Technische Universität München

Fakultät Mathematik

Lehrstuhl M6 – Mathematische Modellbildung

Viscosity Solutions for Control Problems with Hysteresis

Matthias Michael Buchinger

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: Prof. Dr. Marco Cicalese

Prüfer der Dissertation: 1. Prof. Dr. Martin Brokate

2. Prof. Dr. Piermarco Cannarsa

Die Dissertation wurde am 11.12.2017 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 27.03.2018 angenommen.

Abstract

This thesis deals with optimal control problems ("OCP") concerning differential equations with hysteresis, in particular, with application of the dynamic programming method ("DPM") in the case of partial differential equations with hysteresis. The first main part of the thesis focusses on the semilinear case of a heat equation with pointwise applied Play operator. As the solution operator of that equation is hard to handle, we consider different simplifications. First, we replace the Play operator with some regularization, which yields an OCP that can be treated with known methods. Subsequently, we reduce the number of components of the hysteresis variable to finitely many, by consideration of suitable averiges over subsets of the domain. This leads to an abstract problem with hysteresis that can be treated with adapted methods of known results from DPM theory. Further, we show that the considered averiges form a good approximation of the original problem by proving some convergence result. After that, we go into the dicrete DPM by means of some time discretization. The last two sections deal with problems including the time derivative of a Play type hysteresis. We first focus on the treatment of the hysteresis variable and discuss this with the help of some ode model problem. After that, we use this insight to treat a problem belonging to some quasilinear pde with hysteresis.

Acknowledgment

First of all, I want to thank my advisor Professor Brokate for giving me the opportunity to work on this thesis, and for the support and freedom during its preparation. Further, I want to thank Dr. Carl-Friedrich Kreiner, who helped me a lot during the time when I was just a guest at the chair.

I also want to mention the other members of M6, who made the bad times easier and the good ones even happier. A special thanks to Patrick – who had to share an office with me for several years – for all the mathematical, non-mathematical and sometimes also absurd discussions.

In this connection, I also want to thank everyone of the “coffee and canteen groups” for the funny time.

Last but not least, a special thanks goes to my family, especially my parents, for their support over all the years.

Table of Contents

PART ONE

1 Optimal control theory **2**
1.1 Some examples of control problems 2
1.2 The dynamic programming method 3

2 Hysteresis effects and operators **5**
2.1 Hysteresis phenomena 5
2.2 Some examples of hysteresis operators 5
2.3 Properties of the play operator 8
2.4 Approximation of the play operator 11
2.5 A general approximation result 16
2.6 Pointwise applied play operators and weak differentiability 24
2.7 A representation result for the one dimensional stop operator via projections in a Hilbert space 29

PART TWO

3 The heat equation with hysteresis **38**
3.1 Weak solutions and regularity results 38
3.2 Heat equation with regularized play operator 47
3.3 A control problem corresponding to the dynamics of the regularized partial differential equation 54
3.4 Dynamic programming method - Application of known results 56
3.5 Convergence of optimal value functions 58

4 An Abstract model and general assumptions **64**
4.1 Special Semigroup 64
4.2 Dynamic and further assumptions 65

5 The control problem **72**

6 Dynamic programming and the value function as viscosity solution **75**
6.1 The dynamic programming principle 75
6.2 The HJB inclusion and existence of solutions 76

7 Comparison result and uniqueness of viscosity solutions **85**

8 A convergence theorem **96**
8.1 The function space $L^2(\Omega; H^\theta(0,T))$ 96
8.2 Further properties of solutions and the corresponding solution operator 100
8.3 Convergence of solutions of problems (P), (P') 112
8.4 Convergence of optimal value functions for problem (P') 123

9 Time discrete dynamic programming and approximative optimal feedback controls **127**

PART THREE

---

10 A model problem corresponding to an ordinary differential equation including the time derivative of a hysteresis nonlinearity **144**
10.1 Properties of the o.d.e. 144
10.2 The control problem and Dynamic Programming - adaption of the method of [1] 147

11 Dynamic Programming for a problem corresponding to a quasilinear p.d.e. with Play type hysteresis **153**
11.1 The differential equation and properties of solutions 153
11.2 The control problem and properties of the value function 154
11.3 The HJB equation and existence of solutions 157
11.4 Comparison and uniqueness result 165

APPENDIX

A Note on Equations with hysteresis **174**

B Rate of convergence for the regularized equation **180**

C A regularized Hilpert-type inequality **187**

D Some frequently used Theorems **189**

References **191**

List of Problem Definitions and Assumptions **194**

List of Figures **194**

PART ONE

1 Optimal control theory

The goal of optimal control theory is to control some process with minimal effort. Further, it is realistic that there are restrictions on the set of controls, for example due to technical realizability. As we will restrict to mathematical analysis rather than modelling of real systems later, we also restrict ourselfes to some mathematical aspects of the theory in this introduction.

1.1 Some examples of control problems

Let us, as an introduction, consider a dynamical system of the form

$$\dot{y} = f(y, \alpha), \qquad y(0) = y_0,$$

where the control variable α is, for each $t > 0$, allowed to take values in a set $\mathbb{A} \subset \mathbb{R}$, and $y(t) \in \mathbb{R}$, $f : \mathbb{R}^2 \to \mathbb{R}$. Thus, for fixed initial value, the solution would depend only on the control, which will be indicated by y_α. A typical functional would be

$$J(y, \alpha) := \int_0^T (y_\alpha(t))^2 + (\alpha(t))^2 \, dt,$$

meaning that by looking for some minimal value of J, we would aim for some efficient control $\bar{\alpha}$ for which $y_{\bar{\alpha}}$ is kept small, in some sense. As we only consider trajectories on some interval $[0, T]$ of bounded length, a control problem corresponding to such type of functional is called a finite horizon problem. There are also problems with other types of functionals, such as

$$J^*(y, \alpha) := \int_0^\infty e^{-\lambda t} \left((y_\alpha(t))^2 + (\alpha(t))^2 \right) dt.$$

Here, $\lambda \geq 0$ is a so called discount factor. Problems with such types of functionals are called infinite horizon problems, as the whole trajectory has to be considered in the minimization process. Corresponding to the optimal control $\bar{\alpha}$ (if it exists), there is the optimal value of the functional $J(y_{\bar{\alpha}}, \bar{\alpha})$ (resp. $J^*(y_{\bar{\alpha}}, \bar{\alpha})$). It may vary with the initial value, so that we can define the so called **(optimal) value function** via

$$V(y_0) := J(y_{y_0,\bar{\alpha}}, \bar{\alpha}).$$

The dynamic programming method, introduced in the next section, has its focus on how to characterize the optimal value function, which then, in a way, solves simultaneously multiple optimal control problems. Of course, much more complex dynamical systems can be considered in optimal control theory, such as optimal heating processes, optimal chemical synthesis, population dynamics etc., see [2], [3] and many others.

1.2 The dynamic programming method

For this introduction, we will consider our simple system from section 1.1. Assume that for every initial value y_0, there exists an optimal control $\bar{\alpha}$. Then, by definition,

$$V(y_0) = \int_0^\infty e^{-\lambda t} \left((y_{y_0,\bar{\alpha}}(t))^2 + (\bar{\alpha}(t))^2 \right) dt.$$

Since y meets a semigroup property of the form

$$y_{y_{y_0,\bar{\alpha}}(t),\bar{\alpha}(t+\cdot)}(s) = y_{y_0,\bar{\alpha}}(t+s), \qquad t, s > 0,$$

we can derive

$$\begin{aligned} V(y_0) &= \int_0^T e^{-\lambda t} \left((y_{y_0,\bar{\alpha}}(t))^2 + (\bar{\alpha}(t))^2 \right) dt + e^{-\lambda T} V(y_{y_0,\bar{\alpha}}(T)) \\ &= \inf_\alpha \left\{ \int_0^T e^{-\lambda t} \left((y_{y_0,\alpha}(t))^2 + (\alpha(t))^2 \right) dt + e^{-\lambda T} V(y_{y_0,\alpha}(T)) \right\}, \end{aligned}$$

the so called **dynamic programming principle (DPP)** (sometimes also refered to as “Bellman’s principle of optimality”). If V is differentiable at y_0, this implies that

$$\lambda V(y_0) + \sup_{a \in \mathbb{A}} \left\{ -f(y_0, a) \cdot DV(y_0) - y_0^2 - a^2 \right\} = 0,$$

which is a so called **Hamilton Jacobi Bellman equation (HJB equation)**. Thus, if V is everywhere differentiable, this uniquely determines the optimal value function of the control problem. Unfortunately, even in the most simple case of smooth data, we can not expect that V has such regularity; in fact, there are many examples where this is not the case, see e.g. [2, chapter 6, example 1.5] for a finite horizon problem where the value function is not everywhere differentiable. Hence, one has to look for different types of solution concepts. It turned out that viscosity solutions do a great job here (cf. e.g. [4]). However, when dealing with control systems including partial differential equations, further problems occur when one wants to apply the dynamic programming method. This is due to the appearence of some unbounded operator (e.g., some differential operator A of elliptic type). There are also different approaches how to deal with those problems. In their pioneering works, Crandall and Lions [5] used the notion of B-continuity, i.e., they introduced a compact operator B, such that A^*B is bounded and linear, which can then be used to define viscosity solutions for HJB equations in infinite dimensions. This method is picked up by [2], which gave a quite general theorem for the characterization of optimal value functions for control problems in infinite dimensions. But there are also different approaches in the infinite dimensional setting; e.g., [6] introduced a definition for viscosity solutions based on the regularity properties of the value function. We will follow this approach (but adapt it to an infinite horizon problem including hysteresis) in sections 6 and 7. Further, in section 11, we will adapt the method of [7] to handle a control problem associated with some quasilinear p.d.e. with hysteresis.

2 Hysteresis effects and operators

In this section we briefly discuss some aspects of hysteresis, and give examples where such phenomena may appear. Then we turn to the mathematical viewpoint and show how some of the basic hysteresis operators may be described. In more detail, we will discuss the Play and Stop operators, which are important examples, and whose properties will frequently be used in later sections.

2.1 Hysteresis phenomena

A typical example where hysteresis appears is ferromagnetism (cf. [8]). Some materials, such as iron, can be magnetized if an external field is applied. As the power of the external field is raised, more and more of the atoms will abrupt adjust to this, until saturation effects appear. If the external field is then removed, the magnetization of the material will be weakened; but, not all of the atoms will turn back into their original state, so that the material remains magnetized unless an external field of opposite polarisation is applied. This effect of "lacking behind", described by the following figure, is the nature of hysteresis.

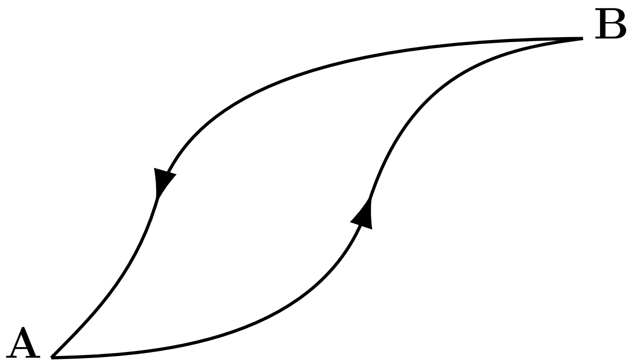

Figure 1: Hysteresis loop

As described, in this example the hysteresis loop originates by the sum of small elements, so called hysterons, which jump immediately from one level to another when certain internal values raise, resp., fall beyond some characteristic point. In this way, superpositions of simple hysteresis operators can create more complex ones. This effect also explains the existence of hysteresis loops in systems describing processes in biology (cf. [9]) or economic sciences (cf. [10]).

2.2 Some examples of hysteresis operators

One example of "simple" hysteresis operators are the above mentioned **hysterons**, also called **Relays**. Assume that at time t_0, the output value is zero, and that this is kept until the input variable $x(t)$ becomes larger than some value a_1. Then, the output jumps from zero to one. However, if the input is decreased again, the output will not change until $x < a_2$ for some $a_2 < a_1$. If it becomes smaller than a_2, it jumps back to zero. This definition (which is somewhat problematic, because there is not really a "best choice" how to define when the output should jump) produces an input output diagram of the following form:

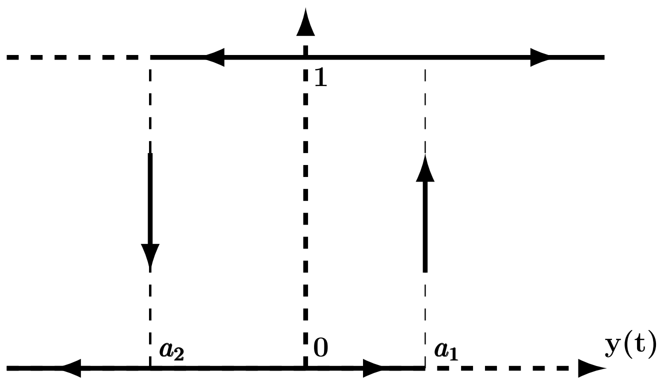

Figure 2: Input-output relation of Relay operator

The next example is the so called **Play Operator**, which has its name from the idea that one considers the play between two mechanical elements.

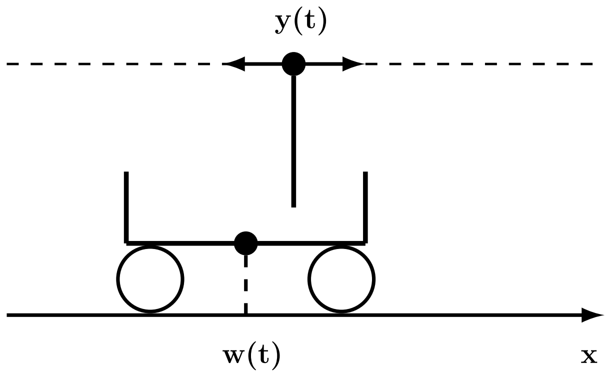

Figure 3: Play between mechanical elements

As long as element one (input "y(t)") moves in the interior of element two (output "w(t)"), this will not change the output. This will only happen, if it touches the boundary. So, if element two has diameter $2r$, then an input-output diagram of the form

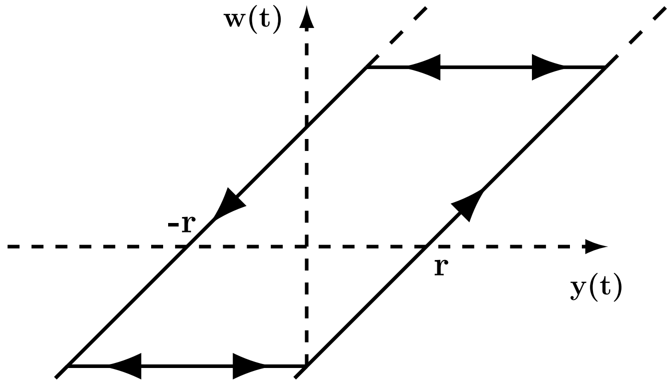

Figure 4: Input-output relation of Play operator

is produced. Closely related to the Play is the so called **Stop operator**. If the second element is fixed, then element one can only move inside this interval of length $2r$, even if some force is applied. So, the output "e(t)" is stopped at the boundary, producing a diagram of the form

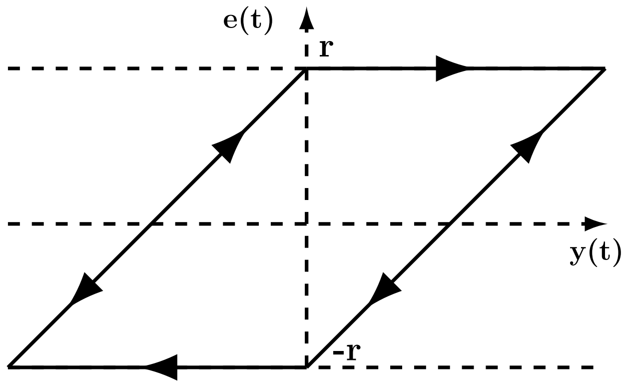

Figure 5: Input-output relation of Stop operator

Intuitively, we extract from the stop just the movement of the input while it is in the interior of element two, whereas the play only reacts if the boundary is moved; so we should have that the outputs sum up to the input function. This property is shown in the next section. Further, each of those elementary hysteresis operators can be used as building blocks for more complex ones; e.g., one might think of superpositions of play operators with different values of r. At this point we note that there are also various different types of hysteresis operators such as generalised Plays, see e.g. [11].

2.3 Properties of the play operator

We collect some properties of the play and stop operators; those (elementary) results are taken from [12, chapter 2.3]. From the interpretation as mechanical play, it is clear that the extreme points of input functions play an important role for the output. We therefore start with defining the operator on strings, and then extend this to continuous functions.

Definition 2.1 *The function* $f_r : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$*,*

$$f_r(v, w) := \max\{v - r, \min\{v + r, w\}\},$$

where $r \geq 0$*, is called* ***update function*** *(of the play operator).*

□

The following simple lemma (cf. [12, lemma 2.3.1]) is useful.

Lemma 2.2 *The update functions satisfy, for any* $r_j \geq 0$*,* $v_j, w_j \in \mathbb{R}$*,* $j = 1, 2$*, the inequality*

$$|f_{r_1}(v_1, w_1) - f_{r_2}(v_2, w_2)| \leq \max\{|v_1 - v_2| + |r_1 - r_2|, |w_1 - w_2|\}.$$

Proof: For $a, b, c, d \in \mathbb{R}$, it is easy to see that

$$|\max\{a, b\} - \max\{c, d\}| \leq \max\{|a - c|, |b - d|\},$$

and the inequality holds analogously, if max is replaced by min on the left hand side. Hence,

$$\begin{aligned} &|f_{r_1}(v_1, w_1) - f_{r_2}(v_2, w_2)| \\ &\qquad \leq \max\{|(v_1 - r_1) - (v_2 - r_2)|, |(v_1 + r_1) - (v_2 + r_2)|, |w_1 - w_2|\}, \end{aligned}$$

which implies the assertion.

□

To sufficiently describe the output on the set S of all strings, we introduce the notion of final value mappings.

Definition 2.3 *Let* $w_{-1} \in \mathbb{R}$ *be the* ***internal configuration before the operator is applied****. Define* $w_0 := f_r(s_0, w_{-1})$*, and iteratively,* $w_{k+1} := f(s_{k+1}, w_k)$*. The string* w *is the output of the play operator applied to* $s \in S$*. The* ***final value mapping*** *corresponding to the play operator on the set of strings is then defined through*

$$\mathcal{F}_{r,f}(s) := w_n, \qquad s = (s_0, \dots, s_n).$$

□

Next assume that $v \in C_{pm}[0,T]$, the set of all continuous functions on the interval $[0,T]$ for which there exists a decomposition $t_0 < \dots < t_n$ of $[0,T]$ such that v is monotonic on each subinterval (t_{i-1}, t_i). Then we can define, for every $t \in [0,T]$, the output $\mathcal{F}_r[v; w_{-1}](t)$ of the play through application of its discrete version to the string consisting of all extreme values of v on $[0,t]$. Hence, for such functions, the inequality in lemma 2.2 takes the form

$$\begin{aligned} &|\mathcal{F}_{r_1}[v_1; w_{-1,1}](t) - \mathcal{F}_{r_2}[v_2; w_{-1,2}](t)| \\ &\qquad \le \max\left\{ |r_1 - r_2| + \sup_{0\le\tau\le t} |v_1(\tau) - v_2(\tau)|, |w_{-1,1} - w_{-1,2}| \right\}. \end{aligned}$$

Using a density argument, we can thus define the play operator for continuous inputs. It has the following properties.

Theorem 2.4 *For any $r \ge 0$, the operator $\mathcal{F}_r$ can be extended uniquely to a Lipschitz continuous operator $\mathcal{F}_r : C[0,T] \times \mathbb{R} \to C[0,T]$, and it holds, for all $v, v_1, v_2 \in C[0,T]$, $w_{-1}, w_{-1,1}, w_{-1,2}, y_{-1} \in \mathbb{R}$, for all $s \ge 0$ and $0 \le t' < t \le T$,*

$$\begin{aligned} \|\mathcal{F}_r[v_1; w_{-1,1}] - \mathcal{F}_r[v_2; w_{-1,2}]\|_{C[0,T]} &\le \max\left\{ \|v_1 - v_2\|_{C[0,T]}, |w_{-1,1} - w_{-1,2}| \right\}, \\ |\mathcal{F}_r[v, w_{-1}](t) - \mathcal{F}_r[v, w_{-1}](t')| &\le \sup_{t'\le\tau\le t} |v(\tau) - v(t')|, \\ \mathcal{F}_r[v; w_{-1}] &= \mathcal{F}_r[v - w_{-1}; 0] + w_{-1}, \\ \mathcal{F}_r[v_1; w_{-1,1}] &\le \mathcal{F}_r[v_2; w_{-1,2}], \quad \text{if } v_1 \le v_2 \text{ and } w_{-1,1} \le w_{-1,2}, \\ \mathcal{F}_r[\mathcal{F}_s[v; y_{-1}]; w_{-1}] &= \mathcal{F}_{r+s}[v; w_{-1}], \quad \text{if } |y_{-1} - w_{-1}| \le r. \end{aligned}$$

Further, it meets the semigroup property

$$\mathcal{F}_r[v(t + \cdot); \mathcal{F}_r[v; w_{-1}](t)](\tau) = \mathcal{F}_r[v; w_{-1}](t + \tau), \qquad \forall t, \tau > 0.$$

Proof: The semigroup property for continuous and piecewise monotone functions follows immediately from the definition, so that the general case may be shown using a density argument. For the other statements, we refer to [12, theorem 2.3.2].

□

We will need the connection between play and stop operators established in the following result.

Theorem 2.5 *Let $e_r : \mathbb{R} \to \mathbb{R}$, be the update function of the stop operator, which is defined through*

$$e_r(v) := \min\{r, \max\{-r, v\}\}.$$

Then, for any $w_{-1} \in \mathbb{R}$, the mapping $\mathcal{E}_{r,f} : S \to \mathbb{R}$,

$$\mathcal{E}_{r,f}(v_0) = e_r(v_0 - w_{-1}),$$
$$\mathcal{E}_{r,f}(v_0, \ldots, v_n) = e_r(v_N - v_{N-1} + \mathcal{E}_{r,f}(v_0, \ldots, v_{N-1})),$$

defines a hysteresis operator $\mathcal{E}_r[\,\cdot\,; w_{-1}]$. It extends to a Lipschitz continuous operator $\mathcal{E}_r : C[0,T] \times \mathbb{R} \to C[0,T]$ and satisfies, for all functions $v, v_1, v_2 \in C[0,T]$, initial internal values $w_{-1}, w_{-1,1}, w_{-1,2} \in \mathbb{R}$, and $0 \le t' < t \le T$, (where $\mathcal{E}_r[v] := \mathcal{E}_r[v; 0]$),

$$\|\mathcal{E}_r[v_1] - \mathcal{E}_r[v_2]\|_{C[0,T]} \le 2\,\|v_1 - v_2\|_{C[0,T]},$$
$$|\mathcal{E}_r[v](t) - \mathcal{E}_r[v](t')| \le 2 \sup_{t' \le \tau \le t} |v(\tau) - v(t')|,$$
$$\mathcal{E}_r[v; w_{-1}] + \mathcal{F}_r[v; w_{-1}] = v,$$
$$\mathcal{E}_r[v; w_{-1}] = \mathcal{E}_r[v - w_{-1}],$$
$$\mathcal{E}_r[v_1; w_{-1,1}] \le \mathcal{E}_r[v_2; w_{-1,2}], \text{ if } v_1 \le v_2 \text{ and } w_{-1,1} \le w_{-1,2}.$$

Further, it meets the semigroup property

$$\mathcal{E}_r[v(t+\cdot); \mathcal{E}_r[v; w_{-1}](t)](t) = \mathcal{E}_r[v; w_{-1}](t+\tau) \qquad \forall t, \tau > 0.$$

Proof: The semigroup property follows again by some density argument; for the other statements, we refer to [12, proposition 2.3.4].

□

We finish this section with a result corresponding to differentiability properties of the stop and play operators.

Proposition 2.6 *Let $r \ge 0$, $v \in W^{1,1}(0,T)$ be given, and define the sets*

$$A_\pm := \{t \in [0,T] : \mathcal{E}_r[v](t) = \pm r\},\ A_0 := \{t \in [0,T] : |\mathcal{E}_r[v](t)| < r\}.$$

Then,

$$\mathcal{E}_r[v]' = 0, \quad \mathcal{F}_r[v]' = v' \ge\ (\le)\ 0, \text{ a.e. on } A_+\ (A_-),$$
$$\mathcal{E}_r[v]' = v', \quad \mathcal{F}_r[v]' = 0 \text{ a.e. on } A_0.$$

In particular, a.e. on $(0,T)$,

$$|\mathcal{F}_r[v]'| \le |v'|, \qquad |\mathcal{E}_r[v]'| \le |v'|.$$

Further, the property of piecewise monotonicity (i.e., the final value mapping $\mathcal{W}_f$ fulfills $v_N \ge v_{N-1} \Rightarrow \mathcal{W}_f(v_0, \ldots, v_{N-1}) \le \mathcal{W}_f(v_0, \ldots, v_N)$) attains the form

$$v'\mathcal{F}_r[v]' \ge 0, \qquad v'\mathcal{E}_r[v]' \ge 0, \text{ a.e. on } (0,T).$$

Proof: See [12, Lemma 2.3.8].

□

The latter proposition also gives rise to the representation of the Play operator via differential inclusions, cf. [11] for more details.

2.4 Approximation of the play operator

In this and the following section, we derive regularisation results for the Play operator. We note that a similar result for the Stop operator is given in [13] via slightly different methods. For $\varepsilon > 0$, we investigate the solution of the ordinary differential equation (which is proposed in [14])

$$\begin{aligned} \varepsilon \dot{z} &= G(z-u), \\ z(0) &= z_0, \end{aligned} \tag{2.1}$$

where the function $G : \mathbb{R} \to \mathbb{R}$ is defined by

$$G(x) := -(-x+r)_- + (-x-r)_+. \tag{2.2}$$

Here, $r > 0$, and $(y)_- := -\min(0, y)$ denotes the negative part, $(y)_+ := \max(0, y)$ the positive part of $y \in \mathbb{R}$.

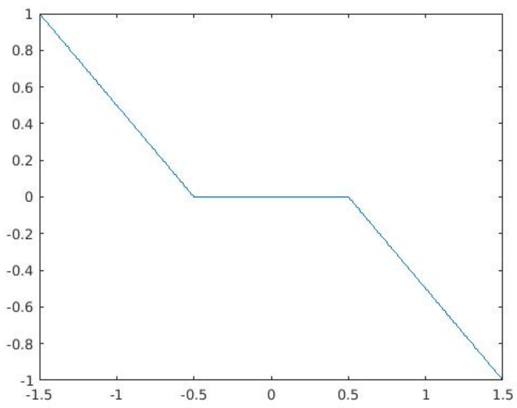

Figure 6: graph of G for r=0.5

Though the solution depends on the value of the parameter ε, we will most of the time drop the index ε to simplify notations. We make the following general assumptions.

Assumption 2.7

(V1) $u^\varepsilon \to u^\infty$ *(strongly) in* $C[0,T]$ *as* $\varepsilon \downarrow 0$,

(V2) u^ε *is bounded in* $H^1(0,T)$ *as* $\varepsilon \downarrow 0$,

(V3) z_0, u_0 *do not depend on* $\varepsilon > 0$, *and* $|u_0 - z_0| \le r$.

□

Our goal is to show that in the limit $\varepsilon \downarrow 0$, the sequence of solutions z^ε converges to the output $\mathcal{F}_r[u^\infty; z_0]$ of the play operator. To this end, we first need to show that there exists a unique solution to (2.1).

Lemma 2.8 *Let assumption 2.7 hold. Then there exists a unique solution to* (2.1) *which is an element of* $H^2(0,T)$ *and bounded w.r.t.* $\|\cdot\|_{C[0,T]}$ *as* $\varepsilon \downarrow 0$.

Proof: The function $s \mapsto G(s)$ is piecewise linear with derivative not larger than one. Hence,

$$\left|\frac{1}{\varepsilon}G(x_1 - y) - \frac{1}{\varepsilon}G(x_2 - y)\right| \le \frac{1}{\varepsilon}\left|x_1 - x_2\right|,\ \forall x_1, x_2, y \in \mathbb{R}.$$

Since $u \in C[0,T]$, there is some constant $c > 0$, such that (locally in $x \in B_\delta(z_0)$),

$$\left|\frac{1}{\varepsilon}G(x-u)\right| \le c,$$

and c depends only on ε, r, δ and T. By continuity, all functions appearing are also measurable, and we can use theorem 1.4.3 and Corollary 1.4.4 of [12] to conclude that there exists a unique local solution on $(0,T)$ in the sense of Carathéodory. Since z and u are continuous as well as the function G, the solution is actually classical, with derivative $\dot z \in H^1(0,T)$ (as $u \in H^1(0,T)$), and may be continued to the whole interval $(0,T)$. Next, note that $G(x)x \le 0$ for all $x \in \mathbb{R}$. Thus, if we multiply (2.1) by $(z-u)$ and then integrate, we get

$$\int_0^\tau \varepsilon \dot z (z-u)ds = \int_0^\tau G(z-u)(z-u)ds \le 0 \implies \int_0^\tau \dot z z ds \le \int_0^\tau \dot z u ds.$$

Using $\dot z z = \frac{1}{2}\frac{d}{dt}(z^2)$, partial integration yields

$$\frac{1}{2}(z(\tau)^2 - z_0^2) \le z(\tau)u(\tau) - z_0 u_0 - \int_0^\tau z(s)\dot u(s)ds.$$

We can estimate further by application of absolute value to the right hand side, and Young's inequality (lemma D.1, with $p = q = 2$, $\delta = \sqrt{2}$), to get

$$\frac{1}{2}(z(\tau)^2 - z_0^2) \leq \frac{1}{4}z(\tau)^2 + u(\tau)^2 + z_0 u_0 + \frac{1}{4}\int_0^\tau z(s)^2 ds + \int_0^\tau \dot{u}(s)^2 ds.$$

Introducing

$$C := 2z_0^2 + 4\|u\|_{C[0,T]}^2 + 4|z_0 u_0| + 4\|u\|_{H^1(0,T)},$$

we may thus write

$$z(\tau)^2 \leq C + \int_0^\tau z(s)^2 ds.$$

Thanks to Gronwall's inequality (theorem D.3), this implies

$$z(\tau)^2 \leq Ce^\tau \leq Ce^T, \ \forall \tau \in [0,T]. \tag{2.3}$$

Noting that assumption 2.7 implies that C is bounded for $\varepsilon \downarrow 0$, we infer from (2.3) that $\|z^\varepsilon\|_{C[0,T]}$ is bounded as $\varepsilon \downarrow 0$, which concludes the proof.

□

Lemma 2.9 *Under assumption 2.7, there exists $C > 0$ independent of $\varepsilon > 0$, such that*

$$0 \leq \frac{1}{\varepsilon^2}\int_0^T G(z^\varepsilon(s) - u^\varepsilon(s))^2 ds \leq C.$$

Proof: Define the auxiliary function

$$V(t) := \frac{1}{2}\dot{z}(t)^2.$$

By lemma 2.8, we may apply the chain rule, in order to calculate its weak derivative. Hence,

$$\dot{V} = \dot{z}\ddot{z} = \frac{1}{\varepsilon}G(z-u)\frac{1}{\varepsilon}G'(z-u)(\dot{z} - \dot{u}).$$

According to its definition, $G(x) \neq 0 \Rightarrow G'(x) = -1$, so that we may infer from the last equality,

$$\dot{V} = -\frac{1}{\varepsilon^2}G(z-u)\left(\frac{1}{\varepsilon}G(z-u) - \dot{u}\right)$$

$$= -\frac{1}{\varepsilon^3}G(z-u)^2 + \frac{1}{\varepsilon^2}G(z-u)\dot{u}, \text{ a.e.} \tag{2.4}$$

Due to the regularity of solutions proven in lemma 2.8 and assumption 2.7, it holds $V(0) = \frac{1}{2}(G(z_0 - u_0))^2 = 0$. Integration of (2.4) now yields

$$\begin{aligned} 0 \le V(t) = \int_0^t \dot{V}(s)ds = & -\frac{1}{\varepsilon^3}\int_0^t G(z(s)-u(s))^2 ds \\ & + \frac{1}{\varepsilon^2}\int_0^t G(z(s)-u(s))\dot{u}(s)ds \\ \Rightarrow \int_0^t G(z(s)-u(s))^2 ds \le \varepsilon & \int_0^t G(z(s)-u(s))\dot{u}(s)ds. \end{aligned}$$

Since the right hand side is smaller than $\varepsilon \|G(z-u)\|_{L^2(0,t)} \|\dot{u}\|_{L^2(0,t)}$, we conclude that

$$0 \le \|G(z-u)\|_{L^2(0,t)} \le \varepsilon \|\dot{u}\|_{L^2(0,t)},$$

for each $0 < t \le T$, which implies the assertion.

□

We turn now to some problem that uniquely determines the output of the play operator (see also [15, 16, 17]).

Problem ($\mathbf{P_0}$): For given $r, T > 0$, $u \in C[0,T]$, $a_0 \in [-r, r]$, find $\xi \in CBV[0,T]$, such that:

1. $u(t) - \xi(t) \in [-r, r]$, for every $t \in [0,T]$,
2. $u(0) - \xi(0) = a_0$,
3. $\int_0^T (u(s) - \xi(s) - y(s))d\xi(s) \ge 0$, for all $y \in C[0,T]$, with $\|y\|_{C[0,T]} \le r$.

□

The integral has to be understood in the sense of Riemann-Stieltjes.

Theorem 2.10 *The unique solution to* ($\mathbf{P_0}$) *is* $\xi = \mathcal{F}_r[u; \xi(0)]$.

Proof: Let us show that $\xi = \mathcal{F}_r[u; \xi(0)]$ is a solution to ($\mathbf{P_0}$). To this end, we only need to check whether the third property is fulfilled, as the others are obviously valid. We first show that 3. is equivalent to

$$\begin{gathered} \forall [a,b] \subset [0,T],\ \forall y \in C[0,T], \|y\|_{C[0,T]} \le r : \\ \int_a^b (u(s) - \xi(s) - y(s))d\xi(s) \ge 0. \end{gathered} \tag{2.5}$$

Proof of the assertion: Assume 3. holds. Choose an admissible sequence $(y_n)_n$ which fulfills

$$y_n(t) = \begin{cases} y(t), & t \in [a,b] \\ u(t) - \xi(t), & t \in [0,T] \setminus [a - \frac{1}{n}, b + \frac{1}{n}] \end{cases} .$$

Then

$$0 \le \int_0^T (u(s) - \xi(s) - y_n(s))d\xi(s) = \int_a^b (u(s) - \xi(s) - y(s))d\xi(s) + \delta(n),$$

where $\delta(n) \to 0$ as $n \to \infty$; so by taking the limit, this yields

$$\int_a^b (u(s) - \xi(s) - y(s))d\xi(s) \ge 0,$$

as claimed. The converse implication follows by simply choosing $a = 0$ and $b = T$. Now we use (2.5) to prove the theorem. If $|u - \mathcal{F}_r[u;\xi(0)]| < r$ on some interval $[a,b] \subset [0,T]$, then $\mathcal{F}_r[u;\xi(0)]$ is constant on that interval, and (2.5) equals zero. Else, if $|u - \mathcal{F}_r[u;\xi(0)]| = r$ on some interval $[c,d]$, then $\mathcal{F}_r[u;\xi(0)]$ must be monotonic on $[c,d]$, as well as u. But then, (2.5) holds, because both $(u(s) - \xi(s) - y(s))$ and $d\mathcal{F}_r[u;\xi(0)]$ must have the same sign. By continuity, (2.5) must be valid.

Now, assume that ξ solves $(\mathbf{P_0})$. If $|u(s) - \xi(s)| < r$, by continuity of u, ξ, there must be some interval $[a,b] \subset [0,T]$ such that $|u - \xi| < r$ on $[a,b]$. Choosing $y = \pm r$ on $[a,b]$, it follows that ξ must be constant on this interval. On the other hand, if $|u(s) - \xi(s)| = r$ on some interval $[c,d]$, choosing $y = 0$, it follows that $d\xi$ must have the same sign as $u - \xi$ on that interval, because we would get a contradiction if this was not the case for some subset of $[c,d]$. By continuity of ξ and $\mathcal{F}_r[u;\xi(0)]$, this implies that both functions coincide.

□

Now we can use the characterization of the play operator in terms of problem $(\mathbf{P_0})$ to prove convergence of z^ε.

Theorem 2.11 *Let assumption 2.7 hold. Then* $z^\varepsilon \to z^\infty = \mathcal{F}_r[u^\infty;\xi(0)]$ *(strongly) in* $C[0,T]$ *and weakly in* $H^1(0,T)$.

Proof: According to lemmata 2.8 and 2.9, for each sequence $(\varepsilon_n)_n$ such that $\varepsilon_n \downarrow 0$ as $n \to \infty$, $(z^{\varepsilon_n})_n$ is bounded in $H^1(0,T)$, so that we can extract a subsequence (for simplicity also denoted by z^{ε_n}) which converges both weakly in $H^1(0,T)$ and strongly in $C[0,T]$ to some z^∞ (as $H^1(0,T) \hookrightarrow C[0,T]$ compact, which is often referred to as Morrey's theorem, cf., e.g., [18, theorem 6.3, pt. II]). We show that z^∞ solves $(\mathbf{P_0})$. Assumption $(V3)$ implies 2. in $(\mathbf{P_0})$, and according to lemma 2.9,

$$\int_0^T G(z^{\varepsilon_n}(s) - u^{\varepsilon_n}(s))^2 ds \downarrow 0, \text{ as } n \to \infty,$$

hence, $z^\infty - u^\infty \in [-r, r]$, by definition of G. Let now $y \in C[0,T]$ be any function such that $\|y\|_{C[0,T]} \leq r$. Then, since $z^n := z^{\varepsilon_n} \in H^1(0,T)$ (denoting $u^n := u^{\varepsilon_n}$):

$$I_n := \int_0^T (u^n(s) - z^n(s) - y(s))dz^n(s) = \int_0^T (u^n(s) - z^n(s) - y(s))\dot{z}^n(s)ds.$$

Since $z^n \to z^\infty$, $u^n \to u^\infty$ strongly, and $\dot{z}^n \to \dot{z}$ weakly in $L^2(0,T)$,

$$I_n \xrightarrow{n\to\infty} I_\infty := \int_0^T (z^\infty(s) - u^\infty(s) - y(s))\dot{z}^\infty(s)ds.$$

Inserting $\dot{z}^n = \frac{1}{\varepsilon_n}G(z-u)$ yields

$$I_n = \frac{1}{\varepsilon_n}\int_0^T (u^n(s) - z^n(s) - y(s))G(z^n(s) - u^n(s)))ds.$$

The integrand is, for all $n \in \mathbb{N}$, a nonnegative function, because

$$G(z^n(t) - u^n(t)) \begin{cases} > 0, & \text{if } z^n(t) - u^n(t) < -r \\ < 0, & \text{if } z^n(t) - u^n(t) > r \\ = 0, & \text{else.} \end{cases}$$

This implies $I_\infty \geq 0$, and by theorem 2.10, $z^\infty = \mathcal{F}_r[u^\infty, z_0]$. In particular, this means that every subsequence has a subsequence that converges to $\mathcal{F}_r[u^\infty, z_0]$, hence, $\lim_{\varepsilon\downarrow 0} z^\varepsilon = \mathcal{F}_r[u^\infty, z_0]$, as claimed.

□

2.5 A general approximation result

In fact, one can use the procedure demonstrated in section 2.4 to prove convergence in the case when G is replaced by some more general function F. However, the proof and the result itself will be slightly different. Let, in the following, z denote the solution of

$$\begin{aligned} \varepsilon\dot{z} &= F(z-u), \\ z(0) &= z_0, \end{aligned} \tag{2.6}$$

and let again assumption 2.7 hold. We will assume later that F should be Lipschitz continuous; then, just as in the proof of lemma 2.8, one can see that there exists a unique solution to (2.6). Next, we refine some of the results stated in section 2.3; in fact, this will be nothing new, but just some other way of writing down how the play operator behaves on shifted initial values.

Definition 2.12 *We call the function*

$$f_{r,s}(v,w) := \max\{v-r, \min\{v+s, w\}\},$$

*where, $r,s>0$, the **update function of the nonsymmetric play operator**, which will be denoted by $\mathcal{F}_{r,s}$.*

□

If $r=s$, then we get the usual update function considered in section 2.3.

Lemma 2.13 *If $r_j, s_j \geq 0$, $v_j, w_j \in \mathbb{R}$, $j=1,2$, then*

$$|f_{r_1,s_1}(v_1,w_1) - f_{r_2,s_2}(v_2,w_2)| \\ \leq \max\{|v_1-v_2| + |r_1-r_2| + |s_1-s_2|, |w_1-w_2|\}.$$

Proof: As noted in the proof of lemma 2.2, for all $a,b,c,d \in \mathbb{R}$,

$$|\min\{a,b\} - \min\{c,d\}| \leq \max\{|a-c|, |b-d|\},$$

as well as

$$|\max\{a,b\} - \max\{c,d\}| \leq \max\{|a-c|, |b-d|\}.$$

We apply this to $|f_{r_1,s_1}(v_1,w_1) - f_{r_2,s_2}(v_2,w_2)|$, to get

$$|f_{r_1,s_1}(v_1,w_1) - f_{r_2,s_2}(v_2,w_2)| \leq \\ \max\{|v_1-r_1-v_2+r_2|, |w_1-w_2|, |v_1-s_1-v_2+s_2|\},$$

which implies the assertion.

□

For continuous inputs, one therefore gets

$$|\mathcal{F}_{r_1,s_1}[v_1;w_{-1,1}](t) - \mathcal{F}_{r_2,s_2}[v_2;w_{-1,2}](t)| \\ \leq \max\left\{|r_1-r_2| + |s_1-s_2| + \sup_{0\leq\tau\leq t}|v_1(\tau)-v_2(\tau)|, |w_{-1,1}-w_{-1,2}|\right\}. \tag{2.7}$$

From inequality (2.7), we learn in particular, that $\mathcal{F}_{r,s}[v;w_{-1}]$ converges uniformly to $\mathcal{F}_{l,l}[v;w_{-1}]$, whenever $r\to l$ and $s\to l$. Next, we want to approximate the nonsymmetric play operator in the manner described in section 2.4. To this end, we introduce the functions

$$G_{(r,s),(m_1,m_2)}(x) := m_1(-x-r)_+ - m_2(-x+s)_-,$$

where m_1, m_2, r, $s > 0$, and the corresponding differential equations

$$\dot{z} = \frac{1}{\varepsilon} G_{(r,s),(m_1,m_2)}(z-u). \tag{2.8}$$

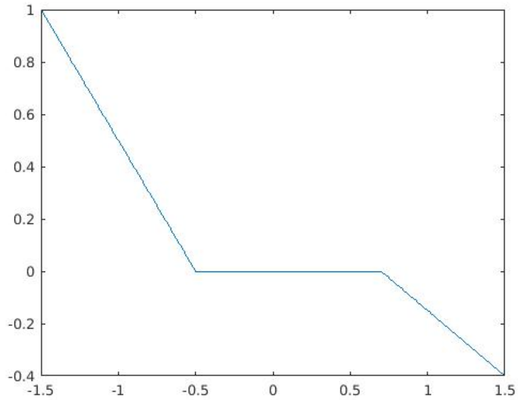

Figure 7: graph of $G_{(r,s),(m_1,m_2)}$ for r=0.5, $m_1 = 1$, s=0.7, $m_2 = 0.5$

Similar to 2.7, we make the following assumptions.

Assumption 2.14

(V1') $u^\varepsilon \to u^\infty$ *strongly in* $C[0,T]$ *as* $\varepsilon \downarrow 0$,

(V2') u^ε *is bounded in* $H^1(0,T)$ *as* $\varepsilon \downarrow 0$,

(V3') z_0, u_0 *do not depend on* $\varepsilon > 0$, *and* $z_0 - u_0 \in [-r, s]$.

□

For simplicity, we will again most of the time neglect the index ε. The function $G_{(r,s),(m_1,m_2)}$ has the same continuity properties as the function G considered in section 2.4, so that existence, uniqueness and regularity properties of solutions to (2.8) may be proven by the same arguments as in lemma 2.8. As a next step, we show a result similar to lemma 2.9.

Lemma 2.15 *Let assumption 2.14 hold. Then there is a constant* $C > 0$, *such that*

$$0 \leq \frac{1}{\varepsilon^2} \int_0^T G_{(r,s),(m_1,m_2)}(z-u)^2 ds \leq C.$$

Proof: Again, we define the auxiliary function

$$V(t) := \frac{1}{2}\dot{z}(t)^2.$$

We may apply the chain rule for Sobolev functions to calculate its weak derivative, which yields

$$\begin{aligned}\dot{V} = \dot{z}\ddot{z} =& \frac{1}{\varepsilon}G_{(r,s),(m_1,m_2)}(z-u)\frac{1}{\varepsilon}G'_{(r,s),(m_1,m_2)}(z-u)(\dot{z}-\dot{u})\\ =& \frac{1}{\varepsilon^3}G_{(r,s),(m_1,m_2)}(z-u)^2G'_{(r,s),(m_1,m_2)}(z-u)\\ &-\frac{1}{\varepsilon^2}G_{(r,s),(m_1,m_2)}(z-u)G'_{(r,s),(m_1,m_2)}(z-u)\dot{u}.\end{aligned}$$

According to assumption $(V3')$, $V(0)=0$, so that

$$0 \le V(t) = \int_0^t \dot{V}(s)ds.$$

Further, since $G'_{(r,s),(m_1,m_2)} \le 0$, we get $G' = -|G|$, so that the latter implies

$$\begin{aligned}0 \le& \int_0^t G_{(r,s),(m_1,m_2)}(z-u)^2 \left|G'_{(r,s),(m_1,m_2)}(z-u)\right| ds\\ \le& \varepsilon \int_0^t G_{(r,s),(m_1,m_2)}(z-u)\left|G'_{(r,s),(m_1,m_2)}(z-u)\right| \dot{u}ds.\end{aligned} \tag{2.9}$$

Note that, by definition of $G_{(r,s),(m_1,m_2)}$,

$$\left|G'_{(r,s),(m_1,m_2)}(z-u)\right| = m_1\chi_1(z-u) + m_2\chi_2(z-u), \text{ a.e.,}$$

where χ_1, χ_2 are suitable characteristic functions. Thus, (2.9) may be written as

$$\begin{aligned}m_1\int_0^t G^2_{(r,s),(m_1,m_2)}\chi_1 ds + m_2\int_0^t G^2_{(r,s),(m_1,m_2)}\chi_2 ds&\\ \le \varepsilon \int_0^t G_{(r,s),(m_1,m_2)}\dot{u}\left(m_1\chi_1 + m_2\chi_2\right)ds.&\end{aligned}$$

Now assume w.l.o.g. that $m_1 \ge m_2$, and multiply the latter inequation by $\frac{1}{m_2}$. This yields

$$\begin{aligned}&\int_0^t G^2_{(r,s),(m_1,m_2)}\chi_1 ds + \int_0^t G^2_{(r,s),(m_1,m_2)}\chi_2 ds\\ &\le \frac{m_1}{m_2}\int_0^t G^2_{(r,s),(m_1,m_2)}\chi_1 ds + \int_0^t G^2_{(r,s),(m_1,m_2)}\chi_2 ds\end{aligned}$$

$$\leq \varepsilon \int_0^t G_{(r,s),(m_1,m_2)} \dot{u} \left(\frac{m_1}{m_2} \chi_1 + \chi_2 \right).$$

The property $G_{(r,s),(m_1,m_2)}(z-u) \neq 0 \Rightarrow \chi_1 + \chi_2 = 1$, together with the Cauchy Schwartz inequality, then yield

$$\begin{aligned}
&\int_0^t G_{(r,s),(m_1,m_2)}(z-u)^2 ds \\
&\leq \varepsilon \left\| G_{(r,s),(m_1,m_2)}(z-u) \left(\frac{m_1}{m_2} \chi_1 + \chi_2 \right) \right\|_{L^2(0,t)} \|\dot{u}\|_{L^2(0,t)} \\
&\leq \varepsilon \mu \left\| G_{(r,s),(m_1,m_2)}(z-u) \right\|_{L^2(0,t)} \|\dot{u}\|_{L^2(0,t)},
\end{aligned}$$

because $\mu := \frac{m_1}{m_2} \geq 1$. Thus,

$$\left\| G_{(r,s),(m_1,m_2)}(z-u) \right\|_{L^2(0,t)} \leq \varepsilon \mu \|\dot{u}\|_{L^2(0,t)},$$

which implies the assertion.

□

Theorem 2.16 *Under assumption 2.14, it holds* $z^\varepsilon \to \mathcal{F}_{r,s}[u^\infty; z_0]$ *(strongly) in* $C[0,T]$ *and weakly in* $H^1(0,T)$.

Proof: Similar to theorem 2.11. Since $\mathcal{F}_{r,s}$ is equal to some shift of $\mathcal{F}_l$, with $l = \frac{r+s}{2}$, an analogous characterization result like theorem 2.10 is valid. Again, by assumption 2.14 and lemma 2.15,

$$\begin{aligned}
I_n :=& \int_0^T \left(u^n(s) - z^n(s) - y(s) \right) \dot{z}^n(s) ds \\
=& \int_0^T \left(u^n(s) - z^n(s) - y(s) \right) \frac{1}{\varepsilon} G_{(r,s),(m_1,m_2)}(z^n(s) - u^n(s)) ds \\
\geq& 0,
\end{aligned}$$

by distinction of cases, which implies $I_\infty \geq 0$. Another application of lemma 2.15 shows that $z^\infty - u^\infty \in [-r, s]$, and the result follows.

□

To build a bridge to equation 2.6, we need a simple well known theorem about solutions of ordinary differential equations (compare [19, theorem 1.3]). Let

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(0) = x_0, \end{cases} \quad \text{and} \quad \begin{cases} \dot{y}(t) = g(t, y(t)), \\ y(0) = y_0. \end{cases} \tag{2.10}$$

Theorem 2.17 *Let $x, y : [0, \tau) \to \mathbb{R}$ be piecewise continuously differentiable solutions to* (2.10)*, with f, g such that there exists $L > 0$:*

$$f(t, \xi) - g(t, \eta) \geq -L\,|\xi - \eta| \text{ for all } t \in [0, \tau) \text{ and } \xi, \eta \in \mathbb{R}.$$

If $x_0 > y_0$, then

$$x(t) > y(t) \text{ for } t \in [0, \tau).$$

Else, if $x_0 \geq y_0$, then

$$x(t) \geq y(t) \text{ for } t \in [0, \tau).$$

Proof: We start with the case $x_0 > y_0$. Assume that the assertion was false. Then, by continuity of the functions x and y, there must be some $t_0 > 0$ such that $(x-y)(t_0) = 0$ and $(x - y)(t) > 0$ for all $0 \leq t < t_0$. By assumption, it holds for almost every $t \in [0, t_0]$,

$$\dot{x}(t) - \dot{y}(t) = f(t, x(t)) - g(t, y(t)) \geq -L(x(t) - y(t)),$$

so that

$$\frac{d}{dt}\left(e^{Lt}\,(x(t) - y(t))\right) = e^{Lt}\,((\dot{x}(t) - \dot{y}(t)) + L\,(x(t) - y(t))) \geq 0.$$

But then

$$e^{Lt_0}\,(x(t_0) - y(t_0)) - (x_0 - y_0) = \int_0^{t_0} \frac{d}{dt}\left(e^{Lt}\,(x(t) - y(t))\right) dt \geq 0,$$

which implies

$$x(t_0) - y(t_0) \geq e^{-Lt_0}(x_0 - y_0) > 0,$$

a contradiction to $(x-y)(t_0) = 0$, and the first part of the theorem is proven. In order to prove the second part, we only have to check the case $x_0 = y_0$ (because otherwise, the assertion follows by the first part of the theorem). Assume for contradiction that there exists $t_1 \in [0, \tau)$, such that $x(t_1) < y(t_1)$. The continuity of x and y implies that there exists some interval $[t_0, t_1] \subset [0, \tau)$ with $x(t_0) = y(t_0)$ and $x(t) < y(t)$ for all $t \in (t_0, t_1]$. Since in that case, by assumption,

$$\dot{x}(t) - \dot{y}(t) = f(t, x(t)) - g(t, y(t)) \geq -L\,(y(t) - x(t))\,, \text{ a.e.},$$

it holds for almost every $t \in [t_0, t_1]$,

$$\frac{d}{dt}\left(e^{-Lt}\,(x(t) - y(t))\right) = e^{-Lt}\,((\dot{x}(t) - \dot{y}(t)) - L\,(x(t) - y(t))) \geq 0.$$

But this implies

$$e^{-Lt_1}(x(t_1)-y(t_1)) = e^{-Lt_1}(x(t_1)-y(t_1)) - e^{-Lt_0}(x(t_0)-y(t_0)) = \int_{t_0}^{t_1} \frac{d}{dt}\left(e^{-Lt}(x(t)-y(t))\right) dt \geq 0,$$

a contradiction to the definition of the interval $[t_0, t_1]$.

□

Now, for given $T > 0$, $y \in H^1(0,T)$ and $\lambda > 0$, we study the behaviour of the initial value problems

$$\begin{cases} \dot{z}(t) = \lambda F(z(t)-y(t)),\ z(0) \in \mathbb{R}, \\ \dot{z}_1(t) = \lambda G_{(r_1,s_1),(m_1,m_2)}(z_1(t)-y(t)),\ z_1(0) \in \mathbb{R}, \\ \dot{z}_2(t) = \lambda G_{(r_2,s_2),(n_1,n_2)}(z_2(t)-y(t)),\ z_2(0) \in \mathbb{R}, \end{cases} \tag{2.11}$$

under the condition

$$G_{(r_1,s_1),(m_1,m_2)}(x) \leq F(x) \leq G_{(r_2,s_2),(n_1,n_2)}(x),\ \forall x \in \mathbb{R}. \tag{2.12}$$

From theorem 2.17, we infer the following result.

Corollary 2.18 *Let* (2.12) *hold, and F be some Lipschitz continuous function. If $z_1(0) \leq z(0) \leq z_2(0)$, then the solutions of* (2.11) *fulfill*

$$z_1(t) \leq z(t) \leq z_2(t),\ \forall t \in [0,T].$$

Proof: We use the abbreviations $G_1 := G_{(r_1,s_1),(m_1,m_2)}$, $G_2 := G_{(r_2,s_2),(n_1,n_2)}$. Note that G_1 is Lipschitz continuous with constant $L_1 := \max\{m_1, m_2\}$ and G_2 is Lipschitz with constant $L_2 := \max\{n_1, n_2\}$; let L_F denote the Lipschitz constant of F. Existence and uniqueness of solutions to (2.11) can be shown as in lemma 2.8. Now, (2.12) implies that

$$\begin{aligned} \lambda F(\xi - y(t)) - \lambda G_1(\eta - y(t)) \geq & \lambda\left(G_1(\xi - y(t)) - G_1(\eta - y(t))\right) \\ \geq & -\lambda L_1 \left|\xi - y(t) - (\eta - y(t))\right| \\ = & -\lambda L_1 \left|\xi - \eta\right|, \end{aligned}$$

and thus $z \geq z_1$ by theorem 2.17. On the other hand, we also have

$$\begin{aligned} \lambda G_2(\eta - y(t)) - \lambda F(\xi - y(t)) \geq & \lambda\left(F(\eta - y(t)) - F(\xi - y(t))\right) \\ \geq & -\lambda L_F \left|\eta - y(t) - (\xi - y(t))\right| \\ = & -\lambda L_F \left|\xi - \eta\right|, \end{aligned}$$

so that $z_2 \geq z$ by theorem 2.17, and the proof is complete. □

Now we come to the main theorem of this section.

Theorem 2.19 *Let assumption 2.14 hold, and F be some Lipschitz continuous function, such that for every $\delta > 0$, there exist $G_1^\delta, G_2^\delta : \mathbb{R} \to \mathbb{R}$ satisfying*

$$\begin{aligned} G_1^\delta(x) :=& G_{(r+\delta,r),(m_1(\delta),m_2)}(x) \le F(x), \\ G_2^\delta(x) :=& G_{(r,r+\delta),(n_1,n_2(\delta))}(x) \ge F(x), \end{aligned}$$

for all $x \in \mathbb{R}$, where $m_1(\delta), m_2, n_1, n_2(\delta) > 0$. Then the solutions z^ε of (2.6) converge to $\mathcal{F}_r[u^\infty; z_0]$, strongly in $C[0,T]$ as $\varepsilon \downarrow 0$.

Proof: Let $\mu > 0$ be arbitrary. We choose admissible initial values, i.e.,

$$z_1^\varepsilon(0) = z_2^\varepsilon(0) = z^\varepsilon(0) = z_0 \in [u_0 - r, u_0 + r], \ \forall \varepsilon > 0.$$

Now, if $\varepsilon, \delta > 0$ are fixed, we may apply corollary 2.18, to get $z_1^\varepsilon(t) \le z^\varepsilon(t) \le z_2^\varepsilon(t)$, for all $t \in [0,T]$, which implies

$$0 \le z^\varepsilon(t) - z_1^\varepsilon(t) \le z_2^\varepsilon(t) - z_1^\varepsilon(t), \ \forall t \in [0,T],$$

and hence,

$$\|z^\varepsilon - z_1^\varepsilon\|_{C[0,T]} \le \|z_1^\varepsilon - z_2^\varepsilon\|_{C[0,T]} .$$

We choose $\delta = \frac{\mu}{9}$. By theorem 2.16, $z_1^\varepsilon \to \mathcal{F}_{r+\mu/9,r}[u^\infty; z_0]$ and $z_2^\varepsilon \to \mathcal{F}_{r,r+\mu/9}[u^\infty; z_0]$ in $C[0,T]$ as $\varepsilon \downarrow 0$. Further, (2.7) implies

$$\left\|\mathcal{F}_{r+\mu/9,r}[u^\infty; z_0] - \mathcal{F}_{r,r+\mu/9}[u^\infty; z_0]\right\|_{C[0,T]} \le \frac{2}{9}\mu.$$

Altogether, this implies that

$$\begin{aligned} \|z_1^\varepsilon - z_2^\varepsilon\|_{C[0,T]} \le& \left\|z_1^\varepsilon - \mathcal{F}_{r+\mu/9,r}[u^\infty; z_0]\right\|_{C[0,T]} + \left\|z_2^\varepsilon - \mathcal{F}_{r,r+\mu/9}[u^\infty; z_0]\right\|_{C[0,T]} \\ &+ \left\|\mathcal{F}_{r+\mu/9,r}[u^\infty; z_0] - \mathcal{F}_{r,r+\mu/9}[u^\infty; z_0]\right\|_{C[0,T]} \\ \le& \frac{\mu}{3}, \end{aligned}$$

if $\varepsilon > 0$ is chosen such that the first two summands both are smaller than $\frac{\mu}{18}$. But for those ε then holds

$$\begin{aligned} \|z^\varepsilon - \mathcal{F}_r[u^\infty; z_0]\|_{C[0,T]} \le& \|z^\varepsilon - z_1^\varepsilon\|_{C[0,T]} + \|z_1^\varepsilon - z_2^\varepsilon\|_{C[0,T]} \\ &+ \left\|z_2^\varepsilon - \mathcal{F}_{r,r+\mu/9}[u^\infty; z_0]\right\|_{C[0,T]} \\ &+ \left\|\mathcal{F}_{r,r+\mu/9}[u^\infty; z_0] - \mathcal{F}_r[u^\infty; z_0]\right\|_{C[0,T]} \\ \le& 2\|z_1^\varepsilon - z_2^\varepsilon\|_{C[0,T]} + \frac{\mu}{18} + \frac{\mu}{9} \\ <& \mu. \end{aligned}$$

As $\mu > 0$ was arbitrary, the result follows. □

Remark 2.20 *We can not conclude as in theorem 2.11, that z^ε converges weakly in $H^1(0,T)$, because from*

$$z_1 - u^\varepsilon \le z^\varepsilon - u^\varepsilon \le z_2 - u^\varepsilon,$$

it only follows (if F is monotone)

$$F(z_1 - u^\varepsilon) \ge F(z^\varepsilon - u^\varepsilon) \ge F(z_2 - u^\varepsilon),$$

and thus, by definition,

$$G_2(z_1 - u^\varepsilon) \ge F(z^\varepsilon - u^\varepsilon) \ge G_1(z_2 - u^\varepsilon).$$

From this inequality, we can derive

$$\frac{1}{\varepsilon}\left|F(z^\varepsilon - u^\varepsilon)\right| \le \frac{1}{\varepsilon}\left(\left|G_2(z_1 - u^\varepsilon)\right| + \left|G_1(z_2 - u^\varepsilon)\right|\right),$$

but we do not know whether the right hand side is bounded as $\varepsilon \downarrow 0$ or not.

□

2.6 Pointwise applied play operators and weak differentiability

As seen in section 2.3, the play operator on strings is completely described by the update function

$$f_r(y,w) := \max\{y - r, \min\{y + r, w\}\},$$

and this can be used to completely describe the operator on the vector space $C_{pl}[0,T]$ of continuous piecewise linear functions. Then, with some continuity arguments, one can define the output for every input function in $C[0,T]$. In later sections, when the play operator appears in the partial differential equation, we will have pointwise, for almost every $x \in \Omega$ (some open, bounded domain), a trajectory $y(\cdot,x) \in C[0,T]$, and the play operator applied to it, i.e., $\mathcal{F}_r[y(\cdot,x); w_0(x)]$. To this end, we now change perspectives. We start by writing the update function in another way.

Lemma 2.21 *It holds*

$$f_r(y,w) = \big(y - r + (w - y - r)_+ - w\big)_+ - (w - y - r)_+ + w.$$

Proof: By distinction of cases, one easily sees that

$$\begin{aligned}\min\{y+r,w\} &= \begin{cases} y+r, & \text{if } y+r \le w \Leftrightarrow w-y-r \ge 0 \\ w, & \text{if } w \le y+r \Leftrightarrow w-y-r \le 0 \end{cases} \\ &= -\left[w-y-r\right]_+ + w,\end{aligned}$$

and

$$\begin{aligned}\max\{a,b\} &= \begin{cases} a, & \text{if } a \ge b \Leftrightarrow a-b \ge 0 \\ b, & \text{if } b \ge a \Leftrightarrow a-b \le 0 \end{cases} \\ &= [a-b]_+ + b.\end{aligned}$$

Choosing $a = y - r$ and $b = \min\{y+r,w\}$ proves the claim.

□

Let now $\Omega \subset \mathbb{R}^n$ be some open, bounded domain with smooth boundary (at least Lipschitzian). For $y \in L^2(\Omega; C[0,T])$ and (a.e.) $x \in \Omega$, we consider the maps $t \mapsto y(x)(t) = y(t,x) \in C[0,T]$. As seen in the construction of the play operator, it is not really important how the decomposition of the interval $[0,T]$ looks like, to get convergence of the output functions. So, we choose, for all x and corresponding functions $y(\cdot,x)$, the same decomposition. This leads us to the space of continuous piecewise linear functions with values in some space X (here, e.g., $X = L^2(\Omega)$).

Definition 2.22 *We define the space of piecewise linear functions with values in some Banach space X, by*

$$\begin{aligned}C_{pl}(0,T;X) := \{y \in C([0,T];X) :\ & \exists n \in \mathbb{N}, 0 = t_0 < t_1 < \cdots < t_n = T, \\ & \text{and } x_i \in X, i = 0,\dots,n, \text{ such that,} \\ & y(t) = \frac{t-t_j}{t_{j+1}-t_j} x_{j+1} + \frac{t_{j+1}-t}{t_{j+1}-t_j} x_j,\ \forall\, t \in [t_j,t_{j+1}]\}\,.\end{aligned}$$

□

$C_{pl}(0,T;X)$ is a subspace of $C(0,T;X)$, which is not closed, but dense.

Lemma 2.23 *Let X be a seperable Banach space. $C_{pl}([0,T];X)$ is dense in $C(0,T;X)$ and in the Bochner spaces $L^p(0,T;X)$ for $p \in [1,\infty)$.*

Proof: Let $y \in C(0,T;X)$, $\varepsilon > 0$, and $t_0 = 0$. Since the mapping $f_0 : t \mapsto \|y(0) - y(t)\|_X$ is a continuous real function,

$$t_1 := \inf\{\tau \in [0,T] \mid f_0(\tau) > \varepsilon\}$$

is well defined whenever the set is nonempty (if it is, we choose $t_1 = T$). By continuity, $t_1 > 0$, $f_0(t_1) = \varepsilon$ and $f_0(t) \leq \varepsilon \ \forall t \in [t_0, t_1]$. We iterate this procedure, i.e., for given t_{i-1}, we define $f_{i-1} : t \mapsto \|y(t_{i-1}) - y(t)\|_X$, $t \in [t_{i-1}, T]$, and choose

$$t_i := \inf \{\tau \in [t_{i-1}, T] \mid f_{i-1}(\tau) > \varepsilon\} ;$$

if the set is empty, we choose $t_i := T$. Assume that this procedure would not stop after finitely many steps. Then, the $(t_i)_i$ would form a strictly monotonic increasing sequence, which is bounded by T, and thus convergent. Let t^* denote this limit. Then, since the algorithm did not converge after finitely many steps, it must hold that

$$\|y(t_{i+1}) - y(t_i)\| = \varepsilon > 0, \ \forall i \in \mathbb{N}.$$

But this is a contradiction to the continuity of y, as both $y(t_{i+1})$, $y(t_i)$ converge to $y(t^*)$ as $i \to \infty$. Hence, the procedure stops after finitely many steps. Then, by choosing $y(t_j) := x_j$ in definition 2.22, we get an element in $C_{pl}(0, T; X)$, which has distance at most ε from y w.r.t. $\|\cdot\|_{C(0,T;X)}$, and the first part of the lemma is proven.

