On rate–independent hysteresis models

Alexander Mielke
Mathematisches Institut A
Universität Stuttgart
D–70569 Stuttgart
mielke@mathematik.uni-stuttgart.de

Florian Theil
Mathematics Institute
University of Warwick
Coventry CV4 7AL, U.K.
theil@maths.warwick.ac.uk

5 April 2001 / revised 21 September 2001

1 Introduction

This paper deals with a general approach to the modeling of rate–independent processes which may display hysteretic behavior. Such processes play an important role in many applications like plasticity and phase transformations in elastic solids, electromagnetism, dry friction on surfaces, or in pinning problems in superconductivity, cf. [Vis94, BrS96].

The evolution equations which govern those processes constitute the limit problems if the influence of inertia and relaxation times vanishes, i.e. the system rests unless the external loading is varied. Only the stick–slip dynamics is present in the Cauchy problem, this means that the evolution equations are necessarily non–autonomous. Although the solutions often exhibit quite singular behavior, the reduced framework offers great advantages.

Firstly the amount of modelling can be reduced to its absolute minimum. More importantly, our approach is only based on energy principles. This allows us to treat the Cauchy problem by mainly using variational techniques. This robustness is necessary in order to study problems which come from continuum mechanics like plasticity, cf. [Mie00, CHM01, Mie01]. There the potential energy is invariant under the group of rigid body rotations $\mathrm{SO}(d)$ where $d \in \{1, 2, 3\}$ is the dimension. This invariance implies that convexity can almost never be expected and more advanced lower semicontinuity results (like polyconvexity) are required to assure the existence of solutions for a time discretized version of the problem.

An example which illustrates this remark is a problem from phase transformations in solids, see [MTL00]. Although none of the classical methods from Section 7 can be applied, we are able to prove the existence of solutions by establishing weak lower semicontinuity of certain critical quantities.

Here we present an abstract framework which is based on two energy functionals, namely the potential energy $I(t, z)$ and the dissipation $\Delta(\dot{z})$. Here $z \in X$, X a separable, reflexive Banach space with dual X^*, is the variable describing the process, and $\dot{z}$ is the time derivative. The central feature of rate–independence means that a solution $z : [0, T] \to X$ remains a solution if the time is rescaled. This leads to a dissipation functional $\Delta : X \to [0, \infty)$ which is homogeneous of degree 1, i.e., $\Delta(\alpha v) = \alpha\Delta(v)$ for $\alpha \geq 0$ and $v \in X$.

Special cases of this situation are well studied in the theory of variational inequalities or as sweeping processes. There we have the potential energy

$$I(t,z)=\frac{1}{2}\langle Az,z\rangle-\langle g^*(t),z\rangle \tag{1.1}$$

with $A\in \mathrm{Lin}(X,X^*)$ symmetric and positive definite and the dissipation has the form

$$\Delta(v)=\max\{\langle w^*,v\rangle : w^*\in F^*\} \tag{1.2}$$

where $F^*\subset X^*$ is a bounded, closed, convex set containing 0 in its interior. The variational inequality problem then is to find $z\in W^{1,1}([0,T],X)$ such that for a.a. $t\in[0,T]$

$$\langle Az{-}g,v{-}\dot z\rangle+\Delta(v)-\Delta(\dot z)\geq 0 \text{ for all } v\in X. \tag{1.3}$$

The equivalent sweeping process formulation is

$$\dot z\in\partial\chi_{F^*}(g{-}Az)\quad (\text{or } -\dot z\in\partial\chi_{g^*-F^*}(Az)). \tag{1.4}$$

Here χ_{F^*} is the (convex) characteristic function of F^* taking the value 0 on F^* and ∞ else, and ∂ denotes the subdifferential. We refer to [Mor77, Mon93, KuM97, KuM98] for more details on this matter.

We treat these problems in Section 7 under the assumption, that I satisfies certain smoothness assumptions. The purpose of this work is to broaden the model class significantly. On the one hand we want to be able to deal with nonconvex or not strictly convex potentials. This leads to the possibility that the solutions are discontinuous, i.e. jumps can occur. Such singularities indicate that our purely energetic formulation is too simplistic to describe physical phenomena since the dynamics is not slow anymore. As it stands, our formulation favours discontinuities, i.e. the state will jump as soon as possible, independently of the shape of the energy landscape between the left hand and the right hand limit. We have to pay this price to obtain energy conservation independently from the regularity of the solutions.

On the other hand we want to allow for restrictions of the state space as well, i.e. we want to be able to restrict $z(t)$ to a closed convex set $E\subset X$. A large part of this work is new even in the case of finite dimensional spaces X.

Our point of departure is the following energy formulation of the problem: Find $z:[0,T]\to E$ with

(S) For a.a. $t\in[0,T]$ we have
$I(t,z(t))\leq I(t,y)+\Delta(y{-}z(t))$ for all $y\in E$, (1.5)
(E) $I(t,z(t))+\int_s^t\Delta(\mathrm{d}z)\leq I(s,z(s))+\int_s^t\partial_t I(\tau,z(\tau))\mathrm{d}\tau$ for all $0\leq s\leq t\leq T$.

The first condition (S) can be interpreted as global stability of the state $z(t)$ at time t: by changing $z(t)\in E$ into $y\in E$ the release of potential energy $I(t,z(t))-I(t,y)$ is never larger than the associated dissipated energy. The second condition (E) is the integrated energy balance between final, dissipated, initial energy and the "work done by external forces" in every time interval $[s,t]\subset[0,T]$ (often $\partial_t I(t,z)=-\langle\dot g^*(t),z(t)\rangle$).

In Section 3 we show that formulation (1.5) is more general than the local formulation (1.3) but both formulations are equivalent if $I(t,\cdot):X\to\mathbb{R}$ is convex and satisfies further technical conditions. Moreover, for convex problems we establish equivalence to the subdifferential formulation

$$0 \in \partial\Delta(\dot{z}) + \partial I(t,z), \tag{1.6}$$

which was introduced in [CoV90, Vis01] and called "doubly nonlinear problem".

However, these formulations have to be treated with care since in the nonconvex or not strictly convex case we have to allow for jumps of z. Hence, the right solution space is $\mathrm{BV}_-([0,T],X)$, the space of left–continuous functions having finite total variation. For an exact formulation of (1.6) we need to replace the derivative $\dot{z}$ by a reduced derivative $\mathrm{rd}(z) : [0,T] \to \{v \in X : \|v\| \le 1\}$ and an associated Radon measure $\mu_z \in \mathcal{M}([0,T])$ such that

$$z(t_2) - z(t_1) = \int_{[t_1,t_2)} \mathrm{rd}(z)(r)\mu_z(\mathrm{d}r). \tag{1.7}$$

holds which means that $\mathrm{rd}(z)(t)\mu_z(\mathrm{d}t)$ plays the role of the time–derivative which is a vector–valued measure. The correct replacement of (1.6) is then

$$0 \in \partial\Delta(\mathrm{rd}(z)(t)) + \partial I(t,z(t)) \text{ for } \mu_z\text{–a.a. } t \in [0,T]. \tag{1.8}$$

In Section 4 we analyse a general method for obtaining piecewise constant in time approximations which works for nonconvex problems as well. It is based on the incremental problem for the time discretization $0 = t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N = T$

$$\left\{\begin{array}{l} \text{Let } z_0 = z(0). \text{ For } k = 1, \ldots N \text{ find } z_k \in E \text{ with} \\ I(t_k, z_k) \le \inf\{I(t_k,y) + \Delta(y - z_{k-1}) : y \in E\}. \end{array}\right. \tag{1.9}$$

This approach leads in a natural way to discretized versions of energy formulation (1.5) (see Theorem 4.1) and, in the convex case, of the variational inequality (1.3) (see (4.3)).

A central object in this study are the sets $\mathcal{S}(t)$ of stable states at time t, namely

$$\mathcal{S}(t) = \{z \in E : I(t,z) \le I(t,y) + \Delta(y - z) \text{ for all } y \in E\}.$$

The major problem arises from $\mathcal{S}(t)$ not being weakly closed since even for convex $I(t,\cdot)$ the set $\mathcal{S}(t)$ may not be convex.

Assuming weak closedness of the set of stable states $\mathcal{S}$ plus some regularity conditions on I we give an existence proof in Theorem 6.3. This proof relies on the incremental problem and on an abstract version of Helly's selection principle (cf. [BaP86]). Here we do not require that $I(t,\cdot)$ is smooth, especially the assumption $E = X$ which is central for Theorem 7.1 is not necessary.

In Section 7 for $E = X$ and $I(t,\cdot)$ uniformly convex we prove existence, uniqueness (Theorem 7.1) and discuss the question of temporal smoothness of solutions. To this end we show (Theorem 7.3) that the solutions of the incremental problem converge strongly, by generalizing ideas in [HaR95]. As far as the authors know there is no uniqueness result even for nontrivial $E \subset X$ $X = \mathbb{R}^2$. We propose a new "structure condition" which implies uniqueness, cf. Section 7.2 and Appendix C.

We note that in [HaR95] a convergence result of the incremental problem (for the case (1.1) only) is given only under the assumption $z \in W^{2,1}((0,T),X)$, i.e. $\ddot{z} \in L^1((0,T),X)$.

Our above–mentioned convergence result Theorem 7.3 is independent of this assumption. From very simple examples we see that, even for the case $E = X$, we can only expect $\dot{z} \in \mathrm{BV}((0,T),X)$. Our Theorem 7.8 proves this under the additional assumption that the boundary of F^* is C^2.

