$$

Proposition 9.5. *Let $\psi : \mathcal{Y} \to (-\infty, \infty]$ be convex, lower-semicontinuous and 1-homogeneous, i.e.*

$$\forall x \in D(\psi), \lambda > 0 : \;\; \lambda x \in D(\psi) \;\wedge\; \psi(\lambda x) = \lambda\psi(x)\,.$$

Then the Fitzpatrick of the subdifferential of ψ coincides with the sum of ψ and its convex conjugate, in formula

$$\forall(\xi, \xi^*) \in \mathcal{Y} \times \mathcal{Y}^* : f_{\partial\psi}(\xi, \xi^*) = \psi(\xi) + \psi^*(\xi^*)\,.$$

Proof. For all ψ convex, lower-semicontinuous and 1-homogeneous there exists a closed convex set $0 \in K \subset \mathcal{Y}$ such that $\psi(\xi) = M_K(\xi)$. Therefore we know that $\psi^*(\xi^*) = I_{K^*}(\xi^*)$. Notice that if $\xi^* \in \partial M_K(x)$ then

$$\langle x, x^* \rangle_{\mathcal{Y}^*,\mathcal{Y}} = M_K(x) + I_{K^*}(x)\,. \tag{9.16}$$

By definition of the Fitzpatrick we obtain

$$f_{\partial\psi}(\xi, \xi^*) = \sup\{\underbrace{\langle \xi^*, \xi_0 \rangle_{\mathcal{Y}^*,\mathcal{Y}} - M_K(\xi_0)}_{\le I_{K^*}(\xi^*)} + \underbrace{\langle \xi_0^*, \xi \rangle_{\mathcal{Y}^*,\mathcal{Y}} - I_K(\xi_0^*)}_{\le M_K(\xi)} : \; \xi_0^* \in \partial M_K(\xi_0)\} \tag{9.17}$$

This implies that $f_{\partial\psi}(\xi, \xi^*) \le M_K(\xi) + I_{K^*}(\xi^*)$. For the opposite inequality first assume that $\xi^* \notin K^*$. Then there exists an $\xi_0 \in K$ such that $\langle \xi^*, \xi_0 \rangle_{\mathcal{Y}^*,\mathcal{Y}} > 1$. Choose an arbitrary $\xi_0^* \in \partial M_K(\xi) = \partial M_K(\lambda\xi)$ for any positive $\lambda > 0$. Then

$$\begin{aligned} &\langle \xi^*, \lambda\xi_0 \rangle_{\mathcal{Y}^*,\mathcal{Y}} - M_K\lambda\xi_0 + \langle \xi_0^*, \xi \rangle_{\mathcal{Y}^*,\mathcal{Y}} - I_K\xi_0^* \\ &\ge \;\; \lambda\Big(\langle \xi^*, \xi_0 \rangle_{\mathcal{Y}^*,\mathcal{Y}} - 1\Big) + \langle \xi_0^*, \xi \rangle_{\mathcal{Y}^*,\mathcal{Y}} \to +\infty \;\; (\lambda \to \infty) \end{aligned}$$

Therefore $f_{\partial\psi}(\xi, \xi^*) \ge I_{K^*}(\xi^*)$. On the other hand choosing $\xi_0 = 0$ we obtain $\partial M_K(\xi_0) = K^*$ and deduce that

$$f_{\partial\psi}(\xi, \xi^*) \ge \sup\Big\{\langle \xi_0^*, \xi \rangle_{\mathcal{Y}^*,\mathcal{Y}} : \xi_0^* \in K^*\Big\} = M_K(\xi)\,. \tag{9.18}$$

In total we now obtain that $f_{\partial\psi}(\xi, \xi^*) \ge M_K(\xi) + I_{K^*}(\xi^*)$, which concludes the proof. □

The use of representative functions for our approach will be elaborated in the following. First we shall introduce another notion of variational convergence, the so called Γ-convergence. This particular notion of convergence was introduced by by Ennio de Giorgi in the 70s [30, 29]. For a thorough introduction to Γ-convergence and its applications we refer to the monographs [16, 24]. For our purposes we need the following definitions

Definition 9.6. Let $\mathcal{Y}$ be a Banach space. Let $(f_n)_{n\in\mathbb{N}}$ be a family of functions $f_n : \mathcal{Y} \to (-\infty, \infty]$.

- A function f is called the weak (strong) $\Gamma - \liminf$ of the sequence $(f_n)_{n\in\mathbb{N}}$, in formula

$$f = \Gamma_w - \liminf_{n\to\infty} f_n \qquad \left(f = \Gamma_s - \liminf_{n\to\infty} f_n\right) \tag{9.19}$$

if it holds for all $y \in \mathcal{Y}$

$$\begin{aligned} f(y) &= \inf\left\{\liminf_{n\to\infty} f_n(y_n) : (y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}\,,\ y_n \rightharpoonup y\right\} \quad \text{or} \\ f(y) &= \inf\left\{\liminf_{n\to\infty} f_n(y_n) : (y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}\,,\ y_n \to y\right\} \quad \text{respectively.} \end{aligned}$$

- A function f is called the weak (strong) Γ-$\limsup$, in formula

$$f = \Gamma_w - \limsup_{n\to\infty} f_n \qquad \left(f = \Gamma_s - \limsup_{n\to\infty} f_n\right) \tag{9.20}$$

if for all $y \in \mathcal{Y}$ it holds

$$\begin{aligned} f(y) &= \inf\left\{\limsup_{n\to\infty} f_n(y_n) : (y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}\,,\ y_n \rightharpoonup y\right\} \quad \text{or} \\ f(y) &= \inf\left\{\limsup_{n\to\infty} f_n(y_n) : (y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}\,,\ y_n \to y\right\} \quad \text{respectively.} \end{aligned}$$

- f is called the Γ-limit of (f_n) with respect to the weak (strong) topology on $\mathcal{Y}$ if

$$\begin{aligned} \Gamma_w - \liminf_{n\to\infty} f_n &= f = \Gamma_w - \limsup_{n\to\infty} f_n \quad \text{or} \\ \Gamma_s - \liminf_{n\to\infty} f_n &= f = \Gamma_s - \limsup_{n\to\infty} f_n \quad \text{respectively.} \end{aligned}$$

Remark 9.7. Notice that if $f = \Gamma_s - \limsup f_n$ then it holds for all $y \in \mathcal{Y}$ that

$$f(y) = \min\left\{\limsup_{n\to\infty} f_n(y_n) : (y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}\,,\ y_n \to y\right\}.$$

In other words the minimum is always attained. To this end let $y \in \mathcal{Y}$ such that $f(y) < \infty$. If $f(y) = \infty$ the claim is trivial. Then for each $k \in \mathbb{N}$ there exists a sequence $(y_{n,k})_{n\in\mathbb{N}}$ such that $y_{n,k} \to y$ as $n \to \infty$ and

$$\limsup_{n\to\infty} f_n(y_{n,k}) \le f(y) + 2^{-k}.$$

Showing that the minimum is always attained is now a simple diagonal sequence argument.

Theorem 9.8. *Let $\mathcal{Y}$ be a reflexive Banach space and $(\alpha_n)_{n\in\mathbb{N}}$ be a family of maximal monotone operators $\alpha_n : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$. Let $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ be also maximal monotone. Then the following are equivalent*

(1) $\alpha_n \xrightarrow{g} \alpha$

(2) $f_\alpha \le \Gamma_w - \liminf\limits_{n\to\infty} f_{\alpha_n}$

(3) $\rho_\alpha \ge \Gamma_s - \limsup\limits_{n\to\infty} \rho_{\alpha_n}$

Here $\alpha_n \xrightarrow{g} \alpha$ indicates graph convergence as introduced in Definition 2.43. Up to the authors knowledge the above theorem is new. A similar result that may spring to mind is the classical theorem that mosco convergence of a convex functions is equivalent to the graph convergence of their derivative, see e.g. [4, Theorem 3.66]. For a special class of representative functions (so called selfdual representatives) Ghossoub and coworkers moreover proved that Γ-convergence to a selfdual function implies Graph convergence of the represented operators to a maximal monotone operator [38].

Proof. **Step 1.** We first show that (1) ⇒ (2). To this end let $(y_n, y_n')_{n\in\mathbb{N}} \subset \mathcal{Y} \times \mathcal{Y}^*$ be a family such that

$$(y_n, y_n') \rightharpoonup (y, y') \quad \text{in } \mathcal{Y} \times \mathcal{Y}^* . \tag{9.21}$$

This is equivalent to $y_n \rightharpoonup y$ and $y_n', \rightharpoonup y$. Moreover choose $(x, x') \in \alpha$. Then as $\alpha_n \xrightarrow{g} \alpha$ there exists a sequence $(x_n, x_n')_{n\in\mathbb{N}}$, $(x_n, x_n') \in \alpha_n$ such that $x_n \to x$ and $x_n' \to x'$. We hence can deduce that

$$\begin{aligned}\liminf_{n\to\infty} f_{\alpha_n}\left(y_n, y_n'\right) &\geq \liminf_{n\to\infty}\Big(\left\langle y_n', x_n\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \left\langle x_n', x_n - y_n\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}\Big)\\ &= \left\langle y', x\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \left\langle x', x - y\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} .\end{aligned}$$

Taking the sup over all $(x, x') \in \alpha$ we deduce that (2) holds.

Step 2. To prove that (1) ⇒ (3) we use a little trick. First notice that $\alpha_n \xrightarrow{g} \alpha$ implies $\alpha_n^{-1} \xrightarrow{g} \alpha^{-1}$. Hence we deduce from the prior step that

$$\Gamma_w - \liminf_{n\to\infty} f_{\alpha_n^{-1}} \geq f_{\alpha^{-1}}.$$

Furthermore notice that as $\mathcal{Y} \times \mathcal{Y}^*$ is reflexive we can simply take the convex conjugate to obtain that

$$\rho_\alpha \geq \left(\Gamma_w - \liminf_{n\to\infty} f_{\alpha_n^{-1}}\right)^*$$

It then follows from [4, Theorem III.3.7] that

$$\rho_\alpha \geq \Gamma_s - \limsup_{n\to\infty} \rho_{\alpha_n}$$

Step 3. We will now show that (3) ⇒ (1). To this end fix $(y, y') \in \alpha$. Then there exists a sequence $(y_n, y_n')_{n\in\mathbb{N}} \subset \mathcal{Y} \times \mathcal{Y}^*$ such that $y_n \to y$, $y_n' \to y'$ and

$$\limsup_{n\to\infty} \rho_{\alpha_n}(y_n', y_n) \leq \rho_\alpha(y, y') = \left\langle y', y\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} . \tag{9.22}$$

Such a sequence exists due to the prior remark. In particular there exists a sequence $(\varepsilon_n)_{n\in\mathbb{N}} \subset [0, \infty)$, $\varepsilon_n \to 0$ such that

$$\rho_{\alpha_n}(y_n', y_n) \leq \left\langle y', y\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} + \varepsilon_n \tag{9.23}$$

Now notice that for all $n \in \mathbb{N}$ it holds $\rho^*_{\alpha_n} = f_{\alpha_n^{-1}}$ and thus

$$\rho^*_{\alpha_n}(y, y') \geq \left\langle y', y\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} \quad \wedge \quad f_{\alpha_n^{-1}}(y', y) \geq \left\langle y, y'\right\rangle_{\mathcal{Y},\mathcal{Y}^*} . \tag{9.24}$$

A generalization of the Bronstedt-Rockafellar approximation Lemma due to M. Marques Alves and B. F. Svaiter [67, Theorem 3.4] now grants that for all $n \in \mathbb{N}$ there exists $(x_n, x_n') \in \alpha_n$ such that

$$|x_n - y_n| \leq \sqrt{\varepsilon_n} \quad \wedge \quad |x_n' - y_n'| \leq \sqrt{\varepsilon_n}\,. \tag{9.25}$$

Then one can easily see that $x_n \to y$ and $x_n' \to y'$ which completes the proof.

Step 4. It remains to show that $(2) \Rightarrow (1)$. Here we once again use the little trick from before. Indeed by a similar argument as we used in Step 2 we know that (2) implies

$$\rho_{\alpha^{-1}} \geq \Gamma_s - \limsup_{n\to\infty} \rho_{\alpha_n^{-1}} \tag{9.26}$$

and we thus obtain from our most recent step that $\alpha_n^{-1} \xrightarrow{g} \alpha^{-1}$. This however already implies $\alpha_n \xrightarrow{g} \alpha$. □

9.3. Doubly nonlinear evolution equations as minimizations problems

The idea to write doubly nonlinear evolution equations as minimization problems is not new. It has been done for rate independent evolutions e.g. by U. Stefanelli [101] as a generalization of the Brezis-Ekeland principle. We will illustrate the idea behind this in a simplified example.

Let $\mathcal{E} \in C^1([0,T] \times \mathcal{Y})$ and $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ be maximal monotone. Assume furthermore that $u \in W^{1,1}(0,T;\mathcal{Y})$. Then by the properties of the Fitzpatrick function

$$\alpha\left(\dot{u}(t)\right) \ni -\nabla\mathcal{E}_t(u(t)) \quad \Longleftrightarrow \quad f_\alpha\left(\dot{u}(t), -\nabla\mathcal{E}_t(u(t))\right) \leq \langle -\nabla\mathcal{E}_t(u(t)), \dot{u}(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$$

Then as for all $(y,y') \in \mathcal{Y}$ it holds $f_\alpha(y,y') - \langle y', y\rangle_{\mathcal{Y}^*,\mathcal{Y}} \geq 0$ we obtain that

$$\begin{aligned} & f_\alpha\left(\dot{u}(t), -\nabla\mathcal{E}_t(u(t))\right) \leq \langle -\nabla\mathcal{E}_t(u(t)), \dot{u}(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} \quad \text{a.e. in } [0,T] \\ \Longleftrightarrow \quad & \int_0^T f_\alpha\left(\dot{u}(t), -\nabla\mathcal{E}_t(u(t))\right) + \langle \nabla\mathcal{E}_t(u(t)), \dot{u}(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}\ \mathrm{d}t \leq 0\,. \end{aligned}$$

Finally the integrating over $\langle \nabla\mathcal{E}_t(u(t)), \dot{u}(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}}$ we deduce that u solves (9.1) if and only if

$$\mathcal{E}_T(u(T)) + \int_0^T f_\alpha\left(\dot{u}(t), -\nabla\mathcal{E}_t(u(t))\right)\,\mathrm{d}t - \mathcal{E}_0(u(0)) - \int_0^T \partial_t\mathcal{E}_t(u(t))\,\mathrm{d}t \leq 0\,.$$

Moreover it holds for all $v \in W^{1,1}(0,T;\mathcal{Y})$ that

$$\mathcal{E}_T(v(T)) + \int_0^T f_\alpha\left(\dot{v}(t), -\nabla\mathcal{E}_t(v(t))\right)\,\mathrm{d}t - \mathcal{E}_0(u(0)) - \int_0^T \partial_t\mathcal{E}_t(v(t))\,\mathrm{d}t \geq 0\,.$$

Now we can define the functional $\mathcal{I} : W^{1,1}(0,T;\mathcal{Y}) \to [0,\infty]$ by

$$\mathcal{I} : v \mapsto \mathcal{E}_T(v(T)) + \int_0^T f_\alpha\left(\dot{v}(t), -\nabla\mathcal{E}_t(v(t))\right)\,\mathrm{d}t - \mathcal{E}_0(u(0)) - \int_0^T \partial_t\mathcal{E}_t(v(t))\,\mathrm{d}t\,.$$

Then finding a solution to (9.1) is equivalent to searching for a minimizer of $\mathcal{I}$ with the additional condition that the minimizer has to attain the value 0.

9.4. Remark on notation

An avid reader will notice that there is a change of notation in this part compared to the previous part. Whereas in the previous part the external load was described by the variable $u(t)$ and the state of the system by $\xi(t)$ in this part it is different. Here the state of the system is denoted by $u(t)$ and the adjoint state by $\xi(t)$. It is certainly uncommon to change the notation between two parts of the same work. So why has it been done here?

The main reason for this choice is to adhere to the conventions of the respective research areas. Though both parts are discussing rate independent processes the former part concerning 'classical' approaches is mainly investigate by the group around Pavel Krejčí and Martin Brokate. The approach in the present part using variational techniques has been shaped by the Alexander Mielke and coworkers. Thus slightly different notations arose. Most readers are - at least in the authors mind - probably only interested in one part. In order to make it easier for those readers to follow this thesis we decided for the change in notation.

To summarize, in the following the unknown state of the system will be denoted by $u(t)$. By $\xi(t)$ we will denote an element of the (generalized) subdifferential of $\mathcal{E}_t(u(t))$ with respect to the state u.

10. Approximation of maximal monotone operators on reflexive Banach spaces

Throughout this section we assume that $\mathcal{Y}$ is a separable, reflexive Banach space. Our aim is to give a very general setting of assumptions. Let $\mathcal{E} : [0,T] \times \mathcal{Y} \to (-\infty,\infty]$ be an energy function, $F : [0,T] \times \mathcal{Y} \to \mathcal{Y}^*$ and $P : [0,T] \times \mathcal{Y} \to \mathbb{R}$ be its generalized state and time derivatives respectively. We are going to study

$$\alpha\left(\dot{u}(t)\right) + F_t(u(t)) \ni 0 \quad \text{a.e. in } [0,T] \tag{10.1}$$

In the next section we will provide the framework in which we want to study the problem. Especially we are going to define what precisely is considered to be a solution of the above inclusion. We will then state the main result of this chapter. Roughly speaking it states that if $(\alpha_n)_{n\in\mathbb{N}}$ is a family of maximal monotone operators with $\alpha_n \xrightarrow{g} \alpha$ and $(u_n)_{n\in\mathbb{N}}$ is a corresponding family of solutions, then every limit point u will solve the limit equation. Before this result can be shown some preparations have to be done. This will be content of the third section. We will conclude this chapter by proving the main theorem.

10.1. Basic assumptions

For any function $g : [0,T] \times X \to Y$ where spaces X, Y are some spaces we are going to write $g_t(\cdot)$ when we are referring to $g(t,\cdot)$. Let us start by introducing the setting in which we are going to study the problem. We first formulate our precise assumptions on the energy functional.

Assumption 10.1 (Assumptions on the energy)**.** Define $\mathcal{G}(u) := \sup\{\mathcal{E}_t(u) : t \in [0,T]\}$. The triple $(\mathcal{E}, F, P)$ has the following properties:

Non degeneracy and lower semi continuity: The domain of $\mathcal{E}$ is of the form $D(\mathcal{E}) = [0,T]\times D$ and $F : [0,T] \times D \rightrightarrows \mathcal{Y}^*$ for some Borel set D. Furthermore we assume

$$\begin{gathered} u \mapsto \mathcal{E}_t(u) \text{ is l.s.c. for all } t \in [0,T], \\ \exists C_0 > 0 : \ \forall (t,u) \in [0,T] \times D : \mathcal{E}_t(u) \geq -C_0 \text{ and} \\ \text{graph}(F) \text{ is a Borel set of } [0,T] \times \mathcal{Y} \times \mathcal{Y}^* \,. \end{gathered} \tag{10.$\mathcal{E}_0$}$$

Coercivity: For all $t \in [0,T]$ it holds

$$u \mapsto \mathcal{E}_t(u) \text{ has compact sublevels} \tag{10.$\mathcal{E}_1$}$$

Lipschitz continuity: There exists $C_1 > 0$ such that for all $u \in D$ and $s,t \in [0,T]$ it holds

$$|\mathcal{E}_t(u) - \mathcal{E}_s(u)| \leq C_1 \mathcal{E}_t(u) |t-s| \,. \tag{10.$\mathcal{E}_2$}$$

Conditioned one-sided time-differentiability: There exists a function $P : \text{graph}(F) \to \mathbb{R}$ and a constant $C_2 > 0$ such that

$$\forall(t,u,\xi) \in \text{graph}(F) : \quad \liminf_{h \downarrow 0} \frac{\mathcal{E}_{t+h}(u) - \mathcal{E}_t(u)}{h} \leq P_t(u,\xi) \leq C_2 \mathcal{G}(u) \tag{10.$\mathcal{E}_3$}$$

Chain rule inequality: Let $u \in \text{BV}(0,T;\mathcal{Y})$ and $\xi \in L^1(0,T;\mathcal{Y}^*)$ such that

$$\sup_{t \in [0,T]} \mathcal{E}_t(u(t)) < \infty \quad \text{and} \quad \xi(t) \in F_t(u(t)) \text{ for a.a. } t \in [0,T]\,. \tag{10.2}$$

Then the map $e : t \mapsto \mathcal{E}_t(u(t))$ is of bounded variation. Furthermore let $\mathrm{d}u$ denote the Radon measure induced by u and $\mathrm{d}e$ the Radon measure induced by e then

$$\mathrm{d}e \ll \mathrm{d}u + \mathcal{L}|_{[0,T]} \tag{10.$\mathcal{E}_4$}$$

Moreover for almost all Lebesgue points t_0 of $\dot{u}_{ac}$ and $\dot{e}_{ac}$ it holds

$$\dot{e}_{ac}(t_0) \geq \langle \xi(t_0), \dot{u}_{ac}(t_0) \rangle_{\mathcal{Y}^*,\mathcal{Y}} + P_{t_0}\left(u(t_0), \xi(t_0)\right) \tag{10.$\mathcal{E}_5$}$$

Weak closedness: For all $t \in [0,T]$ and for all sequences $(u_n)_{n\in\mathbb{N}} \subset \mathcal{Y}$, $\xi_n \in F_t(u_n)$, $\mathbb{E}_n = \mathcal{E}_t(u_n)$ and $p_n = P_t(u_n, \xi_n)$ with

$$u_n \to u \text{ in } \mathcal{Y},\ \xi_n \rightharpoonup \xi \text{ in } \mathcal{Y}^*,\ p_n \to p \text{ and } \mathbb{E}_n \to \mathbb{E} \text{ in } \mathbb{R} \tag{10.3}$$

it holds

$$(t,u) \in D(F),\ \xi \in F_t(u),\ p \leq P_t(u,\xi) \text{ and } \mathbb{E} = \mathcal{E}_t(u)\,. \tag{10.$\mathcal{E}_6$}$$

Remark 10.2. We should take notice of the following facts

- There exists $C > 0$ such that for all $(t,u) \in [0,T] \times D$ it holds $\mathcal{G}(u) \leq C\mathcal{E}_t(u)$.
- Moreover there exists come $C > 0$ such that for all $(t,u,\xi) \in \text{graph}(F)$ it holds $P_t(u,\xi) \leq C\mathcal{G}(u)$. To prove this it suffices to show that there exists $C \in \mathbb{R}$ such that $P_t(u,\xi) \geq C|\mathcal{G}(u)|$. This however is a direct consequence of combining (10.$\mathcal{E}_2$) and (10.$\mathcal{E}_3$).
- The set of all Lebesgue points of both $\dot{u}_{ac}$ and $\dot{e}_{ac}$ has full measure.
- The above set of assumptions implies the assumptions of [69, Section 4.1] bar the *variational sum rule*. Indeed it needs to be shown that the chain rule inequality (4.E_5) of the aforementioned paper holds. To this end let $u \in W^{1,1}(0,T;\mathcal{Y})$ and $\xi \in L^1(0,T;\mathcal{Y}^*)$ such that (10.2) is satisfied. Then due to (10.$\mathcal{E}_4$) the map $t \mapsto \mathcal{E}_t(u(t))$ is absolutely continuous. The set of Lebesgue points of both $\dot{u}_{ac}$ and $\dot{e}_{ac}$ has full measure and especially the weak derivatives $\frac{\mathrm{d}}{\mathrm{d}t}u(t)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t))$ coincide almost everywhere with $\dot{u}_{ac}$ and $\dot{e}_{ac}$ respectively. Therefore (10.$\mathcal{E}_5$) implies the chain rule inequality of [69].