For the second part, let $y \in L^p(0, T; X)$ and $\varepsilon > 0$. By definition of Bochner's integral, there is a sequence of simple functions $y_n = \sum_{i=1}^n \chi_{A_i} x_i$ such that $A_i \subset [0, T]$, $x_i \in X$, which converges to y in $L^p(0, T; X)$. We choose $n \in \mathbb{N}$, such that $\|y_n - y\|_{L^p(0,T;X)} < \frac{\varepsilon}{2}$. Note that, by definition, $y_n|_{A_i} = x_i$ is constant in t, with $x_i \in X$. But the function y_n can be approximated by functions in $C(0, T; X)$ (e.g., by mollification w.r.t. time), and by the first part, also by functions in $C_{pl}(0, T; X)$. Hence, there exists $y^* \in C_{pl}(0, T; X)$, such that $\|y_n - y^*\|_{L^p(0,T;X)} < \frac{\varepsilon}{2}$. But then,

$$\begin{aligned}\|y - y^*\|_{L^p(0,T;X)} &\leq \|y - y_n\|_{L^p(0,T;X)} + \|y_n - y^*\|_{L^p(0,T;X)} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,\end{aligned}$$

which concludes the proof.

□

As positive and negative part are weakly differentiable functions, the update function f_r also has this property. In what follows, we allow the parameter r to depend on $x \in \Omega$.

Proposition 2.24 *Let $k \in \mathbb{N}$, D denote some (weak) partial derivative, i.e., $D = \partial_{x_i}$ for some i. If $r, w_0 \in W^{1,p}(\Omega)$, $y = (y_1, \ldots, y_{k+1})^T \in (W^{1,p}(\Omega))^{k+1}$, $p \in [1, \infty]$, there exist functions $\lambda_j^{k+1}(x)$, $0 \leq j \leq k$, with $0 \leq \lambda_j^{k+1} \leq 1$ and $\sum_{j=1}^k \lambda_j^{k+1} \leq 1$, such that (a.e. in Ω),*

$$|D\mathcal{F}_{r,f}(y; w_0)| \leq \left(\sum_{j=1}^{k+1} \lambda_j^{k+1} |Dy_j| \right) + |Dr| + |Dw_0| .$$

Proof: Let H denote the Heavyside function. Thanks to lemma 2.21 and the chain rule for Sobolev functions,

$$\begin{aligned} &Df_r(y_1,w_0)\\ &\quad= H(y_1-r+[w_0-y_1-r]_+-w_0)(Dy_1-Dr-Dw_0)\\ &\qquad+H(y_1-r+[w_0-y_1-r]_+-w_0)H(w_0-y_1-r)(Dw_0-Dy_1-Dr)\\ &\qquad-H(w_0-y_1-r)(Dw_0-Dy_1-Dr)+Dw_0\\ &\quad=\chi_1\,(Dy_1-Dr-Dw_0)+\chi_1\chi_2\,(Dw_0-Dy_1-Dr)\\ &\qquad-\chi_2\,(Dw_0-Dy_1-Dr)+Dw_0, \end{aligned}$$

where χ_i are defined through

$$\chi_1 := H(y_1-r+[w_0-y_1-r]_+-w_0),\ \chi_2 := H(w_0-y_1-r).$$

In the case $\chi_2 \neq 0 \Leftrightarrow w_0-y_1-r>0$, we infer that

$$\chi_1 = H\left(y_1-r+w_0-y_1-r-w_0\right)=H(-2r)=0,$$

hence, $\chi_1\chi_2=0$, and the formula for the derivative simplifies to

$$Df_r(y_1,w_0) =(\chi_1+\chi_2)Dy_1+(1-\chi_1-\chi_2)Dw_0+(\chi_2-\chi_1)Dr.$$

This may be rewritten by introducing $\lambda_0 := \chi_1+\chi_2$ and $\mu_0 := \chi_2-\chi_1$, and then has the form

$$Df_r(y_1,w_0)=\lambda_0 Dy_1+(1-\lambda_0)Dw_0+\mu_0 Dr,$$

which is actually even stronger than the claimed inequality, for $k=0$. We continue by induction. Assume that the claim holds for k. An analogous computation shows that there exist characteristic functions χ_1^k, χ_2^k with $\chi_1^k\chi_2^k=0$, and such that

$$Df_r(y_{k+1},w_k)=(\chi_1^k+\chi_2^k)Dy_{k+1}+(1-\chi_1^k-\chi_2^k)Dw_k+(\chi_2^k-\chi_1^k)Dr.$$

Again, we rewrite the equation, using $\lambda_k := \chi_1^k+\chi_2^k$ and $\mu_k := \chi_2^k-\chi_1^k$, and get

$$Df_r(y_{k+1},w_k)=\lambda_k Dy_{k+1}+(1-\lambda_k)Dw_k+\mu_k Dr.$$

Now, by definition, $w_k=f_r(y_k,w_{k-1})$, so that we can estimate

$$\begin{aligned} |Df_r(y_{k+1},w_k)| \leq&\lambda_k\left|Dy_{k+1}\right|+(1-\lambda_k)\left(\sum_{j=0}^{k-1}\lambda_j^k\left|Dy_{j+1}\right|\right)\\ &+(1-\lambda_k)\left(|Dr|+|Dw_0|\right)+|\mu_k|\,|Dr|, \end{aligned}$$

where, we used the induction hypothesis. Next, note that $\mu_k(1-\lambda_k)=0$, because

$$\begin{aligned}\chi_2^k = 1 \Rightarrow \chi_1^k = 0 &\Rightarrow (1-\lambda_k) = 0 \ \wedge \ \mu_k = 1\\ \chi_2^k = 0 \ \wedge \ \chi_1^k = 0 &\Rightarrow (1-\lambda_k) = 1 \ \wedge \ \mu_k = 0\\ \chi_2^k = 0 \ \wedge \ \chi_1^k = 1 &\Rightarrow (1-\lambda_k) = 0 \ \wedge \ \mu_k = -1.\end{aligned}$$

But then $(1-\lambda_k)\,|Dr| + |\mu_k|\,|Dr| \leq |Dr|$. Further, $\lambda_k + (1-\lambda_k)(\sum_{j=0}^{k-1}\lambda_j^k) \leq \lambda_k + (1-\lambda_k) = 1$, so that

$$|Df_r(y_{k+1}, w_k)| \leq \lambda_k\,|Dy_{k+1}| + \left(\sum_{j=0}^{k-1}(1-\lambda_k)\lambda_j^k\,|Dy_{j+1}|\right) + |Dw_0| + |Dr|\,.$$

Hence, the claim holds with iterative defined

$$\lambda_k^{k+1} := \lambda_k, \quad \lambda_j^{k+1} := (1-\lambda_k)\lambda_j^k,\ j \in \{0,\ldots,k-1\}\,,$$

where λ_k is given through the characteristic functions χ_1^k, χ_2^k.

□

What does proposition 2.24 mean for functions $y \in C_{pl}([0,T], H^1(\Omega))$? On the one hand, one can easily complement the values of $\mathcal{F}_r[y; w_0]$ between the gridpoints t_i, which yields the solution $w = \mathcal{F}_r[y; w_0] \in L^2(\Omega; C[0,T])$. On the other hand, by linearity of differentiation operators and the property, that interim values can be represented as convex combinations of grid points, an inequality as in proposition 2.24 must hold (almost) everywhere for w. We thus have the following result.

Corollary 2.25 *Let $p \in [1,\infty]$, $y \in C_{pl}([0,T]; W^{1,p}(\Omega))$ and k the number of grid points for y in $[0,T]$. Let r as in proposition 2.24, and $w = \mathcal{F}_r[y; w_0]$, where $w_0 := w(0,x) \in W^{1,p}(\Omega)$. Then there exist functions $\lambda_j(t,x)$, $j = 1,\ldots,k$, with the property $0 \leq \lambda_j \leq 1$ and $\sum_{j=1}^{k}\lambda_j \leq 1$, such that (a.e.),*

$$|Dw| \leq \left(\sum_{j=1}^{k}\lambda_j\,|Dy_j|\right) + |Dr| + |Dw_0|\,.$$

Proof: Inserting a linear combination of y_{i+1} and y_i in the inequality shows the property for interim values, corresponding to time points between t_i and t_{i+1}. The λ_j with $j > i+1$ equal zero at those time points.

□

The following theorem will later be used to prove some additional regularity result for solutions of a special partial differential equation.

Theorem 2.26 *Let $p \in [1, \infty]$ and $(y_m)_m \subset C_{pl}([0,T]; W^{1,p}(\Omega))$ be a sequence which converges in $L^p(\Omega; C[0,T])$ to y. Let w_0, r be as in corollary 2.25, $w := \mathcal{F}_r[y; w_0]$ and $w_m := \mathcal{F}_r[y_m; w_0]$. If there exists a constant $C > 0$, such that, for any weak partial derivative $D \in \{\partial_{x_1}, \dots, \partial_{x_n}\}$, it holds*

$$\max_{i \in \{1,\dots,k(m)\}} |Dy_{m,i}| \leq C,$$

for all $m \in \mathbb{N}$ and almost every $x \in \Omega$, then $w \in L^\infty(0,T; W^{1,p}(\Omega))$. Here, $k(m)$ denotes the number of grid points of y_m.

Proof: By corollary 2.25, for every $t \in [0,T]$ and almost every $x \in \Omega$,

$$\begin{aligned} |Dw|^p &\leq \left(\left(\sum_{j=1}^{k} \lambda_j \, |Dy_j| \right) + |Dr| + |Dw_0| \right)^p \\ &\leq \left(C \left(\sum_{j=1}^{k} \lambda_j \right) + |Dr| + |Dw_0| \right)^p, \end{aligned}$$

which is bounded in $L^1(\Omega)$ by assumption. The sequence $(w_m)_m$ is then bounded in $L^\infty(0,T; W^{1,p}(\Omega))$, so that there exists a weak star convergent subsequence. By uniqueness of limits, every convergent subsequence must converge to w, and thus $w \in L^\infty(0,T; W^{1,p}(\Omega))$.

□

2.7 A representation result for the one dimensional stop operator via projections in a Hilbert space

In this section, we consider the one dimensional stop operator on the Hilbert space $H^1(0,T)$. It may be defined as solution operator to the following problem (cf., e.g. [17], where a more general problem is discussed).

Problem (S): Given $z_0 \in [-r, r] =: Z$ and $y \in H^1(0,T)$, find $x \in H^1(0,T)$, such that

$$\begin{array}{lrl} (S1) & x(t) \in Z & \forall t \in [0,T], \\ (S2) & (x(t) - z)\,(\dot{y}(t) - \dot{x}(t)) \geq 0 & \forall z \in Z, \text{ a.e. } t \in [0,T], \\ (S3) & x(0) = z_0. & \end{array}$$

We note that condition $(S2)$ is remindful of a property that projections onto convex sets in Hilbert spaces have.

Theorem 2.27 *Let A be some convex closed subset of some Hilbert space $\mathcal{H}$, and let $\mathcal{P}$ denote the corresponding projection mapping, i.e.,*

$$\mathcal{P}(x) = \left\{ y \in A \mid \|y - x\|_{\mathcal{H}} = \inf_{a \in A} \{\|x - a\|_{\mathcal{H}}\} \right\}.$$

Then $\mathcal{P}(x)$ is a singleton for every $x \in \mathcal{H}$, so that we may define the projection operator P via $\mathcal{P}(x) =: \{P(x)\}$. Further, we have the inequality

$$\langle x - P(x), a - P(x) \rangle \leq 0, \ \forall a \in A.$$

Scetch of the proof (for more details, see, e.g., [20, theorem 3.14]): Existence follows by weak compactness of balls in a Hilbert space, and the weak lower semicontinuity of the norm. For uniqueness, one needs the convexity of A and the Hilbert space $\mathcal{H}$. The inequality for the operator P may be shown by some separation argument: By convexity, there exists a separating hyperplane through $P(x)$. But, by convexity of A, the vector $a - P(x)$ must point into the other halfspace compared to $x - P(x)$. Thus, the scalar product must be nonpositive.

□

Consider the bilinear form defined by

$$\langle x, y \rangle := \int_0^T \dot{x}(t)\dot{y}(t)dt,$$

and the space

$$\mathcal{H}_0^1(0, T) := \left\{ y \in H^1(0, T) \mid y(0) = 0 \right\}.$$

This is a Hilbert space if considered w.r.t. $\langle \cdot, \cdot \rangle$, because $(x, y)^*_{H^1(0,T)} := x(0)y(0) + \int_0^T \dot{x}(t)\dot{y}(t)dt$ is a scalar product on $H^1(0, T)$ that reduces to $\langle \cdot, \cdot \rangle$ on $\mathcal{H}_0^1$. We are looking for $\mathcal{E}_r[y; 0]$ in the case $y \in \mathcal{H}_0^1(0, T)$, which will, in that case, again be an element of this space. Let us introduce some notations.

Definition 2.28 *The set of all functions $a \in \mathcal{H}_0^1(0, T)$ such that $a(t) \in Z$ for all $t \in [0, T]$ will be denoted by A. Further, for the rest of this section, x, y will denote elements in $\mathcal{H}_0^1$, such that $x = P(y)$ is the projection of y in $\mathcal{H}_0^1(0, T)$ to the closed convex subset A.*

□

By definition, $x(t) \in Z$ for all $t \in [0, T]$, and $x(0) = 0 = \mathcal{E}_r[y; 0](0)$. One might ask, whether condition $(S2)$ is fulfilled by x. The answer is no.

Example 2.29 *Let* $r = 1$*,* $T = 2$*,* $f(t) := t$*. Then*

$$\mathcal{E}_r[f;0](t) = \min\{1,t\}.$$

But the minimizer for $\min_{v\in A}\left\{\int_0^2 (\dot{v}(t) - \dot{f}(t))^2 dt\right\}$ *is* $g(t) := \frac{1}{2}t$*. Actually,*

$$\int_0^2 (\dot{g}(t) - \dot{f}(t))^2 dt = \int_0^2 \left(\frac{1}{2}\right)^2 dt = \frac{1}{2},$$

whereas

$$\int_0^2 (\dot{\mathcal{E}}_r[f;0](t) - \dot{f}(t))^2 dt = \int_1^2 1dt = 1.$$

□

It is noticable, however, that $g(2) = \mathcal{E}_r[f;0](2)$ in example 2.29. We will next analyze the properties of the projection x.

Theorem 2.30 *Let* x, y *as in definition 2.28. If* T *is a Lebesgue point (see [21, def. 8.4.8, thm. 8.4.7] for details) of* $\dot{x}$ *and* $\dot{y}$*, then*

$$\dot{x}(T)\left(\dot{y}(T) - \dot{x}(T)\right) \geq 0.$$

Further, let by x_T*,* y_T *denote the dependence on* T*, and assume that* $x_T(T) \in \operatorname{int}(Z)$*. Then there is* $\varepsilon > 0$*, such that*

$$x_{T+\tau}(T) = x_T(T),\ \forall\ 0 \leq \tau < \varepsilon.$$

Proof: By theorem 2.27, the minimizer x fulfills

$$\int_0^T (\dot{y}(t) - \dot{x}(t))\dot{x}(t)dt \geq \int_0^T (\dot{y}(t) - \dot{x}(t))\dot{a}(t)dt,\ \forall a \in A. \tag{2.13}$$

Consider, for $h > 0$, the test function

$$a_h(t) := \begin{cases} x(t), & t \in [0, T-h], \\ x(T-h), & t \in [T-h, T]. \end{cases}$$

Then, we infer from the latter inequality,

$$\int_{T-h}^T (\dot{y}(t) - \dot{x}(t))\dot{x}(t)dt \geq 0.$$

Since T is, by assumption, a Lebesgue point of the integrand, we may divide by h and then let $h \downarrow 0$, to get the assertion. To prove the second part, note that if $x_T(T) \in \text{int}(Z)$, then there is $\varepsilon > 0$, such that

$$a^*(t) := \begin{cases} x_T(t), & t \in [0, T], \\ x_T(T) + \int_T^t \dot{y}_{T+\varepsilon}(s)ds, & t \in [T, T+\varepsilon], \end{cases}$$

is an element of $A_{T+\varepsilon}$. We claim that $x_{T+\varepsilon} = a^*$. Since a.e.,

$$\dot{a}^*(t) = \begin{cases} \dot{x}_T(t), & t \in (0, T), \\ \dot{y}_{T+\varepsilon}(t), & t \in (T, T+\varepsilon), \end{cases}$$

it holds that

$$\|a^* - y_{T+\varepsilon}\|_{\mathcal{H}_0^1(0,T+\varepsilon)} = \|x_T - y_T\|_{\mathcal{H}_0^1(0,T)}.$$

But the function

$$T \mapsto \|x_T - y_T\|_{\mathcal{H}_0^1(0,T)}$$

must be monotonic increasing, as the ristriction of each function in A_{T+h} to the interval $[0, T]$ is an element of A_T. Thus, a^* must be a minimizer, and by uniqueness, $x_{T+\varepsilon} = a^*$ as claimed. By definition, $a^*(T) = x_T(T)$, and the proof is complete.

□

As mentioned before the last theorem, there seems to be a connection between $x_T(T)$ and $\mathcal{E}_r[y; 0](T)$. For $h > 0$ and arbitrary $z \in Z$, consider the test function

$$a_h(t) := \begin{cases} x_T(t), & t \in [0, T-h], \\ \frac{T-t}{h} x_T(T-h) + \frac{t-(T-h)}{h} z, & t \in [T-h, T]. \end{cases}$$

It holds $a_h \in A$, by convexity of Z, and

$$\dot{a}_h(t) = \begin{cases} \dot{x}_T(t), & t \in (0, T-h) \\ -\frac{1}{h} x_T(T-h) + \frac{1}{h} z, & t \in (T-h, T), \end{cases}$$

almost everywhere. But then, inequality (2.13) yields

$$\frac{1}{h} \int_{T-h}^{T} (\dot{y}(t) - \dot{x}_T(t))(x_T(T-h) - z)dt \geq -\int_{T-h}^{T} (\dot{y}(t) - \dot{x}_T(t))\dot{x}_T(t)dt.$$

The right hand side converges to zero as $h \downarrow 0$, hence, if T is a Lebesgue point of $\dot{y}(\cdot) - \dot{x}_T(\cdot)$, we arrive at

$$(\dot{y}(T) - \dot{x}_T(T))(x_T(T) - z) \geq 0, \tag{2.14}$$

which is close to $(S2)$. This proves the first part of the following result.

Proposition 2.31 *If T is a Lebesgue point of the function $\dot{y}(\cdot) - \dot{x}_T(\cdot)$, then for all $z \in Z$, inequality (2.14) holds. Further, if T is a Lebesgue point of the derivative of $t \mapsto x_t(t)$ and $\dot{y}$, then (2.14) also holds in the sense $\dot{x}_T(T) := \frac{d}{dt}x_t(t)|_{t=T}$.*

Proof: Only the second part is left; to this end, we consider the function $t \mapsto x_t(t)$ on the interval $[0, T+h]$, and do the same construction as before (assuming for the moment that the function is an element of $\mathcal{H}_0^1(0, T+h)$, which is proved in the next proposition). Then, letting again $h \downarrow 0$ yields the result.

□

We still need some further properties of the function $t \mapsto x_t(t)$, concerning measurability, differentiability, etc.

Proposition 2.32 *If $y \in C^k[0,T]$, $k > 0$, then the function $t \mapsto x_t(t)$ is piecewise C^k. If $y \in \mathcal{H}_0^1(0,T)$, then $t \mapsto x_t(t)$ is also an element of $\mathcal{H}_0^1(0,T)$.*

Proof: We first show that $x_t(t)$ is continuous. As the initial value is always $0 \in \text{int}(Z)$, we infer from theorem 2.30 that $x_t(t)$ is continuous at $t = 0$. Let now $T > 0$ and $\varepsilon > 0$. From the transformation theorem, we get that if $x_{T+\varepsilon}$ is the projection of $y_{T+\varepsilon}$ in $\mathcal{H}_0^1(0, T+\varepsilon)$, then $x^\varepsilon := x_{T+\varepsilon}(\frac{T}{T+\varepsilon}\cdot)$ is the projection of $y^\varepsilon := y_{T+\varepsilon}(\frac{T}{T+\varepsilon}\cdot)$ in $\mathcal{H}_0^1(0,T)$. Since $y^\varepsilon \to y_T$ as $\varepsilon \downarrow 0$ in $\mathcal{H}_0^1(0,T)$, the lipschitz continuity of the projection mapping implies that $x^\varepsilon \to x_T$. Thus, in particular, $x_{T+\varepsilon}(T+\varepsilon) \to x_T(T)$ as $\varepsilon \downarrow 0$, i.e., $x_t(t)$ is right continuous. Using similar arguments, one can show that this function is also left continuous, and hence, continuity follows.

As seen in the proof of theorem 2.30, if $x_t(t) \in \text{int}(Z)$, then we can (locally) explicitly write down the minimizer $x_{T+\tau} = x_T(T) + \int_0^\tau \dot{y}(s)ds$; hence, in that case, $x_t(t)$ has the same regularity as y on that interval. In particular, x is C^k on that set of time points if y is. Whenever $x_\cdot(\cdot)$ is on the boundary of Z for some dense set of time points, then, by continuity, it is constant on some closed interval, and thus in particular C^∞ at the interior of that interval.

□

Theorem 2.33 *If $y \in \mathcal{H}_0^1(0,T)$, then $t \mapsto x_t(t)$ equals $t \mapsto \mathcal{E}_r[y;0](t)$.*

Proof: For $h > 0$, consider

$$\begin{aligned} x_{T+h}(T+h) - x_T(T) &= \int_0^{T+h} \dot{x}_{T+h}(t)dt - \int_0^T \dot{x}_T(t)dt \\ &= \int_0^T \dot{x}_{T+h}(t) - \dot{x}_T(t)dt + \int_T^{T+h} \dot{x}_{T+h}(t)dt. \end{aligned}$$

Then, for every $z \in Z$ and $T > 0$,

$$
\begin{aligned}
&\left(\dot{y}(T) - \frac{1}{h}\left[\int_0^T \dot{x}_{T+h}(t) - \dot{x}_T(t)dt + \int_T^{T+h} \dot{x}_{T+h}(t)dt\right]\right)(x_T(T) - z) \\
&= \left(\dot{y}(T) - \frac{1}{h}\int_T^{T+h} \dot{x}_{T+h}(t)dt\right)(x_T(T) - z) \\
&\quad - \frac{1}{h}\left(x_{T+h}(T) - x_T(T)\right)\left(x_T(T) - z\right).
\end{aligned}
$$

If T is a Lebesgue point of $x_t(t)$, the first expression converges, for $h \downarrow 0$, to $(\dot{y}(T) - \dot{x}_T(T))\,(x_T(T) - z)$, which is nonnegative (a.e.) by proposition 2.31. The second term is also nonnegative: if $x_T(T) \in \operatorname{int}(Z)$, then $x_T(T) = x_{T+h}(T)$ for all h small enough; else, if $x_T(T) \in \partial Z$, then $x_T(T) = r$ or $x_T(T) = -r$. In the first case, $x_{T+h}(T) - x_T(T) \leq 0$, and $x_T(T) - z \geq 0$, so that the product is nonpositive. In the second case, $x_{T+h}(T) - x_T(T) \geq 0$ and $x_T(T) - z \leq 0$, and the product is again nonpositive.

Altogether, this implies that, if T is a Lebesgue point of $t \mapsto \dot{x}_t(t)$ and $\dot{y}$, then

$$
(\dot{y}(T) - \dot{x}_T(T))\,(x_T(T) - z) \geq 0.
$$

By proposition 2.32 and the Lebesgue differentiation theorem (cf. [21, theorem 8.4.7]), almost every point is a Lebesgue point, hence, $(S2)$ is valid.

□

Remark 2.34 *Most of the argumentation holds for the case of multidimensional stop operator. It is, however, not clear, what happens with the factor $\frac{1}{h}(x_{T+h}(T) - x_T(T))$ in the case $x_T(T) \in \partial Z$. The simulation of a two dimensional example seemed not to show convergence:*
We used the unit circle as characteristic set Z. As reference curve served an approximation via the standard discretization scheme. This is, in the one dimensional case, given by the update function, and in higher dimensions, by the projection to the set Z (in fact, the update function is the projection to an interval). In order to calculate the projection, we used polar coordinates. The curves corresponding to those approximations are called StringStop 1-4, each one for different time discretization. The curve denoted by Projectioninspace is the approximation through the scheme presented in this section applied in two dimensions. We used a regular time discretization, which leads then to the problem of minimizing

$$
\sum_{k=1}^{n} \Delta t \left\| \frac{v_k - v_{k-1}}{\Delta t} - \frac{y_k - y_{k-1}}{\Delta t} \right\|^2 = \frac{1}{\Delta t} \sum_{k=1}^{n} \left\| (v_k - v_{k-1}) - (y_k - y_{k-1}) \right\|^2,
$$

over all collections of two dimensional vectors $v_0, \ldots, v_n$*. The norm is the euclidian one. Together with the conditions* $(S1)$ *and* $(S3)$*, this problem may be written in the form*

$$\min \begin{pmatrix} v_0 \\ | \\ v_n \end{pmatrix}^T A_M \begin{pmatrix} v_0 \\ | \\ v_n \end{pmatrix} - b_M^T \begin{pmatrix} v_0 \\ | \\ v_n \end{pmatrix} + c_M,$$

$$\text{w.r.t.} \begin{cases} v_1^T v_1 & \leq 1, \\ & \vdots \\ v_n^T v_n & \leq 1, \\ v_0 & = 0, \end{cases}$$

where A_M *is a matrix,* b_M *a vector, and* c_M *is scalar. Then, we used the Matlab solver fmincon to solve this problem. Note also that by the procedure presented in this section, we have to solve such a minimization problem for every time point.*

The following figure 8 shows the resulting curves. It seems that they divide whenever the boundary is reached. Hence, the multidimensional problem seems to be more complex.

□

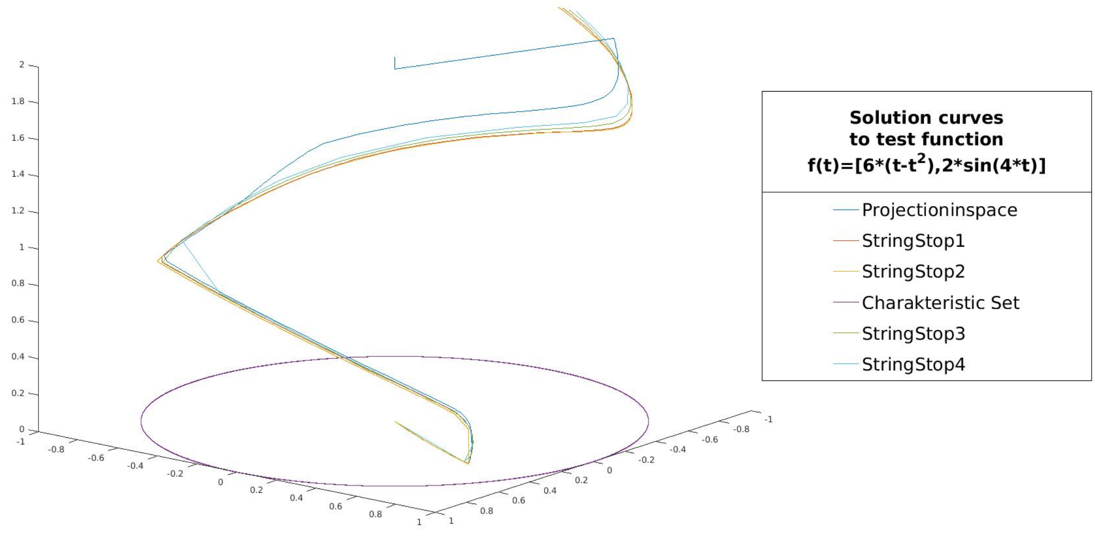

Figure 8: Numerical approximation of Stop operator

PART TWO

3 The heat equation with hysteresis

In this section we present the existence and uniqueness results from [11, chapter X.1] for the heat equation with hysteresis. After that, we give some additional regularity results for solutions in the case of regular initial values. Then, we use our approximation result for the play operator from section 2.4, to find an approximation for the solutions of the heat equation, in the case, when we restrict ourselves to pointwise applied play operators. We show that this regularized partial differential equation has in fact better properties than the one containing hysteresis nonlinearity; actually, one can, quite easily, show existence and uniqueness of mild solutions of that equation. This is then used to build a control problem w.r.t. the regularized equation, and known results concerning the dynamic programming method are applied. As mild and weak solutions agree for regular initial values, we then turn to showing convergence of weak solutions of the regularized equation to solutions of the original one, and use this to prove pointwise convergence of the optimal value function of the control problem.

3.1 Weak solutions and regularity results

In this section, we present the existence and uniqueness proof of Visintin based on time discretization (cf. [11, ch. X.1, thm. 1.1]; in parts of the proof, we also follow [12, thm. 3.3.2]) for the heat equation with hysteresis,

$$\begin{aligned}
&\frac{\partial y}{\partial t} + w - \Delta y = f, \text{ in } \Omega, \\
&w(x,\cdot) = \mathcal{W}[y(x,\cdot),x](\cdot), \ x \in \Omega, \\
&y(x,t) = 0, \text{ on } \partial\Omega \times (0,T), \\
&y(x,0) = y_0(x), \ x \in \Omega.
\end{aligned} \tag{3.1}$$

The hysteresis operator may depend on $x \in \Omega$; usually, this is due to different initial values, but one might also think of other influences. Throughout this section, we will always make the following assumptions.

Assumption 3.1

- $\Omega \subset \mathbb{R}^n$ *is some open, bounded domain with smooth boundary (i.e., at least* C^2*). Let* T *denote the endtime. For* $t \in (0,T]$*, we set* $\Omega_t := \Omega \times (0,T)$*.*
- *For all* $x \in \Omega$*,* $\mathcal{W}[\cdot;x]$ *is continuous on* $C[0,T]$ *and piecewise monotonic, and the parametrised final value mapping*

$$(s,x) \mapsto \mathcal{W}_f(s,x), \ s = (v_0,\dots,v_M) \in S,$$

is measurable for all $M \in \mathbb{N}$*, and fulfills*

$$|\mathcal{W}_f(s,x)| \le c_0(x) + c_1 \|s\|_\infty ,$$

for some $c_0 \in L^2(\Omega)$*,* $c_1 > 0$*, independent of* s *and* M*; additionally,* c_1 *is assumed to be independent of* $x \in \Omega$*.*

□

Remark 3.2

- *By assumption 3.1, there exists* $L > 0$ *and* $g \in L^2(\Omega)$*, such that, for all* $v \in \mathcal{M}(\Omega; C[0,T])$*,*

$$\|\mathcal{W}[v(x,\cdot),x]\|_{C[0,T]} \le L \|v(x,\cdot)\|_{C[0,T]} + g(x), \text{ a.e. in } \Omega.$$

- *We will show that there exists a solution* y *of* (3.1) *in the space*

$$Y := L^\infty(0,T; H_0^1(\Omega)) \cap H^1(0,T; L^2(\Omega)).$$

- *For* $s \in (0, \frac{1}{2})$*, there is the chain of continuous imbeddings (cf. [11])*

$$Y \hookrightarrow H^1(\Omega_T) \hookrightarrow H^s(\Omega; H^{1-s}(0,T)) \hookrightarrow L^2(\Omega; C[0,T]).$$

  Moreover, the last imbedding is compact, so that Y *is compactly imbedded in* $L^2(\Omega; C[0,T])$ *(this can be proved, e.g., with arguments from interpolation theory; we refer to the works [22, 23]).*

- *The conditions on the hysteresis operator are, e.g., fulfilled by Preisach operators (cf. [12, remark 3.3.1]); it is clear that the play operator has these properties.*

□

Theorem 3.3 (Existence of solutions) *Let* $y_0 \in H_0^1(\Omega)$*,* $w_0 \in L^2(\Omega)$ *and* $f \in L^2(\Omega_T)$*. If assumption 3.1 holds, then there exists a weak solution* (y,w) *to* (3.1)*, in the sense that*

$$\int_0^T \int_\Omega (y_t(x,t) + w(x,t))\, \varphi(x,t) dx dt + \int_0^T \int_\Omega \nabla y(x,t) \cdot \nabla \varphi(x,t) dx dt$$
$$= \int_0^T \int_\Omega f(x,t) \varphi(x,t) dx dt, \quad \forall \varphi \in L^2(0,T; H_0^1(\Omega)),$$

and the other equalities of (3.1) *hold pointwise almost everywhere. Moreover,*

$$y \in Y, \quad w \in L^2(\Omega; C[0,T]).$$

Proof: Via time discretization (we follow [12, thm. 3.3.2] for the solution of the time discretized problem). For $M \in \mathbb{N}$, we divide the interval $[0, T]$ into M parts of length $h := \frac{T}{M}$. For the rest of the proof, let C_i denote positive constants that may depend on Ω, T, f and the initial value, but not on $m \in \{1, \ldots, M\}$.

Now, for such m, consider the semidiscrete problem at timestep $t = mh$, for the unknowns $y^m, w^m : \Omega \to \mathbb{R}$,

$$\begin{aligned} \frac{1}{h}\int_\Omega \left(y^m - y^{m-1}\right)\varphi dx + \int_\Omega w^m \varphi dx + \int_\Omega \nabla y^m \cdot \nabla \varphi dx \\ = \int_\Omega f^m \varphi dx, \qquad \forall \varphi \in H_0^1(\Omega), \\ w^m(x) = \mathcal{W}_f\left(\left(y^0(x), \ldots, y^m(x)\right); x\right), \text{ a.e. } x \in \Omega, \end{aligned} \tag{3.2}$$

and

$$f^m(x) := \frac{1}{h}\int_{(m-1)h}^{mh} f(x,t)dt, \qquad y^0(x) := y_0(x),$$
$$w^0(x) := \mathcal{W}_f\left(\left(y^0(x), x\right)\right).$$

One can rewrite the semilinear variational equality in the form

$$a(y^m, \varphi) + \int_\Omega b^m(x, y^m(x))\varphi(x)dx = 0, \quad \forall \varphi \in H_0^1(\Omega),$$

where $a(\cdot,\cdot)$ denotes the scalar product in $H_0^1(\Omega)$, and the function $b^m : \Omega \times \mathbb{R} \to \mathbb{R}$ is defined as

$$b^m(x,y) := \frac{1}{h}y + \mathcal{W}_f\left(\left(y^0(x), \ldots, y^{m-1}(x), y\right), x\right) - \frac{1}{h}y^{m-1}(x) - f^m(x).$$

This function is, by assumption 3.1, measurable w.r.t. x, continuous w.r.t. y, and $b^m(x,\cdot)$ is strictly monotonic increasing for all $x \in \Omega$ by the piecewise monotonicity assumption on $\mathcal{W}$. Further, an inequality of the form

$$|b^m(x,y)| \le c_1^m(x) + c_2\left(\sum_{k=0}^{m-1}\left|y^k(x)\right|\right) + c_3|y|$$

holds, where c_2, c_3 are positive constants and $c_1^m \in L^2(\Omega)$. Application of [12, theorem 1.3.2] yields then the existence of unique $y^m \in H_0^1(\Omega), w^m \in L^2(\Omega)$. By iteration, we thus get a solution for the discretized problem.

Our next goal is to derive suitable a priori estimates for the solutions of the discretized problem (here, we follow [11, ch. X.1, thm. 1.1]). As all y^m are elements of

$H_0^1(\Omega)$, we may test the variational equality with $\varphi = y^m - y^{m-1}$. Summation over $m \in \{1, \ldots, k\}$, $k \in \{1, \ldots, M\}$ then leads to

$$\sum_m h \int_\Omega \left(\frac{y^m - y^{m-1}}{h}\right)^2 dx + \sum_m \int_\Omega \nabla y^m \cdot \nabla(y^m - y^{m-1}) dx$$
$$\leq \sum_m \int_\Omega \left|f^m(y^m - y^{m-1})\right| dx + \sum_m \int_\Omega \left|w^m(y^m - y^{m-1})\right| dx.$$

With Young's inequality, the right hand side may be estimated by

$$\frac{1}{2\alpha} \sum_m h \int_\Omega (f^m)^2 dx + \frac{1+\alpha}{2} \sum_m \int_\Omega \frac{(y^m - y^{m-1})^2}{h} dx + \frac{1}{2} \sum_m h \int_\Omega (w^m)^2 dx;$$

we choose $0 < \alpha < 1$. Further,

$$\sum_m h \int_\Omega (f^m)^2 dx \leq \int_0^T \int_\Omega |f(x,t)|^2 \, dx dt.$$

We still need to estimate the term $\sum \int (w^m)^2 dx$. To this end, define

$$z^m := \left|y^0\right| + \sum_{j=1}^m \left|y^j - y^{j-1}\right|, \; m > 1, \; z^0 := \left|y^0\right|.$$

Assumption 3.1 allows us then to estimate

$$|w^m| \leq L \max_{1 \leq j \leq m} \left|y^j\right| + g \leq L z^m + g, \quad \text{a.e. in } \Omega.$$

Noting that

$$z^m - z^{m-1} = \left|y^m - y^{m-1}\right|,$$

as well as

$$\sum_m \int_\Omega \nabla y^m \cdot \nabla(y^m - y^{m-1}) dx$$
$$= \frac{1}{2} \sum_m \left\|\nabla y^m - \nabla y^{m-1}\right\|^2 + \frac{1}{2} \left\|\nabla y^k\right\|^2 - \frac{1}{2} \left\|\nabla y^0\right\|^2,$$

we thus derive

$$\sum_m \int_\Omega h \left(\frac{z^m - z^{m-1}}{h}\right)^2 dx + \sum_m \left\|\nabla y^m - \nabla y^{m-1}\right\|^2 + \left\|\nabla y^k\right\|^2$$

$$\leq c_3 \sum_m h \int_\Omega (z^m)^2 dx + c_4.$$

Thanks to

$$\left\| z^k \right\| - \left\| y^0 \right\| \leq \sqrt{T} \left(\sum_m h \left\| \frac{z^m - z^{m-1}}{h} \right\|^2 \right)^{\frac{1}{2}},$$

we can get to

$$\left\| z^k \right\| - \left\| y^0 \right\| \leq c_5 \left(h \sum_{m=1}^{k} \left\| z^m \right\|^2 \right)^{\frac{1}{2}} + c_6.$$

Now, one can use Gronwall's inequality, which implies

$$\left\| z^k \right\| \leq c_7, \quad \forall k \in \{1, \ldots, M\}, \ \forall M \in \mathbb{N}.$$

and therefore

$$\begin{gathered} \sum_{m=1}^{M} h \left\| \frac{y^m - y^{m-1}}{h} \right\|^2 + \max_{1 \leq k \leq M} \left\| \nabla y^k \right\|^2 + \sum_{m=1}^{M} \left\| \nabla y^m - \nabla y^{m-1} \right\|^2 \leq c_8, \\ \sum_{m=1}^{M} h \int_\Omega \left| w^m(x) \right|^2 dx \leq c_9. \end{gathered} \tag{3.3}$$

The rest of the proof deals with the passage to the limit, $M \to \infty$ (we follow again [12, thm. 3.3.2]). To this end, we introduce, for $M \in \mathbb{N}$ and $m \in \{1, \ldots, M\}$, the notation y_M^m, w_M^m, f_M^m, and piecewise linear interpolates

$$\begin{aligned} y_M(x, (m+\tau)h) &:= \tau y_M^{m+1}(x) + (1-\tau) y_M^m(x), \quad \tau \in [0,1], \\ w_M(x, (m+\tau)h) &:= \tau w_M^{m+1}(x) + (1-\tau) w_M^m(x), \quad \tau \in [0,1], \end{aligned}$$

as well as the piecewise constant interpolates

$$\begin{aligned} \tilde{y}_M(x, (m+\tau)h) &:= y_M^{m+1}(x), \quad \tau \in (0,1], \\ \tilde{w}_M(x, (m+\tau)h) &:= w_M^{m+1}(x), \quad \tau \in (0,1], \\ \tilde{f}_M(x, (m+\tau)h) &:= f_M^{m+1}(x), \quad \tau \in (0,1]. \end{aligned}$$

Due to those definitions, the variational equality reads

$$\int_0^T \int_\Omega (y_{M,t} + \tilde{w}_M) \varphi dx dt + \int_0^T \int_\Omega \nabla \tilde{y}_M \cdot \nabla \varphi dx dt$$

$$= \int_0^T \int_\Omega \tilde{f}_M \varphi dxdt, \quad \forall \varphi \in L^2(0,T;H_0^1(\Omega)),$$

and the a priori estimate (3.3) implies

$$\int_0^T \int_\Omega y_{M,t}^2 dxdt + \sup_{0\le t\le T} \left(\|\nabla \tilde{y}_M(t)\|^2 + \|\nabla y_M(t)\|^2 \right) \le c_{10}.$$

Another application of (3.3) shows

$$\|y_M - \tilde{y}_M\|^2_{L^2(0,T;H_0^1(\Omega))} = \frac{T}{3M} \sum_m \left\| \nabla y_M^{m+1} - \nabla y_M^m \right\|^2 \xrightarrow{M\to\infty} 0, \tag{3.4}$$

and

$$\|w_M\|^2_{L^2(\Omega_T)} \le 2 \|\tilde{w}_M\|^2_{L^2(\Omega_T)} \le 2c_9.$$

Hence, we may, at least for some suitable subsequences (which in that case will again be denoted in the same way), conclude that

$$\begin{aligned} &y_M \to y, \text{ weak} \star \text{ in } Y, \\ &\tilde{y}_M \to \tilde{y}, \text{ weak} \star \text{ in } L^\infty(0,T;H_0^1(\Omega)), \\ &w_M \to w, \text{ weak in } L^2(\Omega_T), \\ &\tilde{w}_M \to \tilde{w}, \text{ weak in } L^2(\Omega_T), \end{aligned}$$

as $M \to \infty$. From (3.4), we infer $y = \tilde{y}$. Further, $\tilde{f}_M \to f$ in $L^2(\Omega)$, hence, we may take the limit of the variational equality, and arrive at

$$\int_0^T \int_\Omega (y_t + \tilde{w}) \varphi dxdt + \int_0^T \int_\Omega \nabla y \cdot \nabla \varphi dxdt = \int_0^T \int_\Omega f \varphi dxdt,$$

for all $\varphi \in L^2(0,T;H_0^1(\Omega))$. We still have to show that $w = \tilde{w} = \mathcal{W}$ holds (a.e.). Let us define the functions

$$\begin{aligned} \lambda_M(x,t) &:= \mathcal{W}[y_M(x,\cdot),x](t), \ M \in \mathbb{N}, \\ \lambda(x,t) &:= \mathcal{W}[y(x,\cdot),x](t). \end{aligned}$$

From the compactness of the imbedding $Y \hookrightarrow L^2(\Omega;C[0,T])$ (cf. remark 3.2 and the reference given there), we infer that $y_M(x,\cdot) \to y(x,\cdot)$ in $C[0,T]$ for almost every $x \in \Omega$. By the continuity assumption on $\mathcal{W}$, also $\lambda_M(x,\cdot) \to \lambda$ in $C[0,T]$, a.e. $x \in \Omega$. Assumption 3.1 makes it possible to estimate

$$\sup_{0\le t\le T} |\lambda_M(x,t)| \le c_0(x) + c_1 \sup_{0\le t\le T} |y_M(x,t)|, \quad \text{a.e. } x \in \Omega,$$

and, since the right hand side converges in $L^2(\Omega)$, we infer that $\lambda_M \to \lambda$ strongly in $L^2(\Omega; C[0,T])$, from dominated convergence. Note that w_M is the piecewise linear interpolate of λ_M, so that one can show analogously $w_M - \lambda_M \to 0$ in $L^2(\Omega; C[0,T])$, hence, $w_M \to w = \lambda$ in $L^2(\Omega; C[0,T])$. In the same way one can show that $w = \tilde{w}$ almost everywhere, which completes the proof.

□

For the uniqueness result, we add a Lipschitz type condition for the hysteresis operator.

Assumption 3.4 *There exists $L > 0$ such that, for every $t \in (0,T]$ and all $v_1, v_2 \in L^2(\Omega; C[0,T])$:*

$$\|\mathcal{W}[v_1;\cdot] - \mathcal{W}[v_2;\cdot]\|_{L^2(\Omega;C[0,t])} \leq L \|v_1 - v_2\|_{L^2(\Omega;C[0,t])}.$$

□

Remark 3.5 *The condition in assumption 3.4 is quite natural for a large group of hysteresis operators. We have seen in section 2.3 that the play operator is one example. Hence, also superpositions of the latter belong to the class of operators for which 3.4 is fulfilled.*

□

The following uniqueness result is due to Visintin (cf., [11, ch. X.1, thm. 1.2]).

Theorem 3.6 *Under assumptions of theorem 3.3 and 3.4, there is exactly one solution to* (3.1).

Proof: By theorem 3.3, there exists at least one solution. Assume now that (y^1, w^1), (y^2, w^2) are, respectively, two solutions of (3.1), i.e.,

$$y_t^i + w^i - \Delta y^i = f, \quad i = 1, 2,$$

holds in the sense of theorem 3.3. Consider the differences $y := y^1 - y^2$, $w := w^1 - w^2$; those then solve the equation

$$y_t + w - \Delta y = 0,$$

in the sense of theorem 3.3 with initial value $y(0) = 0$. Next note that, since $y_t, w \in L^2(\Omega_T)$, the same must hold for Δy, so that the equality holds, in particular, almost everywhere in Ω_T. Hence, we may test with y_t: for arbitrary $\tau \in (0,T)$, we then get

$$\int_{\Omega_\tau} |y_t|^2 \, d\mathcal{L} - \int_{\Omega_\tau} \Delta y y_t d\mathcal{L} \leq \|w\|_{L^2(\Omega_\tau)} \|y_t\|_{L^2(\Omega_\tau)} .$$

Using partial integration (see the next lemma for a proof) and $y(0) = 0$, the latter implies

$$\int_{\Omega_\tau} |y_t|^2 \, d\mathcal{L} + \frac{1}{2} \int_\Omega |\nabla y(x,\tau)|^2 \, dx \leq \|w\|_{L^2(\Omega_\tau)} \|y_t\|_{L^2(\Omega_\tau)} . \tag{3.5}$$

From assumption 3.4, we infer that

$$\|w\|_{L^2(\Omega_\tau)} \leq \sqrt{\int_{\Omega_\tau} \|w(x,\cdot)\|^2_{C[0,\tau]} \, d\mathcal{L}} \leq L\sqrt{\tau} \, \|y\|_{L^2(\Omega;C[0,\tau])} ,$$

and, thanks to Jensen's inequality and $y(x,0) = 0$ almost everywhere,

$$\|y(x,\cdot)\|^2_{C[0,\tau]} \leq \tau \int_0^\tau |y_t(x,s)|^2 \, ds, \quad \text{a.e. } x \in \Omega.$$

Altogether, we thus get

$$\|w\|_{L^2(\Omega_\tau)} \leq L\tau \, \|y_t\|_{L^2(\Omega_\tau)} .$$

Plugging this into (3.5) yields

$$\|y_t\|^2_{L^2(\Omega_\tau)} + \frac{1}{2} \|\nabla y(\cdot,\tau)\|^2_{L^2(\Omega)} \leq L\tau \, \|y_t\|^2_{L^2(\Omega_\tau)} .$$

For $\tau \in (0, \frac{1}{L})$, we infer from this that $y = 0$ almost everywhere in Ω_τ, i.e., $y^1 = y^2$ in Ω_τ, and thereby $w^1 = w^2$. By iteration, we may further conclude $y^1 = y^2$, as well as $w^1 = w^2$ on the whole Ω_T, which completes the proof.

□

In the proof, the following lemma was used with f replaced by $f - w$.

Lemma 3.7 *Let u be an L^2-solution of*

$$u_t - \Delta u = f,$$

i.e., the equation is valid a.e. in Ω_T, and all functions are square integrable. Further, let $u \in Y$, and $\langle \cdot, \cdot \rangle$ denote the scalar product on $L^2(\Omega)$.

1. If $\partial\Omega$ is of class C^1, then $u \in C(0,T;L^2(\Omega))$, and for almost every $t \in (0,T)$,

$$\frac{d}{dt}\|u(t)\|^2_{L^2(\Omega)} = 2\langle u(t), \partial_t u(t)\rangle .$$

2. If $\partial\Omega$ is of class C^2, then $\nabla u \in C(0,T;L^2(\Omega))$, and

$$-\int_0^T\int_\Omega \Delta u(x,s)u_t(x,s)dxds = \frac{1}{2}\left(\|\nabla u(t)\|^2_{L^2(\Omega)} - \|\nabla u(0)\|^2_{L^2(\Omega)}\right).$$

Proof: *1.* For $\Omega \subset\subset V$, there exists a continuation operator E, such that Eu has compact support in V, see [24, 5.4 theorem 1]. Since $u \in Y$, because of the Gelfand triple $H_0^1(V) \hookrightarrow L^2(V) \hookrightarrow H^{-1}(V)$, we get the regularities $Eu \in L^2(0,T;H_0^1(V))$, $\partial_t Eu \in L^2(0,T;L^2(V)) \subset L^2(0,T;H^{-1}(V))$. The well known interpolation theorem (cf. [24, 5.9 theorem 3]) yields then $Eu \in C(0,T;L^2(V))$, which implies $u \in C(0,T;L^2(\Omega))$. In order to prove the claimed identity, it suffices to note that $u, u_t \in L^2(0,T;L^2(\Omega))$, because then

$$\frac{d}{dt}\langle u(t), u(t)\rangle = \frac{d}{dt}\left\langle u_0 + \int_0^t \dot{u}(r)dr, u_0 + \int_0^t \dot{u}(s)ds\right\rangle = 2\langle u(t), \partial_t u(t)\rangle, \quad \text{a.e.}$$

2. As $\partial\Omega$ is C^2, results from regularity theory of elliptic partial differential equations imply that $u(\cdot,t) \in H^2(\Omega)$, for almost every t [24], and thus, by measurability and integrability also $u \in L^2(0,T;H^2(\Omega))$. Since $u_t \in L^2(0,T;L^2(\Omega))$, similar to part *1.*, one can show that $u \in C(0,T;H^1(\Omega))$, such that $\nabla u \in C(0,T;L^2(\Omega))$. Next, consider the trivial continuation (i.e. by zero) of u w.r.t. t on $\mathbb{R}$, and define $u_\varepsilon := \eta_\varepsilon * u$, the convolution of u with the standard mollifier. Then, for all t,

$$\begin{aligned}\frac{d}{dt}\|\nabla u_\varepsilon(t)\|^2_{L^2(\Omega)} =& 2\langle \nabla u_\varepsilon(t), \nabla u_{\varepsilon,t}(t)\rangle = 2\int_\Omega \nabla u_\varepsilon(t)\cdot\nabla u_{\varepsilon,t}(t)dx \\ =& -2\int_\Omega \Delta u_\varepsilon(t)u_{\varepsilon,t}(t)dx + 2\int_{\partial\Omega}\nabla u_\varepsilon(t)u_{\varepsilon,t}(t)dS(x). \end{aligned} \tag{3.6}$$

$u_\varepsilon \in H^2(\Omega)$ implies that the restriction of ∇u_ε to $\partial\Omega$ is an $L^2(\partial\Omega)$ function, and $u_{\varepsilon,t} = \eta_{\varepsilon,t} * u = \eta_{\varepsilon,t} * 0 = 0$ on $\partial\Omega$, because $u(t) \in H_0^1(\Omega)$ for almost every t. Hence,

$$\int_{\partial\Omega}\nabla u_\varepsilon(t)u_{\varepsilon,t}(t)dS(x) = 0, \quad \text{for a.e. } t \in (0,T).$$

Further, $\Delta u, u_t \in L^2(\Omega_T)$ implies

$$\int_\Omega \Delta u_\varepsilon(t)u_{\varepsilon,t}(t)dx \xrightarrow{\varepsilon\downarrow 0} \int_\Omega \Delta u(t)u_t(t)dx,$$

for a.e. t. Thus, integrating (3.6) w.r.t. t, and then letting $\varepsilon \downarrow 0$ yields the desired equality. □

At the end of this chapter, we collect some regularity results.

Proposition 3.8 *Under assumption 3.1, each solution y of* (3.1) *corresponding to initial values $(y_0, w_0) \in H_0^1(\Omega) \times L^2(\Omega)$ has the regularity*

$$y \in Y \cap C(0,T;H^1(\Omega)), \quad w \in H^1(0,T;L^2(\Omega)).$$

Further, consider the special case $\mathcal{W}[\cdot, x] = \mathcal{F}_{r(x)}[\,\cdot\,; w_0(x)]$ with $r \in H^1(\Omega)$, and let $1 < p \leq \infty$, $f \in W^{1,1}(0,T;L^p(\Omega))$, $w_0, \Delta y_0 \in L^p(\Omega)$. If $p < \infty$, then

$$y \in W^{1,1}(0,T;L^p(\Omega)) \cap L^\infty(0,T;W^{2,p}(\Omega)), w \in L^p(\Omega;C[0,T]).$$

If, in addition, $p > n$ holds, then we also have $w \in L^\infty(0,T;W^{1,\infty}(\Omega))$.

Proof: The first part of the proposition is shown in the proofs of theorem 3.3 and lemma 3.7. For the second part, we refer to [11, Prop. 1.3, page 301], where the result is given by refinements of the a priori estimates of the time discrete problem. Especially, it is shown there, that for the solution of the discretized problem it holds $\|\Delta y_j\|_{L^p(\Omega)} \leq C$ for some positive constant C, uniformly in j and the length of the time steps. Hence, if $p > n$, we infer from Sobolev's inequality [18, theorem 4.12] that ∇y_j is uniformly bounded in $L^\infty(\Omega)$, so that theorem 2.26 applies, which yields $\nabla w \in L^\infty(\Omega_T)$.

□

3.2 Heat equation with regularized play operator

Our next goal is to find an approximation of solutions of (3.1) in the special case when the hysteresis operator is a pointwise applied play operator, i.e., $\mathcal{W}[\cdot, x] = \mathcal{F}_{r(x)}[\,\cdot\,; w_0(x)]$. To this end, we use our approximation result from section 2.3, and show that the weak solution of the regularized problem

$$\begin{aligned}
&\frac{\partial y}{\partial t} + z - \Delta y = f, \text{ in } \Omega, \\
&\dot{z}(x,t) = \frac{1}{\varepsilon} G\left(z(x,t) - y(x,t)\right), \ x \in \Omega, \\
&y(x,t) = 0, \text{ on } \partial\Omega \times (0,T), \\
&y(x,0) = y_0(x), \ x \in \Omega, \\
&z(x,0) = z_0(x) \in \left[y_0(x) - r(x), y_0(x) + r(x)\right], \ x \in \Omega,
\end{aligned} \tag{3.7}$$

converges, for $\varepsilon \downarrow 0$, to the weak solution of (3.1). Note that the parameter $r = r(x)$ may depend on the space variable, so that the same actually holds for G, too. So, with G we always mean a function of the form (2.2), where we allow of

$r \in L^2(\Omega)$ whenever G is applied to functions in $L^2(\Omega; C[0,T])$. Further, we will discuss existence and uniqueness of mild solutions to (3.7). The proofs will not follow the discretization method used in theorem 3.3, but the (simpler) method suggested in a remark of [11, page 300], which is based on the Lipschitz type property of the hysteresis operator, cf. assumption 3.4. We start with collecting some additional properties of the approximation operator.

Lemma 3.9 *Let P denote the solution operator*

$$P : H^1(0,T) \to H^1(0,T),\ P(y) = z = z_0 + \frac{1}{\varepsilon}\int_0^{\cdot} G(z-y)d\mathcal{L}.$$

For every pair of functions $v_1, v_2 \in H^1(0,T)$ and $t \in (0,T]$, it holds

$$\|P(v_1) - P(v_2)\|_{C[0,t]} \le \frac{T}{\varepsilon}e^{\frac{T}{\varepsilon}} \|v_1 - v_2\|_{C[0,t]}.$$

Hence, if $v_1, v_2 \in L^2(\Omega; H^1(0,T))$, we get the inequality

$$\|P(v_1) - P(v_2)\|_{L^2(\Omega;C[0,t])} \le \frac{T}{\varepsilon}e^{\frac{T}{\varepsilon}} \|v_1 - v_2\|_{L^2(\Omega;C[0,t])};$$

moreover, if $v_1(0) = v_2(0)$, then it holds

$$\|P(v_1) - P(v_2)\|_{L^2(\Omega_t)} \le \frac{2t^2}{\varepsilon} \|\dot{v}_1 - \dot{v}_2\|_{L^2(\Omega_t)},$$

for all $t \in [0, \min\{\frac{\varepsilon}{2}, T\}]$.

Proof: Recall that G is Lipschitz continuous with constant $L = 1$. Then, for $t \in (0,T]$,

$$\begin{aligned}|P(v_1)(t) - P(v_2)(t)| \le& \frac{1}{\varepsilon}\int_0^t |G(P(v_1)(s) - v_1(s)) - G(P(v_2)(s) - v_2(s))|\,ds \\ \le& \frac{1}{\varepsilon}\int_0^t |P(v_1)(s) - v_1(s) - P(v_2)(s) + v_2(s)|\,ds \\ \le& \frac{T}{\varepsilon}\|v_1 - v_2\|_{C[0,t]} + \frac{1}{\varepsilon}\int_0^t |P(v_1)(s) - P(v_2)(s)|\,ds.\end{aligned}$$

From Gronwall's inequality, we infer, for any $\tau \in (0,t)$

$$|P(v_1)(\tau) - P(v_2)(\tau)| \le \frac{T}{\varepsilon}e^{\frac{T}{\varepsilon}} \|v_1 - v_2\|_{C[0,t]},$$

so the claimed inequality follows by application of $\sup_{\tau\in(0,t)}$. The second part follows directly from the first one. In order to prove the last inequality, note that for almost every $x \in \Omega$,

$$\varepsilon\left(\dot{z}_1 - \dot{z}_2\right) \leq \left|G(z_1 - v_1) - G(z_2 - v_2)\right| \leq |z_1 - z_2| + |v_1 - v_2|\,.$$

Denoting, for short, $Z := z_1 - z_2,\ V := v_1 - v_2$, integration of the latter implies, due to $Z(0) = 0$,

$$\varepsilon Z(x,t) \leq \int_0^t |Z(x,s)| + |V(x,s)|\,ds,$$

for all $t \in [0,T]$, almost everywhere in Ω. The same holds if we interchange the roles of v_1 and v_2, so that

$$\varepsilon\left|Z(x,t)\right| \leq \int_0^t |Z(x,s)| + |V(x,s)|\,ds, \quad \text{a.e. } x \in \Omega.$$

As $Z, V \in L^2(\Omega; C[0,T])$, this implies

$$\varepsilon\left\|Z(x,\cdot)\right\|_{C[0,t]} \leq t\left\|Z(x,\cdot)\right\|_{C[0,t]} + \int_0^t |V(x,s)|\,ds, \quad \text{a.e. } x \in \Omega.$$

So, restricting ourselves to the interval $t \in [0, \min\left\{\frac{\varepsilon}{2}, T\right\}] =: I$, we get

$$\left\|Z(x,\cdot)\right\|_{C[0,t]} \leq \frac{2}{\varepsilon}\int_0^t |V(x,s)|\,ds, \quad \text{a.e. } x \in \Omega$$

$$\Rightarrow \left\|Z\right\|^2_{L^2(\Omega;C[0,t])} \leq \frac{4}{\varepsilon^2} t \left\|V\right\|^2_{L^2(\Omega_t)}\,.$$

Hence, for all $t \in I$, it holds (thanks to $V(0) = 0$),

$$\begin{aligned}\left\|Z\right\|^2_{L^2(\Omega_t)} &\leq t\left\|Z\right\|^2_{L^2(\Omega;C[0,t])}\\ &\leq \frac{4t^2}{\varepsilon^2}\left\|V\right\|^2_{L^2(\Omega_t)}\\ &\leq \frac{4t^4}{\varepsilon^2}\left\|\dot{V}\right\|^2_{L^2(\Omega_t)},\end{aligned}$$

and we can conclude by taking the square root.

□

Theorem 3.10 (Existence & uniqueness of weak solutions for (3.7)) *Let $\Omega \subset \mathbb{R}^n$ be an open, bounded domain with smooth boundary, $y_0 \in H_0^1(\Omega)$, $z_0 \in L^2(\Omega)$ and $f \in L^2(\Omega_T)$. Then there exists a unique weak solution to (3.7) in the sense that*

$$\int_0^T \int_\Omega (y_t(x,t) + z(x,t))\,\varphi(x,t)dxdt + \int_0^T \int_\Omega \nabla y(x,t) \cdot \nabla \varphi(x,t)dxdt$$
$$= \int_0^T \int_\Omega f(x,t)\varphi(x,t)dxdt, \quad \forall \varphi \in L^2(\Omega; H_0^1(\Omega)),$$

and the other equations from (3.7) hold almost everywhere. Further, the solution has the regularity

$$y \in Y, \quad z \in L^2(\Omega; H^1(0,T)).$$

Proof: We reformulate the problem as fixed point problem and show that the solution operator is (locally) a contraction map. To this end, let $v_1, v_2 \in L^2(\Omega; H^1(0,T))$ be arbitrary, and define, for $i \in \{1,2\}$, z_i as the unique solution of

$$z_i(x,t) := z_0(x) + \frac{1}{\varepsilon}\int_0^t G\left(z_i(x,s) - v_i(x,s)\right) ds.$$

From lemma 2.8, we infer that the solution exists pointwise for almost every $x \in \Omega$, and by the assumption on v_i, we also get $z_i \in L^2(\Omega; H^1(0,T))$, so that the operator

$$P : L^2(\Omega; H^1(0,T)) \to L^2(\Omega; H^1(0,T)),$$
$$P(v)(x,t) = z_0(x) + \int_0^t G\left(P(v)(x,s) - v(x,s)\right) ds,$$

is well defined. We may then write $z_i = P(v_i)$, and lemma 3.9 applies to this operator. Further, for all $v \in L^2(\Omega; H^1(0,T))$, there exists a unique weak solution to

$$\int_0^T \int_\Omega (y_t(x,t) + P(v)(x,t))\,\varphi(x,t)dxdt + \int_0^T \int_\Omega \nabla y(x,t) \cdot \nabla \varphi(x,t)dxdt$$
$$= \int_0^T \int_\Omega f(x,t)\varphi(x,t)dxdt, \quad \forall \varphi \in L^2(\Omega; H_0^1(\Omega)),$$

with initial and boundary values as in (3.7) (as this is just a simple heat equation). Since the solution thereof has particularly the regularity $y \in L^2(\Omega; H^1(0,T))$, we may define the solution operator $S : L^2(\Omega; H^1(0,T)) \to L^2(\Omega; H^1(0,T))$, which maps v to the solution of that pde. Let us show that, at least for small $T > 0$, the operator S is a contraction mapping on $L^2(\Omega; H^1(0,T))$. By the standard improved regularity result for the heat equation, $y_i := S(v_i) \in Y$ must be an $L^2(\Omega_T)$ solution. Then, the difference $y := y_1 - y_2$ fulfills

$$y_t + P(v_1) - P(v_2) + \Delta y = 0,$$

almost everywhere. Testing with $y_t \cdot \chi_{(0,\tau)}$, and applying lemma 3.7 yields (note $y_0 = 0$)

$$\|y_t(\tau)\|^2_{L^2(\Omega_\tau)} + \frac{1}{2}\|\nabla y(\tau)\|^2_{L^2(\Omega)} \leq \frac{1}{2}\|y_t(\tau)\|^2_{L^2(\Omega_\tau)} + \frac{1}{2}\int_0^\tau \int_\Omega |P(v_1)(x,t) - P(v_2)(x,t)|^2 \, dxdt$$

Then, together with lemma 3.9 (note that we have to restrict ourselves to the interval $\tau \in I := (0, \min\left\{\frac{\varepsilon}{2}, T\right\}])$, if $v_1(0) = v_2(0)$, the latter inequality yields

$$\|y_t(\tau)\|^2_{L^2(\Omega_\tau)} + \|\nabla y(\tau)\|^2_{L^2(\Omega)} \leq \frac{4\tau^4}{\varepsilon^2}\|\dot{v}_1 - \dot{v}_2\|^2_{L^2(\Omega_\tau)}.$$

With Poincaré's inequality [18, theorem 6.30], we may thus find, for each $\tau \in I$, some constant $c = c(\tau)$ such that $c(\tau) \downarrow 0$ as $\tau \downarrow 0$, and

$$\|S(v_1) - S(v_2)\|_{L^2(\Omega;H^1(0,\tau))} \leq c(\tau)\|v_1 - v_2\|_{L^2(\Omega;H^1(0,\tau))}.$$

So, if $\tau > 0$ is small enough, then $c(\tau) < \frac{1}{2}$, and S forms a contraction on the set of functions in $L^2(\Omega; H^1(0,\tau))$ with initial values $v(0) = y_0$. We therefore get, by classical arguments, a unique solution to the fixed point problem, which is then also the unique solution to the system (3.7) for small $\tau \in I$. By iteration, the solution may be continued to the whole interval; the higher regularity $y \in Y$ carries over from the regularity of the approximations to which Banach's fixed point theorem is applied, by continuous dependence on the data.

□

Next we show that (3.7) has the nice property, that one can quite easily establish mild solutions. At this point, recall that the Laplacian with domain $H^1_0(\Omega) \cap H^2(\Omega)$ generates an analytic semigroup on $L^2(\Omega)$, if $\Omega \subset \mathbb{R}^n$ is open and bounded with smooth boundary (cf. [25, chapter 3.1.1]). The corresponding one parameter semigroup will be denoted by $e^{t\Delta}$. Note also, that the conditions of the Caratheodory existence and uniqueness results for z are still fulfilled, if y is merely in $L^1(0,T)$.