For many applications (see e.g. [KuM00, MTL00, GoM00]) the assumption of reflexitivity of X is too restrictive. We note that the formulation (1.5) via energy functionals provides us with the advantage that the time derivative $\dot{z}$ is not required. Moreover, the existence theorem 6.3 uses only the assumption of weak sequential compactness which is considerably weaker, see Remark 6.4.

Throughout this work we have assumed that the underlying geometry is the linear space X, in particular the dissipation functional Δ does not depend on the position $z \in E \subset X$. A more general treatment should involve Banach manifolds and the dissipation is then a Finsler metric. For applications in this context see [Mie00, Mie01].

2 Notations and setup of the problem

Let X be a separable, reflexive Banach space and E a closed convex subset. For each $z \in E$ the (inward) tangential cone $\mathrm{T}_z E$ is defined via

$$\mathrm{T}_z E = \overline{\{w \in X : \exists r > 0 : z + rw \in E\}}, \tag{2.1}$$

where $\overline{A}$ is the closure of A in the strong topology. Later on we will also use the (outward) normal cone $\mathrm{N}^*_z E \subset X^*$ defined via

$$\mathrm{N}^*_z E = \{y^* \in X^* : \langle y^*, w\rangle \le 0 \,\forall\, w \in \mathrm{T}_z E\}.$$

The dissipation is implemented by a function $\Delta : X \to [0,\infty)$ which is convex, homogeneous of degree 1 (i.e. $\Delta(\alpha z) = \alpha\Delta(z)$ for $\alpha \ge 0$ and $z \in X$) and satisfies, for some $C^{(2)}_\Delta \ge C^{(1)}_\Delta \ge 1$,

$$C^{(1)}_\Delta \|z\| \le \Delta(z) \le C^{(2)}_\Delta \|z\| \text{ for all } z \in X. \tag{2.2}$$

Convexity and homogeneity of degree 1 imply the triangle inequality

$$\Delta(z+\tilde{z}) \le \Delta(z) + \Delta(\tilde{z}) \text{ for all } z, \tilde{z} \in X \tag{2.3}$$

which will be used often subsequently.

These assumptions are equivalent to the existence of a convex closed set F^* in X^* with $\{z^* \in X^* : \|z^*\| \le C^{(1)}_\Delta\} \subset F^* \subset \{z^* \in X^* : \|z^*\| \le C^{(2)}_\Delta\}$ such that

$$\Delta(z) = \max\{\langle z^*, z\rangle : z^* \in F^*\}. \tag{2.4}$$

We continue to use $\langle\cdot,\cdot\rangle$ for the duality pairing on $X^* \times X$.

The space of functions of bounded variations is defined here to be

$$\mathrm{BV}([0,T],X) = \{z : [0,T] \to X : \mathrm{Var}(z,[0,T]) < \infty\}$$

where $\mathrm{Var}(z,[0,T]) = \sup\{\sum_{k=1}^N \|z(t_k)-z(t_{k-1})\| : 0\le t_0<t_1<\ldots<t_N\le T\}$. Functions in $\mathrm{BV}([0,T],X)$ are continuous except for an at most countable number of jump points at which the right and left limits $z_+(t) = \lim_{s\searrow t} z(s)$ and $z_-(t) = \lim_{s\nearrow t} z(s)$ exist. We define the closed subspaces

$$\mathrm{BV}_\pm([0,T],X) = \{z \in \mathrm{BV}([0,T],X) : z = z_\pm\},$$

which will be used subsequently. For functions $z = z_+$ we attach artificially a limit from the left which we denote by $z(0-)$. This will allow us to impose an initial condition $z_+(0-) = z_0$ at time $t = 0$ even when $z_+(0) \neq z_0$. For $z = z_-$ a corresponding construction is not necessary as $z_0 = z(0) = z(0-) \neq z_+(0)$ is allowed, which also corresponds to a jump at time 0.

As explained in detail in Appendix A, to each $z \in \mathrm{BV}_\pm([0,T],E)$ we can associate a derivative which is the product $\mathrm{rd}(z)\mu_z$ of the reduced derivative $\mathrm{rd}(z) : [0,T] \to \{\, v \in X : \|v\| \le 1\,\}$ and the differential measure $\mu_z \in \mathcal{M}([0,T])$ (the set of Radon measures in $[0,T]$) . The two components are defined via

$$\mu_z([s,t)) = t-s + \int_{[s,t)} \|\mathrm{d}z\| \quad \text{and} \quad z(t)-z(s) = \int_{[s,t)} \mathrm{rd}(z)(r)\mu_z(\mathrm{d}r).$$

Here $\mathrm{rd}(z) : [0,T] \to X$ is defined only μ_z–almost everywhere (a.e.). Note that our definition of the differential measure differs from that in [Mon93, Sect.0.1] where $\nu_z([s,t)) = \int_{[s,t)} \|\mathrm{d}z\|$ and $\|\mathrm{nd}(z)(t)\| = 1$ ν_z–a.e. in $[0,T]$ such that $\mathrm{d}z = \mathrm{nd}(z)\nu_z = \mathrm{rd}(z)\mu_z$. The derivative doesn't distinguish between right and left continuous versions, i.e., $\mathrm{rd}(z_+) = \mathrm{rd}(z_-)$ and $\mu_{z_-} = \mu_{z_+}$. For $z = z_+$ with initial datum $z(0-) = z_0 \neq z(0)$ we have $\mu_z(\{0\}) = \|z(0)-z_0\| =: r > 0$ and $\mathrm{rd}(z)(0) = \frac{1}{r}[z(0)-z_0]$.

For $z \in \mathrm{BV}_-([0,T],X)$ the Δ–variation on the interval $J \subset [0,T]$ is defined by

$$\mathrm{Var}_\Delta(z,J) = \int_J \Delta(\mathrm{d}z) \stackrel{\text{def}}{=} \int_J \Delta(\mathrm{rd}(z)(t))\mu_z(\mathrm{d}t).$$

This is the same as the supremum over all sums of the form $\sum_{k=1}^N \Delta(z(t_k)-z(t_{k-1}))$ where $N \in \mathbb{N}$, $t_0, t_N \in J$ and $t_0<t_1<\ldots<t_N$. Note that we have to be careful about jumps at the boundary of J if J contains the corresponding boundary point. In particular, for $z = z_+$ special care has to be taken for the left limit $z(0-)$, since our definition implies $\mathrm{Var}_\Delta(z,[0,T]) = \Delta(z(0)-z(0-)) + \mathrm{Var}_\Delta(z,(0,T])$.

On E the time–dependent energy functional $I(t,z)$ is defined such that

$$I \in \mathrm{C}^1([0,T]\times E,[0,\infty)),$$

where implicitly we have assumed that I is bounded from below by some constant, which was set to 0 without loss of generality. The main assumption on $I(t,\cdot)$ is weak lower semi–continuity:

$$z_n \rightharpoonup z \text{ implies } I(t,z) \le \liminf_{n\to\infty} I(t,z_n). \tag{2.5}$$

With respect to the time dependence we assume that there exist constants $C_1, C_2 > 0$ such that

$$\begin{array}{ll} (a) & |I(t,z)-I(\widehat{t},z)| \le C_1|t-\widehat{t}|, \\ (b) & |\partial_t I(t,z)-\partial_t I(t,\widehat{z})| \le C_2\|z-\widehat{z}\|, \\ (c) & z_n \rightharpoonup z \text{ implies } \partial_t I(t,z_n) \to \partial_t I(t,z), \end{array} \tag{2.6}$$

for all $t, \widehat{t} \in [0,T]$ and $z, \widehat{z} \in E$.

Further structural assumptions on I are convexity assumptions on $I(t,\cdot)$. We say that I is convex, strictly convex or α–uniformly convex if for all $t \in [0,T]$ and all z_0, z_1 the following conditions hold:

convex:	$I(t,z_\theta) \le (1-\theta)I(t,z_0) + \theta I(t,z_1)$ for all $\theta \in [0,1]$;
strictly convex:	$I(t,z_\theta) < (1-\theta)I(t,z_0) + \theta I(t,z_1)$ for all $\theta \in (0,1)$;
α–unif. convex:	$I(t,z_\theta) \le (1-\theta)I(t,z_0) + \theta I(t,z_1) - \frac{\alpha}{2}(\theta-\theta^2)\|z_0-z_1\|^2$ for all $\theta \in [0,1]$;

where $z_\theta = (1-\theta)z_0 + \theta z_1$. These convexity conditions are never assumed without stating explicitly.

The most general formulation for our rate–independent processes is global in time. We have a set of *stable* states which depends on time, in addition a solution satisfies the *energy inequality*. This formulation does not rely on convexity assumptions for I and is particularly useful for problems where only generalized convexity notions like quasiconvexity hold, since weak lower semi–continuity is still valid. Moreover it does not involve derivatives of I with respect to z.