At some points stronger results can be shown if the chain rule inequality (10.$\mathcal{E}_5$) is replaced by the stricter chain rule equality

$$\dot{e}_{ac}(t_0) = \langle \xi(t_0), \dot{u}_{ac}(t_0) \rangle_{\mathcal{Y}^*,\mathcal{Y}} + P_{t_0}\left(u(t_0), \xi(t_0)\right) \tag{10.$\mathcal{E}_5^*$}$$

We will now precisely define what we understand to be a solution of the doubly nonlinear differential inclusion.

Definition 10.3. Let $(\mathcal{E}, F, T)$ satisfy Hypothesis 10.1 and assume that $\alpha : \mathcal{Y} \rightrightarrows \mathcal{Y}^*$ is a maximal monotone operator. Choose $u_0 \in D$. A function tuple $(u, \xi) \in W^{1,1}(0, T; \mathcal{Y}) \times L^1(0, T; \mathcal{Y}^*)$ is called a solution to the Cauchy problem

$$\alpha\left(\dot{u}(t)\right) + F_t(u(t)) \ni 0, \; u(0) = u_0$$

if and only if it satisfies $u(0) = u_0$, $\xi(t) \in F_t(u(t))$ for almost all $t \in [0, T]$ and the energy identity

$$\mathcal{E}_t(u(t)) + \int_0^t f_\alpha\left(\dot{u}(s), -\xi(s)\right) \,\mathrm{d}s = \mathcal{E}_0(u(0)) + \int_0^t P_s(u(s), \xi(s))\,\mathrm{d}s \tag{10.4}$$

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

Proposition 10.4. *If the tuple $(u, \xi) \in W^{1,1}(0, T; \mathcal{Y}) \times L^1(0, T; \mathcal{Y}^*)$ satisfies the above definition then for almost every $t \in [0, T]$ it holds*

$$\alpha\left(\dot{u}(t)\right) \ni -\xi(t)\,. \tag{10.5}$$

Proof. First notice that since $u \in W^{1,1}$ due to $(10.\mathcal{E}_4)$ the map $t \mapsto \mathcal{E}_t(u(t))$ is absolutely continuous as well. By definition we know that

$$\begin{aligned} &\int_0^T f_\alpha\left(\dot{u}(t), -\xi(t)\right) \,\mathrm{d}t \\ \leq \;\; &\mathcal{E}_0\left(u(0)\right) - \mathcal{E}_T\left(u(T)\right) + \int_0^T P_t\left(u(t), \xi(t)\right) \,\mathrm{d}t \leq \int_0^T \left\langle -\xi(t), \dot{u}(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} \,\mathrm{d}t\,. \end{aligned}$$

The last inequality is a result of $(10.\mathcal{E}_5)$. Hence due to (9.10) it holds that for almost all $t \in [0, T]$ we have

$$f_\alpha\left(\dot{u}(t), -\xi(t)\right) = \left\langle -\xi(t), \dot{u}(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}$$

which is equivalent to our assertion. □

The above result can be strengthened if one uses a stronger assumption on the chain rule.

Proposition 10.5. *If the tuple $(\mathcal{E}, F, P)$ satisfies Assumption 10.1 where $(10.\mathcal{E}_5)$ is replaced by $(10.\mathcal{E}_5^*)$ then a tuple $(u, \xi) \in W^{1,1}(0, T; \mathcal{Y}) \times L^1(0, T; \mathcal{Y}^*)$ is a solution in the sense of Definition 10.3 if and only if*

$$\xi(t) \in -\alpha\left(\dot{u}(t)\right) \cap F_t(u(t)) \quad \textit{for a.e. } t \in [0, T]\,.$$

Proof. One direction of the proof can be directly taken from above. The other direction uses the fact that now indeed

$$\begin{aligned} \int_0^t f_\alpha\left(\dot{u}(s), -\xi(s)\right) \,\mathrm{d}s \;&=\; \int_0^t \left\langle -\xi(s), \dot{u}(s)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} \,\mathrm{d}s \\ &=\; \mathcal{E}_0(u(0)) - \mathcal{E}_t(u(t)) + \int_0^t P_s\left(u(s), \xi(s)\right) \,\mathrm{d}s \end{aligned}$$

The second equality is due to ($10.\mathcal{E}_5^*$). This implies that (u, ξ) satisfies the conditions of Definition 10.3. □

Remark 10.6. Definition 10.3 can be understood as a minimization problem. Indeed it holds for all $(v, \eta) \in W^{1,1}(0,T;\mathcal{Y}) \times L^1(0,T;\mathcal{Y}^*)$ that

$$\mathcal{E}_t(v(t)) + \int_0^t f_\alpha\left(\dot{v}(s), -\eta(s)\right)\,\mathrm{d}s \geq \mathcal{E}_0(v(0)) + \int_0^t P_s(v(s), \eta(s))\,\mathrm{d}s$$

Hence when we subtract the right hand side, Definition 10.3 can be interpreted as a 0-minimization problem.

10.2. Main result

As a last prerequisite before stating the main theorem of this chapter we need to state our assumptions on the family of maximal monotone operators (α_n).

Assumption 10.7. Let $(\alpha_n)_{n\in\mathbb{N}}$, $\alpha_n : \mathcal{Y} \to \mathcal{Y}^*$ be a maximal monotone operator with $0 \in \alpha_n(0)$ for all $n \in \mathbb{N}$. Then we assume there exists $c_1, c_2, c_3 > 0$, $p \geq 1$ and $q > 1$ independent of $n \in \mathbb{N}$ such that

$$\forall n \in \mathbb{N} : \forall x \in D\left(\alpha_n\right), y \in \alpha_n(x) : \langle y, x\rangle_{\mathcal{Y}^*,\mathcal{Y}} \geq c_1|x|^p + c_2|y|^q - c_3\,. \tag{10.α_0}$$

Let us shortly make the following remark: If the family (α_n) complies with Assumption 10.7 and $\alpha_n \xrightarrow{g} \alpha$ than also α satisfies the growth condition

$$\forall x \in D\left(\alpha\right), y \in \alpha(x) : \langle y, x\rangle_{\mathcal{Y}^*,\mathcal{Y}} \geq c_1|x|^p + c_2|y|^q - c_3\,. \tag{10.6}$$

Our main result reads as follows.

Theorem 10.8. *Let the triplet $(\mathcal{E}, F, P)$ comply with Assumption 10.1. Furthermore let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of maximal monotone operators such that Assumption 10.7 is satisfied. Let $(u_n, \xi_n)_{n\in\mathbb{N}} \subset W^{1,1}(0,T;\mathcal{Y}) \times L^1(0,T;\mathcal{Y})$ be a family of functions such that (u_n, ξ_n) solves the differential inclusion*

$$\alpha_n(\dot{u}_n(t)) + F_t(u_n(t)) \ni 0, u_n(0) = u_0 \tag{10.7}$$

in the sense of Definition 10.3. Assume that $\alpha_n \xrightarrow{g} \alpha$. Then there exists a pair of functions $u \in \mathrm{BV}(0,T;\mathcal{Y})$ and $\xi \in L^1(0,T;\mathcal{Y}^*)$ with $u(0) = u_0$, $\xi(t) \in F_t(u(t))$ for almost all $t \in [0,T]$ such that up to a subsequence

$$u_n(t) \to u(t)\;\forall t \in [0,T]\,,\;\dot{u}_n \cdot \mathcal{L}|_{[0,T]} \overset{*}{\rightharpoonup} \mathrm{d}u \in \mathcal{M}(0,T;\mathcal{Y}^*) \tag{10.8}$$

Notice that the generalized derivative $\mathrm{d}u$ is a Banach space valued measure. The pair (u, ξ) satisfies for almost all $t \in [0,T]$ the energy inequality

$$\mathcal{E}_t(u(t)) + \int_0^t f_\alpha\left(\dot{u}_{ac}(\tau), -\xi(\tau)\right)\,\mathrm{d}\tau + \int_0^t f_\alpha^\infty\left(\dot{u}_s(\tau), 0\right)\|(\,\mathrm{d}\dot{u})_s\|(\tau)$$
$$\leq \mathcal{E}_0(u(0)) + \int_0^t P_\tau(u(\tau), \xi(\tau))\,\mathrm{d}\tau\,. \tag{10.9}$$

Here $(\mathrm{d}\dot{u})_s$ is the singular component of the Radon measure induced by u.

Notice that since f_α^∞ is 1-homogeneous (see Theorem 2.34) the result holds irrespective of the parametrization of the measure $(\,\mathrm{d}\dot{u})_s$. One can replace $\dot{u}_s$ and $\|(\,\mathrm{d}\dot{u})_s\|$ by $v : [0,T] \to \mathcal{Y}$ and $\nu \in \mathcal{M}(0,T)$ respectively as long as $(\,\mathrm{d}\dot{u})_s \ll \nu$ and $(\,\mathrm{d}\dot{u})_s = v \cdot \nu$.

Remark 10.9. Notice that we do not state any existence results for solutions to (9.1). This is not in the scope of our analysis. The author is aware of only two existence results which extend the original setting by Colli and Visintin [21] for general maximal monotone operators. Those are due to A. Segatti [99] and G. Akagi [1].

10.3. A selection argument on Young measures

Before we can proceed to the proof of Theorem 10.8 we need to introduce two Lemmas. First we present a small variation of Theorem B.1 from [69].

Lemma 10.10. *Let $u \in \mathrm{BV}(0,T;\mathcal{Y})$ satisfy*

$$\sup_{t\in[0,T]} \mathcal{E}_t(u(t)) < \infty\,,\ (t,u(t)) \in \mathrm{dom}(F) \textit{ for a.a. } t \in (0,T) \tag{10.10}$$

Moreover let $(\sigma_t)_{t\in[0,T]} \subset \mathcal{M}(\mathcal{Y}^* \times \mathbb{R})$ be a family Young measures such that

$$\textit{for a.a. } t \in [0,T] \quad \forall (\xi,p) \in \mathrm{supp}(\sigma_t) : \xi \in F_t(u(t))\,,\ p \le P_t(u(t),\xi(t)) \tag{10.11}$$

$$\int_0^T \int_{\mathcal{Y}^*\times\mathbb{R}} f_\alpha(\dot{u}_{ac}(s), -\zeta)\,\mathrm{d}\sigma(\zeta,p) ds \le \infty \tag{10.12}$$

Then the map $e : t \mapsto \mathcal{E}_t(u(t))$ has bounded variation and it holds for almost all $t \in [0,T]$ that t is Lebesgue point of $\dot{e}_{ac}$ and $\dot{u}_{ac}$ and it holds

$$\dot{e}_{ac}(t) \ge \int_{\mathcal{Y}^*\times\mathbb{R}} \langle \zeta, \dot{u}_{ac}(t_0)\rangle_{\mathcal{Y}^*,\mathcal{Y}} + p\,\mathrm{d}\sigma_t(\zeta,p)\,. \tag{10.13}$$

Proof. Let $K(t,u(t)) \subset \mathcal{Y} \times \mathbb{R}$ be the set of all $(\xi,p) \in \mathcal{Y} \times \mathbb{R}$ such that $\xi \in F_t(u(t))$ and $p \le P_t(u(t),\xi(t))$. Then there exists a family ξ_n, p_n of strongly measurable functions such that for almost all $t \in [0,T]$ it holds

$$\{\xi_n(t), p_n(t)\} \subset K(t,u(t)) \subset \overline{\{\xi_n(t), p_n(t)\}}\,. \tag{10.14}$$

This has been shown in the first step of the proof of [69, Theorem B.1]. Furthermore it can be shown that the family (ξ_n, p_n) can be chosen such that

$$\forall n \in \mathbb{N} : \xi_n \in L^1(0,T;\mathcal{Y}^*) \ \text{ and } \sup_{n\in\mathbb{N}} \int_0^T f_\alpha(\dot{u}_{ac}(t), -\xi_n(t))\,\mathrm{d}t < \infty\,. \tag{10.15}$$

To this end we define the function

$$g(t) := \inf\{f_\alpha(\dot{u}_{ac}(t), -\zeta) : (\zeta,p) \in K(t,u(t))\} \ \text{ for almost all } t \in [0,T] \tag{10.16}$$

Notice that due to (10.14) it holds

$$g(t) := \inf_{n\in\mathbb{N}}\{f_\alpha(\dot{u}_{ac}(t), -\xi_n(t))\} \text{ for a.a. } t \in [0,T] \tag{10.17}$$

and hence g is measurable. Moreover

$$\int_0^T g(t)\,\mathrm{d}t \leq \int_0^T \int_{\mathcal{Y}\times\mathbb{R}} f_\alpha(\dot{u}_{ac}(t), -\zeta)\,\mathrm{d}\sigma_t(\zeta, p)\,\mathrm{d}t < \infty. \tag{10.18}$$

Following the argument of [69, Theorem B.1] and [94, Lemma 3.4] we obtain (10.15). From there we deduce by using the chain rule inequality that the map $e : t \mapsto \mathcal{E}_t(u(t))$ is of bounded variation. Now there exists a set $\mathcal{T}_n$ of full measure such that

$$\dot{e}_{ac}(t) \geq \langle \xi_n(t), \dot{u}_{ac}(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} + p_n(t)\,. \tag{10.19}$$

Setting $\mathcal{T} = \bigcap_{n\in\mathbb{N}} \mathcal{T}_n$, which still has full measure, we deduce by density and convexity that for any $t \in \mathcal{T}$

$$\dot{e}_{ac}(t) \geq \langle \zeta, \dot{u}_{ac}(t)\rangle_{\mathcal{Y}^*,\mathcal{Y}} + p \tag{10.20}$$

for all $(\zeta, p) \in \overline{\text{conv}\, K(t, u(t))}$ which readily implies (10.13).

□

The following lemma is a variation of [69, Lemma B.2]

Lemma 10.11. *Let the triplet* $(\mathcal{E}, F, P)$ *comply with Assumptions 10.1 and the dissipation operator* α *satisfy the growth condition* (10.6). *Furthermore let* $u \in \mathrm{BV}(0, T; \mathcal{Y})$ *such that* (10.10) *holds. Suppose that the set*

$$S\,(t, u(t), \dot{u}_{ac}(t)) := \{(\zeta, p) \in \mathcal{Y}^* \times \mathbb{R} : \zeta \in F_t(u(t))\,,\ -\zeta \in \alpha(\dot{u}_{ac}(t))\,, p \leq P_t(u(t), \zeta)\} \tag{10.21}$$

is nonempty for almost all $t \in (0, T)$. *Then there exist measurable functions* $\xi : (0, T) \to \mathcal{Y}^*$ *and* $p : (0, T) \to \mathbb{R}$ *such that for almost all* $t \in [0, T]$

$$(\xi(t), p(t)) \in \text{argmin}\,\{f_\alpha(\dot{u}_{ac}(t), -\zeta) - p : (\zeta, p) \in S\,(t, u(t), \dot{u}_{ac}(t))\}\,. \tag{10.22}$$

Proof. The proof follows closely the lines of the proof of Lemma B.2 in [69] and only takes small deviations in order to account for the fact that α is not a cyclic operator. We shall first show that

$$\text{argmin}\,\{f_\alpha(\dot{u}_{ac}(t), -\zeta) - p : (\zeta, p) \in S\,(t, u(t), \dot{u}_{ac}(t))\} \neq \emptyset \ \text{ for a.a. } t \in [0, T] \tag{10.23}$$

To this end choose $(\zeta_n, p_n)_{n\in\mathbb{N}} \subset S\,(t, u(t), \dot{u}_{ac}(t))$ such that

$$\lim_{n\to\infty} f_\alpha(\dot{u}_{ac}(t), -\zeta_n) - p_n = \inf\,\{f_\alpha(\dot{u}_{ac}(t), -\zeta) - p : (\zeta, p) \in S\,(t, u(t), \dot{u}_{ac}(t))\} \tag{10.24}$$

Then there exist some constants $C, C' \geq 0$ such that

$$f_\alpha(\dot{u}_{ac}(t), -\zeta_n) \leq C + p_n \leq C + P_t\,(u(t), \zeta_n) \leq C' \tag{10.25}$$

The first inequality is due to the definition of (ζ_n, p_n), the second follows from the fact that $(\zeta_n, p_n) \in S(t, u(t), \dot{u}_{ac}(t))$. To see the last inequality notice that due to (10.$\mathcal{E}_3$) it holds $P_t\,(u(t), \zeta_n) \leq C_2\mathcal{G}(u(t))$. As $\mathcal{E}_t(u(t)) < \infty$ we can apply (10.$\mathcal{E}_2$) in order to see that $\mathcal{G}(u(t)) < \infty$. Applying the growth condition (10.6) we can now deduce that the family $(\zeta_n)_{n\in\mathbb{N}}$

is uniformly bounded in $\mathcal{Y}^*$. Also the family $(p_n)_{n\in\mathbb{N}}$ is bounded in $\mathbb{R}$ as one easily obtains from (10.25) that

$$-C \leq p_n \leq C' - C\,.$$

Hence there exists a (non relabeled) subsequence $(\zeta_n, p_n)_{n\in\mathbb{N}}$ and $\left(\overline{\zeta}, \overline{p}\right) \in \mathcal{Y}^* \times \mathbb{R}$ such that $\zeta_n \rightharpoonup \overline{\zeta}$ and $p_n \to \overline{p}$. Notice that due to the weak closedness condition (10.$\mathcal{E}_6$) it holds $\overline{\zeta} \in F_t(u(t))$ and $\overline{p} \leq P_t\left(u(t), \zeta\right)$. Moreover $-\overline{\zeta} \in \alpha\left(\dot{u}_{ac}(t)\right)$ as the graph of a maximal monotone operator is strongly-weakly closed. Hence we obtain that $\left(\overline{\zeta}, \overline{p}\right) \in S\left(t, u(t), \dot{u}_{ac}(t)\right)$. As the function f_α is weakly lower semicontinuous we deduce that

$$f_\alpha\left(\dot{u}_{ac}(t), -\overline{\zeta}\right) - \overline{p} \leq \lim_{n\to\infty} f_\alpha(\dot{u}_{ac}(t), -\zeta_n) - p_n$$

Hence we obtain that

$$\left(\overline{\zeta}, \overline{p}\right) \in \operatorname{argmin}\left\{f_\alpha(\dot{u}_{ac}(t), -\zeta) - p : (\zeta, p) \in S\left(t, u(t), \dot{u}_{ac}(t)\right)\right\} \tag{10.26}$$

Moreover notice that the set

$$\mathcal{S} := \left\{(t, u, v, \zeta, p) \in [0,T] \times \mathcal{Y} \times \mathcal{Y} \times \mathcal{Y}^* \times \mathbb{R} : \zeta \in F_t(u) \cap \left(-\alpha(v)\right), p \leq P_t(u, \zeta)\right\} \tag{10.27}$$

is a Borel set of $[0,T] \times \mathcal{Y} \times \mathcal{Y} \times \mathcal{Y}^* \times \mathbb{R}$. This is due to the fact that both graph(F) and α are Borel sets of $[0,T] \times \mathcal{Y} \times \mathcal{Y}^*$ and $\mathcal{Y} \times \mathcal{Y}^*$ respectively and P : graph$(F) \to \mathbb{R}$ is a Borel function. Moreover by assumption there exists a set $\mathcal{T} \subset [0,T]$ with full measure such that $S(t, u(t), \dot{u}_{ac}(t)) \neq \emptyset$. We define the graph of the multivalued function $t \in \mathcal{T} \mapsto S\left(t, u(t), \dot{u}_{ac}(t)\right) \subset \mathcal{Y}^* \times \mathbb{R}$ by

$$\mathrm{S} := \left\{(t, \zeta, p) \in \mathcal{T} \times \mathcal{Y}^* \times \mathbb{R} : (\zeta, p) \in S\left(t, u(t), \dot{u}_{ac}(t)\right)\right\}\,. \tag{10.28}$$

Using the definition of $\mathcal{S}$ from above we can see that the following identity holds

$$\mathrm{S} = \left\{(t, \zeta, p) \in \mathcal{T} \times \mathcal{Y}^* \times \mathbb{R} : (t, u(t), \dot{u}_{ac}(t)\zeta, p) \in \mathcal{S}\right\}\,. \tag{10.29}$$

Moreover notice that the function $u : (0,T) \to \mathcal{Y}$ and $\dot{u}_{ac} : (0,T) \to \mathcal{Y}^*$ are Borelian up to choosing a suitable representative in the equivalence class of $\dot{u}_{ac}$. Then the existence of a measurable selection (ξ, p) as defined in (10.22) is a consequence of [20, Theorem III.6 and Corollary III.3]. □

10.4. Proof of the main result

The proof of Theorem 10.8 combines ideas of the finite dimensional lower semicontinuity theorem [37, Theorem 5.27] and selection arguments from the proof of Theorem 4.4 in [69]. As usual we are going to use a constant C which may change from line to line but is independent of n.