Theorem 3.11 (Existence & uniqueness of mild solutions for (3.7)**)** *Let $\Omega \subset \mathbb{R}^n$ be some open, bounded domain with smooth boundary, $y_0, z_0 \in L^2(\Omega)$,*

$f \in L^2(\Omega_T)$. *There exists a unique mild solution to* (3.7), *i.e.,* $y \in C(0,T;L^2(\Omega))$ *and* $z \in H^1(0,T;L^2(\Omega))$ *which satisfy the* L^2 *integral equation*

$$y(t) = e^{t\Delta}y_0 + \int_0^t e^{(t-s)\Delta}\left(f(s) - z(s)\right)ds,$$
$$z(t) = z_0 + \frac{1}{\varepsilon}\int_0^t G(z(s) - y(s))ds.$$

Proof: Again we rewrite the problem in the form of some fixed point problem. For any $v \in L^2(0,T;L^2(\Omega)) = L^2(\Omega;L^2(0,T))$, denote by z_v the solution of

$$z_v(t) = z_0 + \frac{1}{\varepsilon}\int_0^t G(z_v(s) - v(s))ds \in H^1(0,T;L^2(\Omega)).$$

Then, in particular, the map

$$S : C(0,T;L^2(\Omega)) \to C(0,T;L^2(\Omega)),$$
$$S(v)(t) = e^{t\Delta}y_0 + \int_0^t e^{(t-s)\Delta}\left(f(s) - z_v(s)\right)ds$$

is well defined and satisfies (recall that $e^{t\Delta}$ is a semigroup of contractions)

$$\|S(v_1)(t) - S(v_2)(t)\| \leq \int_0^t \|z_{v_1}(s) - z_{v_2}(s)\|\, ds. \tag{3.8}$$

where $\|\cdot\| := \|\cdot\|_{L^2(\Omega)}$. Further, a simple calculation shows

$$\begin{aligned}\|z_{v_1}(t) - z_{v_2}(t)\|^2 \leq& \int_\Omega \left(\frac{1}{\varepsilon}\int_0^t |G(z_{v_1}(s) - v_1(s)) - G(z_{v_2}(s) - v_2(s))|\, ds\right)^2 dx\\ \leq& \frac{t}{\varepsilon^2}\int_\Omega\int_0^t 2\,|z_{v_1}(s) - z_{v_2}(s)|^2 + 2\,|v_1(s) - v_2(s)|^2\, ds\\ =& \frac{2t}{\varepsilon^2}\|v_1 - v_2\|^2_{L^2(\Omega_t)} + \frac{2t}{\varepsilon^2}\int_0^t \|z_{v_1}(s) - z_{v_2}(s)\|^2\, ds.\end{aligned}$$

We apply a version of Gronwall's lemma (cf. D.3), which implies

$$\begin{aligned}\|z_{v_1}(t) - z_{v_2}(t)\|^2 \leq& \frac{2t}{\varepsilon^2}\|v_1 - v_2\|^2_{L^2(\Omega_t)}\left(1 + \frac{2t}{\varepsilon^2}\int_0^t \exp\left\{\int_s^t 2\tau d\tau\right\}ds\right)\\ \leq& 2ct\,\|v_1 - v_2\|^2_{L^2(\Omega_t)},\end{aligned}$$

with $c = \frac{1}{\varepsilon^2}\left(1 + \frac{2T^2}{\varepsilon^2}\exp(2T^2)\right)$. Due to that inequality, we infer from (3.8),

$$\|S(v_1)(t) - S(v_2)(t)\| \leq \sqrt{2ct^3}\,\|v_1 - v_2\|_{L^2(\Omega_t)}$$

$$\leq \sqrt{2ct^4} \, \|v_1 - v_2\|_{C(0,T;L^2(\Omega))} \, ,$$

and we have a contraction mapping on $C(0, T; L^2(\Omega))$, if $T > 0$ is small enough. Thus there exists locally a unique solution, which can be continued to the whole interval $[0, T]$, because z_v is bounded whenever v is.

□

In order to show that weak and mild solutions coincide whenever $y_0 \in H_0^1(\Omega)$, we use the following results.

Theorem 3.12 (cf. [26, proposition 3.8, p. 145, corollary 3.1, p. 145]) *Let H be a Hilbert space and $(A, D(A))$ be the generator of an analytic semigroup on H. Assume that $f \in L^2(0, T; H)$ and $y_0 \in H$. Then there exists a unique classical solution in $L^2(0, T; H)$ to the problem*

$$\dot{y}(t) = Ay(t) + f, \quad t \in [0, T],$$
$$y(0) = y_0,$$

which is given by the mild solutions formula. Moreover, this solution is an element of $W^{1,2}(0, T; H) \cap L^2(0, T; D(A))$ (i.e., a strict solution), if and only if $y_0 \in D_A(\frac{1}{2}, 2)$.

□

Theorem 3.13 (cf. [26, p. 169, 170]) $D_A(\frac{1}{2}, 2) \cong D((-A)^{\frac{1}{2}})$.

Proof: Combine equation (6.4) on page 169 with theorem 6.1(i) on page 170 of [26].

□

In our case, since we have a solution y, we may apply theorem 3.12 with f replaced by $f - z_y \in L^2(0, T; L^2(\Omega))$, so that the mild solution y is in fact strict, if $y_0 \in D_A(\frac{1}{2}, 2)$. As this space is isometrically isomorphic to $D((-A)^{\frac{1}{2}})$ by theorem 3.13, it suffices to characterize this space. But in the case $A = \Delta$, $D(A) = H_0^1(\Omega) \cap H^2(\Omega)$, it holds

$$\left\| (-\Delta)^{\frac{1}{2}} x \right\|^2 = \langle -\Delta x, x \rangle = \int_\Omega |\nabla x|^2 \, d\mathcal{L},$$

for all $x \in D(A)$. Hence, $D((-\Delta)^{\frac{1}{2}}) = H_0^1(\Omega)$ here. But, by the standard higher regularity result for the heat equation, also the weak solution from theorem 3.10 is strict. Hence, it suffices to show:

Proposition 3.14 *There is at most one strict solution to (3.7) in $L^2(\Omega_T)$.*

Proof: Let y_1, y_2 be two strict solutions. Testing the difference of the (almost everywhere valid) equations with $\dot{y}_1 - \dot{y}_2$, we may infer from lemma 3.9 that $y_1 = y_2$ on Ω_τ, for $\tau \in I = (0, \min\left\{\sqrt{\frac{\varepsilon}{2}}, T\right\})$. The claim follows then by iteration.

□

3.3 A control problem corresponding to the dynamics of the regularized partial differential equation

Our next goal is to apply theorems from [2, chapter 6] about dynamic programming method in infinite dimensions. Their results are close to the ones of [5], who did pioneer work in that area. In this section, we state a control problem that fits the assumptions of [2, chapter 6,§ 6]. We reformulate (3.7) as integral problem

$$\begin{pmatrix} y_x(t) \\ z_x(t) \end{pmatrix} = \begin{pmatrix} e^{t\Delta} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \int_0^t \begin{pmatrix} e^{(t-s)\Delta} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} \alpha(s) - z_x(s) \\ \frac{1}{\varepsilon} G(z_x(s) - y_x(s)) \end{pmatrix} ds \tag{3.9}$$

on the space $L^2(\Omega) \times L^2(\Omega) =: X$. We denote with $\alpha \in \mathcal{A} \subset L^\infty(0, \infty; L^2(\Omega))$ the function that was formerly denoted by f; those functions will serve as controls for the dynamical system. Here, $x = (x_1, x_2) \in X$ is the initial value, and the operators

$$t \mapsto e^{tA} := \begin{pmatrix} e^{t\Delta} & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \qquad t \geq 0,$$

form a family of strongly continuous semigroups of contractions on X, because for all $x \in X$,

$$\begin{aligned} \left\| e^{tA} x \right\|_X :=& \sqrt{\|e^{t\Delta} x_1\|_{L^2(\Omega)}^2 + \|x_2\|_{L^2(\Omega)}^2} \\ \leq& \sqrt{\|x_1\|_{L^2(\Omega)}^2 + \|x_2\|_{L^2(\Omega)}^2} \\ =& \|x\|_X . \end{aligned}$$

Next, we introduce a cost functional for an infinite horizon problem via

$$J_x(\alpha) := \int_0^\infty f^0(y_x(t), z_x(t), \alpha(t)) e^{-\lambda t} dt, \tag{3.10}$$

where $\lambda > 0$ is a so called discount factor, and $f^0 : X \times U \to \mathbb{R}$. From theorem 3.11, we infer that for every $\alpha \in \mathcal{A}$, there exists a unique solution to (3.9). Then, for quite general functions f^0, J_x is well defined. Let

$$V(x) := \inf_{\alpha \in \mathcal{A}} J_x(\alpha)$$

denote the **value function**, which yields, for every initial point $x \in X$, the optimal value (at least, if it is attained) of the minimization problem. For our model problem, we make the following assumptions.

Assumption 3.15 *For $U \subset L^2(\Omega)$ bounded, let $\mathcal{A}$ be the set of measurable functions $\alpha : [0, \infty) \to U$. Further, we assume that there exist constants $C, m > 0$ with*

$0 \le m < \frac{\lambda\varepsilon}{\varepsilon+1}$ *and a local modulus of continuity* $\omega(\cdot,\cdot)$ *(i.e., a continuous function* $\omega : (\mathbb{R}^+)^2 \to \mathbb{R}^+$ *such that for every fixed* $a \ge 0$, $\omega(\cdot, a)$ *is a modulus of continuity), such that for all* $x, \bar{x} \in X$ *and* $u \in U$,

$$\begin{cases} |f^0(x,u)| \le C\left(1 + \|x\|_X\right)^m, \\ |f^0(x,u) - f^0(\bar{x},u)| \le \omega\left(\|x - \bar{x}\|_X, \max\left\{\|x\|_X, \|\bar{x}\|_X\right\}\right). \end{cases}$$

□

Lemma 3.16 *Let denote*

$$f : X \times U \to X, \quad f(x,u) := \begin{pmatrix} u - x_2 \\ \frac{1}{\varepsilon} G(x_2 - x_1) \end{pmatrix}.$$

Then,

$$|f(x,u) - f(\bar{x},u)| \le \sqrt{\frac{\varepsilon^2 + 2}{\varepsilon^2}} \|x - \bar{x}\|_X,$$

and there is $L > 0$ *such that* $|f(0,u)| \le L \ \forall u \in U$. *Further, the generator of* e^{tA} *is*

$$(A, D(A)) := \left(\begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix}, \left(H_0^1(\Omega) \cap H^2(\Omega)\right) \times L^2(\Omega) \right),$$

and there exists $B \in \mathcal{L}(X)$ *positive, self adjoint, such that* $\mathcal{R}(B) \subset D(A^*)$ *(thus,* $A^*B \in \mathcal{L}(X)$*), and for some* $c_0 > 0$,

$$\langle A^*Bx, x\rangle_X \le c_0 \langle Bx, x\rangle_X - \|x\|_X^2, \quad \forall x \in X.$$

Proof: From the Lipschitz continuity of G, we infer that

$$\begin{aligned} \|f(x,u) - f(\bar{x},u)\|_X =& \sqrt{\|x_2 - \bar{x}_2\|_{L^2(\Omega)}^2 + \frac{1}{\varepsilon^2}\|x_2 - x_1 - \bar{x}_2 + \bar{x}_1\|_{L^2(\Omega)}^2} \\ \le& \sqrt{\left(1 + \frac{2}{\varepsilon^2}\right)\|x_2 - \bar{x}_2\|_{L^2(\Omega)}^2 + \frac{2}{\varepsilon^2}\|x_1 - \bar{x}_1\|_{L^2(\Omega)}^2} \\ \le& \sqrt{\frac{2 + \varepsilon^2}{\varepsilon^2}} \|x - \bar{x}\|_X. \end{aligned}$$

The generator of $e^{t\Delta}$ is Δ, and the identity is a bounded linear operator, constant in t; hence, the form of the generator of e^{tA} follows. Now, with $B_1 := (\mathbb{1} - \Delta)^{-1} : L^2(\Omega) \to H_0^1(\Omega) \cap H^2(\Omega)$, since $\Delta = \Delta^*$,

$$\langle \Delta B_1 x_1, x_1\rangle_{L^2(\Omega)} = \langle B_1 x_1, x_1\rangle_{L^2(\Omega)} - \|x_1\|_{L^2(\Omega)}^2.$$

Then, since $A = A^*$, with

$$B := \begin{pmatrix} B_1 & 0 \\ 0 & \mathbb{1} \end{pmatrix} : \left(L^2(\Omega)\right)^2 \to \left(H_0^1(\Omega) \cap H^2(\Omega)\right) \times L^2(\Omega),$$

we get, for $x \in X$,

$$\begin{aligned} \langle A^* Bx, x \rangle_X = \langle \Delta B_1 x_1, x_1 \rangle_{L^2(\Omega)} &= \langle B_1 x_1, x_1 \rangle_{L^2(\Omega)} - \|x_1\|^2_{L^2(\Omega)} \\ &= \langle Bx, x \rangle_X - \|x\|^2_X \,. \end{aligned}$$

As $\mathbb{1}$ as well as $-\Delta$ are positive operators, we may choose B together with $c_0 = 1$.

□

3.4 Dynamic programming method - Application of known results

We are now ready to apply the results from [2, chapter 6].

Proposition 3.17 (cf. [2, chapter 6, proposition 6.1]) *If assumption 3.15 holds, Then the value function V is locally uniformly continuous and for some constant $K > 0$,*

$$|V(x)| \leq K(1 + \|x\|_X)^m, \quad \forall x \in X.$$

Proof: See [2], proposition 6.1 (iii), chapter 6.

□

As usual, one needs the dynamic programming principle, which takes here the following form.

Proposition 3.18 (cf. [2, chapter 6, proposition 6.2]) *If assumption 3.15 holds, then for any $x \in X$ and $t > 0$,*

$$V(x) = \inf_{\alpha \in \mathcal{A}} \left\{ \int_0^t f^0(y_x(s), z_x(s), \alpha(s)) e^{-\lambda s} ds + V(y_x(t), z_x(t)) e^{-\lambda t} \right\}.$$

Proof: See [2, chapter 6,proposition 6.1].

□

Consider the (formal) HJB equation

$$\begin{aligned} &\lambda V(x) - \langle \nabla V(x), Ax \rangle_X - H(x, \nabla V(x)) = 0, \\ &H(x,p) := \inf_{u \in U} \left\{ \langle p, f(x,u) \rangle + f^0(x,u) \right\}, \quad x, p \in X, \end{aligned} \tag{3.11}$$

and the sets of test functions

$$\begin{aligned} \Phi_0 &:= \{ \varphi \in C^1(X) | \varphi \text{ is weakly sequentially} \\ &\qquad\qquad \text{lower semicontinuous}, \quad A^* \nabla \varphi \in C(X) \}, \\ \mathcal{G}_0 &:= \{ g \in C^1(X) | \exists \rho \in C^1(\mathbb{R}), \ \rho' \geq 0, \\ &\qquad\qquad \rho'(0) = 0, \ g(x) = \rho(\|x\|_X), \quad \forall x \in X \}. \end{aligned}$$

Definition 3.19 (cf. [2, chapter 6, proposition 6.3]) *A function $v \in C(X)$ is called **viscosity subsolution** (resp. **supersolution**) of (3.11), if for all $\varphi \in \Phi_0$ and $g \in \mathcal{G}_0$, whenever the function $v - \varphi - g$ attains a local maximum (resp., $v + \varphi + g$ attains a local minimum) at $x \in X$, it holds*

$$\lambda v(x) - \langle A^* \nabla \varphi(x), x \rangle_X - H(x, \nabla \varphi(x) + \nabla g(x)) \leq 0,$$

(respectively,

$$\lambda v(x) + \langle A^* \nabla \varphi(x), x \rangle_X - H(x, -\nabla \varphi(x) - \nabla g(x)) \geq 0. \text{)}$$

□

Theorem 3.20 (cf. [2, chapter 6, proposition 6.4]) *Assume that 3.15 holds. Then the value function V is a viscosity solution of (3.11).*

Proof: See [2, chapter 6, proposition 6.4].

□

To show uniqueness in this setting, one needs the notion of B-continuity.

Definition 3.21 (cf. [2, chapter 6,definition 2.3]) *Let $B \in \mathcal{L}(X)$ be self- adjoint and positive, and define the seminorm induced by B as*

$$\|x\|_B = \langle Bx, x \rangle_X^{\frac{1}{2}}, \quad \forall x \in X.$$

*A function $v : X \to \mathbb{R}$ is said to be **B-continuous** at $x_0 \in X$, if for all $x_n \in X$ with $x_n \to x_0$ weakly and $\|Bx_n - Bx_0\| \to 0$, it holds $v(x_n) \to v(x_0)$.*

□

Proposition 3.22 (cf. [2, chapter 6, proposition 6.5]) *If assumption 3.15 holds, then there exists a local modulus of continuity $\omega(\cdot,\cdot)$, such that*

$$|V(x) - V(\bar{x})| \leq \omega(\|x - \bar{x}\|_B, \max\{\|x\|_X, \|\bar{x}\|_X\}), \quad \forall x, \bar{x} \in X.$$

Proof: See [2] chapter 6, proposition 6.5, part two; the remaining assumptions were shown in lemma 3.16.

□

With this proposition, one can then characterize the value function in the following way.

Theorem 3.23 (cf. [2, chapter 6, theorem 6.6]) *Let assumption 3.15 hold, and denote by B the operator from lemma 3.16. Then the value function V is the unique B-continuous viscosity solution of* (3.11) *satisfying*

$$|V(x)| \leq K(1 + \|x\|_X)^m, \quad \forall x \in X.$$

Proof: See [2], chapter 6, theorem 6.6.

□

3.5 Convergence of optimal value functions

We conclude the investigation with showing convergence of trajectories of (3.7), which will then be used to show pointwise convergence of the value function for "regular" initial values. We call a pair of initial values (y_0, z_0) **admissible**, if $|z_0(x) - y_0(x)| \leq r(x)$ holds for almost every $x \in \Omega$. We will need the following statement.

Lemma 3.24 *Let y_0, z_0 be a pair of admissible initial values, z being the solution of*

$$\varepsilon \dot{z} = G(z - y), \quad z(0) = z_0,$$

and $y \in H^1(0, T; L^2(\Omega))$. Then, almost everywhere,

$$|z(x,t)| \leq r(x) + |y(x,t)| + \int_0^t |\dot{y}(x,s)|\, ds.$$

Proof: Note that, since $H^1(0,T; L^2(\Omega)) = L^2(\Omega; H^1(0,T))$, for almost every $x \in \Omega$, it holds $y(x,\cdot), z(x,\cdot) \in H^1(0,T)$. For such x, define $w := z - y$. If $|w| \leq r$, then the "reversed triangle inequality" yields

$$|z(x,t)| \leq r(x) + |y(x,t)| \leq r(x) + |y(x,t)| + \int_0^t |\dot{y}(x,s)|\, ds.$$

If not, then $G(w(x,t)) = -w(x,t) - r(x)$, or $G(w(x,t)) = -w(x,t) + r(x)$. In the first case, there is, by continuity, some $t_0 \geq 0$ such that $w(x,s) < -r(x)$ for all $t_0 < s \leq t$, and $w(x,t_0) = -r(x)$; further, it holds

$$\dot{w}(x,s) = \dot{z}(x,s) - \dot{y}(x,s) = -\frac{1}{\varepsilon}w(x,s) - \frac{1}{\varepsilon}r(x) - \dot{y}(x,s),$$

for alomst every $s \in (t_0, t)$. The variation of constants formula yields

$$w(x,t) = -\, r(x) \exp\left(-\frac{1}{\varepsilon}(t-t_0)\right) - \int_{t_0}^{t} \exp\left(-\frac{1}{\varepsilon}(t-s)\right)\left(\frac{1}{\varepsilon}r(x) + \dot{y}(x,s)\right) ds.$$

Noting that

$$r(x)\int_{t_0}^{t} \exp\left(-\frac{1}{\varepsilon}(t-s)\right)\frac{1}{\varepsilon}ds = r(x) - r(x)\exp\left(-\frac{1}{\varepsilon}(t-t_0)\right),$$

it follows

$$w(x,t) = -r(x) - \int_{t_0}^{t} \exp\left(-\frac{1}{\varepsilon}(t-s)\right)\dot{y}(x,s)ds,$$

and thus

$$|w(x,t)| \leq r(x) + \int_{t_0}^{t} |\dot{y}(x,s)|\, ds,$$

for all such $t > t_0$. The same inequality can be derived in the second case, using the same arguments. Hence, in particular,

$$|z(x,t)| \leq r(x) + |y(x,t)| + \int_{0}^{t} |\dot{y}(x,s)|\, ds$$

for all t and almost every $x \in \Omega$ as claimed.

□

Proposition 3.25 *Let y_0, z_0 be admissible initial values, $T > 0$, $\partial\Omega$ be of class C^2, $r \in L^2(\Omega)$, and $y^\varepsilon \in Y$, for every $\varepsilon > 0$, denote the weak solution to* (3.7) *(see theorem 3.10). Then y^ε is bounded in Y for $\varepsilon \downarrow 0$.*

Proof: As $y^\varepsilon \in Y$ is also a strict solution, it holds, in particular,

$$\dot{y}^\varepsilon + z^\varepsilon - \Delta y^\varepsilon = f \quad \text{a.e. in } \Omega_T.$$

From lemma 3.7, we infer that, testing the equation with $\dot{y}^\varepsilon \chi_{(0,t)}$ for $t \in (0,T]$ leads to

$$\int_{\Omega_t} (\dot{y}^\varepsilon)^2 \, d\mathcal{L} + \frac{1}{2} \int_\Omega |\nabla y^\varepsilon(x,t)|^2 \, dx = \int_{\Omega_t} (f - z^\varepsilon) \, \dot{y}^\varepsilon d\mathcal{L} + \frac{1}{2} \int_\Omega |\nabla y^\varepsilon(x,0)|^2 \, dx.$$

Let us denote $K := \frac{1}{2} \int_\Omega |\nabla y^\varepsilon(x,t)|^2 \, dx$. Lemma 3.24 and the triangle inequality followed by Jensen's inequality yield

$$\begin{aligned} \|z^\varepsilon\|_{L^2(\Omega_t)} \leq & \sqrt{t} \, \|r\|_{L^2(\Omega)} + \|y^\varepsilon\|_{L^2(\Omega_t)} + \sqrt{\int_0^t \int_\Omega \left(\int_0^s |\dot{y}^\varepsilon(x,r)| \, dr \right)^2 dx ds} \\ \leq & \sqrt{t} \, \|r\|_{L^2(\Omega)} + \|y^\varepsilon\|_{L^2(\Omega_t)} + t \, \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)} . \end{aligned}$$

Further, using similar arguments, we infer from

$$y^\varepsilon(x,t) = y_0^\varepsilon(x) + \int_0^t \dot{y}^\varepsilon(x,s) ds \quad \text{a.e.},$$

that

$$\|y^\varepsilon\|_{L^2(\Omega_t)} \leq \sqrt{t} \, \|y_0^\varepsilon\|_{L^2(\Omega)} + t \, \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)} .$$

Hence, the last two inequalities imply

$$\begin{aligned} -\int_{\Omega_t} z^\varepsilon \dot{y}^\varepsilon d\mathcal{L} & \leq \|z^\varepsilon\|_{L^2(\Omega_t)} \, \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)} \\ & \leq \left(\sqrt{t} \, \|r\|_{L^2(\Omega)} + \|y^\varepsilon\|_{L^2(\Omega_t)} + t \, \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)} \right) \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)} \\ & \leq \sqrt{t} \, \|r\|_{L^2(\Omega)} \, \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)} + \sqrt{t} \, \|y_0^\varepsilon\|_{L^2(\Omega)} \, \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)} + 2t \, \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)}^2 . \end{aligned}$$

With Young's inequality and $x \leq 1 + x^2 \; \forall x \in \mathbb{R}$, the right hand side may be further estimated by

$$\begin{aligned} & \sqrt{t} \, \|r\|_{L^2(\Omega)} \left(1 + \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)}^2 \right) + \frac{\sqrt{t}}{2} \left(\|y_0^\varepsilon\|_{L^2(\Omega)}^2 + \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)}^2 \right) + 2t \, \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)}^2 \\ & \leq C_1 + C_2 \sqrt{t} \, \|\dot{y}^\varepsilon\|_{L^2(\Omega_t)}^2 , \end{aligned}$$

where

$$\begin{aligned} C_1 := & \sqrt{T} \, \|r\|_{L^2(\Omega)} + \frac{1}{2} \sqrt{T} \, \|y_0^\varepsilon\|_{L^2(\Omega)}^2 , \\ C_2 := & \|r\|_{L^2(\Omega)} + 2\sqrt{T} + \frac{1}{2}. \end{aligned}$$

We may thus infer from the tested equality, that

$$\left(\frac{1}{2} - C_2\sqrt{t}\right) \|\dot{y}^\varepsilon\|^2_{L^2(\Omega_t)} + \frac{1}{2}\int_\Omega |\nabla y^\varepsilon(x,t)|^2\, dx \le \frac{1}{2}\|f\|^2_{L^2(\Omega_T)} + K + C_1,$$

and hence, boundedness of y^ε on $(0,\tau)$, whenever $\tau < (4C_2^2)^{-1}$. As C_2 only depends on r and T, we may iterate this procedure and conclude boundedness on the whole interval.

□

Corollary 3.26 *Let the assumptions of proposition 3.25 and theorem 3.10 hold. Then $y^\varepsilon \to y$ weak $\star$ in Y and $z^\varepsilon \to w$ weak in $L^2(\Omega_T)$ as $\varepsilon \downarrow 0$, where (y,w) is the weak solution to* (3.1) *(see theorem 3.10).*

Proof: First note that by proposition 3.25, y^ε is bounded in Y as $\varepsilon \downarrow 0$, so that there exists a weak $\star$ convergent (sub-) sequence $y^{\varepsilon_n} =: y^n \to y \in Y$. Next, we make again use of the compactness of the imbedding $Y \hookrightarrow L^2(\Omega; C[0,T])$. This enables us to find a subsequence (for simplicity again denoted by y^n) that converges strong in $L^2(\Omega; C[0,T])$ to y. But then, it is necessary that $y^n(x,\cdot)$ converges in $C[0,T]$ to $y(x,\cdot)$, for almost every $x \in \Omega$. Due to theorem 2.11, for almost all $x \in \Omega$, $z^\varepsilon(x,\cdot) \to w(x,\cdot)$ in $C[0,T]$, where $w = \mathcal{F}_r[y; z_0]$. From lemma 3.24, we get additionally, that (at least a subsequence, which will then be denoted in the same way) z^n converges weakly in $L^2(\Omega_T)$, and, by uniqueness of limits, $z^n \rightharpoonup w$. We may then take the limit of the variational equality, which shows that (y,w) solves (3.1) in the sense of theorem 3.3. As we can argue in the same way for every subsequence, the proof is complete.

□

We end this section by showing that, under certain conditions, convergence of the trajectories implies pointwise convergence of the value functions. We add the following, more restrictive, assumptions.

Assumption 3.27 *The function f^0 is bounded and Lipschitz continuous, and only depends on the x_1-coordinate, i.e., $f^0(x,u) = h(x_1,u)$ for some function $h : L^2(\Omega) \times U \to \mathbb{R}$, and there are constants $c, c_L > 0$ such that*

$$\left|f^0(x,u)\right| \le c, \quad \left|f^0(x,u) - f^0(\bar{x},u)\right| \le c_L \|x_1 - \bar{x}_1\|_X, \quad \forall x, \bar{x} \in X,\ u \in U.$$

□

Remark 3.28 *Due to the boundedness assumption, the value function is automatically well defined because of the appearence of the discount factor. Thus, we may use the same target functional and set of controls to create a control problem for* (3.1)*. The corresponding value function will, in the following, be denoted by V^0; the one corresponding to the approximate problem by V^ε. We also remark that the dynamic programming principle (cf. proposition 3.18) holds for V^0, even we have not defined trajectories for all $x \in X$; this is because trajectories are unique and thus the solution operator has a semigroup property.*

□

Theorem 3.29 *Let assumptions 3.15 and 3.27 hold, together with the ones of corollary 3.26. Then, for every admissible $x \in H_0^1(\Omega) \times L^2(\Omega)$,*

$$V^\varepsilon(x) \stackrel{\varepsilon\downarrow 0}{\longrightarrow} V^0(x).$$

Proof: Let $\delta > 0$ be arbitrary. By definition, there exists a control $\alpha \in \mathcal{A}$, such that

$$V^0(x) \geq \int_0^\infty e^{-\lambda t} h(y(t), \alpha(t)) dt - \frac{\delta}{2}.$$

Then, the Lipschitz and boundedness conditions imply

$$\begin{aligned} V^\varepsilon(x) - V^0(x) &\leq \int_0^\infty e^{-\lambda t} \left(h(y^\varepsilon(t), \alpha(t)) - h(y(t), \alpha(t)) \right) dt + \frac{\delta}{2} \\ &\leq \int_0^T e^{-\lambda t} \left\| y^\varepsilon(t) - y(t) \right\|_{L^2(\Omega)} dt + 2c \int_T^\infty e^{-\lambda t} dt + \frac{\delta}{2}. \end{aligned}$$

We may now choose $T > 0$ so that the second term is smaller than $\frac{\delta}{2}$. But then, with corollary 3.26, we get that

$$\limsup_{\varepsilon\downarrow 0} \left(V^\varepsilon(x) - V^0(x) \right) \leq \delta,$$

and, since $\delta > 0$ was arbitrary,

$$\limsup_{\varepsilon\downarrow 0} \left(V^\varepsilon(x) - V^0(x) \right) \leq 0.$$

Now, as second step, let again $\delta > 0$ be arbitrary, and note that for all $\varepsilon > 0$, there exists a control $\alpha \in \mathcal{A}$ such that

$$V^\varepsilon(x) \geq \int_0^\infty e^{-\lambda t} h(y^\varepsilon(t), \alpha^\varepsilon(t)) dt - \frac{\delta}{2}.$$

We may then, similar to the first part, infer that

$$\liminf_{\varepsilon \downarrow 0} \big(V^\varepsilon(x) - V^0(x)\big) \geq 0.$$

Altogether, this implies

$$\lim_{\varepsilon \downarrow 0} \big(V^\varepsilon(x) - V^0(x)\big) = 0.$$

□

4 An Abstract model and general assumptions

4.1 Special Semigroup

Let X be a Hilbert space with norm $\|\cdot\|$ induced by the inner product $\langle\cdot,\cdot\rangle$, let A be a (possibly unbounded) operator with domain $D(A) \subset X$, which is closed, densely defined and selfadjoint in X, such that $\langle Ax, x\rangle \le -\omega \|x\|^2$ for all $x \in D(A)$, where $\omega > 0$. Then the following holds (see e.g. [27, appendix], or [6]):

- $(A, D(A))$ is the infinitesimal generator of a strongly continuous analytic semigroup of contractions on X, which will be denoted by e^{tA}, for $t \ge 0$;
- for all $\theta \in \mathbb{R}$, $(-A)^\theta$ exists and it holds

$$(-A)^\alpha(-A)^\beta x = (-A)^{\alpha+\beta}x = (-A)^\beta(-A)^\alpha x,\ x \in D((-A)^{\max(\alpha,\beta,\alpha+\beta)}); \quad (4.1)$$

- for all $\delta \ge 0$ there exists a constant $M_\delta > 0$ such that, for every $t > 0$:

$$\left\|(-A)^\delta e^{tA}\right\|_{\mathcal{L}(X)} \le M_\delta t^{-\delta} e^{-\omega t}; \quad (4.2)$$

- if $0 \le \delta \le 1$ and $x \in D((-A)^\delta)$, then there exists $N_\delta > 0$ such that, for all $t \ge 0$:

$$\left\|\left(e^{tA} - I\right)x\right\| \le N_\delta t^\delta \left\|(-A)^\delta x\right\|; \quad (4.3)$$

- for all $0 \le \delta < \gamma \le 1$ and all $x \in D((-A)^\gamma)$, there exists $M_{\delta,\gamma} > 0$ such that the following interpolation inequality between the corresponding abstract Sobolev spaces holds:

$$\left\|(-A)^\delta x\right\| \le M_{\delta,\gamma} \|(-A)^\gamma x\|^{\frac{\delta}{\gamma}} \|x\|^{1-\frac{\delta}{\gamma}}. \quad (4.4)$$

Further, if $f \in L^1(0,T;X)$, then there exists a unique mild solution for

$$\dot{y} - Ay = f \quad \text{in } (0,T), \qquad y(0) = x \in X, \quad (4.5)$$

which has to be understood in the integral form

$$y(t) = e^{tA}x + \int_0^t e^{(t-s)A} f(s)ds \ \in C(0,T;X),$$

see e.g. [25, chapter 4].

4.2 Dynamic and further assumptions

In this section we will study existence and uniqueness of solutions to the differential equation

$$\begin{aligned} \dot{y} - Ay + F(y) &= \alpha, \\ F(y) &= \sum_{k=1}^{n} B_k(f_k(y))g_k, \end{aligned} \tag{4.6}$$

under **assumption (A1)**:
A is an operator of the type described in section 4.1 and for any $T > 0$, each of the operators $B_k : C[0,T] \to C[0,T]$ is Lipschitz continuous with Lipschitz constant not larger than $L > 0$, and $g_k, \alpha \in L^\infty(0,T;X)$, $k \in \{1,\dots,n\}$, $f_k \in X$ $(f_k(y) := \langle f_k, y \rangle)$.

Theorem 4.1 *If **(A1)** holds, then for every $x = y(0) \in X$ there exists a unique mild solution $y \in C(0,T;X)$ to* (4.6)*, in the sense that*

$$y(t) = e^{tA}x + \int_0^t e^{(t-s)A}\left(\alpha(s) - F(y(s))\right) ds.$$

If, in addition, $x \in D((-A)^\theta)$ for some $\theta \in (0,1)$, then $y \in C(0,T;D((-A)^\theta))$, and for any $\varphi \in C^1(X)$ such that $(-A)^{1-\theta}\nabla\varphi(\cdot) \in C(X)$, the formula

$$\begin{aligned} \varphi(y(t)) - \varphi(x) = \int_0^t &- \left\langle (-A)^{1-\theta}\nabla\varphi(y(s)), (-A)^\theta y(s) \right\rangle ds \\ &+ \int_0^t \langle \varphi(y(s)), \alpha(s) - F(y(s)) \rangle \, ds \end{aligned}$$

holds.

Proof: Let $0 < \tau \le T$, and $v \in C(0,\tau;X)$ be arbitrary. By assumption, the function defined by

$$f_v(s) := \alpha(s) - F(v(s)) \ \in L^\infty(0,\tau;X),$$

hence there exists a unique mild solution z of $\dot{z} - Az = f_v$, i.e.,

$$z(t) = e^{tA}x + \int_0^t e^{(t-s)A} f_v(s) ds, \quad t \in [0,\tau].$$

Thus, the assignment $v \mapsto z$ defines an operator $J : C(0,\tau;X) \to C(0,\tau;X)$ such that y solves (4.6) for T replaced by τ if and only if y is a fixed point of J. Our goal is now to apply Banach's fixed point theorem, at least for small τ, and then to show that the solution can be continued. To this end, let $v_1, v_2 \in C(0,\tau;X)$ and $z_i = J(v_i)$, $i = 1,2$. Then,

$$\begin{aligned}\|z_1(t) - z_2(t)\| =& \left\| \int_0^t e^{(t-s)A} \left(F(v_2(s)) - F(v_1(s))\right) ds \right\| \\ \leq & L \sum_{k=1}^n \int_0^t \|\langle f_k, v_2(\cdot) - v_1(\cdot)\rangle\|_{C[0,t]} \cdot \|g_k(s)\| \\ \leq & \tau L \left(\sum_{k=1}^n \|f_k\| \, \|g_k\|_{L^\infty(0,\tau;X)} \right) \|v_1 - v_2\|_{C(0,\tau;X)} .\end{aligned}$$

Choosing $\bar{\tau} = \min(T, \frac{1}{2}(L\sum_{k=1}^n \|f_k\| \, \|g_k\|_{C(0,T;X)})^{-1})$, we find that J is a contraction mapping on $C(0,\bar{\tau};X)$, and we can apply Banach's fixed point theorem, which yields a unique local solution of (4.6) on $[0,\bar{\tau}]$. Since

$$\begin{aligned}\|y(t)\| \leq \|x\| + T\|\alpha\|_{C(0,T;X)} + L \sum_{k=1}^n |B_k(0)| \, \|g_k\|_{L^\infty(0,T;X)} \\ + L \sum_{k=1}^n \|f_k\| \, \|g_k\|_{C(0,T;X)} \int_0^t \|y(\cdot)\|_{C(0,s;X)} \, ds,\end{aligned}$$

Gronwall's lemma implies that solutions are bounded on bounded intervals. Hence, the local solution can be continued to the whole interval $[0,T]$.

Now, if $x \in D((-A)^\theta)$, the higher regularity follows with (4.2), (4.3) and the semigroup property by standard arguments. Further, let $\varepsilon > 0$, $x \in D((-A)^{\theta+\varepsilon})$, and φ with the above properties be given; denote $f(s) := \alpha(s) - F(y(s))$. For $\tau > 0$, we introduce (let $T(t) := e^{tA}$)

$$\bar{y}(t) := T(\tau)y(t) = T(t)T(\tau)x + \int_0^t T(t-s)T(\tau)f(s)ds,$$

the solution of

$$\begin{aligned}\bar{y}(0) =& T(\tau)x, \\ \dot{\bar{y}}(t) =& A\bar{y}(t) + T(\tau)f(t).\end{aligned}$$

It holds $\bar{y} \in C(0,T;D(A))$, and thus it fulfills the differential equation for almost every t and its derivative is integrable. Hence, we may compute

$$\varphi(T(\tau)y(t)) - \varphi(T(\tau)x) = \int_0^t \langle \nabla\varphi(T(\tau)y(s)), AT(2\tau)y(s) + T(2\tau)f(s)\rangle \, ds$$

$$= -\int_0^t \big\langle (-A)^{1-\theta}\nabla\varphi(T(\tau)y(s)), (-A)^{\theta}T(2\tau)y(s)\big\rangle\, ds$$
$$+ \int_0^t \langle \nabla\varphi(T(\tau)y(s)), T(2\tau)f(s)\rangle\, ds.$$

Letting $\tau \downarrow 0$ yields the result for $x \in D((-A)^{\theta+\varepsilon})$, because (4.3) implies that $(-A)^{\theta}T(\tau)y(s) = T(\tau)(-A)^{\theta}y(s) \to (-A)^{\theta}y(s)$ uniformly in $s \in [0,t]$. Then, by continuity, we can let $\varepsilon \downarrow 0$ to get the desired formula.

□

We are mainly interested in operators B_k which are of hysteresis type. Let $\mathcal{F}_\rho$ denote the play operator corresponding to the interval $[-\rho,\rho]$, and define a Prandtl-Ishlinskii operator of play type via

$$P_k : C[0,T] \times (f_k(x) - Z_k) \to C[0,T],$$
$$P_k[y;\zeta_k] := \int_{R_k} \mathcal{F}_\rho[y;\zeta_k]d\mu_k(\rho),$$

where, for each $k \in \{1,\dots,n\}$, R_k is a subset of $\mathbb{R}^+$, μ_k is a probability measure on R_k, and $\zeta_k = f_k(x) - \xi_k$,

$$\xi_k \in Z_k := \{\xi \mid \xi : R_k \to \mathbb{R} \text{ measurable } \wedge\ |\xi(\rho)| \le \rho\ \forall\rho \in R_k\}.$$

We assume further that

$$\int_{R_k} \rho^2 d\mu_k(\rho) < \infty,$$

which implies that each Z_k is a subset of $L^2(R_k,\mu_k)$ (this is obviously fulfilled when R_k is a bounded set, which is typically the case). Note that since μ_k is a probability measure, by Jensen's inequality

$$\left(\int_{R_k} \xi(\rho)d\mu_k(\rho)\right)^2 \le \int_{R_k} \xi(\rho)^2 d\mu_k(\rho) \le \int_{R_k} \rho^2 d\mu_k(\rho) < \infty,$$

and therefore (by interpolation) $Z_k \subset L^p(R_k,\mu_k)$, $p \in [1,2]$. For such p, we may thus define, for $\xi \in Z := \prod_{k=1}^n Z_k$, the norms

$$\|\xi\|_{Z,p} := \left(\sum_{k=1}^n \int_{R_k} \xi_k(\rho)^p d\mu_k(\rho)\right)^{\frac{1}{p}},\ \|\xi\|_Z := \|\xi\|_{Z,2}. \tag{4.7}$$

Now, for fixed $\xi = (\xi_1, \ldots, \xi_n)^T \in Z$, each of the operators $P_k[\cdot, \zeta_k]$ is Lipschitz continuous in $C[0,T]$, because for all $t \in [0,T]$,

$$\begin{aligned} |P_k[y_1;\zeta_k](t) - P_k[y_2;\zeta_k](t)| &\leq \int_{R_k} |\mathcal{F}_\rho[y_1;\zeta_k(\rho)](t) - \mathcal{F}_\rho[y_2;\zeta_k(\rho)](t)| \, d\mu_k(\rho) \\ &\leq \int_{R_k} \sup_{t\in[0,T]} |y_1(t) - y_2(t)| \, d\mu_k(\rho) \\ &= \sup_{t\in[0,T]} |y_1(t) - y_2(t)| \, . \end{aligned}$$

Hence, theorem 4.1 is applicable for $B_k(\cdot) := P_k[\cdot, \zeta_k]$. We will refer to this special case as **problem (P)**:

$$\begin{aligned} \dot{y} - Ay + F(y) &= \alpha, \\ F(y) &= \sum_{k=1}^{n} B_k(f_k(y))g_k, \end{aligned}$$

with assumption **(A1)**, and assumption **(A2)** :

$$B_k = P_k \; \forall \; k \in \{1, \ldots, n\} \, , \; \xi \in Z \subset \prod_{k=1}^{n} L^2(R_k, \mu_k), \; \mu_k \text{ probability measures.}$$

Moreover, we define the vector valued function

$$w(t) = (w_1(t), \ldots, w_n(t)) := (P_1[f_1(y);\xi_1](t), \ldots, P_n[f_n(y);\xi_n](t)),$$

whose coordinates might be interpreted as some kind of "localized mean hysteresis" (when f_k, g_k are chosen such that P_k only operates on some disjoint subsets Ω_k of Ω). Next we present a continuity result for the solution operator corresponding to problem **(P)**.

Proposition 4.2 *Let y_1, y_2 denote, respectively, the solutions to problem **(P)** corresponding to initial values $(x,\xi),(z,\zeta) \in X \times (\prod_{k=1}^{n}(f_k(x) - Z_k))$, and assume that for some $0 \leq \theta < 1$ it holds that $f_k \in D((-A)^\theta)$, for every $k \in \{1, \ldots, n\}$. Then there exist a constant $C_1 = C_1(\theta, f_1, \ldots, f_n)$ and a locally integrable function $C_2 = C_2(T, \theta, g_1, \ldots, g_n, f_1, \ldots, f_n)$ such that, for any $t \in [0,T]$,*

$$\begin{aligned} |w_1(t) - w_2(t)|_1 &\leq C_1 \left(\left\| (-A)^{-\theta}(x - z) \right\| + \|\xi - \zeta\|_{Z,1} \right) \cdot \exp(t \sum_{k=1}^{n} \Gamma_k), \\ \|y_1(t) - y_2(t)\| &\leq C_2 \left(\left\| (-A)^{-\theta}(x - z) \right\| + \|\xi - \zeta\|_{Z,1} \right) \cdot \exp(t \sum_{k=1}^{n} \Gamma_k). \end{aligned}$$

C_1, C_2, $\Gamma_1, \ldots, \Gamma_n$ are given by (4.10), (4.11) and (4.9).

Proof: Let $T > 0$. For fixed $k \in \{1, \ldots, n\}$ and $\rho \in R_k$, the Lipschitz continuity of $\mathcal{F}_\rho$ in $C(0,T)$ allows us to estimate

$$
\begin{aligned}
\Delta\mathcal{F}_\rho^k(T) :=& \left|\mathcal{F}_\rho[f_k(y_1);\xi_k(\rho)](T) - \mathcal{F}_\rho[f_k(y_2);\zeta_k(\rho)](T)\right| \\
\leq& \sup_{t\in[0,T]} |f_k(y_1)(t) - f_k(y_2)(t)| + |\xi_k(\rho) - \zeta_k(\rho)| \\
=& \sup_{t\in[0,T]} \Big| \big\langle f_k, e^{tA}(x-z)\big\rangle \\
&+ \int_0^t \big\langle f_k, e^{(t-s)A}\left(F[y_2;\zeta](s) - F[y_1;\xi](s)\right)\big\rangle\, ds \Big| + |\xi_k(\rho) - \zeta_k(\rho)| \\
\leq& \sup_{t\in[0,T]} \left|\big\langle (-A)^\theta f_k, e^{tA}(-A)^{-\theta}(x-z)\big\rangle\right| \\
&+ \int_0^t \sum_{j=1}^n \left|\big\langle f_k, e^{(t-s)A} g_j(s)\big\rangle\right| \cdot \int_{R_j} \Delta\mathcal{F}_\rho^j(s) d\mu_j(\rho) ds \\
&+ |\xi_k(\rho) - \zeta_k(\rho)| \\
\leq& \left\|(-A)^\theta f_k\right\| \left\|(-A)^{-\theta}(x-z)\right\| \\
&+ \Gamma_k \int_0^T \sum_{j=1}^n \int_{R_j} \Delta\mathcal{F}_\rho^j(s) d\mu_j(\rho) ds + |\xi_k(\rho) - \zeta_k(\rho)|, \qquad (4.8)
\end{aligned}
$$

$$
\Gamma_k := \sup_{t>0}\left\{ \sup_{0\leq s\leq t}\left\{ \max_{1\leq j\leq n}\left\{ \left|\big\langle f_k, e^{(t-s)A} g_j(s)\big\rangle\right| \right\}\right\}\right\}. \qquad (4.9)
$$

Note that (4.2) implies that $\Gamma_k < \infty$. Denoting $\left\|(-A)^\theta f_k\right\| =: a_k$, integration w.r.t. μ_k followed by summation over k yields

$$
\begin{aligned}
\sum_{k=1}^n \int_{R_k} \Delta\mathcal{F}_\rho^k(T) d\mu_k(\rho) \leq& \left(\sum_{k=1}^n a_k\right) \left\|(-A)^{-\theta}(x-z)\right\| \\
&+ \left(\sum_{k=1}^n \Gamma_k\right) \int_0^T \sum_{j=1}^n \int_{R_j} \Delta\mathcal{F}_\rho^j(s) d\mu_j(\rho) ds \\
&+ \sum_{k=1}^n \int_{R_k} |\xi_k(\rho) - \zeta_k(\rho)|\, d\mu_k.
\end{aligned}
$$

Hence, applying the standard version of Gronwall's lemma (cf. theorem D.3) to the function $u(t) := \sum_{k=1}^n \int_{R_k} \Delta\mathcal{F}_\rho^k(t) d\mu_k(\rho)$ yields

$$
\sum_{k=1}^n \int_{R_k} \Delta\mathcal{F}_\rho^k(t) d\mu_k(\rho) \leq \left(\sum_{k=1}^n a_k\right) \left\|(-A)^{-\theta}(x-z)\right\| \exp\left(t\sum_{k=1}^n \Gamma_k\right)
$$

$$+ \|\xi - \zeta\|_{Z,1} \exp\left(t \sum_{k=1}^{n} \Gamma_k\right).$$

Therefore, one can choose

$$C_1 = C_1(\theta, f_1, \dots, f_n) := 1 + \sum_{k=1}^{n} a_k. \tag{4.10}$$

Now, using the definition of mild solutions, estimate (4.2) and (4.9), we get for all $t > 0$

$$\begin{aligned}
\|y_1(t) - y_2(t)\| \leq & \frac{M_\theta}{t^\theta} \left\|(-A)^{-\theta}(x - z)\right\| \\
& + \int_0^t \left(\max_{1\leq k\leq n} \|g_k(s)\|\right) \sum_{k=1}^{n} \int_{R_k} \Delta\mathcal{F}_\rho^k(s) d\mu_k(\rho) ds \\
\leq & \frac{M_\theta}{t^\theta} \left\|(-A)^{-\theta}(x - z)\right\| \\
& + \left(\sum_{k=1}^{n} a_k\right) \left\|(-A)^{-\theta}(x - z)\right\| \int_0^t G_n(s) \exp\left(s \sum_{k=1}^{n} \Gamma_k\right) \\
& + \|\xi - \zeta\|_{Z,1} \int_0^t G_n(s) \exp\left(s \sum_{k=1}^{n} \Gamma_k\right),
\end{aligned}$$

where $G_n(s) := \max_{1\leq k\leq n} \|g_k(s)\| \leq \max_{1\leq k\leq n} \|g_k\|_{L^\infty(0,T;X)} =: G$, $t \in [0, T]$. Hence, for

$$C_2 := C_2(t, \theta, G, f_1, \dots, f_n) = \left(1 + t^{-\theta}\right) \left(M_\theta + tG\left(1 + \sum_{k=1}^{n} a_k\right)\right) \tag{4.11}$$

it holds that

$$\|y_1(t) - y_2(t)\| \leq C_2 \left(\left\|(-A)^{-\theta}(x - z)\right\| + \|\xi - \zeta\|_{Z,1}\right) \exp\left(t \sum_{k=1}^{n} \Gamma_k\right).$$

Since $\theta \in [0, 1)$, C_2 is integrable over $(0, T)$, and the proof is complete.

□

Remark 4.3 *Proposition 4.2 holds for general operators fulfilling a Lipschitz condition, i.e., one might for example replace the family of play operators $\mathcal{F}_\rho$ by some Lipschitz continuous operator $B : C[0, T] \to C[0, T]$, and choose $\mu = \mathcal{L}|_{[0,1]}$, the Lebesgue measure on $[0, 1]$. Then a similar estimate holds with $\xi = \zeta$ (i.e., $\|\xi - \zeta\|_{Z,1} = 0$); however, the Lipschitz constant (which equals one for the play operator) will appear in the exponential function.*

□

Remark 4.4 *Using* (4.2), *one can estimate*

$$\Gamma_k \leq \max_{1\leq j\leq n} \left\{ \|f_k\| \, \|g_j\|_{L^\infty(0,T;X)} \right\}.$$

□

With the well known identity $\mathcal{F}_r = I - \mathcal{E}_r$, where I denotes the identical mapping and $\mathcal{E}_r$ denotes the stop operator corresponding to the characteristic set $[-r, r]$, problem **(P)** may alternatively be written as **problem (Q)**:

$$\begin{gathered} \dot{y} - Ay + F(y;\xi) = \alpha, \\ F(y;\xi) = \sum_{k=1}^{n} f_k(y) \cdot g_k - \int_{R_k} \mathcal{E}_\rho[f_k(y);\xi_k] d\mu_k(\rho) \cdot g_k, \\ \text{together with assumptions } (A1),\ (A2), \end{gathered}$$

with initial values in $X \times Z$. Since in later sections, we will only work with this formulation of the problem, we translate proposition 4.2. It is just a matter of how initial values are defined - one gets $\xi_{k,P} = f_k(x) - \xi_{k,Q}$, and therefore, since μ_k are probability measures,

$$\|\xi_P - \zeta_P\|_{Z,1} \leq \|\xi_Q - \zeta_Q\|_{Z,1} + \sum_{k=1}^{n} |\langle f_k, x - z\rangle| \, .$$

By the same arguments as in the proof of proposition 4.2, we get analogous estimates with slightly bigger constants

$$\bar{C}_1 := C_1 \cdot \left(1 + \sum_{k=1}^{n} a_k\right), \ \bar{C}_2 := C_2 \cdot \left(1 + \sum_{k=1}^{n} a_k\right).$$

Corollary 4.5 *Let the assumptions of proposition 4.2 hold. There are analogous estimates for problem **(Q)**, with C_1, C_2 replaced by $\bar{C}_1$, $\bar{C}_2$.*

□

5 The control problem

We are going to introduce an infinite horizon control problem related to problem **(Q)**. To this end, note that $T > 0$ was arbitrary in section 4, so that the existence and uniqueness of solutions is guaranteed for arbitrary large time. Further, corollary 4.5 holds.

Let us now make some further assumptions, which should hold generally in sections 5 to 7. Let $\mathbb{A} \subset X$ be nonempty, bounded and closed, and define

$$\mathcal{A} := \{\beta | \beta : [0, \infty) \to \mathbb{A} \text{ measurable}\},$$

the set of controls for problem **(Q)**, i.e., we will assume $\alpha \in \mathcal{A}$. In addition, let $g_k \in X$ (i.e., constant in time), for each $k \in \{1, \ldots, n\}$. In order to define the cost functional, we consider

$$\begin{aligned} &L : X \times \mathbb{R}^n \times \mathbb{A} \longrightarrow \mathbb{R}^+, \text{ such that} \\ &|L(x_1, w_1, a) - L(x_2, w_2, a)| \leq C_L \left(\|x_1 - x_2\| + |w_1 - w_2|_2\right), \\ &\forall x_1, x_2 \in X, w_1, w_2 \in \mathbb{R}^n, a \in \mathbb{A}. \end{aligned} \tag{5.1}$$

Then, we define a cost functional

$$J(x, \xi, \alpha) := \int_0^\infty e^{-\lambda t} L(y_{x,\xi,\alpha}(t), w_{x,\xi,\alpha}(t), \alpha(t)) dt,$$

with discount factor

$$\lambda > \Gamma := \sum_{k=1}^n \Gamma_k. \tag{5.2}$$

The functions $y_{x,\xi,\alpha}$, $w_{x,\xi,\alpha}$ denote the solution to problem **(Q)** and its hysteresis part corresponding to inital values $(x, \xi) \in X \times Z$ and w.r.t. the control $\alpha \in \mathcal{A}$. More precisely,

$$w_{x,\xi,\alpha,k}(t) := f_k(y_{x,\xi,\alpha}(t)) - \int_{R_k} \mathcal{E}_\rho[f_k(y_{x,\xi,\alpha}); \xi_k](t) d\mu_k(\rho).$$

To shorten notations, we will also write

$$\xi_{x,\xi,\alpha}(t) := (\mathcal{E}.[f_1(y_{x,\xi,\alpha}); \xi_1(\cdot)](t), \ldots, \mathcal{E}.[f_n(y_{x,\xi,\alpha}); \xi_n(\cdot)](t)) \in Z,$$

so that $\xi_{x,\xi,\alpha}(0) = \xi$. By (4.7), we may view Z as subset of $\prod_{k=1}^n L^2(R_k, \mu_k)$ w.r.t. the norm $\|\xi\|_Z$, which is induced by the scalar product

$$\langle \xi, \zeta \rangle_Z := \sum_{k=1}^n \int_{R_k} \xi(\rho) \zeta(\rho) d\mu_k(\rho).$$

We also introduce the vector

$$\mathbb{1} := (\mathbb{1}_{R_1}, \ldots, \mathbb{1}_{R_n}),$$

wherefore we may write

$$\sum_{k=1}^{n} \int_{R_k} \xi_k(\rho) d\mu_k(\rho) = \langle \xi, \mathbb{1} \rangle_Z .$$

Note that according to these notations, we have

$$w_0 = \left(f_1(x) - \int_{R_1} \xi_1(\rho) d\mu_1(\rho), \ldots, f_n(x) - \int_{R_n} \xi_n(\rho) d\mu_n(\rho) \right). \tag{5.3}$$

As usual, the value function is defined by

$$V(x, \xi) := \inf_{\alpha \in \mathcal{A}} J(x, \xi, \alpha). \tag{5.4}$$

It has the following Lipschitz type property.

Theorem 5.1 *Let the assumptions of problem **(Q)** and* (5.2) *hold. Then, for every* $0 \leq \theta < 1$, *for which* $f_k \in D((-A)^\theta)$ *holds for all* $k \in \{1, \ldots, n\}$, *there is* $C > 0$, *such that, for all* $(x, \xi), (z, \zeta) \in X \times Z$,

$$|V(x, \xi) - V(z, \zeta)| \leq C \left(\left\| (-A)^{-\theta}(x - z) \right\| + \|\xi - \zeta\|_Z \right). \tag{5.5}$$

Proof: By the definition of V (5.4) and the Lipschitz condition (5.1) on L, we have

$$|V(x, \xi) - V(z, \zeta)|$$
$$\leq \sup_{\alpha \in \mathcal{A}} \left\{ \int_0^\infty e^{-\lambda t} \left| L(y_{x,\xi,\alpha}(t), w_{x,\xi,\alpha}(t), \alpha(t)) - L(y_{z,\zeta,\alpha}(t), w_{z,\zeta,\alpha}(t), \alpha(t)) \right| dt \right\} \tag{5.6}$$
$$\leq \sup_{\alpha \in \mathcal{A}} \left\{ C_L \int_0^\infty e^{-\lambda t} \left(\|y_{x,\xi,\alpha}(t) - y_{z,\zeta,\alpha}(t)\| + |w_{x,\xi,\alpha}(t) - w_{z,\zeta,\alpha}(t)|_2 \right) dt \right\}. \tag{5.7}$$

Since the norms $|\cdot|_1$, $|\cdot|_2$ are equivalent on $\mathbb{R}^n$, we may, after having enlarged the constant, apply corollary 4.5, so that for some $C > 0$

$$|V(x, \xi) - V(z, \zeta)|$$
$$\leq C \left(\left\| (-A)^{-\theta}(x - z) \right\| + \|\xi - \zeta\|_{Z,1} \right) \int_0^\infty e^{(\Gamma - \lambda)t}(1 + t + t^{-\theta}) dt.$$

The integral on the right hand side converges for $\lambda > \Gamma$, and the norm $\|\xi - \zeta\|_{Z,1}$ can be estimated by some constant times $\|\xi - \zeta\|_Z$ (use equivalence of norms in $\mathbb{R}^n$ and Jensen's inequality). Thus, (5.5) follows. □

Remark 5.2 *If $\lambda \leq \Gamma$, one can still show that the value function is locally Hölder continuous: It is not hard to see that solutions to problem **(Q)** can not leave a ball in X whose radius depends only on the initial value $x \in X$ (not even on $\xi \in Z$). Therefore, $L(y, w, \alpha)$ can be estimated by some function depending only on the initial value. Splitting the term in absolute value of* (5.6) *into $\cdot^{1-\beta}$ and $\cdot^{\beta}$ (and β small enough), one can conclude, using the Lipschitz continuity of L.*

□

6 Dynamic programming and the value function as viscosity solution

6.1 The dynamic programming principle

Recall that each of the operators $\mathcal{E}_\rho$ has the semigroup property

$$\mathcal{E}_\rho[y;\xi(\rho)](t+\tau) = \mathcal{E}_\rho\left[v;\mathcal{E}_\rho[y;\xi(\rho)](t)\right](\tau),$$

with $v(s) := y(t+s),\ s \in [0,\tau]$. The following theorem, sometimes also called the Bellman principle of optimality, holds in quite general situations. It formalizes the intuition that an optimal control can be found by splitting the interval $[0,\infty)$ and then find optimal solutions to the subproblems. More precisely, it holds:

Theorem 6.1 *(dynamic programming principle) For all $(x,\xi) \in X \times Z$ and $t > 0$, it holds that*

$$V(x,\xi) = \inf_{\alpha\in\mathcal{A}} \left\{ \int_0^t e^{-\lambda s} L(y_{x,\xi,\alpha}(s), w_{x,\xi,\alpha}(s), \alpha(s))ds + e^{-\lambda t} V(y_{x,\xi,\alpha}(t), \xi_{x,\xi,\alpha}(t)) \right\}. \tag{6.1}$$

Proof: Since g_k is assumed to be constant in time, y and w have a semigroup property. Let $W(x,\xi)$ denote the right hand side of (6.1), and I the integral term thereof. We begin with showing that $V(x,\xi) \geq W(x,\xi)$. For all $\alpha \in \mathcal{A}$,

$$J(x,\xi,\alpha) = I + \int_t^\infty e^{-\lambda s} L(y_{x,\xi,\alpha}(s), w_{x,\xi,\alpha}(s), \alpha(s))ds.$$

After the change of variables $r = s - t$, the equation reads

$$J(x,\xi,\alpha) = I + e^{-\lambda t}\int_0^\infty e^{-\lambda r} L(y_{x,\xi,\alpha}(t+r), w_{x,\xi,\alpha}(t+r), \alpha(t+r))dr. \tag{6.2}$$

Using the semigroup property, (6.2) is equivalent to

$$J(x,\xi,\alpha) = I + e^{-\lambda t} J(y_{x,\xi,\alpha}(t), \xi_{x,\xi,\alpha}(t), \bar{\alpha}),$$

with $\bar{\alpha}(r) := \alpha(t+r)$. Since $J(y_{x,\xi,\alpha}(t), \xi_{x,\xi,\alpha}(t), \bar{\alpha}) \geq V(y_{x,\xi,\alpha}(t), \xi_{x,\xi,\alpha}(t))$, we can take the infimum to conclude $V(x,\xi) \geq W(x,\xi)$.

To prove the other inequality, fix $t > 0$, $\alpha \in \mathcal{A}$, $\epsilon > 0$, and choose $\alpha' \in \mathcal{A}$ such that

$$V(y_{x,\xi,\alpha}(t))w_{x,\xi,\alpha}(t)) \geq J(y_{x,\xi,\alpha}(t), w_{x,\xi,\alpha}(t), \alpha') - \epsilon.$$

Define the control

$$\bar{\alpha}(s) := \begin{cases} \alpha(s) & \text{if } s \le t, \\ \alpha'(s-t) & \text{else.} \end{cases}$$

Setting $\bar{x} := y_{x,\xi,\alpha}(t)$ and $\bar{\xi} := \xi_{x,\xi,\alpha}(t)$, we find

$$V(x,\xi) \le J(x,\xi,\bar{\alpha}) = I + e^{-\lambda t} J(\bar{x},\bar{\xi},\alpha') \le I + e^{-\lambda t} V(\bar{x},\bar{\xi}) + \epsilon.$$

Since ϵ and α are arbitrary, this shows $V(x,\xi) \le W(x,\xi)$.

□

Remark 6.2 *The proof is completely standard and can, e.g., also be found in [1, 4] with only some notational differences.*

□

6.2 The HJB inclusion and existence of solutions

For $p \in X$, $q \in L^2(R,\mu) := \prod_{k=1}^n L^2(R_k,\mu_k)$, $x \in X$, $\xi \in Z$, we define the Hamiltonian

$$H(x,\xi,p,q) := \sup_{a \in \mathbb{A}} \left\{ -\langle p, a\rangle - L(x,w_0,a) - \sum_{k=1}^n \langle q_k, \mathbb{1}\rangle_{R_k} \cdot \langle f_k, a\rangle \right\},$$

with $w_0 = w_0(x,\xi)$ as defined in (5.3). We are interested in solutions to the following formal HJB inclusion:

$$\begin{aligned} &\lambda\Phi(x,\xi) - \langle \nabla_x \Phi(x,\xi), Ax\rangle + \sum_{k=1}^n \langle \nabla_x \Phi(x,\xi), g_k\rangle \left(f_k(x) - \langle \xi_k, \mathbb{1}\rangle_{R_k}\right) \\ &\quad + \sum_{j,k=1}^n \langle \nabla_{\xi_k} \Phi(x,\xi), \mathbb{1}\rangle_{R_k} \langle f_k, g_j\rangle \left(f_j(x) - \langle \xi_j, \mathbb{1}\rangle\right) \\ &\quad - \sum_{k=1}^n \langle \nabla_{\xi_k} \Phi(x,\xi), \mathbb{1}\rangle_{R_k} \langle f_k, Ax\rangle \\ &\quad + \sum_{k=1}^n \langle \nabla_{\xi_k} \Phi(x,\xi), N_{Z_k}(\xi_k)\rangle_{R_k} + H(x,\xi,\nabla_x\Phi(x,\xi),\nabla_\xi\Phi(x,\xi)) \ni 0, \end{aligned} \tag{6.3}$$

where $N_{Z_k}(\xi_k)$ stands for the normal cone to Z_k at ξ_k. The following definitions are inspired by [6] for the x variable, [1, 28] for the hysteresis part.

Definition 6.3 *Let $C^1_\Phi(X)$ be the set of all test functions φ satisfying*

(Φ1) $\varphi \in C^1(X)$.

(Φ2) *For all $\theta \in [0, \frac{1}{2}]$, $\nabla\varphi(x) \in D((-A)^\theta)$ if and only if $x \in D((-A)^\theta)$.*

(Φ3) *The mapping $x \mapsto \nabla\varphi(x)$ is continuous from $D((-A)^\theta)$ into itself.*

□

Definition 6.4 *A function $v : X \times Z \to \mathbb{R}$ satisfying the Lipschitz estimate (5.5) with $\theta \geq \frac{1}{2}$ is called* ***viscosity subsolution*** *of (6.3), if for every $\varphi \in C^1_\Phi(X)$, $\psi \in C^1(Z)$ and $(x, \xi) \in \left(D((-A)^{\frac{1}{2}}) \times Z\right) \cap \arg\max(v - \varphi - \psi)$, there exist $p_k \in N_{Z_k}(\xi_k)$, $k \in \{1, \ldots, n\}$, such that*

$$\begin{aligned}
&\lambda v(x,\xi) + \left\langle (-A)^{\frac{1}{2}} \nabla\varphi(x), (-A)^{\frac{1}{2}} x \right\rangle \\
&\quad + \sum_{k=1}^{n} \langle \nabla\varphi(x), g_k \rangle \left(f_k(x) - \langle \xi_k, \mathbb{1} \rangle_{R_k} \right) \\
&\quad + \sum_{j,k=1}^{n} \langle \nabla_k \psi(\xi), \mathbb{1} \rangle_{R_k} \langle f_k, g_j \rangle \left(f_j(x) - \langle \xi_j, \mathbb{1} \rangle_{R_j} \right) \\
&\quad + \sum_{k=1}^{n} \langle \nabla_k \psi(\xi), \mathbb{1} \rangle_{R_k} \left\langle (-A)^{\frac{1}{2}} f_k, (-A)^{\frac{1}{2}} x \right\rangle \\
&\quad + \sum_{k=1}^{n} \langle \nabla_k \psi(\xi), p_k \rangle_{R_k} + H(x, \xi, \nabla\varphi(x), \nabla\psi(\xi)) \leq 0.
\end{aligned} \tag{6.4}$$

Similarly, a function $v : X \times Z \to \mathbb{R}$ satisfying (5.5) with $\theta \geq \frac{1}{2}$ is called ***viscosity supersolution*** *of (6.3), if for every $\varphi \in C^1_\Phi(X)$, $\psi \in C^1(Z)$ and $(x, \xi) \in \left(D((-A)^{\frac{1}{2}}) \times Z\right) \cap \arg\min(v - \varphi - \psi)$, there exist $q_k \in N_{Z_k}(\xi_k)$, $k \in \{1, \ldots, n\}$,* *such that*

$$
\begin{aligned}
&\lambda v(x,\xi) + \left\langle (-A)^{\frac{1}{2}} \nabla\varphi(x), (-A)^{\frac{1}{2}} x \right\rangle \\
&\quad + \sum_{k=1}^{n} \langle \nabla\varphi(x), g_k\rangle \left(f_k(x) - \langle \xi_k, \mathbb{1}\rangle_{R_k} \right) \\
&\quad + \sum_{j,k=1}^{n} \langle \nabla_k \psi(\xi), \mathbb{1}\rangle_{R_k} \langle f_k, g_j\rangle \left(f_j(x) - \langle \xi_j, \mathbb{1}\rangle_{R_j} \right) \\
&\quad + \sum_{k=1}^{n} \langle \nabla_k \psi(\xi), \mathbb{1}\rangle_{R_k} \left\langle (-A)^{\frac{1}{2}} f_k, (-A)^{\frac{1}{2}} x \right\rangle \\
&\quad + \sum_{k=1}^{n} \langle \nabla_k \psi(\xi), q_k\rangle_{R_k} + H(x,\xi,\nabla\varphi(x),\nabla\psi(\xi)) \geq 0.
\end{aligned}
\tag{6.5}
$$

Finally, a function is called ***viscosity solution*** *of* (6.3)*, if it is a viscosity subsolution and a viscosity supersolution of* (6.3).