(GF) [Global Formulation] *Find $z \in \mathrm{BV}_\pm([0,T],X)$ with $z(0-) = z_0$ and $z(t) \in E$ such that conditions (S) and (E) hold:*

(S) for λ–a.a. $t \in [0,T]:\ I(t,z(t)) \le I(t,y) + \Delta(y-z(t))$ for all $y \in E$,
(E) for all $0 \le t_1 \le t_2 \le T:\ I(t_2,z(t_2)) + \int_{t_1}^{t_2} \Delta(\mathrm{d}z) \le I(t_1,z(t_1)) + \int_{t_1}^{t_2} \partial_t I(s,z(s))\mathrm{d}s$,

where $\int_{t_1}^{t_2} \Delta(\mathrm{d}z) = \int_{[t_1,t_2)} \Delta(\mathrm{d}z)$ in the case $z \in \mathrm{BV}_-([0,T],X)$ and accordingly if $z \in \mathrm{BV}_+([0,T],X)$

Here (S) is the condition of global stability in the whole state space E, and λ denotes the one–dimensional Lebesgue measure.

The definition of the energy inequality (E) is such that it implies the two natural requirements for evolutionary problems, namely *restrictions* and *concatenations* of solutions remain solutions. To be more precise consider a solution $z : [0,T] \to E$ and any subinterval $[s,t] \subset [0,T]$, then the restriction $z|_{[s,t]}$ solves (GF) with initial datum $z(s-)$. Moreover, if $z_1 : [0,t] \to E$ and $z_2 : [t,T] \to E$ solve (GF) on the respective intervals and if $z_1(t-) = z_2(t-)$ then the concatenation $z : [0,T] \to E$ solves (GF) as well.

In particular, (S) and (E) imply that if z jumps at time t from z_- to z_+ then

$$I(t,z_+) + \Delta(z_+-z_-) = I(t,z_-). \tag{2.7}$$

3 Three alternative formulations of the problem

There exist three different formulations which are equivalent to (GF) if the potential energy I is convex. In the nonconvex case only certain implications are correct. The global formulation has the big advantage that it doesn't need the differentiability of the potential energy $I(t,\cdot)$. The other three formulations involve the derivative $\mathrm{D}I(t,z(t))$.

The first formulation (LF) localizes both, the definition of stable sets and the energy inequality. It is somehow impractical since the structure of a constrained evolution is still present; in particular, the tangent spaces $\mathrm{T}_{z(t)}E$ depend discontinuously on $z(t)$.

The second formulation (VI) is a variational inequality which is a standard rewriting of the local formulation (LF). Hence, it is always equivalent to (LF).

The third formulation (SF) is independent from these ideas. It is a single evolution equation without constraints which has the difficulty that the time derivative appears inside of a strong nonlinearity.

The following formulations will make essential use of knowledge on the reduce derivative which is contained in the following lemma. This result explains why certain of our formulations are more natural for $z = z_-$ while others are more natural for $z = z_+$. We will always indicate this by adding a subscript $_-$ or $_+$ to the formulation. However, for the global formulation (GF) this is not necessary, as it is easily seen that $z = z_-$ solves $(\mathrm{GF})_-$ if and only if $w = z_+$ solves $(\mathrm{GF})_+$.

Lemma 3.1 *For $z \in \mathrm{BV}_\pm([0,T],X)$ with $z : [0,T] \to E$ there exists a set $\mathcal{T} \subset [0,T]$ of full μ_z–measure such that for all $t \in \mathcal{T}$ we have*

$$\mathrm{rd}(z)(t) \in \mathrm{T}_{z_-(t)}E \quad \textit{and} \quad -\mathrm{rd}(z)(t) \in \mathrm{T}_{z_+(t)}E.$$

Proof. We use the stretched function $\widehat{z} \in \mathrm{C}^{\mathrm{Lip}}([0,\widehat{T}],X)$ associated to z which is defined in (A.1). Define the set $\widetilde{\mathcal{T}} \subset [0,\widehat{T}]$ to be the set of points τ where we have

$$\widehat{z}'(\tau) = \tfrac{\mathrm{d}}{\mathrm{d}\tau}\widehat{z}(\tau) = \lim_{\rho\searrow 0} \tfrac{1}{\rho}[\widehat{z}(\tau{+}\rho) - \widehat{z}(\tau)] = \lim_{\rho\searrow 0} \tfrac{1}{\rho}[\widehat{z}(\tau) - \widehat{z}(\tau{-}\rho)].$$

By Lebesgue's lemma $\widehat{\mathcal{T}}$ has full Lebesgue measure, i.e., $\lambda(\widehat{\mathcal{T}}) = \lambda([0,\widehat{T}]) = \widehat{T}$, cf. [Mon93]. Moreover, $\widehat{z}(\tau) \pm \rho\widehat{z}'(\tau) + o(\rho) = \widehat{z}(\tau{\pm}\rho) \in E$ implies $\pm\widehat{z}'(\tau) \in \mathrm{T}_{\widehat{z}(\tau)}E$ for all $\tau \in \widehat{\mathcal{T}}$.

The desired result is now obtained by undoing the stretching using the mapping $\widehat{t} : \tau \mapsto t$ and the set $\mathcal{T} = \widehat{t}(\widehat{\mathcal{T}}) \subset [0,T]$. We find $\mu_z(\mathcal{T}) = \lambda(\widehat{\mathcal{T}}) = \widehat{T} = \mu_z([0,T])$. For continuity points of z we have $\mathrm{rd}(z)(t) = \widehat{z}'(\widehat{\tau}(t))$, see (A.2). At jump points the desired result is obvious as $\mathrm{rd}(z)$ points in the jump direction from $z(t) = z_-(t)$ to $z_+(t)$. ∎

The local formulation needs the assumption that I is a C^1 function. The measure μ_z also appears and thus jump points can be treated suitably. Recall that the set of jump points has Lebesgue measure 0, but each jump point has a positive μ_z–measure.

(LF)$_\pm$ [Local Formulation] *Find $z \in \mathrm{BV}_\pm([0,T],X)$ with $z(0) = z_0$ and $z(t) \in E$ such that conditions ($\mathrm{S_{loc}}$) and ($\mathrm{E_{loc}}$) hold μ_z–a.e. in $[0,T]$:*

$$\begin{aligned} &(\mathrm{S_{loc}}) \quad \langle \mathrm{D}I(t,z(t)), v\rangle + \Delta(v) \geq 0 \textit{ for all } v \in \mathrm{T}_{z(t)}E,\\ &(\mathrm{E_{loc}}) \qquad \langle \mathrm{D}I(t,z(t)), \mathrm{rd}(z)(t)\rangle + \Delta(\mathrm{rd}(z)(t)) \leq 0. \end{aligned}$$

Now restrict to the case $z = z_-$. Using $\mathrm{rd}(z)(t) \in \mathrm{T}_{z(t)}E$ (from Lemma 3.1) and subtracting the two conditions we are lead to the single variational inequality

$$\langle \mathrm{D}I(t,z(t)), v{-}\mathrm{rd}(z)(t)\rangle + \Delta(v) - \Delta(\mathrm{rd}(z)(t)) \geq 0 \quad \text{ for all } v \in \mathrm{T}_{z(t)}E$$

which is equivalent to ($\mathrm{S_{loc}}$) & ($\mathrm{E_{loc}}$). Introducing the characteristic function $\chi_{\mathrm{T}_{z(t)}E}$ we can incorporate the side condition into the variational inequality and obtain the third formulation.

(VI)$_-$ ***[Variational Inequality]*** *Find* $z \in \mathrm{BV}_-([0,T],X)$ *with* $z(0) = z_0$ *and* $z(t) \in E$ *such that*

$$\begin{aligned}\langle \mathrm{D}I(t,z(t)), w - \mathrm{rd}(z)(t)\rangle + \Delta(w) - \Delta(\mathrm{rd}(z)(t))\\ +\chi_{\mathrm{T}_{z(t)}E}(w) - \chi_{\mathrm{T}_{z(t)}E}(\mathrm{rd}(z)(t)) \geq 0 \quad \text{for all } w \in X\end{aligned}$$

holds μ_z–a.e. *in* $[0,T]$.

The fourth formulation uses subdifferentials denoted by ∂. For a characteristic function $\chi_E : X \to [0,\infty]$ we find the normal cone mapping, i.e., $\partial\chi_E(z) = \mathrm{N}^*_z E$ for $z \in X$. The subdifferential of the dissipation functional Δ is given via

$$\partial\Delta(v) = \underset{z^* \in F^*}{\operatorname{argmax}} \langle z^*, v\rangle = \begin{cases} F^* & \text{for } v = 0, \\ \{\, \delta^* \in F^* \;:\; \Delta(v) = \langle \delta^*, v\rangle \,\} & \text{for } v \neq 0. \end{cases} \tag{3.1}$$

(SF)$_\pm$ ***[Subdifferential Formulation]*** *Find* $z \in \mathrm{BV}_\pm([0,T],X)$ *with* $z(0-) = z_0$ *and* $z(t) \in E$ *such that*

$$0 \in \partial\Delta(\mathrm{rd}(z)(t)) + \mathrm{D}I(t,z(t)) + \partial\chi_E(z(t)) \quad \text{for } \mu_z\text{–a.a. } t \in [0,T].$$

This formulation is especially useful for general, convex $I(t,\cdot)$ (not necessarily differentiable, but lower semicontinuous). In writing $\widetilde{I}_E(t,z) = I(t,z) + \chi_E(z)$ we obtain the short form

$$0 \in \partial\Delta(\mathrm{rd}(z)(t)) + \partial\widetilde{I}_E(z(t)) \quad \text{for } \mu_z\text{–a.a. } t \in [0,T]. \tag{3.2}$$

Assuming $\widetilde{I}_E(t,z) = J(z) - \langle g^*(t), z\rangle$ this is exactly the doubly nonlinear formulation of Colli & Visintin [CoV90, Vis01], namely $g^*(t) \in \partial\Delta(\mathrm{rd}(z)(t)) + \partial J(z(t))$.