Proof of Theorem 10.8. **Step 1 - A priori estimates and compactness:** Due to the assumption that $0 \in \alpha_n(0)$ it holds for all $n \in \mathbb{N}$ that $f_{\alpha_n} \geq 0$. Writing $E_n(t) = \mathcal{E}_t(u_n(t))$ we hence obtain

$$E_n(t) \leq \mathcal{E}_0(u_0) + \int_0^t C_3 E_n(s)\,\mathrm{d}s \quad \forall t \in [0,T]\,.$$

Employing Gronwall's Lemma we get $\sup\{E_n(t) : t \in [0,T]\} \leq C$. Using both assertions from Remark 10.2 we obtain that $\operatorname{ess\,sup}\{|P_t(u_n(t), \xi(t))| : t \in [0,T]\} \leq C$. Moreover due to (10.$\mathcal{E}_0$) and (10.$\alpha_0$) and the above estimates

$$\int_0^T c_1|\dot{u}_n(s)| + c_2|\xi_n(s)|^q \,\mathrm{d}s \leq \mathcal{E}_0(u_0) + \int_0^T C_3 E_n(t) + C_0 + c_3\,.$$

Therefore it holds $|\dot{u}_n|_{L^1(0,T;\mathcal{Y})} + |\xi_n(s)|_{L^q(0,T;\mathcal{Y}^*)} \leq C$. Using $u_n(0) = u_0$ we deduce that $|u_n|_{\mathrm{BV}(0,T;Vc)}$ is bounded. For the sake of simplicity we set $p_n(s) := P_s(u_n(s), \xi_n(s))$. Now notice that

$$h_n : t \mapsto \mathcal{E}_t(u_n(t)) - \int_0^t p_n(s)\,\mathrm{d}s$$

is monotone decreasing. To see this choose $0 \leq s < t \leq T$ and subtract the energy identity (10.4) for s from the energy equality for t. We then obtain

$$h_n(t) - h_n(s) = -\int_s^t f_{\alpha_n}(\dot{u}_n(\tau), -\xi_n(\tau))\,\mathrm{d}\tau \leq 0$$

Therefore we can easily compute

$$\mathrm{Var}(h_n) = \mathcal{E}_0(u_0) - \mathcal{E}_T(u_n(T)) + \int_0^T p_n(s)\,\mathrm{d}s \leq C\,.$$

Moreover since p_n is uniformly bounded in $L^\infty(0,T)$ we know that $\|\mathcal{E}_t(u_n(t))\|_{\mathrm{BV}} \leq C$.

Applying the Helly principle we can choose a subsequence $(n_k)_{k\in\mathbb{N}}$ and $u \in \mathrm{BV}(0,T;\mathcal{Y})$, $E \in \mathrm{BV}(0,T)$ such that

$$(u_{n_k}, \mathcal{E}_t(u_{n_k}(t))) \to (u(t), E(t)) \ \text{ in } \mathcal{Y} \times \mathbb{R}\,.$$

Passing to an additional (not relabeled) subsequence we know that there exists $\lambda \in \mathcal{M}(0,T;\mathcal{Y})$ such that

$$\dot{u}_{n_k} \cdot \mathcal{L} \overset{*}{\rightharpoonup} \lambda\,. \tag{10.30}$$

Here L denotes the Lebesgue measure on the interval $[0,T]$. Due to a generalization of Lebesgue's decomposition theorem [32, Theorem 9] there exist $\lambda_{ac}, \lambda_s \in \mathcal{M}(0,T;\mathcal{Y})$ such that

$$\|\lambda_{ac}\| \ll \mathcal{L} \quad \wedge \quad \|\lambda_s\| \perp \mathcal{L} \quad \wedge \quad \lambda = \lambda_{ac} + \lambda_s\,. \tag{10.31}$$

Since $\mathcal{Y}$ is reflexive it has the Radon-Nykodym property. We define the following Radon-Nykodym derivatives

$$\dot{u}_{ac} := \frac{\mathrm{d}\lambda_{ac}}{\mathrm{d}\mathcal{L}}\,,\ \dot{u}_s := \frac{\mathrm{d}\lambda_s}{\mathrm{d}\|\lambda_s\|}\,.$$

Finally an additional (not relabeled) subsequence grants a Young measure $(\sigma_t)_{t\in[0,T]}$, $\sigma_t \in \mathcal{M}(\mathcal{Y} \times \mathbb{R})$ such that for all $t \in [0,T]$ it holds that $\sigma_t(\mathcal{Y} \times \mathbb{R}) \leq 1$ and for almost all $t \in [0,T]$

it holds

$$\begin{aligned} \sigma_t(\mathcal{Y}\times\mathbb{R}) &= 1 \\ \text{supp}\,(\sigma_t) &\subset \bigcap_{j\in\mathbb{N}} \text{cl}\left(\{(\xi_{n_k}(t), p_{n_k}(t)) : k \geq j\}\right) . \end{aligned}$$

Here cl is the closure with respect to the weak topology on $\mathcal{Y}^* \times \mathbb{R}$. Moreover it holds

$$\begin{aligned} \xi_{n_k}(t) &\rightharpoonup \int_{\mathcal{Y}^*\times\mathbb{R}} \zeta \,\mathrm{d}\sigma_t(\zeta,p) =: \widehat{\xi}(t) \quad \text{in } L^q(0,T;\mathcal{Y}^*) \text{ and} \\ \partial_t\mathcal{E}_t(u_n(t)) &\overset{*}{\rightharpoonup} \int_{\mathcal{Y}^*\times\mathbb{R}} p \,\mathrm{d}\sigma_t(\zeta,p) =: \widehat{p}(t) \quad \text{in } L^\infty(0,T) . \end{aligned}$$

Step 2 - Existence of generalized derivatives: From now on, for the sake of simplicity, we shall write instead of n_k plainly k. Choose t such that $u_k(t) \to u(t)$, $\mathcal{E}_t(u_k(t)) \to E(t)$ and

$$\text{supp}\,(\sigma_t) \subset \bigcap_{j\in\mathbb{N}} \text{cl}\left(\{(\xi_{n_k}(t), p_{n_k}(t)) : k \geq j\}\right) .$$

The set of all $t \in [0,T]$ violating any of the above conditions has Lebesgue measure 0. Choose $(\zeta, p) \in \text{supp}\,(\sigma_t)$. Then there exists a further subsequence k_j such that $(\xi_{k_j}(t), p_{k_j}(t)) \rightharpoonup (\zeta, p)$. Due to (10.$\mathcal{E}_6$) we then obtain that for almost any $t \in [0,T]$ it holds

$$\begin{aligned} (t, u(t)) &\in D(F) \\ \mathcal{E}_t(u(t)) &= E(t) \\ \text{supp}\,(\sigma_t) &\subset \{(\zeta,p) \in \mathcal{Y}^* \times \mathbb{R} : \zeta \in F_t(u(t))\,,\ p \leq P_t(u(t),\zeta)\} . \end{aligned}$$

Hence for almost all $t \in [0,T]$ it holds that

$$\{(\zeta,p) \in \mathcal{Y}^* \times \mathbb{R} : \zeta \in F_t(u(t))\,,\ p \leq P_t(u(t),\zeta)\} \neq \emptyset .$$

Step 3 - $\liminf$ **result for the Fitzpatrick:** Notice that

$$t \mapsto f_{\alpha_k}\left(\dot{u}_k(t), -\xi_k(t)\right)$$

is bounded in $L^1(0,T)$. Therefore there exists a measure $\mu \in \mathcal{M}(0,T)$ such that (up to a subsequence)

$$f_{\alpha_k}\left(\dot{u}_k(\cdot), -\xi_k(\cdot)\right) \cdot \mathcal{L} \overset{*}{\rightharpoonup} \mu \ \text{ in } \mathcal{M}(0,T)$$

Employing [37, Corollary 1.116] we obtain that there exist μ_{ac}, μ_s, $\mu_\perp$ in $\mathcal{M}(0,T)$ such that

$$\begin{gathered} \mu_{ac} \ll \mathcal{L}\,,\ \mu_s \ll \|\lambda_s\|\,,\ \mu_\perp \perp \mathcal{L} + \|\lambda_s\| \text{ and} \\ \mu = \mu_{ac} + \mu_s + \mu_\perp \end{gathered}$$

Since $f_{\alpha_k}(\dot{u}_k(t), -\xi_k(t)) \geq 0$ we obtain $\mu_\perp \geq 0$. For any $t \in [0,T]$ we define

$$Q(t,\varepsilon) := \left(t_0 - \frac{\varepsilon}{2}, t_0 + \frac{\varepsilon}{2}\right) .$$

Choose $t_0 \in (0,T)$ such that $\sigma_{t_0}(\mathcal{Y}^* \times \mathbb{R}) = 1$,

$$\begin{aligned} \frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) &= \lim_{\varepsilon\to 0} \frac{\mu\left(Q(t_0,\varepsilon)\right)}{\mathcal{L}\left(Q(t_0,\varepsilon)\cap[0,T]\right)} < \infty\,, \\ \widehat{\xi}(t_0) &= \lim_{\varepsilon\to 0} \frac{1}{\mathcal{L}\left(Q(t_0,\varepsilon)\cap[0,T]\right)} \int\limits_{Q(t_0,\varepsilon)} \widehat{\xi}(t)\,\mathrm{d}t \qquad \text{and} \\ \dot{u}_{ac}(t_0) &= \lim_{\varepsilon\to 0} \frac{\lambda\left(Q(t_0,\varepsilon)\cap[0,T]\right)}{\mathcal{L}\left(Q(t_0,\varepsilon)\cap[0,T]\right)}\,. \end{aligned}$$

The set of t_0 for which at least one of these condition fails is a Lebesgue null set. Since $\alpha_k \overset{g}{\longrightarrow} \alpha$ also $\alpha_k^{-1} \overset{g}{\longrightarrow} \alpha^{-1}$ in the graph sense and we can apply Theorem 9.8 to obtain that for $\rho_{\alpha_k^{-1}} = f^*_{\alpha_k}$ the $\Gamma_s - \limsup$ inequality holds as $k \to \infty$. Therefore for any $(x,x^*) \in \mathcal{Y}\times\mathcal{Y}^*$ there exists a sequence $(x_k,x_k^*)_{k\in\mathbb{N}}$ such that $(x_k,x_k^*) \to (x,x^*)$ and

$$\limsup_{n\to\infty} \rho_{\alpha_k^{-1}}(x_k,x_k^*) \le \rho_{\alpha^{-1}}(x,x^*)\,. \tag{10.32}$$

Choose a sequence ε_m such that for all $m\in\mathbb{N}$

$$\mu\left(\partial Q(t_0,\varepsilon)\cap[0,T]\right) = \lambda\left(\partial Q(t_0,\varepsilon)\cap[0,T]\right) = 0\,. \tag{10.33}$$

Also notice that

$$\frac{\mathcal{L}\left(Q(t_0,\varepsilon)\right)}{\varepsilon} \to 1$$

for all $t_0 \in (0,T)$. Choose $(x,x^*) \in \mathcal{Y}\times\mathcal{Y}^*$. Without loss of generality we assume that $\rho_{\alpha^{-1}}(x,x^*) < \infty$. Furthermore choose a sequence $(x_k,x_k^*)_{k\in\mathbb{N}} \subset \mathcal{Y}\times\mathcal{Y}^*$ such that $(x_k,x_k^*) \to (x,x^*)$ and (10.32) is satisfied. We now may compute

$$\begin{aligned} &\frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) \\ &= \lim_{m\to 0}\lim_{k\to\infty} \varepsilon_m^{-1} \int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} f_{\alpha_k}(\dot{u}_k, -\xi_k(t))\,\mathrm{d}t \\ &\ge \liminf_{m\to 0}\liminf_{k\to\infty} \varepsilon_m^{-1} \int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \langle x_k^*, \dot{u}_k\rangle_{X^*,X} + \langle -\xi_k(t), x_k\rangle_{X^*,X} - \rho_{\alpha_k^{-1}}(x_k,x_k^*)\,\mathrm{d}t \end{aligned}$$

From the weak star convergence (10.30) we get

$$\int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \langle x_k^*, \dot{u}_k\rangle_{X^*,X} \to \int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \langle x^*, \mathrm{d}\lambda(t)\rangle_{X^*,X}\,. \tag{10.34}$$

Since $x_k \to x$ we also deduce

$$\int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \langle \xi_k(t), x_k\rangle_{X^*,X}\,\mathrm{d}t \to \int\limits_{Q(\varepsilon_m)} \left\langle \widehat{\xi}(t), x\right\rangle_{X^*,X}\,\mathrm{d}t$$

and thus obtain by using the definition of $\widehat{\xi}$ that

$$\frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) \geq \langle x^*, \dot{u}_{ac}(t_0)\rangle_{X^*,X} + \int\limits_{\mathcal{Y}^*\times\mathbb{R}} \langle \zeta, x\rangle_{X^*,X} \, \mathrm{d}\sigma_{t_0}(\zeta,p) - \rho_{\alpha^{-1}}(x,x^*)$$

for any $(x,x^*) \in \mathcal{Y}\times\mathcal{Y}^*$. Therefore

$$\frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) \geq \int\limits_{\mathcal{Y}^*\times\mathbb{R}} f_\alpha\left(\dot{u}_{ac}(t_0), -\zeta\right) \, \mathrm{d}\sigma_{t_0}(\zeta,p)\,. \tag{10.35}$$

Furthermore choose $t_0 \in [0,T]$ such that it satisfies

$$\begin{aligned}
\frac{\mathrm{d}\mu_s}{\mathrm{d}\|\lambda_s\|}(t_0) &= \lim_{\varepsilon\to 0}\frac{\mu\left([Q(t_0,\varepsilon)\cap[0,T])\right)}{\|\lambda_s\|\left(Q(t_0,\varepsilon)\cap[0,T]\right)} < \infty\,,\\
u_s(t_0) &= \lim_{\varepsilon\to 0}\frac{\lambda\left(Q(t_0,\varepsilon)\cap[0,T]\right)}{\|\lambda_s\|\left(Q(t_0,\varepsilon)\cap[0,T]\right)} \quad\text{and}\\
0 &= \lim_{\varepsilon\to 0}\frac{\mathcal{L}\left(Q(t_0,\varepsilon)\cap[0,T]\right)}{\|\lambda_s\|\left(Q(t_0,\varepsilon)\cap[0,T]\right)}\,.
\end{aligned}$$

The set of all t_0 failing any of the above assumptions is a $\|\lambda_s\|$-null set. As before we may choose a vanishing sequence $(\varepsilon_m)_{m\in\mathbb{N}}$ such that (10.33) is satisfied. Similarly choose $(x,x^*) \in \mathcal{Y}\times\mathcal{Y}^*$. Then there exists a sequence $(x_k,x_k^*)_{k\in\mathbb{N}} \subset \mathcal{Y}\times\mathcal{Y}^*$ such that $(x_n,x_n^*) \to (x,x^*)$ and (10.32) is satisfied. We also can assume that $\rho_{\alpha^{-1}}(x,x^*) < \infty$. We now compute that

$$\begin{aligned}
&\frac{\mathrm{d}\mu_s}{\mathrm{d}\|\lambda_s\|}(t_0)\\
&= \lim_{m\to\infty}\lim_{k\to\infty}\left(\|\lambda_s\|\left(Q(t_0,\varepsilon_m)\cap[0,T]\right)\right)^{-1} \int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} f_{\alpha_k}(\dot{u}_k, -\xi_k(t))\\
&\geq \liminf_{m\to\infty}\liminf_{k\to\infty}\frac{1}{\|\lambda_s\|\left(Q(t_0,\varepsilon_m)\cap[0,T]\right)}\,\cdot\\
&\quad\cdot\left(\int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \langle x_k^*, \dot{u}_k\rangle_{X^*,X} + \langle \xi_k(t), x_k\rangle_{X^*,X}\,\mathrm{d}t - \mathcal{L}(Q(t_0,\varepsilon_m)\cap[0,T])\rho^*_{\alpha_k}(x_k,x_k^*)\right)
\end{aligned}$$

Once again we can see that

$$\begin{aligned}
&\int\limits_{Q(\varepsilon_m)} \langle x_k^*, \dot{u}_k\rangle_{X^*,X} \to \int\limits_{Q(\varepsilon_m)} \langle x^*, \,\mathrm{d}\lambda(t)\rangle_{X^*,X} \text{ and}\\
&\int\limits_{Q(\varepsilon_m)} \langle \xi_k(t), x_k\rangle_{X^*,X}\,\mathrm{d}t \to \int\limits_{Q(\varepsilon_m)} \left\langle \widehat{\xi}(t), x\right\rangle_{X^*,X}\mathrm{d}t\,.
\end{aligned}$$

By assumption we then know that

$$\lim_{m\to\infty} \frac{1}{\|\lambda_s\|\left(Q(t_0,\varepsilon_m)\cap[0,T]\right)} \int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \langle x^*,\, \mathrm{d}\lambda(t)\rangle_{X^*,X} = \langle x^*, \dot u_s\rangle_{X^*,X}\ ,$$

$$\lim_{m\to\infty} \frac{1}{\|\lambda_s\|\left(Q(t_0,\varepsilon_m)\cap[0,T]\right)} \int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \left\langle -\widehat{\xi}(t), x\right\rangle_{X^*,X} \mathrm{d}t = 0 \quad \text{and}$$

$$\lim_{m\to\infty} \frac{\mathcal{L}(Q(t_0,\varepsilon_m)\cap[0,T])}{\|\lambda_s\|\left(Q(t_0,\varepsilon_m)\cap[0,T]\right)} \rho^*_{\alpha^{-1}}(x,x^*) = 0\,.$$

Hence we obtain

$$\frac{\mathrm{d}\mu_s}{\mathrm{d}\|\lambda_s\|}(t_0) \geq \langle x^*, \dot u_s(t_0)\rangle_{X^*,X}$$

for all (x,x^*) such that $\rho_A^{-1}(x,x^*)<\infty$. Due to [37, Proposition 4.77] we then obtain

$$\frac{\mathrm{d}\mu_s}{\mathrm{d}\|\lambda_s\|}(t_0) \geq f_\alpha^\infty(\dot u_s(t_0),0)\,.$$

We hence may deduce that

$$\begin{aligned} &\liminf_{k\to\infty} \int_0^t f_{\alpha_k}\left(\dot u_k(s), -\xi_k(s)\right)\,\mathrm{d}s \\ &\geq \int_0^t \int\limits_{\mathcal{Y}^*\times\mathbb{R}} f_\alpha\left(\dot u_{ac}(s),-\zeta\right)\,\mathrm{d}\sigma_s(\zeta,p)\,\mathrm{d}s + \int_0^t f_\alpha^\infty\left(\dot u_s(s),0\right)\,\mathrm{d}\|\lambda_s\| \end{aligned} \tag{10.36}$$

We thus have now proven that for almost all $t\in[0,T]$

$$\begin{aligned} \mathcal{E}_t(u(t)) + \int_0^t \int\limits_{\mathcal{Y}^*\times\mathbb{R}} f_\alpha\left(\dot u_{ac}(s),-\zeta\right)\,\mathrm{d}\sigma_s(\zeta,p)\,\mathrm{d}s + \int_0^t f_\alpha^\infty\left(\dot u_s(s),0\right)\,\mathrm{d}\|\lambda_s\| \\ \leq \mathcal{E}_0(u(0)) + \int_0^t \int\limits_{\mathcal{Y}^*\times\mathbb{R}} p\,\mathrm{d}\sigma_s(\zeta,p)\,\mathrm{d}s\,. \end{aligned} \tag{10.37}$$

Step 4 - Selection argument: We are now going to prove that for almost all $t\in[0,T]$ it holds

$$\operatorname{supp}(\sigma_t) \subset \{(\zeta,p)\in\mathcal{Y}^*\times\mathbb{R} : \zeta\in F_t(u(t)) \wedge -\zeta\in\alpha(u(t))\,,\ p\leq\partial_t\mathcal{E}_t(u(t))\}\,.$$

