Nonsmooth Evolutions *

Martin Brokate **

Contents

*Course Notes, University of Graz, November 5 - 15, 2019
**TU Munich and WIAS Berlin

1 Some Basic Notions of Convex Analysis

An introductory example. We consider the initial value problem

$$u' + \operatorname{sign}(u) = 0\,, \quad u(0) = u_0\,. \tag{1.1}$$

Here, $\operatorname{sign}(x) = 1$ if $x > 0$ and $\operatorname{sign}(x) = -1$ if $x < 0$. For the moment we do not fix the value of $\operatorname{sign}(0)$.

We look for solutions $u : \mathbb{R}_+ \to \mathbb{R}$. The only candidate which makes sense is

$$u(t) = \begin{cases} \max\{u_0 - t, 0\}\,, & \text{if } u_0 \geq 0\,, \\ \min\{u_0 + t, 0\}\,, & \text{if } u_0 \leq 0\,. \end{cases} \tag{1.2}$$

Indeed, in the region $\{(t, v) : t \in \mathbb{R},\, v \neq 0\} \subset \mathbb{R}^2$ every solution of $u' + \operatorname{sign}(u) = 0$ satisfies $(d/dt)|u| = -1$. Since $u(|u_0|) = 0$, for $t > |u_0|$ we cannot have $u(t) \neq 0$.

The function (1.2) is Lipschitz continuous, but not differentiable at $t = u_0$. Moreover, it depends continuously on u_0. We obtain it as a solution of (1.1), almost everywhere in $\mathbb{R}_+$, if and only if we set $\operatorname{sign}(0) = 0$.

We can avoid having to find the correct value of $\operatorname{sign}(0)$ if we consider sign as a set-valued function,

$$\operatorname{sign}(x) = \begin{cases} 1\,, & x > 0\,, \\ -1\,, & x < 0\,, \\ [-1, 1]\,, & x = 0\,, \end{cases} \tag{1.3}$$

and write the differential equation as a differential inclusion

$$u' + \operatorname{sign}(u) \ni 0\,, \quad u(0) = u_0\,. \tag{1.4}$$

The solution (1.2) satisfies $u'(t) + \operatorname{sign}(u(t)) \ni 0$ almost everywhere in $\mathbb{R}_+$.

We will see in a moment that $\operatorname{sign}(x) = \partial\varphi(x)$ for $\varphi(x) = |x|$.

The definition (1.2) of the sign function has the property that every function $\alpha : \mathbb{R} \to \mathbb{R}$ with $\alpha(x) \in \operatorname{sign}(x)$ for all $x \in \mathbb{R}$ is nondecreasing.

The situation is different for the initial value problem

$$u' - \operatorname{sign}(u) = 0\,, \quad u(0) = u_0\,. \tag{1.5}$$

Here we have

$$u(t; u_0) = \begin{cases} u_0 + t\,, & \text{if } u_0 > 0\,, \\ u_0 - t\,, & \text{if } u_0 < 0\,. \end{cases} \tag{1.6}$$

Since

$$\lim_{\varepsilon \downarrow 0} u(t; \varepsilon) = t\,, \quad \lim_{\varepsilon \uparrow 0} u(t; \varepsilon) = -t\,, \tag{1.7}$$

it is not possible to define a solution $u(\cdot, 0) : \mathbb{R}_+ \to \mathbb{R}$ such that both uniqueness and continuous dependence on the initial value hold.

Some basic notions of convex analysis. In the following, all vector spaces are vector spaces over the real numbers.

Definition 1.1 (Epigraph)
Let V be a vector space, $\varphi : V \to (-\infty, \infty]$. The subset

$$\operatorname{epi}\varphi = \{(v, \mu) : v \in V,\ \mu \in \mathbb{R},\ \mu \geq \varphi(v)\} \tag{1.8}$$

*of $V \times \mathbb{R}$ is called the **epigraph** of φ.*

$$D(\varphi) = \{v : v \in V,\ \varphi(v) < +\infty\} \tag{1.9}$$

*is called the **effective domain** or simply **domain** of φ.*

Denoting by $p_V : V \times \mathbb{R} \to V$ the projection to the first component, we get

$$D(\varphi) = p_V(\operatorname{epi}\varphi)\,. \tag{1.10}$$

We have $\operatorname{epi}\varphi = \emptyset$ if and only if $\varphi(v) = +\infty$ for all $v \in V$. If this is not the case, φ is called **proper**.

Definition 1.2 (Convex function)
*Let V be a vector space. A function $\varphi : V \to (-\infty, \infty]$ is called **convex**, if $\operatorname{epi}\varphi$ is convex.*

Proposition 1.3 *Let V be a vector space. A function $\varphi : V \to (-\infty, \infty]$ is convex if and only if*

$$\varphi(\lambda v + (1-\lambda)w) \leq \lambda\varphi(v) + (1-\lambda)\varphi(w) \tag{1.11}$$

for all $v, w \in V$ and all $\lambda \in [0, 1]$.

Proof: Direct from the definitions. $\square$

Definition 1.4 *Let V be a vector space, $K \subset V$ convex. A function $\varphi : K \to (-\infty, \infty]$ is called convex, if the function $\tilde{\varphi} : V \to (-\infty, \infty]$,*

$$\tilde{\varphi}(v) = \begin{cases} \varphi(v)\,, & v \in K\,, \\ +\infty\,, & v \notin K\,, \end{cases} \tag{1.12}$$

is convex.

Let $K \subset V$. The function $I_K : V \to [0, \infty)$,

$$I_K(v) = \begin{cases} 0\,, & v \in K \\ +\infty\,, & v \notin K\,, \end{cases} \tag{1.13}$$

is called the **indicator function** of K. This function is convex if and only if K is convex, since $\operatorname{epi} I_K = K \times \mathbb{R}_+$.

Lemma 1.5 *Let V be a vector space, $\varphi : V \to (-\infty, \infty]$ convex. Then the **sublevel sets** $\{v : v \in V,\ \varphi(v) \leq \alpha\}$ and $\{v : v \in V,\ \varphi(v) < \alpha\}$ are convex for all $\alpha \in (-\infty, \infty]$.*

Proof: Direct from the definitions. $\square$

Definition 1.6 (Lower semicontinuity)
*Let V be a Banach space. A function $\varphi : V \to (-\infty, \infty]$ is called **(weakly) lower semicontinuous** if the sublevel sets*

$$M_\alpha = \{v : v \in V,\, \varphi(v) \le \alpha\} \tag{1.14}$$

are (weakly) closed for all $\alpha \in \mathbb{R}$.

Proposition 1.7 *Let V be a Banach space. A function $\varphi : V \to (-\infty, \infty]$ is (weakly) lower semicontinuous if and only if* $\operatorname{epi}\varphi$ *is (weakly) closed in $V \times \mathbb{R}$.*

Proof: "$\Leftarrow$": Let $\alpha \in \mathbb{R}$. Then

$$F_\alpha = \{(v, \alpha) : v \in V,\, \varphi(v) \le \alpha\} = \operatorname{epi}\varphi \cap (V \times \{\alpha\}) \tag{1.15}$$

is (weakly) closed in $V \times \mathbb{R}$, therefore, too, the sublevel set $M_\alpha = j_\alpha^{-1}(F_\alpha)$; here $j_\alpha : V \to V \times \mathbb{R}$ denotes the embedding $j_\alpha(v) = (v, \alpha)$.
"$\Rightarrow$": We show that the complement of $\operatorname{epi}\varphi$ is open. Let $(v, \alpha) \notin \operatorname{epi} f$, so $\varphi(v) > \alpha$. Choose an $\varepsilon > 0$ such that $\varphi(v) > \alpha + \varepsilon$. By assumption, the set $U = \{v : \varphi(v) > \alpha + \varepsilon\}$ is open in V, and $W = U \times (\alpha - \varepsilon, \alpha + \varepsilon)$ is an open neighborhood of (v, α) with $W \cap \operatorname{epi}\varphi = \emptyset$, since $\varphi(w) > \alpha + \varepsilon > \beta$ for all $(w, \beta) \in W$. $\square$

Corollary 1.8 *Let V be a Banach space. A subset $K \subset V$ is closed if and only if I_K is lower semicontinuous.*

Proof: "$\Rightarrow$": $\operatorname{epi} I_K = K \times [0, \infty)$.
"$\Leftarrow$": $K = \{v : I_K(v) \le 0\}$. $\square$

Corollary 1.9 *Let V be a Banach space, $\varphi : V \to (-\infty, \infty]$ convex. Then φ is lower semicontinuous if and only if it is weakly lower semicontinuous.*

Proof: In Banach space, a convex set is closed if and only if it is weakly closed. $\square$

Lemma 1.10 *Let V be a Banach space, let $\varphi : V \to (-\infty, \infty]$ be convex and lower semicontinuous. Then*

$$\varphi(v) \le \liminf_{n \to \infty} \varphi(v_n) \tag{1.16}$$

for all sequences $v_n \rightharpoonup v$.

Proof: Assume that $v_n \rightharpoonup v$, but $\varphi(v) > \liminf \varphi(v_n)$. Then there exists a subsequence $\{v_{n_k}\}$ and an $\varepsilon > 0$ such that $\varphi(v_{n_k}) \le \varphi(v) - \varepsilon =: \alpha$. As φ is lower semicontinuous, the sublevel set M_α is weakly closed, so $\varphi(v) \le \alpha$, a contradiction. $\square$

Proposition 1.11 *Let V be a reflexive Banach space, let $\varphi : V \to (-\infty, \infty]$ be convex, lower semicontinuous and proper. Moreover, we assume that*

$$\lim_{\|v\| \to \infty} \varphi(v) = \infty\,. \tag{1.17}$$

Then there exists a $u \in V$ such that

$$\varphi(u) = \min_{v \in V} \varphi(v)\,. \tag{1.18}$$

Proof: Let $\{u_n\}$ be a minimal sequence for φ in V, that is, $\varphi(u_n) \downarrow \inf_{v\in V} \varphi(v)$. Due to (1.17), all sublevel set of φ and, hence, the sequence $\{u_n\}$ are bounded in V. Since V is reflexive, there exists a subsequence $\{u_{n_k}\}$ with $u_{n_k} \rightharpoonup u$ for some $u \in V$. From Lemma 1.10 it follows that

$$\varphi(u) \leq \liminf_{k\to\infty} \varphi(u_{n_k}) = \inf_{v\in V} \varphi(v)\,.$$

□

Lemma 1.12 *Let V be a Banach space, let $\varphi_i : V \to (-\infty, \infty]$ be convex and lower semicontinuous for all $i \in I$. Then $\sup_{i\in I} \varphi_i$ is convex and lower semicontinuous.*

Proof:

$$\operatorname{epi}(\sup_{i\in I} \varphi_i) = \bigcap_{i\in I} \operatorname{epi} \varphi_i\,.$$

□

Proposition 1.13 *Let V be a Banach space, $\varphi : V \to (-\infty, \infty]$ convex, lower semicontinuous and proper. Then*

$$\varphi = \sup\{g|\ g : V \to \mathbb{R} \text{ affine and continuous, } g \leq \varphi\}\,. \tag{1.19}$$

Proof: "$\geq$": Obvious.
"$\leq$": It suffices to show: If $(v, a) \in V \times \mathbb{R}$ with $a < \varphi(v)$, then there exists an affine continuous function $g : V \to \mathbb{R}$ such that $a \leq g(v)$ and $g \leq \varphi$. This is proved as a consequence of the separation theorem, applied in the space $V \times \mathbb{R}$ to such a point $(\varphi(v), a)$ and the closed convex set $\operatorname{epi}\varphi$. □

Definition 1.14 (Subdifferential)
*Let V be a Banach space, $\varphi : V \to (-\infty, \infty]$. A functional $u^* \in V^*$ is called a **subgradient** of φ in $u \in V$, if $\varphi(u) < \infty$ and*

$$\varphi(v) \geq \varphi(u) + \langle u^*, v - u\rangle\ , \quad \text{for all } v \in H. \tag{1.20}$$

The set

$$\partial\varphi(u) = \{u^* : u^* \in V^*,\ w \text{ is a subgradient for } \varphi \text{ in } u\} \tag{1.21}$$

*is called the **subdifferential** of φ in u.*
We set $\varphi(u) = \emptyset$ if $\varphi(u) = \infty$.

If $V = H$ is a Hilbert space, according to the Riesz isomorphism between H and H^* we also call $w \in H$ a subgradient of φ in $u \in H$, if $\varphi(u) < \infty$ and

$$\varphi(v) \geq \varphi(u) + \langle w, v - u\rangle\ , \quad \text{for all } v \in H. \tag{1.22}$$

The subdifferential is then given by

$$\partial\varphi(u) = \{w : w \in H,\ w \text{ is a subgradient for } \varphi \text{ in } u.\} \tag{1.23}$$

Example 1.15

(i) For $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi(v) = |v|$, we have $\partial\varphi(u) = \{1\}$ if $u > 0$, $\partial\varphi(u) = \{-1\}$ if $u < 0$, and $\partial\varphi(0) = [-1, 1]$. Thus, the subdifferential of the absolute value function equals the set-valued sign function.

(ii) For $\varphi : \mathbb{R} \to \mathbb{R}$,

$$\varphi(v) = \begin{cases} 1\,, & v > 0\,, \\ 0\,, & v \le 0\,, \end{cases}$$

we have $\partial\varphi(0) = \{0\}$. But if we set

$$\varphi(v) = \begin{cases} 1\,, & v \ge 0\,, \\ 0\,, & v < 0\,, \end{cases}$$

we have $\partial\varphi(0) = \emptyset$.

Definition 1.16 (Normal cone)
Let V be a Banach space, $K \subset V$ convex, $u \in K$. An $u^* \in V^*$ is called a ***support functional*** *for K in u if*

$$\langle u^*, u - v\rangle \ge 0\,, \qquad \textit{for all } v \in K. \tag{1.24}$$

The set

$$N_K(u) = \{u^* : u^* \in V^*,\ u^* \textit{ is a support functional for } K \textit{ in } u\} \tag{1.25}$$

is called the ***normal cone*** *at K in u. For $u \notin K$ we set $N_K(u) = \emptyset$.*

Again, if $V = H$ is a Hilbert space, one may replace the support functionals $u^* \in H^*$ by elements $w \in H$ with $\langle w, u - v\rangle \ge 0$ for all $v \in K$. Then $N_K(u)$ becomes a subset of H instead of H^*.

Lemma 1.17 *Let V be a Banach space, $K \subset V$ convex. Then*

$$\partial I_K(u) = N_K(u)\,, \quad \textit{if } u \in K, \tag{1.26}$$

and $\partial I_K(u) = \emptyset$ otherwise.

Proof: Direct from the definitions. □

Proposition 1.18 *Let V be a Banach space, $\varphi : V \to (-\infty, \infty]$, let $u \in V$ with $\varphi(u) < \infty$. Then*

$$\varphi(u) = \min_{v \in V} \varphi(v) \qquad \Leftrightarrow \qquad 0 \in \partial\varphi(u)\,. \tag{1.27}$$

Proof: Direct from the definition. □

Lemma 1.19 *Let $\varphi : \mathbb{R} \to (-\infty, \infty]$ convex, $u \in D(\varphi)$. Then*

$$d(t) = \frac{\varphi(u+t) - \varphi(u)}{t} \tag{1.28}$$

defines a nondecreasing function $d : \mathbb{R} \setminus \{0\} \to (-\infty, \infty]$. Moreover, $d(-t) \leq d(t)$ for all $t > 0$.

Proof: We first consider d on $(0, \infty)$. For $0 < s < t$ we have

$$u + s = \frac{t-s}{t} u + \frac{s}{t}(u+t)\,,$$

therefore

$$\varphi(u+s) \leq \frac{t-s}{t}\varphi(u) + \frac{s}{t}\varphi(u+t)\,.$$

We subtract $\varphi(u)$ and divide by s to obtain

$$d(s) = \frac{\varphi(u+s) - \varphi(u)}{s} \leq \frac{\varphi(u+t) - \varphi(u)}{t} = d(t)\,.$$

Now we consider d on $(-\infty, 0)$. For this purpose, we define $\tilde{\varphi} : V \to (-\infty, \infty]$ by $\tilde{\varphi}(r) = \varphi(2u - r)$ Then $\tilde{\varphi}$ is convex, and the corresponding difference quotient becomes

$$\tilde{d}(t) = \frac{\tilde{\varphi}(u+t) - \tilde{\varphi}(u)}{t} = \frac{\varphi(u-t) - \varphi(u)}{t} = -d(-t)\,.$$

By what we have proved above, $\tilde{d}$ is nondecreasing on $(0, \infty)$. Therefore, d is nondecreasing on $(-\infty, 0)$.

Finally, we show that $d(-t) = d(t)$ for all $t > 0$. Indeed,

$$\varphi(u) \leq \frac{1}{2}\varphi(u-t) + \frac{1}{2}\varphi(u+t)\,.$$

Thus $\varphi(u) - \varphi(u-t) \leq \varphi(u+t) - \varphi(u)$ and therefore

$$d(-t) = \frac{\varphi(u-t) - \varphi(u)}{-t} \leq \frac{\varphi(u+t) - \varphi(u)}{t} = d(t)\,.$$

□

Proposition 1.20 *Let H be a Hilbert space, let $\varphi : H \to (-\infty, \infty]$ be convex, lower semicontinuous and proper, $f \in H$. Then the function $J : H \to (-\infty, \infty]$,*

$$J(v) = \frac{c}{2}\|v - f\|^2 + \varphi(v)\,, \tag{1.29}$$

has a unique minimum $u \in H$, and

$$c(f - u) \in \partial\varphi(u)\,. \tag{1.30}$$

Proof: As φ is convex, lower semicontinuous and proper, the same is true for J. By Proposition 1.13, J has an affine minorant, that is, there exists $w \in H$ and $\alpha \in \mathbb{R}$ such that

$$\varphi(v) \geq \langle w, v\rangle - \alpha\,, \quad \text{for all } v \in H. \tag{1.31}$$

It follows that

$$J(v) \geq \|v\| \Big(\frac{c}{2}\|v\| - c\|f\| - \|w\|\Big) - \alpha\,, \quad \text{for all } v \in H.$$

Thus, $J(v) \to \infty$ for $\|v\| \to \infty$. Now Proposition 1.11 implies that J has a minimum $u \in H$; this minimum is unique since J is strictly convex.