The first aim of this section is to show that the global formulation (GF) always implies the local formulations (LF)$_-$ and (LF)$_+$ but not vice versa. For the case that $I(t,\cdot)$ is convex and satisfies the above technical assumption we show that all three formulations are equivalent. Second we compare the formulations (SF)$_\pm$ and (LF)$_\pm$ and at the end of the section we present the associated formulation as a sweeping process.

Example 3.2 *The formulation (SF)*$_-$ *is not very useful, as is seen by this simple example. Take* $E = [0,l] \subset \mathbb{R}$ *and* $I(t,z) = \alpha z^2/2 - tz$ *with* $\alpha > 0$. *Choosing* $z_0 = 0$ *the unique solution of (GF), (LF)*$_-$*, (VI)*$_-$ *and (SF)*$_+$ *is given by* $z_{\mathrm{sln}}(t) = \max\{0, \min\{l, (t-1)/\alpha\}\}$ *for* $t \geq 0$. *This function is also a solution of (SF)*$_-$*, however, there are many more solutions: Choose* $\tau \in (0,1]$ *and* $Z \in (0, \max\{l, (t+1)/\alpha\})$. *Then, the function* z_- *with* $z_-(t) = 0$ *for* $t \leq \tau$ *and* $z_-(t) = \max\{Z, z_{\mathrm{sln}}(t)\}$ *for* $t > \tau$ *solves (SF)*$_-$ *as well.*

Example 3.3 *We consider* $E = X = \mathbb{R}$, $\Delta(v) = |v|$ *and a general nonconvex, differentiable potential* $I : [0,T] \times \mathbb{R} \to [0,\infty)$. *We discuss what conditions are imposed on possible jumps from* z_0 *to* z_1 *by the different formulations (GF), (LF)*$_-$ *and (LF)*$_+$*, respectively. In all cases stability leads to the necessary condition* $|\mathrm{D}I(t,z_j)| \leq 1$. *For (GF) we find* $\mathrm{D}I(t,z_0) = \mathrm{D}I(t,z_1) = -\mathrm{sign}(z_1 - z_0)$ *together with the global condition* $I(t,z_0) = I(t,z_1) + |z_1 - z_0| \leq I(t,y) + |y - z_0|$ *for all* $y \in \mathbb{R}$. *For (LF)*$_-$ *we find*

$\mathrm{D}I(t,z_0) = -\mathrm{sign}(z_1-z_0)$ *and* $|\mathrm{D}I(t,z_1)| \leq 1$ *whereas (LF)*$_+$ *gives* $|\mathrm{D}I(t,z_0)| \leq 1$ *and* $\mathrm{D}I(t,z_1) = -\mathrm{sign}(z_1-z_0)$*. The two local formulations (LF)*$_\pm$ *make no statement on the jump of* I*, so generally (E) cannot be recovered.*

Theorem 3.4 (Relation between (GF) and (LF))
(a) For general I *we have (GF)* $\Rightarrow$ *(LF*$^{\text{aver}}$*)*$_\pm$*, where (E*$_{\text{loc}}$*) is replaced by*

$$(E^{\text{aver}}_{\text{loc}}) \qquad \langle B^*(t), \mathrm{rd}(z)(t)\rangle + \Delta(\mathrm{rd}(z)(t)) \leq 0 \quad \mu_z\text{–a.e. on } [0,T], \tag{3.3}$$

with $B^*(t) = \int_0^1 \mathrm{D}I(t,(1-\theta)z_-(t)+\theta z_+(t))\mathrm{d}\theta$.

(b) for general I *we have (LF)*$_-$ $\Leftrightarrow$ *(VI)*$_-$*;*

(c) for convex I *we have (GF)* $\Leftrightarrow$ *(LF)*$_-$ $\Leftrightarrow$ *(LF)*$_+$*.*

Theorem 3.5 (Relation between (LF) and (SF))
(a) For general I *we have (SF)*$_+$ $\Leftrightarrow$ *(LF)*$_+$*;*
(b) for general I *we have (LF)*$_-$ $\Rightarrow$ *(SF)*$_-$*;*
(c) for convex I *we have (GF)* $\Leftrightarrow$ *(LF)*$_\pm$ $\Leftrightarrow$ *(SF)*$_+$*.*

The proofs of both theorems use several intermediate results which are developed now.

Lemma 3.6 *(S) implies (S*$_{\text{loc}}$*).*

Proof. For any $v \in \mathrm{T}_{z(t)}E \setminus \{0\}$ we find a sequence of $w_k \in X$ and a sequence $r_k > 0$ such that $\|w_k\| = \|v\|$, $w_k \to v$, $r_k \to 0$ and $z_k = z(t)+r_k w_k \in E$. From stability (S) and differentiability we conclude

$$\begin{aligned} 0 &\leq \tfrac{1}{r_k}\Big[I(t,z(t)+r_k w_k) + \Delta(r_k w_k) - I(t,z(t))\Big] \\ &= \langle \mathrm{D}I(t,z(t)), w_k\rangle + \Delta(w_k) + o(r_k)_{k\to\infty}. \end{aligned}$$

With $k \to \infty$ we obtain (S$_{\text{loc}}$). ■

The next lemma shows that the energy inequality (E) can be replaced by an energy identity under natural circumstances. In fact, considering the energy $I(t,z(t)) + \int_{[0,t)} \Delta(\mathrm{d}z) - \int_{[0,t)} \partial_t I(\tau,z(\tau))\mathrm{d}\tau$ we see that (S) implies that this energy cannot decrease. The same conclusion follows from (S$_{\text{loc}}$) at continuity points of z.

Lemma 3.7 (Energy conservation) *A process* $z \in \mathrm{BV}_-([0,T],E)$ *satisfies for all* $0 \leq s < t \leq T$ *the energy identity*

$$I(t,z(t)) + \int_{[s,t)} \Delta(\mathrm{d}z) = I(s,z(s)) + \int_{[s,t)} \partial_t I(\tau,z(\tau))\,\mathrm{d}\tau. \tag{3.4}$$

if one of the following two conditions is satisfied:
(1) z *satisfies (S*$_{\text{loc}}$*) and (E*$_{\text{loc}}$*) and* $z \in \mathrm{C}([0,T],E)$*;*
(2) z *satisfies (S) on all of* $[0,T]$ *and (E) holds for* $t_1 = 0$ *and* $t_2 = T$ *only.*

Proof. We define the quantity

$$e(s,t) = I(t,z(t)) + \int_{[s,t)} \Delta(\mathrm{d}z) - I(s,z(s)) - \int_{[s,t)} \partial_t I(\tau,z(\tau))\,\mathrm{d}\tau.$$

By definition we have $e(r,t) = e(r,s) + e(s,t)$ for any $r < s < t$. We use part (c) of Theorem A.1:

$$I(t,z(t)) - I(s,z(s)) = \int_s^t \partial_t I(r,z(r))\,\mathrm{d}r + \int_{[s,t)} \langle B^*(t), \mathrm{rd}(z)(r)\rangle \mu_z(\mathrm{d}r) \tag{3.5}$$

with $B^*(t) = \int_0^1 \mathrm{D}I(t,(1-\theta)z(t)+\theta z_+(t))\mathrm{d}\theta$. This yields

$$e(s,t) = \int_{[s,t)} \Big[\langle B^*(t), \mathrm{rd}(z)(t)\rangle + \Delta(\mathrm{rd}(z)(t))\Big]\mu_z(\mathrm{d}t). \tag{3.6}$$

If z is continuous we have $B^*(t) = \mathrm{D}I(t,z(t))$, and ($\mathrm{S}_{\mathrm{loc}}$) & ($\mathrm{E}_{\mathrm{loc}}$) together with Lemma 3.1 give the first claim.

For the second claim we have that $e(0,T) \leq 0$. We use ($\mathrm{S}_{\mathrm{loc}}$) to show $e(s,t) \geq 0$ for all $s < t$ which gives the desired result $e(s,t) = 0$. Observe that the integrand in (3.6) is nonnegative μ_z–a.e. in $[0,T]$ by (S). Indeed, if z is continuous at t then this follows from ($\mathrm{S}_{\mathrm{loc}}$), and if z has a jump from z_- to z_+ at time t, then

$$\begin{aligned}\textstyle\int_{\{t\}} \Big[\langle B^*(t), \mathrm{rd}(z)(t)\rangle + \Delta(\mathrm{rd}(z)(t))\Big]\mu_z(\mathrm{d}t) &= \langle B^*(t), z_+ - z_-\rangle + \Delta(z_+ - z_-)\\ &= I(t,z_+) - I(t,z_-) + \Delta(z_+ - z_-)\end{aligned}$$

which is 0 by the global stability of z_- at time t, see (2.7). ■

In the proof of Lemma 3.7 we saw that it might be more natural to replace in (LF) the local energy condition ($\mathrm{E}_{\mathrm{loc}}$) by an averaged version (3.3). This unsymmetry disappears in the case of convex $I(t,\cdot)$, as the following result implies $B^*(t) = \mathrm{D}I(t,z(t))$ μ_z–a.e. in $[0,T]$.