Especially we have to prove that $-\zeta\in\alpha(u(t))$ for almost every $t\in[0,T]$. To this end notice that for all $0\leq s<t\leq T$

$$E(t)-E(s)+\mu([s,t]) = \int\limits_0^s p(s)\,\mathrm{d}s\,. \tag{10.38}$$

Notice that due to Lemma 10.10 $t\mapsto\mathcal{E}(u(t))$ is of bounded variation. For the sake of notational simplicity set $e: t\mapsto\mathcal{E}_t(u(t))$. Let $\mathcal{T}\subset[0,T]$ be the set of all t_0 such that t_0 is a Lebesgue point of p,

$$
\begin{aligned}
\frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) &= \lim_{\varepsilon\to 0}\frac{\mu\left([t_0-\varepsilon,t_0+\varepsilon]\cap[0,T]\right)}{\varepsilon}<\infty\,,\\
\xi(t_0) &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int\limits_{t_0-\varepsilon/2}^{t_0+\varepsilon/2}\xi(t)\,\mathrm{d}t\,,\\
\dot u_{ac}(t_0) &= \lim_{\varepsilon\to 0}\frac{\lambda\left([t_0-\varepsilon,t_0+\varepsilon]\cap[0,T]\right)}{\varepsilon} \qquad \text{and}\\
\dot e_{ac}(t_0) &= \lim_{\varepsilon\to 0}\frac{e\left(t+\frac{\varepsilon}{2}\right)-e\left(t-\frac{\varepsilon}{2}\right)}{\varepsilon}\,.
\end{aligned}
$$

Then $\mathcal{T}$ has full measure. Choose a sequence (ε_n), $\varepsilon_n\downarrow 0$ such that (10.33) holds and $\mathcal{E}_{t_0\pm\varepsilon_n/2}(u(t_0\pm\varepsilon_n/2))=E(t_0\pm\varepsilon_n/2)$. Then from (10.38), (10.35) and Lemma 10.10 we obtain that for almost any $t_0\in\mathcal{T}$ it holds

$$
\begin{aligned}
&\int\limits_{\mathcal{Y}^*\times\mathbb{R}} f_\alpha(u_{ac}(t_0),-\zeta)\,\mathrm{d}\sigma_{t_0}(\zeta,p)\\
&\quad\le\; -\dot e_{ac}(t_0)+\int\limits_{\mathcal{Y}^*\times\mathbb{R}} p\,\mathrm{d}\sigma_{t_0}(\zeta,p)\le\int\limits_{\mathcal{Y}^*\times\mathbb{R}}\langle-\zeta,\dot u_{ac}(t_0))\rangle_{X^*,X}\,\mathrm{d}\sigma_{t_0}(\zeta,p)
\end{aligned}
$$

and hence deduce that for almost any $t_0\in\mathcal{T}$ (and hence for almost any $t_0\in(0,T)$) it holds

$$
\int\limits_{\mathcal{Y}^*\times\mathbb{R}}\langle-\zeta,\dot u_{ac}t_0)\rangle_{X^*,X}-f_\alpha(u_{ac}(t_0),-\zeta)\,\mathrm{d}\sigma_{t_0}(\zeta,p)\ge 0
$$

This enables us to deduce that for almost all $t\in(0,T)$

$$
\operatorname{supp}(\sigma_t)\subset\left\{(\zeta,p)\in\mathcal{Y}^*\times\mathbb{R}:\zeta\in F_t(u(t))\,,\ -\zeta\in\alpha(\dot u_{ac}(t))\text{ and }p\le P_t(u(t),\zeta)\right\}.
$$

Thus we meet the conditions of Lemma 10.11 and deduce that there exist measurable functions $\xi:(0,T)\to\mathcal{Y}$ and $p:(0,T)\to\mathbb{R}$ such that for almost all $t\in(0,T)$

$$
\xi(t)\in F_t(u(t))\,,-\xi(t)\in\alpha(\dot u_{ac}(t))\text{ and }p(t)\le P_t(u(t),\xi(t))\,.
$$

such that (10.22) holds. From there we deduce that for almost all $t\in[0,T]$

$$
f_\alpha(\dot u_{ac}(t),-\xi(t))-p(t)\le\int\limits_{\mathcal{Y}^*\times\mathbb{R}} f_\alpha(\dot u_{ac}(t),-\zeta)-p\,\mathrm{d}\sigma_t(\zeta,p)\,.
$$

Applying this result to (10.37) we then deduce that

$$
\begin{aligned}
&\mathcal{E}_t(u(t))+\int_0^t f_\alpha\left(\dot u_{ac}(s),-\xi(s)\right)\mathrm{d}s+\int_0^t f_\alpha^\infty\left(\dot u_s(s),0\right)\,\mathrm{d}\|\lambda_s\|\\
&\qquad\le\mathcal{E}_0(u(0))+\int_0^t p(s)\,\mathrm{d}s\le\mathcal{E}_0(u(0))+\int\limits_0^t P_s\left(u(s),\xi(s)\right)\,\mathrm{d}s\,.
\end{aligned}
$$

This completes the proof. □

Let us finish this chapter by making a few remarks. First, by a simple subsequence argument for any limit point u of the family $(u_n)_{n\in\mathbb{N}}$ there exists a function $\xi \in L^q(0,T;\mathcal{Y}^*)$ such that (10.9) is satisfied. Moreover if u is an absolutely continuous limit point of the family $(u_n)_{n\in\mathbb{N}}$ then there is a corresponding function $\xi \in L^q(0,T;\mathcal{Y})$ such that (u,ξ) solves (10.1) in the sense of Definition 10.3.

11. Approximation of dissipation potentials on non-reflexive Banach spaces

In the case that $\mathcal{Y}$ is a possibly non-reflexive Banach space we are not able to provide results in the full generality as above. The Fitzpatrick formalism is no longer valid. Therefore we need to confine ourselves to the case of cyclic operators. That is we are going to study doubly nonlinear differential inclusions of the type

$$\partial\Psi\left(\dot{u}(t)\right) + \partial_u \mathcal{E}_t\left(u(t)\right) \ni 0\,. \tag{11.1}$$

These types of equations are of considerable interest in applications as one will often find dissipation functionals of the type $\|\cdot\|_{L^1(\Omega)}$ for some open bounded set Ω. The prototypical application we have in mind are approximations of the form $\frac{1}{1+\varepsilon}\|\cdot\|^{1+\varepsilon}_{L^{1+\varepsilon}(\Omega)}$ as $\varepsilon \to 0$. Whereas these functionals can all be defined on the space $L^1(\Omega)$ the natural spaces for their conjugates to be defined on are $L^{(1+\varepsilon)/\varepsilon}(\Omega)$. Another example would be considering the limit in $\|\cdot\|_{L^1(\Omega)} + \varepsilon\|\cdot\|^2_{L^2(\Omega)}$. For these functions the natural space the dual can be defined on would be the Hilbert space $L^2(\Omega)$. We will use these two examples as a means to test our assumptions for applicability.

In order to be able to deal with these types of problems, especially the fact that different 'state spaces' are required, we need to provide a more involved setting. This will be done in the first section. However still a number of problems arise when we allow for non-reflexive Banach spaces. Most notoriously among them is the loss of the Radon-Nikodym property. We shall present a way how this problem can be overcome. To do so a relaxation of the state space is needed and the concept of metric derivatives is introduced. Also useful properties of the Mosco convergence are lost. We will address these obstacles in the third section. Based on our analysis in these two sections we can then state our main result. Its proof will be part of the last section of this chapter.

11.1. Basic setting of the problem

Our aim is to allow for solutions of the approximating equations to be in different spaces. Therefore we start by defining what families of 'state spaces' are deemed to be admissible for our analysis.

Definition 11.1 (Admissible families of state spaces)**.** Let H be a separable Hilbert space and $\mathcal{Y}$ a separable Banach space. Moreover let $(\mathcal{Y}_n)_{n\in\mathbb{N}}$ be a family of Banach spaces. The triple $\big(\mathcal{Y}, H, (\mathcal{Y}_n)_{n\in\mathbb{N}}\big)$ is admissible if and only if

(i) For all $n \in \mathbb{N}$ it holds

$$H \overset{\mathrm{i}_n}{\hookrightarrow} \mathcal{Y}_n \overset{\mathrm{j}_n}{\hookrightarrow} \mathcal{Y} \tag{11.2}$$

where each embedding is continuous.

(ii) What is more, there exists a $C > 0$ such that for all $n \in \mathbb{N}$ it holds that

$$|\mathfrak{j}_n(y)|_{\mathcal{Y}} \leq C\,|y|_{\mathcal{Y}_n} \tag{11.3}$$

(iii) Moreover for all $m, n \in \mathbb{N}$ it holds

$$\mathfrak{j}_n \circ \mathfrak{i}_n = \mathfrak{j}_m \circ \mathfrak{i}_m\,. \tag{11.4}$$

We denote the thus uniquely defined embedding $H \hookrightarrow \mathcal{Y}$ by $\mathfrak{i}$ and assume that $\mathfrak{i}(H)$ is strongly dense in $\mathcal{Y}$.

Throughout this chapter we will identify an element $y \in \mathcal{Y}_n$ with the corresponding element $\mathfrak{j}_n(y) \in \mathcal{Y}$ if no confusion can arise.

We will now present our assumptions on the energy functional. These assumptions are similar to (10.1) but need to be refined in some points in order to account for the more general setting introduced above.

Assumption 11.2 (Assumptions on the energy)**.** Let $\mathcal{E} : [0,T] \times H \to [-\infty, \infty]$ be and energy functional, $F : D(F) \to H$ be its generalized state derivative and $P : D(P) \to \mathbb{R}$ be its generalized time derivative. For a function $u : [0,T] \to H$ define $\mathcal{G}(u) := \sup\{\mathcal{E}_t(u) : t \in [0,T]\}$. The triple $(\mathcal{E}, F, P)$ has the following properties:

Non degeneracy and lower semi continuity: The domain of $\mathcal{E}$ is of the form $D(\mathcal{E}) = [0,T] \times D$ and $F : [0,T] \times D \rightrightarrows H$ for some Borel set D. Furthermore we assume

$$\begin{gathered} u \mapsto \mathcal{E}_t(u) \text{ is l.s.c. for all } t \in [0,T], \\ \exists C_0 > 0 : \ \forall (t,u) \in [0,T] \times D : \mathcal{E}_t(u) \geq -C_0 \text{ and} \\ \operatorname{graph}(F) \text{ is a Borel set of } [0,T] \times \mathcal{Y} \times \mathcal{Y}^*\,. \end{gathered} \tag{11.$\mathcal{E}_0$}$$

Coercivity: For all $t \in [0,T]$

$$u \mapsto \mathcal{E}_t(u) \text{ has compact sub-levels} \tag{11.$\mathcal{E}_1$}$$

Lipschitz continuity: There exists $C_1 > 0$ such that for all $u \in D$ and $s, t \in [0,T]$ it holds

$$|\mathcal{E}_t(u) - \mathcal{E}_s(u)| \leq C_1 \mathcal{E}_t(u) |t-s|\,. \tag{11.$\mathcal{E}_2$}$$

Conditioned one-sided time-differentiability: There exists a function $P : \operatorname{graph}(F) \to \mathbb{R}$ and a constant $C_2 > 0$ such that

$$\forall (t,u,\xi) \in \operatorname{graph}(F) : \ \liminf_{h \downarrow 0} \frac{\mathcal{E}_{t+h}(u) - \mathcal{E}_t(u)}{h} \leq P_t(u,\xi) \leq C_2 \mathcal{G}(u) \tag{11.$\mathcal{E}_3$}$$

Chain rule inequality: Let Z be either $\mathcal{Y}_n$ for some $n \in \mathbb{N}$ or $Z = \mathcal{Y}$. Let $u \in W^{1,1}(0,T;Z)$ and $\xi \in L^1(0,T;Z^*)$ such that

$$\sup_{t \in [0,T]} \mathcal{E}_t(u(t)) < \infty \ \text{ and } \ \xi(t) \in F_t(u(t)) \text{ for a.a. } t \in [0,T]\,. \tag{11.5}$$

Then the map $e \mapsto \mathcal{E}_t(u(t))$ is absolutely continuous and for almost all $t \in [0,T]$ it then holds

$$\dot{e}(t) \geq \langle \xi(t), \dot{u}_{ac}(t) \rangle_{\mathcal{Y}^*,\mathcal{Y}} + P_t\left(u(t), \xi(t)\right) \tag{11.$\mathcal{E}_5$}$$

Weak closedness: For all $t \in [0,T]$ and for all sequences $(u_n)_{n\in\mathbb{N}} \subset H$, $\xi_n \in F_t(u_n)$, $\mathbb{E}_n = \mathcal{E}_t(u_n)$ and $p_n = P_t(u_n, \xi_n)$ with

$$u_n \to u \text{ in } \mathcal{Y}, \xi_n \rightharpoonup \xi \text{ in } H, p_n \to p \text{ and } \mathbb{E}_n \to \mathbb{E} \text{ in } \mathbb{R} \tag{11.6}$$

it holds

$$(t,u) \in D(F),\ \xi \in F_t(u),\ p \leq P_t(u,\xi) \text{ and } \mathbb{E} = \mathcal{E}_t(u)\,. \tag{11.$\mathcal{E}_6$}$$

Similar to the prior chapter we also need to state what we precisely understand to be a solution of a doubly nonlinear differential inclusion (11.1). Once again let Z be either $\mathcal{Y}_n$ for some $n \in \mathbb{N}$ or $Z = \mathcal{Y}$.

Definition 11.3. Let $(\mathcal{E}, F, T)$ satisfy Hypothesis 10.1 and assume that $\Psi : \mathcal{Y} \to [0,+\infty]$ be a convex, lower semi-continuous function. Choose $u_0 \in D$. A function tuple $(u,\xi) \in W^{1,1}(0,T;Z) \times L^1(0,T;Z^*)$ is called a solution to the Cauchy problem

$$\partial\Psi\left(\dot{u}(t)\right) + F_t(u(t)) \ni 0,\ u(0) = u_0 \tag{11.7}$$

if and only if it satisfies $u(0) = u_0$, $\xi(t) \in F_t(u(t))$ and the energy identity

$$\mathcal{E}_t(u(t)) + \int_0^t \Psi\left(\dot{u}(s)\right) + \Psi^*\left(-\xi(s)\right)\,\mathrm{d}s = \mathcal{E}_0(u(0)) + \int_0^t P_s(u(s), \xi(s))\,\mathrm{d}s \tag{11.8}$$

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

Notice that if the space Z is reflexive any solution to (11.1) in the sense of the above definition also solves (11.1) in the sense of Definition 10.3. We are also able to state an analogue to Proposition 10.4 also in this setting it holds that

Proposition 11.4. *If the tuple $(u,\xi) \in W^{1,1}(0,T;Z) \times L^1(0,T;Z^*)$ satisfies the above definition then for almost every $t \in [0,T]$ it holds*

$$\partial\Psi\left(\dot{u}(t)\right) \ni -\xi(t)\,. \tag{11.9}$$

Moreover if (11.$\mathcal{E}_5$) holds with '=' then also the converse assertion is true.

We refrain from giving a proof here as it exactly follows the lines of the prior proof of Proposition 10.4. Instead we will state the general conditions on the family of dissipation potentials $(\Psi_n)_{n\in\mathbb{N}}$, that we are going to impose in the sequel. Once again the choice of our assumptions is driven by the two model examples we have in mind.

Assumption 11.5. Let $(\mathcal{Y}, H, (\mathcal{Y}_n))$ be an admissible triple. We assume that the family of dissipation potentials $(\Psi_n)_{n\in\mathbb{N}}$ satisfies the following conditions.

Non degeneracy. For all $n \in \mathbb{N}$ it holds

$$\Psi_n : \mathcal{Y}_n \to [0,\infty) \quad \text{is convex and lower semi-continuous.} \tag{11.10}$$

Uniform growth. There exist $c_1, c_2, c_3 > 0$, $p \geq 1$ and $q > 1$ such that for all $n \in \mathbb{N}$ it holds

$$\forall (x, x^*) \in \mathcal{Y}_n \times \mathcal{Y}_n^* : \quad \Psi_n(x) + \Psi_n^*(x^*) \geq c_1 \, |\mathrm{j}_n(x)|_{\mathcal{Y}}^p + c_2 \, |\mathrm{i}_n^* (x^*)|_{H^*}^q - c_3 \tag{11.11}$$

Moreover we assume that $\Psi_n(0) = 0$.

Remark 11.6. Assume that we consider the family of potentials $\frac{1}{p_n} \| \cdot \|_{L^{p_n}}^{p_n}$ for some family $(p_n) \in [1,2]$ with $p_n \downarrow 1$. Then we can choose $\mathcal{Y} = L^2$, $\mathcal{Y}_n = L^{p_n}$ and $H = L^2$ and it is easy to see that the potentials satisfy the conditions of Assumption 11.5. If one chooses $\Psi_n(\cdot) = \| \cdot \|_{L^1} + \frac{1}{n} \| \cdot \|_{L^2}^2$ then a natural choice would be $\mathcal{Y} = L^1$ and $H = \mathcal{Y}_n = L^2$ for all $n \in \mathbb{N}$.

Assume that $(\Psi_n)_{n \in \mathbb{N}}$ was a family of convex, lower semi continuous functions on $\mathcal{Y}$. If $\mathcal{Y}$ were reflexive then

$$\partial \Psi_n \xrightarrow{g} \partial \Psi \quad \Longleftrightarrow \quad \Psi_n \xrightarrow{M} \Psi \quad \Longleftrightarrow \quad \Psi_n^* \xrightarrow{M} \Psi^* \, . \tag{11.12}$$

A major part of the proof of Theorem 10.8 was to show that equation (10.36), a $\liminf$ inequality for the Fitzpatrick function, holds. A similar result will be needed in our context. One can easily replace the Fitzpatrick of $f_{\partial \Psi_n}$ by the sum $\Psi_n + \Psi_n^*$. If $\mathcal{Y}$ was reflexive we would now aim to prove that

$$\liminf_{n \to \infty} \int_0^T \Psi_n \left(\dot{u}_n(t) \right) \, \mathrm{d}t \quad \geq \quad \int_0^T \Psi \left(\dot{u}(t) \right) \, \mathrm{d}t \quad \text{and} \tag{11.13}$$

$$\liminf_{n \to \infty} \int_0^T \Psi_n^* \left(-\xi_n(t) \right) \, \mathrm{d}t \quad \geq \quad \int_0^T \Psi^* \left(-\xi(t) \right) \, \mathrm{d}t \, . \tag{11.14}$$

for any limit point (u, ξ) of the family of solutions $(u_n . \xi_n)_{n \in \mathbb{N}}$.Especially showing (11.13) gives rise to some concern. The first obstacle is that even if u is absolutely continuous there need not to exist a weak derivative of u in $L^1(0, T, \mathcal{Y})$. The second obstacle is the fact that the implication

$$\Psi_n \xrightarrow{M} \Psi \quad \Longrightarrow \Psi_n^* \xrightarrow{M} \Psi^* \tag{11.15}$$

is in general no longer valid. In fact it has been shown by G. Beer and J. M. Borwein in [11] that if the above implication holds for all families of convex and lower semi continuous functions, then the underlying Banach space is reflexive. In order to overcome these problems a number of preparations are needed. We start by stating our precise assumptions on the energy functional. We will address these problems in the following two sections.

11.2. Embedding of the state space and metric derivatives

In order to overcome the lack of the Radon-Nikodym property one can embed the state space into a larger space which has a weak version of the Radon-Nikodym property. The results in this section are are due to be published in an forthcoming article by A. Mielke, R. Rossi and G. Savaré [74]. However as at the time this thesis was written the result were not yet published available we shall present proofs here **for which we claim no originality whatsoever**.

11.2.1. An embedding result of reflexive Banach spaces

Proposition 11.7. *Let $\mathcal{Y}$ be a separable Banach space. Then there exists a separable space $X \subset \mathcal{Y}^*$ and an isometric embedding $\mathrm{j} : \mathcal{Y} \hookrightarrow X^*$ such that $\mathrm{j}\,(\mathcal{Y})$ is weakly - star dense in X^*.*

Proof. As $\mathcal{Y}$ is separable there exists a countable subset $\{y_n' : n \in \mathbb{N}\} \subset \mathcal{Y}^*$ such that

$$|y| = \sup \left\{ \left\langle y_n', y \right\rangle_{\mathcal{Y}^*,\mathcal{Y}} : n \in \mathbb{N} \right\} . \tag{11.16}$$

We then define $X := \overline{\mathrm{span}\,\{y_n' : n \in \mathbb{N}\}}$. Clearly X is a closed linear subspace of $\mathcal{Y}^*$, moreover X is separable. Define the embedding $\mathrm{j} : \mathcal{Y} \hookrightarrow X^*$ by

$$\langle \mathrm{j}(y), x \rangle_{X^*,X} = \langle x, y \rangle_{\mathcal{Y}^*,\mathcal{Y}} \quad \forall x \in X \, .$$

We will now show that $\mathrm{j}(\mathcal{Y})$ is weakly-star dense in X^*. Let $\widehat{X}$ be the weak-star closure of $\mathrm{j}(\mathcal{Y})$. Notice that the set of all continuous linear functionals on X^* with respect to its weak-star topology is exactly X. Now assume that $x' \notin \widehat{X}$ then there exists some $x \in X$ and $\alpha \in \mathbb{R}$ such that

$$\left\langle x', x \right\rangle_{X^*,X} > \alpha \geq \langle \widehat{x}, x \rangle_{X^*,X} \quad \widehat{x} \in \widehat{X} \, .$$

As $\widehat{X}$ is a linear space we obtain that

$$\langle \widehat{x}, x \rangle_{X^*,X} \leq 0 \quad \forall \widehat{x} \in \widehat{X}$$

Hence for all $y \in \mathcal{Y}$ it holds

$$\langle x, y \rangle_{\mathcal{Y}^*,\mathcal{Y}} = 0 \forall$$

and thus $x = 0$ which is a contradiction. Moreover the embedding $j : \mathcal{Y} \hookrightarrow X^*$ is isometric. This is due to the choice of the subspace X, especially (11.16). □

Why does this result help? As X^* is the dual of a separable space, it possesses a weak-star Radon Nikodym theorem, see Section 4.2. Hence if $y : [0,T] \to \mathcal{Y}$ is absolutely continuous then also the map $\widehat{y} : t \mapsto \mathrm{j}\,(y(t))$ is absolutely continuous. Hence it has a weak star derivative $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{y} \in L^1(0,T;X')$. To see where this leads to we need to introduce the so called metric derivatives.