The optimality condition (1.30) can be obtained from the sum rule for subdifferentials. Alternatively, an elementary proof is the following. Let $v \in H$ be arbitrary. We set

$$v_t = u + t(v-u)\,, \quad t \in [0,1]\,.$$

For all $t \in [0,1]$ we have

$$\begin{aligned} 0 \leq J(v_t) - J(u) &= \frac{c}{2}\|(u-f) + t(v-u)\|^2 - \frac{c}{2}\|u-f\|^2 + \varphi(v_t) - \varphi(u) \\ &= ct\,\langle u-f, v-u\rangle + \frac{ct^2}{2}\|v-u\|^2 + \varphi(v_t) - \varphi(u)\,. \end{aligned}$$

Dividing by t yields

$$\begin{aligned} 0 &\leq c\,\langle u-f, v-u\rangle + \frac{ct}{2}\|v-u\|^2 + \frac{\varphi(v_t)-\varphi(u)}{t} \\ &\leq c\,\langle u-f, v-u\rangle + \frac{ct}{2}\|v-u\|^2 + \varphi(v) - \varphi(u)\,, \end{aligned}$$

the latter since the difference quotient is monotone according to Lemma 1.19. Passing to the limit $t \downarrow 0$ yields

$$0 \leq c\,\langle u-f, v-u\rangle + \varphi(v) - \varphi(u)\,.$$

As $v \in H$ was arbitrary, $c(f-u) \in \partial\varphi(u)$ follows. $\square$

2 Monotone Operators

Let V, W be sets, $R \subset V \times W$ a relation. We may interpret a relation as a **set-valued mapping** in the following way. Given a relation R, we define

$$A : V \to \mathcal{P}(W)\,, \tag{2.1}$$

by

$$Av = \{w : w \in W,\ (v, w) \in R\}\,. \tag{2.2}$$

Instead of (2.1), we write

$$A : V \rightrightarrows W\,. \tag{2.3}$$

Conversely, any set-valued mapping $A : V \rightrightarrows W$ defines a relation R if we set

$$(v, w) \in R \quad \Leftrightarrow \quad w \in Av\,.$$

The **domain** and the **range** of A are defined by

$$\begin{aligned} D(A) &= \{v : v \in V,\ Av \neq \emptyset\} \\ \mathrm{im}\,(A) &= \bigcup_{v \in V} Av \end{aligned} \tag{2.4}$$

Let $A, B : V \rightrightarrows W$ be set-valued mappings which arise from relations R and S, respectively. B is called an **extension** of A if $R \subset S$; if moreover $R \neq S$, the extension is called **proper**.

The **inverse** $A^{-1} : W \rightrightarrows V$ of $A : V \rightrightarrows W$ is defined as

$$A^{-1}w = \{v : v \in V,\ w \in Av\}\,. \tag{2.5}$$

We have $D(A^{-1}) = \mathrm{im}\,(A)$ since

$$w \in Av \qquad \Leftrightarrow \qquad v \in A^{-1}w$$

holds for all $v \in V$, $w \in W$.

Let W be a vector space, let $A, B : V \rightrightarrows W$ and $\lambda \in \mathbb{R}$. We define

$$\begin{aligned} \lambda A &= \{(v, \lambda w) : v \in V\,,\ w \in Av\}\,, \\ A + B &= \{(v, w + z) : v \in V\,,\ w \in Av\,,\ z \in Bv\}\,, \\ \mathrm{cl}\,(\mathrm{conv}\,A) &= \{(v, w) : v \in V\,,\ w \in \mathrm{cl}\,(\mathrm{conv}\,Av)\}\,. \end{aligned}$$

For the sum we have $D(A + B) = D(A) \cap D(B)$, as $M + \emptyset = \emptyset$ for every subset M of H.

Definition 2.1 (Monotone operator)
(i) Let V be a Banach space. An operator $A : V \rightrightarrows V^*$ is called ***monotone*** *if*

$$\langle v_2^* - v_1^*, v_2 - v_1\rangle \geq 0\,, \quad \textit{for all } v_1, v_2 \in H,\ v_1^* \in Av_1,\ v_2^* \in Av_2. \tag{2.6}$$

(ii) Let H be a Hilbert space. An operator $A : H \rightrightarrows H$ is called ***monotone*** *if*

$$\langle w_2 - w_1, v_2 - v_1\rangle \geq 0\,, \quad \textit{for all } v_1, v_2 \in H,\ w_1 \in Av_1,\ w_2 \in Av_2. \tag{2.7}$$

In both cases, A is called ***maximal monotone*** *if A does not have a proper extension which is monotone.*

Let $f : \mathbb{R} \to \mathbb{R}$ be monotone nondecreasing. Then

$$\tilde{f}(r) = [f(r-), f(r+)] , \quad f(r-) := \sup_{t<r} f(t) , \quad f(r+) := \inf_{t>r} f(t) , \tag{2.8}$$

defines a maximal monotone operator $\tilde{f} : \mathbb{R} \rightrightarrows \mathbb{R}$.

Using Zorn's lemma one can show that every monotone operator has a maximal monotone extension.

Lemma 2.2 *Let V be a Banach space, let $A, B : V \rightrightarrows V^*$ monotone, $\lambda \geq 0$.*
(i) A^{-1}, λA, $A + B$ and $\mathrm{cl}\,(\mathrm{conv}\, A)$ are monotone.
(ii) If A is maximal monotone, then Av is closed and convex for all $v \in H$.
The same is true if H is a Hilbert space and $A, B : H \rightrightarrows H$ are monotone.

Proof: (i) follows directly from the definitions. (ii) holds since $\mathrm{cl}\,(\mathrm{conv}\, A)$ is a monotone extension of A by (i), so $A = \mathrm{cl}\,(\mathrm{conv}\, A)$ if A is maximal monotone. □

Definition 2.3 (Accretive operator)
Let V be a Banach space, $A : V \rightrightarrows V$.
*(i) A is called **accretive** if*

$$\|v_2 - v_1\| \leq \|(v_2 + \lambda w_2) - (v_1 + \lambda w_1)\| \tag{2.9}$$

holds for all $v_1, v_2 \in V$, $w_1 \in Av_1$, $w_2 \in Av_2$ and all $\lambda > 0$.
*(ii) A is called **maximal accretive** if it is accretive and has no proper extension which is accretive.*
*(iii) A is called **m-accretive** if it is accretive and $\mathrm{im}\,(I + A) = V$.*

Proposition 2.4 *Let H be a Hilbert space, $A : H \rightrightarrows H$.*
(i) Let $v, w \in H$. Then

$$\langle v, w\rangle \geq 0 \qquad \Leftrightarrow \qquad \|v\| \leq \|v + \lambda w\| \quad \text{for all } \lambda > 0. \tag{2.10}$$

(ii) A is monotone if and only if A is accretive.

Proof: (i) This follows from the equality

$$\|v + \lambda w\|^2 - \|v\|^2 = 2\lambda\langle v, w\rangle + \lambda^2\|w\|^2 .$$

(ii) We apply (i) with $v = v_2 - v_1$ and $w = w_2 - w_1$, where $w_i \in Av_i$.
□

Proposition 2.5 *Let H be a Hilbert space, $\varphi : H \to (-\infty, \infty]$ convex, lower semicontinuous and proper. Then $\partial\varphi : H \rightrightarrows H$ is m-accretive.*

Proof: Let $v_1, v_2 \in H$ and $w_1 \in \partial\varphi(v_1)$, $w_2 \in \partial\varphi(v_2)$. Adding the inequalities

$$\begin{aligned} \varphi(v_2) - \varphi(v_1) &\geq \langle w_1, v_2 - v_1\rangle \\ \varphi(v_1) - \varphi(v_2) &\geq \langle w_2, v_1 - v_2\rangle \end{aligned}$$

yields $\langle w_2 - w_1, v_2 - v_1 \rangle \geq 0$, so $\partial\varphi$ is accretive by Proposition 2.4. It remains to show that

$$\mathrm{im}\,(I + \partial\varphi) = H\,.$$

Let $f \in H$ be arbitrary. We define $J : H \to (-\infty, \infty]$ by

$$J(v) = \frac{1}{2}\|v - f\|^2 + \varphi(v)\,.$$

According to Proposition 1.20, J has a unique minimum $u \in H$, and $f - u \in \partial\varphi(u)$. Thus $f \in \mathrm{im}\,(I + \partial\varphi)$. □

Example 2.6 On $H = L^2(\Omega)$, $\Omega \subset \mathbb{R}^n$ open, we define

$$Av = -\Delta v\,, \quad D(A) = H_0^1(\Omega) \cap H^2(\Omega)\,. \tag{2.11}$$

We want to prove that $A : H \rightrightarrows H$ is m-accretive. Indeed, $A : D(A) \to H$ is monotone (hence, accretive) because

$$\langle Av, v \rangle = \int_\Omega -v(x) \cdot \Delta v(x)\, dx = \int_\Omega \|\nabla v(x)\|^2\, dx \geq 0\,, \quad \text{for all } v \in D(A).$$

Therefore, A is m-accretive if and only if the boundary value problem

$$-\Delta u + u = f$$

has a solution $u \in D(A)$ for every $f \in L^2(\Omega)$. This is the case if Ω is bounded and $\partial\Omega$ is sufficiently regular. We refer to the variational theory of elliptic partial differential equations, see e.g. [8]. □

Definition 2.7 (Nonexpansive operator)
*Let H be a Hilbert space. A set-valued operator $A : H \rightrightarrows H$ is called **nonexpansive** if*

$$\|w_2 - w_1\| \leq \|v_2 - v_1\| \tag{2.12}$$

holds for all $v_1, v_2 \in H$, $w_1 \in Av_1$, $w_2 \in Av_2$.

Setting $v_2 = v_1$ in (2.12) we see that, for all $v \in H$, Av has at most one element. Nonexpansive operators are therefore single-valued mappings $A : D(A) \to H$.

Lemma 2.8 *Let H be a Hilbert space, $A : H \rightrightarrows H$. Then A is accretive if and only if $(I + \lambda A)^{-1}$ is nonexpansive for all $\lambda \geq 0$.*

Proof: This is a direct consequence of the definitions 2.3 and 2.7. □

Lemma 2.9 *Let H be a Hilbert space, $A : H \rightrightarrows H$ accretive. Then there are equivalent:*
(i) $\mathrm{im}\,(I + \lambda A) = H$ *for some* $\lambda > 0$,
(ii) $\mathrm{im}\,(I + \lambda A) = H$ *for all* $\lambda > 0$.

Proof: Let $\mathrm{im}\,(I+\lambda A)=H$ for some $\lambda>0$. Fix $\mu>\lambda/2$. For arbitrary $w\in H$ we define $T_w:H\to H$ by

$$T_w(v)=(I+\lambda A)^{-1}\Big(\frac{\lambda}{\mu}w+(1-\frac{\lambda}{\mu})v\Big)\,.$$

Since $0<\lambda/\mu<2$ and $(I+\lambda A)^{-1}$ is nonexpansive, T_w is a contraction on H. Let $v\in H$ be the fixed point of T_w. Then

$$\frac{\lambda}{\mu}w+(1-\frac{\lambda}{\mu})v\in(I+\lambda A)v\,.$$

Subtracting v and multiplying by μ/λ gives $w\in v+\mu Av$. As w was arbitrary, $\mathrm{im}\,(I+\mu A)=H$ for $\mu>\lambda/2$. Now (ii) follows by induction. $\square$

Proposition 2.10 (Characterization of maximal monotone operators)
Let H be a Hilbert space, $A:H\rightrightarrows H$. The following are equivalent:

(i) *A is maximal monotone.*

(ii) *A is maximal accretive.*

(iii) *A is m-accretive.*

(iv) *A is accretive and $\mathrm{im}\,(I+\lambda A)=H$ for all $\lambda>0$.*

Proof: “(i)⇔(ii)”: This follows, since by Proposition 2.4(ii) an extension of A is monotone if and only if it is accretive.
“(iii)⇔(iv)”: This is a direct consequence of Lemma 2.9.
“(iii)⇒(ii)”: Let B be an accretive extension of A, let $w\in Bv$. Since A is m-accretive, there exists $u\in D(A)$ such that $w+v\in u+Au$. Thus

$$w+v\in u+Bu\,,\quad w+v\in v+Bv\,,$$

so $u,v\in(I+B)^{-1}(w+v)$. Since $(I+B)^{-1}$ is nonexpansive, we must have $u=v$, thus $w+v\in v+Av$ and finally $w\in Av$. It follows that $A=B$.
“(i)⇒(iii)”: This proof is long and delicate, it relies on a minimax theorem (an existence result for saddle points) which in turn is based on Brouwer’s fixed point theorem. We refer to [5] and to [2], Theorem 2.2. $\square$

Lemma 2.11 (Lipschitz perturbation)
Let $A:H\rightrightarrows H$ be m-accretive, $B:H\to H$ accretive and Lipschitz continuous with $D(B)=H$. Then $A+B$ is m-accretive.

Proof: Choose $\lambda>0$ small enough such that $\lambda B:H\to H$ is a contraction. For every $w\in H$ we have

$$u+\lambda Au+\lambda Bu\ni w\quad\Leftrightarrow\quad u=T_wu:=(I+\lambda A)^{-1}(w-\lambda Bu)$$

As $(I+\lambda A)^{-1}$ is nonexpansive, $T_w:H\to H$ is a contraction and thus has a fixed point $u\in H$. Therefore, $w\in\mathrm{im}\,(I+\lambda(A+B))$. Since w was arbitrary, it follows from Proposition 2.10 that $A+B$ is m-accretive. $\square$

Part (ii) of the following lemma is a variant of what is known as **Minty’s trick**. It allows to pass to the limit in the scalar product although both factors converge only weakly.

Lemma 2.12 *Let $A : H \rightrightarrows H$ be m-accretive, let $w_n \in Av_n$, $v_n \rightharpoonup v$ and $w_n \rightharpoonup w$ in H.*
(i) If

$$\liminf_{n\to\infty} \langle w_n, v_n\rangle \le \langle w, v\rangle$$

then $w \in Av$.
(ii) If

$$\limsup_{n\to\infty} \langle w_n, v_n\rangle \le \langle w, v\rangle$$

then $w \in Av$ and $\langle w_n, v_n\rangle \to \langle w, v\rangle$.

In particular, if one of the sequences converges strongly, then the scalar product converges and $w \in Av$ holds.

Proof: (i) Let $\tilde{v} \in D(A)$ and $\tilde{w} \in A\tilde{v}$. As A is accretive,

$$0 \le \langle w_n - \tilde{w}, v_n - \tilde{v}\rangle = \langle w_n, v_n\rangle - \langle w_n, \tilde{v}\rangle - \langle \tilde{w}, v_n\rangle + \langle \tilde{w}, \tilde{v}\rangle \ .$$

Passing to the limit inferior for $n \to \infty$ yields

$$0 \le \liminf_{n\to\infty} \langle w_n, v_n\rangle - \langle w, \tilde{v}\rangle - \langle \tilde{w}, v\rangle + \langle \tilde{w}, \tilde{v}\rangle \le \langle w - \tilde{w}, v - \tilde{v}\rangle \ . \tag{2.13}$$

Thus $\tilde{A} : H \rightrightarrows H$ defined by $\tilde{A}v = Av \cup \{w\}$ and $\tilde{A}\tilde{v} = A\tilde{v}$ for $\tilde{v} \neq v$ is monotone. As A is maximal, we must have $w \in Av$.
(ii) Setting $\tilde{v} = v$ and $\tilde{w} = w$ in (2.13) yields

$$\langle w, v\rangle \le \liminf_{n\to\infty} \langle w_n, v_n\rangle$$

which, together with the assumption, implies $\langle w_n, v_n\rangle \to \langle w, v\rangle$. □

An important tool for the analysis of accretive resp. monotone operators are approximations by single-valued mappings, in particular the following one.

Let $A : H \rightrightarrows H$ be accretive. By Lemma 2.8,

$$J_\lambda := (I + \lambda A)^{-1}$$

is a nonexpansive operator for every $\lambda > 0$, with $D(J_\lambda) = H$ if A is m-accretive.

Definition 2.13 (Yosida regularization)
Let $A : H \rightrightarrows H$ be accretive, $\lambda > 0$. The ***Yosida regularization*** *$A_\lambda : H \to H$ of A is defined as*

$$A_\lambda = \frac{1}{\lambda}(I - J_\lambda)\,. \tag{2.14}$$

Let $\varphi : H \to (-\infty, \infty]$ be convex, lower semicontinuous and proper. We have seen in Proposition 1.20 that

$$v \mapsto \frac{1}{2\lambda}\|v - u\|^2 + \varphi(v)$$

has a unique minimum in H.