Lemma 3.8 *Assume $z \in \mathrm{BV}_-([0,T],E)$ solves (GF) and that I is convex. Then, the map $\big([0,T] \to X^*; t \mapsto \mathrm{D}I(t,z(t))\big)$ is continuous. Moreover, if z jumps at time t from z_- to z_+, then $\mathrm{D}I(t,\cdot)$ is constant along the straight jump line.*

Proof. The result is trivial at points where z is continuous. Hence, we consider the case of a jump. Let $z_\theta = (1-\theta)z_- + \theta z_+$ and $B^*_- = \mathrm{D}I(t,z_-)$. On the one hand convexity implies $I(t,z_\theta) \geq I(t,z_-) + \theta\langle B^*_-, z_+ - z_-\rangle$. On the other hand global stability gives $I(t,z_-) \leq I(t,z_\theta) + \theta\Delta(z_+ - z_-)$. Together with (2.7) we find that the function $I(t,\cdot)$ is affine on the straight line connecting z_- and z_+. Since $I(t,\cdot)$ is convex and differentiable, at all points of the jump line the tangent planes are the same. This proves the result. ■

We now relate the subdifferential formulations $(\mathrm{SF})_\pm$ to the local formulations $(\mathrm{LF})_\pm$. The basic result is obtained from the following lemma which uses simple arguments from convex analysis.

Lemma 3.9 *Let $C \subset X$ be a closed convex cone and $C^* \subset X^*$ the dual cone, see Appendix B. Moreover, assume that $\Delta : X \to [0, \infty)$ is as above. Then, the following three conditions on $w \in X$ and $\beta^* \in X^*$ satisfy the implications (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii).*

(i) $-w \in C$ *and* $0 \in \partial\Delta(w) + \beta^* + C^*$;

(ii) $\langle \beta^*, v\rangle + \Delta(v) \geq 0$ *for all* $v \in C$ *and* $\langle \beta^*, w\rangle + \Delta(w) \leq 0$;

(iii) $0 \in \partial\Delta(w) + \beta^* + C^*$.

This result has a strange unsymmetry which cannot be avoided. Already the simplest case $X = \mathbb{R}$, $C = [0, \infty)$ and $\Delta(v) = |v|$ shows this. The values $\beta^* = 1$ and $w = 1$ satisfy the condition (iii) but not (ii). Moreover, $\beta^* = -1$ and $w = 1$ satisfies (ii) but not (i). Assuming $w \in C$ in addition to (iii) doesn't help to imply (ii).

Proof. Condition (iii) is equivalent to $-\beta^* \in \partial\Delta(w) + C^*$. From the form of the subdifferential $\partial\Delta$ in (3.1) we find the equivalent formulation

$$\text{(iii)'} \quad \Big[\forall\, v \in C \colon \langle \beta^*, v\rangle + \Delta(v) \geq 0\Big] \text{ and } \Big[\exists\, \nu^* \in C^* \colon \langle \beta^*, w\rangle + \Delta(w) + \langle \nu^*, w\rangle = 0\Big].$$

Adding the condition $-w \in C$ we obtain (i), and the implication (i) $\to$ (ii) follows directly from $\langle \nu^*, -w\rangle \leq 0$.

To prove that (ii) implies (iii) we consider two cases. If $\langle \beta^*, w\rangle + \Delta(w) = 0$, then we choose $\nu^* = 0$ and (iii)' holds. If $\langle \beta^*, w\rangle + \Delta(w) < 0$ then the first condition in (ii) tells us that $w \notin C$. Thus there exists a $\gamma^* \in C^*$ with $\langle \gamma^*, w\rangle > 0$. Choosing $r > 0$ suitably the vector $\nu^* = r\gamma^*$ satisfies the second condition in (iii)'. ∎

After the preparatory work, it is easy to prove the Theorems 3.4 & 3.5.

Proof of Theorem 3.4. It is essential that for any solution z of the four formulations we can apply Lemma 3.1 which guarantees $\pm\mathrm{rd}(z)(t) \in \mathrm{T}_{z_\mp(t)}E$ μ_z–a.e..

Part (a) and the direction "$\Rightarrow$" in parts (b) and (c) are proved above.

ad (b) "$\Leftarrow$": Inserting $w = 0$ as a test state in (VI) we obtain ($\mathrm{E}_{\mathrm{loc}}$). Next insert $w = \lambda v + \mathrm{rd}(z)(t)$, divide by λ and take the limit $\lambda \to \infty$. By the lower semicontinuity of $\Delta(\cdot)$ and $\chi_{\mathrm{T}_{z(t)}}(\cdot)$ we arrive at ($\mathrm{S}_{\mathrm{loc}}$).

ad (c) "$\Leftarrow$": In the case of convexity the global stability (S) immediately follows from the local one ($\mathrm{S}_{\mathrm{loc}}$). Integrating ($\mathrm{E}_{\mathrm{loc}}$) and using (3.5) together with Lemma 3.8 gives the global energy condition (E). ∎

Proof of Theorem 3.5. We observe the correspondence between $(\mathrm{SF})_+$, $(\mathrm{LF})_\pm$ and $(\mathrm{SF})_-$ and the conditions (i), (ii) and (iii) in Lemma 3.9, respectively, where $-\mathrm{rd}(z)(t) \in \mathrm{T}_{z_+(t)}E$ is used for (i) but $\mathrm{rd}(z)(t) \in \mathrm{T}_{z_-(t)}E$ is of no help in (iii).

This proves (a) and (b), where the equivalence in (a) follows as $-\mathrm{rd}(z)(t) \in \mathrm{T}_{z_+(t)}E$ is known also in $(\mathrm{LF})_+$. Part (c) is a consequence of (a) and Theorem 3.4(c). ∎

Finally we connect our formulation to the so–called sweeping processes as discussed in [KuM97, KuM98, Mon93]. Again we assume convexity of I. Our subdifferential formulation (SF) is posed in the dual space X^* and we need to employ duality arguments (see Appendix B) involving the Legendre–Fenchel transform $\mathcal{L}$ to return to an equation in the space X. In the case $E=X$, our equation reads $-\mathrm{D}I(t,z(t)) \in \partial\Delta(\mathrm{rd}(z)(t))$, and by Theorem B.1 this is equivalent to

$$\mathrm{rd}(z)(t) \in \partial\chi_{F^*}(-\mathrm{D}I(t,z(t))) \quad \text{for } \mu_z\text{–a.a. } t\in[0,T], \tag{3.7}$$

since χ_{F^*} is the Legendre transform $\mathcal{L}\Delta$ of Δ.

In the case of $E \subsetneqq X$ the situation is a little more involved as we cannot insert a set–valued function into a subdifferential. We start from the variational inequality (VI)$_-$ which can be rewritten as $m_t(w) \ge m_t(\mathrm{rd}(z)(t)) = 0$ for all $w\in X$ where $m_t(w) = \langle \mathrm{D}I(t,z(t)),w\rangle + \Delta(w) + \chi_{\mathrm{T}_{z(t)}E}(w)$. This is equivalent to $0\in\partial m_t(\mathrm{rd}(z)(t))$ or $-\mathrm{D}I(t,z(t)) \in \partial f_t(\mathrm{rd}(z)(t))$ where $f_t(w) = \Delta(w) + \chi_{\mathrm{T}_{z(t)}E}(w)$. By the duality theorem B.1 we have $\mathrm{rd}(z)(t) \in \partial(\mathcal{L}f_t)(-\mathrm{D}I(t,z(t)))$. Using Proposition B.3 we obtain an explicit form for $\mathcal{L}$ which leads to the final result.

(SP)$_-$* *[Sweeping Process Formulation] *Find $z\in \mathrm{BV}_-([0,T],X)$ with $z(0)=z_0$ and $z(t)\in E$ such that*

$$\mathrm{rd}(z)(t) \in \partial\chi_{F^*+\mathrm{N}_{z(t)}E}(-\mathrm{D}I(t,z(t))) \quad \textit{for } \mu_z\text{–a.a. } t\in[0,T]. \tag{3.8}$$

4 Time discretization

One of the standard methods to obtain solutions of nonlinear evolution equations is that of approximation by time discretizations. To this end we choose discrete times $0=t_0 < t_1 < \ldots < t_N = T$ and consider the incremental problem.

(IP) *For given $z_0\in E$ find $z_1,\ldots,z_N\in E$ such that*

$$z_k \in \ \mathit{argmin}\,\{I(t_k,z) + \Delta(z-z_{k-1}) : z\in E\} \tag{4.1}$$

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

By weak lower semi–continuity and boundedness from below of $I(t,\cdot)$ we obtain the following result.

Theorem 4.1 *The incremental problem (4.1) always has a solution. Any solution satisfies for $k=1,\ldots,N$*
(i) z_k is stable for time t_k
(ii) $\int_{[t_{k-1},t_k)} \partial_t I(s,z_k)\,\mathrm{d}s \le I(t_k,z_k) - I(t_{k-1},z_{k-1}) + \Delta(z_k - z_{k-1}) \le \int_{[t_{k-1},t_k)} \partial_t I(s,z_{k-1})\,\mathrm{d}s$
(iii) $\sum_{k=1}^N \Delta(z_k-z_{k-1}) \le I(0,z_0) + C_1T$
(iv) $\|z_k\| \le \|z_0\| + (I(0,z_0)+C_1T)/C_\Delta^{(1)}$.

Remark: The assertions (i) and (ii) are the best replacements of the conditions (S) and (E) in the time–continuous case.