11.2.2. Metric derivatives

Metric derivatives can be used to extend certain differential equations to metric spaces. We refer to the recent monograph by L. Ambrosio, N. Gigli and G. Savaré [3] for a thorough introduction into gradient flows in metric spaces. This book also provides a solid introduction into the necessary concepts for such a study. We also point out two articles by A. Mielke, R. Rossi and G. Savaré [93, 72] who studied doubly nonlinear differential inclusions and especially rate independent evolutions on metric spaces. These are topics of interest in their own right. Here however the concept of metric derivatives serves 'merely' as a technical tool allowing us to overcome the problems at hand.

For our study we require the possibility of non-symmetric distance functions. To this end let us shortly define our assumptions. Let $\mathfrak{X}$ be a set and $d : \mathfrak{X} \times \mathfrak{X} \to [0,\infty]$. We say that $(\mathfrak{X}, d)$ is a (possibly asymmetric) metric space if

$$\forall u, v \in \mathfrak{X} : \qquad d(u,v) = 0 \Longleftrightarrow u = v \tag{11.17}$$

$$\forall u, v, w \in \mathfrak{X} : \qquad d(u,w) \leq d(u,v) + d(u,w)\,. \tag{11.18}$$

Definition 11.8 (Metric derivative)**.** Let $(\mathfrak{X}, d)$ be a metric space and $v : [0,T] \to \mathfrak{X}$. The metric derivative of v at a point $t \in [0,T]$ is given by

$$|v'|_d(t) = \lim_{h\to 0} \frac{d\left(v(t), v(t+h)\right)}{h} \tag{11.19}$$

should this limit exist.

A function $v : [0,T] \to \mathfrak{X}$ is called absolutely continuous if there exists a function $m \in L^1(0,T)$ such that

$$\forall 0 \leq s < t \leq T : d(v(s), v(t)) \leq \int_s^t m(\tau)\,\mathrm{d}\tau\,. \tag{11.20}$$

Theorem 11.9 ([3, Theorem 1.1.2], [93, Proposition 2.2])**.** *Let $v \in AC(0,T;\mathfrak{X})$. Then for almost every $t \in [0,T]$ the metric derivative $|v'|_d(t)$ exists. The function $t \mapsto |v'|_d(t)$ is integrable and*

$$\forall 0 \leq s < t \leq T : d(v(s), v(t)) \leq \int_s^t |v'|_d(\tau)\,\mathrm{d}\tau\,.$$

Moreover for any function $m \in L^1(0,T)$ satisfying (11.20) *it holds*

$$|v'|_d(t) \leq m(t) \quad \textit{for a.e. } t \in [0,T]\,.$$

To see how this relates to our problem let d be the metric induced by the norm on $\mathcal{Y}$.

Proposition 11.10. *Let $\mathcal{Y}$ be a reflexive space and $X \subset \mathcal{Y}^*$ be the separable space constructed in Proposition 11.7. Let $y : [0,T] \to \mathcal{Y}$ be absolutely continuous. Then the map $t \mapsto \mathfrak{j}\left(y(t)\right)$ is in $AC(0,T;X^*)$. It thus has an weak-star derivative denoted by $\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{j}\left(y(t)\right)$ and it holds*

$$\left\|\frac{\mathrm{d}}{\mathrm{d}t}\widehat{y}(t)\right\|_{X'} = \left|y'\right|_d(t) \quad \textit{for a.e. } t \in [0,T]\,.$$

Proof. Indeed let $y : [0,T] \to \mathcal{Y}$ be absolutely continuous. Then as the embedding $\mathfrak{j} : \mathcal{Y} \to X^*$ is an isometry also $t \mapsto \mathfrak{j}\left(y(t)\right)$ is absolutely continuous. For all $0 \leq s < t \leq T$ it holds

$$\begin{aligned} |y(t) - y(s)|_{\mathcal{Y}} &= \|\mathfrak{j}\left(y(t)\right) - \mathfrak{j}\left(y(s)\right)\|_{X^*} \\ &= \sup\left\{\langle \mathfrak{j}\left(y(t)\right) - \mathfrak{j}\left(y(s)\right), x_n\rangle_{X^*,X} : n \in \mathbb{N}\right\}\,, \end{aligned}$$

where the family $(x_n)_{n\in\mathbb{N}}$ can be taken from the proof of Proposition 11.7. Notice that for every n the set of all $t \in [0,T]$ such that

$$\lim_{h\to 0} \frac{1}{h} \int_t^{t+h} \left\langle \frac{\mathrm{d}}{\mathrm{d}s}\widehat{y}(s), x_n \right\rangle_{X^*,X} \mathrm{d}s = \left\langle \frac{\mathrm{d}}{\mathrm{d}s}\widehat{y}(t), x_n \right\rangle_{X^*,X}$$

has full measure. Hence for almost all $t \in [0,T]$ it holds that

$$\liminf_{h\to 0} \frac{1}{h} |y(t+h) - y(t)|_{\mathcal{Y}} \geq \left\| \frac{\mathrm{d}}{\mathrm{d}t}\widehat{y}(t) \right\|_{X'}$$

At the same time it holds for any $0 \leq r < s \leq T$ that

$$\int_r^s \left\| \frac{\mathrm{d}}{\mathrm{d}t}\widehat{y}(t) \right\|_{X'} \mathrm{d}t \geq \|\mathrm{j}(y(s)) - \mathrm{j}(y(r))\|_{X'} = |y(s) - y(r)|_{\mathcal{Y}}$$

This shows the claim. □

Example 11.11. Assume that $\mathcal{Y} = L^1(\Omega)$ for some open, bounded set Ω. Then $\mathcal{Y}^* = L^\infty(\Omega)$. A natural choice for X would be the set of all bounded continuous functions on Ω, $C_b(\Omega)$. The corresponding dual is the set of all finite Radon measure $\mathcal{M}(\Omega)$.

11.3. Mosco convergence in non-reflexive Banach spaces

Let us interpret u as a function on X^* and denote by $\frac{\mathrm{d}}{\mathrm{d}t}u$ its weak-star derivative. If we try to prove an analogue of (11.13) we immediately run into the problem that Ψ is only defined on $\mathcal{Y}$. Hence one needs to find a way to extend this function to X^*. As we are trying to establish a $\liminf$ result it appears natural to relax Ψ to its lower semi continuous closure $\Psi_{\text{rel}} : X^* \to (-\infty, \infty]$ which is given by

$$\Psi_{\text{rel}}(x') = \sup\left\{ \langle x', x\rangle_{X^*,X} - \Psi^*(x) : x \in X \subset \mathcal{Y}^* \right\} . \tag{11.21}$$

We would then replace (11.13) by

$$\liminf_{n\to\infty} \int_0^T \Psi_n\left(\dot{u}_n(t)\right)\, \mathrm{d}t \geq \int_0^T \Psi_{\text{rel}}\left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{j}u(t)\right)\right) \mathrm{d}t\,. \tag{11.22}$$

To prove such an inequality an analogue of the $liminf$-inequality (2.24) with respect to Ψ_{rel} and the weak star topology of X^* is needed. To be precise we would desire that

$$\forall (y_n) \subset \mathcal{Y},\ \mathrm{j}(x_n) \overset{*}{\rightharpoonup} x' \text{ in } X^* : \quad \liminf_{n\to\infty} \Psi_n(y_n) \geq \Psi_{\text{rel}}(x')\,. \tag{11.23}$$

However even if $\Psi_n \xrightarrow{M} \Psi$ the above inequality does not need to hold, as the following example will show. The idea behind this example is due to G. Beer and J. M. Borwein [11].

Example 11.12. Assume that $\mathcal{Y} = L^1(-1,1)$. Then $\mathcal{Y}^* = L^\infty(-1,1)$ and one may choose $X = C_b(-1,1)$ and consequently $X^* = \mathcal{M}(-1,1)$. Weierstrass theorem grants that X is separable and it is simple to check that $L^1 \hookrightarrow \mathcal{M}(-1,1)$ is weakly star dense. Choose $g \in C_b(-1,1)$ to be $g(x) = 1 - |x|$. Then for all $f \in L^1(-1,1)$ with $\|f\|_{L^1(-1,1)} = 1$ it holds

$$\langle g, f\rangle_{\mathcal{Y}^*,\mathcal{Y}} = \int_{-1}^1 f(x)g(x)\,\mathrm{d}x < 1\,. \tag{11.24}$$

The reason therefore is that an L^1 function cannot concentrate its mass on one point. Now define $(A_n)_{n\in\mathbb{N}}$ by

$$A_n = \left\{ f \in L^1(\Omega) : \ \|f\|_{L^1(-1,1)} = 1 \, , \ \langle g, f \rangle_{\mathcal{Y}^*,\mathcal{Y}} \geq 1 - \frac{1}{n} \right\} .$$

Notice that due to (11.24) the intersection of all A_n is empty, i.e.

$$\bigcap_{n\in\mathbb{N}} A_n = \emptyset \, .$$

Furthermore we define the sets C_n by

$$C_n = \mathrm{co}\left(\{0\} \cup A_n\right) \, .$$

Notice that A_n is closed and hence so is C_n. For $n \in \mathbb{N}$ we define the functional $\Psi_n : X \to [0, +\infty]$ by $\Psi_n(x) := M_{C_n}(x)$. First notice that $M_{C_n} \xrightarrow{M} M_{\{0\}}$. First $C_n \xrightarrow{M} \{0\}$ in the sense of Mosco convergence of convex sets. First, as $0 \in C_n$ for all $n \in \mathbb{N}$ a strong recovery sequence exits. Second let $(x_n)_{n\in\mathbb{N}} \subset X$ with $x_n \in C_n$ and $x_n \rightharpoonup x$. Then as $(C_n)_{n\in\mathbb{N}}$ is a decreasing sequence of closed convex sets, we obtain that $x \in C_n$ for all $n \in \mathbb{N}$. Due to [Borwein - Lemma 3.1] we obtain that $x = 0$. As Mosco convergence of a family of convex functions is equivalent to Mosco convergence of their epigraphs we deduce that $M_{C_n} \xrightarrow{M} M_{\{0\}}$.
Define the sequence $(f_n)_{n\in\mathbb{N}} \subset L^1(-1,1)$ by $f_n(x) = \frac{n}{2}\chi_{\left(-\frac{1}{n},\frac{1}{n}\right)}$ be a sequence of functions. It is simple to see that $\|f_n\|_{L^1} \equiv 1$ and

$$f_n \cdot \mathcal{L} \overset{*}{\rightharpoonup} \delta_0 \quad \text{in } \mathcal{M}(-1,1) \, ,$$

where δ_0 is the Dirac measure on 0. Moreover for all $n \in \mathbb{N}$ it holds

$$\int_{-1}^{1} f(x)g(x) \, \mathrm{d}x = \int_{-1/n}^{1/n} \frac{n}{2} - \frac{n}{2}|x| \, \mathrm{d}x = 1 - \frac{1}{2n}$$

Then as $f_n \in A_n$ it also holds $f_n \in C_n$. Moreover for all $n \in \mathbb{N}$ and $\lambda > 1$ it holds $\lambda f_n \neq C_n$, as $\|\lambda f_n\|_{L^1} = \lambda > 1$. Thus

$$\Psi_n\left(f_n\right) \equiv 1 \, .$$

However as $\Psi = M_{\{0\}}$ we obtain that also $\Psi_{\text{rel}} = M_{\{0\}}$, and hence $\Psi_{\text{rel}}\left(\delta_0\right) = +\infty$. Therefore the $\liminf$ inequality is not satisfied.

Before we further discuss this topic note that for any $y \in \mathcal{Y}$ it holds that $\Psi_{\text{rel}}(\mathrm{j}(y)) \leq \Psi(y)$. However one can not be certain that also equality holds. This can be overcome by adapting the choice of X.

Proposition 11.13. *Let $\mathcal{Y}$ be a separable Banach space and $\Psi : \mathcal{Y} \to [0, \infty]$ be a proper, convex, lower semi continuous function such that $\Psi(x) = 0$ iff $x = 0$ and*

$$\forall \left(y, y'\right) \in \mathcal{Y} \times \mathcal{Y}^* : \quad \Psi(y) + \Psi^*(y') \geq c_1 |y|_{\mathcal{Y}}^p + c_2 |y'|_{\mathcal{Y}^*}^q - c_3 \, . \tag{11.25}$$

for some $c_1, c_2, c_3 > 0$, $p \geq 1$ and $q > 1$. Then there exists a separable subspace $X \subset \mathcal{Y}^*$ and an isometric embedding $\mathrm{j} : \mathcal{Y} \hookrightarrow X^*$ such that $\mathrm{j}\left(\mathcal{Y}\right)$ is weakly star dense in X^* and

$$\forall y \in \mathcal{Y} : \Psi(y) = \sup \left\{ \left\langle y', y \right\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*(y') : y' \in X \right\} . \tag{11.26}$$

Proof. **Step 1 - Convex conjugate.** First we are going to show that for all $y \in \mathcal{Y}$ there exists some $y' \in \mathcal{Y}^*$ such that

$$\Psi(y) = \langle y', y\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*(y') .$$

To this end notice that as Ψ is proper, convex, and lower semi continuous it holds that

$$\Psi(y) = \sup\left\{\langle y', y\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*(y') : y' \in \mathcal{Y}^*\right\} .$$

Let $y \in \mathcal{Y}^*$ be fixed. Then for $n \in \mathbb{N}$ choose $y_n' \in \mathcal{Y}^*$ such that

$$\Psi(y) \le \langle y_n', y\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*\left(y_n'\right) + \frac{1}{n} .$$

The family $(y_n')_{n\in\mathbb{N}}$ is uniformly bounded in $\mathcal{Y}^*$ because

$$\langle y', y\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*\left(y'\right) \le \left|y'\right| |y| - c_2 \left|y'\right|^q + c_3 \to -\infty \quad \text{as } \left|y'\right| \to \infty .$$

Hence there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ such that $y_{n_k}' \stackrel{*}{\rightharpoonup} y'$ for some $y' \in \mathcal{Y}^*$. As Ψ^* is lower semi continuous we obtain that

$$\Psi(y) \le \limsup_{k\to\infty} \langle y_{n_k}', y\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*\left(y_{n_k}\right) + \frac{1}{n_k} \le \langle y', y\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*\left(y'\right) \le \Psi(y) .$$

Step 2 - Choice of subspace. We will now prove that there exists a separable subset $X \subset \mathcal{Y}^*$ satisfying (11.26). To this end let $(y_n)_{n\in\mathbb{N}}$ be strongly dense in $\mathcal{Y}$. For all $n \in \mathbb{N}$ choose y_n' such that

$$\Psi\left(y_n\right) = \langle y_n', y_n\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*\left(y_n'\right) .$$

Moreover let $(\widehat{y}_n')_{n\in\mathbb{N}}$ be the sequence satisfying (11.16). We then define

$$X = \overline{\operatorname{span}\left(\{y_n' : n \in \mathbb{N}\} \cup \{\widehat{y}_n' : n \in \mathbb{N}\}\right)}$$

Obviously X is a closed, linear and separable subspace of $\mathcal{Y}^*$. Moreover for all $y \in \mathcal{Y}$ it holds that

$$\Psi(y) \ge \sup\left\{\langle x, y\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*(x) : x \in X\right\} .$$

For any $y \in \mathcal{Y}$ there exists a subsequence $(y_{n_k})_{k\in\mathbb{N}}$ such that $y_{n_k} \to y$. Let $\left(y_{n_k}'\right)_k$ be the corresponding sequence in X. Due to the lower semi continuity

$$\Psi(y) \le \liminf_{k\to\infty} \Psi\left(y_{n_k}\right) \le \liminf_{k\to\infty} \langle y_{n_k}', y_{n_k}\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*\left(y_{n_k}'\right) .$$

Notice that the family $\left(y_{n_k}'\right)_k$ is uniformly bounded in $\mathcal{Y}$. This is due to the observation that

$$0 \le \Psi\left(y_{n_k}\right) \le \left|y_{n_k}\right| \left|y_{n_k}'\right| - c_2 \left|y_{n_k}'\right|^q + c_3$$

and the fact that $|y_{n_k}| < C$ for some $C > 0$. Hence there exists a subsequence $\left(y_{n_{k_j}}'\right)_k$ such that

$$y_{n_{k_j}}' \stackrel{*}{\rightharpoonup} y' \quad \wedge \quad \liminf_{k\to\infty} \langle y_{n_k}', y_{n_k}\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*\left(y_{n_k}'\right) = \liminf_{j\to\infty} \left\langle y_{n_{k_j}}', y_{n_{k_j}}\right\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*\left(y_{n_{k_j}}'\right)$$

Moreover we have that $y' \in X$. Hence we can deduce that

$$\Psi(y) \leq \left\langle y', y \right\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*\left(y'\right) \leq \sup\left\{ \langle x, y \rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi^*(x) : x \in X \right\} .$$

Step 3 - Inclusion. We define inclusion $\mathrm{j} : \mathcal{Y} \hookrightarrow X^*$ by

$$\forall x \in X : \langle \mathrm{j}(y), x \rangle_{X^*,X} = \langle x, y \rangle_{\mathcal{Y}^*,\mathcal{Y}} .$$

Showing that this inclusion is weakly star dense in X^* has already been done in the proof of Proposition 11.7. The space defined there is contained in the space constructed here. It also follows that the embedding is isometric. □

We are now going to show that asking for (11.23) is equivalent to requiring that for all elements of X there exists a strong recovery sequence for Ψ^*.

Proposition 11.14. *Let $(\Psi_n)_{n\in\mathbb{N}}$ be a family of convex, uniformly proper and lower semi continuous functions $\Psi_n : \mathcal{Y} \to [0, \infty)$ such that $\Psi_n(0) = 0$. Let $g : X \to [0, \infty)$ be convex, proper and lower semi continuous. Then the following two are equivalent*

i) For all $(y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}$ with $\mathrm{j}\,(y_n) \overset{*}{\rightharpoonup} x'$ in X^* it holds

$$\liminf_{n\to\infty} \Psi_n(y_n) \geq g^*(x') . \tag{11.27}$$

ii) For all $x \in X$ there exists a sequence $y_n' \in \mathcal{Y}^*$ such that

$$y_n' \to x \quad \text{in } \mathcal{Y}^* \quad \wedge \quad \limsup_{n\to\infty} \Psi_n^*\left(y_n'\right) \leq g(x) \tag{11.28}$$

Proof. **ii) ⇒ i).** This direction is the easier one. Choose $(y_n)_{n\in\mathbb{N}} \subset \mathcal{Y}$ such that $\mathrm{j}(y_n) \overset{*}{\rightharpoonup} x'$ for some $x' \in X$. Moreover choose $x \in X$ and a corresponding sequence $(y_n')_{n\in\mathbb{N}} \subset \mathcal{Y}^*$ such that (11.28) is satisfied. Then

$$\liminf_{n\to\infty} \Psi_n\left(y_n\right) \geq \liminf_{n\to\infty} \left\langle y_n', y_n \right\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi_n^*\left(y_n'\right)$$

Notice that by assumption

$$\liminf_{n\to\infty} -\Psi_n^*\left(y_n'\right) = -\limsup_{n\to\infty} \Psi_n^*\left(y_n'\right) \geq g(x) .$$

Moreover as $\mathrm{j}\,(y_n) \overset{*}{\rightharpoonup} x'$ in X^* we have $|\mathrm{j}\,(y_n)|_X$ is uniformly bounded and as j is an isometry so is $|y_n|_{\mathcal{Y}}$. We furthermore obtain that

$$\begin{aligned} &\left| \left\langle y_n', y_n \right\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \left\langle x', x \right\rangle_{X^*,X} \right| \\ &\quad \leq \; \left| y_n' - x \right| \left| y_n \right| + \left| \left\langle \mathrm{j}(y_n) - x', x \right\rangle_{X^*,X} \right| \to 0 . \end{aligned}$$

Hence we deduce that

$$\liminf_{n\to\infty} \Psi_n\left(y_n\right) \geq \left\langle x', x \right\rangle_{X^*,X} - g(x) .$$

Taking the sup over all $x \in X$ we obtain the claim.

i) ⇒ ii). In order to establish this direction we follow an approach by A. De Acosta [28, Appendix B] who showed for separable, possibly non-reflexive Banach spaces that $\Psi_n^* \xrightarrow{M} \Psi^* \Rightarrow \Psi_n \xrightarrow{M} \Psi$. In fact he assumed that Ψ_n* was an increasing sequence. However following the lines of his proof one can easily see that it already suffices to assume Mosco convergence.