Definition 2.14 (Moreau regularization)
Let $\varphi : H \to (-\infty, \infty]$ convex, lower semicontinuous and proper, let $\lambda > 0$. The function

$$\varphi_\lambda(u) = \min_{v \in H} \left(\frac{1}{2\lambda} \|v - u\|^2 + \varphi(v) \right) \tag{2.15}$$

*is called the **Moreau regularization** of φ.*

Example 2.15 (i) Let $H = \mathbb{R}$, $\varphi = I_{\{0\}}$, that is, $\varphi(r) = 0$ for $r = 0$ and $\varphi(r) = \infty$ for $r \neq 0$. The Moreau regularization of φ is

$$\varphi_\lambda(r) = \min_{s \in \mathbb{R}} \left(\frac{1}{2\lambda} |s - r|^2 + \varphi(s) \right) = \frac{1}{2\lambda} r^2 .$$

The subdifferential $\beta = \partial\varphi : \mathbb{R} \rightrightarrows \mathbb{R}$ is

$$\beta(r) = \begin{cases} \mathbb{R} , & r = 0 , \\ \emptyset , & r \neq 0 . \end{cases}$$

We have $r \in (I + \lambda\beta)(s)$ if and only if $s = 0$ and $r \in \mathbb{R}$, thus $(I + \lambda\beta)^{-1} = 0$. The Yosida regularization of β becomes

$$\beta_\lambda(r) = \frac{r}{\lambda} .$$

(ii) Let $H = \mathbb{R}$, $\varphi(r) = |r|$. The Moreau regularization of φ is given by

$$\varphi_\lambda(r) = \begin{cases} r - \frac{\lambda}{2} , & r > \lambda , \\ \frac{r^2}{2\lambda} , & |r| \leq \lambda , \\ -r - \frac{\lambda}{2} , & r < -\lambda . \end{cases}$$

The subdifferential $\beta = \partial\varphi$ equals the set-valued sign function

$$\beta(r) = \begin{cases} 1 , & r > 0 , \\ [-1, 1] , & r = 0 , \\ -1 , & r < 0 . \end{cases}$$

We have $r \in (I + \lambda\beta)(s)$ if and only if $r = s + \lambda$ (if $s > 0$) resp. $r = s - \lambda$ (if $s < 0$) resp. $r \in s + [-\lambda, \lambda] = [-\lambda, \lambda]$ (if $s = 0$). It follows that

$$J_\lambda(r) = \begin{cases} r - \lambda , & r > \lambda , \\ 0 , & r \in [-\lambda, \lambda] , \\ r + \lambda , & r < -\lambda . \end{cases}$$

Therefore, the Yosida regularization of the sign function is

$$\beta_\lambda(r) = \frac{r - J_\lambda(r)}{\lambda} = \begin{cases} 1 , & r > \lambda , \\ \frac{r}{\lambda} , & r \in [-\lambda, \lambda] , \\ -1 , & r < -\lambda . \end{cases}$$

(iii) Let K be a closed convex subset of a Hilbert space H, let $\varphi = I_K$. Its Moreau regularization is given by

$$\varphi_\lambda(u) = \min_{v \in H} \left(\frac{1}{2\lambda} \|v-u\|^2 + I_K(v) \right) = \frac{1}{2\lambda} (\mathrm{dist}\,(u,K))^2 .$$

According to Lemma 1.17, $\partial I_K(u) = N_K(u)$, the normal cone. It follows that

$$\begin{aligned} v = J_\lambda(u) \quad &\Leftrightarrow \quad u \in (I + \lambda N_K)(v) \quad \Leftrightarrow \quad \frac{1}{\lambda}(u-v) \in N_K(v) \\ &\Leftrightarrow \quad \frac{1}{\lambda} \langle u-v, v-z \rangle \geq 0 \qquad \text{for all } z \in K \\ &\Leftrightarrow \quad v = P_K(u) . \end{aligned}$$

In particular, $J_\lambda = P_K$ does not depend on λ. The Yosida regularization of the normal cone mapping $u \mapsto N_K(u)$ becomes

$$(N_K)_\lambda(u) = \frac{1}{\lambda}(u - P_K(u)) .$$

$\square$

Let A be m-accretive. According to Proposition 2.10 and Lemma 2.2(ii), the sets Av are closed and convex. Denoting by $P_{Av} : H \to H$ the projection onto Av, we define $A^0 : D(A) \to H$ by

$$A^0 v = P_{Av}(0) . \tag{2.16}$$

Thus, $A^0 v$ is the norm-minimal element of Av, $\|A^0 v\| \leq \|w\|$ for all $w \in Av$.

Proposition 2.16 *Let $A : H \rightrightarrows H$ be m-accretive.*
(i) For all $\lambda > 0$, its Yosida regularization A_λ is m-accretive.
(ii) We have

$$A_\lambda v \in A J_\lambda v \,, \quad \textit{for all } v \in H, \tag{2.17}$$

$$\|A_\lambda u - A_\lambda v\| \leq \frac{1}{\lambda} \|u - v\| \,, \quad \textit{for all } u, v \in H, \tag{2.18}$$

$$(A_\lambda)_\mu = A_{\lambda+\mu} \,, \quad \textit{for all } \lambda, \mu > 0. \tag{2.19}$$

(iii) For all $v \in D(A)$ we have

$$\|A_\lambda v\| \uparrow \|A^0 v\| \,, \quad A_\lambda v \to A^0 v \,, \quad \textit{for } \lambda \downarrow 0, \tag{2.20}$$

and

$$\|A_\lambda v - A^0 v\|^2 \leq \|A^0 v\|^2 - \|A_\lambda v\|^2 . \tag{2.21}$$

(iv) For all $v \notin D(A)$ we have

$$\|A_\lambda v\| \uparrow +\infty . \tag{2.22}$$

Proof: (ii) For $J_\lambda v = (I + \lambda A)^{-1} v$ we have

$$v \in (I + \lambda A)(J_\lambda v) = J_\lambda v + \lambda A(J_\lambda v) .$$

This implies (2.17) as $A_\lambda = (I - J_\lambda)/\lambda$. For $u, v \in H$ we compute, using (2.17) for the last estimate,

$$\begin{aligned}\|A_\lambda v - A_\lambda u\|\|v-u\| &\geq \langle A_\lambda v - A_\lambda u, v - u\rangle \\ &= \langle A_\lambda v - A_\lambda u, \lambda A_\lambda v - \lambda A_\lambda u\rangle + \langle A_\lambda v - A_\lambda u, J_\lambda v - J_\lambda u\rangle \\ &\geq \lambda\|A_\lambda v - A_\lambda u\|^2 .\end{aligned}$$

This proves (2.18) and that A_λ is monotone (hence, accretive). In order to prove (2.19) we first note that

$$\begin{aligned} w = A_\lambda v \quad &\Leftrightarrow \quad v - \lambda w = J_\lambda v = (I+\lambda A)^{-1} v \\ &\Leftrightarrow \quad v \in (I+\lambda A)(v - \lambda w) \\ &\Leftrightarrow \quad v \in v - \lambda w + \lambda A(v-\lambda w) \\ &\Leftrightarrow \quad w \in A(v - \lambda w) .\end{aligned}$$

It follows that

$$\begin{aligned} w = A_{\mu+\lambda} v \quad &\Leftrightarrow \quad w \in A(v - \mu w - \lambda w) \quad \Leftrightarrow \quad w = A_\lambda(v - \mu w) \\ &\Leftrightarrow \quad w = (A_\lambda)_\mu v .\end{aligned}$$

(i) As A_λ is accretive by (ii), by Proposition 2.10 it suffices to show that $I + (\lambda/2)A_\lambda$ is surjective. Now

$$w = \big(I + \frac{\lambda}{2}A_\lambda\big)v \qquad \Leftrightarrow \qquad v = T_w(v) := w - \frac{\lambda}{2}A_\lambda(v) .$$

Since T_w is a contraction on H by (2.18), for every $w \in H$ these equations have a solution $v \in H$.

(iii) At first, let $v \in H$ be arbitrary. As A_μ is accretive and single-valued, setting $J^\mu_\lambda = (I + \lambda A_\mu)^{-1}$ we have $(A_\mu)_\lambda v = A_\mu J^\mu_\lambda v$ by (2.17) applied to A_μ, and

$$0 \leq \langle A_\mu v - (A_\mu)_\lambda v, v - J^\mu_\lambda v\rangle = \lambda \langle A_\mu v - (A_\mu)_\lambda v, (A_\mu)_\lambda v\rangle \; .$$

Consequently, by (2.19)

$$\|A_{\lambda+\mu}v\|^2 = \|(A_\mu)_\lambda v\|^2 \leq \langle A_\mu v, A_{\lambda+\mu} v\rangle \; , \quad \|A_{\lambda+\mu}v\| \leq \|A_\mu v\| . \tag{2.23}$$

In order to prove (iii), let now $v \in D(A)$. Since $A^0 v \in Av$ and $A_\lambda v \in AJ_\lambda v$, we get as before

$$0 \leq \left\langle A^0 v - A_\lambda v, v - J_\lambda v\right\rangle = \lambda \left\langle A^0 v - A_\lambda v, A_\lambda v\right\rangle \, ,$$

so

$$\|A_\lambda v\|^2 \leq \left\langle A^0 v, A_\lambda v\right\rangle \, , \quad \|A_\lambda v\| \leq \|A^0 v\| . \tag{2.24}$$

It follows from (2.23) and (2.24) that $\{A_\lambda v\}$ is bounded in H and $\|A_\lambda v\| \uparrow \gamma$ as $\lambda \to 0$ for some $\gamma \geq 0$. Moreover

$$\begin{aligned}\|A_{\lambda+\mu}v - A_\lambda v\|^2 &= \|A_{\lambda+\mu}v\|^2 + \|A_\lambda v\|^2 - 2\langle A_{\lambda+\mu}v, A_\lambda v\rangle \, , \\ &\leq \|A_{\lambda+\mu}v\|^2 + \|A_\lambda v\|^2 - 2\|A_{\lambda+\mu}v\|^2 \\ &= \|A_\lambda v\|^2 - \|A_{\lambda+\mu}v\|^2 \end{aligned} \tag{2.25}$$

Since H is complete, $A_\lambda v \to w$ as $\lambda \to 0$ for some $w \in H$. Now $v - J_\lambda v = \lambda A_\lambda v$ implies that $J_\lambda v \to v$. From Lemma 2.12 we conclude that $w \in Av$. As $\|w\| \leq \|A^0 v\|$ and since $A^0 v$ is the unique element of minimum norm in Av ist, we must have $w = A^0 v$, so $\|A_\lambda v\| \uparrow \|A^0 v\|$. This proves (2.20). Passing to the limit $\lambda \to 0$ in (2.25) proves (2.21).
(iv) We have seen in (2.23) that $\|A_\lambda v\|$ is nondecreasing as $\lambda \to 0$, for all $v \in H$. Moreover, the computation (2.25) showed that if $\|A_\lambda v\|$ is bounded as $\lambda \to 0$, then $A_\lambda v \to w \in Av$ for some $w \in H$, so $v \in D(A)$. Therefore, $v \notin D(A)$ implies that $\|A_\lambda v\| \uparrow +\infty$. $\square$

3 The Bochner Integral

In this chapter, $[a, b]$ always denotes a compact interval in $\mathbb{R}$. For $A \subset [a, b]$ we denote its characteristic function by χ_A,

$$\chi_A(t) = \begin{cases} 1\,, & t \in A\,, \\ 0\,, & t \notin A\,. \end{cases} \tag{3.1}$$

Definition of the Bochner integral. It is based on the notion of a simple function, which is a generalization of the notion of a step function.

Definition 3.1 (Simple function) *Let V be a Banach space. A function $u : [a, b] \to V$ is called **simple** if it has the form*

$$u(t) = \sum_{i=1}^{n} \chi_{A_i}(t) v_i\,, \tag{3.2}$$

where $n \in \mathbb{N}$, $A_i \subset [a, b]$ measurable and $v_i \in V$ for $1 \le i \le n$.

Lemma 3.2 *Let V be a Banach space, $u : [a, b] \to V$ simple. Then there exists a unique representation of u in the form (3.2) satisfying*

$$\bigcup_i A_i = [a, b]\,, \quad A_i \cap A_j = \emptyset \text{ and } v_i \ne v_j \text{ for } i \ne j\,. \tag{3.3}$$

*It is called the **canonical representation** of u.*

Proof: Omitted. □

Definition 3.3 (Bochner measurability)
*Let V be a Banach space. A function $u : [a, b] \to V$ is called **Bochner measurable** if there exists a sequence of simple functions $u_n : [a, b] \to V$ such that*

$$\lim_{n\to\infty} u_n(t) = u(t) \tag{3.4}$$

for almost all $t \in [a, b]$.

Definition 3.4 *Let V be a Banach space, $u : [a, b] \to V$ a simple function*

$$u(t) = \sum_{i=1}^{n} \chi_{A_i}(t) v_i\,. \tag{3.5}$$

*The **Bochner integral** of u is defined as*

$$\int_a^b u(t)\,dt = \sum_{i=1}^{n} \operatorname{meas}(A_i) v_i\,. \tag{3.6}$$

For measurable $A \subset [a, b]$ we define

$$\int_A u(t)\,dt = \int_a^b \chi_A(t) u(t)\,dt\,. \tag{3.7}$$

Definition 3.4 makes sense since the value of the right side of (3.6) does not depend on which representation of u we choose.

As a direct consequence of the definition we obtain that for simple functions $u, v : [a, b] \to V$ and numbers $\alpha, \beta \in \mathbb{R}$

$$\int_a^b \alpha u(t) + \beta v(t)\,dt = \alpha \int_a^b u(t)\,dt + \beta \int_a^b v(t)\,dt\,, \tag{3.8}$$

as well as

$$\left\| \int_a^b u(t)\,dt \right\| \le \int_a^b \|u(t)\|\,dt\,. \tag{3.9}$$

Lemma 3.5 *Let V be a Banach space, $u_n : [a, b] \to V$ a sequence of simple functions satisfying $u_n \to u$ almost everywhere. Then for every $n \in \mathbb{N}$ the function $f : [a, b] \to \mathbb{R}$ defined by*

$$f(t) = \|u_n(t) - u(t)\| \tag{3.10}$$

is measurable.

Proof: We have

$$f(t) = \lim_{m\to\infty} f_m(t)\,, \quad f_m(t) := \|u_n(t) - u_m(t)\|\,, \tag{3.11}$$

and f_m is a simple function for all $m \in \mathbb{N}$. □

Let now $u_n : [a, b] \to V$ be a sequence of simple functions with $u_n \to u$ pointwise a.e., satisfying

$$\lim_{n\to\infty} \int_a^b \|u_n(t) - u(t)\|\,dt = 0\,. \tag{3.12}$$

(By Lemma 3.5 the integrand is measurable.) Due to

$$\begin{aligned} \left\| \int_a^b u_n(t)\,dt - \int_a^b u_m(t)\,dt \right\| &\le \int_a^b \|u_n(t) - u_m(t)\|\,dt \\ &\le \int_a^b \|u_n(t) - u(t)\|\,dt + \int_a^b \|u_m(t) - u(t)\|\,dt\,, \end{aligned} \tag{3.13}$$

setting

$$y_n = \int_a^b u_n(t)\,dt \tag{3.14}$$

we obtain a Cauchy sequence $\{y_n\}$ in V. If $v_n : [a, b] \to V$ defines another sequence with the same properties as $\{u_n\}$,

$$\begin{aligned} \left\| \int_a^b v_n(t)\,dt - \int_a^b u_n(t)\,dt \right\| &\le \int_a^b \|v_n(t) - u_n(t)\|\,dt \\ &\le \int_a^b \|v_n(t) - u(t)\|\,dt + \int_a^b \|u_n(t) - u(t)\|\,dt\,. \end{aligned} \tag{3.15}$$

Therefore, the limit of $\{y_n\}$ does not depend on the choice of the sequence $\{u_n\}$.

Definition 3.6 (Bochner integral)
A function $u : [a,b] \to V$ is called ***Bochner integrable*** *if there exists a sequence of simple functions $u_n : [a,b] \to V$ such that $u_n \to u$ pointwise a.e. and*

$$\lim_{n\to\infty} \int_a^b \|u_n(t) - u(t)\| \, dt = 0 \,, \tag{3.16}$$

In this case, the ***Bochner integral*** *of u is defined as*

$$\int_a^b u(t) \, dt = \lim_{n\to\infty} \int_a^b u_n(t) \, dt \,. \tag{3.17}$$

Lemma 3.7 *Let V be a Banach space, let $u, v : [a,b] \to V$ Bochner integrable and $\alpha, \beta \in \mathbb{R}$. Then $\alpha u + \beta v$ is Bochner integrable, and*

$$\int_a^b \alpha u(t) + \beta v(t) \, dt = \alpha \int_a^b u(t) \, dt + \beta \int_a^b v(t) \, dt \,. \tag{3.18}$$

Proof: This follows directly from the definitions. □

Proposition 3.8 *Let V be a Banach space. A function $u : [a,b] \to V$ is Bochner integrable if and only if u is Bochner measurable and the function $t \mapsto \|u(t)\|$ is integrable. In this case,*

$$\left\| \int_a^b u(t) \, dt \right\| \le \int_a^b \|u(t)\| \, dt \,. \tag{3.19}$$

Proof: "⇒": Let (u_n) be a sequence of simple functions with $u_n \to u$ pointwise a.e. and

$$\int_a^b \|u(t) - u_n(t)\| \, dt = 0 \,. \tag{3.20}$$

Since $\|u_n(t)\| \to \|u(t)\|$ for a.e. $t \in [a,b]$, the function $t \mapsto \|u(t)\|$ is measurable. Then

$$\int_a^b \|u(t)\| \, dt \le \int_a^b \|u(t) - u_n(t)\| \, dt + \int_a^b \|u_n(t)\| \, dt < \infty \,. \tag{3.21}$$

"⇐": Let (u_n) be a sequence of simple functions with $u_n \to u$ pointwise almost everywhere. For any given $\varepsilon > 0$ we define $v_n : [a,b] \to V$ by

$$v_n(t) = \begin{cases} u_n(t) \,, & \text{if } \|u_n(t)\| \le (1+\varepsilon)\|u(t)\| \,, \\ 0 \,, & \text{otherwise.} \end{cases} \tag{3.22}$$

v_n is a simple function, since $\{t : \|u_n(t)\| \le (1+\varepsilon)\|u(t)\|\}$ is measurable. For

$$f_n(t) = \|v_n(t) - u(t)\| \tag{3.23}$$

we have $f_n \to 0$ pointwise a.e. and

$$0 \le f_n(t) \le (2+\varepsilon)\|u(t)\| \,. \tag{3.24}$$

From Lebesgue's theorem on dominated convergence we obtain

$$\lim_{n\to\infty} \int_a^b \|v_n(t) - u(t)\| \, dt = \lim_{n\to\infty} \int_a^b f_n(t) \, dt = 0 \,. \tag{3.25}$$

Therefore, u is Bochner integrable. With (3.9) it follows that

$$\left\| \int_a^b v_n(t) \, dt \right\| \le \int_a^b \|v_n(t)\| \, dt \le (1+\varepsilon) \int_a^b \|u(t)\| \, dt \tag{3.26}$$

and thus

$$\left\| \int_a^b u(t) \, dt \right\| = \left\| \lim_{n\to\infty} \int_a^b v_n(t) \, dt \right\| \le (1+\varepsilon) \int_a^b \|u(t)\| \, dt \,. \tag{3.27}$$

Since $\varepsilon > 0$ was arbitrary, (3.19) follows. $\square$

Function spaces. We now consider functions $u : [a,b] \to V$ for which

$$\int_a^b \|u(t)\|^p \, dt < \infty \tag{3.28}$$

holds.

Definition 3.9 *Let V be a Banach space, $1 \le p < \infty$. We define*

$$L^p(a,b;V) = \{[u] \mid u : [a,b] \to V \textit{ is Bochner measurable and satisfies (3.28) } \} \,. \tag{3.29}$$

Here, $[u]$ denotes the equivalence class of u with respect to the equivalence relation

$$u \sim v \quad \Leftrightarrow \quad u = v \textit{ almost everywhere } . \tag{3.30}$$

Due to Proposition 3.8, $L^1(a,b;V)$ coincides with the vector space of all Bochner integrable functions on $[a,b]$.