Proof. From $I(t,z)\ge 0$ we have

$$I(t_k,z) + \Delta(z-z_{k-1}) \ge C_\Delta^{(1)}\|z-z_{k-1}\|.$$

Hence any minimizing sequence $(z_k^j)_{j\in\mathbb{N}}$ is bounded and a subsequence converges weakly to $z_k \in E \subset X$. Convexity of Δ and weak lower semi–continuity of $I(t_k,\cdot)$ gives

$$\begin{aligned} I(t_k,z_k) + \Delta(z_k - z_{k-1}) &\le \liminf_{j\to\infty} I(t_k,z_k^j) + \liminf_{j\to\infty} \Delta(z_k^j - z_{k-1}) \\ \le \liminf_{j\to\infty} \Big[I(t_k,z_k^j) + \Delta(z_k^j - z_{k-1}) \Big] &= \inf\{\, I(t_k,y) + \Delta(y - z_{k-1}) \;:\; y\in E\,\}. \end{aligned}$$

This is equivalent to $z_k \in \operatorname{argmin} I(t,\cdot) + \Delta(\cdot - z_{k-1})$.

The stability (i) is obtained by the minimization property and the triangle inequality (2.3) as follows. For all $w\in E$ we have

$$\begin{aligned} I(t_k,w) + \Delta(w - z_k) &= I(t_k,w) + \Delta(w - z_{k-1}) + \Delta(w - z_k) - \Delta(w - z_{k-1}) \\ &\ge I(t_k,z_k) + \Delta(z_k - z_{k-1}) + \Delta(w - z_k) - \Delta(w - z_{k-1}) \ge I(t_k,z_k). \end{aligned}$$

The lower estimate in the energy estimate (ii) is deduced from the stability of z_{k-1} with respect to z_k:

$$\begin{aligned} I(t_k,z_k) + \Delta(z_k - z_{k-1}) &= I(t_{k-1},z_k) + \Delta(z_k - z_{k-1}) + \int_{[t_{k-1},t_k)} \partial_t I(s,z_k)\,\mathrm{d}s \\ &\ge I(t_{k-1},z_{k-1}) + \int_{[t_{k-1},t_k)} \partial_t I(s,z_k)\,\mathrm{d}s. \end{aligned}$$

The upper estimate in (ii) follows since z_k is a minimizer:

$$I(t_k,z_k) + \Delta(z_k - z_{k-1}) \le I(t_k,z_{k-1}) = I(t_{k-1},z_{k-1}) + \int_{[t_{k-1},t_k)} \partial_t I(s,z_{k-1})\,\mathrm{d}s.$$

Adding up (ii) for $k=1,\dots,N$ we find

$$I(T,z_N) - I(0,z_0) + \sum_{k=1}^N \Delta(z_k - z_{k-1}) \le \sum_{k=1}^N \int_{[t_{k-1},t_k)} \partial_t I(s,z_{k-1})\,\mathrm{d}s.$$

Using $I(t,z)\ge 0$ and $|\partial_t I(t,z)| \le C_1$ we obtain (iii). Now (iv) follows from $C_\Delta^{(1)}\|z_k - z_0\| \le \Delta(z_k - z_0) \le \sum_{j=1}^N \Delta(z_j - z_{j-1})$ and (iii). ∎

For each discretization $P=\{0,t_1,\dots,t_{N-1},T\}$ of the interval $[0,T]$ and each incremental solution $(z_k)_{k=1,\dots,N}$ of (IP) we denote by Z_P a piecewise constant function with

$$Z_P(0) = z_0 \quad \text{and} \quad Z_P(t) = z_k \text{ for } t\in(t_{k-1},t_k]. \tag{4.2}$$

Summing (iii) in Theorem 4.1 over $k=1,\dots,N$ we find the following result.

Corollary 4.2 *If $(z_k)_{k=1\dots N}$ solves (IP), then Z_P satisfies the energy inequality*

$$I(T,Z_P(T)) + \int_{[0,T]} \Delta(\mathrm{d}Z_P) \le I(0,z_0) + \int_0^T \partial_t I(s,Z_P(s))\mathrm{d}s.$$

The minimization property of z_k leads to a necessary local condition:

Proposition 4.3 *If z_k solves (4.1) then*

$$\langle \mathrm{D}I(t_k,z_k), w{-}z_k{+}z_{k-1}\rangle + \Delta(w) - \Delta(z_k{-}z_{k-1}) \geq 0 \text{ for all } w \in \mathrm{T}_{z_k}E. \tag{4.3}$$

If $I(t_k,\cdot)$ is strictly convex and $E = X$, then (4.3) has a unique solution. Then (4.1) and (4.3) are equivalent.

Proof. Let $z_\theta = (1{-}\theta)z_k + \theta z_{k-1}$ where $\theta \in [0,1]$. The minimization property of z_k gives

$$\begin{aligned} I(t_k,z_k) + \Delta(z_k{-}z_{k-1}) &\leq I(t_k,z_\theta) + \Delta(z_\theta{-}z_{k-1}) \\ &= I(t_k,z_k) + \theta\langle \mathrm{D}I(t_k,z_k), z_{k-1}{-}z_k\rangle\rangle + o(\theta) + (1{-}\theta)\Delta(z_k{-}z_{k-1}), \end{aligned}$$

for $\theta \to 0$. Subtracting the terms of order θ^0, dividing by θ and taking the limit $\theta \to 0$ yields $\langle \mathrm{D}I(t_k,z_k), z_k{-}z_{k-1}\rangle\rangle + \Delta(z_k{-}z_{k-1}) \leq 0$.

From Theorem 4.1 we know that z_k is stable and comparing with $z = z_k + \theta w$, $w \in \mathrm{T}_{z_k}E$, gives similarly $\langle \mathrm{D}I(t_k,z_k), w\rangle + \Delta(w) \geq 0$. Subtracting the previous inequality gives (4.3).

Let $E = X$ and let $z_k^{(j)}$, $j = 1,2$ be two solutions of (4.3). Then, we can use $w = z_k^{(3-j)} - z_{k-1}^{(3-j)}$ as test–function in (4.3) and the estimates for $j = 1$ and 2 to obtain $\langle \mathrm{D}I(t_k,z_k^{(1)}) {-} \mathrm{D}I(t_k,z_k^{(2)}), z_k^{(2)}{-}z_{k-1}^{(2)} - (z_k^{(1)}{-}z_{k-1}^{(1)})\rangle \geq 0$. Induction over k together with strict convexity implies that $z_k^{(1)} = z_k^{(2)}$. ■

For the uniformly convex case we obtain a Lipschitz bound for the incremental problem (4.1).

Theorem 4.4 *If $I(t,\cdot)$ is α–uniformly convex then any solution of (4.1) satisfies*

$$\|z_k{-}z_{k-1}\| \leq \frac{C_2}{\alpha}|t_k{-}t_{k-1}| \text{ for } k = 1,\ldots,N.$$

Proof. The stability of z_{k-1} at t_{k-1} implies via ($\mathrm{S}_{\mathrm{loc}}$) and $z_k{-}z_{k-1} \in \mathrm{T}_{z_{k-1}}E$ the estimate

$$\langle \mathrm{D}I(t_{k-1},z_{k-1}), z_k - z_{k-1}\rangle + \Delta(z_k{-}z_{k-1}) \geq 0. \tag{4.4}$$

Adding this to (4.3) with $w = 0$ we have $\langle \mathrm{D}I(t_{k-1},z_{k-1}) {-} \mathrm{D}I(t_k,z_k), z_k - z_{k-1}\rangle \geq 0$. With the uniform convexity and assumption (2.6) we continue

$$\begin{aligned} 0 \;&\geq\; \langle \mathrm{D}I(t_k,z_k){-}\mathrm{D}I(t_k,z_{k-1}), z_k{-}z_{k-1}\rangle + \langle \mathrm{D}I(t_k,z_{k-1}){-}\mathrm{D}I(t_{k-1},z_{k-1}), z_k{-}z_{k-1}\rangle \\ &\geq\; \alpha\|z_k{-}z_{k-1}\|^2 - C_2|t_k{-}t_{k-1}|\|z_k{-}z_{k-1}\|; \end{aligned}$$

and the result is established. ■

5 Stable sets

The sets of stable points play an important role in the analysis. For $t \in [0,T]$ we let

$$\mathcal{S}(t) = \{z \in E : I(t,z) \leq I(t,y) + \Delta(y{-}z) \text{ for all } y \in E\},$$

which is the set of all stable points at time t. The condition (S) now reads "$z(t) \in \mathcal{S}(t)$".

Lemma 5.1 *Let $I(\cdot,z):[0,T]\to[0,\infty)$ be continuous and $I(t,\cdot):E\to[0,\infty)$ be lower semicontinuous. Then,*
(a) for each t the set $\mathcal{S}(t)$ is closed;
(b) if $t_n\to t$ and $z_n\in\mathcal{S}(t_n)$ with $z_n\to z$, then $z\in\mathcal{S}(t)$.

Proof. Let $H(t,z,w)=I(t,w)+\Delta(w-z)-I(t,z)$ and $\mathcal{H}(t,z)=\inf\{H(t,z,w):w\in E\}$. Clearly, $\mathcal{H}(t,z)\le 0$ and $z\in\mathcal{S}(t)\Leftrightarrow\mathcal{H}(t,z)=0$.

Assume $z_n\in\mathcal{S}(t)$ and $z_n\to z$, then $H(t,z_n,w)\ge 0$ for all w. By continuity of $I(t,\cdot)$ and $\Delta(\cdot)$ we find $H(t,z,w)\ge 0$ and conclude $z\in\mathcal{S}(t)$. This proves (a).