For $c > 0$ we define $\Psi_{n,c}$ by

$$\Psi_{n,c}(y) = \Psi_n(y) + \frac{c}{2}|y|^2 .$$

Moreover we set $g_c : X \to [0, \infty)$ to be

$$g_c(x) = \sup \left\{ \langle x', x \rangle_{X^*,X} - g^*(x') - \frac{c}{2}|x'|^2 : x' \in X^* \right\} .$$

Claim 1. It holds that

$$\forall x \in X : \quad \limsup_{n\to\infty} \Psi_{n,c}^*(x) \le g_c(x) .$$

Without loss of generality we may assume that $g_c(x) < \infty$. Moreover we choose n_k to be the subsequence such that $\limsup \Psi_{n,c}^*(x) = \lim \Psi_{n_k,c}^*(x)$. We can also assume that

$$\lim_{k\to\infty} \Psi_{n_k,c}^*(x) > -\infty .$$

Otherwise the claim is trivially satisfied. Choose $y_n \in \mathcal{Y}$ such that

$$0 \le \Psi_{n_k,c}^*(x) \le \langle x, y_{n_k} \rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi_n(y_{n_k}) - \frac{c}{2}|y_{n_k}|^2 + \frac{1}{n_k} \le |x||y_{n_k}| - \frac{c}{2}|y_{n_k}|^2 + \frac{1}{n_k} .$$

Hence we may deduce that

$$\frac{c}{4}|y_{n_k}|^2 \le \frac{1}{c}|x|^2 - \frac{1}{n_k} \le C ,$$

where the constant C is independent of k but depends on x and c. We therefore deduce that y_{n_k} is uniformly bounded in $\mathcal{Y}$ and hence as well in X^*. Therefore there exists a $x' \in X^*$ such that up to a further subsequence $y_{n_k} \overset{*}{\rightharpoonup} x'$. Using the weak star lower semi continuity of the norm on X^* and the liminf inequality we hence deduce that

$$\begin{aligned}
\limsup_{n\to\infty} \Psi_{n,c}^*(x) &\le \limsup_{k\to\infty} \langle x, y_{n_k} \rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi_{n_k}(y_{n_k}) - \frac{c}{2}|y_{n_k}|^2 + \frac{1}{n_k} \\
&\le \langle x', x \rangle_{X^*,X} - g^*(x') - \frac{c}{2}|x'|^2 \le g_c(x) .
\end{aligned}$$

Moreover it also holds that $\sup_{c>0} g_c(x) = g(x)$. Indeed

$$\begin{aligned}
\sup_{c>0} g_c(x) &= \sup_{c>0} \sup_{x'\in X^*} \langle x', x \rangle_{X^*,X} - g^*(x') - \frac{c}{2}|x'|^2 \\
&= \sup_{x'\in X^*} \sup_{c>0} \langle x', x \rangle_{X^*,X} - g^*(x') - \frac{c}{2}|x'|^2 = g(x) .
\end{aligned}$$

The latter equality is due to the fact that for any Banach space X and any convex, proper, lower semi continuous function g it holds $g^{**} = g$. We hence obtain that

$$\limsup_{c\to 0} \lim_{n\to\infty} \Psi_{n,c}^*(x) \le g(x) .$$

Using a diagonalization argument there exists a function $c = c(n)$, $c(n) \to 0$ such that

$$\limsup_{n\to\infty} \Psi^*_{n,c(n)}(x) \le g(x)\,. \tag{11.29}$$

Notice that for all $n \in \mathbb{N}$ there exists a $y'_n \in \mathcal{Y}$ such that

$$\Psi^*_n\left(y'_n\right) + \frac{1}{2c(n)}\left|y'_n - x\right|^2 \le \Psi^*_{n,c(n)}(x) + \frac{1}{n}\,. \tag{11.30}$$

This readily implies that

$$\limsup_{n\to\infty} \Psi^*_n(y_n) \le g(x)\,.$$

It remains to show that $y'_n \to x$. This is a direct consequence of (11.29) and (11.30) because otherwise

$$\limsup_{n\to\infty}\left(\Psi^*_n\left(y'_n\right) + \frac{1}{2c(n)}\left|y'_n - x\right|^2\right) = \infty$$

This concludes the proof. □

If Ψ is a 1-homogeneous function then Ψ is a pseudo norm on $\mathcal{Y}$. We also can define a possibly non-symmetric metric on $\mathcal{Y}$ by

$$d_\Psi(u,v) = \Psi(v-u) \qquad \forall u,v \in \mathcal{Y}\,. \tag{11.31}$$

Now assume that the function $y : [0,T] \to \mathcal{Y}$ is absolutely continuous w.r.t the norm on $\mathcal{Y}$. Then due to the isometry of j the map $t \mapsto \mathrm{j}\,(y(t))$ is absolutely continuous with respect to the norm on X^*.

Proposition 11.15. *Let $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{j}\,(y(t))$ be the weak-star derivative of $\mathrm{j}\,(u(t))$. Then it holds that*

$$\Psi_{rel}\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{j}\,(y(t))\right) = \left|u'\right|_\Psi(t) \quad \textit{for a.e. } t \in [0,T]\,. \tag{11.32}$$

Here $\left|u'\right|_\Psi(t)$ denotes the metric derivative with respect to the metric d_Ψ.

Proof. Showing this is similar to before. If Ψ is 1-homogeneous then there exists some $K \subset X$ such that $\Psi = M_K$. Define $\widehat{K} = K^* \cap X$ and due to (11.26) it holds

$$\Psi(y) = \sup\left\{\langle \mathrm{j}(y), x\rangle_{X^*,X} : x \in \widehat{K}\right\}$$

We hence obtain for all $0 \le t < t+h \le T$ and $x \in \widehat{K}$ that

$$\Psi(y(t+h) - y(t)) \ge \langle \mathrm{j}\,(y(t+h)) - \mathrm{j}\,(y(t))\,, x\rangle_{X^*,X} = \int_t^{t} +h\left\langle \frac{\mathrm{d}}{\mathrm{d}s}\mathrm{j}(y(s)), x\right\rangle_{X^*,X} \mathrm{d}s\,.$$

Taking the limit $h \to 0$ we deduce for almost every $t \in [0,T]$ that

$$\liminf_{h\downarrow 0} \frac{\Psi(y(t+h) - y(t))}{h} \ge \Psi_{\mathrm{rel}}\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{j}\,(y(t))\right)\,.$$

Moreover one has for $0 \le r \le s \le T$ that

$$\Psi\left(y(s) - y(r)\right) \le \int_r^s \Psi_{\mathrm{rel}}\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{j}\,(y(t))\right)\mathrm{d}t\,.$$

This suffices to conclude that (11.32) holds. □

11.4. Statement of the main result

We will now state our main result. We especially have to define which precise assumption we have to make for the convergence $\Psi_n \to \Psi$. We are not able to prove a result while assuming only Mosco convergence. Because of the problems mentioned in the prior section, we instead have to develop a stronger notion of convergence. To see that this notion is still feasible for applications we are going to show that they are met for the two model examples we introduced at the beginning. As a last preparation we will state our assumption on the limiting potential Ψ.

Assumption 11.16 (Assumption on the limit dissipation potential)**.** We assume that the potential $\Psi : \mathcal{Y} \to [0, +\infty)$ is a proper, convex and lower semi continuous function. Moreover it satisfies $\Psi(0) = 0$ and there exist $c_1, c_2, c_3 \geq 0$, $p \geq 1$ and $q > 1$ such that

$$\forall \left(y, y'\right) \in \mathcal{Y} \times \mathcal{Y}^* : \Psi(y) + \Psi^*(y') \geq c_1 |y|_{\mathcal{Y}}^p + c_2 |y'|_{\mathcal{Y}^*}^q - c_3 \,. \tag{11.33}$$

Let us now state the main theorem of this chapter.

Theorem 11.17. *Let $\left(\mathcal{Y}, (\mathcal{Y}_n)_{n\in\mathbb{N}}, H\right)$ be an admissible family of state spaces. Let the triplet $(\mathcal{E}, F, P)$ satisfy Assumption 11.2. Let $(\Psi_n)_{n\in\mathbb{N}}$ be a family of convex, lower semi continuous potentials $\Psi_n : \mathcal{Y} \to [0, +\infty]$ satisfying Assumption 11.5. Let $(u_n, \xi_n)_{n\in\mathbb{N}}$ be a family of functions $(u_n, \xi_n) \in W^{1,1}(0, T; \mathcal{Y}_n) \times L^1(0, T, \mathcal{Y}_n^*)$ such that (u_n, ξ_n) is a solution to*

$$\partial\Psi_n\left(\dot{u}_n(t)\right) + F_t\left(u_n(t)\right) \ni 0\,, \quad u_n(0) = u_0$$

in the sense of Definition 11.3. Moreover choose Ψ such that Assumption 11.16 holds. Define $X \subset \mathcal{Y}^*$ such that the conditions of Proposition 11.14 are met. If

$$\forall x' \in X^*, (y_n)_{n\in\mathbb{N}} \subset \mathcal{Y} : \mathrm{j}(y_n) \overset{*}{\rightharpoonup} x : \qquad \liminf \Psi_n(y_n) \geq \Psi_{rel}(x) \tag{11.34}$$

$$\forall h \in H : \exists\, (h_n)_{n\in\mathbb{N}} \subset H,\ |h_n - h|_H \to 0 : \qquad \limsup \Psi_n(h_n) \leq \Psi(h) \tag{11.35}$$

then there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ and a tuple $(u, \xi) \in \mathrm{BV}_L(0, T; \mathcal{Y}) \times L^q(0, T; \mathcal{Y}^*)$ such that

$$u_{n_k}(t) \to u(t) \quad \forall t \in [0, T] \quad \wedge \quad \xi(t) \in F_t(u(t)) \quad \textit{a.e. in } [0, T]\,.$$

Moreover the tuple (u, ξ) satisfies $u(0) = u_0$ and it holds for all $t \in [0, T]$ that

$$\mathcal{E}_t(u(t)) + \int_0^t \Psi_{rel}\left(\dot{\hat{u}}_{ac}(\tau)\right) \mathrm{d}\tau + \int_0^t \Psi_{rel}^\infty\left(\dot{\hat{u}}_s(\tau)\right) \mathrm{d}\|\lambda_s\|(\tau) + \int_0^t \Psi^*\left(-\xi(\tau)\right) \mathrm{d}\tau$$
$$\leq \mathcal{E}_0\left(u(0)\right) + \int_0^t P_\tau\left(u(\tau), \xi(\tau)\right) \mathrm{d}\tau\,. \tag{11.36}$$

Certainly a few remarks are in order to explain the convergence conditions (11.34) and (11.35). First of all we shall point out that these two conditions imply that $\Psi_n \overset{M}{\longrightarrow} \Psi$. For the $\liminf$-inequality just note that $y_n \rightharpoonup y$ implies $\mathrm{j}(y_n) \overset{*}{\rightharpoonup} \mathrm{j}(y)$. For the existence of a strong recovery sequence notice that $D(\Psi) = \mathcal{Y}$. Hence Ψ is continuous on $\mathcal{Y}$. Moreover $H \hookrightarrow \mathcal{Y}$ is dense. For any $y \in \mathcal{Y}$ choose a sequence $(h_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} |h_n - y|_{\mathcal{Y}} = 0\,.$$

For each h_n there exists a sequence $(h_{n,k})_{k\in\mathbb{N}}$ such that $\lim |h_{n,k} - h_n|_H = 0$ and the $\limsup$ inequality (11.35) is satisfied. Now remember that convergence with respect to the H-norm implies convergence with respect to the $\mathcal{Y}$-norm. A strong recovery sequence for y can hence be obtained by extracting a diagonal sequence.

As the convergence conditions (11.34) and (11.35) are stronger than Mosco convergence one may ask whether these are actually met in 'interesting' cases. To argue in favor of these two conditions let us shortly show that the two examples from above satisfy these conditions.

Proposition 11.18. *Let $\Omega \subset \mathbb{R}^n$ be an open, bounded and connected set. Define $\mathcal{Y} = L^1(\Omega)$ and $H = L^2(\Omega)$. Choose $\Psi = \|\cdot\|_{L^1(\Omega)}$. The set $X = C_b(\Omega)$ satisfies the condition of Proposition 11.14. For any sequence $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,1]$ such that $\varepsilon_n \to 0$ it holds that both families*

$$\Psi_n^1(\cdot) = \frac{1}{\varepsilon_n}\|\cdot\|_{L^{1+\varepsilon_n}(\Omega)}^{1+\varepsilon_n} \quad \textit{and} \quad \Psi_n^2(\cdot) = \|\cdot\|_{L^1(\Omega)} + \varepsilon_n\|\cdot\|_{L^2(\Omega}^2$$

satisfy the convergence conditions (11.34) *and* (11.35) *with respect to* Ψ.

Proof. **ad Ψ_n^1.** To see that condition $\liminf$-condition holds choose a measure $\mu \in \mathcal{M}(\Omega) = X^*$. Take any family of function $(y_n)_{n\in\mathbb{N}} \subset L^1(\Omega)$ such that $y_n \cdot \mathcal{L} \overset{*}{\rightharpoonup} \mu$. If $y_n \notin L^{1+\varepsilon_n}(\Omega)$ then $\Psi_n^1(y_n) = +\infty$. Otherwise

$$\Psi_n^1(y_n) = \frac{1}{1+\varepsilon_n}\int_\Omega |y_n(x)|^{1+\varepsilon_n}\,\mathrm{d}x \geq \int_\Omega |y_n(x)|\,\mathrm{d}x - \frac{\varepsilon_n}{1+\varepsilon_n}\mathcal{L}(\Omega)\,.$$

This is a consequence of Hölders inequality. Notice that $\frac{\varepsilon_n}{1+\varepsilon_n}\mathcal{L}(\Omega) \to 0$ as $n \to \infty$. Moreover it holds that

$$\|y_n\|_{L^1(\Omega)} = \|y_n \cdot \mathcal{L}\|_{\mathcal{M}(\Omega)}\,.$$

Then (11.34) is a consequence of the weak-star lower semi continuity of the norm on $\mathcal{M}(\Omega)$. In order to find a recovery sequence choose any $h \in L^2(\Omega)$. Then $\Psi_n(h) < \infty$. Moreover it holds that

$$|h(x)|^{1+\varepsilon} \to |h(x)| \quad \text{for a.e. } x \in \Omega\,.$$

From the dominated convergence principle we then deduce that $\Psi_n(h) \to \Psi(h)$.

ad Ψ_n^2. The proof is similar. The liminf condition follows immediately from the fact that

$$\Psi_n^2(y_n) \geq \|y_n\|_{L^1(\Omega)}\,.$$

For the recovery sequence one may once more choose any $h \in L^2(\Omega)$. Then

$$\Psi_n^2(h) = \|h\|_{L^1(\Omega)} + \varepsilon_n\,\|h\|_{L^2(\Omega)} \to \|h\|_{L^1(\Omega)} \quad (\text{as } n \to \infty)\,.$$

□

11.5. Proof of the main theorem

Before we can prove Theorem 11.17 we need to prove another selection principle.

Proposition 11.19. *Assume $(\mathcal{E}, F, P)$ satisfies the Assumption 11.2 and Ψ satisfies Assumption 11.16. Moreover let $u \in BV(0,T;\mathcal{Y})$ such that*

$$\text{for a.a. } t \in [0,T] : (t,u(t)) \in \mathrm{dom}(F) \quad \wedge \quad \sup\{\mathcal{E}_t(u(t)) : t \in [0,T]\} < +\infty\,. \tag{11.37}$$

Assume that the set

$$S(t,u(t)) = \{(\zeta,p) \in H \times \mathbb{R} : \zeta \in F_t(u(t))\,, p \leq P_t(u(t),\zeta)\} \tag{11.38}$$

is non-empty for almost all $t \in [0,T]$. Then there exist two measurable functions $\xi : (0,T) \to H$, $p : (0,T) \to \mathbb{R}$ such that

$$(\xi(t),p(t)) \in \mathrm{argmin}\{\Psi^*(\zeta) - p : (\zeta,p) \in S(t,u(t))\} \quad \text{a.e. in } [0,T] \tag{11.39}$$

This is in fact a simplification of [69, Lemma B.2]. Hence the proof is simpler. The reader might be familiar with the structure of the proof also from Lemma 10.11.

Proof. We start by showing that

$$\mathrm{argmin}\{\Psi^*(\zeta) - p : (\zeta,p) \in S(t,u(t))\} \neq \emptyset \quad \text{for a.e. } t \in [0,T] \tag{11.40}$$

Choose some $t \in [0,T]$ such that $S(t,u(t)) \neq \emptyset$. Without loss of generality we assume that

$$\inf\{\Psi^*(\zeta) - p : (\zeta,p) \in S(t,u(t))\} < +\infty\,.$$

Otherwise the argmin coincides with the set $S(t,u(t))$. Now let (ζ_n,p_n) be an infimizing sequence. Then analogous to (10.25) we obtain that both p_n and $\Psi^*(\zeta_n)$ are uniformly bounded. We hence obtain that ζ_n is uniformly bounded in $\mathcal{Y}^*$ and there exists a subsequence $(\zeta_{n_k},p_{n_k})_{k\in\mathbb{N}}$ such that

$$\zeta_{n_k} \overset{*}{\rightharpoonup} \zeta \quad \text{and} \quad p_{n_k} \to p\,.$$

Especially it holds that $\zeta_{n_k} \rightharpoonup \zeta$ with respect to the weak topology in H as well. Thus due to $(11.\mathcal{E}_6)$ we have $(\zeta,p) \in S(t,u(t))$ and also obtain that

$$(\zeta,p) \in \mathrm{argmin}\{\Psi^*(\zeta) - p : (\zeta,p) \in S(t,u(t))\}$$

Now showing that the existence of measurable selection follows the same lines as in the proof of [69, Lemma B.2]. It is even simpler as there is no dependence on $\dot{u}(t)$. First we define the set

$$\mathcal{S} := \{(t,u,\zeta,p) \in [0,T] \times \mathcal{Y} \times H \times \mathbb{R} : \zeta \in F_t(u)\,, p \leq P_t(u,\zeta)\}$$

This is a Borel set as graph(F) is a Borel set and the map $t :\mapsto P_t(u,\zeta)$ is a Borel function. By assumption there exists a set $\mathcal{T} \subset [0,T]$ of full measure such that

$$S(t,u(t)) \neq \emptyset \quad \forall t \in \mathcal{T}\,.$$

We then define the multivalued map $t \in \mathcal{T} \mapsto S(t,u(t))$. Using the fact that u is a Borel function we then obtain once again from [20, Corollary III.3 and Theorem III.6] that there exists a measurable selection which satisfies 11.39. □

We are now able to prove our main result.

Proof of Theorem 11.17. The steps we follow are similar to the proof of Theorem 10.8.

Step 1 - A priori estimates and compactness. For $n \in \mathbb{N}$ we define

$$E_n : [0,T] \to \mathbb{R}\,, \quad t \mapsto \mathcal{E}_t\left(u_n(t)\right)\,. \tag{11.41}$$

As $\Psi_n, \Psi_n^* \geq 0$ we deduce that

$$E_n(t) \leq C_0 + C_1 \int_0^t E_n(s)\,\mathrm{d}s\,. \tag{11.42}$$

Employing Gronwall's Lemma we obtain $\sup\{E_n(t) : t \in [0,T]\} \leq C$ for some C independent of n. Consequently we have

$$\text{ess sup}\left\{|P_t(u_n(t), \xi_n(t))| : t \in [0,T]\right\} \leq C\,. \tag{11.43}$$

From this we get

$$\int_0^T \Psi_n\left(\dot{u}_n(t)\right) + \Psi_n^*\left(-\xi_n(t)\right)\,\mathrm{d}t \leq C\,. \tag{11.44}$$

The growth condition (11.11) now implies

$$\int_0^T c_1\left|\dot{u}_n(t)\right|_{\mathcal{Y}} + c_2\left|\xi(t)\right|_{H^*}\,\mathrm{d}t \leq C \tag{11.45}$$

Therefore $(u_n)_{n\in\mathbb{N}}$ and $(\xi_n)_{n\in\mathbb{N}}$ are uniformly bounded in $L^1\left(0,T;\mathcal{Y}\right)$ and $L^q(0,T;H)$ respectively. Finally analogous to the proof of Theorem 10.8 we obtain that $t \mapsto E_n(t)$ is uniformly bounded in $BV(0,T)$.
Hence passing to an appropriate subsequence $(n_k)_{k\in\mathbb{N}}$ we find functions $E \in BV(0,T)$, $u \in BV(0,T;\mathcal{Y})$ such that

$$(E_n(t), u_n(t)) \to (E(t), u(t)) \quad \forall t \in [0,T]\,. \tag{11.46}$$

Additionally there exists a family of Young measures $(\sigma_t)_{t\in[0,T]}$, $\sigma_t \in \mathcal{M}\left(H \times \mathbb{R}\right)$ satisfying Theorem 4.22. Especially we can deduce that $\|\sigma_t\| = 1$ almost everywhere and

$$\text{supp}\left(\sigma_t\right) \subset \bigcap_{j\in\mathbb{N}} \text{cl}\left\{(\xi_{n_k}(t), p_{n_k}(t)) : k \geq j\right\} \tag{11.47}$$

where cl is meant to be the weak closure in $H \times \mathbb{R}$. Finally notice that $\mathcal{Y}_n \hookrightarrow X^*$ for all $n \in \mathbb{N}$. Define for all $t \in [0,T]$ and $n \in \mathbb{N}$

$$\widehat{u}_n(t) := \mathfrak{j}\left(u_n(t)\right)\,. \tag{11.48}$$

Any absolutely continuous function on $\mathcal{Y}$ is absolutely continuous on X^*. Furthermore let $\dot{\widehat{u}}_n$ be the weak-star derivative of $\widehat{u}_n$ which exists due to Theorem 4.18. Then $\dot{\widehat{u}}_n(t) = \mathfrak{j}\left(\dot{u}_n(t)\right)$

for almost every $t \in [0,T]$. Furthermore the family $\lambda_n = \dot{\hat{u}}_n \cdot \mathcal{L}$ is uniformly bounded in $\mathcal{M}(0,T,X^*)$. To see this let $0 \le r \le s \le t$. Then

$$|\lambda_n([r,s])|_{X^*} = \sup\left\{\langle \lambda_n([r,s]), x\rangle_{X^*,X} : |x| \le 1\right\}. \tag{11.49}$$

We then can estimate that for all $x \in X$ with $|x| \le 1$ it holds

$$\begin{aligned}\langle \lambda_n([r,s]), x\rangle_{X^*,X} &= \langle \mathfrak{j}(u_n(s)) - \mathfrak{j}(u_n(r)), x\rangle_{X^*,X} \\ &= \langle x, u_n(s) - u_n(r)\rangle_{\mathcal{Y}^*,\mathcal{Y}} \le \int_r^s |\dot{u}_n(t)|_{\mathcal{Y}}\, \mathrm{d}t\end{aligned}$$