Proposition 3.10 *Let V be a Banach space, $1 \le p < \infty$. The space $L^p(a,b;V)$ is a Banach space when equipped with the norm*

$$\|u\|_{L^p(a,b;V)} = \left(\int_a^b \|u(t)\|_V^p \, dt \right)^{\frac{1}{p}} \,. \tag{3.31}$$

If V is a Hilbert space, then $L^2(a,b;V)$ becomes a Hilbert space when equipped with the scalar product

$$\langle u, v \rangle = \int_a^b \langle u(t), v(t) \rangle_V \, dt \,. \tag{3.32}$$

Proof: Omitted. For a given Cauchy sequence, one constructs a limit in the same way as in the scalar case $V = \mathbb{R}$. In order to prove that this limit is Bochner measurable, one uses a characterization of measurability due to Pettis. $\square$

Definition 3.11 *For $u : [a,b] \to \mathbb{R}$ we define*

$$\operatorname*{ess\,sup}_{t\in[a,b]} u(t) = \inf\{M : M \in \mathbb{R},\ u(t) \le M \textit{ for almost all } t \in [a,b]\} \,. \tag{3.33}$$

We now consider functions $u : [a,b] \to V$ with values in a Banach space V for which

$$\operatorname*{ess\,sup}_{t\in[a,b]} \|u(t)\|_V < \infty\,. \tag{3.34}$$

Definition 3.12 *Let V be a Banach space. We define*

$$L^\infty(a,b;V) = \{[u] \,|\, u : [a,b] \to V \text{ is Bochner measurable and (3.34) holds}\}\,. \tag{3.35}$$

Proposition 3.13 *Let V be a Banach space. Then $L^\infty(a,b;V)$ is a Banach space.*

Proof: Again, this is proved in the same manner as in the case $V = \mathbb{R}$. □

Lemma 3.14 *Let V be a Banach space. Then for all $1 \le p \le q \le \infty$ we have*

$$L^q(a,b;V) \subset L^p(a,b;V)\,. \tag{3.36}$$

Proof: As in the case $V = \mathbb{R}$. □

Definition 3.15 *Let V be a Banach space. We define*

$$C([a,b];V) = \{u \,|\, u : [a,b] \to V \text{ continuous}\}\,. \tag{3.37}$$

Definition 3.16 (Oscillation)
Let V be a Banach space, $u : [a,b] \to V$. We define the oscillation of u by

$$\operatorname*{osc}_{[a,b]}(u;\delta) = \sup\{\|u(t) - u(s)\| : s,t \in [a,b],\ |t-s| \le \delta\}\,. \tag{3.38}$$

Lemma 3.17 *Let V be a Banach space, $u : [a,b] \to V$ continuous. Then*

$$\lim_{\delta\to 0} \operatorname*{osc}_{[a,b]}(u;\delta) = 0\,. \tag{3.39}$$

Proof: The statement (3.39) is equivalent to the uniform continuity of u. □

For a continuous function $u : [a,b] \to V$ we have

$$\operatorname*{ess\,sup}_{t\in[a,b]} \|u(t)\| = \max_{t\in[a,b]} \|u(t)\|\,, \tag{3.40}$$

since continuity of u implies that $\|u(t)\| \le \operatorname{ess\,sup}_{s\in[a,b]} \|u(s)\|$ holds for all t.

Proposition 3.18 *Let V be a Banach space. Then $C([a,b];V)$ is a Banach space when equipped with the norm*

$$\|u\|_{C([a,b];V)} = \max_{t\in[a,b]} \|u(t)\|\,. \tag{3.41}$$

Moreover, $C([a,b];V)$ can be identified with a closed subspace of $L^\infty(a,b;V)$.

Proof: If $u : [a,b] \to V$ is continuous, it is Bochner measurable: We define a sequence of simple functions $u_n : [a,b] \to V$ by $u_n(t) = u(ih)$, if $t \in [ih,(i+1)h)$, $h = (b-a)/n$ is. Then

$$\|u_n(t) - u(t)\| \le \operatorname*{osc}_{[a,b]}(u;h)\,, \quad h = \frac{b-a}{n}\,, \tag{3.42}$$

therefore $u_n \to u$ uniformly (pointwise convergence would already be sufficient for our purpose). Moreover: Let (u_n) be a sequence in C satisfying $[u_n] \to [u]$ in L^∞. We choose a subset N in $[a,b]$ of zero measure such that $u_n \to u$ uniformly in $M = [a,b] \setminus N$. Then u is continuous on M. For an arbitrary given $t \in N$ we choose a sequence $(t_k)_{k\in\mathbb{N}}$ in M such that $t_k \to t$. Then

$$\|u_n(t) - u_m(t)\| \le \|u_n(t) - u_n(t_k)\| + \|u_n - u_m\|_{L^\infty(a,b;V)} + \|u_m(t_k) - u_m(t)\|\,,$$

thus $(u_n(t))_{n\in\mathbb{N}}$ is a Cauchy sequence. For $t \in N$, let $\tilde{u}(t)$ be the limit of this Cauchy sequence, and set $\tilde{u}(t) = u(t)$ for $t \in M$. Then $\tilde{u} : [a,b] \to V$ is continuous and $[\tilde{u}] = [u]$. $\square$

Derivatives. For functions $u : [a,b] \to V$, the limit, if it exists,

$$u'(t) = \lim_{h\to 0} \frac{u(t+h) - u(t)}{h}$$

is called the derivative of u at t.

Definition 3.19 *Let V be a Banach space, $u \in L^1(a,b;V)$. A function $f \in L^1(a,b;V)$ is called a* ***weak derivative*** *of u if*

$$\int_a^b u(t)\psi'(t)\,dt = -\int_a^b f(t)\psi(t)\,dt \qquad \text{for all } \psi \in C_0^\infty(a,b). \tag{3.43}$$

As in the scalar case one can show that the weak derivative (if it exists) is unique.

Definition 3.20 *Let V be a Banach space, $1 \le p \le \infty$. The Sobolev space $W^{1,p}(a,b;V)$ is defined as the space of all functions $u \in L^p(a,b;V)$ whose weak derivative exists and is an element of $L^p(a,b;V)$.*

Proposition 3.21 *Let V be a Banach space, let $u \in L^p(a,b;V)$, $1 \le p \le \infty$. Then there are equivalent:*
(i) $u \in W^{1,p}(a,b;V)$.
(ii) There exists a function $\tilde{u} : [a,b] \to V$ with $\tilde{u} = u$ a.e., which is differentiable a.e. in (a,b), whose pointwise derivative $\tilde{u}'$ belongs to $L^p(a,b;V)$ and which satisfies

$$\tilde{u}(t) - \tilde{u}(a) = \int_a^t \tilde{u}'(s)\,ds \qquad \text{for all } t \in [a,b]. \tag{3.44}$$

In this case, $\tilde{u}'$ is a.e. equal to the weak derivative of u.

For the proof we refer to [5], Proposition A.6, and [2], Theorem 1.17.

Consequently, one writes u' both for the weak derivative and for the pointwise a.e. derivative.

Proposition 3.22 *The Sobolev space $W^{1,p}(a,b;V)$ becomes a Banach space when equipped with the norm*

$$\|u\|_{W^{1,p}(a,b;V)} = \|u\|_{L^p(a,b;V)} + \|u'\|_{L^p(a,b;V)} \,.$$

If V is a Hilbert space, $W^{1,2}(a,b;V)$ becomes a Hilbert space with the scalar product

$$\langle u, v\rangle = \int_a^b \langle u(t), v(t)\rangle \, dt + \int_a^b \langle u'(t), v'(t)\rangle \, dt \,.$$

Proof: Omitted. $\square$

4 Evolution Equations

In this chapter, H always denotes a real Hilbert space.

We investigate the initial value problem

$$u' + Au = 0\,, \quad u(0) = u_0\,, \tag{4.1}$$

as well as the corresponding integral equation

$$u(t) = u_0 - \int_0^t Au(s)\,ds\,. \tag{4.2}$$

Proposition 4.1 *Let $A : H \to H$ be accretive and Lipschitz continuous with $D(A) = H$.*
(i) The integral equation (4.2) has for every $u_0 \in H$ a unique solution $u \in C([0,\infty); H)$ which moreover is continuously differentiable.
(ii) Let u and $\hat{u}$ be solutions of (4.2) for the initial values u_0 and $\hat{u}_0$. Then the function $t \mapsto \|u(t) - \hat{u}(t)\|$ is nonincreasing, and in particular

$$\|u(t) - \hat{u}(t)\| \le \|u_0 - \hat{u}_0\|\,. \tag{4.3}$$

The function $t \mapsto \|u'(t)\|$ is nonincreasing.

Proof: (i) Unique solvability in $C([0,T]; H)$ (and thus, in $C([0,\infty); H)$) is a consequence of Banach's fixed point theorem. Since $t \mapsto Au(t)$ is continuous, u' is continuous, too.
(ii) Since A is accretive,

$$\frac{\mathrm{d}}{\mathrm{dt}}\frac{1}{2}\|u(t) - \hat{u}(t)\|^2 = \langle u'(t) - \hat{u}'(t), u(t) - \hat{u}(t)\rangle = -\langle Au(t) - A\hat{u}(t), u(t) - \hat{u}(t)\rangle \le 0\,.$$

Setting $\hat{u}(t) = u(t+h)$ for given $h > 0$ we see that the function $t \mapsto \|u(t+h) - u(t)\|$ is nonincreasing. Dividing by h and passing to the limit $h \to 0$ shows that $t \mapsto \|u'(t)\|$ is nonincreasing. □

Lemma 4.2 *Let H be a Hilbert space, $A : H \rightrightarrows H$ accretive (resp. m-accretive), let $\tilde{A} : L^2(0,T;H) \rightrightarrows L^2(0,T;H)$ be defined by*

$$w \in \tilde{A}v \qquad \Leftrightarrow \qquad w(t) \in A(v(t)) \quad \textit{a.e. in } [0,T]. \tag{4.4}$$

Then $\tilde{A}$ is accretive (resp. m-accretive).

Proof: Let A be accretive, let $v_1, v_2 \in L^2(0,T;H)$, $w_2 \in \tilde{A}v_2$, $w_1 \in \tilde{A}v_1$. Then $w_i(t) \in A(v_i(t))$ for a.a. t. Therefore

$$\langle w_2 - w_1, v_2 - v_1\rangle_{L^2(0,T;H)} = \int_0^T \langle w_2(t) - w_1(t), v_2(t) - v_1(t)\rangle_H \, dt \ge 0\,.$$

Thus $\tilde{A}$ is accretive. Let now A be m-accretive. We have to show that $(I + \tilde{A})u = w$ has a solution $u \in L^2(0,T;H)$ for every $w \in L^2(0,T;H)$. Given $w \in L^2(0,T;H)$, we define $u : [0,T] \to H$ by

$$u(t) = (I + A)^{-1}(w(t))\,.$$

As $(I+A)^{-1}$ is nonexpansive and w is Bochner measurable, u is Bochner measurable, too. Setting $v_0 = (I+A)^{-1}(0)$ we get

$$\|u(t) - v_0\|_H \le \|w(t) - 0\|_H = \|w(t)\|_H \quad \text{a.e. in } [0,T].$$

Therefore, $u \in L^2(0,T;H)$ and $u + \tilde{A}u \ni w$. □

We now consider the initial value problem

$$u' + Au \ni 0\,, \quad u(0) = u_0\,, \tag{4.5}$$

in the case that $A : H \rightrightarrows H$ is m-accretive. We recall that $A^0 v$ denotes the unique element of minimal norm in Av. First we present an existence result. This goes back to Kato [9] and Komura [10, 11].

Proposition 4.3 *Let $A : H \rightrightarrows H$ be m-accretive, let $u_0 \in D(A)$. Then there exists a Lipschitz continuous solution $u : [0,\infty) \to H$ of (4.5) with $u(t) \in D(A)$ for all $t > 0$. Its weak derivative u' belongs to $L^\infty(0,\infty;H)$ and satisfies*

$$u'(t) + Au(t) \ni 0 \quad \textit{a.e. in } (0,\infty), \quad u(0) = u_0\,, \tag{4.6}$$

$$\|u'\|_{L^\infty((0,\infty);H)} \le \|A^0 u_0\|\,, \tag{4.7}$$

$$\|A^0 u(t)\| \le \|A^0 u_0\| \quad \textit{for all } t > 0. \tag{4.8}$$

Proof: We consider the auxiliary problem

$$u' + A_\lambda u = 0\,, \quad u(0) = u_0\,, \tag{4.9}$$

where we have replaced A by its Yosida regularization A_λ. By Proposition 4.1, (4.9) has a continuously differentiable solution $u_\lambda : [0,\infty) \to H$ which satisfies

$$\|A_\lambda u_\lambda(t)\| = \|u_\lambda'(t)\| \le \|u_\lambda'(0)\| = \|A_\lambda u_\lambda(0)\| = \|A_\lambda u_0\| \le \|A^0 u_0\|\,, \tag{4.10}$$

the last inequality was proved in Proposition 2.16. We want to show that $\{u_\lambda\}_\lambda$ has the Cauchy property in $C([0,T];H)$ for every $T > 0$, that is,

$$\lim_{\lambda,\mu\to 0} \max_{t\in[0,T]} \|u_\lambda(t) - u_\mu(t)\| = 0\,. \tag{4.11}$$

For $\lambda, \mu > 0$ we have (we omit the argument "t")

$$u_\lambda' - u_\mu' + A_\lambda u_\lambda - A_\mu u_\mu = 0\,.$$

We multiply by $u_\lambda - u_\mu$ and get

$$\frac{\mathrm{d}}{\mathrm{dt}}\frac{1}{2}\|u_\lambda - u_\mu\|^2 + \langle A_\lambda u_\lambda - A_\mu u_\mu, u_\lambda - u_\mu\rangle = 0\,. \tag{4.12}$$

We decompose

$$u_\lambda - u_\mu = \lambda A_\lambda u_\lambda + J_\lambda u_\lambda - \mu A_\mu u_\mu - J_\mu u_\mu\,. \tag{4.13}$$

By Proposition 2.16 we have $A_\lambda v \in AJ_\lambda v$ for all $v \in H$. Since A is accretive, it follows from (4.13) that

$$\langle A_\lambda u_\lambda - A_\mu u_\mu, u_\lambda - u_\mu\rangle \geq \langle A_\lambda u_\lambda - A_\mu u_\mu, \lambda A_\lambda u_\lambda - \mu A_\mu u_\mu\rangle$$

and furthermore, using (4.10)

$$\begin{aligned} &\geq \lambda\|A_\lambda u_\lambda\|^2 + \mu\|A_\mu u_\mu\|^2 - (\lambda+\mu)\|A_\lambda u_\lambda\|\|A_\mu u_\mu\| \\ &\geq -(\lambda+\mu)\|A_\lambda u_\lambda\|\|A_\mu u_\mu\| \geq -(\lambda+\mu)\|A^0 u_0\|^2 . \end{aligned}$$

Therefore, (4.12) implies

$$\frac{\mathrm{d}}{\mathrm{dt}}\|u_\lambda - u_\mu\|^2 \leq 2(\lambda+\mu)\|A^0 u_0\|^2 .$$

Integrating over $[0,t]$ and taking the square root yields

$$\|u_\lambda(t) - u_\mu(t)\| \leq \sqrt{2(\lambda+\mu)t}\|A^0 u_0\| .$$

Thus, (4.11) is proved. As $C([0,T];H)$ is complete, there exists a $u \in C([0,T];H)$ such that $u_\lambda \to u$ and

$$\|u_\lambda(t) - u(t)\| \leq \sqrt{2\lambda t}\|A^0 u_0\| . \tag{4.14}$$

Moreover, $J_\lambda u_\lambda \to u$ in $C([0,T];H)$ as $\lambda \to 0$, since

$$\|J_\lambda u_\lambda(t) - u_\lambda(t)\| = \lambda\|A_\lambda u_\lambda(t)\| \leq \lambda\|A^0 u_0\| .$$

Let now $t \in [0,T]$ be arbitrary. Since $\{A_\lambda u_\lambda(t)\}_\lambda$ is bounded by (4.10), there exists a weakly convergent subsequence $\{A_{\lambda_k} u_{\lambda_k}(t)\}_k \rightharpoonup z \in H$. As $A_{\lambda_k} u_{\lambda_k}(t) \in AJ_{\lambda_k} u_{\lambda_k}(t)$, Lemma 2.12 implies that $z \in Au(t)$. Thus $u(t) \in D(A)$ and

$$\|A^0 u(t)\| \leq \|z\| \leq \|A^0 u_0\| . \tag{4.15}$$

This proves (4.8). Since

$$\|u_\lambda'(t)\| = \|A_\lambda u_\lambda(t)\| \leq \|A^0 u_0\| , \tag{4.16}$$

$\{u_\lambda'\}$ is bounded in $L^\infty((0,T);H)$, so in $L^2(0,T;H)$, too. Therefore, for some subsequence and some $w \in L^2(0,T;H)$ we have

$$u_{\lambda_k}' \rightharpoonup w \text{ in } L^2(0,T;H), \quad u_{\lambda_k} \to u \text{ in } C([0,T];H).$$

Passing to the limit in

$$u_{\lambda_k}(t) = u_0 + \int_0^t u_{\lambda_k}'(s)\,ds$$

gives

$$u(t) = u_0 + \int_0^t w(s)\,ds .$$

According to Proposition 3.21, u has a derivative (weak and pointwise a.e.), namely $u' = w$. As the derivative is uniquely determined, it follows that $u_\lambda' \rightharpoonup u'$ in $L^2(0,T;H)$.

In order to prove (4.7) we observe that the subset $K = \{v : \|v(t)\|_H \le \|A^0 u_0\|_H \text{ a.e.}\}$ of $L^2(0,T;H)$ is convex and closed, hence weakly closed in $L^2(0,T;H)$. Since $u'_\lambda \in K$ and $u'_\lambda \rightharpoonup u'$, it follows that $u' \in K$. This proves (4.7).

It remains to prove that $-u'(t) \in Au(t)$ a.e. in $(0,T)$. Let $\tilde{A} : L^2(0,T;H) \to L^2(0,T;H)$ be the mapping from Lemma 4.2. As $A_\lambda(u_\lambda(t)) \in AJ_\lambda(u_\lambda(t))$ for all $t > 0$, we get

$$A_\lambda \circ u_\lambda \in \tilde{A}(J_\lambda \circ u_\lambda)\,.$$

Since $\tilde{A}$ is m-accretive and $A_\lambda \circ u_\lambda \rightharpoonup -u'$ as well as $J_\lambda \circ u_\lambda \to u$ hold in $L^2(0,T;H)$, we obtain $-u' \in \tilde{A}u$ by Lemma 2.12, so $-u'(t) \in Au(t)$ a.e. in $(0,T)$. □

The next proposition provides uniqueness.