Part (b) follows from (a) and the (strong) continuity of $H(\cdot,z,w):[0,T]\to\mathbb{R}$. ∎

In the case $E=X$, $I(t,\cdot)$ convex and in $\mathrm{C}^1(X,\mathbb{R})$ the stable set is simply characterized by $\mathcal{S}(t)=\{z:-\mathrm{D}I(t,z)\in F^*\}$. In the general case we have

$$z\in\mathcal{S}(t)\Rightarrow-\mathrm{D}I(t,z)\in F^*+\mathrm{N}_zE \tag{5.1}$$

with equivalence if $I(t,\cdot)$ is convex.

We will see in the following that weak closedness of $\mathcal{S}(t)$ is a very desirable property. A natural way to obtain weak closedness of the stable sets is to show convexity by using that F^* is sufficiently round. We say F^* is γ–round if

$$\begin{aligned}&\forall\theta\in[0,1]\forall z_0^*,z_1^*\in F^*:\\&\{w^*\in X^*:\|w-(\theta z_1^*+(1-\theta)z_0^*)\|\le\gamma\theta(1-\theta)\|z_0^*-z_1^*\|^2\}\subset F^*.\end{aligned}\tag{5.2}$$

In a Hilbert space the ball $B_R(z^*)$ is $(2R)^{-1}$–round.

Theorem 5.2 *Assume $E=X$ and that F^* is γ–round. Moreover, assume that $I(t,\cdot)\in\mathrm{C}^3(X,\mathbb{R})$ with $\|\mathrm{D}^3I(t,z)\|_{X\times X\times X\to\mathbb{R}}\le M$ for all z and that I is α–uniformly convex. Then the inequality $M/(2\alpha^2)\le\gamma$ implies that $\mathcal{S}(t)$ is convex.*

Proof. Take $z_0,z_1\in\mathcal{S}(t)$. By (5.1) we have $\sigma_j^*=\mathrm{D}I(t,z_j)\in F^*$ for $j=0$ and 1. For $\theta\in[0,1]$ we let $z_\theta=(1-\theta)z_0+\theta z_1$, then it suffices to show

$$\|\mathrm{D}I(t,z_\theta)-(1-\theta)\mathrm{D}I(t,z_0)-\theta\mathrm{D}I(t,z_1)\|\le\gamma\theta(1-\theta)\|\sigma_0^*-\sigma_1^*\|^2.$$

The left–hand side takes the form

$$\left\|\theta(1-\theta)\int_{[0,1]}\left[\mathrm{D}^2I(t,z_1+s(z_\theta-z_1))-\mathrm{D}^2I(t,z_0+s(z_\theta-z_0))\right](z_0-z_1)\mathrm{d}s\right\|$$

and thus can be estimated by $\frac{M}{2}\theta(1-\theta)\|z_0-z_1\|^2$. Uniform convexity gives

$$\|\sigma_0^*-\sigma_1^*\|\ge\frac{\langle\mathrm{D}I(t,z_0)-\mathrm{D}I(t,z_1),z_0-z_1\rangle}{\|z_0-z_1\|}\ge\alpha\|z_0-z_1\|$$

and the result is established. ∎

An important special case is $E=X$ and I being a convex quadratic functional.

Corollary 5.3 *Assume $E=X$ and $I(t,z)=\frac{1}{2}\langle A(t)z,z\rangle-\langle g(t),z\rangle$ with $A(t)\in\mathrm{Lin}(X,X^*)$ and $\langle A(t)z,z\rangle>0$ for all $z\in X$. Then $\mathcal{S}(t)$ is convex.*

For convenience, we give a direct proof. However, this result is a special case of Theorem 5.2 and of Theorem 5.4.

Proof. Since $I(t,\cdot)$ is convex and smooth we have $z \in \mathcal{S}(t) \Leftrightarrow -\mathrm{D}I(t,z) = -A(t)z+g(t) \in F^*$. Since F^* is convex and $\mathrm{D}I$ is linear we conclude convexity of $\mathcal{S}(t)$. ■

Theorem 5.4 *If one of the following conditions holds, then the stable sets $\mathcal{S}(t)$ are weakly closed :*
(1) $\mathcal{S}(t)$ is convex.
(2) For all $w \in E$ the mapping $\Delta(w-\cdot): E \to \mathbb{R}$ is weakly continuous.
(3) $E = X$ and the mapping $\mathrm{D}I(t,\cdot): E \to X^*$ is weakly continuous.

Proof. Part (1) is clear since closed convex sets are weakly closed.

Part (2) follows by using weak continuity of $\Delta(\cdot)$ and weak lower semi–continuity of $I(t,\cdot)$ as follows. Assume $z_n \rightharpoonup z$ and $z_n \in \mathcal{S}(t)$. With the notation as in the proof of Lemma 5.1 we have

$$\begin{aligned} 0 \le \limsup_{n\to\infty} H(t,w,z_n) &= I(t,w) + \lim_{n\to\infty} \Delta(w-z_n) - \liminf_{n\to\infty} I(t,z_n) \\ &\le I(t,w) + \Delta(w-z) - I(t,z) = H(t,w,z) \end{aligned}$$

for all $w \in E$. Thus, $z \in \mathcal{S}(t)$.

Part (3) follows with (5.1) where $\mathrm{N}_z E = \{0\}$. ■

Note that the condition (2) together with $0 < C_\Delta^{(1)} \le C_\Delta^{(2)} < \infty$ in (2.2) is rather restrictive; it means that $E \cap B_r(0)$ is compact for each $r > 0$.

In the following examples we give simple functionals $I(t,z)$ and $\Delta(z)$ such that the stable set $\mathcal{S}(t)$ is nonconvex and not weakly closed. See also Example C.3.

Example 5.5 *Let $E = X = \mathbb{R} \times H$ where H is a Hilbert space. Let $z = (a,h) \in X$ and*

$$I(t,z) = \frac{1}{4}(a^2+\|h\|^2)^2 - \gamma(t)a, \quad \Delta(z) = \sqrt{a^2+\|h\|^2}.$$

Then $z \in \mathcal{S}(t) \Leftrightarrow \|\mathrm{D}I(t,z)\| \le 1$, where $\mathrm{D}I(t,(a,h)) = (a^2+\|h\|^2)\binom{a}{h} - \binom{\gamma(t)}{0}$. Now, assume $\gamma(t_1) = 2$, then $(a,0) \in \mathcal{S}(t_1)$ if and only if $a \in [1, 3^{1/3}]$. Consider $z_* = (a_*, h_*)$ with $a_* = (3^5/2^8)^{1/3} < 1$ and $\|h_*\| = (3\cdot 5^3/2^{16})^{1/6}$, then a straight forward calculation gives $\|\mathrm{D}I(t_1,z_*)\| = 1$. In fact, these z_* are the ones having the smallest a–component.

Clearly, $\mathcal{S}(t_1)$ cannot be convex, since (a_*,h_*) and $(a_*,-h_*)$ are in $\mathcal{S}(t_1)$ but $(a_*,0) \notin \mathcal{S}(t_1)$, see Fig. 1 for a visualization. Moreover, if H is infinite dimensional then $\mathcal{S}(t_1)$ is not weakly closed. In fact, take any sequence h_k with $\|h_k\| = \|h_*\|$ and $h_k \rightharpoonup 0$, then $(a_*,h_k) \in \mathcal{S}(t_1)$ and $(a_*,h_k) \rightharpoonup (a_*,0) \notin \mathcal{S}(t_1)$.

Example 5.6 *In this example I is quadratic but $E \subsetneqq X$. Consider $E = B_R(0) \subset X$ Hilbert space,*

$$I(t,z) = \frac{\alpha}{2}\|z\|^2 - \langle g(t), z\rangle, \quad \Delta(z) = \|z\|.$$

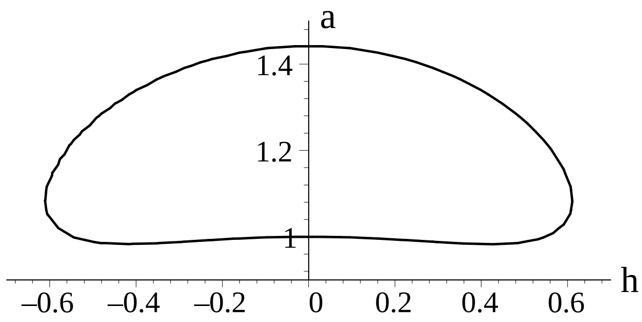

Figure 1: The visualization of the stable set in Example 5.5 in the case $H = \mathbb{R}$ clearly shows that convexity can not be expected.

Then z with $\|z\| < R$ is stable if and only if $\|\alpha z - g(t)\| \leq 1$. For z with $\|z\| = R$ the boundary of E can stabilize; and stability holds if there exists $\gamma \in [\alpha, \infty)$ such that $\|\gamma z - g(t)\| \leq 1$.

Thus, in the case $\|g(t)\| \leq \sqrt{1+\alpha^2 R^2}$ we have the convex stable set $\mathcal{S}(t) = \{z \in E : \|\alpha z - g(t)\| \leq 1\} \cap B_R(0)$, which is the intersection of two balls. In the case $\|g(t)\| > \sqrt{1+\alpha^2 R^2}$ we have

$$\mathcal{S}(t) = \{z \in E : \|\alpha z - g(t)\| \leq 1\} \cup \{z \in E : \|z\| = R, \| \left(\|g(t)\|^2 - 1\right)^{1/2} z - R g(t) \| \leq R\}$$

which always contains a nonconvex part of the boundary of the sphere.