Applying (11.3) we hence know that for any closed set $[r,s] \subset [0,T]$ it holds

$$|\lambda_n([r,s])|_{X^*} \le C \int_r^s |\dot{u}_n(t)|_{\mathcal{Y}_n}\, \mathrm{d}t \tag{11.50}$$

It is thus an immediate consequence of (11.45) that λ_n is uniformly bounded in $\mathcal{M}(0,T;X^*)$. Hence there exists a further not relabeled subsequence $(n_k)_{k\in\mathbb{N}}$ such that

$$\lambda_{n_k} \overset{*}{\rightharpoonup} \lambda \quad \text{in } \mathcal{M}(0,T;X^*). \tag{11.51}$$

Step 2 - $\liminf$**-inequality for the dissipation potential.** We are now going to show that

$$\liminf_{n\to\infty} \int_0^t \Psi_n(\dot{u}_n(t))\, \mathrm{d}t \ge \int_0^t \Psi_{\mathrm{rel}}\left(\dot{\hat{u}}_{ac}(s)\right) \mathrm{d}s + \int_0^t \Psi_{\mathrm{rel}}^{\infty}\left(\dot{\hat{u}}_s(s)\right) \mathrm{d}\|\lambda_s\|(s) \tag{11.52}$$

We shall first pass to a non-relabeled subsequence such that the $\liminf$ is indeed a limit. Notice first that due to (11.44) the family

$$f_n : t \mapsto \Psi_n(\dot{u}_n(t)) \tag{11.53}$$

is uniformly bounded in $L^1(0,T)$. There exists a subsequence $(n_k)_{k\in\mathbb{N}}$ and a measure $\mu \in \mathcal{M}(0,T)$ such that

$$f_n \cdot \mathcal{L} \overset{*}{\rightharpoonup} \mu \tag{11.54}$$

Hence there exist measures $\lambda_{ac}, \lambda_s \in \mathcal{M}(0,T;X^*)$ and $\mu_{ac}, \mu_s, \mu_\perp \in \mathcal{M}(0,T)$ such that

$$\|\lambda_{ac}\| \ll \mathcal{L} \quad \wedge \quad \|\lambda_s\| \perp \mathcal{L} \quad \wedge \quad \lambda = \lambda_s + \lambda_{ac} \quad \text{and} \tag{11.55}$$

$$\mu_{ac} \ll \mathcal{L} \quad \wedge \quad \mu_s \ll \|\lambda_s\| \quad \wedge \quad \mu_\perp \perp \mathcal{L} + \|\lambda_s\| \quad \wedge \quad \mu = \mu_{ac} + \mu_s + \mu_\perp \tag{11.56}$$

For the sake of notational simplicity we shall simply write the index k instead of (n_k). Define as before $Q(t_0,\varepsilon) := \left(t_0 - \frac{\varepsilon}{2}, t_0 + \frac{\varepsilon}{2}\right)$. For any countable family $(x_i)_{i\in\mathbb{N}} \subset X$ there exists set $\mathcal{T} \subset [0,T]$ of full measure such that for all $t_0 \in \mathcal{T}$ and all $i \in \mathbb{N}$ it holds

$$\lim_{\varepsilon\to 0} \frac{\langle \lambda(Q(t_0,\varepsilon) \cap [0,T]), x_i\rangle_{X^*,X}}{\mathcal{L}(Q(t_0,\varepsilon))} = \left\langle \dot{\hat{u}}_{ac}(t_0), x_i\right\rangle_{X^*,X}. \tag{11.57}$$

If one chooses $(x_i)_{i\in\mathbb{N}}$ to be strongly dense in X a simple limit argument shows that the above equality holds for all $x \in X$. Moreover for almost every $t \in [0,T]$ it holds that

$$\frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) = \lim_{\varepsilon\to 0} \frac{\mu\left(Q(t_0,\varepsilon)\cap[0,T]\right)}{\mathcal{L}\left(Q(t_0,\varepsilon)\cap[0,T]\right)} < \infty\,. \tag{11.58}$$

Now choose $x \in X$ such that $\Psi^*(x) < +\infty$. Then as a consequence of Proposition 11.13 there exists a strong recovery sequence $(y_k') \subset \mathcal{Y}^*$ such that $y_k' \to x$ and $\Psi_n^*\left(y_k'\right) \to \Psi^*(x)$. Furthermore we can choose a subsequence ε_m such that $\varepsilon_m \to 0$ and $\lambda\left(\partial Q(t_0,\varepsilon_m)\cap[0,T]\right) = \mu\left(\partial Q(t_0,\varepsilon_m)\right) = 0$. It then holds that

$$\begin{aligned}\frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) &= \lim_{m\to\infty}\lim_{k\to\infty} \frac{1}{\mathcal{L}\left(Q(t_0,\varepsilon_m)\cap[0,T]\right)} \int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \Psi_k(\dot u_k(t))\,\mathrm{d}t \\ &\geq \liminf_{m\to\infty} \frac{1}{\mathcal{L}\left(Q(t_0,\varepsilon_m)\cap[0,T]\right)} \liminf_{k\to\infty} \int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \left\langle y_k', \dot u_k(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,\mathrm{d}t - \Psi_k^*\left(y_k'\right)\cdot\mathcal{L}\left(Q(t_0,\varepsilon_m)\cap\right.\end{aligned}$$

We then can write

$$\begin{aligned}\int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \left\langle y_k', \dot u_k(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,\mathrm{d}t &= \int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \left\langle x, \dot u_k(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,\mathrm{d}t \\ &\quad + \int\limits_{Q(t_0,\varepsilon_m)} \left\langle y_k' - x, \dot u_n(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,\mathrm{d}t\end{aligned}$$

For the first term we obtain

$$\begin{aligned}&\int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \left\langle x, \dot u_k(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,\mathrm{d}t \\ &\quad = \int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \left\langle \mathrm{j}\left(\dot u_k(t)\right), x\right\rangle_{X^*,X}\,\mathrm{d}t \to \left\langle \lambda\left(Q\left(t_0,\varepsilon_m\right)\cap[0,T]\right), x\right\rangle_{X^*,X}\,.\end{aligned}$$

The second term can simply be estimated by

$$\int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]} \left|\left\langle y_k' - x, \dot u_k(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}\right|\,\mathrm{d}t \leq \left|y_k' - x\right|_{\mathcal{Y}^*} \left\|\dot u_k(t)\right\|_{L^1(0,T;\mathcal{Y}^*)} \to 0\,. \tag{11.59}$$

Letting $m \to \infty$ we finally deduce that

$$\frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) \geq \left\langle \dot{\hat u}_{ac}(t_0), x\right\rangle_{X^*,X} - \Psi^*(x) \tag{11.60}$$

and taking the supremum over all $x \in X$ we obtain

$$\frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) \geq \Psi_{\text{rel}}\left(\dot{\hat u}_{ac}(t_0)\right)\,. \tag{11.61}$$

Now let $\mathcal{T}_s \subset [0,T]$ be the set of all t_0 such that

$$\lim_{\varepsilon\to 0} \frac{\mu\left(Q\left(t_0,\varepsilon\right)\cap[0,T]\right)}{\|\lambda_s\|\left(Q\left(t_0,\varepsilon\right)\cap[0,T]\right)} = \frac{\mathrm{d}\mu_s}{\mathrm{d}\|\lambda_s\|}(t_0) \tag{11.62}$$

$$\lim_{\varepsilon\to 0} \frac{\left\langle\lambda\left(Q\left(t_0,\varepsilon\right)\cap[0,T]\right),x\right\rangle_{X^*,X}}{\|\lambda_s\|\left(Q\left(t_0,\varepsilon\right)\cap[0,T]\right)} = \left\langle\dot{\hat{u}}_s(t_0),x\right\rangle_{X^*,X} \quad \forall x\in X \qquad \text{and} \tag{11.63}$$

$$\lim_{\varepsilon\to 0} \frac{\mathcal{L}\left(Q\left(t_0,\varepsilon\right)\cap[0,T]\right)}{\|\lambda_s\|\left(Q\left(t_0,\varepsilon\right)\cap[0,T]\right)} = 0\,. \tag{11.64}$$

Then the set $[0,T]\setminus\mathcal{T}_s$ is a $\|\lambda_s\|$-null set. Once again let us choose a sequence $(\varepsilon_m)_{m\in\mathbb{N}}$, $\varepsilon_m\to 0$ such that $\lambda\left(\partial Q(t_0,\varepsilon_m)\right) = \mu\left(\partial Q(t_0,\varepsilon_m)\cap[0,T]\right) = 0$. Choosing $x\in X$ with $\Psi^*(x) < +\infty$ and a corresponding strong recovery sequence $(y_k)_{k\in\mathbb{N}}$ we now may compute

$$\begin{aligned}\frac{\mathrm{d}\mu_s}{\mathrm{d}\mathcal{L}}(t_0) &= \lim_{m\to\infty}\lim_{k\to\infty}\frac{1}{\|\lambda_s\|\left(Q(t_0,\varepsilon_m)\right)}\int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]}\Psi_k(\dot{u}_k(t))\,\mathrm{d}t\\ &\geq \liminf_{m\to\infty}\liminf_{k\to\infty}\left(\|\lambda_s\|\left(Q(t_0,\varepsilon_m)\cap[0,T]\right)\right)^{-1}\cdot\\ &\qquad\left(\int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]}\left\langle y_k',\dot{u}_k(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,\mathrm{d}t - \mathcal{L}\left(Q(t_0,\varepsilon_m)\cap[0,T]\right)\Psi_k^*\left(y_k'\right)\right).\end{aligned}$$

Analogous to our prior calculations we obtain that

$$\int\limits_{Q(t_0,\varepsilon_m)\cap[0,T]}\left\langle y_k',\dot{u}_k(t)\right\rangle_{\mathcal{Y}^*,\mathcal{Y}}\,\mathrm{d}t \to \left\langle\lambda\left(Q\left(t_0,\varepsilon_m\right)\cap[0,T]\right),x\right\rangle_{X^*,X}. \tag{11.65}$$

Moreover as $m\to\infty$ we deduce that

$$\liminf_{m\to\infty}\frac{\left\langle\lambda\left(Q\left(t_0,\varepsilon_m\right)\cap[0,T]\right),x\right\rangle_{X^*,X} - \mathcal{L}\left(Q(t_0,\varepsilon_m)\cap[0,T]\right)\Psi^*\left(x\right)}{\|\lambda_s\|\left(Q(t_0,\varepsilon_m)\cap[0,T]\right)} = \left\langle\dot{\hat{u}}_s(t_0),x\right\rangle_{X^*,X}. \tag{11.66}$$

Hence taking the supremum over all $x\in X$ we finally deduce that

$$\frac{\mathrm{d}\mu_s}{\mathrm{d}\mathcal{L}}(t_0) \geq \Psi_{\mathrm{rel}}^{\infty}\left(\dot{\hat{u}}_s(t_0)\right). \tag{11.67}$$

Inequality (11.52) now is a consequence of integration over $[0,T]$ and the fact that as $\Psi_n\geq 0$ also $\mu_\perp\geq 0$.

Step 3 - $\liminf$**-inequality for the Young measure** It follows from the weak strong closedness (11.$\mathcal{E}_6$) that

$$\operatorname{supp}(\sigma_t)\subset\left\{(\zeta,p):\zeta\in F_t(u(t)), p\leq P_t(u(t),\zeta)\right\} \quad \text{a.e. in } [0,T]\,. \tag{11.68}$$

Moreover we obtain that for almost every $t\in[0,T]$ it holds that $(t,u(t))\in D(F)$ and $E(t) = \mathcal{E}_t(u(t))$. It remains to show that

$$\liminf_{n\to\infty}\int\limits_0^t\Psi_n^*\left(-\xi_n(s)\right)\,\mathrm{d}s \geq \int\limits_0^t\int_{H\times\mathbb{R}}\Psi^*(-\zeta)\,\mathrm{d}\sigma_s(\zeta,p)\,\mathrm{d}s\,. \tag{11.69}$$

This is a consequence of Theorem 4.22, if we can show that for all families $(h'_n)_{n\in\mathbb{N}} \subset H^* \cong H$, $h'_n \rightharpoonup h'$ it holds that

$$\liminf_{n\to\infty} \Psi_n^*(h'_n) \geq \Psi^*(h') \,. \tag{11.70}$$

To see this choose any $h \in H$. Then there exists a family $(h_n)_{n\in\mathbb{N}} \subset H$, $h_n \to h$ such that (11.35) is satisfied. Then it holds that

$$\liminf_{n\to\infty} \Psi_n^* \left(h'_n\right) \geq \liminf_{n\to\infty} \left\langle h'_n, h_n \right\rangle - \Psi_n \left(h_n\right) \geq \left\langle h', h \right\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi(h) \,.$$

Taking the fact that $H \hookrightarrow \mathcal{Y}$ dense and the fact that Ψ is continuous as it is defined everywhere on $\mathcal{Y}$ we deduce that

$$\forall h' \in H^* \setminus \mathcal{Y}^* : \quad \sup_{h\in H} \left\langle h', h \right\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi(h) = +\infty \,.$$

On the other hand for all $h' \in \mathcal{Y}^* \subset H^*$ it holds that

$$\sup_{h\in H} \left\langle h', h \right\rangle - \Psi(h) = \sup_{y\in\mathcal{Y}} \left\langle h', y \right\rangle_{\mathcal{Y}^*,\mathcal{Y}} - \Psi(y) = \Psi^*(h') \,.$$

This implies (11.70).

Step 4 - Selection argument and conclusion of the proof. Finally we obtain from Proposition 11.19 that there exists a pair of functions (ξ, p) such that

$$\Psi^* \left(\xi(t)\right) - p(t) \leq \int_{H\times\mathbb{R}} \Psi^*(\zeta) - p \,\mathrm{d}\sigma_t(\zeta, p)$$

Now combining all the above results and integrating and integrating over $[0, t)$ we deduce that for almost all $t \in [0, T]$ it holds

$$\begin{aligned} \mathcal{E}_t(u(t)) + \int_0^t \Psi_{\text{rel}} \left(\dot{\widehat{u}}_{ac}(\tau)\right) \mathrm{d}\tau + \int_0^t \Psi_{\text{rel}}^\infty \left(\dot{\widehat{u}}_s(\tau)\right) \mathrm{d}\|\lambda_s\|(\tau) + \int_0^t \Psi^* \left(-\xi(\tau)\right) \mathrm{d}\tau \\ \leq \mathcal{E}_0 \left(u(0)\right) + \int_0^t P_s \left(u(s), \xi(s)\right) \mathrm{d}s \,. \end{aligned}$$

□

11.6. The case of rate-independent limits

We will finish this chapter with a few remarks concerning the rate independent case. Assume that Ψ is a 1-homogeneous function. Then there exists a convex set $K \subset \mathcal{Y}$ such that $\Psi = M_K$ and it holds that $\Psi^* = I_{K^*}$. It follows from

$$\int_0^T \Psi^* \left(-\xi(t)\right) \mathrm{d}t = \int_0^T I_{K^*} \left(-\xi(t)\right) \mathrm{d}t < +\infty$$

that $-\xi(t) \in K^*$ for almost every $t \in [0,T]$. In the terminology of rate-independent processes this means that the function u satisfies the local stability condition. Moreover also Ψ_{rel} is also 1-homogeneous and it holds $\Psi_{\text{rel}}^\infty = \Psi_{\text{rel}}$. Then the limit inequality (11.36) reads

$$\begin{aligned}\mathcal{E}_t(u(t)) + \int_0^t \Psi_{\text{rel}}\left(\dot{\hat{u}}_{ac}(\tau)\right)\,\mathrm{d}\tau + \int_0^t \Psi_{\text{rel}}\left(\dot{\hat{u}}_s(\tau)\right)\,\mathrm{d}\|\lambda_s\|(\tau) + \int_0^t I_{K^*}\left(-\xi(\tau)\right)\,\mathrm{d}\tau \\ \leq \mathcal{E}_0\left(u(0)\right) + \int_0^t P_\tau\left(u(\tau),\xi(\tau)\right)\,\mathrm{d}\tau\,.\end{aligned}$$

In order to give context to this equation in the purview of solution concepts for rate independent systems we need the following proposition.

Proposition 11.20. *Let $\mathcal{Y}$ be Banach space and $\Psi : \mathcal{Y} \to [0,+\infty]$ be a proper, 1-homogeneous, convex and lower semi continuous functional satisfying Assumption 11.16. Let $X \subset \mathcal{Y}^*$ satisfy the conditions of Proposition 11.14. Let Ψ_{rel} be the relaxation of Ψ to X^* according to* (11.21). *For a function $u \in \mathrm{BV}(0,T;\mathcal{Y})$ define*

$$\begin{aligned}&\mathrm{Diss}_\Psi\left(u,[0,t)\right) := \\ &\quad\sup\left\{\sum_{i=1}^n \Psi\left(u(t_i-) - u(t_{i-1}+)\right) + \sum_{i=0}^n \Psi\left(u(t_i+) - u(t_i-)\right) : (t_i)_{i=0}^n \subset [0,t),\ t_{i-1} < t_i\right\}.\end{aligned}$$

Then it holds that

$$\mathrm{Diss}_\Psi(u,[0,t)) = \int_{[0,t)} \Psi_{rel}\left(\dot{\hat{u}}_{ac}(\tau)\right)\,\mathrm{d}\tau + \int_{[0,t)} \Psi_{rel}\left(\dot{\hat{u}}_s(\tau)\right)\,\mathrm{d}\|\lambda_s\|(\tau) \tag{11.71}$$

where $\lambda, \lambda_s, \lambda_{ac}$ and $\dot{\hat{u}}_{ac}$, $\dot{\hat{u}}_s$ are defined as in the above proof.

We now can rewrite the limit equation in the rate-independent case as

$$\mathcal{E}_t(u(t)) + \mathrm{Diss}_\Psi(u,[0,t)) + \int_0^t I_{K^*}\left(-\xi(\tau)\right)\,\mathrm{d}\tau \leq \mathcal{E}_0\left(u(0)\right) + \int_0^t P_\tau\left(u(\tau),\xi(\tau)\right)\,\mathrm{d}\tau\,.$$

In the language of rate independent systems such a function pair is called a local solution to the evolution equation (11.1), see e.g. [73]. It is one of the weakest notions of solutions for rate independent systems. Especially in general uniqueness cannot be expected [102]. Approximating rate independent processes has been considered in a number of works, see e.g. [61, 72, 73, 70]. Whereas in these cases a more precise understanding of what happens at jump points has been obtained the form of the approximation was limited to specific families of dissipation potentials. The benefit of the above result is that the family of dissipation potentials is rather arbitrary. However one pays for this by loosing information on what happens when the solution jumps. One could argue that the above theorem shows that the minimal requirement for a solution of a rate independent system should be that it is a local solution.

As the final act of this chapter let us prove Proposition 11.20.