Proposition 4.4 *Let $A : H \rightrightarrows H$ be accretive, let $f, \hat{f} \in L^1(0,T;H)$, let $u, \hat{u} \in W^{1,1}(0,T;H)$ solutions of*

$$u' + Au \ni f\,, \quad \hat{u}' + A\hat{u} \ni \hat{f}\,, \tag{4.17}$$

on some interval $[s,t]$. Then

$$\|u(t) - \hat{u}(t)\| \le \|u(s) - \hat{u}(s)\| + \int_s^t \|f(\tau) - \hat{f}(\tau)\|\, d\tau\,. \tag{4.18}$$

Proof: Since A is accretive, we obtain a.e. on $[s,t]$ the estimate

$$\begin{aligned}\|u(\tau) - \hat{u}(\tau)\| \frac{\mathrm{d}}{\mathrm{d}\tau} \|u(\tau) - \hat{u}(\tau)\| &= \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{1}{2}\|u(\tau) - \hat{u}(\tau)\|^2 = \langle u(\tau) - \hat{u}(\tau), u'(\tau) - \hat{u}'(\tau)\rangle \\ &\le \left\langle u(\tau) - \hat{u}(\tau), f(\tau) - \hat{f}(\tau)\right\rangle \le \|u(\tau) - \hat{u}(\tau)\| \|f(\tau) - \hat{f}(\tau)\|\,.\end{aligned}$$

Therefore,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\|u(\tau) - \hat{u}(\tau)\| \le \|f(\tau) - \hat{f}(\tau)\|$$

holds a.e. on $[s,t]$. Integrating over $[s,t]$ yields (4.18). □

Here are some further properties of the solution of $u' + Au \ni 0$.

Proposition 4.5 *In the situation of Proposition 4.3 the following holds:*
(i) The solution u has a right derivative $u'_+(t)$ in every point $t \ge 0$ which satisfies

$$u'_+(t) + A^0 u(t) = 0\,. \tag{4.19}$$

(ii) The function $t \mapsto A^0 u(t)$ is right-continuous.
(iii) The function $t \mapsto \|A^0 u(t)\|$ is nonincreasing.

Proof: (ii) We first show that $t \mapsto A^0 u(t)$ is right-continuous at 0. Let $t_n \downarrow 0$ such that $A^0 u(t_n) \rightharpoonup \xi$. As $u(t_n) \to u_0$ we have $\xi \in Au_0$ by Lemma 2.12. Moreover, using (4.8),

$$\|\xi\| \le \liminf_{n\to\infty} \|A^0 u(t_n)\| \le \|A^0 u_0\|\,.$$

It follows that $\xi = A^0 u_0$. As therefore ξ does not depend upon the choice of the sequence $\{t_n\}$, we see that $A^0 u(t) \rightharpoonup A^0 u_0$ for $t \downarrow 0$. Using again that $\|A^0 u(t)\| \le \|A^0 u_0\|$ we

conclude that $A^0 u(t) \to A^0 u_0$ as $t \downarrow 0$. Let now $t_0 > 0$ be arbitrary. The translate $t \mapsto u(t_0 + t)$ uniquely solves the initial value problem for the initial value $u(t_0)$; by what we have just proved, u is right-continuous at t_0.
(iii) Again considering the translate just mentioned we see that $\|A^0 u(t_0 + t)\| \le \|A^0 u(t_0\|$ for all $t > 0$ by (4.8).
(i) Again using the translate we see that for all $t \ge 0$, $h > 0$

$$\frac{1}{h}\|u(t+h) - u(t)\| = \left\| \frac{1}{h} \int_t^{t+h} u'(s)\, ds \right\| \le \|A^0 u(t)\| \,.$$

For $h \to 0$ the middle term converges to $\|u'_+(t)\|$ a.e. in t. Thus

$$\|u'(t)\| \le \|A^0 u(t)\| \,, \quad -u'(t) \in Au(t) \,, \quad \text{a.e. in } t.$$

Therefore, $-u'(t) = A^0 u(t)$ a.e. in t. This implies

$$\frac{u(t+h) - u(t)}{h} + \frac{1}{h} \int_t^{t+h} A^0 u(s)\, ds = 0$$

for all $t \ge 0$, $h > 0$. By (ii), the second term converges to $A^0 u(t)$ as $h \to 0$. Therefore, $u'_+(t)$ exists for all t, and (iii) is proved. □

We now consider the initial value problem with nonzero right side,

$$u' + Au \ni f(t) \,, \quad u(0) = u_0 \,. \tag{4.20}$$

Proposition 4.6 *Let $A : H \rightrightarrows H$ be m-accretive, let $u_0 \in D(A)$ and $f \in W^{1,1}(0,T;H)$. Then there exists a unique $u \in W^{1,1}(0,T;H)$ which solves (4.20) a.e. in $(0,T)$. Moreover, u is Lipschitz continuous and, for all $t \ge 0$, $u(t) \in D(A)$ and $u'_+(t)$ exists.*

Proof: One considers solutions u_λ of the auxiliary problems

$$u' + A_\lambda u = f(t) \,, \quad u(0) = u_0 \,.$$

and modifies the proof of Proposition 4.3 in a suitable manner. We refer to [20], Theorem IV.4.1. □

In order to treat more general right sides and initial values, we generalize the notion of a solution on the basis of the estimate (4.18). Let $f \in L^1(0,T;H)$ and $u_0 \in \overline{D(A)}$ be given. Let $f_n \in W^{1,1}(0,T;H)$ and $u_0^n \in D(A)$ such that $f_n \to f$ in $L^1(0,T;H)$ and $u_0^n \to u_0$ in H. Let $u_n \in W^{1,1}(0,T;H)$ be the solution to f_n and u_0^n according to Proposition 4.6. Applying (4.18) with $s = 0$ and $t = T$ to solutions u_n and u_m we obtain

$$\|u_n - u_m\|_{C([0,T];H)} \le \|u_0^n - u_0^m\|_H + \|f_n - f_m\|_{L^1(0,T;H)} \,. \tag{4.21}$$

Therefore, $\{u_n\}$ is a Cauchy sequence in $C([0,T];H)$, so $u_n \to u$ for some $u \in C([0,T];H)$. This limit does not depend on the choice of the approximating sequence, so u is uniquely defined by this process. It is called the **generalized solution** of (4.20) for $f \in L^1(0,T;H)$ and $u_0 \in \overline{D(A)}$.

In the case where $A = \partial\varphi$, the generalized solution is more regular than just an element of $C([0,T];H)$, if f belongs to L^2.

Proposition 4.7 *Let $A = \partial\varphi$ where $\varphi : H \to (-\infty, \infty]$ is convex, lower semicontinuous and proper. Let $f \in L^2(0,T;H)$ and $u_0 \in \overline{D(A)}$. Then the generalized solution u of (4.20) satisfies $u(t) \in D(A)$ a.e. in $(0,T)$ as well as*

$$\varphi \circ u \in L^1(0,T)\,, \quad \sqrt{t}u' \in L^2(0,T;H)\,. \tag{4.22}$$

If moreover $u_0 \in D(\varphi)$, then

$$\varphi \circ u \in L^\infty(0,T)\,, \quad u' \in L^2(0,T;H)\,. \tag{4.23}$$

Proof: See Theorem IV.4.3 in [20]. □

Example 4.8 In a Hilbert space H we consider

$$u' + \partial\varphi(u) \ni f\,, \quad u(0) = u_0\,, \tag{4.24}$$

where $\varphi = I_K$ is the indicator function of a closed convex subset K of H. Since $\partial\varphi = N_K$, the normal cone, the differential inclusion becomes

$$f(t) - u'(t) \in N_K(u(t))\,.$$

This is equivalent to the evolution variational inequality

$$\begin{aligned} \langle u'(t), v - u(t)\rangle \geq \langle f(t), v - u(t)\rangle \qquad &\text{for all } v \in K, \\ u(t) \in K\,. \end{aligned}$$

□

Example 4.9 In $H = L^2(\Omega)$ we consider

$$u' + Au + \partial\varphi(u) \ni f\,, \quad u(0) = u_0\,, \tag{4.25}$$

where $A = -\Delta$ (or a more general elliptic operator) and again $\varphi = I_K$ with $K \subset H$ closed and convex. One can show that under certain conditions the sum $A + \partial\varphi$ is m-accretive. The differential inclusion becomes

$$f(t) - (u'(t) - \Delta u(t)) \in N_K(u(t))\,.$$

This is equivalent to the evolution variational inequality

$$\begin{aligned} \langle u'(t) - \Delta u(t), v - u(t)\rangle \geq \langle f(t), v - u(t)\rangle \qquad &\text{for all } v \in K, \\ u(t) \in K\,. \end{aligned}$$

Written more explicitly, the inequality has the form

$$\int_\Omega (\partial_t u(t,x) - \Delta u(t,x))(v(x) - u(t,x))\,dx \geq \int_\Omega f(t,x)(v(x) - u(t,x))\,dx\,, \quad \text{for all } v \in K.$$

□

5 The play and the stop operator

In this section we discuss the definition and some basic properties of the so-called play and stop operators.

Rate independence. Let $\mathcal{W}$ be an operator which maps time-dependent functions u with values in an arbitrary set X to time-dependent functions $w = \mathcal{W}[u]$ with values in some set Y. Such an operator $\mathcal{W}$ is called **rate independent** if it commutes with time transformations $\psi : [a,b] \to [a,b]$ of the underlying time interval $[a,b]$,

$$\mathcal{W}[u \circ \psi] = (\mathcal{W}[u]) \circ \psi \,. \tag{5.1}$$

The time transformations are assumed to be nondecreasing,

$$s \le t \qquad \Rightarrow \qquad \psi(s) \le \psi(t) \,,$$

and surjective; in particular, $\psi(a) = a$, $\psi(b) = b$ and ψ is continuous.

The operator $\mathcal{W}$ is said to possess the **Volterra property**, if for every t the value $w(t)$ does not depend upon the future values $u(s)$, $s > t$, of the function u. A rate-independent operator which possesses the Volterra property is called a **hysteresis operator**.

In the period 1965 to 1985, a basic theory of hysteresis operators was developed by Krasnosel'skiĭ and his group, see the monograph [12]. Other monographs in this tradition are [15, 21, 6, 13]. There is also the collection [4].

The scalar play. It arises when the diagonal $w = u$ in the u-w-plane (which represents the identity operator $w = Iu$ on functions) is split into two separate straight lines $w = u - r$ und $w = u + r$, where $r > 0$ is given. On the right line $w = u - r$ one can only ascend, on the left line one can only descend, and in the region in between w has to remain constant. If the input function u is continuous and monotone, this behaviour is described by

$$w(t) = f_r(u(t), w(a)) \,, \quad t \ge a \,,$$

where

$$f_r(u, w) = \max\{u - r, \min\{u + r, w\}\} \,. \tag{5.2}$$

If $u : [a,b] \to \mathbb{R}$ is continuous and piecewise monotone with respect to the partition $\{t_i\}$ of $[a,b]$ (that is, u is monotone on each subinterval $[t_i, t_{i+1}]$), we set

$$w(t) = f_r(u(t), w(t_i)) \,, \quad t_i < t \le t_{i+1} \,, \tag{5.3}$$

starting with

$$w(a) = f_r(u(a), w_a) \tag{5.4}$$

$w_a \in \mathbb{R}$ being a given initial value. In this manner we obtain an operator

$$w = \mathcal{P}_r[u; w_a] \,, \quad \mathcal{P}_r : C_{pm}[a,b] \times \mathbb{R} \to C_{pm}[a,b] \,. \tag{5.5}$$

Here, $C_{pm}[a,b]$ denotes the space of all continuous and piecewise monotone functions on $[a,b]$ with values in $\mathbb{R}$.

For $r = 0$ we get back the identity $\mathcal{P}_0 = I$ since $f_0(u,w) = u$.

Maximum norm estimate. The basic maximum estimate for the scalar play arises from a corresponding estimate for the function f_r.

Lemma 5.1 *We have*

$$|f_{r_1}(u_1,w_1)-f_{r_2}(u_2,w_2)| \le \max\{|u_1-u_2|+|r_1-r_2|\,,\,|w_1-w_2|\} \tag{5.6}$$

for all $r_j \ge 0$, $u_j, w_j \in \mathbb{R}$.

Proof. We have

$$|\max\{x_1,y_1\}-\max\{x_2,y_2\}| \le \max\{|x_1-x_2|,|y_1-y_2|\}\,,$$
$$|\min\{x_1,y_1\}-\min\{x_2,y_2\}| \le \max\{|x_1-x_2|,|y_1-y_2|\}\,,$$

for all $x_1,x_2,y_1,y_2 \in \mathbb{R}$. Consequently

$$|f_{r_1}(u_1,w_1)-f_{r_2}(u_2,w_2)| \le \max\{|(u_1-r_1)-(u_2-r_2)|\,,\,|(u_1+r_1)-(u_2+r_2)|\,,\,|w_1-w_2|\}\,.$$

This implies the assertion. □

Proposition 5.2 *The operator $\mathcal{P}_r$ defined in (5.2) – (5.5) can be extended uniquely to a Lipschitz continuous operator*

$$\mathcal{P}_r : C[a,b]\times\mathbb{R} \to C[a,b]\,,$$

and there holds

$$\|\mathcal{P}_{r_1}[u_1;w_{a,1}]-\mathcal{P}_{r_2}[u_2;w_{a,2}]\|_\infty \le \max\{\|u_1-u_2\|_\infty+|r_1-r_2|\,,\,|w_{a,1}-w_{a,2}|\} \tag{5.7}$$

for all $u_1,u_2 \in C[a,b]$ and all $w_{a,1},w_{a,2}\in\mathbb{R}$.

Proof. Since $C_{pm}[a,b]$ is dense in $C[a,b]$ it suffices to show that (5.7) holds for $u_1,u_2 \in C_{pm}[0,T]$. Let $\{t_i\}$ be a partition of $[a,b]$ such that both u_1 and u_2 are monotone on all subintervals $[t_i,t_{i+1}]$. By Lemma 5.1, on each subinterval, setting $w_j(t)=\mathcal{P}_{r_j}[u_j;w_{a,j}](t)$,

$$|w_1(t)-w_2(t)| \le \max\{\max_{s\in[a,t]}|u_1(s)-u_2(s)|+|r_1-r_2|\,,\,|w_1(t_i)-w_2(t_i)|\}$$

for all $t\in[t_i,t_{i+1}]$, therefore

$$|w_1(t_{i+1})-w_2(t_{i+1})| \le \max\{\|u_1-u_2\|_\infty+|r_1-r_2|\,,\,|w_1(t_i)-w_2(t_i)|\}\,.$$

The assertion now follows by induction over i. □

The scalar stop. Let $r\ge 0$. For a given input $u\in C[a,b]$ and initial value $z_a\in[-r,r]$, the stop is defined by

$$\mathcal{S}_r[u;z_a] = u-\mathcal{P}_r[u;w_a]\,, \tag{5.8}$$

where $w_a = u(a)-z_a$. Thus, the output functions

$$z=\mathcal{S}_r[u;z_a]\,,\quad w=\mathcal{P}_r[u;w_a]$$

are related by

$$u(t)=w(t)+z(t)\,,\quad t\in[a,b]\,. \tag{5.9}$$

Variation norm estimate. Let $u_1,u_2:[a,b]\to\mathbb{R}$ be piecewise linear, let

$$w_j=\mathcal{P}_r[u_j;w_{a,j}]\,,\quad z_j=\mathcal{S}_r[u_j;z_{a,j}]\,,\quad j=1,2\,,$$

such that

$$w_j(a)+z_j(a)=u_j(a)\,.$$

Lemma 5.3 *We have*

$$(\dot{w}_1(t) - \dot{w}_2(t))(z_1(t) - z_2(t)) \ge 0 \tag{5.10}$$

for all except finitely many $t \in (a, b)$.

Proof. By construction, the functions w_j and z_j are piecewise linear. Let t be a point of continuity of $\dot{w}_1$ and $\dot{w}_2$. Assume without loss of generality that $z_1(t) > z_2(t)$. Then

$$z_1(t) > -r\,, \quad z_2(t) < r\,.$$

We have two cases:

$$z_1(t) < r \quad \Rightarrow \quad \dot{w}_1(t) = 0\,, \quad z_1(t) = r \quad \Rightarrow \quad \dot{w}_1(t) \ge 0\,.$$

In both cases

$$\dot{w}_1(t)(z_1(t) - z_2(t)) \ge 0\,.$$

Analogously, $z_2(t) < r$ implies that $\dot{w}_2(t)(z_1(t) - z_2(t)) \le 0$. □

Lemma 5.4 *If u_1, u_2 are piecewise linear, we have*

$$|\dot{w}_1(t) - \dot{w}_2(t)| + \frac{\mathrm{d}}{\mathrm{dt}}|z_1(t) - z_2(t)| \le |\dot{u}_1(t) - \dot{u}_2(t)|\,. \tag{5.11}$$

for all except finitely many $t \in (a, b)$.

Proof. On intervals where $z_1 = z_2$ we have

$$w_1 - w_2 = u_1 - z_1 + z_2 - u_2 = u_1 - u_2\,,$$

thus (5.11) holds with equality. On intervals where $z_1 \ne z_2$ we obtain from (5.10) that

$$|\dot{w}_1(t) - \dot{w}_2(t)| = (\dot{w}_1(t) - \dot{w}_2(t))\mathrm{sign}\,(z_1(t) - z_2(t))\,.$$

Moreover

$$\frac{\mathrm{d}}{\mathrm{dt}}|z_1(t) - z_2(t)| = (\dot{z}_1(t) - \dot{z}_2(t))\mathrm{sign}\,(z_1(t) - z_2(t))\,.$$

Adding the previous two equations yields the assertion. □

For $u \in W^{1,1}(a, b)$ we define

$$\|u\|_{BV} = |u(a)| + \mathrm{var}\,(u) = |u(a)| + \int_a^b |\dot{u}(t)|\,dt \tag{5.12}$$

Proposition 5.5 *Let $u_1, u_2 \in W^{1,1}(a, b)$, $w_{a,1}, w_{a,2} \in \mathbb{R}$. Then $w_j = \mathcal{P}_r[u_j; w_{a,j}]$ satisfy*

$$\mathrm{var}\,(w_2 - w_1) \le \mathrm{var}\,(u_2 - u_1) + |u_2(a) - u_1(a)| + |w_2(a) - w_1(a)|\,. \tag{5.13}$$

Consequently,

$$\mathcal{P}_r : W^{1,1}(a, b) \times \mathbb{R} \to W^{1,1}(a, b)$$

is Lipschitz continuous, and the same holds for $\mathcal{S}_r$.