□

The above definition makes sense because of the following statements (see [6, 29] and the references therein):

Lemma 6.5 *Let* $(x_0, \xi_0) \in X \times Z$ *and* $v : X \times Z \to \mathbb{R}$ *such that* (5.5) *holds for some* $\theta > 0$. *Then both the subdifferential* $D_x^- v(x_0,\xi_0)$ *and the superdifferential* $D_x^+ v(x_0,\xi_0)$ *are included in* $D((-A^*)^\theta) = D((-A)^\theta)$. *Moreover, if* C *denotes the Lipschitz constant in* (5.5)*, then*

$$\left\| (-A)^\theta p \right\| \leq C,\ \forall p \in D_x^+ v(x_0,\xi_0) \cup D_x^- v(x_0,\xi_0).$$

Proof: For fixed $\xi_0 \in Z$ define the function $\tilde{v} : X \to \mathbb{R}$, $\tilde{v}(x) := v(x,\xi_0)$. Then, from (5.5),

$$|\tilde{v}(x) - \tilde{v}(z)| \leq C \left\| (-A)^\theta (x-z) \right\|, \tag{6.6}$$

hence, we can apply [29, corollary 3.4] to $\tilde{v}$, which yields the result for functions in one variable. Since A is assumed to be self adjoint, we can write A instead of A^*, and the proof is complete.

□

Lemma 6.6 *Let* $\varphi,\ \psi \in C^1(X \times Z)$ *and* $w,\ v \in C(X \times Z)$. *If* (x_0,ξ_0) *is a local maximum for* $w - \varphi$, *then* $\nabla_x \varphi(x_0,\xi_0) \in D_x^+ w(x_0,\xi_0)$. *Similarly, if* (z_0,ζ_0) *is a local minimum for* $v - \psi$, *then* $\nabla_x \psi(z_0,\zeta_0) \in D_x^- v(z_0,\zeta_0)$.

Proof: As in the proof of the last lemma, we use (almost the same) arguments as [6, lemma 3.6] for the statement in one variable. If (x_0, ξ_0) is a local maximum for $w - \varphi$, then there is in particular some open set $U \subset X$ such that

$$\begin{aligned} w(x_0,\xi_0) - \varphi(x_0,\xi_0) \geq & w(x,\xi_0) - \varphi(x,\xi_0) \\ \Leftrightarrow w(x,\xi_0) - w(x_0,\xi_0) - & [\varphi(x,\xi_0) - \varphi(x_0,\xi_0)] \leq 0, \ \forall x \in U. \end{aligned} \tag{6.7}$$

Since φ is, by assumption, Frechet differentiable, it holds

$$\varphi(x,\xi_0) - \varphi(x_0,\xi_0) - \langle \nabla_x \varphi(x_0,\xi_0), x - x_0 \rangle = o(\|x - x_0\|). \tag{6.8}$$

Further, we can write

$$\begin{aligned} & \frac{w(x,\xi_0) - w(x_0,\xi_0) - \langle \nabla_x \varphi(x_0,\xi_0), x - x_0 \rangle}{\|x - x_0\|} \\ = & \frac{w(x,\xi_0) - w(x_0,\xi_0) - [\varphi(x,\xi_0) - \varphi(x_0,\xi_0)]}{\|x - x_0\|} \\ & + \frac{\varphi(x,\xi_0) - \varphi(x_0,\xi_0) - \langle \nabla_x \varphi(x_0,\xi_0), x - x_0 \rangle}{\|x - x_0\|}. \end{aligned}$$

By (6.7) and (6.8), application of $\limsup_{x \to x_0}$ shows $\nabla_x \varphi(x_0,\xi_0) \in D_x^+ w(x_0,\xi_0)$. The case when (z_0,ζ_0) is a local minimum for $v - \psi$ follows by similar arguments.

□

Since in definition 6.4 we assume that (5.5) holds with $\theta \geq \frac{1}{2}$, lemma 6.5 implies that the sub- and superdifferential of any viscosity subsolution (and supersolution, resp.) are included in $D((-A)^{\frac{1}{2}})$, and thus, by lemma 6.6, $\nabla \varphi(x) \in D((-A)^{\frac{1}{2}})$, if x is a local extremum. Noting assumption ($\Phi 2$), this implies $x \in D((-A)^{\frac{1}{2}})$, so that inequations (6.4) and (6.5) are meaningful.

Theorem 6.7 *Let the assumptions of theorem 5.1 hold, and $f_k \in D((-A)^\theta)$ for all $1 \leq k \leq n$ and $\theta \geq \frac{1}{2}$. Then the value function V (5.4) is a viscosity solution of (6.3) in the sense of definition 6.4.*

Proof: We first proof that V is a viscosity subsolution. To this end, let $\varphi \in C^1_\Phi(X)$, $\psi \in C^1(Z)$ and $(x,\xi) \in D((-A)^{\frac{1}{2}}) \times Z$ be a local maximum for $V - \varphi - \psi$. Then, particularly, for $t > 0$ small enough,

$$V(x,\xi) - \varphi(x) - \psi(\xi) \geq V(y_{x,\xi,a}(t), \xi_{x,\xi,a}(t)) - \varphi(y_{x,\xi,a}(t)) - \psi(\xi_{x,\xi,a}(t)).$$

The subscript a means that we choose a constant control $\alpha \equiv a \in \mathbb{A}$. Then, by the dynamic programming principle (theorem 6.1),

$$\begin{aligned}&\varphi(x) - \varphi(y_{x,\xi,a}(t)) + \psi(\xi) - \psi(\xi_{x,\xi,a}(t))\\ &\le V(x,\xi) - V(y_{x,\xi,a}(t), \xi_{x,\xi,a}(t))\\ &\le \int_0^t e^{-\lambda s} L(y_{x,\xi,a}(s), w_{x,\xi,a}(s), a) ds + \left(e^{-\lambda t} - 1\right) V(y_{x,\xi,a}(t), \xi_{x,\xi,a}(t)).\end{aligned} \tag{6.9}$$

Due to the regularity of φ and solutions to problem **(Q)**, using $(\Phi 3)$,

$$\varphi(x) - \varphi(y_{x,\xi,a}(t)) \tag{6.10}$$

$$\begin{aligned}&= \int_0^t \left\langle (-A)^{\frac{1}{2}} \nabla\varphi(y_{x,\xi,a}(s)), (-A)^{\frac{1}{2}} y_{x,\xi,a}(s) \right\rangle ds\\ &\quad + \sum_{k=1}^n \int_0^t \langle \nabla\varphi(y_{x,\xi,a}(s)), f_k(y_{x,\xi,a}(s)) g_k(s) \rangle ds\\ &\quad - \sum_{k=1}^n \int_0^t \left\langle \nabla\varphi(y_{x,\xi,a}(s)), \int_{R_k} \mathcal{E}_\rho[f_k(y_{x,\xi,a}); \xi](s) d\mu_k \cdot g_k(s) \right\rangle ds\\ &\quad - \int_0^t \langle \nabla\varphi(y_{x,\xi,a}(s)), a \rangle ds,\end{aligned} \tag{6.11}$$

see theorem 4.1. For ψ, we get

$$\psi(\xi) - \psi(\xi_{x,\xi,a}(t)) = -\int_0^t \left\langle \nabla\psi(\xi_{x,\xi,a}(s)), \left(\frac{d}{ds} f(y_{x,\xi,a}(s))\right) - \eta(s) \right\rangle_Z ds, \tag{6.12}$$

where $f(y) = (f_1(y), \ldots, f_n(y))^T$, and $\eta(s) = (\eta_1(s), \ldots, \eta_n(s))^T$, with $\eta_k(s) \in N_{Z_k}(\xi_{x,\xi,a}(s)) \cap B(0, \left|\frac{d}{ds} f_k(y_{x,\xi,a}(s))\right|)$ for each k (here, $B(0,r)$ stands for the closed ball around 0 with radius r in $L^2(R_k, \mu_k)$). This is because for almost every ρ, the derivative is bounded by $\left|\frac{d}{ds} f_k(y_{x,\xi,a}(s))\right|$, and μ_k is a probability measure. Note that

$$\begin{aligned}\frac{d}{ds} f_k(y_{x,\xi,a}(s)) &= \langle f_k, A y_{x,\xi,a}(s) - F(y_{x,\xi,a}; \xi)(s) + a \rangle\\ &= -\left\langle (-A)^{\frac{1}{2}} f_k, (-A)^{\frac{1}{2}} y_{x,\xi,a}(s) \right\rangle + \langle f_k, a - F(y_{x,\xi,a}; \xi)(s) \rangle,\end{aligned} \tag{6.13}$$

which continuously depends on s. Hence, it is bounded for $s \downarrow 0$, and (6.12) is meaningful. Moreover, this shows that there exists $M > 0$ (depending only on f_k and $\|x\|_{D((-A)^{\frac{1}{2}})}$) such that, for every k and $s > 0$ not too large, $\eta_k(s) \in N_{Z_k}(\xi_{x,\xi,a}(s)) \cap$

$B(0, M)$. We will need this property later. Altogether, we may write (6.12) as

$$\begin{aligned}
&\psi(\xi) - \psi(\xi_{x,\xi,a}(t)) \\
&= \sum_{k=1}^{n} \int_0^t \langle \nabla_k \psi(\xi_{x,\xi,a}(s)), \mathbb{1} \rangle_{R_k} \left\langle (-A)^{\frac{1}{2}} f_k, (-A)^{\frac{1}{2}} y_{x,\xi,a}(s) \right\rangle ds \\
&\quad - \sum_{k=1}^{n} \int_0^t \langle \nabla_k \psi(\xi_{x,\xi,a}(s)), \mathbb{1} \rangle_{R_k} \langle f_k, a - F(y_{x,\xi,a}; \xi)(s) \rangle \, ds \\
&\quad + \sum_{k=1}^{n} \int_0^t \langle \nabla_k \psi(\xi_{x,\xi,a}(s)), \eta(s) \rangle_{R_k} \, ds.
\end{aligned} \tag{6.14}$$

Plugging (6.11), (6.14) in (6.9) and dividing by t, we get the inequality

$$\begin{aligned}
&\frac{1}{t} \int_0^t \left\langle (-A)^{\frac{1}{2}} \nabla \varphi(y_{x,\xi,a}(s)), (-A)^{\frac{1}{2}} y_{x,\xi,a}(s) \right\rangle ds \\
&\quad + \sum_{k=1}^{n} \frac{1}{t} \int_0^t \langle \nabla \varphi(y_{x,\xi,a}(s)), f_k(y_{x,\xi,a}(s)) g_k \rangle \, ds \\
&\quad - \sum_{k=1}^{n} \frac{1}{t} \int_0^t \left\langle \nabla \varphi(y_{x,\xi,a}(s)), \int_{R_k} \mathcal{E}_\rho[f_k(y_{x,\xi,a}); \xi](s) d\mu_k \cdot g_k \right\rangle ds \\
&\quad - \frac{1}{t} \int_0^t \langle \nabla \varphi(y_{x,\xi,a}(s)), a \rangle \, ds \\
&\quad + \sum_{k=1}^{n} \frac{1}{t} \int_0^t \langle \nabla_k \psi(\xi_{x,\xi,a}(s)), \mathbb{1} \rangle_{R_k} \left\langle (-A)^{\frac{1}{2}} f_k, (-A)^{\frac{1}{2}} y_{x,\xi,a}(s) \right\rangle ds \\
&\quad - \sum_{k=1}^{n} \frac{1}{t} \int_0^t \langle \nabla_k \psi(\xi_{x,\xi,a}(s)), \mathbb{1} \rangle_{R_k} \langle f_k, a - F(y_{x,\xi,a}; \xi)(s) \rangle \, ds \\
&\quad + \sum_{k=1}^{n} \frac{1}{t} \int_0^t \langle \nabla_k \psi(\xi_{x,\xi,a}(s)), \eta_k(s) \rangle_{R_k} \, ds \\
&\leq \frac{1}{t} \int_0^t e^{-\lambda s} L(y_{x,\xi,a}(s), w_{x,\xi,a}(s), a) ds \\
&\quad + \left(\frac{e^{-\lambda t} - 1}{t} \right) V(y_{x,\xi,a}(t), \xi_{x,\xi,a}(t)).
\end{aligned} \tag{6.15}$$

We now take a closer look at the term containing η (the one in front of "$\leq$"). To this end, let $\{t_m\}_{m \in \mathbb{N}}$ be any sequence satisfying $t_m \downarrow 0$ as $m \to \infty$ and t_m small enough, so that (6.15) is valid for all m. For every k,

$$\begin{aligned}&\left|\frac{1}{t_m}\int_0^{t_m}\langle\nabla_k\psi(\xi_{x,\xi,a}(s)),\eta_k(s)\rangle_{R_k}\,ds-\frac{1}{t_m}\int_0^{t_m}\langle\nabla_k\psi(\xi),\eta_k(s)\rangle_{R_k}\,ds\right|\\&\le\frac{1}{t_m}\int_0^{t_m}\|\nabla_k\psi(\xi_{x,\xi,a}(s))-\nabla_k\psi(\xi)\|_{R_k}\|\eta_k(s)\|_{R_k}\,ds\\&\le M\sup_{s\in[0,t_m]}\|\nabla_k\psi(\xi_{x,\xi,a}(s))-\nabla_k\psi(\xi)\|_{R_k}\to 0,\end{aligned}\tag{6.16}$$

as $m\to\infty$, by continuity. Thus, it suffices to analyze the behavior of

$$\frac{1}{t_m}\int_0^{t_m}\langle\nabla_k\psi(\xi),\eta_k(s)\rangle_{R_k}\,ds=\left\langle\nabla_k\psi(\xi),\frac{1}{t_m}\int_0^{t_m}\eta_k(s)ds\right\rangle_{R_k}\tag{6.17}$$

for $m\to\infty$. Now, an application of Jensen's inequality shows that the sequence $\frac{1}{t_m}\int_0^{t_m}\eta_k(s)ds$ is bounded in $L^2(R_k,\mu_k)$:

$$\int_{R_k}\left(\frac{1}{t_m}\int_0^{t_m}\eta_k(s)ds\right)^2d\mu_k\le\int_{R_k}\frac{1}{t_m}\int_0^{t_m}\eta_k(s)^2dsd\mu_k\le M^2.$$

Thus, there exists a weakly convergent subsequence (for simplicity, we use the same index for the subsequence). Restricting ourselves to this subsequence, we may take the limit of (6.17); hence, what is left to show is that the weak limit η_k^* is an element of the normal cone $N_{Z_k}(\xi_k)$. Assume for contradiction that this was not the case. Then there exists $\epsilon>0$ and a set $E\subset R_k$ such that $\mu_k(E)=\epsilon$ and $\eta_k^*(\rho)\notin N_{[-\rho,\rho]}(\xi_k(\rho))$, for all $\rho\in E$ (note that in our setting, $N_{Z_k}(\xi_k)=\{n_\rho|n_\cdot:R_k\to\mathbb{R},\ n_\rho\in N_{[-\rho,\rho]}(\xi_k(\rho))\}$).

For $\gamma>0$ define the sets

$$\begin{aligned}\mathcal{M}_1^\gamma:=&\{\rho\in R_k|\ (\xi_k(\rho)=\rho\vee\xi_k(\rho)=-\rho)\wedge\gamma<2\rho\},\\\mathcal{M}_2^\gamma:=&\{\rho\in R_k|\ |\xi_k(\rho)-\rho|\ge\gamma\ \wedge\ |\xi_k(\rho)+\rho|\ge\gamma\},\\\mathcal{M}_3^\gamma:=&R_k\setminus(\mathcal{M}_1\cup\mathcal{M}_2^\gamma),\end{aligned}$$

and choose γ such that $\mu_k(\mathcal{M}_3^\gamma)<\frac{\epsilon}{4}$. By continuity, it holds that $(\xi_{x,\xi,a})_k(t_m)$ converges to ξ w.r.t. the $L^2(R_k,\mu_k)$ norm, which implies convergence μ_k almost everywhere, at least for some subsequence [21, theorem 3.3.13] (we may here, w.l.o.g., assume that $(t_m)_m$ was chosen properly). Hence, by Egorov's theorem [21, theorem 2.5.7], we can find a set $N\subset R_k$ of measure $\mu_k(N)<\frac{\epsilon}{4}$ such that $(\xi_{x,\xi,a})_k(t_m)\to\xi$ uniformly on N^c. Thus, there exists $\bar{m}\in\mathbb{N}$ such that $|(\xi_{x,\xi,a})_k(t_m)(\rho)-\xi_k(\rho)|<\frac{\gamma}{2}$, for all $m>\bar{m}$, $\rho\in N^c$. Now, for every $\mathcal{M}_1^\gamma\cap N^c$ and $m>\bar{m}$, $\eta_k(t_m)(\rho)\in N_{[-\rho,\rho]}(\xi_k(\rho))$, which, by convexity of the normal cone, implies $\eta_k^*(\rho)\in N_{[-\rho,\rho]}(\xi_k(\rho))$. Hence, $\eta_k^*(\rho)\in N_{[-\rho,\rho]}(\xi_k(\rho))$ for every $\rho\in\mathcal{M}_1^\gamma\cap N^c$.

Since on $\mathcal{M}_2^\gamma \cap N^c$, $\eta(t_m)$ is constant ($= 0$) for $m > \bar{m}$, we also get $\eta_k^*(\rho) \in N_{[-\rho,\rho]}(\xi_k(\rho))$ for every $\rho \in \mathcal{M}_2^\gamma \cap N^c$.

Altogether, we have shown that $\eta_k^*(\rho) \in N_{[-\rho,\rho]}(\xi_k(\rho))$, for all $\rho \in R_k \setminus (\mathcal{M}_3^\gamma \cup N)$. But $\mu_k(\mathcal{M}_3^\gamma) + \mu_k(N) < \frac{\epsilon}{2} < \mu_k(E)$, a contradiction.

Concerning the other terms in (6.15), we can give estimates of the type (6.16), so that we may replace the time dependent elements. Further, if we also replace t by t_m, we get the inequality

$$\begin{aligned}
&\left\langle (-A)^{\frac{1}{2}} \nabla\varphi(x), (-A)^{\frac{1}{2}} x \right\rangle \\
&\quad + \sum_{k=1}^{n} \langle \nabla\varphi(x), f_k(x) g_k \rangle \\
&\quad - \sum_{k=1}^{n} \left\langle \nabla\varphi(x), \int_{R_k} \xi_k(\rho) d\mu_k \cdot g_k \right\rangle ds \\
&\quad - \langle \nabla\varphi(x), a \rangle \, ds \\
&\quad + \sum_{k=1}^{n} \langle \nabla_k \psi(\xi), \mathbb{1} \rangle_{R_k} \left\langle (-A)^{\frac{1}{2}} f_k, (-A)^{\frac{1}{2}} x \right\rangle \\
&\quad - \sum_{k=1}^{n} \langle \nabla_k \psi(\xi), \mathbb{1} \rangle_{R_k} \langle f_k, a - F(y_{x,\xi,a}; \xi)(0) \rangle \\
&\quad + \sum_{k=1}^{n} \left\langle \nabla_k \psi(\xi), \frac{1}{t_m} \int_0^{t_m} \eta_k(s) ds \right\rangle_{R_k} \\
&\leq L(x, w_0, a) - \lambda V(x, \xi) + m \cdot o\left(\frac{1}{m}\right).
\end{aligned}$$

Thus, letting $m \to \infty$, and then applying $\sup_{a \in \mathbb{A}}$, yields (6.4) and the first part of the proof is complete. To prove that V is a viscosity supersolution, we need similar arguments. Let $\varphi \in C_\Phi^1(X)$, $\psi \in C^1(Z)$ and $(x, \xi) \in D((-A)^{\frac{1}{2}}) \times Z$ be a local minimum for $V - \varphi - \psi$. Then,

$$V(x, \xi) - \varphi(x) - \psi(\xi) \leq V(z, \zeta) - \varphi(z) - \psi(\zeta), \tag{6.18}$$

for all (z, ζ) in some ball around (x, ξ). By the dynamic programming principle (theorem 6.1), there exists, for every $\epsilon > 0$ and $t > 0$ a control $\alpha \in \mathcal{A}$, such that

$$V(x, \xi) + \epsilon t \geq \int_0^t e^{-\lambda s} L(y_{x,\xi,\alpha}(s), w_{x,\xi,\alpha}(s), \alpha(s)) ds + e^{-\lambda t} V(y_{x,\xi,\alpha}(t), \xi_{x,\xi,\alpha}(t)). \tag{6.19}$$

Thus, combining (6.18) and (6.19), for small $t > 0$,

$$\int_0^t e^{-\lambda s} L(y_{x,\xi,\alpha}(s), w_{x,\xi,\alpha}(s), \alpha(s)) ds$$

$$
\begin{aligned}
&+ e^{-\lambda t} V(y_{x,\xi,\alpha}(t), \xi_{x,\xi,\alpha}(t)) - \epsilon t - \varphi(x) - \psi(\xi) \\
&\leq V(y_{x,\xi,\alpha}(t), \xi_{x,\xi,\alpha}(t)) - \varphi(y_{x,\xi,\alpha}(t)) - \psi(\xi_{x,\xi,\alpha}(t)).
\end{aligned}
$$

Rearranging and dividing by t yields

$$
\begin{aligned}
&\frac{e^{-\lambda t} - 1}{t} V(y_{x,\xi,\alpha}(t), \xi_{x,\xi,\alpha}(t)) + \frac{1}{t} \int_0^t e^{-\lambda s} L(y_{x,\xi,\alpha}(s), w_{x,\xi,\alpha}(s), \alpha(s)) ds \\
&+ \frac{1}{t} \left(\varphi(y_{x,\xi,\alpha}(t)) - \varphi(x)\right) + \frac{1}{t} \left(\psi(\xi_{x,\xi,\alpha}(t)) - \psi(\xi)\right) \leq \epsilon.
\end{aligned}
$$

Just as before, (6.11) and (6.14) hold with a replaced by α. Again, we restrict ourself to some sequence $(t_m)_{m \in \mathbb{N}}$ with $t_m \downarrow 0$. Noting that $y_{x,\xi,\alpha}(t_m) \to x$ in $D((-A)^{\frac{1}{2}})$ as $m \to \infty$ uniformly in $\alpha \in \mathcal{A}$ as well as $\xi_{x,\xi,\alpha}(t_m) \to \xi$ uniformly in α, the same estimates and argumentation as in the first part of the proof yield weak convergence of $\eta(\cdot)$ to some element of the normal cone. Hence, replacing the converging time dependent elements (with just some error $o(\frac{1}{m})$) yields

$$
\begin{aligned}
&- \lambda V(x,\xi) - \left\langle (-A)^{\frac{1}{2}} \nabla\varphi(x), (-A)^{\frac{1}{2}} x \right\rangle \\
&- \sum_{k=1}^{n} \langle \nabla\varphi(x), g_k \rangle \left(f_k(x) - \langle \xi_k, \mathbb{1} \rangle_{R_k} \right) \\
&- \sum_{k=1}^{n} \langle \nabla_k \psi(\xi), \mathbb{1} \rangle_{R_k} \left\langle (-A)^{\frac{1}{2}} f_k, (-A)^{\frac{1}{2}} x \right\rangle \\
&- \sum_{j,k=1}^{n} \langle \nabla_k \psi(\xi), \mathbb{1} \rangle_{R_k} \langle f_k, g_j \rangle \left(f_j(x) - \langle \xi_j, \mathbb{1} \rangle_{R_k} \right) \\
&- \sum_{k=1}^{n} \langle \nabla_k \psi(\xi), q_k \rangle - \epsilon + m o\left(\frac{1}{m}\right) \\
&\leq \frac{1}{t_m} \int_0^{t_m} - \langle \nabla\varphi(x), \alpha(s) \rangle - L(x, w_0, \alpha(s)) - \sum_{k=1}^{n} \langle \nabla_k \psi(\xi), \mathbb{1} \rangle \langle f_k, \alpha(s) \rangle \, ds \\
&\leq H(x, \xi, \nabla\varphi(x), \nabla\psi(\xi)).
\end{aligned}
$$

Since $\epsilon > 0$ was arbitrary, we can conclude the proof by similar argumentation as in the subsolution case.

□

Remark 6.8 *The norms of the elements p, q of the normal cone are bounded; in fact, by (6.13), for each k and each ρ,*

$$
|p_k(\rho)|, |q_k(\rho)| \leq \left| \left\langle (-A)^{\frac{1}{2}} f_k, (-A)^{\frac{1}{2}} x \right\rangle \right| + |\langle f_k, a - F(x;\xi) \rangle| .
$$

□

7 Comparison result and uniqueness of viscosity solutions

Theorem 7.1 *Let v, w be, respectively, viscosity subsolution and viscosity supersolution of* (6.3) *in the sense of definition 6.4, and $f_k \in D((-A)^{\frac{3}{2}})$ for each $1 \le k \le n$. If $\lambda \ge 2\sum_{k=1}^{n} \|g_k\| \|f_k\|$, then $v \le w$.*

Proof: Assume for contradiction that $v \nleq w$. Then there exist $(x_0, \xi_0) \in X \times Z$ and $\delta > 0$, such that $v(x_0, \xi_0) - w(x_0, \xi_0) = \delta$. For $\epsilon > 0$ and $\mu > 0$, define the function $\bar{\Phi} : (X \times Z)^2 \to \mathbb{R}$ by

$$\begin{aligned}\bar{\Phi}(x, \xi, y, \zeta) := v(x, \xi) - w(y, \zeta) - \frac{1}{2\epsilon}\left\langle (-A)^{-1}(x - y), x - y \right\rangle \\ - \frac{\mu}{2}\left(\|x\|^2 + \|y\|^2\right) - \frac{1}{2\epsilon}\|\xi - \zeta\|_Z^2 .\end{aligned}$$

Since $\bar{\Phi}(x_0, \xi_0, x_0, \xi_0) = v(x_0, \xi_0) - w(x_0, \xi_0) - \mu\|x_0\|^2$, we have

$$\sup \bar{\Phi} \ge \bar{\Phi}(x_0, \xi_0, x_0, \xi_0) > \frac{\delta}{2}, \text{ for } 0 < \mu < \frac{\delta}{2\|x_0\|^2}, \tag{7.1}$$

with the convention $\frac{1}{0} = \infty$. As ususal, we need to find points where the supremum is attained. By the Lipschitz condition (5.5) (theorem 5.1), we see that $\bar{\Phi}$ is weakly upper semicontinuous w.r.t. the x and y variable, but we don't have such a property for ξ and ζ. To overcome this problem, define the auxiliary function

$$\Phi^*(\xi, \zeta) := \sup_{(x,y) \in X^2} \bar{\Phi}(x, \xi, y, \zeta).$$

It is well defined, because for $\mu > 0$, by the Lipschitz continuity of v and w, to find the supremum, one can restrict oneself to some ball around zero and $\bar{\Phi}$ is locally bounded. Further, Φ^* is Lipschitz continuous, because for any $(\xi_1, \zeta_1) \in Z^2$, $(\xi_2, \zeta_2) \in Z^2$,

$$\begin{aligned}|\Phi^*(\xi_1, \zeta_1) - \Phi^*(\xi_2, \zeta_2)| &\le \sup_{(x,y) \in X^2} \left\{ \left|\bar{\Phi}(x, \xi_1, y, \zeta_1) - \bar{\Phi}(x, \xi_2, y, \zeta_2)\right| \right\} \\ &\le \sup_{(x,y) \in X^2} \left\{ |v(x, \xi_1) - v(x, \xi_2)| + |w(y, \zeta_1) - w(y, \zeta_2)| \right. \\ &\qquad \left. + \frac{1}{2\epsilon}\left| \|\xi_1 - \zeta_1\|_Z^2 - \|\xi_2 - \zeta_2\|_Z^2 \right| \right\}.\end{aligned} \tag{7.2}$$

Using the Lipschitz continuity of v and w, and that

$$\left| \|\xi_1 - \zeta_1\|_Z^2 - \langle \xi_1 - \zeta_1, \xi_2 - \zeta_2 \rangle_Z + \langle \xi_1 - \zeta_1, \xi_2 - \zeta_2 \rangle_Z - \|\xi_2 - \zeta_2\|_Z^2 \right|$$

$$\begin{aligned}
&\leq \|\xi_1 - \zeta_1\|_Z \left(\|\xi_1 - \xi_2\|_Z + \|\zeta_1 - \zeta_2\|_Z\right) \\
&\quad + \|\xi_2 - \zeta_2\|_Z \left(\|\xi_1 - \xi_2\|_Z + \|\zeta_1 - \zeta_2\|_Z\right) \\
&\leq c \cdot \left(\|\xi_1 - \xi_2\|_Z + \|\zeta_1 - \zeta_2\|_Z\right),
\end{aligned}$$

(recall that Z is a bounded subset of $L^2(R,\mu)$), (7.2) thus implies that there is $C > 0$ such that

$$|\Phi^*(\xi_1,\zeta_1) - \Phi^*(\xi_2,\zeta_2)| \leq C\left(\|\xi_1 - \xi_2\|_Z + \|\zeta_1 - \zeta_2\|_Z\right).$$

Since Z is also convex and closed subset of $L^2(R,\mu)$, we may apply the theorem from [30, p. 8], which asserts that for each $\gamma > 0$ there are $\nu_1, \nu_2 \in L^2(R,\mu)$ with $\|\nu_1\|_Z + \|\nu_2\|_Z < \gamma$, such that the function

$$(\xi,\zeta) \mapsto \Phi^*(\xi,\zeta) + \langle \nu_1, \xi\rangle_Z + \langle \nu_2, \zeta\rangle_Z$$

attains its supremum on Z. In particular, there are $\bar{\xi}, \bar{\zeta} \in Z$ for which

$$\begin{aligned}
\Phi_m :=& \Phi^*(\bar{\xi},\bar{\zeta}) + \langle \nu_1, \bar{\xi}\rangle_Z + \langle \nu_2, \bar{\zeta}\rangle_Z \\
\geq& \Phi^*(\xi,\zeta) + \langle \nu_1, \xi\rangle_Z + \langle \nu_2, \zeta\rangle_Z \\
\geq& \bar{\Phi}(x,\xi,y,\zeta) + \langle \nu_1, \xi\rangle_Z + \langle \nu_2, \zeta\rangle_Z,
\end{aligned} \tag{7.3}$$

for all $(x,\xi,y,\zeta) \in (X \times Z)^2$. Now we choose a maximizing sequence for $\bar{\Phi}$, i.e., $((x_n,y_n))_{n\in\mathbb{N}} \in X^2$ and $\bar{\Phi}(x_n,\bar{\xi},y_n,\bar{\zeta}) \to \Phi^*(\bar{\xi},\bar{\zeta})$ as $n \to \infty$. As stated above, such sequence must be bounded and hence, w.l.o.g., converges weakly to some $(\bar{x},\bar{y}) \in X^2$, so that by the weak upper semicontinuity of $(x,y) \mapsto \bar{\Phi}(x,\bar{\xi},y,\bar{\zeta})$,

$$\begin{aligned}
\bar{\Phi}(\bar{x},\bar{\xi},\bar{y},\bar{\zeta}) + \langle \nu_1, \bar{\xi}\rangle_Z + \langle \nu_2, \bar{\zeta}\rangle_Z \geq& \limsup_{n\to\infty} \bar{\Phi}(x_n,\bar{\xi},y_n,\bar{\zeta}) + \langle \nu_1, \bar{\xi}\rangle_Z + \langle \nu_2, \bar{\zeta}\rangle_Z \\
=& \sup_{x,y\in X} \bar{\Phi}(x,\bar{\xi},y,\bar{\zeta}) + \langle \nu_1, \bar{\xi}\rangle_Z + \langle \nu_2, \bar{\zeta}\rangle_Z \\
=& \Phi^*(\bar{\xi},\bar{\zeta}) + \langle \nu_1, \bar{\xi}\rangle_Z + \langle \nu_2, \bar{\zeta}\rangle_Z = \Phi_m.
\end{aligned}$$

So, (7.3) implies that $(\bar{x},\bar{\xi},\bar{y},\bar{\zeta})$ maximizes

$$\Phi(x,\xi,y,\zeta) := \bar{\Phi}(x,\xi,y,\zeta) + \langle \nu_1, \xi\rangle_Z + \langle \nu_2, \zeta\rangle_Z \tag{7.4}$$

on $(X \times Z)^2$. Hence, we have shown that for every $\gamma > 0$ there exist $\nu_1, \nu_2 \in L^2(R,\mu) \cap B(0,\gamma)$ such that the corresponding function Φ defined in (7.4) attains its maximum on $(X \times Z)^2$. In view of (7.1), since Z is bounded, we may choose γ so small that

$$\sup \Phi \geq \frac{\delta}{4}, \text{ for } 0 < \mu \leq \frac{\delta}{\|x_0\|^2} \text{ and } \max\left\{\|\nu_1\|_Z, \|\nu_2\|_z\right\} < \gamma.$$

From the inequality

$$\Phi(\bar{x}, \bar{\xi}, \bar{x}, \bar{\xi}) + \Phi(\bar{y}, \bar{\zeta}, \bar{y}, \bar{\zeta}) \le 2\Phi(\bar{x}, \bar{\xi}, \bar{y}, \bar{\zeta}),$$

we infer that

$$\begin{aligned} &v(\bar{x}, \bar{\xi}) - w(\bar{x}, \bar{\xi}) - \mu \|\bar{x}\|^2 + \langle \nu_1 + \nu_2, \bar{\xi} \rangle_Z + v(\bar{y}, \bar{\zeta}) - w(\bar{y}, \bar{\zeta}) - \mu \|\bar{y}\|^2 \\ &+ \langle \nu_1 + \nu_2, \bar{\zeta} \rangle_Z \le 2v(\bar{x}, \bar{\xi}) - 2w(\bar{y}, \bar{\zeta}) - \frac{1}{\epsilon} \langle (-A)^{-1}(\bar{x} - \bar{y}), \bar{x} - \bar{y} \rangle \\ &- \mu \left(\|\bar{x}\|^2 + \|\bar{y}\|^2 \right) - \frac{1}{\epsilon} \|\bar{\xi} - \bar{\zeta}\|_Z^2 + 2 \langle \nu_1, \bar{\xi} \rangle_Z + 2 \langle \nu_2, \bar{\zeta} \rangle_Z , \end{aligned}$$

which implies

$$\begin{aligned} &\frac{1}{\epsilon} \langle (-A)^{-1}(\bar{x} - \bar{y}), \bar{x} - \bar{y} \rangle + \frac{1}{\epsilon} \|\bar{\xi} - \bar{\zeta}\|_Z^2 \\ &\quad \le v(\bar{x}, \bar{\xi}) - v(\bar{y}, \bar{\zeta}) + w(\bar{x}, \bar{\xi}) - w(\bar{y}, \bar{\zeta}) + \langle \nu_1 - \nu_2, \bar{\xi} - \bar{\zeta} \rangle_Z \\ &\quad \le C \left(\left\| (-A)^{-\frac{1}{2}}(\bar{x} - \bar{y}) \right\| + \|\bar{\xi} - \bar{\zeta}\|_Z \right), \end{aligned} \tag{7.5}$$

for some $C > 0$ independent of μ and ϵ (we have used the Lipschitz property of v, w, and the boundedness of ν_1, ν_2). Noting that

$$\langle (-A)^{-1}(\bar{x} - \bar{y}), \bar{x} - \bar{y} \rangle = \left\| (-A)^{-\frac{1}{2}}(\bar{x} - \bar{y}) \right\|^2 ,$$

application of Young's inequality to the right hand side of (7.5) yields

$$\begin{aligned} \frac{1}{\epsilon} \left\| (-A)^{-\frac{1}{2}}(\bar{x} - \bar{y}) \right\|^2 + \frac{1}{\epsilon} \|\bar{\xi} - \bar{\zeta}\|_Z^2 \le C^2 \epsilon + \frac{1}{2\epsilon} \left\| (-A)^{-\frac{1}{2}}(\bar{x} - \bar{y}) \right\|^2 \\ + \frac{1}{2\epsilon} \|\bar{\xi} - \bar{\zeta}\|_Z^2 , \end{aligned}$$

hence,

$$\frac{1}{\epsilon^2} \left\| (-A)^{-\frac{1}{2}}(\bar{x} - \bar{y}) \right\|^2 + \frac{1}{\epsilon^2} \|\bar{\xi} - \bar{\zeta}\|_Z^2 \le 2C^2. \tag{7.6}$$

Now, we define the functions

$$\begin{aligned} \varphi_1(x) &:= w(\bar{y}, \bar{\zeta}) + \frac{1}{2\epsilon} \langle (-A)^{-1}(x - \bar{y}), x - \bar{y} \rangle + \frac{\mu}{2} \left(\|x\|^2 + \|\bar{y}\|^2 \right), \\ \psi_1(\xi) &:= \frac{1}{2\epsilon} \|\xi - \bar{\zeta}\|_Z^2 - \langle \nu_1, \xi \rangle_Z - \langle \nu_2, \bar{\zeta} \rangle_Z , \\ \varphi_2(y) &:= v(\bar{x}, \bar{\xi}) - \frac{1}{2\epsilon} \langle (-A)^{-1}(\bar{x} - y), \bar{x} - y \rangle - \frac{\mu}{2} \left(\|\bar{x}\|^2 + \|y\|^2 \right), \end{aligned}$$

$$\psi_2(\zeta) := -\frac{1}{2\epsilon}\left\|\bar{\xi}-\zeta\right\|_Z^2 + \left\langle \nu_1, \bar{\xi}\right\rangle + \left\langle \nu_2, \zeta\right\rangle.$$

By construction, $v-\varphi_1-\psi_1$ attains its maximum at $(\bar{x},\bar{\xi})$, and $w-\varphi_2-\psi_2$ attains its minimum at $(\bar{y},\bar{\zeta})$. The derivatives are

$$\begin{aligned}
\nabla\varphi_1(\bar{x}) =& \frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y}) + \mu\bar{x},\\
\nabla\psi_1(\bar{\xi}) =& \frac{1}{\epsilon}(\bar{\xi}-\bar{\zeta}) - \nu_1,\\
\nabla\varphi_2(\bar{y}) =& \frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y}) - \mu\bar{y},\\
\nabla\psi_2(\bar{\zeta}) =& \frac{1}{\epsilon}(\bar{\xi}-\bar{\zeta}) + \nu_2,
\end{aligned}$$

and one immediately sees that $\varphi_1,\varphi_2 \in C^1_\Phi(X)$ (and, in particular, by lemmata 6.5, 6.6, $\bar{x},\bar{y}\in D((-A)^{\frac{1}{2}})$) and $\psi_1,\psi_2\in C^1(Z)$. Hence, since v is a subsolution and w a supersolution,

$$\begin{aligned}
&\lambda v(\bar{x},\bar{\xi}) + \left\langle (-A)^{\frac{1}{2}}\left[\frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y})+\mu\bar{x}\right], (-A)^{\frac{1}{2}}\bar{x}\right\rangle\\
&+\sum_{k=1}^{n}\left\langle \frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y})+\mu\bar{x}, g_k\right\rangle\left(f_k(\bar{x})-\left\langle\bar{\xi}_k,\mathbb{1}\right\rangle_{R_k}\right)\\
&+\sum_{j,k=1}^{n}\left\langle\frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k)-\nu_1^k,\mathbb{1}\right\rangle_{R_k}\langle f_k,g_j\rangle\left(f_j(\bar{x})-\left\langle\bar{\xi}_j,\mathbb{1}\right\rangle_{R_j}\right)\\
&+\sum_{k=1}^{n}\left\langle\frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k)-\nu_1^k,\mathbb{1}\right\rangle_{R_k}\left\langle(-A)^{\frac{1}{2}}f_k,(-A)^{\frac{1}{2}}\bar{x}\right\rangle\\
&+\sum_{k=1}^{n}\left\langle\frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k)-\nu_1^k,p_k\right\rangle_{R_k}\\
&+\sup_{a\in\mathbb{A}}\left\{-\left\langle\frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y})+\mu\bar{x},a\right\rangle - L(\bar{x},\bar{w}_0,a)\right.\\
&\left.-\sum_{k=1}^{n}\left\langle\frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k)-\nu_1^k,\mathbb{1}\right\rangle_{R_k}\langle f_k,a\rangle\right\} \le 0 \le \lambda w(\bar{y},\bar{\zeta})\\
&+\left\langle(-A)^{\frac{1}{2}}\left[\frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y})-\mu\bar{y}\right],(-A)^{\frac{1}{2}}\bar{y}\right\rangle\\
&+\sum_{k=1}^{n}\left\langle\frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y})-\mu\bar{y},g_k\right\rangle\left(f_k(\bar{y})-\left\langle\bar{\zeta}_k,\mathbb{1}\right\rangle_{R_k}\right)
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{j,k=1}^{n}\left\langle \frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k)+\nu_2^k,\mathbb{1}\right\rangle_{R_k}\langle f_k,g_j\rangle\left(f_j(\bar{y})-\left\langle\bar{\zeta}_j,\mathbb{1}\right\rangle_{R_j}\right)\\
&+\sum_{k=1}^{n}\left\langle \frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k)+\nu_2^k,\mathbb{1}\right\rangle_{R_k}\left\langle(-A)^{\frac{1}{2}}f_k,(-A)^{\frac{1}{2}}\bar{y}\right\rangle\\
&+\sum_{k=1}^{n}\left\langle \frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k)+\nu_2^k,q_k\right\rangle_{R_k}\\
&+\sup_{a\in\mathbb{A}}\left\{-\left\langle\frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y})-\mu\bar{y},a\right\rangle-L(\bar{y},\tilde{w}_0,a)\right.\\
&\left.-\sum_{k=1}^{n}\left\langle \frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k)+\nu_2^k,\mathbb{1}\right\rangle_{R_k}\langle f_k,a\rangle\right\}.
\end{aligned}
$$

Consequently,

$$\lambda\left(v(\bar{x},\bar{\xi})-w(\bar{y},\bar{\zeta})\right)\leq-\frac{1}{\epsilon}\left\|\bar{x}-\bar{y}\right\|^2-\mu\left(\left\|(-A)^{\frac{1}{2}}\bar{x}\right\|^2+\left\|(-A)^{\frac{1}{2}}\bar{y}\right\|^2\right) \tag{7.7}$$

$$+\sum_{k=1}^{n}\left\langle\frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y}),g_k\right\rangle\left(f_k(\bar{y}-\bar{x})+\left\langle\bar{\xi}_k-\bar{\zeta}_k,\mathbb{1}\right\rangle_{R_k}\right) \tag{7.8}$$

$$-\mu\sum_{k=1}^{n}\left[\langle\bar{x},g_k\rangle\left(f_k(\bar{x})-\left\langle\bar{\xi}_k,\mathbb{1}\right\rangle_{R_k}\right)+\langle\bar{y},g_k\rangle\left(f_k(\bar{y})-\left\langle\bar{\zeta}_k,\mathbb{1}\right\rangle_{R_k}\right)\right] \tag{7.9}$$

$$+\sum_{j,k=1}^{n}\left\langle\frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k),\mathbb{1}\right\rangle_{R_k}\langle f_k,g_j\rangle\left(f_j(\bar{y}-\bar{x})+\left\langle\bar{\xi}_j-\bar{\zeta}_j,\mathbb{1}\right\rangle_{R_j}\right) \tag{7.10}$$

$$+\sum_{j,k=1}^{n}\langle f_k,g_j\rangle\left[\left\langle\nu_2^k,\mathbb{1}\right\rangle_{R_k}\left(f_j(\bar{y})-\left\langle\bar{\zeta}_j,\mathbb{1}\right\rangle_{R_j}\right)+\left\langle\nu_1^k,\mathbb{1}\right\rangle_{R_k}\left(f_j(\bar{x})-\left\langle\bar{\xi}_j,\mathbb{1}\right\rangle_{R_j}\right)\right] \tag{7.11}$$

$$+\sum_{k=1}^{n}\left\langle\frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k),\mathbb{1}\right\rangle_{R_k}\left\langle(-A)^{\frac{1}{2}}f_k,(-A)^{\frac{1}{2}}(\bar{y}-\bar{x})\right\rangle \tag{7.12}$$

$$+\sum_{k=1}^{n}\left[\left\langle\nu_1^k,\mathbb{1}\right\rangle_{R_k}\left\langle(-A)^{\frac{1}{2}}f_k,(-A)^{\frac{1}{2}}\bar{x}\right\rangle+\left\langle\nu_2^k,\mathbb{1}\right\rangle_{R_k}\left\langle(-A)^{\frac{1}{2}}f_k,(-A)^{\frac{1}{2}}\bar{y}\right\rangle\right] \tag{7.13}$$

$$+\sum_{k=1}^{n}\left\langle\frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k),q_k-p_k\right\rangle_{R_k} \tag{7.14}$$

$$+\sum_{k=1}^{n}\left\langle\nu_1^k,p_k\right\rangle_{R_k}+\left\langle\nu_2^k,q_k\right\rangle_{R_k} \tag{7.15}$$

$$
\begin{aligned}
&+\sup_{a\in\mathbb{A}}\left\{-\left\langle\frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y})-\mu\bar{y},a\right\rangle-L(\bar{y},\tilde{w}_0,a)\right.\\
&\left.-\sum_{k=1}^{n}\left\langle\frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k)+\nu_2^k,\mathbb{1}\right\rangle_{R_k}\langle f_k,a\rangle\right\}
\end{aligned}
\tag{7.16}
$$

$$
\begin{aligned}
&-\sup_{a\in\mathbb{A}}\left\{-\left\langle\frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y})+\mu\bar{x},a\right\rangle-L(\bar{x},\bar{w}_0,a)\right.\\
&\left.-\sum_{k=1}^{n}\left\langle\frac{1}{\epsilon}(\bar{\xi}_k-\bar{\zeta}_k)-\nu_1^k,\mathbb{1}\right\rangle_{R_k}\langle f_k,a\rangle\right\}.
\end{aligned}
\tag{7.17}
$$

We are going to estimate (7.8) to (7.15) and (7.16)+(7.17). To simplify notations, in what follows, C will always stand for some constant independent of ϵ,μ and ν_1,ν_2 (with norms $\leq\gamma$). Since

$$
\begin{aligned}
\left\langle(-A)^{-1}(\bar{x}-\bar{y}),g_k\right\rangle\leq&\left\|(-A)^{-\frac{1}{2}}(-A)^{-\frac{1}{2}}(\bar{x}-\bar{y})\right\|\|g_k\|\\
\leq&C\left\|(-A)^{-\frac{1}{2}}(\bar{x}-\bar{y})\right\|,
\end{aligned}
$$

we get from (7.6) that

$$
\begin{aligned}
(7.8)\leq&\sum_{k=1}^{n}C\left(\left\|(-A)^{\frac{1}{2}}f_k\right\|\left\|(-A)^{-\frac{1}{2}}(\bar{y}-\bar{x})\right\|+\|\bar{\xi}_k-\bar{\zeta}_k\|_{R_k}\right)\\
\leq&C\left(\left\|(-A)^{-\frac{1}{2}}(\bar{y}-\bar{x})\right\|+\|\bar{\xi}-\bar{\zeta}\|_Z\right).
\end{aligned}
$$

Now, (7.6) implies that the right hand side converges to zero for $\epsilon\downarrow 0$. Thus,

$$
(7.8)\leq\beta(\epsilon),\tag{7.18}
$$

where $\beta:\mathbb{R}_+\to\mathbb{R}_+$ is some monotonic increasing function which is continuous in 0 and $\beta(0)=0$. Concerning (7.9), we split the term into

$$
-\mu\sum_{k=1}^{n}\left[\langle\bar{x},g_k\rangle f_k(\bar{x})+\langle\bar{y},g_k\rangle f_k(\bar{y})\right],\tag{7.19}
$$

and

$$
\mu\sum_{k=1}^{n}\left[\langle\bar{x},g_k\rangle\left\langle\bar{\xi}_k,\mathbb{1}\right\rangle_{R_k}+\langle\bar{y},g_k\rangle\left\langle\bar{\zeta}_k,\mathbb{1}\right\rangle_{R_k}\right].\tag{7.20}
$$

Now, the Cauchy Schwartz inequality implies

$$
(7.19)\leq\mu\left(\sum_{k=1}^{n}\|g_k\|\,\|f_k\|\right)\left(\|\bar{x}\|^2+\|\bar{y}\|^2\right),
$$

and, since Z is a bounded set,

$$(7.20) \leq C\mu \left(\|\bar{x}\| + \|\bar{y}\|\right).$$

Using Young's inequality, we can further estimate

$$(7.20) \leq C\mu^{\frac{1}{2}} + \mu^{\frac{3}{2}} \left(\|\bar{x}\|^2 + \|\bar{y}\|^2\right). \tag{7.21}$$

For (7.10), we may use (7.6) to find , similar to the estimation of (7.8), a function β with the same properties as before and

$$(7.10) \leq \beta(\epsilon).$$

Further, it is easy to see that

$$(7.11) \leq C \|\nu_1\|_Z \left(\|\bar{x}\| + 1\right) + C \|\nu_2\|_Z \left(\|\bar{y}\| + 1\right).$$

Hence, if ν_1, ν_2 are chosen such that $\|\nu_1\|_Z, \|\nu_2\|_Z \leq \mu$, we can estimate as in (7.21) to get

$$(7.11) \leq C(\mu + \mu^{\frac{1}{2}}) + \mu^{\frac{3}{2}} \left(\|\bar{x}\|^2 + \|\bar{y}\|^2\right). \tag{7.22}$$

The space $D((-A)^{\frac{1}{2}})$ is imbedded into X, so that there exists $\kappa > 0$, such that

$$\|x\| \leq \kappa \left\|(-A)^{\frac{1}{2}} x\right\|, \; \forall\, x \in D((-A)^{\frac{1}{2}}),$$

(one may also see that such an inequality holds by application of the reverse triangle inequality to (4.3)) hence, from (7.22),

$$(7.11) \leq C(\mu + \mu^{\frac{1}{2}}) + \kappa\mu^{\frac{3}{2}} \left(\left\|(-A)^{\frac{1}{2}}\bar{x}\right\|^2 + \left\|(-A)^{\frac{1}{2}}\bar{y}\right\|^2\right). \tag{7.23}$$

To estimate (7.12), note that $\left\langle \frac{1}{\varepsilon}(\bar{\xi}_k - \bar{\zeta}_k), \mathbb{1} \right\rangle_{R_k}$ is bounded as $\varepsilon \downarrow 0$ by (7.6), and that

$$\begin{aligned}\left|\left\langle (-A)^{\frac{1}{2}} f_k, (-A)^{\frac{1}{2}}(\bar{y} - \bar{x})\right\rangle\right| &= \left|\left\langle (-A)^{\frac{3}{2}} f_k, (-A)^{-\frac{1}{2}}(\bar{y} - \bar{x})\right\rangle\right| \\ &\leq \left\|(-A)^{\frac{3}{2}} f_k\right\| \left\|(-A)^{-\frac{1}{2}}(\bar{x} - \bar{y})\right\|.\end{aligned}$$

Thus, by assumption on f_k and (7.6),

$$(7.12) \leq \beta(\varepsilon).$$

Next, let $z \in X$ be arbitrary. Then, from lemmata 6.5, 6.6, we infer that there is a constant $C > 0$ such that $\left\|(-A)^{\frac{1}{2}}\nabla\varphi_1(\bar{x})\right\| \leq C$, hence,

$$\left|\left\langle (-A)^{\frac{1}{2}}\left[\frac{1}{\epsilon}(-A)^{-1}(\bar{x}-\bar{y})+\mu\bar{x}\right], z\right\rangle\right| \leq C\left\|z\right\|.$$

Thus, the reverse triangle inequality yields

$$\begin{aligned}\left|\left\langle (-A)^{\frac{1}{2}}\bar{x}, z\right\rangle\right| \leq & \frac{1}{\mu}\left[\left\|\frac{1}{\epsilon}(-A)^{-\frac{1}{2}}(\bar{x}-\bar{y})\right\| + C\right]\left\|z\right\| \\ \leq & \frac{C}{\mu}\left\|z\right\|, \end{aligned} \tag{7.24}$$

where we used (7.6) in the last step. A similar inequality holds with $\bar{x}$ replaced by $\bar{y}$. Using (7.24), we can now estimate

$$(7.13) \leq \frac{C}{\mu}\left(\left\|\nu_1\right\|_Z + \left\|\nu_2\right\|_Z\right) \leq C\mu,$$

if ν_1, ν_2 are chosen such that $\left\|\nu_i\right\|_Z \leq \mu^2$ (w.l.o.g., we assume that $\mu < 1$, because then, the latter implies $\left\|\nu_i\right\|_Z \leq \mu$, which is what we used to derive (7.22)). Further,

$$(7.14) \leq 0,$$

because for every k and $\rho \in R_k$, $p_k(\rho)$ is an element of the normal cone of the convex set $[-\rho, \rho]$ at $\xi_k(\rho)$, and analogously for $q_k(\rho)$ and $\zeta_k(\rho)$; hence, the convexity inequality $\langle \xi_k - \zeta_k, q_k - p_k\rangle_{R_k} \leq 0$ holds pointwise almost everywhere. The term (7.15) may be estimated by

$$(7.15) \leq C\mu,$$

since the norms of p_k, q_k are bounded if $\left\|(-A)^{\frac{1}{2}}\bar{x}\right\|$, $\left\|(-A)^{\frac{1}{2}}\bar{x}\right\|$ are bounded (see remark 6.8), and (7.24) holds. So, what is left is the difference of the hamiltonians, (7.16)+(7.17). First note that

$$\begin{aligned}\left|(7.16)+(7.17)\right| \leq \sup_{a\in\mathbb{A}}\{\ & \mu\left|\langle \bar{x}+\bar{y}, a\rangle\right| + \left|L(\bar{x},\bar{w}_0,a) - L(\bar{y},\tilde{w}_0,a)\right| \\ & + \left(\left\|\nu_1\right\|_Z + \left\|\nu_2\right\|_Z\right)\left|\langle f_k, a\rangle\right|\}.\end{aligned}$$

By the Lipschitz continuity of L, the definition of w_0 and the bounds on $\mathbb{A}$ and ν_i, the latter implies

$$\left|(7.16)+(7.17)\right| \leq C\mu(\left\|\bar{x}\right\| + \left\|\bar{y}\right\|) + C\left(\left\|\bar{x}-\bar{y}\right\| + \left\|\xi-\zeta\right\|_Z\right) + C\mu.$$

Using Young's inequality twice, we infer that

$$|(7.16)+(7.17)| \leq C(\mu^{\frac{1}{2}}+\mu)+C\mu^{\frac{3}{2}}\left(\|\bar{x}\|^2+\|\bar{y}\|^2\right)+C\frac{\varepsilon}{2}+\frac{1}{2\varepsilon}\|\bar{x}-\bar{y}\|^2. \quad (7.25)$$

and (7.25) implies

$$|(7.16)+(7.17)| \leq C(\mu^{\frac{1}{2}}+\mu)+C\mu^{\frac{3}{2}}\left(\left\|(-A)^{\frac{1}{2}}\bar{x}\right\|^2+\left\|(-A)^{\frac{1}{2}}\bar{y}\right\|^2\right) + \beta(\varepsilon)+\frac{1}{2\varepsilon}\|\bar{x}-\bar{y}\|^2.$$

Now we plug in all the estimates we have just collected and find

$$\begin{aligned} v(\bar{x},\bar{\xi})-w(\bar{y},\bar{\zeta}) \leq & -\frac{\mu}{\lambda}\left(\left\|(-A)^{\frac{1}{2}}\bar{x}\right\|^2+\left\|(-A)^{\frac{1}{2}}\bar{y}\right\|^2\right)+\beta(\epsilon) \\ & +\frac{\mu}{\lambda}\left(\sum_{k=1}^n\|g_k\|\,\|f_k\|\right)\left(\|\bar{x}\|^2+\|\bar{y}\|^2\right) \\ & +C(\mu+\mu^{\frac{1}{2}})+C\frac{\mu^{\frac{3}{2}}}{\lambda}\left(\left\|(-A)^{\frac{1}{2}}\bar{x}\right\|^2+\left\|(-A)^{\frac{1}{2}}\bar{y}\right\|^2\right) \\ & +C\mu^{\frac{3}{2}}\left(\left\|(-A)^{\frac{1}{2}}\bar{x}\right\|^2+\left\|(-A)^{\frac{1}{2}}\bar{y}\right\|^2\right). \end{aligned}$$

Functions of the form

$$\mu \mapsto -\mu + c\mu^{\frac{3}{2}}$$

with $c>0$ are non-positive for small $\mu>0$, and hence, we can choose $0<\bar{\mu}<\min\left\{\gamma, \frac{\delta}{16\max_{\xi\in Z}\|\xi\|_Z}\right\}$ (note that this implies $\left|\langle\nu_1,\bar{\xi}\rangle\right|<\frac{\delta}{16}$, and similarly for ν_2) such that

$$v(\bar{x},\bar{\xi})-w(\bar{y},\bar{\zeta}) \leq \frac{\delta}{16}+\beta(\epsilon)+\bar{\mu}\frac{\left(\sum_{k=1}^n\|g_k\|\,\|f_k\|\right)}{\lambda}\left(\|\bar{x}\|^2+\|\bar{y}\|^2\right).$$

Choosing $\bar{\epsilon}>0$ small enough, we thus find that for the corresponding maximizer $(\bar{x},\bar{\xi},\bar{y},\bar{\zeta})$ of Φ, the inequality

$$v(\bar{x},\bar{\xi})-w(\bar{y},\bar{\zeta}) \leq \frac{\delta}{8}+\bar{\mu}\frac{\left(\sum_{k=1}^n\|g_k\|\,\|f_k\|\right)}{\lambda}\left(\|\bar{x}\|^2+\|\bar{y}\|^2\right)$$

holds. But then

$$\frac{\delta}{4} \leq \Phi(\bar{x},\bar{\xi},\bar{y},\bar{\zeta}) \leq v(\bar{x},\bar{\xi})-w(\bar{y},\bar{\zeta})-\frac{\bar{\mu}}{2}\left(\|\bar{x}\|^2+\|\bar{y}\|^2\right)+\langle\nu_1,\bar{\xi}\rangle_Z+\langle\nu_2,\bar{\zeta}\rangle_Z$$

$$
\begin{aligned}
&< \frac{\delta}{4} + \bar{\mu}\left(\frac{\left(\sum_{k=1}^{n} \|g_k\| \, \|f_k\|\right)}{\lambda} - \frac{1}{2}\right)\left(\|\bar{x}\|^2 + \|\bar{y}\|^2\right) \\
&\leq \frac{\delta}{4}
\end{aligned}
$$

by assumption on λ, a contradiction.

□

At this point, we note that (7.19) may also be estimated in a different way; in fact, to conclude the proof, it is enough to have

$$
-\frac{\mu}{\lambda}\sum_{k=1}^{n}\left[\langle \bar{x}, g_k\rangle f_k(\bar{x}) + \langle \bar{y}, g_k\rangle f_k(\bar{y})\right] \leq 0. \tag{7.26}
$$

The following special case might be interesting.

Corollary 7.2 *One can drop the assumption on λ in theorem 7.1 if $g_k = c_k f_k$ for some constants $c_k \geq 0$, for each k.*

Proof: If $f_k = c_k g_k$ for some $c_k \geq 0$, then

$$
(7.26) = -\frac{\mu}{\lambda}\sum_{k=1}^{n} c_k \left(g_k(\bar{x})^2 + g_k(\bar{y})^2\right) \leq 0.
$$

□

Together, theorems 6.7 and 7.1 yield the following existence and uniqueness result.

Theorem 7.3 *If assumptions (A1), (A2) hold together with*

- $g_k \in X \ \forall k \in \{1, \dots, n\}$,
- $f_k \in D((-A)^{\frac{3}{2}}), \ \forall k \in \{1, \dots, n\}$,
- (5.1) *holds,*
- $\lambda \geq \max\left\{2\sum_{k=1}^{n} \|g_k\| \, \|f_k\|, \Gamma\right\}$,
- $\mathbb{A} \subset X$ *nonempty, bounded and closed,*

then the value function corresponding to the dynamic ***(Q)*** *and defined by* (5.4) *is the unique viscosity solution of* (6.3) *in the sense of definition 6.4.*

□

Example 7.4 *Let $\Omega \subset \mathbb{R}^m$ be an open and bounded domain. The standard example for an operator that generates an analytic semigroup is the Laplace operator Δ defined on $H_0^1(\Omega) \cap H^2(\Omega)$, which generates such a semigroup on $X = L^2(\Omega)$. We may thus, e.g., take the equation*

$$\dot{y}(t,x) - \partial_{xx} y(t,x) + \sum_{k=1}^{n} \int_0^1 \mathcal{F}_\rho[f_k(y); w_0](t) d\rho \, g_k(x) = \alpha(t,x),$$
$$y(0,x) = y_0(x), \; x \in (0,1),$$

with $y_0 \in L^2(0,1)$, $\Omega := (0,1)$, $\alpha(t,x) \in [-1,1]$, $\forall t \geq 0$, $\forall x \in [0,1]$, and $w_0(\rho)$ admissible for $\mathcal{F}_\rho$. If

$$\tilde{f}_k(x) := \begin{cases} 1, & \text{if } \left|x - \frac{2k-1}{2n}\right| < \frac{1}{4n} \\ 0, & \text{else,} \end{cases}$$

then $f_k := n\left(\tilde{f}_k * \eta_{\frac{1}{4n}}\right)$ is smooth and has support inside $[\frac{k-1}{n}, \frac{k}{n}]$. Applied to y, i.e.,

$$f_k(y)(t) := \int_0^1 f_k(x) y(t,x) dx = n \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(\tilde{f}_k * \eta_{\frac{1}{4n}}\right)(x) y(t,x) dx,$$

this might be seen as some approximation of the mean value of $y(t,\cdot)$ on $[\frac{k-1}{n}, \frac{k}{n}]$. Choosing $g_k = \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n})}$, the operator F may thus be interpreted as a sort of smeared constitutive law. Rewriting the play operator with the formula $Id = \mathcal{F}_\rho + \mathcal{E}_\rho$, one gets an example for problem ***(Q)****. To build an admissible control problem, one might define, for example, the functional*

$$L(y,w,a) := \left(\int_0^1 y(x)^2 dx\right)^{\frac{1}{2}} + \left(\sum_{k=1}^{n} \int_0^1 w_k(\rho)^2 d\rho\right)^{\frac{1}{2}},$$

which is a norm on $L^2(0,1) \times (L^2(0,1)^n)$, and thus has the Lipschitz property (5.1) *by the triangle inequality. Defining the value function as usual, we may thus apply theorem 7.3, if λ is chosen to be large enough; in this example, by remark 4.4, $\Gamma \leq 1$, and*

$$2 \sum_{k=1}^{n} \|g_k\|_{L^2(0,1)} \|f_k\|_{L^2(0,1)} \leq 2 \sum_{k=1}^{n} \frac{1}{\sqrt{n}} \sqrt{n} = 2n.$$

Hence, we might choose $\lambda \geq 2n$.

□

8 A convergence theorem

Example 7.4 shows that with the type of viscosity solution introduced in section 6, we can only handle a finite number of hysteresis elements. However, one may ask whether the solutions of the dynamic converge. This would be plausible as we have taken mean values on small sets which usually gives good discrete approximations of nondiscrete processes. But, as we will see, the question is not easy to answer. The candidate for the limit problem must be regular enough to define hysteresis pointwise at almost every $x \in \Omega$, i.e., we should at least have regularity $L^2(\Omega; C[0,T])$ for solutions. We can not expect that for general initial values in $L^2(\Omega)$ and general controls in $L^2(\Omega \times [0,T])$! We merely have the significantly weaker continuity $C(0,T; L^2(\Omega))$ guaranteed in this case, i.e., for mild solutions to initial values just in $L^2(\Omega)$. We will thus restrict ourselves to more regular initial values. Moreover, we will have to connect the pointwise viewpoint (limit must define hysteresis for almost every $x \in \Omega$) and the global viewpoint of semigroup theory.

8.1 The function space $L^2(\Omega; H^\theta(0,T))$

Let $T > 0$ and $0 < \theta < 1$. The space $H^\theta(0,T)$ is defined by

$$H^\theta(0,T) := \left\{ u \in L^2(0,T) \ : \ \frac{|u(t) - u(s)|}{|t-s|^{\theta + \frac{1}{2}}} \in L^2((0,T) \times (0,T)) \right\}. \tag{8.1}$$

This is a special case of so called **fractional Sobolev Spaces** (sometimes also called **Slobodeckij Spaces**). Endowed with the natural norm

$$\|u\|_{H^\theta(0,T)} := \left(\int_0^T |u(t)|^2 \, dt + \int_0^T \int_0^T \frac{|u(t) - u(s)|^2}{|t-s|^{1+2\theta}} dtds \right)^{\frac{1}{2}}, \tag{8.2}$$

this is an intermediate Banach space (actually Hilbert space) between $L^2(0,T)$ and $H^1(0,T)$, see e.g. [31, page 5,ff.]. The term

$$[u]_{H^\theta(0,T)} := \left(\int_0^T \int_0^T \frac{|u(t) - u(s)|^2}{|t-s|^{1+2\theta}} dtds \right)^{\frac{1}{2}} \tag{8.3}$$

is called **Gagliardo seminorm** of u. There are various Sobolev type inequalities for Slobodeckij spaces, which can be seen as refined versions of the classical Sobolev imbedding theorem. For example, theorem 8.2 of [31] implies the following:

Theorem 8.1 *If $\frac{1}{2} < \theta < 1$, then there exists $C > 0$ depending on θ and T, such that*

$$\|f\|_{C^{0,\alpha}[0,T]} \leq C \|f\|_{H^\theta(0,T)}, \tag{8.4}$$

for every $f \in L^2(0,T)$*, where* $\alpha := \theta - \frac{1}{2}$*.*

From theorem 8.1, we get that $H^\theta(0,T) \hookrightarrow C[0,T]$, because $\|f\|_{C[0,T]} \le \|f\|_{C^{0,\alpha}[0,T]}$ by definition of the Hölder norms. Unfortunately, we couldn't find a reference where explicit embedding constants are given. However, there is an easy way to give estimates for the constants' dependence on the length of the interval. To this end, let $T_0 > 0$ and C_{T_0} such that (8.4) is fulfilled with T replaced by T_0.