Proof. Every $u \in BV(0,T;\mathcal{Y})$ admits left and right hand side limits at every $t \in [0,T]$. For convenience we set $u(0-) := u(0)$ and $u(T+) := u(T)$. Therefore the above definition of Diss_Ψ is sensible. Moreover we define $u^+ : [0,T] \to \mathcal{Y}$ by $u^+ : t \mapsto u(t+)$ and $u^- : [0,T] \to \mathcal{Y}$ by $u^- : t \mapsto u(t-)$. As Ψ is 1-homogeneous there exists a convex set $K \subset \mathcal{Y}$ such that $\Psi = M_K$ and consequently $\Psi^* = I_{K^*}$. Set $\widehat{K} = K \cap X$. Then as a consequence of (11.21) we obtain that $\Psi_{\mathrm{rel}} = \left(I_{\widehat{K}}\right)^*$. Moreover due to (11.26) it holds

$$\mathrm{Diss}_\Psi(u,[0,t]) = \mathrm{Diss}_{\Psi_{\mathrm{rel}}}(\mathrm{j}(u),[0,t]) \,.$$

Note that Diss_Ψ is a left continuous function of bounded variation. Let μ be the Radon measure induced by Diss_Ψ. Notice that as Diss_Ψ is monotone increasing $\mu \geq 0$. Moreover there exist measures $\lambda_{ac}, \lambda_s \in \mathcal{M}(0,T;X^*)$ and $\mu_{ac}, \mu_s, \mu_\perp \in \mathcal{M}(0,T)$ such that

$$\|\lambda_{ac}\| \ll \mathcal{L} \quad \wedge \quad \|\lambda_s\| \perp \mathcal{L} \quad \wedge \quad \lambda = \lambda_s + \lambda_{ac} \quad \text{and}$$
$$\mu_{ac} \ll \mathcal{L} \quad \wedge \quad \mu_s \ll \|\lambda_s\| \quad \wedge \quad \mu_\perp \perp \mathcal{L} + \|\lambda_s\| \quad \wedge \quad \mu = \mu_{ac} + \mu_s + \mu_\perp$$

Notice that $\mu_\perp \geq 0$. Then for $\mathcal{L}$-almost all $t_0 \in [0,T]$ it holds

$$\begin{aligned}
\left\langle \dot{\hat{u}}_{ac}(t_0), x \right\rangle_{X^*,X} &= \lim_{\varepsilon\to 0} \frac{\langle \lambda\left(Q(t_0,\varepsilon)\right), x\rangle_{X^*,X}}{\mathcal{L}\left(Q(t_0,\varepsilon)\right)} \quad \text{and} \\
\frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) &= \lim_{\varepsilon\to 0} \frac{\mu\left(Q(t_0,\varepsilon)\right)}{\mathcal{L}\left(Q(t_0,\varepsilon)\right)} < \infty \,.
\end{aligned}$$

Without loss of generality we may assume that $t_0 \in (0,T)$. Choose ε small enough such that $[t_0 - \varepsilon/2, t_0 + \varepsilon/2] \in [0,T]$. Then

$$\begin{aligned}
\frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) &\geq \liminf_{\varepsilon\to 0} \varepsilon^{-1}\Psi\left(u^-(t_0+\varepsilon/2) - u^+(t_0-\varepsilon/2)\right) \\
&\geq \liminf_{\varepsilon\to 0} \varepsilon^{-1}\left\langle \lambda\left(Q(t_0,\varepsilon)\right), x\right\rangle_{X^*,X} = \left\langle \dot{\hat{u}}_{ac}(t_0), x\right\rangle_{X^*,X} .
\end{aligned}$$

Taking the sup over all $x \in \widehat{K}$ it holds

$$\frac{\mathrm{d}\mu_{ac}}{\mathrm{d}\mathcal{L}}(t_0) \geq \Psi_{\mathrm{rel}}\left(\dot{\hat{u}}_{ac}(t_0)\right).$$

We proceed analogous for $t \in \mathrm{supp}\left(\|\lambda_s\|\right)$. It holds for $\|\lambda_s\|$-almost every $t \in [0,T]$ that

$$\begin{aligned}
\lim_{\varepsilon\to 0} \frac{\mu\left(Q\left(t_0,\varepsilon\right)\cap[0,T]\right)}{\|\lambda_s\|\left(Q\left(t_0,\varepsilon\right)\cap[0,T]\right)} &= \frac{\mathrm{d}\mu_s}{\mathrm{d}\|\lambda_s\|}(t_0) \\
\lim_{\varepsilon\to 0} \frac{\langle\lambda\left(Q\left(t_0,\varepsilon\right)\cap[0,T]\right), x\rangle_{X^*,X}}{\|\lambda_s\|\left(Q\left(t_0,\varepsilon\right)\cap[0,T]\right)} &= \left\langle \dot{\hat{u}}_s(t_0), x\right\rangle_{X^*,X} \quad \forall x \in X \,.
\end{aligned}$$

However in this case we cannot simply assume that $t_0 \in [0,T]$. For $t_0 \in \{0,T\}$ a small case distinction is necessary by considering either $\Psi\left(u(0+\varepsilon/2) - u(0)\right)$. Apart from that the estimate works as above and we deduce that

$$\frac{\mathrm{d}\mu_s}{\mathrm{d}\|\lambda_s\|}(t_0) \geq \Psi_{\mathrm{rel}}\left(\dot{\hat{u}}_s(t_0)\right) .$$

Integration over $[0,t)$ we obtain that

$$\mathrm{Diss}_{\Psi}\left(u,[0,t)\right) \geq \int\limits_{[0,t)} \Psi_{\mathrm{rel}}\left(\dot{\hat{u}}_{ac}(\tau)\right)\,\mathrm{d}\tau + \int\limits_{[0,t)} \Psi_{\mathrm{rel}}\left(\dot{\hat{u}}_{s}(\tau)\right)\,\mathrm{d}\|\lambda_s\|(\tau)\,.$$

The other inequality is simple as it holds for all $0 \leq s < t \leq T$ that

$$\begin{aligned}
&\Psi\left(u(t-)-u(s+)\right)\\
&\quad= \sup_{x\in K^*} \langle \lambda((s,t)), x\rangle_{X^*,X}\\
&\quad= \sup_{x\in K^*} \left(\int\limits_{(s,t)} \left\langle \dot{\hat{u}}_{ac}(\tau), x\right\rangle_{X^*,X}\,\mathrm{d}\tau + \int\limits_{(s,t)} \left\langle \dot{\hat{u}}_{s}(\tau), x\right\rangle_{X^*,X}\,\mathrm{d}\|\lambda_s\|(\tau)\right)\\
&\quad\leq \int\limits_{(s,t)} \sup_{x\in K^*} \left\langle \dot{\hat{u}}_{ac}(\tau), x\right\rangle_{X^*,X}\,\mathrm{d}\tau + \int\limits_{(s,t)} \sup_{x\in K^*} \left\langle \dot{\hat{u}}_{s}(\tau), x\right\rangle_{X^*,X}\,\mathrm{d}\|\lambda_s\|(\tau)
\end{aligned}$$

And a similar argument holds for $\Psi(u(t+) - \Psi(u(t-))$. □

Bibliography

[1] Goro Akagi. Doubly nonlinear evolution equations with non-monotone perturbations in reflexive Banach spaces. *J. Evol. Equ.*, 11(1):1–41, 2011.

[2] Goro Akagi and Ulisse Stefanelli. Weighted energy-dissipation functionals for doubly nonlinear evolution. *J. Funct. Anal.*, 260(9):2541–2578, 2011.

[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

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

[5] Georg Aumann. *Reelle Funktionen*. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, Bd LXVIII. Springer-Verlag, Berlin, 1954.

[6] Ferdinando Auricchio, Alexander Mielke, and Ulisse Stefanelli. A rate-independent model for the isothermal quasi-static evolution of shape-memory materials. *Math. Models Methods Appl. Sci.*, 18(1):125–164, 2008.

[7] Claudio Baiocchi and António Capelo. *Variational and quasivariational inequalities*. A Wiley-Interscience Publication. John Wiley & Sons Inc., New York, 1984. Applications to free boundary problems, Translated from the Italian by Lakshmi Jayakar.

[8] E. J. Balder. A general approach to lower semicontinuity and lower closure in optimal control theory. *SIAM J. Control Optim.*, 22(4):570–598, 1984.

[9] Erik J. Balder. Lectures on Young measure theory and its applications in economics. *Rend. Istit. Mat. Univ. Trieste*, 31(suppl. 1):1–69, 2000. Workshop on Measure Theory and Real Analysis (Italian) (Grado, 1997).

[10] Viorel Barbu. *Nonlinear differential equations of monotone types in Banach spaces*. Springer Monographs in Mathematics. Springer, New York, 2010.

[11] Gerald Beer and Jonathan M. Borwein. Mosco convergence and reflexivity. *Proc. Amer. Math. Soc.*, 109(2):427–436, 1990.

[12] Salomon Bochner. Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind. *Fundam. Math.*, 20:262–276, 1933.

[13] Folkmar Bornemann. Normen zur Charakterisierung der schwach*-Konvergenz beschränkter Folgen. *Math. Semesterber.*, 48(2):247–260, 2002.

[14] Blaise Bourdin, Gilles A. Francfort, and Jean-Jacques Marigo. The variational approach to fracture. *J. Elasticity*, 91(1-3):5–148, 2008.

[15] Richard C. Bradley. An elementary treatment of the Radon-Nikodým derivative. *Amer. Math. Monthly*, 96(5):437–440, 1989.

[16] Andrea Braides. *Γ-convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[17] Martin Brokate. Konvexe analysis. Notes to Lectures held at Technische Universität München, 2008.

[18] Martin Brokate, Pavel Krejčí, and Hans Schnabel. On uniqueness in evolution quasivariational inequalities. *J. Convex Anal.*, 11(1):111–130, 2004.

[19] Martin Brokate and Jürgen Sprekels. *Hysteresis and phase transitions*, volume 121 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1996.

[20] C. Castaing and M. Valadier. *Convex analysis and measurable multifunctions*. Lecture Notes in Mathematics, Vol. 580. Springer-Verlag, Berlin, 1977.

[21] P. Colli and A. Visintin. On a class of doubly nonlinear evolution equations. *Comm. Partial Differential Equations*, 15(5):737–756, 1990.

[22] Pierre Cousin. Sur les fonctions de n variables complexes. *Acta Math.*, 19(1):1–61, 1895.

[23] Bernard Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 1989.

[24] Gianni Dal Maso. *An introduction to Γ-convergence*. Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston Inc., Boston, MA, 1993.

[25] Gianni Dal Maso, Antonio DeSimone, and Maria Giovanna Mora. Quasistatic evolution problems for linearly elastic-perfectly plastic materials. *Arch. Ration. Mech. Anal.*, 180(2):237–291, 2006.

[26] Gianni Dal Maso, Gilles A. Francfort, and Rodica Toader. Quasistatic crack growth in nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 176(2):165–225, 2005.

[27] E. N. Dancer and Brailey Sims. Weak-star separability. *Bull. Austral. Math. Soc.*, 20(2):253–257, 1979.

[28] A. de Acosta. Large deviations for vector-valued functionals of a Markov chain: lower bounds. *Ann. Probab.*, 16(3):925–960, 1988.

[29] Ennio De Giorgi. Convergence problems for functionals and operators. In *Proceedings of the International Meeting on Recent Methods in Nonlinear Analysis (Rome, 1978)*, pages 131–188, Bologna, 1979. Pitagora.

[30] Ennio De Giorgi and Tullio Franzoni. Su un tipo di convergenza variazionale. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 58(6):842–850, 1975.

[31] Wolfgang Desch and Janos Turi. The stop operator related to a convex polyhedron. *J. Differential Equations*, 157(2):329–347, 1999.

[32] J. Diestel and J. J. Uhl, Jr. *Vector measures*. American Mathematical Society, Providence, R.I., 1977. With a foreword by B. J. Pettis, Mathematical Surveys, No. 15.

[33] Nelson Dunford and Jacob T. Schwartz. *Linear Operators. I. General Theory*. With the assistance of W. G. Bade and R. G. Bartle. Pure and Applied Mathematics, Vol. 7. Interscience Publishers, Inc., New York, 1958.

[34] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*. North-Holland Publishing Co., Amsterdam, 1976. Translated from the French, Studies in Mathematics and its Applications, Vol. 1.

[35] Herbert Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[36] S. Fitzpatrick. Representing monotone operators by convex functions. In *Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988)*, volume 20 of *Proc. Centre Math. Anal. Austral. Nat. Univ.*, pages 59–65. Austral. Nat. Univ., Canberra, 1988.

[37] Irene Fonseca and Giovanni Leoni. *Modern methods in the calculus of variations: L^p spaces*. Springer Monographs in Mathematics. Springer, New York, 2007.

[38] Nassif Ghoussoub, Abbas Moameni, and Ramón Zárate Sáiz. Homogenization of maximal monotone vector fields via selfdual variational calculus. *Advanced Nonlinear Studies*, 11(2):323–360, 2011.

[39] T. H. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Ann. of Math. (2)*, 20(4):292–296, 1919.

[40] A. L. Gurson. Continuum theory of ductile rupture by void nucleation and growth: Part i - yield criteria and flow rules for porous ductile media. *J. Engrg. Mater. Tech.*, 99:2–15, 1977.

[41] Weimin Han and B. Daya Reddy. *Plasticity*, volume 9 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 1999. Mathematical theory and numerical analysis.

[42] Ralph Henstock. Definitions of Riemann type of the variational integrals. *Proc. London Math. Soc. (3)*, 11:402–418, 1961.

[43] Ralph Henstock. *Theory of integration*. Butterworths, London, 1963.

[44] Ralph Henstock. A Riemann-type integral of Lebesgue power. *Canad. J. Math.*, 20:79–87, 1968.

[45] Ralph Henstock. Generalized integrals of vector-valued functions. *Proc. London Math. Soc. (3)*, 19:509–536, 1969.

[46] T. H. Hildebrandt. The Borel theorem and its generalizations. *Bull. Amer. Math. Soc.*, 32(5):423–474, 1926.

[47] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of convex analysis*. Grundlehren Text Editions. Springer-Verlag, Berlin, 2001. Abridged version of ıt Convex analysis and minimization algorithms. I [Springer, Berlin, 1993; MR1261420 (95m:90001)] and ıt II [ibid.; MR1295240 (95m:90002)].

[48] Claes Johnson. On plasticity with hardening. *J. Math. Anal. Appl.*, 62(2):325–336, 1978.

[49] Dorothee Knees, Alexander Mielke, and Chiara Zanini. On the inviscid limit of a model for crack propagation. *Math. Models Methods Appl. Sci.*, 18(9):1529–1569, 2008.

[50] V. S. Kozjakin, M. A. Krasnosel′skiĭ, and A. V. Pokrovskiĭ. Vibrationally stable hysterons. *Dokl. Akad. Nauk SSSR*, 206:800–803, 1972.

[51] M. A. Krasnosel′skiĭ, B. M. Darinskiĭ, I. V. Emelin, P. P. Zabreĭko, E. A. Lifšic, and A. V. Pokrovskiĭ. An operator-hysterant. *Dokl. Akad. Nauk SSSR*, 190:34–37, 1970.

[52] M. A. Krasnosel′skiĭ and A. V. Pokrovskiĭ. *Sistemy s gisterezisom*. “Nauka”, Moscow, 1983.

[53] P. Krejčí and A. Vladimirov. Lipschitz continuity of polyhedral Skorokhod maps. *Z. Anal. Anwendungen*, 20(4):817–844, 2001.

[54] Pavel Krejčí. *Hysteresis, convexity and dissipation in hyperbolic equations*. GAKUTO International Series. Mathematical Sciences and Applications, 8. Gakkōtosho Co. Ltd., Tokyo, 1996.

[55] Pavel Krejčí. Evolution variational inequalities and multidimensional hysteresis operators. In *Nonlinear differential equations (Chvalatice, 1998)*, volume 404 of *Chapman & Hall/CRC Res. Notes Math.*, pages 47–110. Chapman & Hall/CRC, Boca Raton, FL, 1999.

[56] Pavel Krejčí and Jaroslav Kurzweil. A nonexistence result for the Kurzweil integral. *Math. Bohem.*, 127(4):571–580, 2002.

[57] Pavel Krejčí and Alexander Vladimirov. Polyhedral sweeping processes with oblique reflection in the space of regulated functions. *Set-Valued Anal.*, 11(1):91–110, 2003.

[58] Pavel Krejčí. Regulated evolution quasivariational inequalities. Notes to Lectures held at the University of Pavia, 2003.

[59] Pavel Krejčí. The Kurzweil integral with exclusion of negligible sets. *Math. Bohem.*, 128(3):277–292, 2003.

[60] Pavel Krejčí and Philippe Laurençot. Generalized variational inequalities. *J. Convex Anal.*, 9(1):159–183, 2002.

[61] Pavel Krejčí and Matthias Liero. Rate independent Kurzweil processes. *Applications of Mathematics*, 54(2):117–145, April 2009.

[62] Pavel Krejčí and Thomas Roche. Lipschitz continuity of the sweeping process in BV-spaces. *Discrete Contin. Dyn. Syst. Ser. B*, 15(3):637–650, 2011.

[63] Martin Kružík and Johannes Zimmer. A model of shape memory alloys taking into account plasticity. *IMA J. Appl. Math.*, 76(1):193–216, 2011.

[64] M. Kunze and M. D. P. Monteiro Marques. On parabolic quasi-variational inequalities and state-dependent sweeping processes. *Topol. Methods Nonlinear Anal.*, 12(1):179–191, 1998.

[65] Jaroslav Kurzweil. Generalized ordinary differential equations and continuous dependence on a parameter. *Czechoslovak Math. J.*, 7 (82):418–449, 1957.

[66] Matthias Liero and Alexander Mielke. An evolutionary elastoplastic plate model derived via Γ-convergence. *Math. Models Methods Appl. Sci.*, 21(9):1961–1986, 2011.

[67] M. Marques Alves and B. F. Svaiter. Brøndsted-Rockafellar property and maximality of monotone operators representable by convex functions in non-reflexive Banach spaces. *J. Convex Anal.*, 15(4):693–706, 2008.

[68] I. D. Mayergoyz. *Mathematical models of hysteresis*. Springer-Verlag, New York, 1991.

[69] A. Mielke, R. Rossi, and G. Savare. Nonsmooth analysis of doubly nonlinear evolution equations. available at arxiv:1105.4296v1, 2011.

[70] A. Mielke and S. Zelik. On the vanishing-viscosity limit in parabolic systems with rate-independent dissipation terms. to appear.

[71] Alexander Mielke and Riccarda Rossi. Existence and uniqueness results for a class of rate-independent hysteresis problems. *Math. Models Methods Appl. Sci.*, 17(1):81–123, 2007.

[72] Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. Modeling solutions with jumps for rate-independent systems on metric spaces. *Discrete Contin. Dyn. Syst.*, 25(2):585–615, 2009.

[73] Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. Bv solutions and viscosity approximations of rate-independent systems. *ESAIM, Control Optim. Calc. Var.*, 18(1):36–80, 2012.

[74] Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. Variational convergence of gradient flows and rate-independent evolutions in metric spaces. *Milan J. Math*, 2013. To appear.

[75] Alexander Mielke and Tomáš Roubíček. A rate-independent model for inelastic behavior of shape-memory alloys. *Multiscale Model. Simul.*, 1(4):571–597 (electronic), 2003.

[76] Alexander Mielke and Florian Theil. On rate-independent hysteresis models. *NoDEA Nonlinear Differential Equations Appl.*, 11(2):151–189, 2004.

[77] Alexander Mielke, Florian Theil, and Valery I. Levitas. A variational formulation of rate-independent phase transformations using an extremum principle. *Arch. Ration. Mech. Anal.*, 162(2):137–177, 2002.

[78] Jean-Jacques Moreau. Problème d'évolution associé à un convexe mobile d'un espace hilbertien. *C. R. Acad. Sci. Paris Sér. A-B*, 276:A791–A794, 1973.

[79] Jean Jacques Moreau. Evolution problem associated with a moving convex set in a Hilbert space. *Journal of Differential Equations*, 26(3):347 – 374, 1977.

[80] Umberto Mosco. Convergence of convex sets and of solutions of variational inequalities. *Advances in Math.*, 3:510–585, 1969.

[81] Stefan Müller. Variational models for microstructure and phase transitions. In *Calculus of variations and geometric evolution problems (Cetraro, 1996)*, volume 1713 of *Lecture Notes in Math.*, pages 85–210. Springer, Berlin, 1999.

[82] B. J. Pettis. On integration in vector spaces. *Trans. Amer. Math. Soc.*, 44(2):277–304, 1938.

[83] I. Picek. Multidimensional hysteresis models. Graduate thesis, Charles University, Prague, 1991. In Czech.

[84] L. Prandtl. Ein Gedankenmodell zur kinetischen Theorie der festen Körper. *Z. f. angew. Math.*, 8:85–106, 1928.

[85] Ferenc Preisach. Über die magnetische Nachwirkung. *Z. Phys*, 94:277–302, 1932.

[86] Vincenzo Recupero. On locally isotone rate independent operators. *Applied Mathematics Letters*, 20(11):1156 – 1160, 2007.

[87] Vincenzo Recupero. BV-extension of rate independent operators. *Math. Nachr.*, 282(1):86–98, 2009.

[88] Vincenzo Recupero. Bv solutions of rate independent variational inequalities. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, accepted for publication, 2010.

[89] B. D. Reddy. Existence of solutions to a quasistatic problem in elastoplasticity. In *Progress in partial differential equations: calculus of variations, applications (Pont-à-Mousson, 1991)*, volume 267 of *Pitman Res. Notes Math. Ser.*, pages 299–311. Longman Sci. Tech., Harlow, 1992.

[90] Thomas Roche. Uniqueness of a quasivariational sweeping process on functions of bounded variation. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, accepted for publication, 2010.

[91] R. Tyrrell Rockafellar. Convex functions, monotone operators and variational inequalities. In *Theory and Applications of Monotone Operators (Proc. NATO Advanced Study Inst., Venice, 1968)*, pages 35–65. Edizioni "Oderisi", Gubbio, 1969.

[92] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[93] Riccarda Rossi, Alexander Mielke, and Giuseppe Savaré. A metric approach to a class of doubly nonlinear evolution equations and applications. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 7(1):97–169, 2008.

[94] Riccarda Rossi and Giuseppe Savaré. Gradient flows of non convex functionals in Hilbert spaces and applications. *ESAIM Control Optim. Calc. Var.*, 12(3):564–614 (electronic), 2006.

[95] Riccarda Rossi and Ulisse Stefanelli. An order approach to a class of quasivariational sweeping processes. *Adv. Differential Equations*, 10(5):527–552, 2005.

[96] Tomáš Roubíček. Thermodynamics of rate-independent processes in viscous solids at small strains. *SIAM J. Math. Anal.*, 42(1):256–297, 2010.

[97] Bernd Schmidt. Partielle Differentialgleichungen 2. Lecture Notes, Technische Universität München, 2009.

[98] H. Schnabel. *Zur Wohlgestelltheit des Gurson-Modells*. PhD thesis, Technische Universität München, 2006.

[99] Antonio Segatti. Global attractor for a class of doubly nonlinear abstract evolution equations. *Discrete Contin. Dyn. Syst.*, 14(4):801–820, 2006.

[100] Ulisse Stefanelli. Nonlocal quasivariational evolution problems. *J. Differential Equations*, 229(1):204–228, 2006.

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

[102] Ulisse Stefanelli. A variational characterization of rate-independent evolution. *Math. Nachr.*, 282(11):1492–1512, 2009.

[103] Marita Thomas and Alexander Mielke. Damage of nonlinearly elastic materials at small strain—existence and regularity results. *ZAMM Z. Angew. Math. Mech.*, 90(2):88–112, 2010.

[104] Michel Valadier. Young measures. In *Methods of nonconvex analysis (Varenna, 1989)*, volume 1446 of *Lecture Notes in Math.*, pages 152–188. Springer, Berlin, 1990.

[105] Augusto Visintin. *Differential models of hysteresis*, volume 111 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 1994.

[106] L. C. Young. On approximation by polygons in the calculus of variations. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 141(844):pp. 325–341, 1933.

[107] L. C. Young. Generalized surfaces in the calculus of variations. *Ann. of Math. (2)*, 43:84–103, 1942.

[108] L. C. Young. Generalized surfaces in the calculus of variations. II. *Ann. of Math. (2)*, 43:530–544, 1942.