Proof. Assume first that u_1, u_2 are piecewise linear. Set $z_j = \mathcal{S}_r[u_j; u_j(a) - w_j(a)]$, then by virtue of (5.11)

$$\begin{aligned}\int_a^b |\dot{w}_2(t) - \dot{w}_1(t)|\, dt &\le \int_a^b |\dot{u}_2(t) - \dot{u}_1(t)|\, dt - |z_2 - z_1|\Big|_a^b \\ &\le \int_a^b |\dot{u}_2(t) - \dot{u}_1(t)|\, dt + |z_2(a) - z_1(a)|\end{aligned}$$

which proves (5.13). This and the estimate

$$|w_2(a) - w_1(a)| \le \max\{|u_2(a) - u_1(a)|, |w_{a,2} - w_{a,1}|\}$$

yields the Lipschitz continuity for piecewise linear input functions. Since those functions are dense in $W^{1,1}(a,b)$, all assertions extend to $W^{1,1}(a,b)$ □

It is no coincidence that the maximum norm and the total variation norm enter the basic estimates. Both norms are (in contrast to other usual norms) invariant w.r.t. time transformations ψ, that is,

$$\|u \circ \psi\|_\infty = \|u\|_\infty \,, \quad \|u \circ \psi\|_{1,1} = \|u\|_{1,1} \,.$$

The scalar stop as a variational inequality. Given $u : [a,b] \to \mathbb{R}$ and $z_a \in [-r,r]$, we look for $z : [a,b] \to \mathbb{R}$ such that

$$\begin{aligned}(\dot{z}(t) - \dot{u}(t))(\zeta - z(t)) \ge 0 \qquad &\forall\, \zeta \in [-r,r], \text{ a.e. in } [a,b], \\ z(t) \in [-r,r] \qquad \forall\, t \in [a,b], \qquad &z(a) = z_a \,.\end{aligned} \tag{5.14}$$

A function $z \in W^{1,1}(a,b)$ solves (5.14) if and only if

$$\begin{aligned}&\dot{z} = \dot{u} \text{ a.e. on } \{|z| < r\}, \\ &\dot{z} = 0 \text{ and } \dot{u} \ge 0 \text{ a.e. on } \{z = r\}, \\ &\dot{z} = 0 \text{ and } \dot{u} \le 0 \text{ a.e. on } \{z = -r\}.\end{aligned} \tag{5.15}$$

This coincides with the properties of the construction given above for the scalar play, taking into account (5.8) and (5.9).

The vector stop and play. Now, the functions u and z take values in some real Hilbert space H instead of $\mathbb{R}$. The interval $[-r,r]$ is replaced by a closed convex set Z, the product in (5.14) by the scalar product in H. Given $u : [a,b] \to H$ and $z_a \in Z$, we look for $z : [a,b] \to Z$ such that

$$\begin{aligned}\langle \dot{z}(t) - \dot{u}(t), \zeta - z(t)\rangle \ge 0 \qquad &\forall\, \zeta \in Z, \text{ a.e. in } [a,b], \\ z(t) \in Z \qquad \forall\, t \in [a,b], \qquad &z(a) = z_a \,.\end{aligned} \tag{5.16}$$

We assume that $u \in W^{1,1}(a,b;H)$, that is, the weak derivative $\dot{u}$ exists and is an element of $L^1(a,b;H)$, the space of H-valued Bochner integrable functions. These functions have the property that

$$u(t) - u(s) = \int_s^t \dot{u}(\tau)\, d\tau$$

holds for all $s, t \in [a,b]$.

It turns out that for any given input $u \in W^{1,1}(0,T;H)$ and initial value z_a there exists a unique solution $z \in W^{1,1}(0,T;H)$ of (5.16). A proof can be found in [14], Theorem 4.1 and Proposition 4.1 as well as in [13]; the uniqueness part is given in Proposition 5.6 below. In the next section, we will present the existence proof as developed in the energetic approach.

Thus, the solution operator $(u, z_a) \mapsto z$ is well-defined; it is called the **vector stop operator**. We write

$$z = \mathcal{S}_Z[u; z_a]\,. \tag{5.17}$$

The corresponding **vector play operator**

$$w = \mathcal{P}_Z[u; z_a] \tag{5.18}$$

should satisfy

$$u(t) = w(t) + z(t)\,, \quad t \in [a,b]\,. \tag{5.19}$$

We achieve this by simply setting

$$\mathcal{P}_Z[u; z_a] = u - \mathcal{S}_Z[u; z_a]\,.$$

The system

$$\begin{gathered} u(t) = w(t) + z(t) \\ \langle \dot{w}(t), z(t) - \zeta \rangle \ge 0 \qquad \forall\, \zeta \in Z, \text{ a.e. in } (a,b), \\ z(t) \in Z \qquad \forall\, t \in [a,b], \qquad z(a) = z_a\,, \end{gathered} \tag{5.20}$$

is obviously equivalent to (5.16).

The rate independence of $\mathcal{S}_Z$ and thus of $\mathcal{P}_Z$ can be checked directly from the definitions. The Volterra property is a direct consequence of Proposition 5.6 below. Thus, $\mathcal{S}_Z$ and $\mathcal{P}_Z$ are hysteresis operators.

Assume that $w, z \in W^{1,1}(a,b)$ are solutions of (5.20). We have

$$\langle \dot{w}(t), \dot{z}(t) \rangle = 0 \quad \text{a.e. in } (a,b). \tag{5.21}$$

Indeed, (5.21) follows from the variational inequality (5.20) by testing with $\zeta = z(t \pm h)$ for $h > 0$, dividing by h and letting h go to 0. The decomposition

$$\dot{u}(t) = \dot{w}(t) + \dot{z}(t)\,, \quad \text{a.e. in } (a,b), \tag{5.22}$$

splits $\dot{u}(t)$ into its normal and its tangential part at $z(t)$; namely,

$$\begin{aligned} \dot{w}(t) &\in N_Z(z(t)) = \{y : \langle y, \zeta - z(t) \rangle \le 0 \ \text{ for all } \zeta \in Z\} \\ \dot{z}(t) &\in T_Z(z(t)) = \overline{\operatorname{cone}(Z - z(t))}\,. \end{aligned} \tag{5.23}$$

As a further consequence of (5.21) and (5.22), $|\dot{w}|^2 = \langle \dot{u}, \dot{w} \rangle$ as well as $|\dot{z}|^2 = \langle \dot{u}, \dot{z} \rangle$, so

$$|\dot{w}(t)| \le |\dot{u}(t)|\,, \quad |\dot{z}(t)| \le |\dot{u}(t)|\,, \quad \text{a.e. in } (a,b). \tag{5.24}$$

We now prove uniqueness and stability of the solution z of (5.16) using the standard monotonicity argument.

Proposition 5.6 *Let $z_1, z_2 \in W^{1,1}(a,b;H)$ be solutions of (5.20) for given functions $u_1, u_2 \in W^{1,1}(a,b;H)$ and initial values $z_{a,1}, z_{a,2}$. Then we have*

$$|z_1(t) - z_2(t)| \le |z_{a,1} - z_{a,2}| + \int_a^t |\dot{u}_1(\tau) - \dot{u}_2(\tau)| \, d\tau \tag{5.25}$$

for all $t \in [a,b]$. In particular, the solution of (5.16) is unique.

Proof: For almost all t we have

$$\begin{aligned} |z_1(t) - z_2(t)| \frac{\mathrm{d}}{\mathrm{dt}} |z_1(t) - z_2(t)| &= \frac{\mathrm{d}}{\mathrm{dt}} \frac{1}{2} |z_1(t) - z_2(t)|^2 = \langle \dot{z}_1(t) - \dot{z}_2(t), z_1(t) - z_2(t) \rangle \\ &= \langle \dot{u}_1(t) - \dot{u}_2(t), z_1(t) - z_2(t) \rangle - \langle \dot{w}_1(t) - \dot{w}_2(t), z_1(t) - z_2(t) \rangle \\ &\le \langle \dot{u}_1(t) - \dot{u}_2(t), z_1(t) - z_2(t) \rangle \le |\dot{u}_1(t) - \dot{u}_2(t)| \, |z_1(t) - z_2(t)| \end{aligned}$$

due to the variational inequality. Therefore

$$\frac{\mathrm{d}}{\mathrm{dt}} |z_1(t) - z_2(t)| \le |\dot{u}_1(t) - \dot{u}_2(t)|$$

for almost all t. Integrating over $[a,t]$ yields the assertion. □

We refer to [14] for several refined stability results for the vector stop and play operators.

6 Energetic Solutions

In the previous section, the existence of a solution of the variational inequality defining the vector stop operator remained open. We now prove existence with the energetic approach, developed by A. Mielke and several co-workers, see the monograph [17] and the references therein.

The energetic approach is based on two potentials which depend on time, typically via a time-dependent function like an external force; their interaction generates a solution, which is a time-dependent function. This approach deals with rate-independent problems in a natural manner. It has created a unifying framework for many different problems arising in mechanics.

The scalar play operator $\mathcal{P}_r$ can be used to illustrate “in a nutshell” some basic ingredients of the energetic approach.

Let $u \in W^{1,1}(a,b)$ be given. The function $w = \mathcal{P}_r[u; w_a]$ satisfies

$$(w(t) - u(t))\dot{w}(t) + r|\dot{w}(t)| = 0\,, \quad \text{a.e. in } (a,b). \tag{6.1}$$

We consider an energy $\mathcal{E} : [a,b] \times \mathbb{R} \to \mathbb{R}$,

$$\mathcal{E}(t,q) = \frac{1}{2}(q - u(t))^2\,, \quad D_q\mathcal{E}(t,q) = q - u(t)\,, \tag{6.2}$$

and a dissipation potential $\mathcal{R} : \mathbb{R} \to \mathbb{R}_+$,

$$\mathcal{R}(v) = r|v|\,, \quad \partial\mathcal{R}(v) = r\mathrm{sign}\,(v)\,. \tag{6.3}$$

Here, “sign” stands for the set-valued sign function with $\mathrm{sign}\,(0) = [-1,1]$, and $\partial\mathcal{R}$ denotes the subdifferential of $\mathcal{R}$ in the sense of convex analysis. We note that

$$\mathcal{R}(v) = \partial\mathcal{R}(v) \cdot v = r|v|\,, \quad \text{for all } v \in \mathbb{R}. \tag{6.4}$$

In terms of $\mathcal{E}$ and $\mathcal{R}$, (6.1) becomes (we omit the phrase “a.e. in (a,b)”)

$$D_q\mathcal{E}(t, w(t))\dot{w}(t) + \mathcal{R}(\dot{w}(t)) = 0\,. \tag{6.5}$$

Since $|w(t) - u(t)| \le r$ holds for the play, (6.2) and (6.3 imply

$$D_q\mathcal{E}(t, w(t))v + \mathcal{R}(v) \ge 0\,, \quad \text{for all } v \in \mathbb{R}. \tag{6.6}$$

Moreover, dividing (6.1) by $\dot{w}(t)$ yields

$$D_q\mathcal{E}(t, w(t)) + \partial\mathcal{R}(\dot{w}(t)) \ni 0\,. \tag{6.7}$$

When applying the energetic approach to problems from mechanics, equations (6.5), (6.6) and (6.7) – for the appropriate energy and dissipation potentials, of course – can be interpreted as a power balance (or local energy balance), a local stability condition, and a force balance, respectively. (The play operator corresponds to the modeling of dry friction in 1D, see [17].)

The local stability condition (6.6) can be replaced by a global stability condition, since $q \mapsto \mathcal{E}(t,q)$ is convex. Indeed, the convex function

$$g(q) = \mathcal{E}(t,q) + \mathcal{R}(q - w(t))$$

satisfies $\partial g(w(t)) = w(t) - u(t) + r \operatorname{sign}(0) \ni 0$. Thus, $w(t)$ is a global minimizer of g, and we obtain the global stability condition

$$\text{(S)} \qquad \mathcal{E}(t, w(t)) \le \mathcal{E}(t,q) + \mathcal{R}(q - w(t))\,, \quad \text{for all } q \in \mathbb{R},\ t \in [a,b]. \tag{6.8}$$

The power balance, too, can be transformed into a global condition. As

$$\mathcal{E}(b, w(b)) - \mathcal{E}(a, w(a)) = \int_a^b D_t\mathcal{E}(t, w(t)) + D_q\mathcal{E}(t, w(t))\dot{w}(t)\,dt$$

holds, using (6.5) we obtain the global energy balance

$$\mathcal{E}(b, w(b)) + \int_a^b \mathcal{R}(\dot{w}(t))\,dt = \mathcal{E}(a, w(a)) + \int_a^b D_t\mathcal{E}(t, w(t))\,dt\,. \tag{6.9}$$

For $\mathcal{R}$ as in (6.3), the integral on the left side equals $r \operatorname{var}(w)$. Thus, we may write (6.9) in the form

$$\text{(E)} \qquad \mathcal{E}(b, w(b)) + \sup_{\Delta} \sum_j \mathcal{R}(w(t_j) - w(t_{j-1})) = \mathcal{E}(a, w(a)) + \int_a^b D_t\mathcal{E}(t, w(t))\,dt\,, \tag{6.10}$$

where the sup ranges over all finite partitions $\Delta = \{t_j\}$ of $[a,b]$. This equation no longer includes derivatives except those of the "driving function" u, as $D_t\mathcal{E}(t,q) = q - \dot{u}(t)$.

The function $w = \mathcal{P}_r[u; z_0]$ solves (6.8) and (6.10). Such solutions are called **energetic solutions** of the system defined by $\mathcal{E}$ and $\mathcal{R}$.

Coercive quadratic energies. This subsection is based on the exposition in Section 3.5 of [17] and Section 2 of [16].

The setting of the problem is as follows. Let Q be a real Hilbert space with dual Q^*. We define the **energy**

$$\mathcal{E}(t,q) = \frac{1}{2}\langle Aq, q\rangle - \langle \ell(t), q\rangle\,, \qquad \mathcal{E} : [a,b] \times Q \to \mathbb{R}\,. \tag{6.11}$$

Here, $\ell \in W^{1,1}(a,b;Q^*)$ is the function which drives the evolution. The operator $A : Q \to Q^*$ is linear, bounded, symmetric and positive definite; in particular, we have

$$\langle Aq, p\rangle = \langle Ap, q\rangle\,, \quad \alpha_0|q|^2 \le \langle Aq, q\rangle \le \alpha_1|q|^2\,, \qquad \text{for all } p, q \in Q \tag{6.12}$$

for some numbers $\alpha_0, \alpha_1 > 0$. The brackets $\langle\cdot,\cdot\rangle$ denote the duality pairing on $Q^* \times Q$, and $|\cdot|$ denotes the norm on Q.

The **dissipation potential** $\mathcal{R} : Q \to [0, +\infty]$ is assumed to be lower semicontinuous, convex and positively 1-homogeneous, that is, $\mathcal{R}(\lambda q) = \lambda\mathcal{R}(q)$ for all $\lambda > 0$, [1] and we assume that $\mathcal{R}(0) = 0$. [2]

[1] Note that for convex functionals on a Hilbert space, the four notions of semicontinuity (weak/strong, sequential/topological) are equivalent. Note also that $\mathcal{R}$ satisfies the triangle inequality, as $\mathcal{R}(p+q) = 2\mathcal{R}((p+q)/2) \le \mathcal{R}(p) + \mathcal{R}(q)$.

[2] Thus, $0 \in \partial\mathcal{R}(0)$; in particular, $\partial\mathcal{R}(0)$ is not empty.

There is no coercivity assumption for $\mathcal{R}$, as the coercivity of A assumed above is all what is needed. Moreover, $A : Q \to Q^*$ is a Hilbert space isomorphism, and

$$|q|_A = \sqrt{\langle Aq, q\rangle} \tag{6.13}$$

defines a norm which by (6.12) is equivalent to $|\cdot|$.

Given an initial condition

$$q(a) = q_a \in Q\,, \tag{6.14}$$

we want to prove the existence of an energetic solution $q \in W^{1,1}(a, b)$ of the problem above; that is, the function q should satisfy the conditions (S) and (E), see (6.8) and (6.10), in place of w, and with Q in place of $\mathbb{R}$. Since we require the condition (S) to hold at the initial time $t = a$, too, this restricts the choice of q_a.

The general frame of the proof is a common one. One discretizes in time, solves a finite sequence of time-discrete problems, proves a priori estimates, goes back to continuous time via interpolation, obtains a candidate for the continuous solution by compactness, and finally shows that this candidate indeed solves the problem. The difficulty, of course, lies in carrying out this program.