Theorem 8.2 *For every* $T > T_0$*,* $C_T := C_{T_0}\left(\frac{T}{T_0}\right)^{\theta-\frac{1}{2}}$ *fulfills*

$$\|f\|_{C[0,T]} \le C_T \|f\|_{H^\theta(0,T)}, \ \forall f \in H^\theta(0,T).$$

Further, for every $0 < T < T_0$*,* $C_T := C_{T_0}\left(\frac{T_0}{T}\right)^{\frac{1}{2}}$ *fulfills*

$$\|f\|_{C[0,T]} \le C_T \|f\|_{H^\theta(0,T)}, \ \forall f \in H^\theta(0,T).$$

Proof: Let $C_T > 0$ denote the smallest constant for which $\|f\|_{C[0,T]} \le C_T \|f\|_{H^\theta(0,T)}$ holds for all f, i.e.,

$$C_T := \sup_{f \in H^\theta(0,T), f \neq 0} \frac{\|f\|_{C[0,T]}}{\|f\|_{H^\theta(0,T)}}.$$

Let $\tau > 0$. There is a one-to-one correspondence between functions defined on $(0,T)$ and functions defined on $(0, T+\tau)$ through the transformation

$$f(s) \mapsto \bar{f}(s) := f\left(\frac{T+\tau}{T}s\right).$$

Then it holds

$$\|\bar{f}\|_{C[0,T]} = \left\| f\left(\frac{T+\tau}{T}\,\cdot\right)\right\|_{C[0,T]} = \|f\|_{C[0,T+\tau]}.$$

Further, the transformation theorem for integrals yields

$$\|\bar{f}\|^2_{L^2(0,T)} = \int_0^T \left(f\left(\frac{T+\tau}{T}s\right)\right)^2 ds = \frac{T}{T+\tau}\int_0^{T+\tau}(f(r))^2 dr = \frac{T}{T+\tau}\|f\|^2_{L^2(0,T+\tau)},$$

and

$$[\bar{f}]^2_{\theta,T} = \int_0^T \int_0^T \frac{\left|\bar{f}(s) - \bar{f}(t)\right|^2}{|s-t|^{1-2\theta}} dsdt$$

$$= \left(\frac{T}{T+\tau}\right)^2 \int_0^{T+\tau} \int_0^{T+\tau} \frac{|f(u)-f(v)|^2}{\left|\frac{T}{T+\tau}u - \frac{T}{T+\tau}v\right|^{1+2\theta}} dudv$$
$$= \left(\frac{T}{T+\tau}\right)^{1-2\theta} [f]^2_{\theta,T+\tau}.$$

Hence, if $\mu := \frac{T}{T+\tau}$, then

$$\begin{aligned}\|\bar{f}\|_{H^\theta(0,T)} &= \left(\mu \|f\|^2_{L^2(0,T+\tau)} + \mu^{1-2\theta}[f]^2_{\theta,T+\tau}\right)^{\frac{1}{2}} \\ &\leq \left(\mu^{1-2\theta} \|f\|^2_{L^2(0,T+\tau)} + \mu^{1-2\theta}[f]^2_{\theta,T+\tau}\right)^{\frac{1}{2}} \\ &= \mu^{\frac{1}{2}-\theta} \|f\|_{H^\theta(0,T+\tau)},\end{aligned}$$

since $0 < \mu < 1$ and $\frac{1}{2} < \theta < 1$. But then

$$\begin{aligned}C_{T+\tau} &= \sup \frac{\|f\|_{C[0,T+\tau]}}{\|f\|_{H^\theta(0,T+\tau)}} \\ &\leq \sup \frac{\|\bar{f}\|_{C[0,T]}}{\|\bar{f}\|_{H^\theta(0,T)}} \mu^{\frac{1}{2}-\theta} = C_T \left(\frac{T+\tau}{T}\right)^{\theta-\frac{1}{2}},\end{aligned}$$

as claimed. The second inequality follows with $\tau < 0$, since then $\mu > 1$ so that

$$\|\bar{f}\|_{H^\theta(0,T)} \leq \mu^{\frac{1}{2}} \|f\|_{H^\theta(0,T+\tau)}.$$

□

Now, in a canonical way, we define the spaces of functions $L^2(\Omega; H^\theta(0,T))$ and $H^\theta(0,T; L^2(\Omega))$ with corresponding norms

$$\|u\|_{H^\theta(0,T;L^2(\Omega))} := \left(\int_0^T \|u(t)\|^2_{L^2(\Omega)} dt + \int_0^T \int_0^T \frac{\|u(t)-u(s)\|^2_{L^2(\Omega)}}{|t-s|^{1+2\theta}} dtds\right)^{\frac{1}{2}},$$
$$\|u\|_{L^2(\Omega;H^\theta(0,T))} := \left(\int_\Omega \|u(x)\|^2_{H^\theta(0,T)} dx\right)^{\frac{1}{2}}.$$

In fact, the two spaces are equal.

Proposition 8.3 *The Hilbert spaces $L^2(\Omega; H^\theta(0,T))$ and $H^\theta(0,T; L^2(\Omega))$ coincide.*

Proof: By Pettis' theorem (see, e.g.,[32, page 131]), noting that $H^\theta(0,T)$ is separable, weak and strong measurability coincide. Hence, $L^2(\Omega; H^\theta(0,T))$ is equal to the space of all $L^2(\Omega\times(0,T))$ functions, for which $\|\cdot\|_{L^2(\Omega;H^\theta(0,T))}$ is finite. Now, using Fubini's theorem, one easily sees that

$$\|u\|_{H^\theta(0,T;L^2(\Omega))} = \|u\|_{L^2(\Omega;H^\theta(0,T))},$$

for every $u \in L^2(\Omega \times (0,T))$ for which at least one of the two norms is finite. This concludes the proof.

□

Proposition 8.3 can be useful when one needs to switch between the viewpoints of space and time variables. We collect some further useful properties of special function spaces in the following lemma.

Lemma 8.4

(i) $L^2(\Omega; C[0,T]) \hookrightarrow C(0,T; L^2(\Omega))$, *and*

$$\|u\|_{C(0,T;L^2(\Omega))} \le \|u\|_{L^2(\Omega;C[0,T])}, \ \forall\, u \in L^2(\Omega; C[0,T]).$$

(ii) Let C *be an embedding constant for* $H^\theta(0,T) \hookrightarrow C[0,T]$, $\frac{1}{2} < \theta < 1$. *Then for all* $u \in L^2(\Omega; H^\theta(0,T))$,

$$\|u\|_{L^2(\Omega;C[0,T])} \le C\,\|u\|_{L^2(\Omega;H^\theta(0,T))}.$$

Proof: *(i)* The space $C_{pl}(0,T;L^2(\Omega))$ of continuous and piecewise linear functions on $[0,T]$ with values in $L^2(\Omega)$ is dense in both $C(0,T;L^2(\Omega))$ (lemma 2.23) and $L^2(\Omega;C[0,T])$: By definition of Bochner spaces, we can approximate every function in $L^2(\Omega;C[0,T])$ by simple functions, i.e. functions of the form

$$u_s(x,t) = \sum_{k=1}^{n} \chi_{\Omega_k}(x) f_k(t),$$

where each $f_k \in C[0,T]$. But those f_k can again be approximated by functions in $C_{pl}(0,T)$. Doing this for each k and putting together those approximations on Ω_k, one gets a function in $C_{pl}(0,T;L^2(\Omega))$ that is a good approximation of the original $L^2(\Omega;C[0,T])$-function.

As a second step, notice that for every function $u \in C_{pl}(0,T;L^2(\Omega))$,

$$\|u\|_{C(0,T;L^2(\Omega))} = \sup_{t\in[0,T]} \left(\int_\Omega |u(t,x)|^2\,dx\right)^{\frac{1}{2}} = \max_j \|u_j\|$$

$$\leq \left(\int_\Omega \left(\max_j u_j(x) \right)^2 dx \right)^{\frac{1}{2}} = \left(\int_\Omega \left(\sup_{t\in[0,T]} |u(t,x)| \right)^2 dx \right)^{\frac{1}{2}}$$
$$= \|u\|_{L^2(\Omega;C[0,T])} .$$

But then, if a sequence of $C_{pl}(0,T;L^2(\Omega))$ functions is a Cauchy sequence w.r.t. $\|\cdot\|_{L^2(\Omega;C[0,T])}$, it is also Cauchy w.r.t. $\|\cdot\|_{C(0,T;L^2(\Omega))}$, and the claim follows by density of $C_{pl}(0,T;L^2(\Omega))$.

(ii) Again we use approximation by simple functions. For such function u_s, we get

$$\begin{aligned} \|u_s\|^2_{L^2(\Omega;C[0,T])} &= \sum_{k=1}^n |\Omega_k| \, \|f_k\|^2_{C[0,T]} \\ &\leq \sum_{k=1}^n |\Omega_k| \, C^2 \, \|f_k\|^2_{H^\theta(0,T)} \\ &= C^2 \, \|u_s\|^2_{L^2(\Omega;H^\theta(0,T))} , \end{aligned}$$

and the result follows.

□

As an application of proposition 8.3 and lemma 8.4, suppose that an operator $F : C[0,T] \to C[0,T]$ is Lipschitz continuous (with Lipschitz constant $L > 0$). Then (denoting $X := L^2(\Omega)$ to shorten notations),

$$\begin{aligned} \|F(y_1(t)) - F(y_2(t))\|_X &\leq \|F(y_1) - F(y_2)\|_{C(0,T;X)} \\ &\leq \|F(y_1) - F(y_2)\|_{L^2(\Omega;C[0,T])} \\ &\leq L \, \|y_1 - y_2\|_{L^2(\Omega;C[0,T])} \\ &\leq CL \, \|y_1 - y_2\|_{L^2(\Omega;H^\theta(0,T))} \\ &= CL \, \|y_1 - y_2\|_{H^\theta(0,T;X)} , \end{aligned}$$

if $\frac{1}{2} < \theta < 1$. As the first expression appears in the definition of $\|\cdot\|_{H^\theta(0,T;X)}$, this can be useful when one is looking for contraction mappings in some setting.

8.2 Further properties of solutions and the corresponding solution operator

We make the following convention:

Definition 8.5 *For the rest of section 8, it will be assumed that $X = L^2(\Omega)$, where $\Omega \subset \mathbb{R}^n$, as usual, is some open, bounded domain with smooth boundary. The corresponding norm will again be denoted by $\|\cdot\|$.* □

Now we can give some further properties of solutions of the dynamical system.

Proposition 8.6 *Let* $T > 0$ *and* y *denote the solution of equation* (4.6)*, which exists according to theorem 4.1. If* $y(0) = x \in D((-A)^\varepsilon)$ *for some* $0 < \varepsilon < 1$*, then it holds that* $y \in H^\theta(0,T;D((-A)^\gamma))$ *for all parameters* θ, γ*, such that* $0 \leq \gamma < \varepsilon$*,* $0 < \theta < 1 - \gamma$ *and* $\theta < \varepsilon - \gamma + \frac{1}{2}$*.*

Proof: According to theorem 4.1, $y \in C(0,T;X)$, so that there exists $C > 0$ (depending on T) such that

$$\|F(y)(t)\| \leq C, \ \forall\, t \in [0,T]. \tag{8.5}$$

Recalling that y solves the integral equation

$$y(t) = e^{tA}x + \int_0^t e^{(t-\tau)A} \underbrace{(\alpha(\tau) - F(y(\tau)))}_{=:f(\tau)} d\tau,$$

since α is also assumed to be bounded, $y \in C(0,T;D((-A)^\gamma))$ for every $\gamma < \min\{\varepsilon, 1\}$, which implies $y \in L^2(0,T;D((-A)^\gamma))$. Hence, we only need to check whether

$$\int_0^T \int_0^T \frac{\|(-A)^\gamma y(t) - (-A)^\gamma y(s)\|^2}{|t-s|^{1+2\theta}} dsdt < \infty.$$

By definition, we get that

$$\|(-A)^\gamma y(t) - (-A)^\gamma y(s)\| \leq \left\|(e^{tA} - e^{sA})(-A)^\gamma x\right\| \tag{8.6}$$

$$+ \int_0^{\min\{t,s\}} \left\|(-A)^\gamma \left(e^{(\max\{t,s\}-\tau)A} - e^{(\min\{t,s\}-\tau)A}\right) f(\tau)\right\| d\tau \tag{8.7}$$

$$+ \int_{\min\{t,s\}}^{\max\{t,s\}} \left\|(-A)^\gamma e^{(\max\{t,s\}-\tau)A} f(\tau)\right\| d\tau. \tag{8.8}$$

Next, we use the properties of the semigroup we collected in section 4.1 to estimate the right hand side of the latter inequality. We start with the right hand side of (8.6). Thanks to (4.3) and (4.2), if $\varepsilon < \gamma + \beta$,

$$\begin{aligned}\left\|(e^{tA} - e^{sA})(-A)^\gamma x\right\| =& \left\|\left(e^{(\max\{s,t\}-\min\{s,t\})A} - I\right) e^{\min\{s,t\}A}(-A)^\gamma x\right\| \\ \leq& N_\beta \,|t-s|^\beta \left\|(-A)^\beta e^{\min\{t,s\}A}(-A)^\gamma x\right\| \\ \leq& M_{\beta-\varepsilon} N_\alpha \,|t-s|^\beta \min\{t,s\}^{\varepsilon-\gamma-\beta} \left\|(-A)^\varepsilon x\right\|.\end{aligned}$$

So if we choose $\beta \in (\theta, \frac{1}{2} + \varepsilon - \gamma)$, then the integral

$$\begin{aligned}&\int_0^T \int_0^T \frac{\left\| \left(e^{tA} - e^{sA}\right)(-A)^\gamma x \right\|^2}{|t-s|^{1+2\theta}} dsdt \\ &\leq M_{\beta-\varepsilon}^2 N_\beta^2 \left\|(-A)^\varepsilon x\right\|^2 \int_0^T \int_0^T |t-s|^{2\beta-2\theta-1} \min\{t,s\}^{2(\varepsilon-\gamma-\beta)} dsdt < \infty.\end{aligned} \tag{8.9}$$

Next we estimate (8.7). Again, we use estimates (4.3) and (4.2), followed by (8.5), to get

$$\begin{aligned}(8.7) \leq & N_\beta M_\beta C |t-s|^\beta \int_0^{\min\{t,s\}} (\min\{t,s\} - \tau)^{-\beta-\gamma} d\tau \\ = & N_\beta M_\beta C |t-s|^\beta \frac{\min\{t,s\}^{1-\beta-\gamma}}{1-\beta-\gamma},\end{aligned}$$

for all $\beta \in (0, 1-\gamma)$. Thus, if we choose $\beta \in (\theta, 1-\gamma)$, it follows that

$$\begin{aligned}&\int_0^T \int_0^T \frac{(8.7)^2}{|t-s|^{1+2\theta}} dsdt \\ &\leq \left(\frac{N_\beta M_\beta C}{1-\beta-\gamma}\right)^2 \int_0^T \int_0^T |t-s|^{2\beta-2\theta-1} \min\{t,s\}^{2-2\beta-2\gamma} dsdt < \infty.\end{aligned} \tag{8.10}$$

Finally, we need to estimate (8.8). We use (4.2) and (8.5) to derive that

$$(8.8) \leq \frac{CM_\gamma}{1-\gamma} |t-s|^{1-\gamma},$$

and thus, since we assumed $\theta + \gamma < 1$,

$$\int_0^T \int_0^T \frac{(8.8)^2}{|t-s|^{1+2\theta}} dsdt \leq C \int_0^T \int_0^T |t-s|^{1-2(\theta+\gamma)} < \infty. \tag{8.11}$$

This completes the proof.

□

Our goal is now to derive some continuity result for the solution operator corresponding to **problem (P)**, i.e., the operator which maps initial values $x \in L^2(\Omega)$, $\xi \in f(x) - Z$ to the function y defined by

$$y(t) = e^{tA}x + \int_0^t e^{(t-s)A} \left(\alpha(s) - \sum_{k=1}^n P_k[f_k(y); \xi_k](s) g_k\right) ds,$$

where we now assume in addition the following **assumptions (A3)**:

- for all $1 \leq k \leq n$, $g_k = \chi_k$ are characteristic functions corresponding to pairwise disjoint subsets $\Omega_k \subset \Omega$.
- for all $1 \leq k \leq n$, $f_k = \frac{1}{|\Omega_k|}\chi_k$.
- $A = \Delta$, the Laplace operator with $D(\Delta) = H_0^1(\Omega) \cap H^2(\Omega)$.

□

To get a result in a useful form concerning convergence to our original problem, we need some more preparing results.

Lemma 8.7 *Let $T > 0$, $\tilde{A}$ be the generator of an analytic semigroup on some Hilbert space X and z denote the mild solution of*

$$\begin{aligned} \dot{z}(t) &= \tilde{A}z(t), \ t \in (0, T), \\ z(0) &= z_0. \end{aligned}$$

Then, for all $\varepsilon > 0$ and $n \in \mathbb{N}$,

$$z \in C^\infty(\varepsilon, T; X) \cap C(\varepsilon, T; D(\tilde{A}^n)).$$

Proof: Existence and uniqueness of the mild solution are standard and can be found, e.g., in [25], as well as the formula

$$z(t) = e^{t\tilde{A}}z_0$$

for the solution (here, as usual, $e^{t\tilde{A}}$ denotes the corresponding operator semigroup, which has all the properties stated in section 4.1). Now, if $t > 0$, then

$$\frac{d^n z}{dt^n}(t) = \tilde{A}^n e^{t\tilde{A}} z_0,$$

and $\tilde{A}^n e^{t\tilde{A}} \in \mathcal{L}(X)$. Moreover, let w.l.o.g. $t > s > \varepsilon > 0$; then,

$$\left\| \left(\tilde{A}^n e^{t\tilde{A}} - \tilde{A}^n e^{s\tilde{A}} \right) z_0 \right\| \leq \left\| \tilde{A}^n e^{\varepsilon\tilde{A}} \right\| \left\| \left(e^{(t-\varepsilon)\tilde{A}} - e^{(s-\varepsilon)\tilde{A}} \right) z_0 \right\| \to 0,$$

if $|t - s| \to 0$. Thus, the result follows.

□

Applying lemma 8.7 to our special case, this yields that for every k and $t > 0$,

$$y_k(t) := e^{tA}\chi_k \in D(A^n) \subset H^{2n}(\Omega), \tag{8.12}$$

for all $n \in \mathbb{N}$. With [24, page 270, theorem 6], we thus get, in particular (recall that $\partial\Omega$ is assumed to be smooth)

$$y_k(t) \in C^2(\bar{\Omega}),$$

for all $t > 0$. Green's formula [24, page 628, theorem 3] thus yields

$$\frac{d}{dt}\int_\Omega y_k(t,x)dx = \int_\Omega \dot{y}_k(t,x)dx = \int_\Omega \Delta y_k(t,x)dx = \int_{\partial\Omega} \nabla y_k(t,x)\cdot \vec{n}(x)dx,$$

where $\vec{n}(x)$ denotes the outer unit normal vector at $x \in \partial\Omega$. Thus, if the latter is nonpositive, this implies

$$\frac{d}{dt}\int_\Omega y_k(t,x)dx \leq 0, \tag{8.13}$$

and hence, by continuity of $t \mapsto \int_\Omega y_k(t,x)dx$, this would yield

$$\int_\Omega y_k(t,x)dx \leq \int_\Omega y_k(0,x)dx, \tag{8.14}$$

for every $t \geq 0$. The former computation will finish the proof of the following proposition.

Proposition 8.8 *Inequality* (8.14) *is valid for all* $t \geq 0$ *and all* k.

Proof: We still have to show that (8.13) holds for $t > 0$. To do so, we introduce an approximation y_k^n of y_k via the equation

$$\dot{y}_k^n(t) = \Delta y_k^n(t),\ t > 0, \qquad y_k^n(0) = \chi_k^n,$$

such that the sequence of initial values χ_k^n fulfills

(1) $\chi_k^n \to \chi_k$ in $L^2(\Omega)$ as $n \to \infty$,

(2) $\chi_k^n \in C_c^\infty(\Omega)$ for all $n \in \mathbb{N}$,

(3) $\chi_k^n \geq 0$ in Ω.

The corresponding (mild) solution must then particularly be a classical solution, hence, the strong maximum principle [24, page 54, theorem 4] for the heat equation holds. As we have chosen Dirichlet boundary conditions, the maximum respectively minimum values of y_k^n must be attained only for $t = 0$, so that (3) implies $y_k^n(t, x) \geq 0$ for all (t, x). Further, (1) implies $y_k^n \to y_k$ in $C(0, T; X)$ as $n \to \infty$:

$$\left\| e^{tA}\chi_k^n - e^{tA}\chi_k \right\| \leq \left\| e^{tA} \right\| \left\| \chi_k^n - \chi_k \right\| \xrightarrow{n\to\infty} 0.$$

But then we must have $y_k(t, x) \geq 0$ for every $t \geq 0$ and almost every $x \in \Omega$. Since lemma 8.7 tells us that $y_k(t) \in C^1(\bar{\Omega})$ for $t > 0$, the directional derivative at the boundary must be nonpositive, i.e., $\nabla y_k(t) \cdot \vec{n} \leq 0$. This completes the proof.

□

Now we are ready to prove an inequality, which will play a role in the main theorem of this section.

Proposition 8.9 *Let y_1, y_2 be, respectively, the solutions to* ***problem (P)*** *(under assumption (A3)) corresponding to initial values (y_1^0, ξ) and (y_2^0, ζ), and such that $y_1^0, y_2^0 \in D((-A)^\varepsilon)$ for some $\varepsilon > 0$. Then,*

$$\left\| \sum_{k=1}^{n} \left(P_k[f_k(y_1); \xi_k](s) - P_k[f_k(y_2); \zeta_k](s) \right) e^{tA} g_k \right\|^2 \\ \leq \|y_1 - y_2\|^2_{L^2(\Omega; C[0,s])} + \sum_{k=1}^{n} |\Omega_k| \int_{R_k} |\xi_k(\rho) - \zeta_k(\rho)|^2 \, d\mu_k(\rho).$$

Proof: We shorten notations by defining

$$\begin{aligned} \lambda_k(t, x) &:= e^{tA}\chi_k, \\ p_k(s) &:= P_k[f_k(y_1); \xi_k](s) - P_k[f_k(y_2); \zeta_k](s). \end{aligned}$$

Since $0 \leq \sum_{k=1}^{n} \lambda_k(0, \cdot) \leq 1$, by the same arguments as used in the proof of proposition 8.8, it must hold that for every t,

$$\Lambda(t, x) := \sum_{k=1}^{n} \lambda_k(t, x) \in [0, 1], \text{ for almost all } x \in \Omega.$$

Let (t, x) be such that $\Lambda(t, x) \in (0, 1]$; then we can apply the discrete version of Jensen's inequality, to derive

$$I(t, s, x) := \left(\sum_{k=1}^{n} p_k(s) \lambda_k(t, x) \right)^2 = \Lambda(t, x)^2 \left(\sum_{k=1}^{n} p_k(s) \frac{\lambda_k(t, x)}{\Lambda(t, x)} \right)^2$$

$$
\begin{aligned}
&\leq \Lambda(t,x)^2 \sum_{k=1}^{n} p_k(s)^2 \frac{\lambda_k(t,x)}{\Lambda(t,x)} \\
&\leq \sum_{k=1}^{n} p_k(s)^2 \lambda_k(t,x).
\end{aligned}
$$

Else, if $\Lambda(t,x) = 0$, then $\lambda_k(t,x) = 0$ for every k, and the inequality holds, too. Hence, it is valid for (almost) every $x \in \Omega$. But then, using proposition 8.8,

$$
\begin{aligned}
\int_\Omega I(t,s,x)dx &\leq \sum_{k=1}^{n} p_k(s)^2 \int_\Omega \lambda_k(t,x)dx \\
&\leq \sum_{k=1}^{n} p_k(s)^2 \int_\Omega \lambda_k(0,x)dx \\
&= \sum_{k=1}^{n} p_k(s)^2 \left|\Omega_k\right| . \qquad (8.15)
\end{aligned}
$$

Now we use the Lipschitz type property of the hysteresis part to estimate

$$
|p_k(s)| \leq \int_{R_k} \max \left\{ |f_k(y_1 - y_2)|_{C[0,s]}, |\xi_k(\rho) - \zeta_k(\rho)| \right\} d\mu_k(\rho).
$$

Introducing characteristic functions through

$$
\begin{aligned}
\mathbb{1}_1(\rho) =& \begin{cases} 1, & \text{if } |f_k(y_1 - y_2)|_{C[0,s]} > |\xi_k(\rho) - \zeta_k(\rho)| \\ 0, & \text{else,} \end{cases} \\
\mathbb{1}_2 =& 1 - \mathbb{1}_1,
\end{aligned}
$$

this may also be written as

$$
|p_k(s)| \leq \int_{R_k} \mathbb{1}_1(\rho) \, |f_k(y_1 - y_2)|_{C[0,s]} + \mathbb{1}_2(\rho) \, |\xi_k(\rho) - \zeta_k(\rho)| \, d\mu_k(\rho).
$$

Recalling that μ_k is, by definition, a probability measure on R_k, the probabilistic version of Jensen's inequality yields

$$
\begin{aligned}
|p_k(s)|^2 &\leq \left(\int_{R_k} \mathbb{1}_1(\rho) \, |f_k(y_1 - y_2)|_{C[0,s]} + \mathbb{1}_2(\rho) \, |\xi_k(\rho) - \zeta_k(\rho)| \, d\mu_k(\rho) \right)^2 \\
&\leq \int_{R_k} \left(\mathbb{1}_1(\rho) \, |f_k(y_1 - y_2)|_{C[0,s]} + \mathbb{1}_2(\rho) \, |\xi_k(\rho) - \zeta_k(\rho)| \right)^2 d\mu_k(\rho) \\
&= \int_{R_k} \mathbb{1}_1(\rho) \, |f_k(y_1 - y_2)|^2_{C[0,s]} + \mathbb{1}_2(\rho) \, |\xi_k(\rho) - \zeta_k(\rho)|^2 \, d\mu_k(\rho)
\end{aligned}
$$

$$\leq |f_k(y_1 - y_2)|^2_{C[0,s]} + \int_{R_k} |\xi_k(\rho) - \zeta_k(\rho)|^2 \, d\mu_k(\rho).$$

Hence,

$$\int_\Omega I(t,s,x)dx \leq \sum_{k=1}^n |\Omega_k| \, |f_k(y_1 - y_2)|^2_{C[0,s]} + \sum_{k=1}^n |\Omega_k| \int_{R_k} |\xi_k(\rho) - \zeta_k(\rho)|^2 \, d\mu_k(\rho).$$

By proposition 8.6 and lemma 8.4, $y_1, y_2 \in L^2(\Omega; C[0,s])$. Recall now that by assumption (A3), $f_k = \frac{1}{|\Omega_k|}\chi_k$, so that the integral version of Jensen's inequality implies

$$\begin{aligned}\sum_{k=1}^n |\Omega_k| \, |f_k(y_1 - y_2)|^2_{C[0,s]} &= \sum_{k=1}^n |\Omega_k| \left(\frac{1}{|\Omega_k|} \max_{\tau \in [0,s]} \left| \int_{\Omega_k} y_1(\tau, x) - y_2(\tau, x) dx \right| \right)^2 \\ &\leq \sum_{k=1}^n |\Omega_k| \left(\int_{\Omega_k} \max_{\tau \in [0,s]} |y_1(\tau, x) - y_2(\tau, x)| \, \frac{dx}{|\Omega_k|} \right)^2 \\ &\leq \sum_{k=1}^n \int_{\Omega_k} \left(\max_{\tau \in [0,s]} |y_1(\tau, x) - y_2(\tau, x)| \right)^2 dx \\ &\leq \|y_1 - y_2\|^2_{L^2(\Omega; C[0,s])} .\end{aligned}$$

We note that the last inequality becomes an equality if $\cup_{k=1}^n \Omega_k = \Omega$.

□

Remark 8.10 *Assumptions (A3) may be weakened to* ***assumption (A3')****:*

- *for all* $1 \leq k \leq n$, $g_k = \chi_k$ *are characteristic functions corresponding to pairwise disjoint subsets* $\Omega_k \subset \Omega$.
- *for all* $1 \leq k \leq n$, $f_k = \frac{1}{|\Omega_k|}\chi_k$.
- *A generates an analytic semigroup of contractions on* $X = L^2(\Omega)$.

In fact, the more classical argumentation of proposition 8.8 can be avoided, if one recognizes that

$$\left\| \sum_{k=1}^n \left(P_k[f_k(y_1); \xi_k](s) - P_k[f_k(y_2); \zeta_k](s) \right) e^{tA} g_k \right\|$$

$$\leq \left\| \sum_{k=1}^{n} \left(P_k[f_k(y_1); \xi_k](s) - P_k[f_k(y_2); \zeta_k](s) \right) g_k \right\|,$$

because e^{tA} is a semigroup of contractions. Then, using the special form of the functions $g_k = \chi_k$, one can directly get inequality (8.15), and can thus avoid the application of classical results such as Green's formula or the maximum principle.

□

Next, we will proof a stability result for regular initial values, which will enable us later to take the limit $n \to \infty$ on bounded intervals $[0, T]$.

Theorem 8.11 *Let $\varepsilon, T > 0$, y_1, y_2 denote solutions to equation (4.6) under assumptions (A1) and (A3') corresponding to initial values (x, ξ), (z, ζ), and such that $x \in D((-A)^\varepsilon)$ as well as $z \in D((-A)^\varepsilon)$. Further, assume that there exist constants $c_1, c_2 > 0$ and $\delta(\xi, \zeta)$, such that*

$$\|F(y_1)(s) - F(y_2)(s)\|^2 \leq c_1 \|y_1 - y_2\|^2_{L^2(\Omega; C[0,s])} + c_2 \delta(\xi, \zeta)^2.$$

Then, for every $\frac{1}{2} < \theta < \varepsilon + \frac{1}{2}$ there is a constant $C = C(T, \theta)$ such that

$$\|y_1 - y_2\|^2_{H^\theta(0,T;X)} \leq \; C \|(-A)^\varepsilon (x - z)\|^2 + C\delta(\xi, \zeta)^2.$$

Proof: By proposition 8.6, $y_1, y_2 \in H^\theta(0, T; X)$ for $\theta < \varepsilon + \frac{1}{2}$. Since in the following, we want to apply lemma 8.4, we have to restrict ourselves further to $\theta > \frac{1}{2}$. So, let $\frac{1}{2} < \theta < \varepsilon + \frac{1}{2}$ and $t \in (0, T]$. To indicate the dependence on ξ, ζ, we will write here

$$F(y_1, \xi)(s) := \sum_{k=1}^{n} P_k(f_k(y_1); \xi_k)(s) g_k,$$

$$F(y_2, \zeta)(s) := \sum_{k=1}^{n} P_k(f_k(y_2); \zeta_k)(s) g_k.$$

Further, we introduce the abbreviations $\delta F(s) := F(y_1, \xi)(s) - F(y_2, \zeta)(s)$ and $\delta y := y_1 - y_2$. Now, note that for $s \in [0, t]$,

$$y_1(s) - y_2(s) = e^{sA}(x - z) + \int_0^s e^{(s-\tau)A} \left(F(y_2, \zeta)(\tau) - F(y_1, \xi)(\tau) \right) d\tau.$$

Thus, using standard arguments (esp. Young's inequality), we can derive an estimate of the form

$$\|y_1 - y_2\|^2_{H^\theta(0,t;X)}$$

$$\leq 2\int_0^t \left\|e^{sA}(x-z)\right\|^2 ds \tag{8.16}$$

$$+2\int_0^t \left(\int_0^s \left\|e^{(s-\tau)A}\left(F(y_2,\zeta)(\tau)-F(y_1,\xi)(\tau)\right)\right\| d\tau\right)^2 ds \tag{8.17}$$

$$+4\int_0^t\int_0^t \frac{\left\|\left(e^{rA}-e^{sA}\right)(x-z)\right\|^2}{|r-s|^{1+2\theta}} drds \tag{8.18}$$

$$+4\int_0^t\int_0^t \frac{\left(\int_0^{\min\{r,s\}} \left\|\left(e^{(\max\{r,s\}-\tau)A}-e^{(\min\{r,s\}-\tau)A}\right)\delta F(\tau)\right\| d\tau\right)^2}{|r-s|^{1+2\theta}} drds \tag{8.19}$$

$$+4\int_0^t\int_0^t \frac{\left(\int_{\min\{r,s\}}^{\max\{r,s\}} \left\|e^{(\max\{r,s\}-\tau)A}\left(F(y_1,\xi)(\tau)-F(y_2,\zeta)(\tau)\right)\right\| d\tau\right)^2}{|r-s|^{1+2\theta}} drds. \tag{8.20}$$

We are going to estimate each of the terms on the right hand side of the inequality, with very similar argumentations as in the proof of proposition 8.6. The first term is easy to handle, we get, e.g., since the semigroup is in particular a contraction semigroup,

$$2\int_0^t \left\|e^{sA}(x-z)\right\|^2 ds \leq 2t\left\|x-z\right\|^2.$$

By assumption, we can write down an estimation of the second term, as

$$ct^3\left(\|\delta y\|^2_{L^2(\Omega);C[0,t]}+\delta(\xi,\zeta)^2\right),$$

where $c>0$ only depends on c_1,c_2. In order to get a good estimation of the third term, we will use properties (4.2) and (4.3) of our semigroup, in the way we did in the proof of proposition 8.6. Since it is exactly the same calculation, we will drop it here; one finds that for some constant $c=c(t,\varepsilon)$ (which goes to zero as $t\downarrow 0$):

$$(8.18)\leq c\left\|(-A)^{\varepsilon}(x-z)\right\|^2.$$

The fourth term can be estimated as follows: we use properties (4.2), (4.3) and the semigroup property to derive

$$\begin{aligned}&\left\|\left(e^{(\max\{r,s\}-\tau)A}-e^{(\min\{r,s\}-\tau)A}\right)\delta F(\tau)\right\|\\ &\qquad\leq \tilde{c}\,|r-s|^{\alpha}\left(\min\{r,s\}-\tau\right)^{-\alpha}e^{-\omega(\min\{r,s\}-\tau)}\left\|\delta F(\tau)\right\|,\end{aligned}$$

with $\alpha\in(\theta,1)$. Since

$$\int_0^a (a-\tau)^{-\alpha}e^{-\omega(a-\tau)}d\tau=\int_0^a x^{-\alpha}e^{-\omega x}dx<\int_0^\infty x^{-\alpha}e^{-\omega x}dx<\infty,$$

for every $a > 0$, there exists some function $\bar{c}$ which grows slower than linear for large s (but is still integrable over every interval of the form $(0, T)$), such that

$$\begin{aligned}(8.19) \leq& \int_0^t \bar{c}(s) \left(\|y_1 - y_2\|^2_{L^2(\Omega;C[0,s])} + \delta(\xi,\zeta)^2 \right) ds \\ \leq& \int_0^t \bar{c}(s) ds \left(\|y_1 - y_2\|^2_{L^2(\Omega;C[0,t])} + \delta(\xi,\zeta)^2 \right).\end{aligned}$$

We note that $\bar{c}$ is actually some function in s, t and explicitly given via

$$\bar{c}(s,t) := \tilde{c} \int_0^t |r - s|^{2\alpha - 2\theta - 1}\, dr,$$

so that

$$\bar{C}(t) := \int_0^t \bar{c}(s) ds$$

is superlinear in t. Next, we estimate (8.20). By assumption,

$$(8.20) \leq 8 \int_0^t \int_0^t \frac{\|y_1 - y_2\|^2_{L^2(\Omega;C[0,\max\{r,s\}])} + \delta(\xi,\zeta)^2}{|r-s|^{2\theta-1}} drds \tag{8.21}$$

Now, for some continuous nonnegative function f, it holds that

$$\begin{aligned}\int_0^t \int_0^t \frac{f(\max\{r,s\})}{|r-s|^\gamma}\, drds =& \int_0^t \int_0^s \frac{f(s)}{(s-r)^\gamma}\, drds + \int_0^t \int_r^t \frac{f(r)}{(r-s)^\gamma}\, drds \\ \leq& \int_0^t \int_0^t \frac{f(s)}{|s-r|^\gamma}\, drds + \int_0^t \int_0^t \frac{f(r)}{|r-s|^\gamma}\, drds \\ =& 2 \int_0^t \int_0^t \frac{f(s)}{|s-r|^\gamma}\, drds.\end{aligned}$$

Applying this to (8.21) with $f = \|y_1 - y_2\|_{L^2(\Omega;C[0,\cdot])}$ yields

$$(8.20) \leq 16\bar{C}(t) \|y_1 - y_2\|^2_{L^2(\Omega;C[0,t])} + 8\bar{C}(t)\delta(\xi,\zeta)^2.$$

So, inserting all five estimates, we get that, for some superlinear continuous function $\tilde{C}(t)$,

$$\begin{aligned}\|\delta y\|^2_{H^\theta(0,t;X)} \leq& 2t \|x - z\|^2 + c(t,\varepsilon) \|(-A)^\varepsilon (x - z)\|^2 \\ &+ \tilde{C}(t) \left(\delta(\xi,\zeta)^2 + \|\delta y\|^2_{L^2(\Omega;C[0,s])} \right).\end{aligned} \tag{8.22}$$

According to lemma 8.4, proposition 8.3 and theorem 8.2, there exists some constant C_T such that

$$\|y\|_{L^2(\Omega;C[0,s])} \leq \|y\|_{H^\theta(0,s;X)} \frac{C_T}{s^{\frac{1}{2}}}$$

for every $s \in (0,T)$. Hence, from this and (8.22) we can derive

$$\begin{aligned}\|\delta y\|^2_{H^\theta(0,t;X)} \leq &2t\,\|x-z\|^2 + \tilde{\tilde{C}}(t)\left(\|(-A)^\varepsilon(x-z)\|^2 + \delta(\xi,\zeta)^2\right)\\ &+ \frac{\tilde{C}(t)}{t}\,\|\delta y\|^2_{H^\theta(0,t;X)}\,.\end{aligned}$$

Since $\tilde{C}(t)$ is superlinear, $\tilde{C}(t)/t$ tends to zero as $t \downarrow 0$. Thus, there is $\tau \in (0,T]$, such that $\tilde{C}(t)/t \leq \frac{1}{2}$ for all $t \in (0,\tau]$, which implies that there is some constant $C > 0$, such that

$$\|\delta y\|^2_{H^\theta(0,t;X)} \leq \|\delta y\|^2_{H^\theta(0,\tau;X)} \leq C\left(\|x-z\|^2 + \|(-A)^\varepsilon(x-z)\|^2 + \delta(\xi,\zeta)^2\right),$$

for every $t \in (0,\tau]$. Our goal is now to get another estimate, where we can use the first part of theorem 8.2. To this end, let $t > \tau$. If we do not use $\|\delta y\|_{L^2(\Omega;C[0,s])} \leq \|\delta y\|_{L^2(\Omega;C[0,t])}$ in our estimations, we arrive at

$$\begin{aligned}\|\delta y\|^2_{H^\theta(0,t;X)} \leq \hat{C}(t)&\left(\|x-z\|^2 + \|(-A)^\varepsilon(x-z)\|^2 + \delta(\xi,\zeta)^2\right)\\ &+ \int_0^t \kappa(s)\,\|\delta y\|^2_{L^2(\Omega;C[0,s])}\,ds,\end{aligned} \tag{8.23}$$

with some continuous function $\hat{C}$ and an integrable map κ. Further,

$$\begin{aligned}\int_0^t \kappa(s)\,\|\delta y\|^2_{L^2(\Omega;C[0,s])}\,ds = &\int_0^\tau \kappa(s)\,\|\delta y\|^2_{L^2(\Omega;C[0,s])}\,ds\\ &+ \int_\tau^t \kappa(s)\,\|\delta y\|^2_{L^2(\Omega;C[0,s])}\,ds,\end{aligned}$$

and, with lemma 8.4, theorem 8.1,

$$\begin{aligned}\int_0^\tau \kappa(s)\,\|\delta y\|^2_{L^2(\Omega;C[0,s])}\,ds &\leq \int_0^t \kappa(s)ds\,\|\delta y\|^2_{L^2(\Omega;C[0,\tau])}\\ &\leq CC_\tau^2 \int_0^\tau \kappa(s)ds\left(\|x-z\|^2 + \|(-A)^\varepsilon(x-z)\|^2 + \delta(\xi,\zeta)^2\right),\end{aligned}$$

as well as (cf., theorem 8.2)

$$\int_\tau^t \kappa(s)\,\|\delta y\|^2_{L^2(\Omega;C[0,s])}\,ds \leq \int_\tau^t \kappa(s)C_\tau^2 s^{2\theta-1}\,\|\delta y\|^2_{H^\theta(0,s;X)}\,ds.$$

Altogether, we thus infer from (8.23),

$$\|\delta y\|^2_{H^\theta(0,t;X)} \le K(t)\left(\|x-z\|^2 + \|(-A)^\varepsilon(x-z)\|^2 + \delta(\xi,\zeta)^2\right) + \int_\tau^t \kappa(s) C_\tau^2 s^{2\theta-1} \|\delta y\|^2_{H^\theta(0,s;X)}\, ds.$$

Applying Gronwall's lemma to the function $f : [\tau, T] \to \mathbb{R}_+$, $t \mapsto \|\delta y\|^2_{H^\theta(0,t;X)}$ then yields the result.

□

8.3 Convergence of solutions of problems (P), (P')

In this section we show some applications of theorem 8.11. We will restrict the discussion to the case where each R_k is an interval, though.

To define "starting curves" for our hysteresis operators pointwise for almost every $x \in \Omega$, we introduce the function space

$$H := L^2(\Omega; L^2(R_\cdot, \mu_\cdot)) \tag{8.24}$$

consisting of all measurable functions f for which $f(x)(\cdot) \in L^2(R_x, \mu_x)$ for almost every $x \in \Omega$. More precisely:

Definition 8.12 *Let $\Omega \subset \mathbb{R}^n$ be an open, bounded domain, and $R^1, R^2 \in L^\infty(\Omega)$ such that $0 < R^1 < R^2$ (almost) everywhere. For every $x \in \Omega$, let μ_x be a probability measure on the interval $R_x := (R^1(x), R^2(x))$ which is absolutely continuous w.r.t. the Lebesgue measure on R_x, i.e., there exists some function $h_x : R_x \to \mathbb{R}_+$ such that $\int_{R^1(x)}^{R^2(x)} h_x(\rho)d\rho = 1$ and $\mu_x = h_x d\mathcal{L}$ for every $x \in \Omega$; in addition, we assume that $h_\cdot(\cdot) > 0$. Then we define*

$$\Omega^\star := \{(x,r) \mid x \in \Omega \text{ and } r \in R_x\},$$
$$H := \{\xi : \Omega^\star \to \mathbb{R} \text{ measurable}\},$$

and H is equipped with the norm

$$\|\xi\|_H := \left(\int_\Omega \int_{R_x} \xi(x,\rho)^2 d\mu_x(\rho) dx\right)^{\frac{1}{2}} = \left(\int_\Omega \int_{R_x} \xi(x,\rho)^2 h_x(\rho) d\rho dx\right)^{\frac{1}{2}}.$$

□

Since $R^1, R^2 \in L^\infty(\Omega)$, we may w.l.o.g. assume that $R^1(x) < R^2(x) \le \|R^2\|_{L^\infty(\Omega)}$ holds for all $x \in \Omega$. Then H can be identified with some closed subspace of some Hilbert space, and is thus itself a Hilbert space.

Proposition 8.13 *Let $\hat{h}$ denote the measurable function*

$$\hat{h} : \Omega \times [0, \left\|R^2\right\|_{L^\infty(\Omega)}] \to \mathbb{R}_+,\ \hat{h}(x,\rho) := \begin{cases} h_x(\rho), & \text{if } (x,\rho) \in \Omega^\star, \\ 1, & \text{else,} \end{cases}$$

and define a measure $\hat{\mu}$ on $\hat{\Omega} := \Omega \times [0, \|R^2\|_{L^\infty(\Omega)}]$ through $d\hat{\mu} = \hat{h}d\mathcal{L}$. If $h_\cdot \in L^\infty(\Omega^\star)$, then $\mathcal{H} := L^2(\hat{\Omega}, \hat{h}d\mathcal{L})$ is a Hilbert space, and the closed subspace

$$\mathcal{H}_s := \left\{\xi \in \mathcal{H} \mid \xi = 0 \text{ on } \hat{\Omega} \setminus \Omega^\star\right\}$$

is isometrically isomorphic to H.

Proof: It is easy to see that $\mathcal{H}$ is a Hilbert space with scalar product

$$\langle \xi, \zeta \rangle_{\mathcal{H}} := \int_{\hat{\Omega}} \xi \zeta \hat{h} d\mathcal{L},$$

as the positivity of the measure implies that the space is a Banach space. It is also clear that the limit of every sequence in $\mathcal{H}_s$ that is convergent in $\mathcal{H}$ will belong to the equivalence class of functions which are zero almost everywhere in $\hat{\Omega} \setminus \Omega^\star$; thus, $\mathcal{H}_s$ is a closed subspace of $\mathcal{H}$. Further, by continuation of functions in H by zero, one gets some element of $\mathcal{H}_s$, and restriction of some function in $\mathcal{H}_s$ to the set $\Omega^\star$ yields an element of H. As this mapping is bijectiv and isometric, the proof is complete.

□

Now, due to proposition 8.13, we can define what we mean when we talk about convergence of some sequence ξ_n in $\mathcal{H}$ converging to some element of H:

Definition 8.14 *We say that some sequence $(\xi_n)_n \subset \mathcal{H}$ converges to some element $\xi \in H$, if ξ_n converges in $\mathcal{H}$ to the continuation*

$$\hat{\xi} := \begin{cases} \xi \text{ on } \Omega^\star, \\ 0 \text{ on } \hat{\Omega} \setminus \Omega^\star, \end{cases}$$

of ξ. Further, we consider sequences of densities $(h_n)_n$, similarly defining spaces H_n, resp. $\mathcal{H}_n = L^2(\hat{\Omega}, \hat{h}_n d\mathcal{L})$. Let us ***assume that $\hat{h} \in L^\infty(\hat{\Omega})$***. *We say that such a sequence of densities converges to h, if the sequence of continuations $(\hat{h}_n)$ is bounded in $L^\infty(\hat{\Omega})$ and converges to $\hat{h}$ in $L^1(\hat{\Omega})$.*

□

Remark 8.15 *(i) With those definitions, it holds that $\hat{h}_n \to \hat{h}$ in $L^p(\hat{\Omega})$ for every $p \in [1, \infty)$ (by dominated convergence).*

(ii) $\hat{\xi}_n \to \xi$ *in* $H :\Leftrightarrow \hat{\xi}_n \to \hat{\xi}$ *in* $\mathcal{H} \Leftrightarrow \hat{\xi}_n \to \hat{\xi}$ *in* $L^2(\hat{\Omega})$.
Proof: Assume first that $\hat{\xi}_n \to \hat{\xi}$ *in* $L^2(\hat{\Omega})$. *Then*

$$\left\|\hat{\xi}_n - \hat{\xi}\right\|_{\mathcal{H}}^2 = \int_{\Omega} \left(\hat{\xi}_n - \hat{\xi}\right)^2 \hat{h} d\mathcal{L} \leq \left\|\hat{h}\right\|_{L^\infty(\hat{\Omega})} \left\|\hat{\xi}_n - \hat{\xi}\right\|_{L^2(\hat{\Omega})}^2 \to 0,$$

as $n \to \infty$. *Now, assume that* $\hat{\xi}_n \to \hat{\xi}$ *in* $\mathcal{H}$. *Then, since* $\hat{h} > 0$, *it is necessary that there exists a subsequence* $\hat{\xi}_{n_k}$ *which converges pointwise a.e. in* $\hat{\Omega}$ *to* $\hat{\xi}$. *By dominated convergence and uniform boundedness, we thus get* $\hat{\xi}_{n_k} \to \hat{\xi}$ *in* $L^p(\hat{\Omega})$ *for all* $p \in [1, \infty)$. *As we can always extract such a subsequence, the result follows.*

□

Since we can view step functions $\sum_{k=1}^n \chi_k \xi_k$ as restrictions of some step function in $\mathcal{H}_n$, we thus can give meaning to convergence of terms such as

$$\sum_{k=1}^n |\Omega_k| \int_{R_k} |\xi_k(\rho) - \zeta_k(\rho)|^2 \, d\mu_k(\rho),$$

which appeared in theorem 8.9. Having this in mind, we can proof a convergence theorem for the solutions of problem **(P)**. As we used initial values of the play operator there, we have to reformulate the result if we want to work with the spaces we have just introduced. To do so, we add the definitions

$$\bar{\mathcal{F}}_\rho[y;\xi] := \mathcal{F}_\rho[y; y(0) - \xi], \qquad \bar{P}_k[y;\xi_k] := \int_{R_k} \bar{\mathcal{F}}_\rho[y; y(0) - \xi_k] d\mu_k,$$

i.e., we give the initial values implicitly via the formula $\mathcal{F}_\rho = I - \mathcal{E}_\rho$. To highlight the different notations, we will refer to this as **problem (P')**.

Corollary 8.16 (Reformulation of proposition 8.9, remark 8.10) *Let* y_1, y_2 *denote solutions to problem* ***(P')*** *corresponding to initial values* $(y_1^0, \xi), (y_2^0, \zeta) \in D((-A)^\varepsilon) \times H_n$, *where* $\varepsilon > 0$ *and* H_n *is determined through the density (step-) function* $\sum_{k=1}^n \chi_k h_k(\rho)$. *Then*

$$\begin{aligned}
&\left\| \sum_{k=1}^n \left(\bar{P}_k[f_k(y_1); \xi_k](s) - \bar{P}_k[f_k(y_2); \zeta_k](s) \right) g_k \right\|^2 \\
&\leq \|y_1 - y_2\|_{L^2(\Omega; C[0,s])}^2 + 2\left\|y_1^0 - y_2^0\right\|^2 + 2\sum_{k=1}^n |\Omega_k| \int_{R_k} |\xi_k(\rho) - \zeta_k(\rho)|^2 \, d\mu_k(\rho) \\
&\leq 3\|y_1 - y_2\|_{L^2(\Omega; C[0,s])}^2 + 2\sum_{k=1}^n |\Omega_k| \int_{R_k} |\xi_k(\rho) - \zeta_k(\rho)|^2 \, d\mu_k(\rho).
\end{aligned}$$

Proof: Since μ_k are probability measures, it holds

$$\sum_{k=1}^{n} |\Omega_k| \int_{R_k} |f_k(y_1(0)) - \xi_k(\rho) - f_k(y_2(0)) + \zeta_k(\rho)|^2 \, d\mu_k(\rho)$$
$$\leq 2\sum_{k=1}^{n} |\Omega_k| \langle f_k, y_1(0) - y_2(0)\rangle^2 + 2\sum_{k=1}^{n} |\Omega_k| \int_{R_k} |\xi_k(\rho) - \zeta_k(\rho)|^2 \, d\mu_k(\rho).$$

From Jensen's inequality, we derive that (denoting $\delta y := y_1(0) - y_2(0)$)

$$\langle f_k, \delta y\rangle^2 = \left(\int_{\Omega_k} \delta y \frac{d\mathcal{L}}{|\Omega_k|}\right)^2 \leq \int_{\Omega_k} \delta y^2 \frac{d\mathcal{L}}{|\Omega_k|},$$

so that the first inequality follows. The second one is obvious.

□

Lemma 8.17 *Assume $\theta > \frac{1}{2}$, $\gamma > 0$ and that there exists $0 < s < \frac{1}{2}$ such that $D((-A)^\gamma) \hookrightarrow H^s(\Omega)$. Then*

$$H^\theta(0, T; D((-A)^\gamma)) \hookrightarrow L^q(\Omega; C[0, T]),$$

for all $q \in [1, \frac{2n}{n-2s}]$, where $\Omega \subset \mathbb{R}^n$.

Proof: By definition, for $y \in H^\theta(0, T; D((-A)^\gamma))$, it must hold that $y(t) \in D((-A)^\gamma)$ for a.e. $t \in [0, T]$. Thus, if C is the imbedding constant for $D((-A)^\gamma) \hookrightarrow H^s(\Omega)$, by definition of the corresponding norms,

$$\|y\|_{H^\theta(0,T;H^s(\Omega))} \leq C \|y\|_{H^\theta(0,T;D((-A)^\gamma))},$$

hence, $H^\theta(0, T; D((-A)^\gamma)) \hookrightarrow H^\theta(0, T; H^s(\Omega))$. Next, from Fubini's theorem, we infer that $H^\theta(0, T; H^s(\Omega)) = H^s(\Omega; H^\theta(0, T))$. Since $H^\theta(0, T) \hookrightarrow C[0, T]$, we also get $H^s(\Omega; H^\theta(0, T)) \hookrightarrow H^s(\Omega; C[0, T])$. Then, for any function $y \in H^s(\Omega; C[0, T])$, it must hold that $\|y(\cdot)\|_{C[0,T]}$ is an element of $H^s(\Omega)$. Hence, applying [31, theorem 6.7] to this function yields $H^s(\Omega; C[0, T]) \hookrightarrow L^q(\Omega; C[0, T])$ for the stated values of q. In particular, the above chain of continuous imbeddings proves the claim.

□

Theorem 8.18 *Let $(x_n, \hat{\xi}_n)_n$ be a sequence of initial values for problem (P') under assumption (A3'), which converges in $D((-A)^\varepsilon) \times L^2(\hat{\Omega})$ to $(x, \hat{\xi})$, and define functions $y_n, \bar{y}_n$ to be the solutions of*

$$y_n(t) := e^{tA}x_n + \int_0^t e^{(t-s)A} \left(\alpha(s) - \sum_{k=1}^{n} \int_{R_k} \bar{\mathcal{F}}_\rho[f_k(y_n); \xi_n^k(\rho)](s)\hat{h}_k(s) g_k d\rho\right) ds,$$

$$\bar{y}_n(t) := e^{tA}x + \int_0^t e^{(t-s)A}\left(\alpha(s) - \int_{R.} \bar{\mathcal{F}}_\rho[\sum_{k=1}^n \chi_k f_k(\bar{y}_n); \xi(\cdot,\rho)](s)\hat{h}(\cdot,\rho)d\rho\right) ds.$$

Further, assume that for every $\gamma > 0$ there exists $\nu > 0$ such that $D((-A)^\gamma) \hookrightarrow H^\nu(\Omega)$, and that

$$\sum_{k=1}^n \chi_{R_k}(\rho)\chi_k \hat{h}_k(\rho) \to \chi_{R.}\hat{h}(\cdot,\rho), \quad \sum_{k=1}^n \chi_k \hat{\xi}_n^k(\rho) \to \hat{\xi}(\cdot,\rho),$$

as $n \to \infty$, pointwise almost everywhere. Then, with

$$\begin{aligned} d := &\sum_{k=1}^n \int_{R_k} \bar{\mathcal{F}}_\rho[f_k(y_n); \xi_n^k(\rho)](s)\hat{h}_k(s)g_k d\rho \\ &- \int_{R.} \bar{\mathcal{F}}_\rho[\sum_{k=1}^n \chi_k f_k(\bar{y}_n); \xi(\cdot,\rho)](s)\hat{h}(\cdot,\rho)d\rho, \end{aligned}$$

it holds

$$\int_\Omega d^2 d\mathcal{L} \leq c\,\|y_n - \bar{y}_n\|^2_{L^2(\Omega;C[0,s])} + \omega\left(\frac{1}{n}\right),$$

where $c > 0$ and ω is a continuous nonnegative function with $\omega(0) = 0$.

Proof: First note that $\bar{y}_n$ is well defined, because we may view the right hand side as one Lipschitz continuous function operating on the whole domain Ω; in this sense, theorem 4.1 is applicable. Let now $I := \left[0, \|R^2\|_{L^\infty(\Omega)}\right]$. We may rewrite the first term in d as

$$d_1 := \int_I \sum_{k=1}^n \chi_{R_k}(\rho)\chi_k \bar{\mathcal{F}}_\rho\left[\sum_{k=1}^n \chi_k f_k(y_n); \sum_{k=1}^n \chi_k \hat{\xi}_n^k(\rho)\right](s)\hat{h}_k(\rho)d\rho,$$

and the second term as

$$d_2 := \int_I \chi_{R.}(\rho)\bar{\mathcal{F}}_\rho\left[\sum_{k=1}^n \chi_k f_k(\bar{y}_n); \hat{\xi}(\cdot,\rho)\right](s)\hat{h}(\cdot,\rho)d\rho.$$

Then, introducing the notation

$$f_1 := \bar{\mathcal{F}}_\rho\left[\sum_{k=1}^n \chi_k f_k(y_n); \sum_{k=1}^n \chi_k \hat{\xi}_n^k(\rho)\right](s),$$

$$f_2 := \bar{\mathcal{F}}_\rho \left[\sum_{k=1}^n \chi_k f_k(\bar{y}_n); \hat{\xi}(\cdot, \rho) \right] (s),$$

we have

$$d = \int_I \sum_{k=1}^n \chi_{R_k}(\rho) \chi_k f_1 \hat{h}_k(\rho) - \chi_{R.}(\rho) f_2 \hat{h}(\cdot, \rho) d\rho.$$

Thus, a.e. in Ω, it holds

$$\begin{aligned} d &= \\ &\int_I \sum_{k=1}^n \chi_{R_k}(\rho)\chi_k f_1 \hat{h}_k(\rho) - \chi_{R.}(\rho) f_1 \hat{h}(\cdot,\rho) + \chi_{R.}(\rho) f_1 \hat{h}(\cdot,\rho) - \chi_{R.}(\rho) f_2 \hat{h}(\cdot,\rho) d\rho \\ &= \int_I f_1 \left[\sum_{k=1}^n \chi_{R_k}(\rho) \chi_k \hat{h}_k(\rho) - \chi_{R.}(\rho) \hat{h}(\cdot,\rho) \right] d\rho \\ &\quad + \int_I \chi_{R.}(\rho) \hat{h}(\cdot,\rho) \left[f_1 - f_2 \right] d\rho. \end{aligned}$$

We need to estimate the $L^2(\Omega)$ norm of this expression. With the abbreviations

$$\mu_1 := \sum_{k=1}^n \chi_{R_k}(\rho) \chi_k \hat{h}_k(\rho) - \chi_{R.}(\rho) \hat{h}(\cdot, \rho), \qquad \mu_2 := \chi_{R.}(\rho) \hat{h}(\cdot, \rho),$$

we have to estimate

$$I_1 := \int_{\hat{\Omega}} f_1^2 \mu_1^2 d\mathcal{L}, \qquad I_2 := \int_{\hat{\Omega}} \mu_2^2 \left[f_1 - f_2 \right]^2 d\mathcal{L}.$$

We start with I_1. From Hölder's inequality, we infer that

$$I_1 \leq \left(\int_{\hat{\Omega}} |f_1|^{2p} \, d\mathcal{L} \right)^{\frac{1}{p}} \left(\int_{\hat{\Omega}} |\mu_1|^{2q} \, d\mathcal{L} \right)^{\frac{1}{q}}, \qquad \frac{1}{p} + \frac{1}{q} = 1, \quad p, q > 1.$$

Note that pointwise a.e.,

$$\begin{aligned} |f_1| &\leq 2 \sup_t \left| \sum_{k=1}^n \chi_k f_k(y_n(t)) \right| + \left| \sum_{k=1}^n \chi_k \hat{\xi}_n^k(\rho) \right| \\ \Rightarrow |f_1|^{2p} &\leq 2^{2p} \left(2^{2p} \sup_t \left| \sum_{k=1}^n \chi_k f_k(y_n(t)) \right|^{2p} + \left| \sum_{k=1}^n \chi_k \hat{\xi}_n^k(\rho) \right|^{2p} \right), \end{aligned}$$

and that

$$\begin{aligned}\int_{\hat{\Omega}} \sup_t \left|\sum_{k=1}^n \chi_k f_k(y_n(t))\right|^{2p} d\mathcal{L} &\leq \int_{\hat{\Omega}} \sup_t \sum_{k=1}^n \chi_k \left|f_k(y_n(t))\right|^{2p} d\mathcal{L} \\ &\leq \sum_{k=1}^n \int_{\hat{\Omega}} \chi_k d\mathcal{L} \sup_t \left|f_k(y_n(t))\right|^{2p} \\ &= \sum_{k=1}^n |I|\,|\Omega_k| \sup_t \left|\int_\Omega \frac{\chi_k}{|\Omega_k|} y_n(t) d\mathcal{L}\right|^{2p} \\ &\leq \sum_{k=1}^n |I|\,|\Omega_k| \sup_t \int_{\Omega_k} |y_n(t)|^{2p} \frac{d\mathcal{L}}{|\Omega_k|} \\ &\leq |I|\, \|y_n\|^{2p}_{L^{2p}(\Omega;C[0,T])},\end{aligned}$$

which is bounded by lemma 8.17 and proposition 8.6 if $p > 1$ is small enough. Since by assumption, $\mu_1 \to 0$ as $n \to \infty$ in $L^1(\hat{\Omega})$, the uniform boundedness implies then that $\mu_1 \to 0$ in $L^q(\hat{\Omega})$ for every $q \in [1, \infty)$; thus choosing q large enough, p becomes small enough. This shows that $I_1 \to 0$ as $n \to \infty$. Next, we estimate I_2. Thanks to the uniform boundedness assumption, $\mu_2 \in L^\infty(\hat{\Omega})$. Hence, for some $c > 0$,

$$I_2 \leq c \int_{\hat{\Omega}} [f_1 - f_2]^2 \, d\mathcal{L}.$$

Using the Lipschitz continuity of $\bar{\mathcal{F}}_\rho$, we see that except for some multiplicative constant, I_2 is estimated by

$$\int_{\hat{\Omega}} \sup_t \left|\sum_{k=1}^n f_k(y_n(t))\chi_k - f_k(\bar{y}_n(t))\chi_k\right|^2 + \left|\sum_{k=1}^n \chi_k \hat{\xi}^k_n(\rho) - \hat{\xi}(\cdot, \rho)\right|^2 d\mathcal{L}.$$

Further,

$$\begin{aligned} I_3 &:= \int_{\hat{\Omega}} \sup_t \left|\sum_{k=1}^n f_k(y_n(t))\chi_k - f_k(\bar{y}_n(t))\chi_k\right|^2 d\mathcal{L} \\ &\leq \int_{\hat{\Omega}} \sum_{k=1}^n \sup_t \chi_k \left(\int_\Omega \left|\frac{\chi_k}{|\Omega_k|}(y_n(t) - \bar{y}_n(t))\right| d\mathcal{L}\right)^2 d\mathcal{L} \\ &\leq |I| \sum_{k=1}^n \sup_t \int_{\Omega_k} |y_n(t) - \bar{y}_n(t)|^2 \, d\mathcal{L} \\ &\leq |I|\, \|y_n - \bar{y}_n\|^2_{L^2(\Omega;C[0,s])},\end{aligned}$$

and

$$I_4 := \int_{\hat{\Omega}} \left| \sum_{k=1}^{n} \chi_k \hat{\xi}_n^k(\rho) - \hat{\xi}(\cdot,\rho) \right|^2 d\mathcal{L} \to 0,$$

as $n \to \infty$, because of our assumptions and Lebesgue's theorem of dominated convergence. Hence, summing up I_1 and I_4 to some $\omega(\frac{1}{n})$, this proves the claim.

□

Remark 8.19 *Under assumption (A3), as remarked in [7, page 4], it holds*

$$D((-A)^\theta) = \begin{cases} H^{2\theta}(\Omega), & \text{if } 0 \le \theta < \frac{1}{4}, \\ \left\{x \in H^{2\theta}(\Omega) : \ x = 0 \text{ on } \partial\Omega\right\}, & \text{if } \frac{1}{4} < \theta < 1, \end{cases}$$

so that an embedding of the type assumed in theorem 8.18 always exists in this case.

□

For the proof of the desired convergence result, we still need a general approximation result for special types of step functions.

Lemma 8.20 *Let $y \in L^2(\Omega; C[0,T])$ and let $\{\Omega_k^n : \ k \in \{1, \dots, n\}\}_n$ be a sequence of decompositions of the domain Ω such that*

- *each Ω_k^n is a measurable subset of Ω,*
- *$\Omega_j^n \cap \Omega_k^n = \{\}$ for all $j \neq k$,*
- *$\max_{1\le k\le n} |\Omega_k^n| \to 0$ as $n \to \infty$.*

Then

$$y^n := \sum_{k=1}^{n} \chi_k \frac{1}{|\Omega_k^n|} \int_{\Omega_k^n} y d\mathcal{L} \to y,$$

in $L^2(\Omega; C[0,T])$, as $n \to \infty$.