The time-discrete problem. Let $\Delta^N = \{t_k\}$ be the equidistant partition of $[a, b]$ with $a = t_0 < \cdots < t_N = b$. Assume that $q_{k-1} \in Q$ is already constructed. Then $q_k \in Q$ is chosen as a solution of

$$\min_{p\in Q} \mathcal{E}(t_k, p) + \mathcal{R}(p - q_{k-1})\,. \tag{6.15}$$

In order to analyze this problem, let

$$J(p) = \mathcal{E}(t, p) + \mathcal{R}(p - q_{in})\,, \quad t \in [a, b]\,,\, q_{in} \in Q\,.$$

Since J is coercive, strictly convex and lower semicontinuous, it has a unique minimizer $q_* \in Q$. An explicit computation which uses the quadratic structure of $\mathcal{E}$ and the convexity of $\mathcal{R}$ yields, for $0 < \lambda \le 1$,

$$\begin{aligned} 0 \le \frac{J(q_* + \lambda(p - q_*) - J(q_*))}{\lambda} \le \\ \le (\mathcal{E}(t, p) + \mathcal{R}(p - q_{in})) - (\mathcal{E}(t, q_*) + \mathcal{R}(q_* - q_{in})) - \frac{1-\lambda}{2}\langle A(p - q_*), p - q_*\rangle\,. \end{aligned}$$

Passing to the limit $\lambda \to 0$, we arrive at

$$\mathcal{E}(t, q_*) + \mathcal{R}(q_* - q_{in}) + \frac{1}{2}\langle A(p - q_*), p - q_*\rangle \le \mathcal{E}(t, p) + \mathcal{R}(p - q_{in})\,, \quad \forall\, p \in Q\,.$$

For problem (6.15), this becomes

$$\mathcal{E}(t_k, q_k) + \mathcal{R}(q_k - q_{k-1}) + \frac{1}{2}|p - q_k|_A^2 \le \mathcal{E}(t_k, p) + \mathcal{R}(p - q_{k-1})\,, \quad \forall\, p \in Q\,. \tag{6.16}$$

This implies (triangle inequality for $\mathcal{R}$)

$$\mathcal{E}(t_k, q_k) \le \mathcal{E}(t_k, p) + \mathcal{R}(p - q_k)\,, \quad \forall\, p \in Q\,. \tag{6.17}$$

Discrete a priori estimates. We insert $p = q_{k-1}$ in (6.16) and obtain

$$\mathcal{E}(t_k, q_k) + \mathcal{R}(q_k - q_{k-1}) + \frac{1}{2}|q_{k-1} - q_k|_A^2 \leq \mathcal{E}(t_k, q_{k-1}) \,.$$

Replacing k with $k+1$ we get

$$\mathcal{E}(t_{k+1}, q_{k+1}) + \mathcal{R}(q_{k+1} - q_k) + \frac{1}{2}|q_k - q_{k+1}|_A^2 \leq \mathcal{E}(t_{k+1}, q_k) \,. \tag{6.18}$$

Inserting $p = q_{k+1}$ in (6.16) gives

$$\mathcal{E}(t_k, q_k) + \mathcal{R}(q_k - q_{k-1}) + \frac{1}{2}|q_{k+1} - q_k|_A^2 \leq \mathcal{E}(t_k, q_{k+1}) + \mathcal{R}(q_{k+1} - q_{k-1}) \,.$$

Rearranging yields, again with the aid of the triangle inequality,

$$\frac{1}{2}|q_{k+1} - q_k|_A^2 \leq \mathcal{E}(t_k, q_{k+1}) - \mathcal{E}(t_k, q_k) + \mathcal{R}(q_{k+1} - q_k) \,,$$

Adding inequality (6.18) to this inequality we next obtain

$$|q_{k+1} - q_k|_A^2 \leq \mathcal{E}(t_k, q_{k+1}) - \mathcal{E}(t_k, q_k) + \mathcal{E}(t_{k+1}, q_k) - \mathcal{E}(t_{k+1}, q_{k+1}) \,. \tag{6.19}$$

It follows that

$$\begin{aligned} |q_{k+1} - q_k|_A^2 &\leq \int_{t_k}^{t_{k+1}} -\partial_t \mathcal{E}(t, q_{k+1}) + \partial_t \mathcal{E}(t, q_k) \, dt = \int_{t_k}^{t_{k+1}} \Big\langle \dot{\ell}(t), q_{k+1} - q_k \Big\rangle \, dt \\ &= \langle \ell(t_{k+1}) - \ell(t_k), q_{k+1} - q_k \rangle \leq \|\ell(t_{k+1}) - \ell(t_k)\|_{Q^*} \cdot |q_{k+1} - q_k| \\ &\leq \|\ell(t_{k+1} - \ell(t_k))\|_{Q^*} \cdot \alpha_0^{-1/2} |q_{k+1} - q_k|_A \,, \end{aligned}$$

so

$$|q_{k+1} - q_k|_A \leq \alpha_0^{-1/2} \|\ell(t_{k+1} - \ell(t_k))\|_{Q^*} \tag{6.20}$$

This yields the discrete a priori estimate (now we write q_k^N for q_k)

$$|q_k^N|_A \leq |q_a|_A + \alpha_0^{-1/2} \int_a^b \|\dot{\ell}(t)\|_{Q^*} \, dt \tag{6.21}$$

which is valid for all k and all N.

Continuous interpolation. Let $\hat{q}^N$ denote the piecewise affine interpolant of $\{t_k^N, q_k^N\}_k$ and $\overline{q}^N$ the piecewise constant interpolant with $\overline{q}^N = q_{k-1}^N$ on $[t_{k-1}^N, t_k^N)$. Then $\hat{q}^N \in C([a,b];Q)$, $\overline{q}^N \in L^\infty(a,b;Q)$, and (6.21) implies

$$\|\hat{q}^N\|_\infty \leq |q_a|_A + C \int_a^b \|\dot{\ell}(t)\|_{Q^*} \, dt \,, \quad \|\overline{q}^N\|_\infty \leq |q_a|_A + C \int_a^b \|\dot{\ell}(t)\|_{Q^*} \, dt \,. \tag{6.22}$$

Moreover, for all $s \leq t$ in $[a,b]$,

$$|\hat{q}^N(t) - \hat{q}^N(s)| \leq C \Big(\int_s^t \|\dot{\ell}(\tau)\|_{Q^*} \, d\tau + \delta_N \Big) \,, \quad \|\hat{q}^N - \overline{q}^N\|_\infty \leq C\delta_N \,, \tag{6.23}$$

for some $\delta_N \to 0$, since $t \mapsto \|\dot{\ell}(t)\|_{Q^*}$ is integrable.

Convergent subsequence. The estimates (6.22) and (6.23) imply that the sequence $\{\hat{q}^N\}_N$ is equicontinuous and bounded in $C([a,b];Q)$. A variant of the Arzela-Ascoli theorem now implies that for some subsequence, denoted by $\{\hat{q}^m\}$, the values $\hat{q}^m(t)$ weakly converge to some $q(t) \in Q$ pointwise in t. By (6.23), also $\overline{q}^m(t) \rightharpoonup q(t)$ pointwise in t. Moreover, the weak lower semicontinuity of the norm in Q and (6.23) imply that

$$|q(t) - q(s)| \le C \int_s^t \|\dot{\ell}(\tau)\|_{Q^*}\, d\tau \tag{6.24}$$

for $s, t \in [a,b]$. Therefore, $q \in W^{1,1}(a,b;Q)$ and $|\dot{q}(t)| \le \|\dot{\ell}(t)\|_{Q^*}$ a.e. in (a,b).

The quadratic trick. When $v_n \rightharpoonup v$ in Q, the quadratic terms $\langle Av_n, v_n\rangle$ need not converge to $\langle Av, v\rangle$. But for fixed $w \in Q$, we nevertheless have

$$\langle A(v_n + w), v_n + w\rangle - \langle Av_n, v_n\rangle \quad \to \quad \langle A(v + w), v + w\rangle - \langle Av, v\rangle \;,$$

since the terms $\langle Av_n, v_n\rangle$ cancel. As a consequence, when $t_n \to t$ in $[a,b]$ and $v_n \rightharpoonup v$ in Q, we get

$$\mathcal{E}(t_n, v_n + w) - \mathcal{E}(t_n, v_n) \quad \to \quad \mathcal{E}(t, v + w) - \mathcal{E}(t, v)\,.$$

This is relevant for the next step.

The limit function *q* satisfies (S). Let $t \in [a,b]$, $t > a$. Given $N \in \mathbb{N}$, choose $k = k(N)$ with $t_k^N \le t < t_{k+1}^N$, thus $q_k^N = \overline{q}^N(t)$. According to (6.17),

$$\mathcal{E}(t_k^N, \overline{q}^N(t)) \le \mathcal{E}(t_k^N, p) + \mathcal{R}(p - \overline{q}^N(t))$$

for all $p \in Q$. We replace p with $p + \overline{q}^N(t) - q(t)$ and obtain

$$\mathcal{E}(t_k^N, \overline{q}^N(t)) \le \mathcal{E}(t_k^N, p + \overline{q}^N(t) - q(t)) + \mathcal{R}(p - q(t))$$

We pass to the subsequence $\overline{q}^m$ with $\overline{q}^m(t) \rightharpoonup q(t)$ and apply the quadratic trick to obtain

$$\mathcal{E}(t, q(t)) \le \mathcal{E}(t, p) + \mathcal{R}(p - q(t)) \quad \text{for all } p \in Q,\ t \in (a,b]. \tag{6.25}$$

The condition on the initial value. In order that (6.25) also holds for $t = a$, the initial value $q_a = q(a)$ has to be a minimizer of

$$J(p) = \mathcal{E}(a, p) + \mathcal{R}(p - q_a) = \frac{1}{2}\langle Ap, p\rangle - \langle \ell(a), p\rangle + \mathcal{R}(p - q_a)\,.$$

As J is convex, this is equivalent to

$$0 \in \partial J(q_a) = Aq_a - \ell(a) + \partial\mathcal{R}(0)\,.$$

This in turn is equivalent to

$$q_a \in A^{-1}(\ell(a) - Z)\,, \quad Z := \partial\mathcal{R}(0)\,. \tag{6.26}$$

The limit function *q* satisfies "≤" in (E). Let $t \in [a,b]$, let $\Delta = \{\tau_j\}$, $0 \le j \le M$ be an arbitrary partition of $[a,t]$, let Δ^N the partition which was used in the time-discrete step. From (6.16) with $p = q_{k-1}^N$ we obtain

$$\mathcal{E}(t_k^N, q_k^N) + \mathcal{R}(q_k^N - q_{k-1}^N) \le \mathcal{E}(t_k^N, q_{k-1}^N) = \mathcal{E}(t_{k-1}^N, q_{k-1}^N) + \int_{t_{k-1}^N}^{t_k^N} D_t\mathcal{E}(s, q_{k-1}^N)\, ds\,. \tag{6.27}$$

If $t \in [t_k^N, t_{k+1}^N)$, we have

$$\mathcal{E}(t, \overline{q}^N(t)) + \sum_{i=1}^{k} \mathcal{R}(q_i^N - q_{i-1}^N) = \mathcal{E}(t_k, q_k^N) + \int_{t_k^N}^{t} D_t\mathcal{E}(s, q_k^N)\, ds + \sum_{i=1}^{k} \mathcal{R}(q_i^N - q_{i-1}^N)\,.$$

On the right side we successively use (6.27) and get

$$\mathcal{E}(t, \overline{q}^N(t)) + \sum_{i=1}^{k} \mathcal{R}(\overline{q}(t_i^N) - \overline{q}(t_{i-1}^N)) \le \mathcal{E}(a, q(a)) + \int_a^t D_t\mathcal{E}(s, \overline{q}^N(s))\, ds\,.$$

The left side does not change if we replace the points from Δ^N by those from $\Delta \cup \Delta^N$; the inequality persists if we only use those from Δ. Thus

$$\mathcal{E}(t, \overline{q}^N(t)) + \sum_{j=1}^{M} \mathcal{R}(\overline{q}^N(\tau_j) - \overline{q}^N(\tau_{j-1})) \le \mathcal{E}(a, q(a)) - \int_a^t \left\langle \dot{\ell}(s), \overline{q}^N(s) \right\rangle\, ds\,.$$

Passing to the limes inferior for the subsequence with $\overline{q}^m(t) \rightharpoonup q(t)$ we obtain

$$\mathcal{E}(t, q(t)) + \sum_{j=1}^{M} \mathcal{R}(q(\tau_j) - q(\tau_{j-1})) \le \mathcal{E}(a, q(a)) - \int_a^t \left\langle \dot{\ell}(s), q(s) \right\rangle\, ds\,.$$

Taking the supremum w.r.t all partitions Δ of $[a, t]$ yields the assertion.

The limit function q satisfies "$\ge$" in (E). Let $t \in [a, b]$, let $\Delta = \{\tau_j\}$, $0 \le j \le M$, be an arbitrary partition of $[a, t]$. For $j \in \{1, \dots, M\}$ we obtain, due to (6.25),

$$\begin{aligned}
&\mathcal{E}(\tau_j, q(\tau_j)) + \mathcal{R}(q(\tau_j) - q(\tau_{j-1})) \\
&\qquad = \int_{\tau_{j-1}}^{\tau_j} D_t\mathcal{E}(s, q(\tau_j))\, ds + \mathcal{E}(\tau_{j-1}, q(\tau_j)) + \mathcal{R}(q(\tau_j) - q(\tau_{j-1})) \\
&\qquad \ge \int_{\tau_{j-1}}^{\tau_j} D_t\mathcal{E}(s, q(\tau_j))\, ds + \mathcal{E}(\tau_{j-1}, q(\tau_{j-1}))\,.
\end{aligned}$$

Let q_Δ be the piecewise constant interpolant of q on Δ with $q_\Delta(s) = q(\tau_j)$ for $s \in (\tau_{j-1}, \tau_j]$. Summing over j we obtain, since $t = \tau_M$,

$$\mathcal{E}(t, q(t)) + \sum_{j=1}^{M} \mathcal{R}(q(\tau_j) - q(\tau_{j-1})) \ge \mathcal{E}(a, q(a)) - \int_a^t \left\langle \dot{\ell}(s), q_\Delta(s) \right\rangle\, ds\,.$$

As $\dot{\ell} \in L^1(a, b; Q^*)$ and q_Δ converges pointwise to the bounded function q as $|Delta| \to 0$, passing to the supremum w.r.t. Δ yields

$$\mathcal{E}(t, q(t)) + \sup_{\Delta} \sum_{j=1}^{M} \mathcal{R}(q(\tau_j) - q(\tau_{j-1})) \ge \mathcal{E}(a, q(a)) + \int_a^t D_t\mathcal{E}(s, q(s))\, ds\,. \tag{6.28}$$

This proves the claim.

We summarize the existence result.

Theorem 6.1 *Let the assumptions on the energy functional $\mathcal{E}$ and the dissipation functional $\mathcal{R}$ be satisfied, as they are described in the subsection above where (6.11) – (6.14) are stated. Then for every $q_a \in A^{-1}(\ell(a) - Z)$, $Z = \partial\mathcal{R}(0)$, there exists an energetic solution $q \in W^{1,1}(a, b : Q)$ with $q(a) = q_a$, that is, a solution which satisfies*

$$\text{(S)} \qquad \mathcal{E}(t, q(t)) \le \mathcal{E}(t, p) + \mathcal{R}(p - q(t)) \,, \quad \textit{for all } p \in Q,\ t \in [a, b], \tag{6.29}$$

as well as

$$\text{(E)} \qquad \mathcal{E}(t, q(t)) + \sup_{\Delta} \sum_j \mathcal{R}(q(t_j) - q(t_{j-1})) = \mathcal{E}(a, q(a)) + \int_a^t D_t\mathcal{E}(s, q(s))\, ds \,, \tag{6.30}$$

for all $t \in [a, b]$, where the supremum is taken over all partitions $\Delta = \{t_j\}$ of $[a, t]$.

In the case considered here (coercive quadratic energy), uniqueness and continuous dependence on data also holds, see [17], Proposition 3.5.5.

From the energetic solution to the vector stop and play.

The starting point is the formula, valid for the energetic solution q,

$$\int_r^t \mathcal{R}(\dot{q}(s))\, ds = \sup_{\Delta} \sum_j \mathcal{R}(q(t_j) - q(t_{j-1})) \tag{6.31}$$

where Δ ranges over all partitions Δ of an arbitrary subinterval $[r, t]$ of $[a, b]$. (See the appendix below after (6.55). We use the chain rule for $t \mapsto \mathcal{E}(t, q(t))$ and get from (E) and (6.31)

$$\int_r^t \langle D_q\mathcal{E}(s, q(s)), \dot{q}(s)\rangle + \mathcal{R}(\dot{q}(s))\, ds = 0 \,. \tag{6.32}$$

Since r and t are arbitrary, we may pass to the pointwise form

$$\langle D_q\mathcal{E}(t, q(t)), \dot{q}(t)\rangle + \mathcal{R}(\dot{q}(t)) = 0 \,, \quad \text{a.e. in } (a, b). \tag{6.33}$$

On the other hand, $q(t)$ minimizes $p \mapsto \mathcal{E}(t, p) + \mathcal{R}(p - q(t))$ due to the stability condition (S), so

$$0 \in D_q\mathcal{E}(t, q(t)) + \partial\mathcal{R}(0) \,. \tag{6.34}$$

The correspondence between the energetic solution and the vector stop resp. play operator is based on a correspondence between the dissipation functional $\mathcal{R} : Q \to [0, +\infty]$ and the closed convex set $\partial\mathcal{R}(0) \subset Q^*$.

At this point we need some results from convex analysis.

Proposition 6.2 *Let $\mathcal{R} : Q \to [0, \infty]$ be convex, lower semicontinuous and positively 1-homogeneous with $\mathcal{R}(0) = 0$. Then*

$$\partial\mathcal{R}(0) = \{\zeta : \langle \zeta, v\rangle \le \mathcal{R}(v) \textit{ for all } v \in Q\} \,, \tag{6.35}$$

The convex conjugate $\mathcal{R}^*$ defined on Q^* by

$$\mathcal{R}^*(\zeta) = \sup_{v \in Q}(\langle \zeta, v\rangle - \mathcal{R}(v)) \tag{6.36}$$

satisfies

$$\mathcal{R}^* = I_{\partial\mathcal{R}(0)} \quad [3] \tag{6.37}$$

and we have

$$\mathcal{R}(v) = \sup_{\zeta\in\partial\mathcal{R}(0)} \langle\zeta, v\rangle , \quad \forall\, v \in Q\,. \tag{6.38}$$

Proof. Since for $\zeta \in Q^*$ and $v \in Q$ we have $\mathcal{R}(v) - \langle\zeta, v\rangle = \mathcal{R}(v) - \mathcal{R}(0) - \langle\zeta, v - 0\rangle$, (6.35) follows. Now, setting $Z = \partial\mathcal{R}(0)$,

$$\begin{aligned}
\zeta \in Z \quad &\Leftrightarrow \quad 0 = \sup_{v\in Q}(\langle\zeta, v\rangle - \mathcal{R}(v)) = \mathcal{R}^*(\zeta) \\
\zeta \notin Z \quad &\Leftrightarrow \quad \sup_{v\in Q}(\langle\zeta, v\rangle - \mathcal{R}(v)) > 0 \quad \Leftrightarrow \quad \sup_{v\in Q,\lambda>0} \lambda(\langle\zeta, v\rangle - \mathcal{R}(v)) = +\infty \\
&\Leftrightarrow \quad \mathcal{R}^*(\zeta) = +\infty\,.
\end{aligned}$$

This proves (6.37). Since $\mathcal{R}$ is proper, convex and lower semicontinous, we have $\mathcal{R}^{**} = \mathcal{R}$ and therefore

$$\mathcal{R}(v) = \mathcal{R}^{**}(v) = I_Z^*(v) = \sup_{\zeta\in Q^*}(\langle\zeta, v\rangle - I_Z(\zeta)) = \sup_{\zeta\in Z}\langle\zeta, v\rangle$$

for all $v \in Q$. $\square$

Conversely, given a closed convex set $Z \subset Q^*$, we may construct a dissipation potential $\mathcal{R}$ on Q with $\partial\mathcal{R}(0) = Z$.

Proposition 6.3 *Let $Z \subset Q^*$ be closed and convex, $Z \neq \emptyset$. Then*

$$\mathcal{R}(v) = I_Z^*(v) = \sup_{\zeta\in Z}\langle\zeta, v\rangle \tag{6.39}$$

defines an $\mathcal{R} : Q \to [0, +\infty]$ which is convex, lower semicontinuous, positively 1-homogeneous and satisfies $\mathcal{R}(0) = 0$ as well as $\partial\mathcal{R}(0) = Z$.