Proof: There is a sequence $(z_m)_{m\in\mathbb{N}}$ of functions such that $z_m \in C(\bar{\Omega} \times [0,T])$ for all m and $z_m \to y$ in $L^2(\Omega; C[0,T])$, as $m \to \infty$. This can be seen, e.g., by density of $C_{pl}(0,T;X)$ functions in $L^2(\Omega; C[0,T])$ (cf. proof of lemma 8.4). Moreover, if $z \in$

$C(\bar{\Omega} \times [0,T])$ and $\Omega^n_{k(n)}$ is a sequence of subsets of Ω such that $\cap_{n\in\mathbb{N}}\Omega^n_{k(n)} = x_0 \in \Omega$, then

$$\begin{aligned}\left\| \frac{1}{\left|\Omega^n_{k(n)}\right|} \int_{\Omega^n_{k(n)}} z(\cdot,x)dx - z(\cdot,x_0) \right\|_{C[0,T]} \\ &\leq \sup_t \frac{1}{\left|\Omega^n_{k(n)}\right|} \int_{\Omega^n_{k(n)}} |z(t,x) - z(t,x_0)|\, dx \\ &\leq \sup_{t,\ \|x-x_0\| \leq \mathrm{diam}(\Omega^n_{k(n)})} |z(t,x) - z(t,x_0)| \xrightarrow{n\to\infty} 0,\end{aligned} \tag{8.25}$$

due to the uniform continuity of z on $\bar{\Omega} \times [0,T]$. Hence, denoting $z_{m,n}(t,x) := \sum_{k=1}^n \chi_k \frac{1}{|\Omega^n_k|} \int_{\Omega^n_k} z_m(t,x)dx$,

$$\begin{aligned}\|y^n - y\|_{L^2(\Omega;C[0,T])} &\leq \|y^n - z_{m,n}\|_{L^2(\Omega;C[0,T])} + \|z_{m,n} - z_m\|_{L^2(\Omega;C[0,T])} \\ &\quad + \|z_m - y\|_{L^2(\Omega;C[0,T])}.\end{aligned}$$

We estimate the first two summands. As usual, from Jensen's inequality, we infer that

$$\begin{aligned}\|y^n - z_{m,n}\|^2_{L^2(\Omega;C[0,T])} &\leq \sum_{k=1}^n \sup_t \int_{\Omega_k} |y(t,x) - z_m(t,x)|^2\, dx \\ &\leq \|y - z_m\|^2_{L^2(\Omega;C[0,T])}.\end{aligned}$$

For the second term, we exploit (8.25) and apply Lebesgue's theorem of dominated convergence, which then yields that

$$\|z_{m,n} - z_m\|_{L^2(\Omega;C[0,T])} \xrightarrow{n\to\infty} 0,$$

for fixed $m \in \mathbb{N}$. Thus,

$$\limsup_{n\to\infty} \|y^n - y\|_{L^2(\Omega;C[0,T])} \leq 2\, \|z_m - y\|_{L^2(\Omega;C[0,T])},$$

and we can conclude by letting $m \to \infty$.

□

Theorem 8.21 *Let the assumptions of theorem 8.18 and lemma 8.20 hold. Then $y_n, \bar{y}_n$ (as defined in theorem 8.18) converge to y in $H^\theta(0,T;X)$ for every $\theta < \min\left\{\frac{1}{2} + \varepsilon, 1\right\}$, where y is the solution of*

$$y(t) = e^{tA}x + \int_0^t e^{(t-s)A} \left(\alpha(s) - \int_{R.} \bar{\mathcal{F}}_\rho\, [y;\xi]\, (s) h(\cdot,\rho) d\rho \right) ds. \tag{8.26}$$

Moreover, if y^1, y^2 *are two solutions of* (8.26) *w.r.t. to initial values* $(x^1,\xi^1),(x^2,\xi^2) \in D((-A)^\varepsilon)\times H$*, then there is* $C>0$ *such that*

$$\begin{aligned}\left\|y^1-y^2\right\|^2_{H^\theta(0,T;X)} \leq & C\left\|(-A)^\varepsilon(x^1-x^2)\right\|^2 \\ &+C\int_\Omega\int_{R.}\left|\xi^1(\cdot,\rho)-\xi^2(\cdot,\rho)\right|^2 h(\cdot,\rho)d\rho d\mathcal{L}.\end{aligned} \tag{8.27}$$

Proof: From theorems 8.18 and 8.11, we infer that it suffices to show that $\bar{y}_n$ converges to y. To this end, we want to apply again theorem 8.11. Hence, we have to estimate

$$\delta_n := \left\|\int_{R.}\left(\bar{\mathcal{F}}_\rho\left[\sum_{k=1}^n \chi_k f_k(\bar{y}_n);\xi(\cdot,\rho)\right](s)-\bar{\mathcal{F}}_\rho\left[y;\xi(\cdot,\rho)\right](s)\right)h(\cdot,\rho)d\rho\right\|.$$

We use the usual arguments. From Jensen's inequality, the Lipschitz continuity of $\bar{\mathcal{F}}_\rho$ and $h\in L^\infty$, we get that there exists $c>0$ constant, such that

$$\begin{aligned}\delta_n^2 &\leq c\int_\Omega \sup_t\left|\sum_{k=1}^n \chi_k f_k(\bar{y}_n)-y\right|^2 d\mathcal{L} \\ &\leq 2c\int_\Omega\sup_t\left|\sum_{k=1}^n\chi_k f_k(\bar{y}_n)-\sum_{k=1}^n\chi_k f_k(y)\right|^2 d\mathcal{L}+2c\int_\Omega\sup_t\left|\sum_{k=1}^n\chi_k f_k(y)-y\right|^2 d\mathcal{L}.\end{aligned}$$

The first summand is not larger than $2c\left\|\bar{y}_n-y\right\|^2_{L^2(\Omega;C[0,s])}$, which can be seen by the usual arguments. The second one converges to zero as $n\to\infty$, by lemma 8.20. Hence, convergence to y follows by application of theorem 8.11. Now, if we have two solutions of (8.26) corresponding to different initial values, we can apply our approximation procedure to find sequences y_n^1, y_n^2, which converge to y^1, y^2 in $H^\theta(0,T;X)$, and for which proposition 8.9, theorem 8.11 are applicable. Then, letting $n\to\infty$, (8.27) follows.

□

Remark 8.22 *In the proof of the last theorem we have neglected the question whether the solution of* (8.26) *actually exists. Under assumption (A3), for initial values* $y_0\in D((-A)^{\frac{1}{2}})$*, this is a consequence of theorem 3.3. However, as* (8.27) *is valid for* ε *smaller than* $\frac{1}{2}$*, we may infer from this existence and uniqueness of solutions for initial values in* $D((-A)^\varepsilon)$*. The case when we just assume assumption (A3') is different though; this can be done with the techniques used to prove theorem 8.11. Since the calculation is quite long, we omit it.*

□

At the end of this chapter, we show that from (8.26) we can go back to the pde we started with, i.e., the equation where the Prandtl Ishlinskii operator is replaced by the usual (pointwise applied) Play Operator. We will do this, roughly spoken, by replacing the density h by some dirac measure. To simplify the analysis, we will assume that

$$\forall x \in \Omega, \rho \in [R_n^1(x), R_n^2(x)], \qquad h_n(x,\rho) = \frac{1}{R_n^2(x) - R_n^1(x)}, \tag{8.28}$$

and assume that R_n^1, R_n^2 converge to some function R in $L^\infty(\Omega)$. If we fix $(y_0, \hat{\xi}) \in (D((-A)^\varepsilon) \times \mathcal{H})$, this gives us, by restriction, for every $n \in \mathbb{N}$, a pair of initial values for problem (P') w.r.t. H_n. If $\hat{\xi} \in C(cl(\hat{\Omega}))$, then

$$\int_{R_n^1(x)}^{R_n^2(x)} \hat{\xi}(\rho, x) \frac{d\rho}{R_n^2(x) - R_n^1(x)} \xrightarrow{n\to\infty} \hat{\xi}(R(x), x), \ x \in \Omega,$$

and even

$$\int_\Omega \int_{R_n^1(x)}^{R_n^2(x)} \left(\hat{\xi}(\rho, x) - \hat{\xi}(R(x), x)\right)^2 \frac{d\rho}{R_n^2(x) - R_n^1(x)} dx \xrightarrow{n\to\infty} 0. \tag{8.29}$$

To avoid technicalities, we will assume in the following theorem that (8.29) holds. Of course, this is true in more general situations than the one when $\hat{\xi}$ is a continuous function.

Theorem 8.23 *Let assumption (A3') hold, and assume that* $y_0 \in D((-A)^\varepsilon)$, $\varepsilon > 0$, *and let* ξ_n *be the restriction to* $\Omega_n^\star$ *(corresponding to* $R_n^1, R_n^2 \in L^\infty(\Omega)$*) of some function* $\hat{\xi} \in \mathcal{H}$. *If both* R_n^1, R_n^2 *converge to* $R \in L^\infty(\Omega)$ *w.r.t.* $\|\cdot\|$, $R(x) \in [R_n^1(x), R_n^2(x)]$ *for a.e.* $x \in \Omega$, *every* $n \in \mathbb{N}$, *and* (8.28), (8.29) *hold, then the corresponding sequence of solutions* y_n *to* (8.26) *converges in* $H^\theta(0,T;L^2(\Omega))$ *(*$\theta < \min\left\{\frac{1}{2} + \varepsilon, 1\right\}$*) to the solution of*

$$\begin{aligned} y(t) =& e^{tA} y_0 + \int_0^t e^{(t-s)A}(\alpha(s) - w(s))ds, \\ w(t) =& \bar{\mathcal{F}}_{R(x)}[y(\cdot, x); \hat{\xi}(R(x), x)](t). \end{aligned} \tag{8.30}$$

Moreover, if $y_0^1, y_0^2 \in D((-A)^\varepsilon)$ *and* $\zeta^1, \zeta^2 \in L^\infty(\Omega)$ *can be represented as limits in the sense of* (8.29), *then the inequality*

$$\left\|y^1 - y^2\right\|^2_{H^\theta(0,T;X)} \le C \left\|(-A)^\varepsilon (y_0^1 - y_0^2)\right\|^2 + C \left\|\zeta^1 - \zeta^2\right\|^2 \tag{8.31}$$

holds, where y^i *denotes the solution of* (8.30) *corresponding to* (y_0^i, ζ^i), *and* C *is the constant appearing in* (8.27).

Proof: We want to apply theorem 8.11. To this end, note that

$$w(x)(t) = \frac{1}{|R_n^2(x) - R_n^1(x)|} \int_{R_n^1(x)}^{R_n^2(x)} \bar{\mathcal{F}}_{R(x)}[y(\cdot, x); \hat{\xi}(R(x), x)](t) d\rho.$$

Hence, denoting for short $I_n(x) := [R_n^1(x), R_n^2(x)]$, from Jensen's inequality and the Lipschitz property of the Play operator, we deduce that

$$\begin{aligned}
&\int_\Omega \left(\int_{I_n(x)} \bar{\mathcal{F}}_\rho \left[y_n(x, \cdot); \hat{\xi}(x, \rho) \right](t) - \bar{\mathcal{F}}_{R(x)} \left[y(x, \cdot); \hat{\xi}(R(x), x) \right](t) \frac{d\rho}{|I_n(x)|} \right)^2 dx \\
&\qquad \leq 3 \int_\Omega \sup_s |y_n(x, s) - y(x, s)|^2 \, dx \\
&\qquad\quad + \int_\Omega \int_{I_n(x)} |\rho - R(x)|^2 \frac{d\rho}{|I_n(x)|} dx \qquad (8.32) \\
&\qquad\quad + 2 \int_\Omega \int_{I_n(x)} \left(\hat{\xi}(x, \rho) - \hat{\xi}(R(x), x) \right)^2 \frac{d\rho}{|I_n(x)|} dx. \qquad (8.33)
\end{aligned}$$

Thus, we only have to show that (8.32), (8.33) converge to zero as $n \to \infty$. For (8.33), this follows immediately from (8.29). Concerning (8.32), a simple calculation shows that

$$(8.32) = \frac{1}{3} \left(\left\| R_n^2 \right\|^2 + \left\langle R_n^2, R_n^1 \right\rangle + \left\| R_n^1 \right\|^2 \right) - \left\langle R_n^2, R \right\rangle - \left\langle R_n^1, R \right\rangle + \|R\|^2,$$

which converges to zero by assumption. Now, application of theorem 8.11 proves the first part of the theorem. Moreover, we can now take the limit in (8.26) to prove (8.31).

□

8.4 Convergence of optimal value functions for problem (P')

We give a general approximation result for optimal value functions of infinite horizon problems, which may then be applied to problem (P') in different situations.

Assumption 8.24 *Let $X = L^2(\Omega)$ and assumption (A3') hold. Let $\mathbb{A} \subset X$ the set where the controls may take values in, and $L : X \times X \times \mathbb{A} \to \mathbb{R}$ a functional that meets the properties*

(L1) $\exists C > 0 : |L(x_1, x_2, a)| \leq C \quad \forall \, (x_1, x_2, a) \in X \times X \times \mathbb{A}$,

(L2) $\exists C_L > 0 \; \forall a \in \mathbb{A} \; \forall x_1, y_1, x_2, y_2 \in X$:

$$|L(x_1, y_1, a) - L(x_2, y_2, a)| \leq C_L \left(\|x_1 - x_2\| + \|y_1 - y_2\| \right).$$

Further, we introduce, for $n \in \mathbb{N} \cup \{\infty\}$, dynamical systems via

$$\begin{aligned} \dot{y}^n(t) =& Ay^n(t) + \alpha(t) - w^n(t) \ \in X, \ a.e.\ t \in [0,T], && (8.34)\\ w^n(t) =& \mathcal{W}^n[y^n; w_0^n](t), \quad t \in [0,T], && (8.35)\\ y^n(0) =& y_0^n, \qquad w^n(0) = w_0^n, && (8.36) \end{aligned}$$

where A is the generator of an analytic semigroup in X, $y_0^n \in D((-A)^\varepsilon) \subset X$, $w_0^n \in L^\infty(\hat{\Omega})$, and $\mathcal{W}^n$ (a sequence of hysteresis-) operators. Moreover, we assume that if $w_0^n \to w_0^\infty \in L^\infty(\hat{\Omega})$ pointwise a.e., and $y^n \to y^\infty$ in $L^2(\Omega; C[0,T])$, then $w^n \to w^\infty$ in $C(0,T;X)$ as $n \to \infty$.

□

As usual, we define the corresponding value functions as

$$V_n(y_0^n, w_0^n) := \inf_{\alpha \in \mathcal{A}} \int_0^\infty e^{-\lambda t} L(y^n(t), w^n(t), \alpha(t))dt,$$

where we assume that $\lambda > 0$. One gets the following result.

Theorem 8.25 *Let assumption 8.24 hold, and assume that for some sequences of initial values y_0^n, w_0^n, it holds that $y_0^n \to y_0^\infty$ in X, $w_0^n \to w_0^\infty$ pointwise a.e., and for the corresponding sequence of trajectories, for every $T > 0$, $y^n \to y^\infty$ in $L^2(\Omega; C[0,T])$ as $n \to \infty$. Then $V_n(y_0^n, w_0^n) \to V_\infty(y_0^\infty, w_0^\infty)$.*

Proof: First note that all V_n are well defined, as we assume that there exist solutions to systems (8.34)-(8.36) and L is bounded. Let C be the constant from (L1). For every $\varepsilon > 0$, we can find $T > 0$ such that

$$2C \int_T^\infty e^{-\lambda t} dt < \varepsilon.$$

Hence, for such T,

$$\begin{aligned} &|V_n(y_0^n, w_0^n) - V_\infty(y_0^\infty, w_0^\infty)| \\ &\leq \sup_\alpha \left\{ \int_0^\infty e^{-\lambda t} \left| L(y^n(t), w^n(t), \alpha(t)) - L(y^\infty(t), w^\infty(t), \alpha(t)) \right| dt \right\} \\ &< \int_0^T C_L \left(\|y^n(t) - y^\infty(t)\| + \|w^n(t) - w^\infty(t)\| \right) dt + \varepsilon. \end{aligned}$$

Now, by assumption 8.24, if $w_0^n \to w_0^\infty$ pointwise a.e. and $y^n \to y^\infty$ in $L^2(\Omega; C[0,T])$, then $w^n \to w^\infty$ in $C(0,T;X)$. Thus,

$$\limsup_{n \to \infty} |V_n(y_0^n, w_0^n) - V_\infty(y_0^\infty, w_0^\infty)| \leq \varepsilon,$$

for every $\varepsilon > 0$, which implies the result. □

Corollary 8.26 *Let assumptions* $(L1), (L2)$ *from 8.24 and the ones from theorem 8.21 hold. Then the corresponding value functions converge pointwise on* $D((-A)^\varepsilon) \times \mathcal{H}$ *for every* $\varepsilon > 0$.

Proof: For sequences of initial values converging in $D((-A)^\varepsilon) \times \mathcal{H}$, we may apply theorem 8.21, which implies convergence of trajectories in $H^\theta(0, T; X)$ for every $T > 0$. Since $H^\theta(0, T; X) \hookrightarrow L^2(\Omega; C[0, T])$, the Lipschitz continuity of the hysteresis operator implies convergence of $w^n \to w^\infty$ in $L^2(\Omega; C[0, T])$, and thus, in particular, convergence in $C(0, T; X)$. This shows that theorem 8.25 is applicable, which concludes the proof.

□

Up to now, we always have assumed in section 8 that $f_k = \frac{\chi_k}{|\Omega_k|}$, which does not meet the smoothness assumption from theorem 7.3. We still want to close this gap at the end of this chapter. We begin with a short remark concerning theorem 8.11.

Remark 8.27 *In theorem 8.11, we actually haven't used the exact form of the operator* F *on the right hand side of equation* (4.6)*, but merely that there is some estimate of a special form. Thus, the theorem is valid for more general operators.*

□

Having this in mind, we can give the following result.

Corollary 8.28 *Let the assumptions and notations of theorem 8.21 hold, and denote by* $y_n^\star$ *the sequence of functions defined through the integral equations*

$$y_n^\star = e^{tA} y_n^\star(0) + \int_0^t e^{(t-s)A} \left(\alpha(s) - \sum_{k=1}^n \chi_k \int_{R_k} \mathcal{F}_\rho \left[f_k^\star(y_n^\star); \xi_k \right](s) d\mu_k(\rho) \right) ds,$$

where $y_n^\star(0) = y(0)$*, and* $f_k^\star = \frac{\nu_k}{|\Omega_k|}$*, where the* ν_k *are nonnegative (smooth) functions such that* $\|\nu_k\|_{L^\infty(\Omega)} \leq 1$ *and having support in* Ω_k*, for all* n*. If*

$$\boldsymbol{\nu}_n := \sum_{k=1}^n \nu_k \to 1, \text{ pointwise a.e.},$$

then $y_n^\star \to y$ *in* $H^\theta(0, T; X)$ *for all* θ *such that* $y_n \to y$ *w.r.t. that norm.*

Proof: Since theorem 8.21 holds, in view of theorem 8.11 and remark 8.27, we only have to find a suitable estimate for

$$d := \sum_{k=1}^n \chi_k \int_{R_k} \mathcal{F}_\rho \left[f_k(y_n); \xi_k \right](s) d\mu_k(\rho) - \sum_{k=1}^n \chi_k \int_{R_k} \mathcal{F}_\rho \left[f_k^\star(y_n^\star); \xi_k \right](s) d\mu_k(\rho).$$

The usual arguments lead to

$$\|d\|^2 \leq \sum_{k=1}^{n} |\Omega_k| \, \|f_k(y_n) - f_k^\star(y_n^\star)\|_{C[0,s]}^2 \leq \sum_{k=1}^{n} \sup_t \int_{\Omega_k} (y_n(t) - \nu_k y_n^\star(t))^2 \, d\mathcal{L}.$$

Next, note that pointwise

$$y_n(t) - \nu_k y_n^\star(t) = (1-\nu_k) y_n(t) + \nu_k (y_n(t) - y_n^\star(t)),$$

with $\nu_k \in [0,1]$. Thus,

$$\begin{aligned}\|d\|^2 &\leq \sum_{k=1}^{n} \sup_t \left[\int_{\Omega_k} (1-\nu_k)\,(y_n(t))^2 + \nu_k\,(y_n(t) - y_n^\star(t))^2 \, d\mathcal{L} \right] \\ &\leq \left\| \sqrt{(1-\boldsymbol{\nu}_n)} y_n \right\|_{L^2(\Omega; C[0,s])}^2 + \|y_n - y_n^\star\|_{L^2(\Omega; C[0,s])}^2 .\end{aligned}$$

As we know from theorem 8.21 that y_n converges in $L^2(\Omega; C[0,T])$, the first summand goes to zero as $n \to \infty$, so that we have found the desired inequality.

□

We can now give a convergence result for the corresponding value functions similar to corollary 8.26.

Corollary 8.29 *Let the assumptions of corollary 8.28 hold together with (L1), (L2) from 8.24. Then the value functions corresponding to the dynamics of $y_n^\star$ converge pointwise on $D((-A)^\varepsilon) \times \mathcal{H}$ to the value function corresponding to the dynamic of y, for every $\varepsilon > 0$.*

Proof: Similar to the one of corollary 8.26.

□

Remark 8.30 *Employing theorem 8.23, a similar convergence result can be proved; thus, we can also go back to the problem discussed in section 3.*

□

9 Time discrete dynamic programming and approximative optimal feedback controls

In this section we will present elements of time discrete dynamic programming and how this might be used to approximate the value function of our original problem corresponding to the heat equation with hysteresis (3.1), and to find approximations of optimal controls via the feedback method for the discretized problem. Whenever possible, we will follow [4, chapter VI], where the method is presented for a finite dimensional model problem. We start with stating the time discretized equation. To not run into difficulties concernig existence of solutions (in particular, for the original equation), we will restrict ourselves from the beginning to weak solutions, and thus to initial values in $H_0^1(\Omega)$. So, let $h > 0$ be the (constant) stepsize; at time level $t = mh$, given $(y_m, w_m) \in D := H_0^1(\Omega) \times L^2(\Omega)$ and $a_m \in L^2(\Omega)$, we are then looking for a solution $y_{m+1} \in H_0^1(\Omega)$ of

$$\begin{aligned}\frac{1}{h}\int_\Omega (y_{m+1} - y_m)\varphi d\mathcal{L} + \int_\Omega w_m \varphi d\mathcal{L} + \int_\Omega \nabla y_{m+1} \cdot \nabla \varphi d\mathcal{L} \\ = \int_\Omega a_m \varphi d\mathcal{L},\ \forall \varphi \in H_0^1(\Omega).\end{aligned} \tag{9.1}$$

As ususal, $w_m := \mathcal{W}(y_0, \dots, y_m, w_0)$ with some (hysteresis) operator $\mathcal{W}$ satisfying the semigroup property $w_{m+1} = \mathcal{W}(y_{m+1}, w_m)$. We remark that in the existence proof in section 3.1, we had the same type of equation but with w_m replaced by w_{m+1}, so that (9.1) is in a way more explicit.

Theorem 9.1 *Let $\Omega \subset \mathbb{R}^n$ with smooth boundary (at least C^2), $m \in \mathbb{N}$, $h > 0$. Given $(y_m, w_m) \in D$ and $a_m \in L^2(\Omega)$, there exists a unique weak solution $y_{m+1} \in H_0^1(\Omega) \cap H^2(\Omega)$ of (9.1).*

Proof: This is a direct consequence of the Lax Milgram theorem and the standard improved regularity result, see e.g. [24, chapter 6].

□

Thus, y_{m+1} solves in fact

$$y_{m+1} = h\Delta y_{m+1} + y_m + ha_m - hw_m, \text{ a.e. in } \Omega, \qquad y_{m+1} = 0 \text{ on } \partial\Omega.$$

We may therefore introduce the solution operator

$$L : D \times L^2(\Omega) \to H_0^1(\Omega) \cap H^2(\Omega) \subset D, \qquad L(y_m, w_m, a_m) = y_{m+1}.$$

Then our time discrete dynamical system takes the form

$$y_{m+1} = L(y_m, w_m, a_m),$$

$$w_{m+1} = \mathcal{W}(L(y_m, w_m, a_m), w_m),$$

or shorter, introducing the operator

$$H(y_m, w_m, a_m) := \mathcal{W}(L(y_m, w_m, a_m), w_m),$$

we may also write

$$\begin{aligned} y_{m+1} &= L(y_m, w_m, a_m), \\ w_{m+1} &= H(y_m, w_m, a_m). \end{aligned} \tag{9.2}$$

Let now $(y_0, w_0) = (x, v) \in D$. To indicate the dependence on initial values and the control, we will from now on write $y_m(x, v, a), w_m(x, v, a)$ for the solutions at iteration level $m \in \mathbb{N}$. Further, for some subset $\mathbb{A}$ of $L^2(\Omega)$, the set $\mathcal{A}$ of admissible controls shall consist of all sequences $\alpha = (a_n)_{n \in \mathbb{N}}$ such that $a_n \in \mathbb{A}$ for all n. Introducing the functional

$$J(x, v, \alpha) := \sum_{n=0}^{\infty} l(y_n(x, v, \alpha), w_n(x, v, \alpha), a_n)\beta^n, \tag{9.3}$$

where $l : (L^2(\Omega))^3 \to \mathbb{R}$, $\beta \in (0, 1)$, we can formulate a (time discrete) control problem via

$$V(x, v) := \inf_{\alpha \in \mathcal{A}} J(x, v, \alpha).$$

As usual, we call $V : D \to \mathbb{R}$ the value function of the control problem.

Lemma 9.2 (time discrete DPP) *Let l be bounded. Then V is well defined and satisfies the dynamic programming equation*

$$V(x, v) = \inf_{a \in \mathbb{A}} \{ l(x, v, a) + \beta V(L(x, v, a), H(x, v, a)) \}.$$

Proof: It is clear from the definition that V is well defined if l is bounded. Now, for $\alpha \in \mathcal{A}$, i.e. $\alpha = (a_0, a_1, a_2, \ldots)$, let $\bar{\alpha} := (a_1, a_2, a_3, \ldots)$ be the corresponding shifted control. Then we may write

$$y_{m+1}(x, v, \alpha) = y_m(y_1(x, v, a_0), w_1(x, v, a_0), \bar{\alpha}) = y_m(L(x, v, a_0), H(x, v, a_0), \bar{\alpha}),$$

and similarly

$$w_{m+1}(x, v, \alpha) = w_m(L(x, v, a_0), H(x, v, a_0), \bar{\alpha}).$$

Thus, from the definition of J, we get that

$$J(x, v, \alpha) = l(x, v, a_0)$$

$$+\beta\sum_{m=0}^{\infty} l(y_m(L(x,v,a_0),H(x,v,a_0),\bar{\alpha}),w_m(L(x,v,a_0),H(x,v,a_0),\bar{\alpha}),a_{m+1})$$
$$= l(x,v,a_0)+\beta J(L(x,v,a_0),H(x,v,a_0),\bar{\alpha}).$$

Hence,

$$J(x,v,\alpha)\geq l(x,v,a_0)+\beta V(L(x,v,a_0),H(x,v,a_0)),$$

which implies that

$$V(x,v)\geq \inf_{a\in\mathbb{A}}\left\{l(x,v,a)+\beta V(L(x,v,a),H(x,v,a))\right\}.$$

Next, let $a\in\mathbb{A}$ be arbitrary, and set

$$z_1 := L(x,v,a),\quad z_2 := H(x,v,a),\quad \hat{\alpha} := (a,\alpha) = (a,a_0,a_1,a_2,\ldots).$$

For every $\varepsilon>0$ there exists $\alpha_\varepsilon = (a_0^\varepsilon,a_1^\varepsilon,a_2^\varepsilon,\ldots)\in\mathcal{A}$ such that

$$V(z_1,z_2)\geq J(z_1,z_2,\alpha_\varepsilon)-\varepsilon.$$

Arguing as before, we get the equality

$$J(x,v,\hat{\alpha}_\varepsilon) = l(x,v,a)+\beta J(L(x,v,a),H(x,v,a),\alpha_\varepsilon),$$

so that

$$\begin{aligned}V(x,v)\leq& J(x,v,\hat{\alpha}_\varepsilon)\\ =& l(x,v,a)+\beta J(L(x,v,a),H(x,v,a),\alpha_\varepsilon)\\ \leq& l(x,v,a)+\beta V(L(x,v,a),H(x,v,a))+\beta\varepsilon.\end{aligned}$$

As $\varepsilon>0$ and $a\in\mathbb{A}$ where arbitrary, the latter implies

$$V(x,v)\leq \inf_{a\in\mathbb{A}}\left\{l(x,v,a)+\beta V(L(x,v,a),H(x,v,a))\right\}.$$

□

Lemma 9.2 means that V solves the dynamic programming equation

$$V(x,v) = \inf_{a\in\mathbb{A}}\left\{l(x,v,a)+\beta V(L(x,v,a),H(x,v,a))\right\}, \tag{9.4}$$

where $(x,v)\in D$.

Definition 9.3 *Let us define*

$$G(u) := \inf_{a \in \mathbb{A}} \{l(x,v,a) + \beta u(L(x,v,a), H(x,v,a))\}.$$

We call u a subsolution of (9.4)*, if*

$$u(x,v) \leq G(u)(x,v), \qquad \forall (x,v) \in D.$$

Similarly, u is called supersolution of (9.4)*, if*

$$u(x,v) \geq G(u)(x,v), \qquad \forall (x,v) \in D.$$

Further, u is called a solution of (9.4)*, if it is both sub- and supersolution.*

□

As for continuous problems, also for the time discrete problem a comparison principle holds.

Proposition 9.4 *Let u_1 be a bounded subsolution of* (9.4) *and u_2 be a bounded supersolution of that equation. Then*

$$u_1(x,v) \leq u_2(x,v), \qquad \forall (x,v) \in D.$$

In particular, there is at most one solution of (9.4)*.*

Proof: Let $(x,v) \in D$. By definition, there exists, for all $\varepsilon > 0$, some $a^\varepsilon = a^\varepsilon(x,v) \in \mathbb{A}$, such that

$$u_2(x,v) \geq G(u_2)(x,v) \geq l(x,v,a^\varepsilon) + \beta u_2(L(x,v,a^\varepsilon), H(x,v,a^\varepsilon)) - \varepsilon.$$

On the other hand, it holds

$$u_1(x,v) \leq G(u_1)(x,v) \leq l(x,v,a^\varepsilon) + \beta u_1(L(x,v,a^\varepsilon), H(x,v,a^\varepsilon)).$$

Thus,

$$\begin{aligned} u_1(x,v) - u_2(x,v) & \\ &\leq \beta\left(u_1(L(x,v,a^\varepsilon), H(x,v,a^\varepsilon)) - u_2(L(x,v,a^\varepsilon), H(x,v,a^\varepsilon))\right) + \varepsilon \\ &\leq \beta \sup_{(\bar{x},\bar{v}) \in D} \{u_1(\bar{x},\bar{v}) - u_2(\bar{x},\bar{v})\} + \varepsilon. \end{aligned}$$

As this is true for every $(x,v) \in D$, the latter implies that

$$\sup_{(x,v) \in D} \{u_1(x,v) - u_2(x,v)\} \leq \beta \sup_{(x,v) \in D} \{u_1(x,v) - u_2(x,v)\} + \varepsilon.$$

As $\varepsilon > 0$ was arbitrary and $\beta \in (0,1)$, we conclude that

$$\sup_{(x,v) \in D} \{u_1(x,v) - u_2(x,v)\} \leq 0,$$

i.e., $u_1 \leq u_2$. □

From lemma 9.2 and proposition 9.4, we can derive that V is the unique solution of (9.4). Before we tackle the problem of convergence for $h \downarrow 0$, we want to establish existence of optimal feedback controls in the time discrete case. To do so, we first need to analyze the properties of V and the solution operators L, H.

Proposition 9.5 *Assume that the hysteresis operator $\mathcal{W}$ satisfies, for all (u_1, v_1), $(u_2, v_2) \in L^2(\Omega) \times L^2(\Omega)$, the Lipschitz type inequality*

$$\|\mathcal{W}(u_1, v_1) - \mathcal{W}(u_2, v_2)\| \leq c_{\mathcal{W}} \left(\|u_1 - u_2\| + \|v_1 - v_2\|\right),$$

where $c_{\mathcal{W}} > 0$. Then, for every $(a_1, b_1, c), (a_2, b_2, c) \in D \times \mathbb{A}$, one has the estimates

$$\begin{aligned} \|L(a_1, b_1, c) - L(a_2, b_2, c)\| &\leq (1 + h) \left(\|a_1 - a_2\| + \|b_1 - b_2\|\right), \\ \|H(a_1, b_1, c) - H(a_2, b_2, c)\| &\leq (2 + h) c_{\mathcal{W}} \left(\|a_1 - a_2\| + \|b_1 - b_2\|\right). \end{aligned}$$

Further, if in addition l satisfies

$$|l(a_1, b_1, c) - l(a_2, b_2, c)| \leq c_l \left(\|a_1 - a_2\| + \|b_1 - b_2\|\right),$$

for some $c_l > 0$, and $0 < \beta < (2 \max\{1 + h, (2 + h) c_{\mathcal{W}}\})^{-1}$, then

$$|V(a_1, b_1) - V(a_2, b_2)| \leq \frac{1}{1 - \eta} \left(\|a_1 - a_2\| + \|b_1 - b_2\|\right),$$

where $\eta := 2\beta \max\{1 + h, (2 + h) c_{\mathcal{W}}\}$.

Proof: For $(a_1, b_1, c), (a_2, b_2, c) \in D \times \mathbb{A}$ consider $y_1 := L(a_1, b_1, c)$ and $y_2 := L(a_2, b_2, c)$. By definition, y_1 solves

$$y_1 = h\Delta y_1 + a_1 + hc - hb_1,$$

whereas y_2 is the solution of

$$y_2 = h\Delta y_2 + a_2 + hc - hb_2.$$

Hence, the difference $y := y_1 - y_2$ satisfies

$$y = h\Delta y + a - hb,$$

where $a := a_1 - a_2$ and $b := b_1 - b_2$. Testing this equation with y yields

$$\|y\|^2 + h \|\nabla y\|^2 = \int_\Omega (a - hb) y \, d\mathcal{L},$$

which implies, due to Young's inequality,

$$\|y\|^2 \leq \|a - hb\|^2.$$

Thus,

$$\|y\| \leq \|a\| + h\,\|b\| \leq (1+h)\,(\|a\| + \|b\|)\,,$$

which proves the first inequality. Next, let us denote $w_1 := H(a_1, b_1, c)$, $w_2 := H(a_2, b_2, c)$, and recall that

$$H(a_j, b_j, c) := \mathcal{W}(L(a_j, b_j, c), b_j).$$

Hence, using the Lipschitz assumption on $\mathcal{W}$ and the inequality for L, we see that

$$\begin{aligned}\|w_1 - w_2\| \leq & c_{\mathcal{W}}\,(\|L(a_1, b_1, c) - L(a_1, b_1, c)\| + \|b_1 - b_2\|)\\ \leq & c_{\mathcal{W}}\,((1+h)\,(\|a_1 - a_2\| + \|b_1 - b_2\|) + \|b_1 - b_2\|)\\ \leq & (2+h)c_{\mathcal{W}}\,(\|a_1 - a_2\| + \|b_1 - b_2\|)\,,\end{aligned}$$

as claimed. Now, let $\varepsilon > 0$ and $\alpha^{\varepsilon} \in \mathcal{A}$ such that

$$V(a_2, b_2) \geq J(a_2, b_2, \alpha^{\varepsilon}) - \frac{\varepsilon}{2}.$$

Then, for every $N \in \mathbb{N}$, we derive from the Lipschitz assumption on l that

$$\begin{aligned}& V(a_1, b_1) - V(a_2, b_2) \leq J(a_1, b_1, \alpha^{\varepsilon}) - J(a_2, b_2, \alpha^{\varepsilon}) + \frac{\varepsilon}{2}\\ & \leq \sum_{n=0}^{N} \beta^n c_l\,(\|y_n(a_1, b_1, \alpha^{\varepsilon}) - y_n(a_2, b_2, \alpha^{\varepsilon})\| + \|w_n(a_1, b_1, \alpha^{\varepsilon}) - w_n(a_2, b_2, \alpha^{\varepsilon})\|)\\ & \quad + \sum_{n=N+1}^{\infty} 2M\beta^n + \frac{\varepsilon}{2},\end{aligned}$$

where M is the bound on l. Choose N such that $2M\sum_{n=N+1}^{\infty}\beta^n \leq \frac{\varepsilon}{2}$. Iteration of the inequalities for L and H yields (denoting for short, $y_n^j := y_n(a_j, b_j, \alpha^{\varepsilon}), w_n^j := w_n(a_j, b_j, \alpha^{\varepsilon})$)

$$\begin{aligned}\left\|y_n^1 - y_n^2\right\| \leq & (1+h)\left(\left\|y_{n-1}^1 - y_{n-1}^2\right\| + \left\|w_{n-1}^1 - w_{n-1}^2\right\|\right)\\ \leq & (1+h)^2\left(\left\|y_{n-2}^1 - y_{n-2}^2\right\| + \left\|w_{n-2}^1 - w_{n-2}^2\right\|\right)\\ & + (1+h)(2+h)c_{\mathcal{W}}\left(\left\|y_{n-2}^1 - y_{n-2}^2\right\| + \left\|w_{n-2}^1 - w_{n-2}^2\right\|\right)\\ \leq & 2\,(\max\{1+h, (2+h)c_{\mathcal{W}}\})^2\left(\left\|y_{n-2}^1 - y_{n-2}^2\right\| + \left\|w_{n-2}^1 - w_{n-2}^2\right\|\right)\\ \leq & \ldots\\ \leq & 2^{n-1}\,(\max\{1+h, (2+h)c_{\mathcal{W}}\})^n\,(\|a_1 - a_2\| + \|b_1 - b_2\|)\,,\end{aligned}$$

and, analogously,

$$\left\|w_n^1 - w_n^2\right\| \leq 2^{n-1}\,(\max\{1+h, (2+h)c_{\mathcal{W}}\})^n\,(\|a_1 - a_2\| + \|b_1 - b_2\|)\,.$$

Therefore we get

$$V(a_1, b_1) - V(a_2, b_2) \leq \sum_{n=0}^{N} (2\beta \max\{1+h, (2+h)c_{\mathcal{W}}\})^n (\|a_1 - a_2\| + \|b_1 - b_2\|) + \varepsilon. \tag{9.5}$$

Hence, if $\eta := 2\beta \max\{1+h, (2+h)c_{\mathcal{W}}\} \in (0,1)$, as $\varepsilon > 0$ was arbitrary,

$$V(a_1, b_1) - V(a_2, b_2) \leq \frac{1}{1-\eta} (\|a_1 - a_2\| + \|b_1 - b_2\|).$$

Interchanging the roles of $(a_1, b_1), (a_2, b_2)$ then proves the claim.

□

Remark 9.6 *The assumption on the hysteresis operator is fulfilled, e.g., in the case of pointwise applied operators meeting a Lipschitz type inequality. Assume that for (a.e.)* $x \in \Omega$,

$$|w_1(x) - w_2(x)| \leq c_{\mathcal{W}} (|L(a_1, b_1, c)(x) - L(a_2, b_2, c)(x)| + |b_1(x) - b_2(x)|).$$

Then, an analogous inequality with $|\cdot|$ *replaced by* $\|\cdot\|$ *must hold. An application of the triangle inequality then yields*

$$\|w_1 - w_2\| \leq c_{\mathcal{W}} (\|L(a_1, b_1, c) - L(a_2, b_2, c)\| + \|b_1 - b_2\|),$$

and one can proceed as shown in the above proof.

□

Next note that since L assigns, for input values y_m, w_m, a, the solution y_{m+1} of the equation

$$y_{m+1} = h\Delta y_{m+1} + y_m + ha - hw_m,$$

it holds $y_{m+1} \in H^2(\Omega)$, which is compactly imbedded into $L^2(\Omega)$. From the definition of weak solutions, we see that for every sequence $(a^k)_{k\in\mathbb{N}} \subset \mathbb{A}$ that converges weakly to some $a^* \in \mathbb{A}$, the corresponding outputs

$$y_{m+1}^k := L(y_m, w_m, a^k)$$

converge weakly in $H^2(\Omega)$, and thus $y_{m+1}^k \to y_{m+1}^*$ in $L^2(\Omega)$ by the compactness of the imbedding. This means that the mapping $a \mapsto L(y_m, w_m, a)$ is weakly sequentially continuous for fixed y_m, w_m when considered as map from $L^2(\Omega)$ into itself. If $\mathcal{W}$ is continuous, then also $a \mapsto H(y_m, w_m, a)$ is weakly sequentially continuous for fixed y_m, w_m. Using this, we are able to prove existence of optimal controls.

Theorem 9.7 *Let the assumptions of proposition 9.5 hold, except for the restriction on β. Then V is continuous from $(L^2(\Omega))^2 \to \mathbb{R}$. If in addition, $\mathbb{A} \subset L^2(\Omega)$ is weakly compact and the mapping $a \mapsto l(x,v,a)$ is weakly lower semicontinuous for every (x,v), there is an optimal control $\alpha^* = (a_n^*)_n$ which can be found by iteratively solving the feedback equations*

$$V(y_n^*, w_n^*) = l(y_n^*, w_n^*, a_n^*) + \beta V(L(y_n^*, w_n^*, a_n^*), H(y_n^*, w_n^*, a_n^*)). \tag{9.6}$$

Proof: Let $(a_n, b_n)_{n\in\mathbb{N}}$ be a sequence of initial values such that $(a_n, b_n) \to (a,b)$ in $L^2(\Omega)$ as $n \to \infty$. From equation (9.5), we get that for every $\varepsilon > 0$,

$$\limsup_{n\to\infty} V(a_n, b_n) - V(a,b) \le \varepsilon,$$
$$V(a,b) - \liminf_{n\to\infty} V(a_n, b_n) \le \varepsilon,$$

which implies

$$\lim_{n\to\infty} V(a_n, b_n) = V(a,b).$$

Next we want to show that for every $(x,v) \in (L^2(\Omega))^2$ there is $a^* \in \mathbb{A}$ such that

$$\begin{aligned} &l(x,v,a^*) + \beta V(L(x,v,a^*), H(x,v,a^*)) \\ &\qquad = \inf_{a\in\mathbb{A}} \left\{ l(x,v,a) + \beta V(L(x,v,a), H(x,v,a)) \right\}. \end{aligned} \tag{9.7}$$

To this end, let $(a_n)_n$ be a minimizing sequence. Due to the weak compactness of $\mathbb{A}$, we may w.l.o.g. assume that a_n converges weakly to some $a_\infty \in \mathbb{A}$. The weak continuity of L and H together with the continuity of V and the weak lower semicontinuity of l imply that a_∞ is in fact optimal in the sense that it meets (9.7). Now choose, for $y_0^* =: x$ and $w_0^* =: v$ some $a_0^* \in \mathbb{A}$ such that (9.7) is valid. Then one can solve the elliptic equation which yields $y_1^* = L(y_0^*, w_0^*, a_0^*)$, and calculate w_1^*. Iterating this method yields trajectories $(y_n^*)_n, (w_n^*)_n$ and a control $\alpha^* = (a_n^*)_n$. We show that α^* is optimal for the control problem. From (9.7) and the discrete dynamic programming principle (lemma 9.2), we find that the equation

$$V(y_n^*, w_n^*) = l(y_n^*, w_n^*, a_n^*) + \beta V(y_{n+1}^*, w_{n+1}^*)$$

holds. Multiplication with β^n leads to

$$\beta^n V(y_n^*, w_n^*) - \beta^{n+1} V(y_{n+1}^*, w_{n+1}^*) = \beta^n l(y_n^*, w_n^*, a_n^*).$$

But then, by definition of J,

$$\begin{aligned} J(x,v,\alpha^*) = \sum_{n=0}^{\infty} l(y_n^*, w_n^*, a_n^*)\beta^n &= \sum_{n=0}^{\infty} \left(\beta^n V(y_n^*, w_n^*) - \beta^{n+1} V(y_{n+1}^*, w_{n+1}^*) \right) \\ &= V(x,v), \end{aligned}$$

which means that α^* is optimal. □

Our next goal is to analyze what happens when $h \downarrow 0$. As the discretization is almost identical to the one used in theorem 3.3, we will show convergence by comparison of the two solutions. As mentioned before, we will restrict to initial values in D. Recall that the weak solution of (9.1) fulfills

$$\begin{aligned}\frac{1}{h}\int_\Omega (y_{m+1}-y_m)\varphi d\mathcal{L} + \int_\Omega w_m \varphi d\mathcal{L} + \int_\Omega \nabla y_{m+1}\cdot\nabla\varphi d\mathcal{L} \\ = \int_\Omega a_m \varphi d\mathcal{L},\ \forall \varphi \in H_0^1(\Omega),\end{aligned}$$

whereas the weak solution z of (3.2) fulfills

$$\begin{aligned}\frac{1}{h}\int_\Omega (z_{m+1}-z_m)\,\varphi d\mathcal{L} + \int_\Omega v_{m+1} \varphi d\mathcal{L} + \int_\Omega \nabla z_{m+1}\cdot\nabla\varphi d\mathcal{L} \\ = \int_\Omega b_m \varphi d\mathcal{L},\ \forall \varphi \in H_0^1(\Omega).\end{aligned}$$

Here, $w_m := \mathcal{W}(y_0,\ldots,y_m)$ and $v_m := \mathcal{W}(z_0,\ldots,z_m)$, and b_m corresponds to a step function $b^N := \sum_{m=0}^N b_m\chi_m$ that converges to some $b \in L^2(\Omega\times(0,T))$ as $N\to\infty$. At this point we remark that weak convergence of b^N would be sufficient, as this restriction would neither influence the a priori estimates nor the passage to the limit of the tested equation. Assume now that the hysteresis operator $\mathcal{W}$ has the property that for all strings $s^1=(s_1^1,\ldots,s_n^1)$, $s^2=(s_1^2,\ldots,s_n^2)$,

$$\left|\mathcal{W}(s^1)-\mathcal{W}(s^2)\right| \le c_\mathcal{W} \max_{0\le k\le n}\left\{\left|s_k^1-s_k^2\right|\right\}. \tag{9.8}$$

Then we may estimate (pointwise)

$$\begin{aligned}|w_m - v_{m+1}| &= |\mathcal{W}(y_0,\ldots,y_m,y_m)-\mathcal{W}(z_0,\ldots,z_{m+1})| \\ &\le c_\mathcal{W}\max\left\{\max_{0\le k\le m}\{|y_k-z_k|\},|y_m-z_{m+1}|\right\} \\ &\le c_\mathcal{W}\max_{0\le k\le m+1}\{|y_k-z_k|\} + c_\mathcal{W}|z_m-z_{m+1}|.\end{aligned}$$

Since we know that each z_k is bounded in L^2 from the arguments in the proof of theorem 3.3, this yields similar a priori estimates for the difference $q_m := y_m - z_m$; thus, we can argue as for theorem 3.3 to show convergence of $\sum_m q_m\chi_m$ to zero. Hence, we have proved the following:

Theorem 9.8 *Let $\Omega\subset\mathbb{R}^n$ be an open bounded domain with smooth boundary (at least C^2), and assume that the (hysteresis-) operator $\mathcal{W}$ satisfies* (9.8)*. If the step function corresponding to the control $\alpha=(a_0,a_1,\ldots)$ converges weakly in $L^2(\Omega_T)$, then the step function corresponding to the solutions of* (9.1) *converge to the unique solution of* (3.1)*, particularly in $L^2(\Omega;L^\infty(0,T))$.*

Proof: We have noticed that q_m can be treated just as y_m; thus, the linear interpolates of z_m must converge to the solution of (3.1) – which is unique due to theorem 3.6, as (9.8) implies that assumption 3.4 is valid for the continuation of the (hysteresis-) operator to continuous functions – w.r.t. the norm of $L^2(\Omega; C[0,T])$ (cf., theorem 3.3). But then the constant interpolates must converge in $L^2(\Omega; L^\infty(0,T))$, which proves the claim.

□

If $\mathbb{A}$ is weakly compact, we have seen that one can find an optimal control $\alpha^* = (a_0^*, a_1^*, \ldots)$ for the discretized problem via iteratively solving the feedback equations (9.6). Further, we can assign step functions $\alpha_h^* := \sum_{m=0}^N a_m^* \chi_m$ to α^*, which then must contain weakly convergent subsequences (in $L^2(\Omega_T)$). As theorem 9.8 tells us that trajectories converge, in that case, to solutions of (3.1), it would be plausible that such weak limit of $\alpha_{h_n}^*$ might be an optimal control for the continuous problem. We will investigate this in the rest of this section. To distinguish between discrete and continuous problem, we will denote the corresponding costs by J_c and J_d, where

$$J_c(x,v,\alpha) := \int_0^\infty \tilde{l}(y_{x,v,\alpha}(t), w_{x,v,\alpha}(t), \alpha(t))e^{-\lambda t}dt, \ \alpha \in \mathcal{A}_c$$

$$J_d(x,v,\alpha) := \sum_{n=0}^\infty h\tilde{l}(y_n(x,v,\alpha), w_n(x,v,\alpha), a_n)(1-\lambda h)^n, \ \alpha \in \mathcal{A}_d.$$

Here, $\mathcal{A}_c$, as usual, consists of all measurable functions α taking values $\alpha(t) \in \mathbb{A}$ for all $t \geq 0$, and $\mathcal{A}_d$ consists of all sequences $(a_0, a_1, \ldots)$ such that $a_j \in \mathbb{A}$ for all $j \in \mathbb{N}$. We note that we arrive at J_d if we set $\beta = 1 - \lambda h$ and $l = h\tilde{l}$ in (9.3). This choice accords to the one of [4], where convergence is shown with the help of HJB equations. As we do not have such for the continuous problem, we need to take a different approach. To be able to compare solutions, we will focus on a special type of controls. Let $T > 0$. By $\mathcal{A}_{pc}^T$, we will denote the set of controls $\bar{\alpha} \in \mathcal{A}_c$ for which there is $k \in \mathbb{N}$ (we will refer to this number as **discretization level**) such that $\bar{\alpha}$ is a.e. equal to a constant on every interval of the form $(\frac{mT}{2^k}, \frac{(m+1)T}{2^k})$, $m \in \mathbb{N}$. The nice thing about this type of function is that if it has such property for $k \in \mathbb{N}$, then also for every $\tilde{k} \geq k$, and we can interprete it as control for the discretized problem. For this type of control, we get the following result.

Theorem 9.9 *Let the assumptions of theorem 9.8 hold and assume in addition that $\tilde{l}$ is bounded by $M > 0$ and has the Lipschitz type property that for all $(a_1, b_1), (a_2, b_2) \in D, c \in \mathbb{A}$,*

$$\left|\tilde{l}(a_1, b_1, c) - \tilde{l}(a_2, b_2, c)\right| \leq c_{\tilde{l}} \left(\|a_1 - a_2\| + \|b_1 - b_2\|\right). \tag{9.9}$$

Then, for every $T > 0$, $(x, v) \in D$ and $(\alpha^k)_k \subset \mathcal{A}_{pc}^T$ with discretization level $k \in \mathbb{N}$ such that α^k converges weakly in $L^2(\Omega_\tau)$ for all $\tau > 0$,

$$\left|J_c(x, v, \alpha^k) - J_d(x, v, \alpha^k)\right| \to 0 \text{ as } k \to \infty,$$

where the step size is associated to k via $h_k = \frac{T}{2^k}$.

Proof: Let $\varepsilon > 0$ and choose $T \in \mathbb{N}$ such that

$$M \sum_{n=T+1}^{\infty} h(1-\lambda h)^n < \frac{\varepsilon}{2}, \qquad M \int_T^{\infty} e^{-\lambda t} dt < \frac{\varepsilon}{2}.$$

The first inequality can be achieved uniformly for $h > 0$ small enough, let's say, for $h < \frac{1}{2\lambda}$, so that we restrict to discretization levels $k \geq \tilde{k}$ where $\frac{T}{2^k} < \frac{1}{2\lambda}$. Let $\alpha^k \in \mathcal{A}_{pc}^T$ with discretization level $k \in \mathbb{N}$, $k \geq \tilde{k}$. The step size is $h_k = T/2^k$, so that we may write

$$J_c(x, v, \alpha) = \sum_{n=0}^{\infty} \int_{\frac{nT}{2^k}}^{\frac{(n+1)T}{2^k}} \tilde{l}(y_{x,v,\alpha}(t), w_{x,v,\alpha}(t), a_n^k) e^{-\lambda t} dt,$$

where $\alpha^k(t) = a_n^k$ (a.e.) on $(\frac{nT}{2^k}, \frac{(n+1)T}{2^k})$. Then, denoting $m := 2^k - 1$ and using the abbreviations $y^k(\cdot) := y_{x,v,\alpha^k}(\cdot)$, $w^k(\cdot) := w_{x,v,\alpha^k}(\cdot)$, $y_n^k := y_n(x, v, \alpha^k)$, $w_n^k := w_n(x, v, \alpha^k)$,

$$\begin{aligned}
&\left|J_c(x, v, \alpha^k) - J_d(x, v, \alpha^k)\right| \\
&\leq \sum_{n=0}^{m} \left| \int_{\frac{nT}{2^k}}^{\frac{(n+1)T}{2^k}} \tilde{l}(y^k(t), w^k(t), a_n^k) e^{-\lambda t} dt - h\tilde{l}(y_n^k, w_n^k, a_n^k)(1-\lambda h)^n \right| + \varepsilon \\
&\leq \sum_{n=0}^{m} \left| \int_{\frac{nT}{2^k}}^{\frac{(n+1)T}{2^k}} \tilde{l}(y^k(t), w^k(t), a_n^k) e^{-\lambda t} - \tilde{l}\left(y^k\left(\frac{nT}{2^k}\right), w^k\left(\frac{nT}{2^k}\right), a_n^k\right) e^{-\lambda \frac{nT}{2^k}} dt \right| \qquad (9.10) \\
&+ \sum_{n=0}^{m} \left| \frac{T}{2^k}\tilde{l}\left(y^k\left(\frac{nT}{2^k}\right), w^k\left(\frac{nT}{2^k}\right), a_n^k\right) e^{-\lambda \frac{nT}{2^k}} - \frac{T}{2^k}\tilde{l}(y_n^k, w_n^k, a_n^k) e^{-\lambda \frac{nT}{2^k}} \right| \qquad (9.11) \\
&+ \sum_{n=0}^{m} \left| \frac{T}{2^k}\tilde{l}(y_n^k, w_n^k, a_n^k) e^{-\lambda \frac{nT}{2^k}} - \frac{T}{2^k}\tilde{l}(y_n^k, w_n^k, a_n^k)(1-\lambda h)^n \right| + \varepsilon. \qquad (9.12)
\end{aligned}$$

We estimate (9.10) to (9.12). From $\left|\tilde{l}\right| \leq M$ and the Lipschitz property,

$$(9.10) \leq \sum_{n=0}^{m} \left| \int_{\frac{nT}{2^k}}^{\frac{(n+1)T}{2^k}} \tilde{l}(y^k(t), w^k(t), a_n^k) e^{-\lambda t} - \tilde{l}(y^k(t), w^k(t), a_n^k) e^{-\lambda \frac{nT}{2^k}} dt \right|$$

$$+\sum_{n=0}^{m}\left|\int_{\frac{nT}{2^k}}^{\frac{(n+1)T}{2^k}} \tilde{l}(y^k(t), w^k(t), a_n^k)e^{-\lambda\frac{nT}{2^k}} - \tilde{l}\left(y^k\left(\frac{nT}{2^k}\right), w^k\left(\frac{nT}{2^k}\right), a_n^k\right)e^{-\lambda\frac{nT}{2^k}}\,dt\right|$$
$$\le \sum_{n=0}^{m} M\int_{\frac{nT}{2^k}}^{\frac{(n+1)T}{2^k}} \left|e^{-\lambda t} - e^{-\lambda\frac{nT}{2^k}}\right| dt$$
$$+\sum_{n=0}^{m}\int_{\frac{nT}{2^k}}^{\frac{(n+1)T}{2^k}} \left|\tilde{l}(y^k(t), w^k(t), a_n^k) - \tilde{l}\left(y^k\left(\frac{nT}{2^k}\right), w^k\left(\frac{nT}{2^k}\right), a_n^k\right)\right| dt$$
$$\le \sum_{n=0}^{m}\frac{MT}{2^k}\left(e^{-\lambda\frac{nT}{2^k}} - e^{-\lambda\frac{(n+1)T}{2^k}}\right)$$
$$+\sum_{n=0}^{m} c_{\tilde{l}}\int_{\frac{nT}{2^k}}^{\frac{(n+1)T}{2^k}}\left(\left\|y^k(t) - y^k\left(\frac{nT}{2^k}\right)\right\| + \left\|w^k(t) - w^k\left(\frac{nT}{2^k}\right)\right\|\right) dt$$
$$\le \frac{MT}{2^k} + c_{\tilde{l}}T\max_{0\le n\le m}\sup_{t\in\left(\frac{nT}{2^k},\frac{(n+1)T}{2^k}\right)}\left\{\left\|y(t) - y\left(\frac{nT}{2^k}\right)\right\| + \left\|w(t) - w\left(\frac{nT}{2^k}\right)\right\|\right\},$$

which converges to zero as $k \to \infty$ due to the regularity of solutions for the heat equation with hysteresis and their strong convergence whenever controls converge weakly, see theorem 9.8. Further,

$$(9.11) \le \sum_{n=0}^{m}\frac{\tilde{c}_l T}{2^k}\left(\left\|y^k\left(\frac{nT}{2^k}\right) - y_n^k\right\| + \left\|w^k\left(\frac{nT}{2^k}\right) - w_n^k\right\|\right),$$

and the right hand side converges to zero as $k \to \infty$ due to theorem 9.8. Next, note that

$$(9.12) \le \sum_{n=0}^{m}\frac{MT}{2^k}\left|e^{-\lambda\frac{nT}{2^k}} - \left(1 - \lambda\frac{T}{2^k}\right)^n\right|.$$

Using the formula $r^n - s^n = (r-s)(r^{n-1} + r^{n-2}s + \ldots + rs^{n-2} + s^{n-1})$ with

$$r = e^{-\lambda\frac{T}{2^k}}, \qquad s = 1 - \lambda\frac{T}{2^k},$$

yields (note that $0 < 1 - \lambda\frac{T}{2^k} < 1$)

$$(9.12) \le M\frac{T}{2^k}\left|e^{-\lambda\frac{T}{2^k}} - 1 + \lambda\frac{T}{2^k}\right|\left[\sum_{j=1}^{m}\sum_{n=1}^{j} e^{-\lambda\frac{(n-1)T}{2^k}}\right]. \tag{9.13}$$

Let us make the definitions

$$\mu_k := \left|e^{-\lambda\frac{T}{2^k}} - 1 + \lambda\frac{T}{2^k}\right|, \qquad \beta_k := e^{-\lambda\frac{T}{2^k}},$$

$$S := \left[\sum_{j=1}^{m}\sum_{n=1}^{j} e^{-\lambda\frac{(n-1)T}{2^k}}\right] = \left[\sum_{j=1}^{m}\sum_{n=1}^{j} \beta_k^{n-1}\right].$$

Then, recalling that $m = 2^k - 1$, a simple calculation shows

$$S = \sum_{j=1}^{m} \frac{1-\beta_k^j}{1-\beta_k} = \frac{1}{1-\beta_k}\left[2^k - 1 - \left(\sum_{j=1}^{m}\beta_k^j\right)\right] = \frac{1}{1-\beta_k}\left[2^k - \frac{1-\beta_k^{m+1}}{1-\beta_k}\right].$$

Next note that

$$\left(1-e^{-\lambda\frac{T}{2^k}}\right)\frac{2^k}{T} = \lambda + \frac{o(\frac{T}{2^k})}{\frac{T}{2^k}} \xrightarrow{k\to\infty} \lambda,$$

which implies

$$\frac{\frac{T}{2^k}}{1-\beta_k} \xrightarrow{k\to\infty} \frac{1}{\lambda} > 0.$$

Therefore, with $\eta_k := \frac{T}{2^k}$,

$$S\eta_k^2 = \frac{\eta_k}{1-\beta_k}\left[\eta_k 2^k - \frac{\eta_k}{1-\beta_k}\left(1-\beta_k^{m+1}\right)\right] \xrightarrow{k\to\infty} \frac{1}{\lambda}\left[\lambda T - \frac{1}{\lambda}\left(1-e^{-\lambda T}\right),\right]$$

which in view of (9.13) implies that

$$(9.12) \le M\left(S\eta_k^2\right)\frac{\mu_k}{\eta_k} \xrightarrow{k\to\infty} 0,$$

because $\mu_k = o(\eta_k)$. Combining all estimates, this shows that

$$\limsup_{k\to\infty}\left|J_c(x,v,\alpha^k) - J_d(x,v,\alpha^k)\right| \le \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, this proves the result.

□

Now we can show optimality of the weak limit α^*.

Theorem 9.10 *Let $\Omega \subset \mathbb{R}^n$ be some open bounded domain with smooth boundary (at least C^2) and $\mathbb{A} \subset L^2(\Omega)$ be weakly compact. Assume that $\mathcal{W}$ satisfies* (9.8), *$\left|\tilde{l}\right| \le M$, $\tilde{l}$ satisfies* (9.9), *and that the map*

$$\eta : \mathbb{A} \to \mathbb{R}, \quad a \mapsto \tilde{l}(x,v,a),$$

is convex for each $(x,v) \in D$. Then every weak limit of step functions $(\alpha_c^n) \subset \mathcal{A}_{pc}^T$ corresponding to $(\alpha^n) \in \mathcal{A}_d$ found by iteratively solving (9.6) *is an optimal control for the continuous optimization problem.*

Proof: Fix $T > 0$. As solutions to the heat equation with hysteresis depend continuously on the data, there exists a sequence of controls $(\beta^k) \subset \mathcal{A}_{pc}^T$ of discretization level $k \in \mathbb{N}$ such that

$$\lim_{n\to\infty} J_c(x, v, \beta^n) = V_c(x, v), \qquad \forall (x, v) \in D.$$

Let, for every $k \in \mathbb{N}$, denote (α^k) an optimal control of the discretized problem at discretization level k. Then, for all k,

$$J_d(x, v, \alpha^k) \leq J_d(x, v, \beta^k),$$

which implies

$$\begin{aligned} J_c(x, v, \alpha^k) + J_d(x, v, \alpha^k) - J_c(x, v, \alpha^k)& \\ \leq J_d(x, v, \beta^k) - J_c(x, v, \beta^k) + J_c(x, v, \beta^k),& \end{aligned}$$

and hence,

$$\begin{aligned} &J_c(x, v, \alpha^k) \\ &\leq \left|J_d(x, v, \alpha^k) - J_c(x, v, \alpha^k)\right| + \left|J_d(x, v, \beta^k) - J_c(x, v, \beta^k)\right| + J_c(x, v, \beta^k). \end{aligned}$$

As the first two expressions on the right hand side converge to zero as $k \to \infty$ by theorem 9.9, it suffices to show that

$$J_c(x, v, \alpha^*) \leq \liminf_{k\to\infty} J_c(x, v, \alpha^k),$$

where α^* denotes the weak limit of the step function corresponding to (α^k), i.e., that the map $\alpha \mapsto J(x, v, \alpha)$ is weak lower sequentially semicontinuous for all $(x, v) \in D$. We show this property now. To this end, let $\varepsilon > 0$ and choose $T > 0$ such that

$$M \int_T^\infty e^{-\lambda t} dt \leq \frac{\varepsilon}{3}.$$

Then, using theorem 9.8 and the Lipschitz property of $\tilde{l}$, we find that

$$\left|\int_0^T e^{-\lambda t}(\tilde{l}(y_{x,v,\alpha^k}(t), w_{x,v,\alpha^k}(t), \alpha^k(t)) - \tilde{l}(y_{x,v,\alpha^*}(t), w_{x,v,\alpha^*}(t), \alpha^k(t)))dt\right| \to 0,$$

as $k \to \infty$. Thus, there exists $\tilde{k} \in \mathbb{N}$ such that the former is less than $\frac{\varepsilon}{3}$ for every $k \geq \tilde{k}$. But the functional

$$J^* : \mathcal{A}|_{(0,T)} \to \mathbb{R}, \quad \alpha \mapsto \int_0^T e^{-\lambda t}\tilde{l}(y_{x,v,\alpha^*}(t), w_{x,v,\alpha^*}(t), \alpha(t))dt,$$

is convex, since η is convex, and therefore also weak sequentially lower semicontinuous. Hence,

$$J^*(\alpha^*) \leq \liminf_{k\to\infty} J^*(\alpha^k).$$

In summary, we get

$$\begin{aligned}
J_c(x, v, \alpha^*) &= J^*(\alpha^*) + J_c(x, v, \alpha^*) - J^*(\alpha^*) \\
&\leq \liminf_{k\to\infty} \left\{ J^*(\alpha^k) \right\} + \frac{\varepsilon}{3} \\
&= \liminf_{k\to\infty} \left\{ J_c(x, v, \alpha^k) + J^*(\alpha^k) - J_c(x, v, \alpha^k) \right\} + \frac{\varepsilon}{3} \\
&\leq \liminf_{k\to\infty} \left\{ J_c(x, v, \alpha^k) + \frac{2\varepsilon}{3} \right\} + \frac{\varepsilon}{3} \\
&= \liminf_{k\to\infty} \left\{ J_c(x, v, \alpha^k) \right\} + \varepsilon.
\end{aligned}$$

As $\varepsilon > 0$ was arbitrary, the result follows.

□

PART THREE

10 A model problem corresponding to an ordinary differential equation including the time derivative of a hysteresis nonlinearity

In this section, we show that the method of [1], i.e., to reformulate problems containing Play hysteresis nonlinearities as reflected control systems, can, in principle, also be used when the differential equation contains the time derivative of the Play. One only has to make little changes.

10.1 Properties of the o.d.e.

We study the ordinary differential equation

$$\begin{aligned} \dot{y} + \dot{w} &= f(y, w, \alpha), \\ w &= \mathcal{F}_r[y; w_0], \\ y(0) &= y_0, \\ w(0) &= w_0 \in [y_0 - r, y_0 + r], \end{aligned} \tag{10.1}$$

where we assume that the continuous function f meets the Lipschitz property

$$|f(a_1, b_1, c) - f(a_2, b_2, c)| \leq c_f \left(|a_1 - a_2| + |b_1 - b_2|\right), \ \forall a_1, a_2, b_1, b_2, c \in \mathbb{R}, \tag{10.2}$$

and the boundedness property

$$|f(a, b, c)| \leq M_f, \ \forall a, b, c \in \mathbb{R}. \tag{10.3}$$

We start with the following existence and uniqueness result for (10.1).

Theorem 10.1 *Under the stated assumptions on f, for every $\alpha \in L^2(0, T)$ and admissible pair of initial values, i.e. $y_0 \in \mathbb{R}$, $w_0 \in [y_0 - r, y_0 + r]$,* (10.1) *has exactly one solution $(y, w) \in W^{1,\infty}(0, T) \times W^{1,\infty}(0, T)$.*

Proof: Via implicit time discretization. For $m \in \mathbb{N}$, let $h = \frac{T}{m}$ be the step size of the discretization of $[0, T]$, and define, for $\alpha \in L^2(0, T)$, the step function

$$\alpha_n(t) := \frac{1}{h} \int_{nh}^{(n+1)h} \alpha(s) ds, \qquad t \in (nh, (n+1)h), 1 \leq n \leq m.$$

Starting at y_0, w_0, we want to iteratively solve

$$\frac{y_{n+1} - y_n}{h} + \frac{w_{n+1} - w_n}{h} = f(y_{n+1}, w_{n+1}, \alpha_{n+1}). \tag{10.4}$$

From (10.2), we infer the two inequalities

$$y_{n+1} - y_n + w_{n+1} - w_n \leq hc_f\left(|y_{n+1} - y_n| + |w_{n+1} - w_n|\right) + hM_f, \tag{10.5}$$
$$y_{n+1} - y_n + w_{n+1} - w_n \geq -hc_f\left(|y_{n+1} - y_n| + |w_{n+1} - w_n|\right) - hM_f. \tag{10.6}$$

If $h < 1/c_f$, then the piecewise monotonicity of the Play operator yields that there are $y_* < y^*$ such that (10.5) is violated if $y_{n+1} > y^*$ and (10.6) is violated if $y_{n+1} < y_*$. By continuity, there exists an intermediate value $\bar{y} \in [y_*, y^*]$ such that (10.4) holds with $y_{n+1} = \bar{y}$. Note also that this fixed point is unique. If this was not the case, there would be $(x_1, x_2), (z_1, z_2)$ such that

$$\frac{x_1 - y_n}{h} + \frac{x_2 - w_n}{h} = f(x_1, x_2, \alpha_{n+1}), \quad \frac{z_1 - y_n}{h} + \frac{z_2 - w_n}{h} = f(z_1, z_2, \alpha_{n+1}),$$

and by the piecewise monotonicity property of the Play, if $x_1 \geq z_1$, then also $x_2 \geq z_2$. So assume w.l.o.g. that $x_1 \geq z_1$. Subtracting the two equations leads then to

$$\frac{|x_1 - z_1|}{h} + \frac{|x_2 - z_2|}{h} \leq c_f\left(|x_1 - z_1| + |x_2 - z_2|\right),$$

which implies $c_f h \geq 1$, a contradiction. Hence, for every $m \in \mathbb{N}$ such that $h = \frac{T}{m} < \frac{1}{c_f}$, the discretized problem is uniquely solvable. To find a priori estimates, note that piecewise monotonicity together with inequalities (10.5), (10.6) imply in fact

$$|y_{n+1} - y_n| + |w_{n+1} - w_n| \leq hc_f\left(|y_{n+1} - y_n| + |w_{n+1} - w_n|\right) + hM_f,$$

so that restricting to just $h < 1/2c_f$, this yields

$$\frac{|y_{n+1} - y_n|}{h} + \frac{|w_{n+1} - w_n|}{h} \leq 2M_f.$$

Hence, the linear interpolates are bounded in $W^{1,\infty}(0,T)$. From the compact imbedding of the latter into $C[0,T]$, we infer that, at least for some subsequence, the linear interpolates converge to some functions $y, w \in W^{1,\infty}$ w.r.t. $\|\cdot\|_\infty$. Since the corresponding constant interpolates $\hat{y}_n, \hat{w}_n$ must then converge to the same limits w.r.t. the L^∞ norm, we get pointwise convergence (a.e.) of $f(\hat{y}_n, \hat{w}_n, \alpha_n)$ to $f(y, w, \alpha)$. The boundedness assumption on f together with Lebesgue's theorem of dominated convergence implies then convergence in $L^2(0,T)$ of $f(\hat{y}_n, \hat{w}_n, \alpha_n)$, so that y, w satisfy the differential equation a.e. in $(0,T)$. Further, from the convergence in $C[0,T]$, we also get that $w = \mathcal{F}_r[y; w_0]$. Hence, y, w solve (10.1). In order to show uniqueness, we assume that there are two solutions (y_1, w_1), (y_2, w_2) corresponding to different initial values (but same "controls" α) and subtract the two differential equations. Denoting $w := w_1 - w_2$, $y := y_1 - y_2$, we get

$$\dot{y} + \dot{w} = f(y_1, w_1, \alpha) - f(y_2, w_2, \alpha),$$

a.e. in $(0, T)$. Now we test this equation with $H(y)$ (where H denotes the Heavyside function) and use Hilpert's inequality (cf. [12, proposition 3.3.3]), which asserts that

$$\frac{d}{dt} w_+(t) \leq w'(t) H(y(t)),$$

for almost every t, where w_+ denotes the positive part of w. Therefore the tested equation yields

$$y_+(t) + w_+(t) \leq y_+(0) + w_+(0) + \int_0^t |f(y_1, w_1, \alpha) - f(y_2, w_2, \alpha)| \, ds.$$

Interchanging the roles of y_1, y_2 and w_1, w_2, we get an analogous inequality for the negative part; summing up those inequalities then yields

$$\begin{aligned} |y(t)| + |w(t)| &\leq |w_0| + |y_0| + 2 \int_0^t |f(y_1, w_1, \alpha) - f(y_2, w_2, \alpha)| \, ds \\ &\leq |w_0| + |y_0| + 2c_f \int_0^t (|y(s)| + |w(s)|) \, ds. \end{aligned}$$

As this is valid for each $t \in (0, T)$, we may apply Gronwall's lemma to get the stability result

$$|y(t)| + |w(t)| \leq (|w_0| + |y_0|) \cdot e^{2c_f t}, \tag{10.7}$$

which implies uniqueness of solutions.

□

If α is continuous, the differential equation shows that $y + w$ is in fact C^1, although one can never expect such regularity for Play hysteresis nonlinearities. What happens here is that $\dot{y}$ and $\dot{w}$ jump at the same time. As $\dot{w}$ equals either zero or $\dot{y}$, we can distinct the cases

1. $\dot{w} = 0 \Rightarrow \dot{y} = f(y, w, \alpha)$,
2. $\dot{w} = \dot{y} \Rightarrow \dot{w} = \dot{y} = \frac{1}{2} f(y, w, \alpha)$.

So when $\dot{y}$ jumps from $f(y, w, \alpha)$ to $\frac{1}{2} f(y, w, \alpha)$, then $\dot{w}$ jumps from zero to $\frac{1}{2} f(y, w, \alpha)$, so that the sum does not jump.

10.2 The control problem and Dynamic Programming - adaptation of the method of [1]

We build a standard infinite horizon control problem to equation (10.1). As usual, let $\lambda > 0$ be the discount factor and $l : \mathbb{R}^3 \to \mathbb{R}$ some continuous and bounded function with the Lipschitz type property

$$|l(a_1, b_1, c) - l(a_2, b_2, c)| \leq c_l \left(|a_1 - a_2| + |b_1 - b_2|\right), \; \forall a_1, a_2, b_1, b_2, c \in \mathbb{R}.$$

The function we want to minimize is then defined through

$$J(x, v, \alpha) := \int_0^\infty l\left(y_{x,v,\alpha}(t), w_{x,v,\alpha}(t), \alpha(t)\right) dt,$$

where $y_0 = x$, $w_0 = v \in [x - r, x + r]$ are the initial values. The corresponding value function is defined via

$$V(x, v) := \inf_{\alpha \in \mathcal{A}} J(x, v, \alpha),$$

where $\mathcal{A}$ is the set of admissible controls; it contains all measurable functions $\alpha : [0, \infty) \to \mathbb{A}$ for some compact set $\mathbb{A} \subset \mathbb{R}$. Let us introduce the set

$$\Omega_r := \left\{(x, v)^T \in \mathbb{R}^2 \; : \; v \in [x - r, x + r]\right\},$$

which contains exactly the admissible pairs of initial values. The value function has the following properties.

Proposition 10.2 *V is bounded and if $\lambda > 2c_f$, then it is also Lipschitz continuous on Ω_r. Moreover, it satisfies the dynamic programming equation*

$$\begin{aligned} &V(x, \xi) \\ &= \inf_{\alpha \in \mathcal{A}} \left\{ \int_0^t e^{-\lambda s} l(y_{x,\xi,\alpha}(s), w_{x,\xi,\alpha}(s), \alpha(s)) ds + e^{-\lambda t} V(y_{x,\xi,\alpha}(t), w_{x,\xi,\alpha}(t)) \right\}, \end{aligned}$$

for all $t > 0$.

Proof: The first part is a simple consequence of the boundedness of l and inequality (10.7). As the proof is standard (and similar to the one of theorem 5.1, e.g.), we omit the details. Further, the DPP follows as in section 6.1 from the semigroup properties.

□

Now recall the method of [1], that we used to handle the hysteresis part in sections 6 and 7. We rewrote the problem in terms of the stop operator "u" and then used that $\dot{u}$ equals $\dot{y}$ plus some element of the normal cone of some characteristic set Z at u. Since we could then replace $\dot{y}$ with the "right hand side" of the differential equation, we were able to treat the problem by switching to the pair of variables (y, u). If we try to do the same here, we will run into the problem of reproducing the term $\dot{w}$ whenever trying to substitute $\dot{y}$. We adapt the method by not replacing w, but only implicitly referring to u; i.e., we will directly replace $\dot{w}$ abstractly by some element of the normal cone of $N_Z(y-w)$, where $Z = [-r, r]$. In this way, we can avoid the explicit change of variables. Then, the existence result takes the following form.

Theorem 10.3 *V is a solution to the differential inclusion*

$$\lambda V(x,\xi) - (V_\xi(x,\xi) - V_x(x,\xi))\, N_Z(x-\xi) + \sup_{a\in\mathbb{A}} \{-V_x(x,\xi) f(x,\xi,a) - l(x,\xi,a)\} \ni 0,$$

in the sense that it is a ***subsolution****, i.e., there exists $p \in N_Z(x-\xi) \cap [-M_f, M_f]$, such that for every $\varphi \in C^1(\Omega_r)$, if $V - \varphi$ has a local maximum at (x,ξ), it holds*

$$\lambda V(x,\xi) - (\varphi_\xi(x,\xi) - \varphi_x(x,\xi))\, p + \sup_{a\in\mathbb{A}} \{-\varphi_x(x,\xi) f(x,\xi,a) - l(x,\xi,a)\} \leq 0,$$

and it is a ***supersolution****, i.e., there exists $q \in N_Z(x-\xi) \cap [-M_f, M_f]$, such that for every $\varphi \in C^1(\Omega_r)$, if $V - \varphi$ has a local minimum at (x,ξ) it holds*

$$\lambda V(x,\xi) - (\varphi_\xi(x,\xi) - \varphi_x(x,\xi))\, q + \sup_{a\in\mathbb{A}} \{-\varphi_x(x,\xi) f(x,\xi,a) - l(x,\xi,a)\} \geq 0.$$

Proof: Let $\varphi \in C^1(\Omega_r)$ and (x,ξ) be a local maximum point for $V - \varphi$. Then, for any constant control $\alpha \equiv a \in \mathbb{A}$, if $t > 0$ is small enough, we have

$$\begin{aligned}\varphi(x,\xi) - \varphi(y_{x,\xi,a}(t), w_{x,\xi,a}(t)) &\leq V(x,\xi) - V(y_{x,\xi,a}(t), w_{x,\xi,a}(t)) \\ &\leq \int_0^t e^{-\lambda s} l(y_{x,\xi,a}(s), w_{x,\xi,a}(s), a) ds \\ &\quad + e^{-\lambda t} \left(V(y_{x,\xi,a}(t), w_{x,\xi,a}(t)) - V(y_{x,\xi,a}(t), w_{x,\xi,a}(t))\right),\end{aligned}$$

where we used the DPP (proposition 10.2) for the second estimate. Next, we rewrite (using the notation $y(t) := y_{x,\xi,a}(t)$, $w(t) := w_{x,\xi,a}(t)$)

$$\varphi(x,\xi) - \varphi(y(t), w(t)) = -\int_0^t \varphi_x(y(s), w(s)) \left(f(y(s), w(s), a) - p(s)\right) ds$$

$$
\begin{aligned}
&- \int_0^t \varphi_\xi(y(s), w(s))p(s)ds \\
&= \int_0^t -\varphi_x(y(s), w(s))f(y(s), w(s), a)ds \\
&+ \int_0^t - \left(\varphi_\xi(y(s), w(s)) - \varphi_x(y(s), w(s))\right) p(s)ds,
\end{aligned}
$$

with some element $p(s) \in N_Z(y(s) - w(s)) \cap [-M_f, M_f]$, for (a.e.) $s \in (0, t)$. Hence, we arrive at

$$
\begin{aligned}
&- \frac{1}{t} \int_0^t \varphi_x(y(s), w(s))f(y(s), w(s), a)ds \\
&- \frac{1}{t} \int_0^t \left(\varphi_\xi(y(s), w(s)) - \varphi_x(y(s), w(s))\right) p(s)ds \\
&\leq \frac{1}{t} \int_0^t e^{-\lambda s} l(y(s), w(s), a)ds + \frac{e^{-\lambda t} - 1}{t} V(y(t), w(t)).
\end{aligned} \tag{10.8}
$$

The only term that we need to take a closer look at is the one containing $p(s)$. We distinguish two cases:

1. Case: $x - \xi \notin \partial Z$: By continuity, $y(t) - w(t) \notin \partial Z$ for all $t > 0$ small enough, so that we may assume $p(t) \equiv 0$.

2. Case: By continuity, $y(t) - w(t)$ is either near r or $-r$ for all small t, so we may assume that either $\frac{1}{t} \int_0^t p(s)ds \in [0, M_f]$ or $\in [-M_f, 0]$ for all such t. Thus, there exists a subsequence (t_n) with $t_n \downarrow 0$ such that $\frac{1}{t_n} \int_0^{t_n} p(s)ds \to p \in N_Z(x - \xi) \cap [-M_f, M_f]$.

Hence, if we switch to a suitable subsequence, we get by the usual arguments that

$$
-\varphi_x(x, \xi)f(x, \xi, a) - \left(\varphi_\xi(x, \xi) - \varphi_x(x, \xi)\right) p \leq l(x, \xi, a) - \lambda V(x, \xi).
$$

Since $a \in \mathbb{A}$ was arbitrary, the subsolution property follows. To prove that V is a supersolution, one assumes that $(x, \xi) \in \Omega_r$ is a local minimum for $V - \varphi$. By the usual arguments, there is then, for every $\varepsilon > 0$ and $t > 0$ small enough, a control $\alpha_\varepsilon \in \mathcal{A}$ such that

$$
\begin{aligned}
&\frac{1}{t} \int_0^t e^{-\lambda s} l(y(s), w(s), \alpha_\varepsilon(s))ds + \frac{e^{-\lambda t} - 1}{t} V(y(t), w(t)) - \varepsilon \\
&\qquad \leq -\frac{1}{t} \int_0^t \varphi_x(y(s), w(s))f(y(s), w(s), \alpha_\varepsilon(s))ds \\
&\qquad - \frac{1}{t} \int_0^t \left(\varphi_\xi(y(s), w(s)) - \varphi_x(y(s), w(s))\right) q(s)ds,
\end{aligned}
$$

where $q(s) \in N_Z(y(s) - w(s)) \cap [-M_f, M_f]$. Now, note that y, w are uniformly continuous w.r.t. the control variable (because $|f| \le M_f$). Hence, we may treat the term containing $q(s)$ as in the subsolution case. Further, we can make use of the inequality

$$-\varphi_x(x,\xi)f(x,\xi,\alpha_\varepsilon(s)) - l(x,\xi,\alpha_\varepsilon(s)) \le \sup_{a\in\mathbb{A}} \{-\varphi_x(x,\xi)f(x,\xi,a) - l(x,\xi,a)\},$$

for every $s > 0$. Then, using the uniform continuity of the appearing functions w.r.t. α_ε, we derive the supersolution property of V via standard arguments.

□

As usual, we prove a comparison result for sub- and supersolutions in order to show uniqueness. Here, it takes the following form.

Theorem 10.4 *Let u_1 be a subsolution and u_2 be a supersolution in the sense of theorem 10.3. If the u_i, $i \in \{1,2\}$, are Lipschitz continuous, then $u_1 \le u_2$.*

Proof: Define the auxiliary function $\Phi : (\Omega_r)^2 \to \mathbb{R}$ via

$$\Phi(x,\xi,y,\zeta) := u_1(x,\xi) - u_2(y,\zeta) - \frac{(x-y)^2}{2\varepsilon} - \frac{(\xi-\zeta)^2}{2\varepsilon} - \mu(x^2+y^2),$$

where ε, μ are positive parameters. Assume for contradiction that there was $(\tilde{x},\tilde{\xi}) \in \Omega_r$ such that

$$u_1(\tilde{x},\tilde{\xi}) - u_2(\tilde{x},\tilde{\xi}) = \delta > 0.$$

Then, there exists $\tilde{\mu} > 0$ such that for all $0 < \mu < \tilde{\mu}$,

$$\Phi(\tilde{x},\tilde{\xi},\tilde{x},\tilde{\xi}) = u_1(\tilde{x},\tilde{\xi}) - u_2(\tilde{x},\tilde{\xi}) - 2\mu\tilde{x} \ge \frac{\delta}{2}.$$

Hence, in particular, $\sup\Phi \ge \frac{\delta}{2} > 0$ for those μ. Further, the Lipschitz continuity implies (note that the "hysteresis variables" are bounded whenever x, y are, as $\max\{|\xi|,|\zeta|\} \le \max\{|x|,|y|\} + r$) the existence of $(\bar{x},\bar{\xi},\bar{y},\bar{\zeta}) \in \Omega_r$ such that

$$\Phi(\bar{x},\bar{\xi},\bar{y},\bar{\zeta}) = \sup\Phi \ge \frac{\delta}{2} > 0.$$

Now, as usual, we take a look at the inequation

$$\Phi(\bar{x},\bar{\xi},\bar{x},\bar{\xi}) + \Phi(\bar{y},\bar{\zeta},\bar{y},\bar{\zeta}) \le 2\Phi(\bar{x},\bar{\xi},\bar{y},\bar{\zeta}),$$

which yields

$$u_1(\bar{x},\bar{\xi}) - u_2(\bar{x},\bar{\xi}) - 2\mu\bar{x}^2 + u_1(\bar{y},\bar{\zeta}) - u_2(\bar{y},\bar{\zeta}) - 2\mu\bar{y}^2$$

$$\leq 2u_1(\bar{x},\bar{\xi}) - 2u_2(\bar{y},\bar{\zeta}) - \frac{(\bar{x}-\bar{y})^2}{\varepsilon} - \frac{(\bar{\xi}-\bar{\zeta})^2}{\varepsilon} - 2\mu(\bar{x}^2+\bar{y}^2).$$

Exploiting the Lipschitz continuity of u_1, u_2 and Young's inequality, this implies

$$\begin{aligned}\frac{(\bar{x}-\bar{y})^2}{\varepsilon} + \frac{(\bar{\xi}-\bar{\zeta})^2}{\varepsilon} &\leq u_1(\bar{x},\bar{\xi}) - u_1(\bar{y},\bar{\zeta}) + u_2(\bar{x},\bar{\xi}) - u_2(\bar{y},\bar{\zeta}) \\ &\leq C\left(|\bar{x}-\bar{y}| + \left|\bar{\xi}-\bar{\zeta}\right|\right) \\ &\leq C^2\varepsilon + \frac{(\bar{x}-\bar{y})^2}{2\varepsilon} + \frac{(\bar{\xi}-\bar{\zeta})^2}{2\varepsilon},\end{aligned}$$

so that there exists $c > 0$ independent of ε, μ such that

$$\frac{(\bar{x}-\bar{y})^2}{\varepsilon^2} + \frac{(\bar{\xi}-\bar{\zeta})^2}{\varepsilon^2} \leq c. \tag{10.9}$$

Next, define the functions

$$\begin{aligned}\varphi_1(x,\xi) :=& u_2(\bar{y},\bar{\zeta}) + \frac{(x-\bar{y})^2}{2\varepsilon} + \frac{(\xi-\bar{\zeta})^2}{2\varepsilon} + \mu(x^2+\bar{y}^2), \\ \varphi_2(y,\zeta) :=& u_1(\bar{x},\bar{\xi}) - \frac{(\bar{x}-y)^2}{2\varepsilon} - \frac{(\bar{\xi}-\zeta)^2}{2\varepsilon} - \mu(\bar{x}^2+y^2).\end{aligned}$$

By definition, $u_1 - \varphi_1$ has a (local) maximum at $(\bar{x},\bar{\xi})$ and $u_2 - \varphi_2$ attains a (local) minimum at $(\bar{y},\bar{\zeta})$. Since u_1 is a subsolution and u_2 a supersolution in the sense of theorem 10.3, this implies that there are $p \in N_Z(\bar{x}-\bar{\xi}) \cap [-M_f, M_f]$, $q \in N_Z(\bar{y}-\bar{\zeta}) \cap [-M_f, M_f]$, such that

$$\begin{aligned}&\lambda u_1(\bar{x},\bar{\xi}) - \left(\frac{\bar{\xi}-\bar{\zeta}}{\varepsilon} - \frac{\bar{x}-\bar{y}}{\varepsilon} - 2\mu\bar{x}\right)p \\ &\quad + \sup_{a\in\mathbb{A}}\left\{\left(-\frac{\bar{x}-\bar{y}}{\varepsilon} - 2\mu\bar{x}\right)f(\bar{x},\bar{\xi},a) - l(\bar{x},\bar{\xi},a)\right\} \leq 0 \\ &\leq \lambda u_2(\bar{y},\bar{\zeta}) - \left(\frac{\bar{\xi}-\bar{\zeta}}{\varepsilon} - \frac{\bar{x}-\bar{\xi}}{\varepsilon} + 2\mu\bar{y}\right)q \\ &\quad + \sup_{a\in\mathbb{A}}\left\{\left(-\frac{\bar{x}-\bar{y}}{\varepsilon} + 2\mu\bar{y}\right)f(\bar{y},\bar{\zeta},a) - l(\bar{y},\bar{\zeta},a)\right\}.\end{aligned}$$

From this inequality, we infer that

$$\begin{aligned}&\lambda\left(u_1(\bar{x},\bar{\xi}) - u_2(\bar{y},\bar{\zeta})\right) \\ &\quad \leq \frac{1}{\varepsilon}\left(\bar{\xi}-\bar{\zeta}-\bar{x}+\bar{y}\right)(p-q) &(10.10)\\ &\quad\quad - 2\mu(\bar{x}p+\bar{y}q) &(10.11)\end{aligned}$$

$$+ 2\mu(|\bar{x}| + |\bar{y}|)M_f \tag{10.12}$$

$$+ c_l \left(|\bar{x} - \bar{y}| + \left|\bar{\xi} - \bar{\zeta}\right|\right) + \frac{|\bar{x} - \bar{y}|}{\varepsilon} \left|f(\bar{x}, \bar{\xi}, a) - f(\bar{y}, \bar{\zeta}, a)\right|. \tag{10.13}$$

We estimate (10.10) to (10.12). As $p \in N_Z(\bar{x} - \bar{\xi})$ and $q \in N_Z(\bar{y} - \bar{\zeta})$,

$$(10.10) = \left((\bar{y} - \bar{\zeta}) - (\bar{x} - \bar{\xi})\right)(p - q) \leq 0.$$

For (10.11), an application of Young's inequality implies

$$(10.11) \leq 2\mu(|\bar{x}| + |\bar{y}|)M \leq c\mu^{\frac{3}{2}} \left(\bar{x}^2 + \bar{y}^2\right) + cM^2\mu^{\frac{1}{2}},$$

for some constant $c > 0$. Analogously, we get the inequality

$$(10.12) \leq c\mu^{\frac{3}{2}}(\bar{x}^2 + \bar{y}^2) + c\mu^{\frac{1}{2}}.$$

To estimate (10.13), we make use of (10.9) and the Lipschitz property of f, to get

$$(10.13) \leq (c_l + cc_f)\left(|\bar{x} - \bar{y}| + \left|\bar{\xi} - \bar{\zeta}\right|\right) = \omega(\varepsilon),$$

where ω is some continuous nonnegative function with the property $\omega(\varepsilon) \to 0$ as $\varepsilon \downarrow 0$, uniformly w.r.t. μ. Plugging in those estimates, it follows that

$$u_1(\bar{x}, \bar{\xi}) - u_2(\bar{y}, \bar{\zeta}) \leq c\mu^{\frac{3}{2}}(\bar{x}^2 + \bar{y}^2) + c\mu^{\frac{1}{2}} + \omega(\varepsilon),$$

for some $c > 0$. But then, our assumption implies that

$$\begin{aligned} 0 < \frac{\delta}{2} &\leq \Phi(\bar{x}, \bar{\xi}, \bar{y}, \bar{\zeta}) \leq u_1(\bar{x}, \bar{\xi}) - u_2(\bar{y}, \bar{\zeta}) - \mu(\bar{x}^2 + \bar{y}^2) \\ &\leq (c\mu^{\frac{3}{2}} - \mu)(\bar{x}^2 + \bar{y}^2) + c\mu^{\frac{1}{2}} + \omega(\varepsilon). \end{aligned}$$

Thus, if we fix $\bar{\mu} \in (0, \tilde{\mu})$ such that $c\bar{\mu}^{\frac{3}{2}} - \bar{\mu} < 0$ and $c\bar{\mu}^{\frac{1}{2}} < \frac{\delta}{4}$, the latter implies

$$0 < \frac{\delta}{4} \leq \omega(\varepsilon),$$

which yields a contradiction when letting $\varepsilon \downarrow 0$. This concludes the proof.

□

Now, we can characterize the value function via the following existence and uniqueness result.

Theorem 10.5 *If $\lambda > 2c_f$, then the value function V is the unique Lipschitz continuous solution in the sense of theorem 10.3 of the corresponding differential inclusion.*

Proof: From theorem 10.3, we infer that V is a solution. By proposition 10.2, V is Lipschitz continuous on Ω_r if $\lambda > 2c_f$, so that we may apply theorem 10.4 with $u_i = V$. Hence, every subsolution is smaller than V, and every supersolution is larger than V; altogether, this shows that every solution is equal to the value function of the control problem.

□

11 Dynamic Programming for a problem corresponding to a quasilinear p.d.e. with Play type hysteresis

Here we want to exploit the insight into problems containing time derivatives of Play type hysteresis. To treat the other variable, we adapt the method of [7] to our problem. To this end, we will have to slightly change the differential equation treated in [11, chapter IX], to "produce a bit more compactness".

11.1 The differential equation and properties of solutions

We start with a special case of an equation considered in [11], [12]:

$$\begin{aligned}
&\dot{y} + \dot{w} - \Delta y = \alpha \quad \text{in } \Omega \times (0,T), \\
&w(\cdot, x) = \mathcal{F}_r[y(\cdot, x); w_0(x)], \quad \text{a.e. } x \in \Omega, \\
&y(t,x) = 0, \quad \text{in } \partial\Omega \times (0,T), \\
&y(0) = y_0,\ w(0) = w_0, \\
&w_0(x) \in [y_0(x) - r, y_0(x) + r],\ \text{a.e. } x \in \Omega,
\end{aligned} \tag{11.1}$$

where $T, r > 0$ and $\Omega \subset \mathbb{R}^n$ is some open, bounded domain with smooth boundary (at least C^2). We will assume further that $y_0, w_0 \in L^2(\Omega)$ and $\alpha \in L^2(\Omega \times (0,T))$. Existence of solutions to (11.1) was first shown by Visintin; proofs can be found (in slightly different presentation) in [11, chapter IX, theorem 1.1], [12, theorem 3.3.2].

Theorem 11.1 *For every $y_0 \in H_0^1(\Omega)$ there exists a weak solution to* (11.1) *in the sense that there are*

$$y \in Y := L^\infty(0,T; H_0^1(\Omega)) \cap H^1(0,T; L^2(\Omega)), \quad w \in H^1(0,T; L^2(\Omega)),$$

such that

$$\int_0^T \int_\Omega (\dot{y}(t,x) + \dot{w}(t,x))\varphi(t,x)dxdt + \int_0^T \int_\Omega \nabla y(t,x) \cdot \nabla\varphi(t,x)dxdt = \int_0^T \int_\Omega \alpha(t,x)\varphi(t,x)dxdt,$$

for all $\varphi \in L^2(0,T; H_0^1(\Omega))$.

Proof: This is a special case of the results in [11], [12]. We remark that the proof is quite similar to the one of the o.d.e. case, theorem 10.1, but with some additional difficulties due to the appearence of the unbounded differential operator.

□

The uniqueness result is again a direct consequence of Hilpert's inequality. But one gets even more, namely a stability result in $L^1(\Omega)$.

Theorem 11.2 *Any pair of weak solutions* $(y_1, w_1), (y_2, w_2)$ *to* (11.1) *corresponding to initial values* $(y_{0,1}, w_{0,1}), (y_{0,2}, w_{0,2})$ *and right hand sides* α_1, α_2 *satisfies*

$$\begin{aligned}\int_\Omega |y_1(t,x) - y_2(t,x)|\,dx + \int_\Omega |w_1(t,x) - w_2(t,x)|\,dx \\ \leq \int_\Omega |y_{0,1}(x) - y_{0,2}(x)|\,dx + \int_\Omega |w_{0,1}(x) - w_{0,2}(x)|\,dx \\ + \int_0^t \int_\Omega |\alpha_1(\tau,x) - \alpha_2(\tau,x)|\,dxd\tau.\end{aligned} \tag{11.2}$$

Proof: See [11, chapter IX, corollary 2.2] or [12, corollary 3.3.6].

□

Inequality (11.2) can also be seen as a continuity property of the solution operator. There is then a unique extension to initial values in $L^1(\Omega)$. For generalised solution concepts which are applicable in this case, we refer to the book of Visintin, [11]. What can be easier done is to derive some regularity properties of solutions. Let (y, w) denote the solution. Then, for $f := \alpha - \dot{w} \in L^2(\Omega_T)$, y is equal to the weak solution of the standard heat equation

$$\begin{aligned}\dot{y} - \Delta y = f \quad \text{in } \Omega \times (0,T), \\ y = 0, \in \partial\Omega \times (0,T).\end{aligned} \tag{11.3}$$

Then, lemma 3.7 implies $y \in C(0,T; H_0^1(\Omega)) \cap L^2(0,T; H^2(\Omega))$. We will make use of this improved regularity result later. Further, we note that the proof of theorem 11.1 (as stated in [12], [11]) shows that $\|y\|_Y$ is bounded by some constant depending only on $|\Omega|$, $\|\alpha\|_{L^2(\Omega)}$, so that, in particular, the improved regularity holds uniformly w.r.t. α chosen from some bounded subset of $L^2(\Omega_T)$.

11.2 The control problem and properties of the value function

As mentioned before, we will need some more compactness properties of the equation later. We will achieve this not by transforming the equation via $y \mapsto By$ with some compact operator $B : L^2(\Omega) \to L^2(\Omega)$ (as done in [7]), as the Play operator only acts in time and thus can "destroy" the regularizing effect. Instead, we restrict to more regular controls by "adding" the operator on the right hand side; i.e., we consider dynamics of the form

$$\begin{aligned} &\dot{y}(t) + \dot{w}(t) - Ay(t) = B\alpha(t), \\ &w = \mathcal{F}_r[y; w_0], \quad w_0 \in [y_0 - r, y_0 + r], \end{aligned} \tag{11.4}$$

where the application of the Play operator has to be understood pointwise in space, and A is selfadjoint and the generator of an analytic semigroup of contractions on $L^2(\Omega)$ ($\Omega \subset \mathbb{R}^n$ some open, bounded domain) such that $D(A)$ is dense in $L^2(\Omega)$. We assume that for every $T > 0$:

V(i): if $y_0 \in D((-A)^{\frac{1}{2}})$, there exists a pair of solutions (y, w) with regularity $(H^1(0,T; L^2(\Omega)) \cap L^2(0,T; D(A))) \times L^2(\Omega; H^1(0,T))$ such that (11.4) holds almost everywhere and $y \in C(0,T; L^2(\Omega))$, uniformly w.r.t. $\alpha \in B_R(0) \subset L^2(\Omega)$, for any (fixed) $R > 0$,

V(ii): inequality (11.2) is valid,

V(iii): an improved regularity result holds, i.e., there is $\beta > 0$ such that $y \in C(0,T; D((-A)^\beta))$ whenever $y_0 \in D(A)$ and uniformly w.r.t. $\alpha \in B_R(0) \subset L^2(\Omega)$, for any (fixed) $R > 0$.

Further, we assume that $(-A)^{1-\beta}B \in \mathcal{L}(L^2(\Omega))$. Moreover, the controls α are to take values in some bounded set $\mathbb{A} \subset L^2(\Omega)$, so that there exists a constant $M_{\mathbb{A}} > 0$ such that $\|Ba\| \le M_{\mathbb{A}}$ for every $a \in \mathbb{A}$ ($\|\cdot\| := \|\cdot\|_{L^2(\Omega)}$, $\langle \cdot, \cdot \rangle := \langle \cdot, \cdot \rangle_{L^2(\Omega)}$). We will refer to these assumptions as **(GA)**.

Next, we introduce the functional we want to minimize. Again, this will be of infinite horizon type. Let $l : D_l := (L^1(\Omega))^2 \times L^2(\Omega) \to \mathbb{R}$ be such that

l(i): $\exists M_l > 0 \; \forall (y, w, a)^T \in D_l$:

$$|l(y, w, a)| \le M_l,$$

l(ii): $\exists C_l > 0 \; \forall (y, w, a)^T, (x, v, a)^T \in D_l$:

$$|l(y, w, a) - l(x, v, a)| \le C_l \left(\|y - x\|_{L^1(\Omega)} + \|w - v\|_{L^1(\Omega)} \right).$$

We will refer to those assumptions as **(lA)**. Then, as usual, for $\lambda > 0$, the cost functional is defined through

$$J(y_0, w_0, \alpha) := \int_0^\infty e^{-\lambda t} l(y_{y_0,w_0,\alpha}(t), w_{y_0,w_0,\alpha}(t), \alpha(t)) dt,$$

and the corresponding value function by

$$V(y_0, w_0) := \inf_{\alpha \in \mathcal{A}} J(y_0, w_0, \alpha).$$

Since Ω is assumed to be bounded, $L^2(\Omega) \subset L^1(\Omega)$. So, inequality (11.2) implies that V is well defined when considered as mapping

$$D_V := \left\{ (y, w) \in (L^2(\Omega))^2 \mid w \in [y - r, y + r] \right\} \to \mathbb{R}.$$

We will always consider this case in what follows. One reason for this is the following regularity result.

Proposition 11.3 *The value function is bounded and Lipschitz continuous, uniformly w.r.t. $\alpha \in \mathcal{A}$; i.e., there exist constants $M_V > 0$ and $C_V > 0$ such that $|V| \leq M_V$ and for every $(y, w), (x, v) \in D_V$,*

$$|V(y, w) - V(x, v)| \leq C_V (\|y - x\| + \|w - v\|).$$

Proof: The boundedness of l directly implies $|V| \leq M_l$. Next, note that

$$\begin{aligned}
|V(y, w) - V(x, v)| \leq & \sup_{\alpha \in \mathcal{A}} |J(y, w, \alpha) - J(x, v, \alpha)| \\
\leq & C_l \int_0^\infty e^{-\lambda t} \left(\|y(t) - x(t)\|_{L^1(\Omega)} + \|w(t) - v(t)\|_{L^1(\Omega)} \right) dt \\
\leq & C_l \left(\|y - x\|_{L^1(\Omega)} + \|w - v\|_{L^1(\Omega)} \right),
\end{aligned}$$

where we used (11.2) in the last step. Hence, if c is an imbedding constant for $L^2(\Omega) \hookrightarrow L^1(\Omega)$, then

$$|V(y, w) - V(x, v)| \leq C_l c (\|y - x\| + \|w - v\|),$$

which completes the proof.

□

As usual, we need the dynamic programming principle.

Lemma 11.4 (DPP) *For every $t > 0$ and $(x, \xi) \in D_V$, it holds*

$$V(x, \xi) = \inf_{\alpha \in \mathcal{A}} \left\{ \int_0^t e^{-\lambda s} l(y_{x,\xi,\alpha}(s), w_{x,\xi,\alpha}(s), \alpha(s)) ds + e^{-\lambda t} V(y_{x,\xi,\alpha}(t), w_{x,\xi,\alpha}(t)) \right\}.$$

Proof: This is an immediate consequence of the semigroup properties of $y(\cdot)$ and $w(\cdot)$.

□

11.3 The HJB equation and existence of solutions

The idea is now to use the method of [7] to handle the y variable, and to integrate the ideas from section 10 to handle the w variable.

Definition 11.5 *For $\delta > 0$ and functions V, φ, ψ, let* $\boldsymbol{M_\delta^+(V, \varphi, \psi)}$ *denote the set consisting of all $(x, \xi) \in \tilde{D}_V := (D((-A)^{\frac{1}{2}}) \times L^2(\Omega)) \cap D_V$ such that*

$$V(x,\xi) - \varphi(x) - \psi(\xi) - \frac{\delta}{2}\left\|(-A)^{\frac{1}{2}}x\right\|^2 \geq V(z,\zeta) - \varphi(z) - \psi(\zeta) - \frac{\delta}{2}\left\|(-A)^{\frac{1}{2}}z\right\|^2,$$

for all $(z, \zeta) \in \tilde{D}_V$. Similarly, let $\boldsymbol{M_\delta^-(V, \varphi, \psi)}$ *denote the set consisting of all $(x, \xi) \in \tilde{D}_V$ such that*

$$V(x,\xi) - \varphi(x) - \psi(\xi) + \frac{\delta}{2}\left\|(-A)^{\frac{1}{2}}x\right\|^2 \leq V(z,\zeta) - \varphi(z) - \psi(\zeta) + \frac{\delta}{2}\left\|(-A)^{\frac{1}{2}}z\right\|^2,$$

for all $(z, \zeta) \in \tilde{D}_V$.

□

Next, we introduce a notion of solution for the formal differential inclusion

$$\begin{aligned}\lambda V(x,\xi) - \langle D_x V(x,\xi), Ax\rangle + \langle D_x V(x,\xi) - D_\xi V(x,\xi), N_Z(x-\xi)\rangle \\ + \sup_{a\in\mathbb{A}}\{-\langle D_x V(x,\xi), Ba\rangle - l(x,\xi,a)\} \ni 0.\end{aligned} \tag{11.5}$$

Definition 11.6 *A bounded continuous function $V : D_V \to \mathbb{R}$ is called* ***viscosity subsolution*** *of* (11.5)*, if for all $\varphi, \psi \in C^1(L^2(\Omega))$,*

- *$M_\delta^+(V, \varphi, \psi) \subset \hat{D}_V := (D(A) \times L^2(\Omega)) \cap D_V$, for all $\delta > 0$,*
- *for every $(x, \xi) \in M_\delta^+(V, \varphi, \psi)$, there is $p \in N_Z(x - \xi) \cap L^2(\Omega)$ such that*

$$\begin{aligned}\lambda V(x,\xi) - \langle D\varphi(x), Ax\rangle + \langle D\varphi(x) - D\psi(\xi), p\rangle + \frac{\delta}{4}\|Ax\|^2 \\ - \frac{\delta}{4}M_{\mathbb{A}} + \sup_{a\in\mathbb{A}}\{-\langle D\varphi(x), Ba\rangle - l(x,\xi,a) + \delta\langle Ax, Ba\rangle\} \leq 0.\end{aligned}$$

Further, a bounded continuous function $V : D_V \to \mathbb{R}$ is called ***viscosity supersolution*** *of* (11.5)*, if for all $\varphi, \psi \in C^1(L^2(\Omega))$,*

- $M_\delta^-(V,\varphi,\psi) \subset \hat{D}_V$, *for all* $\delta > 0$,
- *for every* $(x,\xi) \in M_\delta^-(V,\varphi,\psi)$, *there is* $q \in N_Z(x-\xi) \cap L^2(\Omega)$ *such that*

$$\lambda V(x,\xi) - \langle D\varphi(x), Ax\rangle + \langle D\varphi(x) - D\psi(\xi), p\rangle - \frac{\delta}{4}\|Ax\|^2 + \frac{\delta}{4}M_{\mathbb{A}} + \sup_{a\in\mathbb{A}}\{-\langle D\varphi(x), Ba\rangle - l(x,\xi,a) - \delta\langle Ax, Ba\rangle\} \geq 0.$$

Finally, a function is called ***viscosity solution*** *of* (11.5)*, if it is both viscosity sub- and supersolution.*

□

We remark that by restriction, the function $\varphi + \psi$ may be seen as an element of $C^1(D_V)$, and that the corresponding derivatives coincide with the ones, when considered as mapping $\tilde{D}_V \to \mathbb{R}$ or $\hat{D}_V \to \mathbb{R}$, as the corresponding function spaces are linked by continuous imbeddings.

Theorem 11.7 *If (GA) and (lA) hold, then the value function is a viscosity solution of* (11.5)*.*

Proof: Let us start with proving that the value function V is a viscosity subsolution. To this end, let $\varphi, \psi \in C^1(L^2(\Omega))$, $\delta > 0$ and $(x,\xi) \in M_\delta^+(V,\varphi,\psi)$. Then, for any constant control $\alpha \equiv a \in \mathbb{A}$, denoting for short $y(t) := y_{x,\xi,a}(t)$, $w(t) := w_{x,\xi,a}(t)$,

$$V(x,\xi) - \varphi(x) - \psi(\xi) - \frac{\delta}{2}\left\|(-A)^{\frac{1}{2}}x\right\|^2 \geq V(y(t),w(t)) - \varphi(y(t)) - \psi(w(t)) - \frac{\delta}{2}\left\|(-A)^{\frac{1}{2}}y(t)\right\|^2,$$

for all $t > 0$. From the DPP (lemma 11.4), we then infer that

$$\begin{aligned}&\varphi(x) - \varphi(y(t)) + \psi(\xi) - \psi(w(t)) + \frac{\delta}{2}\left(\left\|(-A)^{\frac{1}{2}}x\right\|^2 - \left\|(-A)^{\frac{1}{2}}y(t)\right\|^2\right)\\ &\leq V(x,\xi) - V(y(t),w(t))\\ &\leq \int_0^t e^{-\lambda s} l(y(s),w(s),a)ds + (e^{-\lambda t} - 1)V(y(t),w(t)).\end{aligned}$$

Using the regularity assumption V(i), we may write

$$\varphi(x) - \varphi(y(t)) = -\int_0^t \langle D\varphi(y(s)), Ay(s) + Ba - \dot{w}(s)\rangle\, ds,$$

$$\psi(\xi) - \psi(w(t)) = -\int_0^t \langle D\psi(w(s)), \dot{w}(s)\rangle \, ds,$$
$$\frac{1}{2}\left(\left\|(-A)^{\frac{1}{2}}x\right\|^2 - \left\|(-A)^{\frac{1}{2}}y(t)\right\|^2\right) = -\int_0^t \langle -Ay(s), Ay(s) + Ba - \dot{w}(s)\rangle \, ds$$
$$= \int_0^t \|Ay(s)\|^2 \, ds + \int_0^t \langle Ay(s), Ba - \dot{w}(s)\rangle \, ds,$$

where $\dot{w}(s) \in N_Z(y(s) - w(s))$ for all s, by definition of the play operator. Plugging these formulas into the above inequality and dividing by t yields

$$\begin{aligned} & -\frac{1}{t}\int_0^t \langle D\varphi(y(s)), Ay(s) + Ba - \dot{w}(s)\rangle \, ds \\ & -\frac{1}{t}\int_0^t \langle D\psi(w(s)), \dot{w}(s)\rangle \, ds \\ & +\frac{\delta}{t}\int_0^t \|Ay(s)\|^2 \, ds + \frac{\delta}{t}\int_0^t \langle Ay(s), Ba - \dot{w}(s)\rangle \, ds \\ & \leq \frac{1}{t}\int_0^t e^{-\lambda s} l(y(s), w(s), a) ds + \frac{e^{-\lambda t} - 1}{t} V(y(t), w(t)). \end{aligned} \tag{11.6}$$

Next note that $\dot{w} = \dot{y}$ at almost every point in Ω_t where $\dot{w} \neq 0$. Hence, (11.4) implies that $|\dot{w}| \leq \frac{1}{2}|Ay + Ba|$ pointwise almost everywhere. We may therefore estimate $\|\dot{w}(s)\| \leq \frac{1}{2}\|Ay(s) + Ba\|$, for a.e. $s \in (0, t)$. Thus, we may further estimate the terms of (11.6) to arrive at

$$\begin{aligned} & \frac{\delta}{t}\int_0^t \|Ay(s)\|^2 \, ds \\ & \leq \frac{1}{t}\int_0^t \|D\varphi(y(s))\| \left(\|Ay(s)\| + \|Ba\| + \frac{1}{2}\|Ay(s)\| + \frac{1}{2}\|Ba\|\right) ds \\ & + \frac{1}{t}\int_0^t \|D\psi(w(s))\| \cdot \frac{1}{2}\left(\|Ay(s)\| + \|Ba\|\right) ds \\ & + \frac{\delta}{t}\int_0^t \|Ay(s)\| \left(\|Ba\| + \frac{1}{2}\left(\|Ay(s)\| + \|Ba\|\right)\right) ds \\ & + \frac{1}{t}\int_0^t e^{-\lambda s} l(y(s), w(s), a) ds + \frac{e^{-\lambda t} - 1}{t} V(y(t), w(t)). \end{aligned} \tag{11.7}$$

By continuity of $D\varphi$, ψ, y, w, l and V, the elements containing only those functions are uniformly bounded w.r.t. small $t > 0$. Hence, for those t, there exists a constant C_δ (depending only on δ and $T > 0$; i.e., we restrict (w.l.o.g.) to $t \in (0, T)$), such that

$$\frac{1}{t}\int_0^t \|Ay(s)\|^2 \, ds \leq C_\delta; \tag{11.8}$$

this can be shown by several applications of Young's inequality to the terms on the right hand side of (11.7) which contain the factor $\|Ay(s)\|$. Now we follow the arguments of [7] to show that (11.8) implies $x \in D(A)$: First of all, (11.8) implies that there exists a sequence of time points $(t_n)_{n\in\mathbb{N}}$ with $t_n \downarrow 0$ as $n \to \infty$, such that $\|Ay(t_n)\| \leq C_\delta$ for all n. Then, for some suitable subsequence, we have $Ay(t_n) \to z$ weakly and $y(t_n) \to x$ strongly. As $y(t_n) = (-A)^{-1}(-A)y(t_n) \to (-A)^{-1}(-z)$, this implies $z = Ax$, so that $x \in D(A)$. In particular, this proves $M_\delta^+(V, \varphi, \psi) \subset \hat{D}_V$. Our goal is now to take the limit $t \downarrow 0$ in (11.6). To this end, we will first replace, where possible, the time dependent elements by suitable constant ones, and then take the limit in the rewritten inequality.

- for $\frac{1}{t}\int_0^t \langle D\varphi(y(s)), Ay(s)\rangle\, ds$:

$$\begin{aligned} m(t) :=& \frac{1}{t}\int_0^t \langle D\varphi(y(s)) - D\varphi(x), Ay(s)\rangle\, ds \\ \leq& \frac{1}{t}\int_0^t \|D\varphi(y(s)) - D\varphi(x)\|\, \|Ay(s)\|\, ds \\ \leq& \left(\max_{s\in[0,t]} \|D\varphi(y(s)) - D\varphi(x)\|\right) \cdot \frac{1}{t}\int_0^t \|Ay(s)\|\, ds. \end{aligned}$$

Defining $\mu(t) := \max_{s\in[0,t]} \|D\varphi(y(s)) - D\varphi(x)\|$, continuity implies $\mu(t) \downarrow 0$ as $t \downarrow 0$, and from Young's and Jensen's inequality, we infer that

$$\begin{aligned} |m(t)| &\leq \frac{1}{2}\frac{\mu(t)^2}{\mu(t)} + \frac{1}{2}\mu(t)\left(\frac{1}{t}\int_0^t \|Ay(s)\|\, ds\right)^2 \\ &\leq \frac{1}{2}\mu(t) + \frac{1}{2}\mu(t)\left(\frac{1}{t}\int_0^t \|Ay(s)\|^2\, ds\right), \end{aligned}$$

which converges to zero as $t \downarrow 0$ due to (11.8). Hence, introducing an error function $\omega(t) = |m(t)|$, we may replace the considered element by

$$\frac{1}{t}\int_0^t \langle D\varphi(x), Ay(s)\rangle\, ds = \left\langle D\varphi(x), \frac{1}{t}\int_0^t Ay(s)ds \right\rangle,$$

with only an error $\omega(t)$, converging to zero as $t \downarrow 0$.

- we claim that for any sequence $(t_n)_{n\in\mathbb{N}}$ with $t_n \downarrow 0$ as $n \to \infty$, it holds

$$\frac{1}{t_n}\int_0^{t_n} Ay(s)ds \rightharpoonup Ax. \tag{11.9}$$

Proof of the claim: As

$$\left\| \frac{1}{t_n} \int_0^{t_n} Ay(s)ds \right\|^2 \leq \frac{1}{t_n} \int_0^{t_n} \|Ay(s)\|^2 \, ds \leq C_\delta,$$

we can extract a weakly convergent subsequence (which will for simplicity again be indexed with "n"). Now take any $\Phi \in D(A)$. Then, continuity of y implies

$$\left\langle \Phi, \frac{1}{t_n} \int_0^{t_n} Ay(s)ds \right\rangle = \frac{1}{t_n} \int_0^{t_n} \langle A\Phi, y(s) \rangle \, ds \to \langle A\Phi, x \rangle = \langle \Phi, Ax \rangle \, .$$

Since $D(A)$ is dense in $L^2(\Omega)$, this uniquely determines the weak limit. Hence, every weakly convergent subsequence has the limit Ax, which implies the assertion.

• (11.9) implies

$$\frac{1}{t_n} \int_0^{t_n} \langle D\varphi(x), Ay(s) \rangle \, ds \to \langle D\varphi(x), Ax \rangle \, ,$$

for any sequence of time points converging to zero from above.

• by continuity, we have

$$\frac{1}{t} \int_0^{t} \langle D\varphi(y(s)), Ba \rangle \, ds \to \langle D\varphi(x), Ba \rangle \, ,$$

as $t \downarrow 0$.

• for $\frac{1}{t} \int_0^t \langle D\varphi(y(s)), \dot{w}(s) \rangle \, ds$: Consider first

$$\begin{aligned} \hat{m}(t) &:= \frac{1}{t} \int_0^{t} \langle D\varphi(y(s)) - D\varphi(x), \dot{w}(s) \rangle \, ds \leq \mu(t) \cdot \frac{1}{t} \int_0^t \|\dot{w}(s)\| \, ds \\ &\leq \frac{1}{2}\mu(t) \, \|Ba\| + \frac{1}{2}\mu(t) \frac{1}{t} \int_0^t \|Ay(s)\| \, ds. \end{aligned}$$

Hence, (11.8) implies that $\hat{m} \to 0$ as $t \downarrow 0$. We may therefore replace the term by

$$\frac{1}{t} \int_0^{t} \langle D\varphi(x), \dot{w}(s) \rangle \, ds = \left\langle D\varphi(x), \frac{1}{t} \int_0^t \dot{w}(s)ds \right\rangle,$$

with only making some error $\omega(t)$.

• we have pointwise (a.e.), that $\dot{w}(s) \in N_Z(y(s) - w(s))$ and

$$\left\| \frac{1}{t_n} \int_0^{t_n} \dot{w}(s)ds \right\|^2 \leq \frac{1}{t_n} \int_0^{t_n} \|\dot{w}\|^2 \, ds \leq C_\delta,$$

for any positive sequence (t_n) with $t_n \downarrow 0$. Hence, we can extract a weakly convergent subsequence; w.l.o.g.,

$$p_n := \frac{1}{t_n} \int_0^{t_n} \dot{w}(s) ds \rightharpoonup p.$$

We claim that $p \in N_Z(x - \xi)$. As in the proof of theorem 6.7, we use Egoroff's theorem; w.l.o.g. (we may switch to a subsequence if necessary), we can assume that $\Delta_n := y(t_n) - w(t_n)$ converges almost uniformly to $\Delta := x - \xi$, i.e., for every $\varepsilon > 0$ there exists a set $S_\varepsilon \subset \Omega$ such that $|\Omega \setminus S_\varepsilon| < \varepsilon$ and $\Delta_n \to \Delta$ uniformly on S_ε. Hence, there exists $N \in \mathbb{N}$ such that for each $b \in \Omega$, we have either $p_n(b) \geq 0$ for all $n \geq N$ or $p_n(b) \leq 0$ for all $n \geq N$. As weak convergence is equivalent to convergence in the mean, this property is preserved when taking the weak limit, so that the restriction $p|_{S_\varepsilon}$ must be an element of $N_Z(x|_{S_\varepsilon} - \xi|_{S_\varepsilon})$. As $\varepsilon > 0$ was arbitrary, this property must hold almost everywhere, so that we can conclude $p \in N_Z(x - \xi)$.

- exploiting the last two points, we find that for some sequence $t_n \downarrow 0$, with some error $\omega(\frac{1}{n})$,

$$\frac{1}{t_n} \int_0^{t_n} \langle D\varphi(y(s)), \dot{w}(s) \rangle \, ds = \langle D\varphi(x), p \rangle + \omega(\frac{1}{n}),$$

where $p \in N_Z(x - \xi)$ and $\|p\| \leq 2(C_\delta + M_{\mathbb{A}})$, due to the bound on $\dot{w}(s)$.

- analogously, we get

$$\frac{1}{t_n} \int_0^{t_n} \langle D\psi(w(s)), \dot{w}(s) \rangle \, ds = \langle D\psi(\xi), p \rangle + \omega(\frac{1}{n}),$$

for some suitable sequence $(t_n)_n$.

- since it is sort of convex combination, for every t_n there exists s_n such that

$$\frac{1}{t_n} \int_0^{t_n} \|Ay(s)\|^2 \, ds \geq \|Ay(s_n)\|^2 .$$

As we may assume that $Ay(s_n) \rightharpoonup Ax$ and since the norm is weakly lower semicontinuous,

$$\liminf_{n \to \infty} \|Ay(s_n)\|^2 \geq \|Ax\|^2 .$$

- From V(iii) and the compactness of B, we get

$$\frac{1}{t} \int_0^t \langle Ay(s), Ba \rangle \, ds = -\frac{1}{t} \int_0^t \left\langle (-A)^\beta y(s), (-A)^{1-\beta} Ba \right\rangle ds$$

$$= -\left\langle \frac{1}{t}\int_0^t (-A)^\beta y(s)ds, (-A)^{1-\beta}Ba \right\rangle.$$

As $(-A)^\beta y(s)$ is continuous by assumption, we may replace the term with

$$\langle Ax, Ba \rangle,$$

with just making some error $\omega(t)$.

- for $\nu_n := \frac{1}{t_n}\int_0^{t_n} \langle Ay(s), \dot{w}(s)\rangle$, we use our estimate on $\|\dot{w}(s)\|$ to get

$$|\nu_n| \leq \frac{1}{2}\frac{1}{t_n}\int_0^{t_n} \|Ay(s)\|^2\, ds + \frac{1}{2}M_{\mathbb{A}}\frac{1}{t_n}\int_0^{t_n} \|Ay(s)\|\, ds.$$

Applying Young's inequality to the second term on the right hand side, we get, due to Jensen's inequality,

$$|\nu_n| \leq \frac{3}{4}\frac{1}{t_n}\int_0^{t_n} \|Ay(s)\|^2\, ds + \frac{1}{4}M_{\mathbb{A}}^2.$$

We can thus charge the terms in equation (11.6),

$$\frac{\delta}{t_n}\int_0^{t_n} \|Ay(s)\|^2\, ds - \delta\nu_n \geq \frac{\delta}{4t_n}\int_0^{t_n} \|Ay(s)\|^2\, ds - \frac{\delta}{4}M_{\mathbb{A}}^2.$$

- by continuity, we get

$$\frac{1}{t_n}\int_0^{t_n} e^{-\lambda s} l(y(s), w(s), a)ds = l(x, \xi, a) + \omega(\frac{1}{n}),$$

and

$$\frac{e^{-\lambda t} - 1}{t} V(y(t), w(t)) = -\lambda V(x, \xi) + \omega(t).$$

Altogether, restricting to some suitable subsequence, we may take the lim inf of (11.6), which yields

$$\lambda V(x,\xi) - \langle D\varphi(x), Ax\rangle + \langle D\varphi(x) - D\psi(\xi), p\rangle$$
$$+ \frac{\delta}{4}\|Ax\|^2 - \frac{\delta}{4}M_{\mathbb{A}} - \langle D\varphi(x), Ba\rangle - l(x,\xi,a) + \delta\langle Ax, Ba\rangle \leq 0.$$

Taking $\sup_{a\in\mathbb{A}}$ then yields the desired inequality. To prove that V is also a supersolution, we use almost the same arguments. First, let $(x,\xi) \in M_\delta^-(V,\varphi,\psi)$, so that, in particular,

$$V(x,\xi) - \varphi(x) - \psi(\xi) + \frac{\delta}{2}\left\|(-A)^{\frac{1}{2}}x\right\|^2$$

$$\leq V(y_{x,\xi,\alpha}(t), w_{x,\xi,\alpha}(t)) - \varphi(y_{x,\xi,\alpha}(t)) - \psi(w_{x,\xi,\alpha}(t)) + \frac{\delta}{2}\left\|(-A)^{\frac{1}{2}}y_{x,\xi,\alpha}(t)\right\|^2,$$

for all $t > 0$ and $\alpha \in \mathcal{A}$. From the DPP, we infer that for every $\varepsilon > 0$ and $t > 0$, there exists some control $\alpha_\varepsilon \in \mathcal{A}$, such that

$$\begin{aligned}
&\varphi(x) - \varphi(y(t)) + \psi(\xi) - \psi(w(t)) + \frac{\delta}{2}\left(\left\|(-A)^{\frac{1}{2}}y(t)\right\|^2 - \left\|(-A)^{\frac{1}{2}}x\right\|^2\right) \\
&\qquad \geq V(x,\xi) - V(y(t), w(t)) \\
&\qquad \geq \int_0^t e^{-\lambda s} l(y(s), w(s), \alpha_\varepsilon(s))ds + (e^{-\lambda t} - 1)V(y(t), w(t)) - \varepsilon t.
\end{aligned}$$

Again, we use the integral representations for $\varphi(x) - \varphi(y(t))$, $\psi(\xi) - \psi(w(t))$ and $\left\|(-A)^{\frac{1}{2}}y(t)\right\|^2 - \left\|(-A)^{\frac{1}{2}}x\right\|^2$; then, dividing by $t > 0$, we get

$$\begin{aligned}
&-\frac{1}{t}\int_0^t \langle D\varphi(y(s)), Ay(s) + B\alpha_\varepsilon(s) - \dot{w}(s)\rangle\, ds \\
&-\frac{1}{t}\int_0^t \langle D\psi(w(s)), \dot{w}(s)\rangle\, ds \\
&-\frac{\delta}{t}\int_0^t \|Ay(s)\|^2\, ds - \frac{\delta}{t}\int_0^t \langle Ay(s), B\alpha_\varepsilon(s) - \dot{w}(s)\rangle\, ds \\
&\geq \frac{1}{t}\int_0^t e^{-\lambda s} l(y(s), w(s), \alpha_\varepsilon(s))ds + \frac{e^{-\lambda t} - 1}{t} V(y(t), w(t)) - \varepsilon.
\end{aligned} \tag{11.10}$$

Now, the uniform continuity of $y(\cdot), w(\cdot)$ on bounded time intervals w.r.t. $\alpha \in \mathcal{A}$ (cf. V(i)) implies that the same estimates as for the subsolution case yield some constant $C_\delta > 0$ such that

$$\frac{1}{t}\int_0^t \|Ay(s)\|^2\, ds \leq C_\delta,$$

and this again implies $x \in D(A)$. Further, the uniform continuity w.r.t. α allows us to establish the same estimates and convergence results as for the subsolution case; the only thing that is different now, is the treatment of the terms which contain α_ε. So, let us take a closer look at those terms. As

$$\begin{aligned}
&-\frac{1}{t}\int_0^t \langle D\varphi(y(s)), B\alpha_\varepsilon(s)\rangle - \frac{\delta}{t}\int_0^t \langle Ay(s), B\alpha_\varepsilon(s)\rangle\, ds \\
&-\frac{1}{t}\int_0^t e^{-\lambda s} l(y(s), w(s), \alpha_\varepsilon(s))ds \\
&= \frac{1}{t}\int_0^t -\langle D\varphi(x), B\alpha_\varepsilon(s)\rangle - \delta\,\langle Ax, B\alpha_\varepsilon(s)\rangle - l(x,\xi,\alpha_\varepsilon(s))ds + \omega(t)
\end{aligned}$$

$$\leq \sup_{a \in \mathbb{A}} \{ - \langle D\varphi(x), Ba \rangle - \delta \langle Ax, Ba \rangle - l(x, \xi, a) \} + \omega(t),$$

we may replace the critical terms in the inequality. Altogether, we get the desired inequality, when taking the lim inf of some suitable sequence in time, and the proof is complete.

□

Remark 11.8 *As mentioned in the proof of the last theorem, it holds*

$$\|p\|, \|q\| \leq 2(C_\delta + M_{\mathbb{A}}),$$

and C_δ depends only on δ, $D\varphi(x)$, $D\psi(\xi)$.

□

11.4 Comparison and uniqueness result

Before giving the comparison result, we recall that from [30], for any upper semi-continuous function $f : D \to \mathbb{R}$, if D is a so called RNP subset of some Banach space X, there exists for every $\kappa > 0$ some linear map T of rank one such that $\|T\|_{\mathcal{L}(X;\mathbb{R})} \leq \kappa$ and $f + T$ attains its supremum over D. As noted in [30, page 4], $D \subset X$ is an RNP set if it is a convex and weakly compact subset of X. Now, let X be some Hilbert space. Then X is reflexive, so that bounded sequences include weakly convergent subsequences. Taking into account Mazur's theorem, we get that every convex, bounded and closed subset of a Hilbert space is weakly compact. Hence, in that case, it suffices to show that D is convex, bounded and closed when we need to prove that D is an RNP subset of X.

Theorem 11.9 *Let u_1, u_2 be bounded and Lipschitz continuous functions $D_V \to \mathbb{R}$ such that u_1 is a viscosity subsolution and u_2 a viscosity supersolution in the sense of definition 11.6. Then $u_1 \leq u_2$.*

Proof: Consider the auxiliary function $\Phi : \tilde{D}_V^2 \to \mathbb{R}$,

$$\begin{aligned}\Phi(x, \xi, y, \zeta) :=& u_1(x, \xi) - u_2(y, \zeta) - \frac{\|x - y\|^2}{2\varepsilon} - \frac{\|\xi - \zeta\|^2}{2\varepsilon} \\ & - \frac{\delta}{2} \left(\left\| (-A)^{\frac{1}{2}} x \right\|^2 + \left\| (-A)^{\frac{1}{2}} y \right\|^2 \right).\end{aligned}$$

Assume for contradiction that there was $(\hat{x}, \hat{\xi}) \in D_V$ such that

$$u_1(\hat{x}, \hat{\xi}) - u_2(\hat{x}, \hat{\xi}) = \tau > 0.$$

Then, by density of $\tilde{D}_V$ in D_V and continuity of u_1, u_2, we can find $(\tilde{x},\tilde{\xi}) \in \tilde{D}_V$, such that

$$u_1(\tilde{x},\tilde{\xi}) - u_2(\tilde{x},\tilde{\xi}) \geq \frac{3}{4}\tau.$$

Thus, as

$$\Phi(\tilde{x},\tilde{\xi},\tilde{x},\tilde{\xi}) = u_1(\tilde{x},\tilde{\xi}) - u_2(\tilde{x},\tilde{\xi}) - \delta \left\|(-A)^{\frac{1}{2}}\tilde{x}\right\|^2,$$

for some $\tilde{\delta} > 0$ it holds

$$\sup_{\tilde{D}_V^2} \Phi \geq \Phi(\tilde{x},\tilde{\xi},\tilde{x},\tilde{\xi}) \geq \frac{\tau}{2} > 0,$$

for every $0 < \delta < \tilde{\delta}$. Now, for every (fixed) parameters δ, ε, the supremum might only be attained on some bounded subset of $\tilde{D}_V$: Since by assumption, there exists a constant $C > 0$ such that $|u_1 - u_2| \leq C$, $\Phi(x,\xi,y,\zeta) \leq 0$ whenever

$$\max\left\{\left\|(-A)^{\frac{1}{2}}x\right\|^2, \left\|(-A)^{\frac{1}{2}}y\right\|^2\right\} \geq \frac{C}{\delta}.$$

Hence, for some $R = R(\delta) > 0$, we may restrict the whole problem to the set

$$\mathbb{D} := \left(\tilde{D}_V \cap \left(\bar{B}_R(0; D((-A)^{\frac{1}{2}})) \times L^2(\Omega)\right)\right)^2,$$

where $\bar{B}_R(0; D((-A)^{\frac{1}{2}}))$ stands for the closed ball with radius R in $D((-A)^{\frac{1}{2}})$. We show that $\mathbb{D}$ is an RNP subset of the Hilbert space $(L^2(\Omega))^2$. Using the remark before theorem 11.9, it suffices to show that $\mathbb{D}$ is a bounded, closed, convex subset of $(L^2(\Omega))^2$.

• Boundedness: From the continuous imbedding $D((-A)^{\frac{1}{2}}) \hookrightarrow L^2(\Omega)$, we infer that, for some $c > 0$,

$$\|x\| \leq c\left\|(-A)^{\frac{1}{2}}x\right\| \leq cR,$$

and thus, since $\xi \in [x-r, x+r]$,

$$\|\xi\| \leq \|x\| + \|r\| \leq cR + \|r\|.$$

As analogous inequalities are valid for (y,ζ), boundedness follows.

• Closedness: Let $(x_n,\xi_n,y_n,\zeta_n)_n \subset \mathbb{D}$ such that $(x_n,\xi_n,y_n,\zeta_n) \to (x,\xi,y,\zeta)$ in $(L^2(\Omega))^4$. We only consider (x_n,ξ_n), as the argument is exactly the same for (y_n,ζ_n).