Proof. All properties of $\mathcal{R}$ except the last follow immediately from its definition in (6.39). We have $Z \subset \partial\mathcal{R}(0)$ by (6.35). If on the other hand $\zeta \notin Z$, by separation there exists a $v \in Q$ such that

$$\langle\zeta, v\rangle > \sup_{\zeta\in Z}\langle\zeta, v\rangle = \mathcal{R}(v)\,.$$

Thus $\zeta \notin \partial\mathcal{R}(0)$. $\square$

Let q be an energetic solution according to Theorem 6.1. We set

$$z(t) = -D_q\mathcal{E}(t, q(t)) = \ell(t) - Aq(t)\,, \quad Z = \partial\mathcal{R}(0)\,. \tag{6.40}$$

From(6.34) we see that

$$z(t) \in Z\,, \quad \text{for all } t \in [a, b]. \tag{6.41}$$

From (6.33) and (6.39) we get

$$\langle z(t), \dot q(t)\rangle = \mathcal{R}(\dot q(t)) = \sup_{\zeta\in Z}\langle\zeta, \dot q(t)\rangle\,, \quad \text{a.e. in } (a, b).$$

[3]The indicator function I_Z of a set Z is defined to be $I_Z(\zeta) = 0$ for $\zeta \in Z$, and $I_Z(\zeta) = +\infty$ otherwise.

Inserting $\dot{q} = A^{-1}(\dot{\ell} - \dot{z})$ yields for a.a. $t \in (a,b)$

$$\Big\langle z(t) - \zeta, A^{-1}(\dot{\ell}(t) - \dot{z}(t)) \Big\rangle_{Q^*Q} \geq 0\,, \quad \text{for all } \zeta \in Z. \tag{6.42}$$

We define a scalar product in Q^* by

$$\langle \zeta, \eta \rangle_{A^{-1}} = \langle \eta, A^{-1}\zeta \rangle_{Q^*Q}\ , \quad \zeta, \eta \in Q^*. \tag{6.43}$$

Therefore $z \in W^{1,1}(a,b;Q^*)$ satisfies

$$\begin{aligned} \Big\langle \dot{z}(t) - \dot{\ell}(t)), \zeta - z(t) \Big\rangle_{A^{-1}} \geq 0\,, \quad \forall\, \zeta \in Z, \text{ for a.a. } t \in (a,b) \\ z(t) \in Z \quad \forall t \in [a,b]\,, \quad z(a) = z_a\,. \end{aligned} \tag{6.44}$$

Thus, if $q \in W^{1,1}(a,b;Q)$ is an energetic solution, the function $z = \ell - Aq$ satisfies

$$z = \mathcal{S}_Z[\ell; z_a] \tag{6.45}$$

in the underlying Hilbert space $(Q^*, \langle \cdot, \cdot \rangle_{A^{-1}})$. As the stop and the play operator always come in pairs, the function $w = \ell - z = Aq$ satisfies, in the same Hilbert space,

$$w = \mathcal{P}_Z[\ell; z_a]\,. \tag{6.46}$$

Consequently,

$$q = A^{-1}\mathcal{P}_Z[\ell; z_a]\,. \tag{6.47}$$

One can also obtain a stop and play pair on the original space Q. Indeed, we have, since A is symmetric,

$$\begin{aligned} \Big\langle z(t) - \zeta, A^{-1}(\dot{\ell}(t) - \dot{z}(t)) \Big\rangle_{Q^*Q} &= \Big\langle A(A^{-1}z(t) - A^{-1}\zeta), A^{-1}(\dot{\ell}(t) - \dot{z}(t)) \Big\rangle_{Q^*Q} \\ = \Big\langle AA^{-1}(\dot{\ell}(t) - \dot{z}(t)), A^{-1}z(t) - A^{-1}\zeta \Big\rangle_{Q^*Q} &= \Big\langle A^{-1}\dot{z}(t) - A^{-1}\dot{\ell}(t)), A^{-1}\zeta - A^{-1}z(t) \Big\rangle_A\,. \end{aligned}$$

In view of (6.44), we see that, on $(Q, \langle \cdot, \cdot \rangle_A)$,

$$A^{-1}z = \mathcal{S}_{A^{-1}Z}[A^{-1}\ell; A^{-1}z_a]\,. \tag{6.48}$$

Since $q + A^{-1}z = A^{-1}\ell$, the corresponding play operator becomes

$$q = \mathcal{P}_{A^{-1}Z}[A^{-1}\ell; A^{-1}z_a]\,. \tag{6.49}$$

From the vector stop to the energetic solution. Let H be a real Hilbert space, let $Z \subset H$ be closed and convex. In order to arrive at the setting above where the $\mathcal{S}_Z$ takes values in the dual Q^* of a space Q where the energy and dissipation potentials are defined, we employ the canonical isomorphism $j : H \to H^{**}$ between H and its bidual, defined by

$$\langle j(h), h^* \rangle_{H^{**}H^*} = \langle h^*, h \rangle_{H^*H}\ , \quad h \in H,\, h^* \in H^*\,.$$

This generates the canonical scalar product on the bidual,

$$\langle j(h), j(g) \rangle_{H^{**}} = \langle h, g \rangle_H\ , \quad h, g \in H\,.$$

With the aid of the Riesz isomorphism $r_H : H \to H^*$,

$$\langle r_H(h), g\rangle_{H^*H} = \langle h, g\rangle_H \ , \quad h, g \in H \, ,$$

we define a scalar product in H^* by

$$\langle r_H(h), r_H(g)\rangle_{H^*} = \langle h, g\rangle_H \ , \quad h, g \in H \, . \tag{6.50}$$

This scalar product gives rise to another Riesz isomorphism $r_{H^*} : H^* \to H^{**}$ defined analogously. One checks from the definitions that

$$j = r_H^* \circ r_H$$

and that

$$\left\langle \tilde{h}, \tilde{g} \right\rangle_{H^{**}} = \left\langle \tilde{h}, r_{H^*}^{-1}(\tilde{g}) \right\rangle_{H^{**}H^*} \ , \quad , \tilde{h}, \tilde{g} \in H^{**} \, . \tag{6.51}$$

Given a stop operator $\mathcal{S}_Z$ on a Hilbert space H, we now construct corresponding energy and dissipation potentials. Using the canonical embedding $j : H \to H^{**}$, we write down the equivalent variational inequality in the bidual,

$$\begin{gathered} \left\langle \dot{z}(t) - \dot{\ell}(t), \zeta - z(t) \right\rangle_{H^{**}} \geq 0 \, , \quad \forall \, \zeta \in Z, \text{ for a.a. } t \in (a, b) \\ z(t) \in Z \quad \forall t \in [a, b] \, , \quad z(a) = z_a \, . \end{gathered} \tag{6.52}$$

Here, we write Z instead of $j(Z)$ and so on. We now set $Q = H^*$, endowed with the scalar product (6.50). We set $A = r_{H^*} : Q \to Q^* = H^{**}$ and $\mathcal{E}(t, p) = 1/2 \, \langle Ap, p\rangle - \langle \ell(t), p\rangle$. We define $\mathcal{R} : Q \to [0, +\infty]$ as above by

$$\mathcal{R}(v) = \sup_{\zeta \in Z} \langle \zeta, v\rangle_{Q^*Q} \ .$$

The variational inequality (6.52) then becomes, by virtue of (6.51),

$$\begin{gathered} \left\langle \dot{z}(t) - \dot{\ell}(t), \zeta - z(t) \right\rangle_{A^{-1}} \geq 0 \, , \quad \forall \, \zeta \in Z, \text{ for a.a. } t \in (a, b) \\ z(t) \in Z \quad \forall t \in [a, b] \, , \quad z(a) = z_a \, . \end{gathered} \tag{6.53}$$

Thus, the energetic solution $q \in W^{1,1}(a, b; Q)$ corresponds to $z = \mathcal{S}_Z[\ell; z_a]$ via $z = \ell - Aq$, compare (6.44).

Existence proof for the vector stop operator. The existence of a solution of the variational inequality defining the vector stop now follows from the existence of an energetic solution obtained in Theorem 6.1.

Uniqueness and stability of the energetic solution. Since any energetic solution q generates a solution $z = \ell - Aq$ of the variational inequality, uniqueness and stability results for the latter can be transferred immediately to the energetic solution. Let $\ell_1, \ell_2 \in W^{1,1}(a, b; Q^*)$ and $z_{a,1}, z_{a,2}$ be given. Let z_1, z_2 be the corresponding solutions of the variational inequality (6.44) and $q_i = A^{-1}(\ell_i - z_i)$, $i = 1, 2$. Let us denote $\Delta\ell = \ell_2 - \ell_1$, $\Delta z = z_2 - z_1$, $\Delta q = q_2 - q_1$. From Proposition 5.6 we obtain

$$|\Delta z(t)|_{A^{-1}} \leq |\Delta z_a|_{A^{-1}} + \int_0^t |\Delta \dot{\ell}(s)|_{A^{-1}} \, ds \, , \quad t \in [a, b] \, . \tag{6.54}$$

We have $A\Delta q = \Delta\ell - \Delta z$ and

$$|\Delta q(t)|_A^2 = \langle A\Delta q(t), \Delta q(t)\rangle = \big\langle \Delta\ell(t) - \Delta z(t), A^{-1}(\Delta\ell(t) - \Delta z(t))\big\rangle = |\Delta\ell(t) - \Delta z(t)|_{A^{-1}}^2 .$$

We then get, using (6.54) and the estimate $|\Delta\ell - \Delta z| \le |\Delta\ell| + |\Delta z|$,

$$|q_2(t) - q_1(t)|_A \le |\ell_2(a) - \ell_1(a)|_{A^{-1}} + |z_{a,2} - z_{a,1}|_{A^{-1}} + 2\int_a^t |\dot\ell_2(s) - \dot\ell_1(s)|_{A^{-1}}\, ds . \quad (6.55)$$

One may compare this with the estimate in Proposition 3.5.5 of [17].

Appendix to Section 6. Let $\mathcal{R} : Q \to [0, +\infty]$ be lower semicontinuous, convex, positively 1-homogeneous with $\mathcal{R}(0) = 0$.

Let $v \in L^1(a, b; Q)$ be arbitrary, let $Z = \partial\mathcal{R}(0)$. Then

$$\begin{aligned}\mathcal{R}\Big(\int_a^b v(t)\, dt\Big) &= \sup_{\zeta\in Z}\Big\langle \zeta, \int_a^b v(t)\, dt\Big\rangle = \sup_{\zeta\in Z}\int_a^b \langle\zeta, v(t)\rangle\ dt \\ &\le \int_a^b \sup_{\zeta\in Z}\langle\zeta, v(t)\rangle\ dt = \int_a^b \mathcal{R}(v(t))\, dt .\end{aligned} \quad (6.56)$$

Let $q \in W^{1,1}(a, b; Q)$. let $\Delta = \{t_j\}$ be a partition of $[a, b]$, $I_j = [t_{j-1}, t_j]$. Then

$$\sum_j \mathcal{R}(q(t_j) - q(t_{j-1})) = \sum_j \mathcal{R}\Big(\int_{I_j} \dot q(t)\, dt\Big) \le \sum_j \int_{I_j} \mathcal{R}(\dot q(t))\, dt = \int_a^b \mathcal{R}(\dot q(t))\, dt . \quad (6.57)$$

As a consequence,

$$\sup_\Delta \sum_j \mathcal{R}(q(t_j) - q(t_{j-1})) \le \int_a^b \mathcal{R}(\dot q(t))\, dt . \quad (6.58)$$

Let $\Delta = \{t_j\}$ be a partition of $[a, b]$, let q_Δ be the piecewise affine interpolant of q w.r.t. Δ, so $q(t_j) = q_\Delta(t_j)$ for all j. Then

$$\begin{aligned}\sum_j \mathcal{R}(q(t_j) - q(t_{j-1})) &= \sum_j \mathcal{R}(q_\Delta(t_j) - q_\Delta(t_{j-1})) = \sum_j (t_j - t_{j-1})\mathcal{R}(\dot q_{\Delta,j}) = \sum_j \int_{I_j} \mathcal{R}(\dot q_{\Delta,j})\, dt \\ &= \int_a^b \mathcal{R}(\dot q_\Delta)\, dt .\end{aligned} \quad (6.59)$$

We now need that for some sequence $\{\Delta^N\}$ of partitions and $q^N := q_{\Delta^N}$ we have $\dot q^N(t) \rightharpoonup \dot q(t)$ a.e. in (a, b). (In fact, we even have $\dot q^N(t) \to \dot q(t)$ a.e., as a consequence of the fact that almost all points in (a, b) are Lebesgue points of $\dot q$.) Then

$$\begin{aligned}\int_a^b \mathcal{R}(\dot q(t))\, dt &\le \int_a^b \liminf_{N\to\infty} \mathcal{R}(\dot q^N(t))\, dt \le \liminf_{N\to\infty} \int_a^b \mathcal{R}(\dot q^N(t))\, dt \\ &= \liminf_{N\to\infty} \sum_j \mathcal{R}(q(t_j^N) - q(t_{j-1}^N)) \le \sup_\Delta \sum_j \mathcal{R}(q(t_j) - q(t_{j-1})) .\end{aligned} \quad (6.60)$$

References

[1] H. Bauschke, P. Combettes: Convex analysis and monotone operator theory in Hilbert spaces. Springer 2011.

[2] V. Barbu: Nonlinear differential equations of monotone types in Banach space. Springer 2010.

[3] V. Barbu: Optimal control of variational inequalities. Pitman 1984.

[4] G. Bertotti, I. Mayergoyz: The science of hysteresis, 3 volumes, Academic Press 2006.

[5] H. Brézis: Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert. North-Holland 1973.

[6] M. Brokate, J. Sprekels: Hysteresis and phase transitions. Springer 1996.

[7] P. Colli, A. Visintin: On a class of doubly nonlinear evolution equations. Comm. Partial Differential Eq. 15 (1990), 737–756.

[8] L.C. Evans: Partial differential equations. American Mathematical Society, Providence, 2010.

[9] T. Kato: Nonlinear semigroups and evolution equations. J. Math. Soc. Japan 19 (1967), 508 – 520.

[10] Y. Komura: Nonlinear semigroups in Hilbert space. J. Math. Soc. Japan 19 (1967), 493 – 507.

[11] Y. Komura: Differentiability of nonlinear semigroups. J. Math. Soc. Japan 21 (1969), 375 – 402.

[12] M. A. Krasnoselskii, A. V. Pokrovskii: Systems with hysteresis. Springer 1989. Russian original: Nauka 1983.

[13] P. Krejčí: Hysteresis, convexity and dissipation in hyperbolic equations. Gakkotosho, Tokyo 1996.

[14] P. Krejčí: Evolution variational inequalities and multidimensional hysteresis operators. In: *Nonlinear Differential Equations* (P. Drábek, P. Krejčí, P. Takáč eds.), Research Notes in Mathematics, Vol. 404, Chapman & Hall/CRC, London, 1999, 47–110.

[15] I. Mayergoyz: Mathematical models of hysteresis. Springer 1991.

[16] A. Mielke: Evolution of rate-independent systems. In: Handbook of Differential Equations, Evolutionary Equations, vol. 2 (eds. C.M. Dafermos and E. Feireisl), Elsevier 2005, 461–559.

[17] A. Mielke, T. Roubíček: Rate-independent systems. Springer 2015.

[18] N.S. Papageorgiu, S.T. Kyritsi-Yiallourou: Handbook of applied analysis. Springer 2009.

[19] T. Roubíček: Nonlinear partial differential equations with applications. Birkhäuser 2013.

[20] R.E. Showalter: Monotone operators in Banach space and nonlinear partial differential equations. American Mathematical Society, 1997.

[21] A. Visintin: Differential models of hysteresis. Springer 1994.

Additional literature related to derivatives and optimal control:

[22] M. Brokate: Optimal control of ordinary differential equations with nonlinearities of hysteresis type. Peter Lang, Frankfurt am Main, 1987. (In German.)

[23] M. Brokate: Newton and Bouligand derivatives of the scalar play and stop operator. arXiv:1607.07344 version 2 (2019). Submitted.

[24] M. Brokate, K. Fellner, M. Lang-Batsching: Weak differentiability of the control-to-state mapping in a parabolic equation with hysteresis. Nonlinear Differ. Equ. Appl. 26 (2019), no.6, 26 – 46.

[25] M. Brokate, P. Krejčí: Optimal control of ODE systems involving a rate independent variational inequality. Discrete Continuous Dyn. Syst. Ser. B 18 (2013), 331–348.

[26] M. Brokate, P. Krejčí: Weak differentiability of scalar hysteresis operators. Discrete Continuous Dyn. Syst. Ser. A 35 (2015), 2405–2421.

[27] M. Delfour, J.-P. Zolesio: Shapes and geometries. Analysis, differential calculus and optimization. SIAM, Philadelphia, 2001. See also the 2nd edition, SIAM 2011.

[28] D. Fraňková: Regulated functions. Math. Bohem. 116 (1991), 20–59.

[29] S. Fitzpatrick, R.R. Phelps: Differentiability of the metric projection in Hilbert space. Trans. Amer. Math. Soc. 270 (1982), 483–501.

[30] I.V. Girsanov: Lectures on mathematical theory of extremum problems. Springer, Berlin, 1972.

[31] M. Hintermüller, K. Ito, K. Kunisch: The primal-dual active set strategy as a semismooth Newton method. SIAM J. Opt. 13 (2003), 865–888.

[32] M. Hintermüller, K. Kunisch: PDE-constrained optimization subject to pointwise constraints on the control, the state, and its derivatives. SIAM J. Opt. 20 (2009), 1133–1156.

[33] R.B. Holmes: Smoothness of certain metric projections on Hilbert space. Trans. Amer. Math. Soc. 184 (1973), 87–100.

[34] S. Hu, N.S. Papageorgiu: Handbook of multivalued analysis, volume I: theory. Kluwer Academic Publishers 1997.

[35] K. Ito, K. Kunisch: Lagrange multiplier approach to variational problems and applications. SIAM Series Advances in Design and Control, SIAM, Philadelphia, 2008.

[36] F. Mignot: Contrôle dans les inéquations variationelles elliptiques. J. Funct. Anal. 22 (1976), 130–185.

[37] M. Ulbrich: Semismooth Newton methods for operator equations in function spaces. SIAM J. Optim. 13 (2003), 805–841.

[38] M. Ulbrich: Semismooth Newton methods for variational inequalities and constrained optimization problems in function spaces. SIAM, Philadelphia, 2011.