As $\left\|(-A)^{\frac{1}{2}}x_n\right\| \leq R$ for all $n \in \mathbb{N}$ by definition of $\mathbb{D}$, we can extract a weakly convergent subsequence $(x_{n_k})_k$; i.e., $x_{n_k} \rightharpoonup x$ in $D((-A)^{\frac{1}{2}})$ as $k \to \infty$. Thus, the weak lower semicontinuity of norms implies

$$R \geq \liminf_{k\to\infty} \left\|(-A)^{\frac{1}{2}}x_{n_k}\right\| \geq \left\|(-A)^{\frac{1}{2}}x\right\|.$$

Further, as we may extract a subsequence $(\xi_{n_k})_k$ which converges pointwise a.e. to ξ, the relation $\xi_{n_k} \in [x_{n_k} - r, x_{n_k} + r]$ for all k implies that $\xi \in [x - r, x + r]$ a.e..
• Convexity: Follows by convexity of balls and sets of the type $[x - r, x + r]$.

Now, as $\Phi : \tilde{D}_V^2 \to \mathbb{R}$ is continuous, the same holds for the restriction to $\mathbb{D}$. Hence, the main theorem of [30, page 7] implies that for every $\nu > 0$, there exists some linear functional $T : (L^2(\Omega))^4 \to \mathbb{R}$ of rank one, such that $\|T\|_{\mathcal{L}((L^2(\Omega))^4;\mathbb{R})} \leq \nu$ and $\Phi + T$ attains its supremum over $\mathbb{D}$. Moreover, from the boundedness of $\mathbb{D}$ (note that the bound depends on δ!) we infer the existence of $\tilde{\nu} = \tilde{\nu}(\delta)$, such that

$$\forall \nu \in (0, \tilde{\nu})\ \exists T_\nu \in (L^2(\Omega))^4 : \quad \sup_{\mathbb{D}} \{\Phi(\cdot) + \langle T_\nu, \cdot\rangle\} \geq \frac{\tau}{4} > 0, \tag{11.11}$$

and there are $(\bar{x}, \bar{\xi}, \bar{y}, \bar{\zeta}) \in \mathbb{D} : \sup_{\mathbb{D}} \{\Phi(\cdot) + \langle T_\nu, \cdot\rangle\} = (\Phi + T_\nu)(\bar{x}, \bar{\xi}, \bar{y}, \bar{\zeta})$. So let $T_\nu = (t_1, t_2, t_3, t_4)$ have such property; denoting $\bar{\Phi} := \Phi + T_\nu$, the usual inequality

$$\bar{\Phi}(\bar{x}, \bar{\xi}, \bar{x}, \bar{\xi}) + \bar{\Phi}(\bar{y}, \bar{\zeta}, \bar{y}, \bar{\zeta}) \leq 2\bar{\Phi}(\bar{x}, \bar{\xi}, \bar{y}, \bar{\zeta})$$

implies that

$$\begin{aligned}
&u_1(\bar{x}, \bar{\xi}) - u_2(\bar{x}, \bar{\xi}) - \delta \left\|(-A)^{\frac{1}{2}}\bar{x}\right\|^2 + \langle t_1 + t_3, \bar{x}\rangle + \left\langle t_2 + t_4, \bar{\xi}\right\rangle \\
&u_1(\bar{y}, \bar{\zeta}) - u_2(\bar{y}, \bar{\zeta}) - \delta \left\|(-A)^{\frac{1}{2}}\bar{y}\right\|^2 + \langle t_1 + t_3, \bar{y}\rangle + \left\langle t_2 + t_4, \bar{\zeta}\right\rangle \\
&\leq 2u_1(\bar{x}, \bar{\xi}) - 2u_2(\bar{y}, \bar{\zeta}) - \frac{\|\bar{x} - \bar{y}\|^2}{\varepsilon} - \frac{\left\|\bar{\xi} - \bar{\zeta}\right\|^2}{\varepsilon} \\
&- \delta \left(\left\|(-A)^{\frac{1}{2}}\bar{x}\right\|^2 + \left\|(-A)^{\frac{1}{2}}\bar{y}\right\|^2\right) + 2\langle t_1, \bar{x}\rangle + 2\left\langle t_2, \bar{\xi}\right\rangle + 2\langle t_3, \bar{y}\rangle + 2\left\langle t_4, \bar{\zeta}\right\rangle.
\end{aligned}$$

Rearranging yields

$$\begin{aligned}
\frac{\|\bar{x} - \bar{y}\|^2}{\varepsilon} + \frac{\left\|\bar{\xi} - \bar{\zeta}\right\|^2}{\varepsilon} \leq &u_1(\bar{x}, \bar{\xi}) - u_1(\bar{y}, \bar{\zeta}) + u_2(\bar{x}, \bar{\xi}) - u_2(\bar{y}, \bar{\zeta}) \\
&+ \langle t_1 - t_3, \bar{x} - \bar{y}\rangle + \left\langle t_2 - t_4, \bar{\xi} - \bar{\zeta}\right\rangle.
\end{aligned}$$

Note that we may, w.l.o.g., restrict to $\nu \leq 1$; but then, exploiting also the Lipschitz continuity of u_1, u_2, there exists some constant $\mathcal{C} > 0$ such that

$$\begin{aligned}\frac{\|\bar{x}-\bar{y}\|^2}{\varepsilon} + \frac{\left\|\bar{\xi}-\bar{\zeta}\right\|^2}{\varepsilon} &\leq \mathcal{C}\left(\|\bar{x}-\bar{y}\| + \left\|\bar{\xi}-\bar{\zeta}\right\|\right) \\ &\leq \mathcal{C}^2\varepsilon + \frac{\|\bar{x}-\bar{y}\|^2}{2\varepsilon} + \frac{\left\|\bar{\xi}-\bar{\zeta}\right\|^2}{2\varepsilon},\end{aligned}$$

where we used Young's inequality as second step. Hence,

$$\frac{\|\bar{x}-\bar{y}\|^2}{\varepsilon^2} + \frac{\left\|\bar{\xi}-\bar{\zeta}\right\|^2}{\varepsilon^2} \leq 2\mathcal{C}^2, \tag{11.12}$$

and the constant $\mathcal{C}$ is independent of ε, δ, ν. Next, consider the "test functions"

$$\begin{aligned}\varphi_1(x) &:= \frac{1}{2\varepsilon}\|x-\bar{y}\|^2 + \frac{\delta}{2}\left\|(-A)^{\frac{1}{2}}\bar{y}\right\|^2 - \langle t_1, x\rangle - \langle t_3, \bar{y}\rangle, \\ \psi_1(\xi) &:= u_2(\bar{y},\bar{\zeta}) + \frac{1}{2\varepsilon}\left\|\xi-\bar{\zeta}\right\|^2 - \langle t_2, \xi\rangle - \left\langle t_4, \bar{\zeta}\right\rangle, \\ \varphi_2(y) &:= -\frac{1}{2\varepsilon}\|\bar{x}-y\|^2 - \frac{\delta}{2}\left\|(-A)^{\frac{1}{2}}\bar{x}\right\|^2 + \langle t_1, \bar{x}\rangle + \langle t_3, y\rangle, \\ \psi_2(\zeta) &:= u_1(\bar{x},\bar{\xi}) - \frac{1}{2\varepsilon}\left\|\bar{\xi}-\zeta\right\|^2 + \left\langle t_2, \bar{\xi}\right\rangle + \langle t_4, \zeta\rangle.\end{aligned}$$

They are all continuously differentiable, and we claim that, if $\nu > 0$ is small enough, then

1. $(\bar{x},\bar{\xi}) \in M_\delta^+(u_1,\varphi_1,\psi_1)$,
2. $(\bar{y},\bar{\zeta}) \in M_\delta^-(u_2,\varphi_2,\psi_2)$.

Proof of the claim: For all $(x,\xi) \in \tilde{D}_V$, we may use $D((-A)^{\frac{1}{2}}) \hookrightarrow L^2(\Omega)$ and the properties $\|\xi\| \leq \|x\| + \|r\|$, $\left\|\bar{\zeta}\right\| \leq \|\bar{y}\| + \|r\|$ to estimate

$$\begin{aligned}&\Phi(x,\xi,\bar{y},\bar{\zeta}) + \langle t_1, x\rangle + \langle t_2, \xi\rangle + \langle t_3, \bar{y}\rangle + \left\langle t_4, \bar{\zeta}\right\rangle \\ &\leq \Phi(x,\xi,\bar{y},\bar{\zeta}) + c\|t_1\|\left\|(-A)^{\frac{1}{2}}x\right\| + \|t_2\|\left(c\left\|(-A)^{\frac{1}{2}}x\right\| + \|r\|\right) \\ &\quad + c\|t_3\|\left\|(-A)^{\frac{1}{2}}\bar{y}\right\| + \|t_4\|\left(c\left\|(-A)^{\frac{1}{2}}\bar{y}\right\| + \|r\|\right).\end{aligned}$$

Hence, if $\nu > 0$ (depending now on δ) is small enough, we can achieve that $\bar{\Phi} \leq \frac{\tau}{8}$ whenever $(x,\xi,\bar{y},\bar{\zeta}) \notin \mathbb{D}$, being smaller than the maximal value on $\mathbb{D}$; hence, $(\bar{x},\bar{\xi}) \in M_\delta^+(u_1,\varphi_1,\psi_1)$; the argument for $M_\delta^-(u_2,\varphi_2,\psi_2)$ is similar.

We can now use $\bar{x}, \bar{y} \in D(A)$ and that u_1, u_2 are, resp., sub- and supersolutions. Noting that

$$D\varphi_1(\bar{x}) = \frac{\bar{x}-\bar{y}}{\varepsilon} - t_1, \qquad D\psi_1(\bar{\xi}) = \frac{\bar{\xi}-\bar{\zeta}}{\varepsilon} - t_2,$$
$$D\varphi_2(\bar{y}) = \frac{\bar{x}-\bar{y}}{\varepsilon} + t_3, \qquad D\psi_2(\bar{\zeta}) = \frac{\bar{\xi}-\bar{\zeta}}{\varepsilon} + t_4,$$

we get that there are $p \in N_Z(\bar{x}-\bar{\xi})$, $q \in N_Z(\bar{y}-\bar{\zeta})$ with norms bounded by some constant depending on δ (cf., remark 11.8), such that

$$\begin{aligned}
&\lambda u_1(\bar{x},\bar{\xi}) - \left\langle \frac{\bar{x}-\bar{y}}{\varepsilon} - t_1, A\bar{x} \right\rangle + \left\langle \frac{\bar{x}-\bar{y}}{\varepsilon} - t_1 - \frac{\bar{\xi}-\bar{\zeta}}{\varepsilon} + t_2, p \right\rangle + \frac{\delta}{4}\|A\bar{x}\|^2 \\
&- \frac{\delta}{4}M_{\mathbb{A}} + \sup_{a\in\mathbb{A}} \left\{ -\left\langle \frac{\bar{x}-\bar{y}}{\varepsilon} - t_1, Ba \right\rangle - l(\bar{x},\bar{\xi},a) + \delta\langle A\bar{x}, Ba\rangle \right\} \le 0 \\
&\le \lambda u_2(\bar{y},\bar{\zeta}) - \left\langle \frac{\bar{x}-\bar{y}}{\varepsilon} + t_3, A\bar{y} \right\rangle + \left\langle \frac{\bar{x}-\bar{y}}{\varepsilon} + t_3 - \frac{\bar{\xi}-\bar{\zeta}}{\varepsilon} - t_4, q \right\rangle - \frac{\delta}{4}\|A\bar{y}\|^2 \\
&+ \frac{\delta}{4}M_{\mathbb{A}} + \sup_{a\in\mathbb{A}} \left\{ -\left\langle \frac{\bar{x}-\bar{y}}{\varepsilon} + t_3, Ba \right\rangle - l(\bar{y},\bar{\zeta},a) - \delta\langle A\bar{y}, Ba\rangle \right\}.
\end{aligned}$$

Rearranging the terms yields

$$\begin{aligned}
\lambda\big(u_1(\bar{x},\bar{\xi}) - u_2(\bar{y},\bar{\zeta})\big) \le{}& -\frac{\delta}{4}\|A\bar{x}\|^2 - \frac{\delta}{4}\|A\bar{y}\|^2 + \frac{\delta}{2}M_{\mathbb{A}} \\
&+ \frac{1}{\varepsilon}\langle \bar{x}-\bar{y}-\bar{\xi}+\bar{\zeta}, q-p\rangle && (11.13)\\
&+ \left\langle \frac{\bar{x}-\bar{y}}{\varepsilon}, A(\bar{x}-\bar{y}) \right\rangle && (11.14)\\
&- \langle t_1, A\bar{x}\rangle - \langle t_3, A\bar{y}\rangle + \langle t_1-t_2, p\rangle + \langle t_3-t_4, q\rangle && (11.15)\\
&+ \sup_{a\in\mathbb{A}} \left\{ -\left\langle \frac{\bar{x}-\bar{y}}{\varepsilon} + t_3, Ba \right\rangle - l(\bar{y},\bar{\zeta},a) - \delta\langle A\bar{y}, Ba\rangle \right\} && (11.16)\\
&- \sup_{a\in\mathbb{A}} \left\{ -\left\langle \frac{\bar{x}-\bar{y}}{\varepsilon} - t_1, Ba \right\rangle - l(\bar{x},\bar{\xi},a) + \delta\langle A\bar{x}, Ba\rangle \right\}. && (11.17)
\end{aligned}$$

We estimate (11.13) to (11.15) and $|(11.16) + (11.17)|$. As $p \in N_Z(\bar{x}-\bar{\xi})$ and $q \in N_Z(\bar{y}-\bar{\zeta})$, we have

$$(11.13) = \frac{1}{\varepsilon}\big\langle (\bar{x}-\bar{\xi}) - (\bar{y}-\bar{\zeta}), q-p \big\rangle \le 0.$$

Further,

$$(11.14) = -\frac{1}{\varepsilon}\left\| (-A)^{\frac{1}{2}}(\bar{x}-\bar{y}) \right\|^2 \le 0.$$

To estimate (11.15), note that $\|A\bar{x}\|, \|A\bar{y}\| \leq R(\delta)$ and $\|p\|, \|q\| \leq C(\delta, \varepsilon)$ by remark 11.8 (note that the derivatives of the test functions also depend on ε); therefore, $|(11.15)|$ can be made smaller than any positive number by fitting $\nu = \nu(\delta, \varepsilon)$, if necessary. Finally,

$$\begin{aligned}
&|(11.16) + (11.17)| \\
&\quad \leq \sup_{a \in \mathbb{A}} \left\{ |\langle t_1 + t_3, Ba \rangle| + \left| l(\bar{x}, \bar{\xi}, a) - l(\bar{y}, \bar{\zeta}, a) \right| + \delta M_{\mathbb{A}} (\|A\bar{x}\| + \|A\bar{y}\|) \right\} \\
&\quad \leq 2\nu M_{\mathbb{A}} + C_l \sqrt{|\Omega|} \left(\|\bar{x} - \bar{y}\| + \left\| \bar{\xi} - \bar{\zeta} \right\| \right) + \delta^{\frac{1}{2}} M_{\mathbb{A}}^2 + \frac{\delta^{\frac{3}{2}}}{2} \left(\|A\bar{x}\|^2 + \|A\bar{y}\|^2 \right) \\
&\quad \leq 2\nu M_{\mathbb{A}} + \omega(\varepsilon) + \delta^{\frac{1}{2}} M_{\mathbb{A}}^2 + \frac{\delta^{\frac{3}{2}}}{2} \left(\|A\bar{x}\|^2 + \|A\bar{y}\|^2 \right),
\end{aligned}$$

where we used the Lipschitz property of l, the imbedding $L^2(\Omega) \hookrightarrow L^1(\Omega)$, Young's inequality and (11.12). Plugging in those estimates then yields

$$\begin{aligned}
\lambda \left(u_1(\bar{x}, \bar{\xi}) - u_2(\bar{y}, \bar{\zeta}) \right) &\leq \left(-\frac{\delta}{4} + \frac{\delta^{\frac{3}{2}}}{2} \right) \left(\|A\bar{x}\|^2 + \|A\bar{y}\|^2 \right) \\
&+ c \left(\frac{\delta}{2} + \delta^{\frac{1}{2}} + \nu(\delta, \varepsilon) \right) + (11.15) + \omega(\varepsilon).
\end{aligned}$$

Now we choose $\hat{\delta} \in (0, \tilde{\delta})$ and $\nu = \nu(\hat{\delta}, \varepsilon) > 0$ such that for all $\varepsilon > 0$,

1. $\left(-\frac{\hat{\delta}}{4} + \frac{\hat{\delta}^{\frac{3}{2}}}{2} \right) \leq 0$,
2. $\frac{1}{\lambda} \left(\frac{\hat{\delta}}{2} + \hat{\delta}^{\frac{1}{2}} + \nu(\hat{\delta}, \varepsilon) \right) + \frac{1}{\lambda}(11.15) \leq \frac{\tau}{8}$,

which is possible as we are free to choose both parameters positive but arbitrary small. As (11.11) implies that $u_1(\bar{x}, \bar{\xi}) - u_2(\bar{y}, \bar{\zeta}) \geq \frac{\tau}{4}$, we get

$$\frac{\tau}{8} \leq \frac{1}{\lambda} \omega(\varepsilon),$$

a contradiction for $\varepsilon \downarrow 0$, and the proof is complete.

□

Now we can state the existence and uniqueness result.

Theorem 11.10 *Under assumptions (GA) and (lA), the value function is the unique bounded and Lipschitz continuous viscosity solution of* (11.5) *in the sense of definition 11.6.*

Proof: Theorem 11.7 shows that V is a viscosity solution and the comparison result, theorem 11.9, implies by the usual argumentation, that every viscosity solution is equal to V.

□

Example 11.11 *Equations* (11.1), (11.2) *imply that we may choose* $A = \Delta$*, the Laplace operator with* $D(A) = H_0^1(\Omega) \cap H^2(\Omega)$*, where* $\Omega \subset \mathbb{R}^n$ *is some bounded and smooth domain. As* $y \in C(0,T;H_0^1(\Omega)) = C(0,T;D((-A)^{\frac{1}{2}}))$ *due to the improved regularity result, we see that a possible choice is* $B = (-A)^{-\frac{1}{2}}$*. Then the assumptions (Vi), (Vii), (Viii) are fulfilled. For* $\mathbb{A}$*, we might choose a set of the form*

$$\mathbb{A} = \left\{a \in L^2(\Omega) \mid f_1(x) \le a(x) \le f_2(x),\ \text{a.e. } x \in \Omega\right\},$$

where $f_1, f_2 \in L^2(\Omega)$*. Then (GA) is satisfied. Now, as* l *must be bounded and Lipschitz continuous* $(L^1(\Omega))^2 \times \mathbb{A} \to \mathbb{R}$*, a possible choice would be*

$$l(y,w,a) = \int_\Omega \min\{R, |y(x)|^p\}\,dx + \int_\Omega \min\{S, |w(x)|^q\}\,dx + \int_\Omega |a(x)|\,dx,$$

for parameters $R, S > 0$ *and* $p, q \ge 1$*. Then (lA) holds. Thus, theorem 11.10 applies, e.g., to the following control problem:*

$$\min_{\alpha \in \mathcal{A}} \left\{ \int_0^\infty e^{-\lambda t} \left(\int_\Omega \min\left\{100, |y(t,x)|^2\right\} dx + \int_\Omega |\alpha(t,x)|\,dx \right) dt \right\},$$

w.r.t.

$$\begin{aligned} &\dot{y}(t) + \dot{w}(t) - \Delta y(t) = (-\Delta)^{-\frac{1}{2}}\alpha(t), \\ &w = \mathcal{F}_r[y; w_0], \quad w_0 \in [y_0 - r, y_0 + r]. \end{aligned}$$

□

APPENDIX

A Note on Equations with hysteresis

When dealing with equations including hysteresis, existence and uniqueness proofs sometimes are quite lengthy, and one could ask whether knowledge about the analogous equation without hysteresis might help. So the following note might be useful. Consider, formally, some equation of the type

$$F(y, u) = 0,$$

where y should be the solution function solving the equation, and u another function, on which the solution y depends, such as, e.g.,

$$F(y, u) := \dot{y} + u - \Delta y - f.$$

In this case, the solution $y = y(u)$ would solve a heat equation with some special data u, f (but f is fixed). If this equation is uniquely solvable for every u (in some suitable set), we can define a solution operator S, which maps u to the corresponding solution $y = S(u)$, so that

$$F(S(u), u) = 0.$$

If we now want to have hysteresis (represented by some operator $\mathcal{W}$) included in that equation, we look for a choice of u such that $u = \mathcal{W}(y) = \mathcal{W}(S(u))$. But that means, in order to solve the equation with hysteresis, we need to find a solution of the fixed point problem

$$u = \mathcal{W}(S(u)). \tag{A.1}$$

The following is an example of an abstract formulation of the contraction method used in some of the proofs in part two (and as proposed in [11, p. 300]).

Theorem A.1 *Let both the hysteresis and solution operator* $\mathcal{W}, S$ *be mappings from* $L^2(\Omega; H^1(0,T))$ *into itself, and Lipschitz continuous with constants* $c_{\mathcal{W}}, c_S > 0$*, such that* $\gamma := c_{\mathcal{W}} \cdot c_S \in (0,1)$*. Then the fixed point equation* (A.1) *has exactly one solution in* $L^2(\Omega; H^1(0,T))$.

Proof: With Banach's fixed point theorem. Let $\|\cdot\|$ denote a suitable norm on $L^2(\Omega; H^1(0,T))$. Then the simple estimate

$$\|\mathcal{W}(S(u_1)) - \mathcal{W}(S(u_2))\| \leq c_{\mathcal{W}} \|S(u_1) - s(u_2)\| \leq \gamma \|u_1 - u_2\|$$

shows that $\mathcal{W} \circ S$ is a contraction mapping.

□

The assumption on S is often fulfilled – at least for small $T > 0$ – when considering weak solutions of partial differential equations depending continuously on the data, where the time derivative of the hysteresis term does not appear in the equation. However, note the following examples.

Example A.2 (1.) *Heat equation with Play hysteresis: Semilinear case.*
Consider again (3.1)*, i.e.,*

$$\dot{y} - \Delta y + \mathcal{F}_r[y; w_0] = f,$$

together with $y_0 \in H_0^1(\Omega)$, $w_0 \in [y_0 - r, y_0 + r]$ and Dirichlet boundary conditions. In order to apply A.1, we then consider the solution operator $S : u \mapsto y = S(u)$ of the standard heat equation

$$\dot{y} - \Delta y + u = f,$$

where $u \in L^2(\Omega; H^1(0,T))$. Testing the equation as done in section 3, we can derive an inequality of the form

$$\|y_1 - y_2\|_{L^2(\Omega;H^1(0,T))} \leq c(T) \|u_1 - u_2\|_{L^2(\Omega;H^1(0,T))},$$

where $c(T)$ is continuous and such that $c(T) \downarrow 0$ as $T \downarrow 0$. Hence, the Lipschitz continuity of the Play implies that we may apply theorem A.1 to find a unique local solution of the equation with hysteresis, which may be continued by standard regularity results.

(2.) *Heat equation with Play hysteresis: Quasilinear case.*
Consider

$$\dot{y} + \frac{\partial}{\partial t}\mathcal{F}_r[y; w_0] - \Delta y = f,$$

with similar sideconditions as in the first example. Now, for ($c \in (-1, 1)$), take a look at

$$\dot{y} + c\dot{u} - \Delta y = f.$$

Considering only functions y_i, u_i with the same initial values, a simple calculation yields

$$\|y_1 - y_2\|_{L^2(\Omega;H^1(0,T))} \leq |c| \|u_1 - u_2\|_{L^2(\Omega;H^1(0,T))},$$

so theorem A.1 applies, and we get a unique solution for the equation with hysteresis for every $c \in (-1, 1)$. Using the continuous dependence on the data result of [11, Proposition 1.4, p. 270], we can let $c \uparrow 1$. However, uniqueness does not carry over here.

□

Another (academic) example where the fixed point method (A.1) might be used is when considering "implicit functions with hysteresis", i.e., equations (without derivatives) depending on variables y, u and t, such that, e.g., $u = \mathcal{F}_r[y]$. As an example, consider the equation

$$y^2 - tu^2 = 0. \tag{A.2}$$

Under which conditions it is assured that (A.2) has a (local) solution $y = y(t)$ and $u = \mathcal{F}_1[y; u_0](t)$ starting at $t_0 = \frac{1}{4}$, $y_0 = \frac{1}{2}$, $u_0 = 1$? We will give an answer using the implicit function theorem and Banach's fixed point theorem. As there might be fixed points which do not have the contraction property, this solution might, however, be improved.

Theorem A.3 *Assume that $y_0 \in \mathbb{R}^n$, $u_0 \in \mathbb{R}^n$ and $t_0 \in \mathbb{R}$ solve the equation*

$$F(y, u, t) = 0, \tag{A.3}$$

and that $F : B_{R_y}(y_0) \times B_{R_u}(u_0) \times B_{R_t}(t_0) \subset \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is continuously differentiable such that the Jacobian $J_yF(y_0, u_0, t_0)$ is invertible and

1. *the Lipschitz conditions*

$$\left|[(J_yF)^{-1}J_uF](y, u_1, t) - [(J_yF)^{-1}J_uF](y, u_2, t)\right| \leq \mu_1(y, u_1, u_2, t)\,|u_1 - u_2|\,,$$
$$\left|[(J_yF)^{-1}J_tF](y, u_1, t) - [(J_yF)^{-1}J_tF](y, u_1, t)\right| \leq \mu_2(y, u_1, u_2, t)\,|u_1 - u_2|\,,$$

   hold, where μ_1, μ_2 are continuous in some neighborhood of (y_0, u_0, u_0, t_0), and

2. $|[(J_yF)^{-1}J_uF](y_0, u_0, t_0)| < 1$.

Further, let P be an operator which is, for some $T^* > t_0$, Lipschitz continuous on $W^{1,1}(t_0, t^*; \mathbb{R}^n)$ with constant smaller than or equal to one, for all $t^* \in (t_0, T^*)$, and such that

$$\|P(y)\|_{W^{1,1}(t_0,t^*;\mathbb{R}^n)} \leq \|y\|_{W^{1,1}(t_0,t^*;\mathbb{R}^n)}\,, \qquad \forall y \in W^{1,1}(t_0, t^*; \mathbb{R}^n).$$

Then there exists $\bar{t} \in (t_0, T^*)$ and continuous functions $y(t)$, $u(t) = P(y)(t)$, such that

$$F(y(t), P(y)(t), t) = 0, \qquad \forall t \in [\,t_0, \bar{t}\,].$$

In the proof of the theorem, we need the following variant of the implicit function theorem, which is in that form a consequence of the contraction mapping principle.

Theorem A.4 ([33], theorem 3.4.10) *Let X, Y, Z be Banach spaces. Let $U \times V$ be an open subset of $X \times Y$. Suppose that $G : U \times V \to Z$ is continuous and has the property that d_2G exists and is continuous at each point of $U \times V$. Assume that the point $(x, y) \in X \times Y$ has the property that $G(x, y) = 0$ and that $d_2G(x, y)$ is invertible. Then there are open balls $M = B_X(x, r)$ and $N = B_Y(y, s)$ such that, for each $\zeta \in M$, there is a unique $\eta \in N$ satisfying $G(\zeta, \eta) = 0$. The function f, thereby uniquely defined near x by the condition $f(\zeta) = \eta$, is continuous.*

□

Proof of theorem A.3: Restricting ourselves to a small suitable time interval containing the initial value t_0 we may assume that F is continuously differentiable. Applying A.4 with $Y, Z = \mathbb{R}^n$, $X = \mathbb{R}$ and $G(t, y) := F(y, u(t), t)$, where u is any function in $W^{1,1}(0, T^*; \mathbb{R}^n)$, we get a local unique solution $y(t) = f(t, u(t))$ satisfying

$$F(f(t, u(t)), u(t), t) = 0, \tag{A.4}$$

for all $t > t_0$ small enough – in fact, this neighborhood depends on the variation of y, u, so that by restriction to a ball around the constant functions y_0, u_0 allows us, in view of (A.1), to formulate our problem as a fixed point problem in $W^{1,1}$ via

$$u(t) = P(y(t)) = P(f(t, u(t))).$$

As mentioned above, we want to apply Banach's fixed point theorem. Since by the assumptions on P,

$$\|P(f(\cdot, u(\cdot)))\|_{W^{1,1}} \leq \|f(\cdot, u(\cdot))\|_{W^{1,1}},$$

it suffices to find a ball B^* in $W^{1,1}$ such that $P \circ f(\cdot, B^*) \subset B^*$ and

$$\|f(\cdot, u_1(\cdot)) - f(\cdot, u_2(\cdot))\|_{W^{1,1}} \leq c \|u_1 - u_2\|_{W^{1,1}}$$

holds for all u_1, u_2 in B^* and some $c \in (0, 1)$. Differentiating (A.4) yields

$$0 = \frac{d}{dt} F(f(t, u(t)), u(t), t) = J_y F \cdot \left[\frac{d}{dt} f(t, u(t))\right] + J_u F \cdot \dot{u}(t) + J_t F,$$

which implies

$$\left[\frac{d}{dt} f(t, u(t))\right] = - [J_y F]^{-1} J_u F \cdot \dot{u}(t) - [J_y F]^{-1} J_t F.$$

Hence, separating also the left hand side of the latter equation into the part that contains u and the remaining variable, we get

$$f_u(t, u(t)) = - [J_y F]^{-1} J_u F, \qquad f_t(t, u(t)) = - [J_y F]^{-1} J_t F. \tag{A.5}$$

Next, note that for any u_1, u_2 in $W^{1,1}$ with the same starting values $u_1(0) = u_2(0)$, we have

$$\begin{aligned}
&\|f(\cdot, u_1(\cdot)) - f(\cdot, u_2(\cdot))\|_{W^{1,1}} \\
&= \int_{t_0}^{T^*} |f_t(t, u_1(t)) - f_t(t, u_2(t)) + f_u(t, u_1(t))\dot{u}_1(t) - f_u(t, u_2(t))\dot{u}_2(t)|\, dt \\
&\leq \int_{t_0}^{T^*} \mu_1 |u_1(t) - u_2(t)| + \mu_2 |u_1(t) - u_2(t)| |\dot{u}_1(t)|\, dt \\
&\quad + \int_{t_0}^{T^*} |f_u(t, u_2(t))| |\dot{u}_1(t) - \dot{u}_2(t)|\, dt.
\end{aligned} \tag{A.6}$$

By the choice of the u_i and the norm on $W^{1,1}$, we also have

$$|u_1(t) - u_2(t)| \leq \int_{t_0}^{T^*} |\dot{u}_1(t) - \dot{u}_2(t)|\, dt = \|u_1 - u_2\|_{W^{1,1}}.$$

This observation implies, together with (A.5) and assumption 2. that if $T^* > t_0$ is small enough, it holds

$$\|f(\cdot, u_1(\cdot)) - f(\cdot, u_2(\cdot))\|_{W^{1,1}} \leq \gamma \|u_1 - u_2\|_{W^{1,1}},$$

for all u_1, u_2 inside a (small) ball around the constant function $\tilde{u}(t) \equiv u_0$ and with the same initial value u_0. Let us denote this set by M. Then, by definition, $P(u)(0) = u_0$ for all u in M, and the boundedness and Lipschitz assumption assure that $P \circ f(M) \subset M$. Thus, the contraction mapping principle applies, and we get a local solution for small $T^* > 0$.

□

Before giving an example, we state a simple corollary.

Corollary A.5 *Let $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ be of class C^2 such that assumption 2. of theorem A.3 holds, and P is an operator of the form as in theorem A.3. Then equation* (A.3) *has a unique local solution.*

Proof: The stronger regularity assumed here implies locally the Lipschitz condition 1. of the previous theorem.

□

Example A.6 *We consider the example from above, i.e., the equation*

$$y^2 - tu^2 = 0,$$

where we now want to have solutions corresponding to $u = P(y)$ for different operators P. Consider the point $t_0 = \frac{1}{4}$, $y_0 = \frac{1}{2}$, $u_0 = 1$ which solves the equation, and the play operator $u = \mathcal{F}_{\frac{1}{2}}[y; u_0]$. Then u stays constant until $y(t) = \frac{3}{2}$ is reached, as the formula

$$y(t) = \sqrt{t}u$$

implies that y must be increasing; hence, we get the solution

$$y(t) = \sqrt{t}, \qquad t \in \left[\frac{1}{4}, \frac{9}{4}\right].$$

As both t and $u(t)$ must then be increasing for larger t, we would have to solve

$$y = (y - \frac{1}{2})\sqrt{t},$$

leading to

$$y(t) = \frac{\sqrt{t}}{2(\sqrt{t} - 1)}.$$

As this function is, however, decreasing for $t \geq \frac{3}{2}$, the equation is not longer solvable for such t. We note that for the derivatives, we get

$$f_u(t, u(t)) = \partial_u y = \partial_u(\sqrt{t}u) = \sqrt{t} = \frac{2tu}{2y} = -\left[J_y F\right]^{-1} J_u F,$$

and that the absolute value of the latter is here smaller than one as long as $t < 1$, which means that condition 2. of the above theorem is in fact too restrictive in this case.

Next, consider the same equation with the operator

$$P(y; u_0)(t) := u_0 + \int_{t_0}^{t} \max\{0, \dot{y}(s)\}\, ds.$$

It is easily seen that it fulfills the requirements of the above theorem. Further, to solve the equation, we now need to find a solution of

$$y = (y + \frac{1}{2})\sqrt{t},$$

which leads to

$$y(t) = \frac{\sqrt{t}}{2(1 - \sqrt{t})}, \qquad u(t) = y(t) + \frac{1}{2}.$$

Hence, we get a solution, as long as $t < 1$, which here coincides with the condition 2. of theorem A.3.

□

B Rate of convergence for the regularized equation

The convergence result in section 3.5 is based on compact imbedding results. However, by a much more careful analyzation of the approximation (2.1), one can give a rate of convergence for the regularized equation (3.7). To do so, we first restrict ourselves to a class of special functions, and then make a distinction of cases, to get an idea of when the approximation and the true value become different, and to what extent. After that, we can give an estimation of how much the two functions can at most differ, in terms of the parameter $\varepsilon > 0$.

Lemma B.1 *Let $v_\varepsilon := y - z_\varepsilon$ denote the corresponding approximation of the stop operator (recall that $\varepsilon \dot{z}_\varepsilon = G(z_\varepsilon - y)$ as defined in (2.1); in particular, we fix a function $y = y_\varepsilon$ for all $\varepsilon > 0$ here). Let denote $v_\varepsilon(t_0) := v_0$, $e(t) := \mathcal{E}_r[y; w_0](t)$, $e_0 := e(t_0) \in [-r, r]$. If $y(t) = c_0 + c_1(t - t_0)$ is an affine map, then*

$$|v_\varepsilon(t) - e(t)| \leq \max\left\{|v_0 - e_0|, |v_0 - e_0| e^{\eta(t)} + |c_1| \varepsilon(1 - e^{\eta(t)})\right\}, \tag{B.1}$$

for all $t \geq t_0$, where

$$\eta(t) := -\frac{1}{\varepsilon}(t - t_0). \tag{B.2}$$

Proof: We start with a distinction of cases. Until the end of this proof, we will drop the index ε.

- **Case 1:** $v(t) \in (-r, r)$.

  This implies that $\dot{v}(t) = \dot{y}(t) = c_1$. Further, we might have $\dot{e}(t) = 0$, or $\dot{e}(t) = c_1$. In the first case, we must have $e_0 = \pm r$ together with $\operatorname{sign}(c_1) = \pm 1$. But then the distance $d(t) := |v(t) - e(t)|$ becomes smaller, as long as we stay in case 1, so that

  $$d(t) = |v(t) - e(t)| \leq |v_0 - e_0| = d(0).$$

  On the other hand, if $\dot{v} = c_1 = \dot{e}$, the distance d does not change.

- **Case 2:** $v(t) \notin (-r, r)$.

  Here, we have to consider several subcases.

  - **Case 2.1:** $v(t) < -r, c_1 > 0$.

    Then $\dot{v} = c_1 - \frac{1}{\varepsilon}G(-v) > c_1$, whereas $\dot{e} = 0$ or $\dot{e} = c_1$. Thus, d gets smaller.

- **Case 2.2:** $v(t) > r, c_1 < 0$.
  Arguing similarly to case 2.1, we find that d gets smaller.
- **Case 2.3:** $v(t) > r, c_1 > 0$.
  Using the variation of constants formula as in lemma 3.24, we get the representation

$$v(t) = r + (v_0 - r)\exp\left(-\frac{1}{\varepsilon}(t - t_0)\right) + c_1\varepsilon\left(1 - \exp\left(-\frac{1}{\varepsilon}(t - t_0)\right)\right). \tag{B.3}$$

  Now, we consider the following subcases.
  * **Case 2.3.a):** $\dot{e} = 0$.
    Then $e(t) = r$ constant, so that

$$\begin{aligned}0 \le v(t) - e(t) = v(t) - r &= (v_0 - r)e^{\eta(t)} + c_1\varepsilon(1 - e^{\eta(t)}) \\ &= (v_0 - e_0)e^{\eta(t)} + c_1\varepsilon(1 - e^{\eta(t)}).\end{aligned}$$

    Hence, we get the estimate

$$|v(t) - e(t)| \le |v_0 - e_0|\, e^{\eta(t)} + |c_1|\, \varepsilon(1 - e^{\eta(t)}). \tag{B.4}$$

  * **Case 2.3b):** $\dot{e} = c_1$.
    Then $e(t) = e_0 + c_1(t - t_0)$. On the other hand, by use of the simple estimate

$$1 - \exp\left(-\frac{1}{\varepsilon}(t - t_0)\right) \le \frac{t - t_0}{\varepsilon}, \qquad \forall t \ge t_0,$$

    we get from (B.3) that

$$v(t) \le r + v_0 - r + c_1\varepsilon\frac{t - t_0}{\varepsilon} = v_0 + c_1(t - t_0) = v_0 - e_0 + e(t),$$

    implying, due to $v > r$, that

$$d(t) = |v(t) - e(t)| \le |v_0 - e_0| = d(0).$$

    Hence, the difference becomes smaller.
- **Case 2.4:** $v(t) < -r, c_1 < 0$.
  By symmetry, or considering subcases similar to 2.3a), 2.3b), we get analogously, that $d(t)$ can only grow via formula (B.4).

Altogether, we are left with analyzing the estimate (B.4). As for

$$f(t) := d(0)e^{\eta(t)} + |c_1|\,\varepsilon(1 - e^{\eta(t)}),$$

it holds

$$f'(t) = \left(|c_1| - \frac{d(0)}{\varepsilon}\right) e^{\eta(t)},$$

the function f is either monotonically increasing or decreasing, and (B.1) follows.

□

With lemma B.1, we can now investigate how one can estimate the difference d for arbitrary continuous and piecewise linear input functions y. Then we use density arguments to get more general assertions.

Proposition B.2 *Let $v_0 = e_0 \in [-r, r]$ be given.*

1. *Assume that $y \in W^{1,1}(0,T)$ has the property that $\dot{y} \in L^\infty(0,T)$. Then*

$$\|v_\varepsilon(\cdot) - e(\cdot)\|_{C[0,T]} \leq \varepsilon \, \|\dot{y}\|_{L^\infty(0,T)} \,. \tag{B.5}$$

2. *Assume that $y \in H^1(0,T)$. Then*

$$\|v_\varepsilon(\cdot) - e(\cdot)\|_{C[0,T]} \leq \frac{\sqrt{\varepsilon}}{2}\left[1 + \|\dot{y}\|^2_{L^2(0,T)}\right]. \tag{B.6}$$

Proof: We start with $y \in C_{pl}[0,T]$; therefore, $\dot{y}$ is simply a step function, corresponding to which there is a partition of the interval $[0,T]$ via time points $0 = t_0 < t_1 < \ldots < t_n = T$, such that $\dot{y}$ exists and is constant on each of the intervals (t_j, t_{j+1}). According to lemma B.1, over each such interval, the difference of the true value and the approximation does not grow or grows via

$$|d(t)| \leq d(t_j) \exp\left(-\frac{1}{\varepsilon}(t - t_j)\right) + |c_j|\,\varepsilon\left(1 - \exp\left(-\frac{1}{\varepsilon}(t - t_j)\right)\right), \tag{B.7}$$

where $t \in [t_j, t_{j+1}]$ and $c_j = \dot{y}$ in (t_j, t_{j+1}). Thus, in the case of growing difference, the maximum would be reached at $t = t_{j+1}$, so that we may introduce the variable $k_{j+1} := d(t_{j+1})$ — it then suffices to give a uniform estimate for all the k_j. Let us further denote

$$\eta_j := -\frac{1}{\varepsilon}(t_{j+1} - t_j).$$

We also introduce the index set I consisting of all $0 \leq i \leq n-1$, for which d actually grows on $[t_i, t_{i+1}]$, i.e., for which (B.7) has to be applied. Then we can estimate, by iteration of (B.7),

$$k_l \leq k_0 \exp\left\{ \sum_{i\in I, i<l} \eta_{l-i} \right\} + \varepsilon \sum_{i\in I, i<l} |c_i| \left(1 - e^{\eta_{l-i}}\right) \exp\left\{ \sum_{m\in I, m\geq i} \eta_{l-m} \right\}. \tag{B.8}$$

As $k_0 = |v_\varepsilon(0) - e(0)| = 0$, we get from (B.8) the estimate

$$k_l \leq \varepsilon \sum_{i\in I, i<l} |c_i| \left(1 - e^{\eta_{l-i}}\right) \exp\left\{ \sum_{m\in I, m\geq i} \eta_{l-m} \right\}. \tag{B.9}$$

Next, we make the following simplification. Since the step functions corresponding to uniform partitions of $[0, T]$ are still dense in $L^1(0, T)$, we may restrict our exposition to the case $t_{j+1} - t_j =: \Delta t$, for all j. Then $\eta_j = -\frac{1}{\varepsilon}\Delta t =: \eta$ for all j, and (B.9) reads

$$\begin{aligned} k_l \leq & \varepsilon \sum_{i\in I, i<l} |c_i| \left(1 - e^{\eta}\right) \exp\left\{ \sum_{m\in I, m\geq i} \eta \right\} \\ = & \varepsilon (1 - e^{\eta}) \sum_{i\in I, i<l} |c_i| \left(e^{\eta}\right)^{\#\{m\in I, m\geq i\}}. \end{aligned} \tag{B.10}$$

Now, (B.10) implies

$$k_l \leq \varepsilon \|\dot{y}\|_{L^\infty(0,T)} \left(1 - e^{\eta}\right) \sum_{i=0}^{\infty} (e^{\eta})^i = \varepsilon \|\dot{y}\|_{L^\infty(0,T)},$$

which holds for all l. Hence, by density of step functions, we can infer inequality (B.5). To prove the second part, we apply Young's inequality D.1 with $p = q = 2$, $\delta^2 = \frac{\Delta t}{\sqrt{\varepsilon}}$, and

$$a = |c_i|, \qquad b = \left(1 - e^{\eta}\right) \left(e^{\eta}\right)^{\#\{m\in I, m\geq i\}}.$$

Thus,

$$\begin{aligned} k_l \leq & \frac{\varepsilon}{2} \sum_{i\in I, i<l} \left\{ \frac{\Delta t}{\sqrt{\varepsilon}} |c_i|^2 + \frac{\sqrt{\varepsilon}}{\Delta t} (1 - e^{\eta})^2 \left(e^{\eta}\right)^{2\#\{m\in I, m\geq i\}} \right\} \\ \leq & \frac{\sqrt{\varepsilon}}{2} \|\dot{y}\|^2_{L^2(0,T)} + \frac{\varepsilon^{\frac{3}{2}}}{2} \frac{1 - e^{\eta}}{\Delta t} (1 - e^{\eta}) \sum_{i\in I, i<l} \left(e^{\eta}\right)^{\#\{m\in I, m\geq i\}}. \end{aligned}$$

Now, we make use of $1 - e^{\eta} \leq \frac{\Delta t}{\varepsilon}$, so that we may further estimate

$$\begin{aligned} k_l \leq & \frac{\sqrt{\varepsilon}}{2} \|\dot{y}\|^2_{L^2(0,T)} + \frac{\sqrt{\varepsilon}}{2}(1 - e^{\eta}) \sum_{i=0}^{n-1} (e^{\eta})^i \\ \leq & \frac{\sqrt{\varepsilon}}{2} \|\dot{y}\|^2_{L^2(0,T)} + \frac{\sqrt{\varepsilon}}{2}(1 - e^{\eta}) \sum_{i=0}^{\infty} (e^{\eta})^i \\ = & \frac{\sqrt{\varepsilon}}{2} \left[1 + \|\dot{y}\|^2_{L^2(0,T)} \right], \end{aligned}$$

for all l. Thus, by density of $C_{pl}[0,T]$ in $H^1(0,T)$, the result follows.

□

We note that if $\dot{y} \in L^{\infty}(0,T)$, then y is continuous on $[0,T]$, and therefore $y \in L^{\infty}(0,T)$. Thus, the functions in point one of the last proposition are exactly the ones of $W^{1,\infty}(0,T)$, which can be shown to be equal to the set of Lipschitz continuous functions on $[0,T]$. Hence, (B.5) may be reformulated via replacing the L^{∞}-norm on the right hand side by the Lipschitz constant of y.

Up to now, we have only considered the approximation v_{ε} for the stop operator, but it is easy to get a similar result for z_{ε}.

Corollary B.3 *Let $z_0 = w_0 \in [y_0 - r, y_0 + r]$ be given.*

1. Assume that $y \in W^{1,1}(0,T)$ has the property that $\dot{y} \in L^{\infty}(0,T)$. Then

$$\|z_{\varepsilon}(\cdot) - w(\cdot)\|_{C[0,T]} \leq \varepsilon \|\dot{y}\|_{L^{\infty}(0,T)} . \tag{B.11}$$

2. Assume that $y \in H^1(0,T)$. Then

$$\|z_{\varepsilon}(\cdot) - w(\cdot)\|_{C[0,T]} \leq \frac{\sqrt{\varepsilon}}{2} \left[1 + \|\dot{y}\|^2_{L^2(0,T)} \right] . \tag{B.12}$$

Proof: By definition, $z_{\varepsilon} + v_{\varepsilon} = y$. Since we also know that $w + e = y$, it holds that

$$\|z_{\varepsilon} - w\|_{C[0,T]} = \|(y - v_{\varepsilon}) - (y - e)\|_{C[0,T]} = \|v_{\varepsilon} - e\|_{C[0,T]} ,$$

and the result follows directly from proposition B.2.

□

Note that those estimates hold for fixed $\varepsilon > 0$ and any y — thus, we may also choose a family of functions y_{ε} which is bounded in $H^1(0,T)$. Then corollary B.3 tells us, that for any such family, we have

$$\|z_{\varepsilon}(\cdot, y_{\varepsilon}) - w(\cdot, y_{\varepsilon})\|_{C[0,T]} \to 0,$$

as $\varepsilon \downarrow 0$. This might look a bit strange at first glance, but it simply reflects the compactness of the imbedding $H^1(0,T) \hookrightarrow C[0,T]$, since any weakly convergent subsequence of y_ε satisfies the assumptions of theorem 2.11.

It is clear that if $y = y_\varepsilon$ depends on $\varepsilon > 0$ such that $y_\varepsilon \to \bar{y}$ as $\varepsilon \downarrow 0$, the rate of convergence of $z_\varepsilon = z_\varepsilon(\cdot, y_\varepsilon)$ will depend on how fast y_ε converges to y. The following estimate, which makes use of the Lipschitz continuity of the play operator, is a very simple way to include this; nevertheless, it will enable us to get a rate of convergence for the p.d.e. (3.7) via Gronwall's lemma.

Corollary B.4 *Assume that* $y_\varepsilon, \bar{y} \in H^1(0,T)$ *with* $y_\varepsilon(0) = \bar{y}(0) =: y_0$*, and* $z_\varepsilon(0) = z_0 = w_0 = w(0,\bar{y}) = w(0,y_\varepsilon) \in [y_0 - r, y_0 + r]$*, for all* $\varepsilon > 0$*. Then,*

$$\|z_\varepsilon(\cdot, y_\varepsilon) - w(\cdot, \bar{y})\|_{C[0,T]} \le \|y_\varepsilon - \bar{y}\|_{C[0,T]} + \frac{\sqrt{\varepsilon}}{2}\left[1 + \|\dot{y}_\varepsilon\|_{L^2(0,T)}\right]. \tag{B.13}$$

Proof: From the triangle inequality, the Lipschitz continuity of the play on $C[0,T]$, and corollary B.3, we get

$$\begin{aligned}\|z_\varepsilon(\cdot, y_\varepsilon) - w(\cdot, \bar{y})\|_{C[0,T]} &\le \|w(\cdot, y_\varepsilon) - w(\cdot, \bar{y})\|_{C[0,T]} + \|z_\varepsilon(\cdot, y_\varepsilon) - w(\cdot, y_\varepsilon)\|_{C[0,T]} \\ &\le \|y_\varepsilon - \bar{y}\|_{C[0,T]} + \frac{\sqrt{\varepsilon}}{2}\left[1 + \|\dot{y}_\varepsilon\|_{L^2(0,T)}\right].\end{aligned}$$

□

Now we are ready to proof a rate of convergence result for (3.7).

Theorem B.5 *Let* y *denote the solution of* (3.1) *w.r.t.* $w := \mathcal{F}_r[y; w_0] = \mathcal{W}[y]$, $w_0 \in [y_0 - r, y_0 + r]$*, and let the assumptions of theorem 3.3 (existence of weak solutions) hold, particularly* $y_0 \in H_0^1(\Omega)$, $f \in L^2(\Omega)$ *(where* Ω *is bounded with smooth boundary). Further, for every* $\varepsilon > 0$*, let* y_ε *denote the weak solution of* (3.2) *(cf. theorem 3.10) w.r.t. the initial value* $y_\varepsilon(0) = y_0$ *and let* z_ε *denote the corresponding regularization of the play, where* $z_\varepsilon(0) := z_0 = w_0$*. Then there exists a constant* $c > 0$ *such that*

$$\|y_\varepsilon - y\|_{L^2(\Omega; C[0,T])} \le c\sqrt{\varepsilon}.$$

Proof: As we have seen in section 3, we may test the difference of the two equations with $\partial_t v := \partial_t(y_\varepsilon - y)$, which yields

$$\|\dot{v}\|^2_{L^2(\Omega_T)} + \int_0^T \int_\Omega (z_\varepsilon - w)\dot{v}\, d\mathcal{L} + \frac{1}{2}\|\nabla v(T, \cdot)\|^2_{L^2(\Omega)} \le 0.$$

Noting the simple estimate

$$\frac{1}{T}\|v\|^2_{L^2(\Omega; C[0,T])} \le \|\dot{v}\|^2_{L^2(\Omega_T)},$$

we easily get to

$$\begin{aligned}\|v\|^2_{L^2(\Omega;C[0,T])} \leq & T^2 \int_0^T \int_\Omega (z_\varepsilon(t,x) - w(t,x))^2 \, dx dt \\ \leq & T^2 \int_0^T \int_\Omega \|z_\varepsilon(\cdot,x) - w(\cdot,x)\|^2_{C[0,t]} \, dx dt,\end{aligned}$$

via application of Young's inequality (with $p = q = 2$, $\delta^2 = T$). Then, for almost every $x \in \Omega$, corollary B.4 applies. Thus,

$$\|v\|^2_{L^2(\Omega;C[0,T])} \leq 2T^2 \int_0^T \|v\|^2_{L^2(\Omega;C[0,t])} \, dt + \varepsilon T^3 \int_\Omega \left[1 + \|\dot{y}_\varepsilon(\cdot,x)\|_{L^2(0,T)}\right]^2 dx.$$

As proposition 3.25 implies that the term in square brackets remains bounded in $L^2(\Omega)$ as $\varepsilon \downarrow 0$, there exists a constant $\hat{c} > 0$ such that

$$\|v\|^2_{L^2(\Omega;C[0,T])} \leq 2T^2 \int_0^T \|v\|^2_{L^2(\Omega;C[0,t])} \, dt + \hat{c}\varepsilon.$$

As this calculation remains true for all $T > 0$, we may apply Gronwall's lemma, which yields

$$\|v\|^2_{L^2(\Omega;C[0,T])} \leq \varepsilon \hat{c} \exp\left(2T^3\right).$$

□

C A regularized Hilpert-type inequality

The Hilpert inequality is, as we have mentioned before, the main tool to proof uniqueness of solutions in the quasilinear case of a heat equation with hysteresis. It can be stated as follows.

Theorem C.1 (Hilpert's inequality, cf. [12, p. 134]) *Consider the hysteresis operator $\mathcal{W}$ given by*

$$\mathcal{W}[y; w_{-1}](t) := q(\mathcal{F}_r[y; w_{-1}](t)), \qquad 0 \le t \le T,$$

with $w_{-1} \in \mathbb{R}$, and where $q \in W^{1,\infty}_{loc}(\mathbb{R})$ is an increasing function. Suppose that $y_1, y_2 \in W^{1,1}(0,T)$ and $w_{-1,1}, w_{-1,2} \in \mathbb{R}$ are given, and let $y := y_2 - y_1$, $w := w_2 - w_1$, where $w_i := \mathcal{W}[y_i; w_{-1,i}]$, $i = 1, 2$. Then

$$\frac{d}{dt} w_+(t) \le w'(t) H(y(t)), \qquad \text{a.e. in } (0,T), \tag{C.1}$$

where $w_+ := \max\{w, 0\}$, and H denotes the Heavyside function

$$H(x) = \begin{cases} 1, & x > 0, \\ 0, & x \le 0. \end{cases}$$

□

We imitate the proof to get a result which looks similar for the regularization z_ε.

Theorem C.2 *Let $y_i \in W^{1,1}(0,T)$ and $z_i := z_{\varepsilon_i}(\cdot, y_i)$, $v_i := v_{\varepsilon_i}(\cdot, y_i)$, $i = 1, 2$, denote the regularizations as in section B. Further, let $w_i = \mathcal{F}_r[y_i; w_{0,i}]$ denote the Play operator, $e_i = \mathcal{E}_r[y_i; e_{0,i}]$ the Stop operator (recall that $w_i + e_i = y_i$) and $w_{0,i} = w_i(0) = z_i(0) \in [y_i(0) - r, y_i(0) + r]$, and $y := y_2 - y_1$, $w := w_2 - w_1$, $\delta_1 := e_1 - v_1$, $\delta_2 := e_2 - v_2$, $\delta := \delta_2 - \delta_1$. Then,*

$$z'(t) H(y(t)) \ge \frac{d}{dt} [w_+(t)] + \delta'(t) H(y(t)), \qquad \text{a.e. } t \in (0,T).$$

Proof: We use the crucial implication for the proof of the Hilpert inequality,

$$w_2(t) < w_1(t), \quad y_2(t) \ge y_1(t) \quad \Rightarrow \quad w_2'(t) \ge 0, \quad w_1'(t) \le 0. \tag{C.2}$$

As noted before, it holds (by definition)

$$z_2(t) - w_2(t) = e_2(t) - v_2(t), \qquad z_1(t) - w_1(t) = e_1(t) - v_1(t). \tag{C.3}$$

We consider the sets of time points

$$M_1 := \{t | w_1(t) > w_2(t)\}, \quad M_2 := \{t | w_2(t) > w_1(t)\}, \quad M := \{t | w_1(t) = w_2(t)\}.$$

Consider $t \in M_1$. Then, if $y_2(t) \geq y_1(t)$, (C.2) applies and yields (subtraction of (C.3))

$$z'(t) \geq \delta'(t);$$

thus, in general, we have

$$z'(t)H(y(t)) \geq \delta'(t)H(y(t)), \qquad a.e.\ t \in M_1. \tag{C.4}$$

Next, consider $t \in M_2$. If $y_1(t) \geq y_2(t)$, by reversing the indices in (C.2), we find that

$$z'(t) \leq \delta'(t),$$

so that in general (note (C.3)),

$$z'(t) \leq \delta'(t) + w'(t)H(y(t)) = \delta'(t)(1 - H(y(t))) + z'(t)H(y(t)),$$

which may be rewritten as

$$z'(t)H(y(t)) \geq w'(t) - \delta'(t)H(y(t)), \tag{C.5}$$

a.e. $t \in M_2$. Now, note that for some (somewhere) dense subset $S \subset M$, by continuity of w, S must be an interval, and thus, $w'(t) = 0$ for a.e. $t \in M$, implying, in view of (C.3),

$$z'(t)H(y(t)) = \delta'(t)H(y(t)), \tag{C.6}$$

a.e. $t \in M$. Summarizing (C.4), (C.5), (C.6) yields the result.

□

D Some frequently used Theorems

A collection of some of the theorems which are applied many times during the exposition.

Lemma D.1 (Young's Inequality)
Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then, for any $\delta > 0$,

$$|ab| \leq \frac{1}{p}\delta^p |a|^p + \frac{1}{q}\delta^{-q} |b|^q, \qquad \forall a, b \in \mathbb{R}.$$

Proof: As shown in [12, lemma 1.2.1], this is a simple consequence of the convexity of the logarithm and the monotonicity of the exponential function. Since

$$\ln\left(\frac{1}{p}\delta^p |a|^p + \frac{1}{q}\delta^{-q} |b|^q\right) \geq \frac{1}{p}\ln\left(\delta^p |a|^p\right) + \frac{1}{q}\ln\left(\delta^{-q} |b|\right),$$

applying the exponential on both sides yields the result.

□

We usually apply lemma D.1 with parameters $p = q = 2$ and $\delta = 1$.

Theorem D.2 (Jensen's inequality, probabilistic version)
Let $(X, \mathcal{A}, \mu)$ be a probability space and $f : X \to J$ (where $J \subset \mathbb{R}$ open) some integrable function. If φ is convex on J, then

$$\varphi\left(\int_X f d\mu\right) \leq \int_X \varphi \circ f d\mu.$$

Proof: See e.g., [34, theorem 6.4.1].

□

This probabilistic version implies the discrete (when μ is a counting measure) and the continuous one (when μ equals Lebesgue measure).

Theorem D.3 (Gronwall's lemma, different versions)

1. *Discrete version: Let $c > 0$, $N \in \mathbb{N}$ and suppose that nonnegative real numbers a_n, b_n, $0 \leq n \leq N$, satisfy*

$$a_n \leq c + \sum_{j=0}^{n-1} a_j b_j, \qquad 0 \leq n \leq N-1.$$

Then

$$a_n \leq c \exp\left(\sum_{j=0}^{n-1} b_j\right), \qquad 0 \leq n \leq N.$$

2. *Continuous version A: Let $c \in C[0,T]$ and $a \in L^1(0,T)$ denote nonnegative functions. If $u \in C[0,T]$ satisfies*

$$0 \leq u(t) \leq c(t) + \int_0^t a(s)u(s)ds, \qquad t \in [0,T],$$

then

$$0 \leq u(t) \leq c(t) + \int_0^t c(s)a(s)\exp\left(\int_s^t a(\tau)d\tau\right)ds, \qquad t \in [0,T].$$

3. *Continuous version B: Let x, k be continuous and a, b Riemann integrable functions on $J := [0,T]$ with b, k nonnegative on J. If*

$$x(t) \leq a(t) + b(t)\int_0^t k(s)x(s)ds, \quad t \in J,$$

then

$$x(t) \leq a(t) + b(t)\int_0^t a(s)k(s)\exp\left(\int_s^t b(r)k(r)dr\right)ds, \quad t \in J.$$

Proof: See, e.g., [35, corollary 2.1.5., page 150] for version 1, [36, III, page 14] for the second and [37, theorem 1, page 356] for the third version.

□

References

[1] M. Quincampoix A. Gudovich. Optimal control with hysteresis nonlinearity and multidimensional play operator. *SIAM J. Control Optim.*, no. 2:788–807, 2011.

[2] J. Yong X. Li. *Optimal Control Theory for Infinite Dimensional Systems.* Systems & Control. Birkhäuser Boston, 1995.

[3] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen.* Vieweg+Teubner, GWV Fachverlage GmbH, Wiesbaden, second edition, 2009.

[4] I. Capuzzo-Dolcetta M. Bardi. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations.* Systems & Control. Springer Science + Business Media New York, first edition, 1997.

[5] P.L. Lions M.G. Crandall. Viscosity solutions of hamilton-jacobi equations in infinite dimensions III-VII. *J. Funct. Anal.*

[6] J.-P Raymond S. Gombao. Hamilton-jacobi equations for control problems of parabolic equations. *ESAIM Control Optim. Calc Var. 12*, no. 2:311–349, 2006.

[7] M.E. Tessitore P. Cannarsa. Infinite-dimensional hamilton-jacobi equations and dirichlet boundary control problems of parabolic type. *SIAM J. Control and Optimization*, 34(6):1831–1847, 1996.

[8] R. Fischer H. Linse. *Elektrotechnik für Maschinenbauer: Grundlagen und Anwendungen.* Vieweg+Teubner Verlag, 2002.

[9] H.R. Noori. *Hysteresis Phenomena in Biology.* Springer Briefs in Applied Sciences and Technology, 2014.

[10] M. Göcke. *Starke Hysteresis im Außenhandel.* Number no. 84 in Wirtschaftswissenschaftliche Beiträge. Physica Verlag Heidelberg, 1993.

[11] A. Visintin. *Differential Models of Hysteresis.* Springer Verlag Berlin Heidelberg, 1994.

[12] J. Sprekels M. Brokate. *Hysterresis and Phase Transitions.* Springer, Applied Mathematical Sciences, volume 121, 1996.

[13] P. Krejčí M. Brokate. Optimal control of ode systems involving a rate independent variational inequality. *Discrete and Continuous Dynamical Systems Series B*, Volume 18(Number 2):pp. 331 – 348, March 2013.

[14] V. Sobolev A. Pokrovskii. A naive view of time relaxation and hysteresis. In *Singular perturbations and hysteresis*, chapter one. SIAM, 2005.

[15] P. Krejčí. Evolution variational inequalities and multidimensional hysteresis operators. *Chapman & Hall/CRC Res. Notes Math. 404*, 1999.

[16] A. Pokrovskii D. Rachinskii P. Krejčí, J.P. O'Kane. Stability results for a soil model with singular hysteretic hydrology. WIAS preprint, 2008.

[17] V. Recupero. The play operator on the rectifiable curves in a hilbert space. *Math. Methods Appl. Sci. 31*, no. 11:1283–1295, 2008.

[18] J. Fournier R. Adams. *Sobolev Spaces.* Pure and applied mathematics ; v. 140. Amsterdam ; Boston, Academic Press, second edition, 2003.

[19] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*, volume 140 of *Graduate Studies in Mathematics.* American Mathematical Society, 2012.

[20] P. L. Combettes H. H. Bauschke. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* CMS Books in Mathematics. Springer Science+Business Media, LLC, 2011.

[21] W. Czaja J.J. Benedetto. *Integration and Modern Analysis.* Birkhäuser Boston, a part of Springer Science + Business Media, LLC, 2009.

[22] E. Magenes J. L. Lions. *Non-Homogeneous Boundary Value Problems And Applications*, volume I. Springer-Verlag Berlin Heidelberg New York, 1972.

[23] E. Magenes J. L. Lions. *Non-Homogeneous Boundary Value Problems And Applications*, volume II. Springer-Verlag Berlin Heidelberg New York, 1972.

[24] L.C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, reprint 2002 edition.

[25] A. Lunardi. *Analytic Semigroups and Optimal Regularity in Parabolic Problems.* Springer Basel, 1995. Reprint of the 1st edition 1995 by Birkhäuser Verlag, Switzerland.

[26] M.C. Delfour S.K. Mitter A. Bensoussan, G. Da Prato. *Representation and Control of Infinite Dimensional Systems.* Systems & Control. Birkhäuser Boston, second edition, 2007.

[27] E. Zeidler. *Nonlinear Functional Analysis and its Applications II/B, Nonlinear Monotone Operators.* Springer Science+Business Media New York, 1990.

[28] F. Bagagiolo. Dynamic programming for some optimal control problems with hysteresis. *NoDEA Nonlinear Differential Equations Appl. 9*, no. 2:149–174, 2002.

[29] H. Frankowska P. Cannarsa. Value function and optimality condition for semilinear control problems. II. parabolic case. *Appl. Math. Opt. 33*, no. 1:1–33, 1996.

[30] C. Stegall. Optimization of functions on certain subsets of banach spaces. *Math. Ann. 236*, no. 2:171–176, 1978.

[31] E. Valdinoci E. Di Nezza, G. Palatucci. Hitchhiker's guide to the fractional sobolev spaces. *Bull. Sci. Math. 136*, no. 5:521–573, 2012.

[32] K. Yosida. *Functional Analysis*. Springer Verlag Berlin Heidelberg, sixth edition, 1995.

[33] H. R. Parks S. G. Krantz. *The Implicit Function Theorem*. Modern Birkhäuser Classics. Springer Science+Business Media New York, 2013.

[34] P.A. Loeb. *Real Analysis*. Birkhäuser, Springer International Publishing Switzerland, 2016.

[35] Y. Qin. *Integral and Discrete Inequalities and their Applications Volume I: Linear Inequalities*. Springer International Publishing Switzerland, 2016.

[36] W. Walter. *Differential and Integral Inequalities*. Number 55 in Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer-Verlag Berlin Heidelberg New York, 1970.

[37] D. S. Mitrinović J. E. Pečarić A. M. Fink. *Inequalities Involving Functions and Their Integrals and Derivatives*, volume 53 of *Mathematics and Its Applications (East European Series)*. Springer-Science+Business Media Dordrecht, 1991.

List of Problem Definitions and Assumptions

Problem (P_0) 14
Problem (S) 29
Assumption (A1) 65
Problem (P) 68
Assumption (A2) 68
Problem (Q) 71
Assumption A3 102
Assumption A3' 107
Problem (P') 114
Assumption (GA) 155
Assumption (lA) 155

List of Figures

1 Hysteresis loop 5
2 Input-output relation of Relay operator 6
3 Play between mechanical elements 6
4 Input-output relation of Play operator 7
5 Input-output relation of Stop operator 7
6 graph of G for r=0.5 11
7 graph of $G_{(r,s),(m_1,m_2)}$ for r=0.5, $m_1 = 1$, s=0.7, $m_2 = 0.5$ 18
8 Numerical approximation of Stop operator 35