However, there are also many nontrivial examples where convexity of the stable set can be shown directly. For instance, consider $X = L^1(\Omega; \mathbb{R}^n)$, $\Delta(z) = \|z\|_1 = \int_\Omega |z(x)| \, \mathrm{d}x$ and $E = L^1(\Omega; K)$ where K is a compact and convex subset of $\mathbb{R}^n$. Let $W : [0,T] \times \Omega \times K \to \mathbb{R}$ be continuous and $W(t, x, \cdot) : K \to \mathbb{R}$ be convex. For $I(t, z) = \int_\Omega W(t, x, z(x)) \, \mathrm{d}x$ the stability of $z \in E$ can be checked pointwise: z is stable at time t if and only if for a.a. $x \in \Omega$ the vector $z(x)$ is stable for $\widehat{I}(t, \xi) = W(t, x, \xi)$ and $\widehat{\Delta}(\xi) = |\xi|$. We refer to [MTL00] for a nontrivial application of this idea.

6 Existence and uniqueness results for general I

The existence theory can be approached in a rather general setting even without convexity. We use the incremental method of Section 4. For a discretization $P = \{0, t_1, \ldots, t_{N-1}, T\}$ of $[0,T]$ the fineness is $\delta(P) = \max\{\, t_k - t_{k-1} \; : \; k=1, \ldots, N \,\}$ and Z_P denotes the left–continuous, piecewise constant solution associates to a solution of (IP), see (4.2). By our assumptions all these functions Z_P are bounded in $\mathrm{BV}([0,T], X)$, since $Z_P(0) = z_0$ and $\sum_{k=1}^N \Delta(z_k - z_{k-1}) \leq I(0, z_0) + C_1 T$, see Lemma 4.1 (iii). To extract a convergent subsequence we use the following generalization of Helly's selection principle, see [BaP86]. We sketch the proof for the convenience of the reader.

Theorem 6.1 *Let $(z_n)_{n \in \mathbb{N}}$ be a bounded sequence in $\mathrm{BV}([0,T], X)$ with $z_n(t) \in E$ for all $n \in \mathbb{N}$ and $t \in [0,T]$. Then there exists a subsequence (n_k) and functions $\delta^\infty \in \mathrm{BV}([0,T], \mathbb{R})$ and $z^\infty \in \mathrm{BV}([0,T], X)$ such that the following holds:*

(a) $\int_{[0,t)} \Delta(\mathrm{d}z_{n_k}) \to \delta^\infty(t)$ *for all* $t \in [0,T]$*;*
(b) $z_{n_k}(t) \rightharpoonup z^\infty(t) \in E$ *for all* $t \in [0,T]$*;*
(c) $\int_{[s,t)} \Delta(\mathrm{d}z^\infty) \le \delta^\infty(t) - \delta^\infty(s)$ *for all* $0 \le s < t \le T$*.*

Proof. First consider the monotone functions $\delta_n(t) = \int_{[0,t)} \Delta(\mathrm{d}z_n)$. By the real–valued version of Helly's principle there exists a subsequence (n_l) and a monotone function $\delta^\infty : [0,T] \to [0,\infty)$ such that for all $t \in [0,T]$ we have $\delta_{n_l}(t) \to \delta^\infty(t)$ for $l \to \infty$.

Clearly, δ^∞ has an at most a countable jump set $J \subset [0,T]$. Now choose a sequence $(t_j)_{j\in\mathbb{N}}$ which is dense in $[0,T]$ and contains J. Since there is an $R > 0$ with $\|z_n(t)\| \le R$ we have weak compactness and can construct a further subsequence (n_k) such that for all $j \in \mathbb{N}$ we have $z_{n_k}(t_j) \rightharpoonup z^\infty(t_j)$ for $k \to \infty$. This is the definition of z^∞ at $t = t_j$.

Using the weak lower semi–continuity of Δ it follows that z^∞ must be continuous except at the jump points of $\delta^\infty(t)$. Thus, we define z^∞ on all of $[0,T]$ by its continuous extension in $[0,T]/J$.

Moreover, by b) we have weak convergence for all $t \in [0,T]$ as follows. The case $t \in J$ is clear, hence assume $t \in [0,T]/J$ and choose $y^* \in X^*$. Then,

$$\begin{aligned}|\langle z_{n_k}(t) - z^\infty(t), y^*\rangle| \le &\|z_{n_k}(t) - z_{n_k}(t_j)\|\|y^*\| \\ &+ |\langle z_{n_k}(t_j) - z^\infty(t_j), y^*\rangle| + \|z^\infty(t_j) - z^\infty(t)\|\|y^*\|.\end{aligned}$$

Now we can choose t_j such that the first and third term are less that $\varepsilon/3$ for all $k \ge k_0$, since both terms can be estimated by $|\delta^\infty(t) - \delta^\infty(t_j)|2\|y^*\|$. Keeping j fixed and increasing k_0 if necessary the second term is less than $\varepsilon/3$ as well as for $k \ge k_0$. ∎

In the above and in the following result the function z is a general BV–function where at jump points $z_+(t)$, $z(t)$ and $z_-(t)$ may all be different. Combining the a– priori estimates in Section 4 and the previous theorem we arrive at the following result.

Proposition 6.2 *Let* $P^{(j)} = \{0, t_1^{(j)}, \ldots, t_{N(j)-1}^{(j)}, T\}$ *be a sequence of discretizations whose fineness* $\delta(P^{(j)})$ *tends to* 0*. Denote by* $z^{(j)} = Z_{P^{(j)}}$ *the associated step functions (4.2) of the incremental problems (IP). Then there exists a subsequence* $(j_l)_{l\in\mathbb{N}}$ *and functions* $\delta^\infty, i^\infty : [0,T] \to \mathbb{R}$ *and* $z^\infty \in \mathrm{BV}([0,T],X)$ *such that*
(i) $z^{(j_l)}(t) \rightharpoonup z^\infty(t)$*,* $\int_{[0,t)} \Delta(\mathrm{d}z^{(j_l)}) \to \delta^\infty(t)$ *for all* $t \in [0,T]$*;*
(ii) $\int_{[s,t)} \Delta(\mathrm{d}z^\infty) \le \delta^\infty(t) - \delta^\infty(s)$ *for all* $s < t$*,*
(iii) $i^\infty(t) = \lim_{l\to\infty} I(t, z^{(j_l)}(t)) \ge I(t, z^\infty(t))$*;*
(iv) $i^\infty(t) + \delta^\infty(t) = I(0, z_0) + \int_{[0,t)} \partial_t I(s, z^\infty(s))\,\mathrm{d}s$*.*

Proof. Applying Theorem 6.1 to the sequence $z^{(j)}$ we immediately obtain (i) and (ii). The weak continuity of $\partial_t I(s,\cdot)$ (see assumption (2.6)) implies

$$\int_{[0,t)} \partial_t I(s, z^{(j_l)}(s))\,\mathrm{d}s \to \int_{[0,t)} \partial_t I(s, z^\infty(s))\,\mathrm{d}s$$

for all $t \in [0,T]$. Adding the two–sided energy estimate (ii) in Lemma 4.1 we have

$$\begin{aligned}\int_{[0,t_n^{(j)})} \partial_t I(s, z^{(j)}(s))\,\mathrm{d}s \;&\le\; I(t_n^{(j)}, z^{(j)}(t_n^{(j)})) - I(0, z_0) + \int_{[0,t_n^{(j)})} \Delta(\mathrm{d}z^{(j)}) \\ &\le\; \int_{[0,t_n^{(j)})} \partial_t I(s, \widehat{z}^{(j)}(s))\,\mathrm{d}s,\end{aligned} \tag{6.1}$$

where $\widehat{z}^{(j)}(t) = \sum_{k=0}^{N(j)-1} i_{[k_k,t_{k+1})}(t) z_k^{(j)}$ is the right–continuous step function. At all continuity points t of z^∞ we have $\widehat{z}^{(j_l)}(t) \rightharpoonup z^\infty(t)$ and hence the upper and lower estimates in (6.1) both converge, for $j_l \to \infty$ and $t_{n_l}^{(j_l)} \to t$, to $\int_{[0,t)} \partial_t I(s, z^\infty(s))\,\mathrm{d}s$.

With (i) and (6.1) we conclude that the limit $i^\infty(t)$ exists for all $t \in [0,T]$ and (iv) holds. Lower semi–continuity of $I(t,\cdot)$ implies (iii). ■

The remaining task is to show that the limit function gives rise to a solution of our problem. The functions $z = (z^\infty)_\pm$ always satisfy the global energy inequality (E), which follows from

$$I(t, z(t)) + \int_{[0,t)} \Delta(\mathrm{d}z) \le i^\infty(t) + \delta^\infty(t) = I(0, z_0) + \int_{[0,t)} \partial_t I(s, z(s))\,\mathrm{d}s.$$

The major problem is to obtain stability of $z(t)$ for all $t \in [0,T]$. To derive this from the stability of $z^{(j)} = Z_{P^{(j)}}$ at the times $t \in P^{(j)}$ we have two choices. Either we improve the weak convergence in Proposition 6.2(i) to strong convergence and use the closedness of the stable sets, see Lemma 5.1. Or we stay with the weak convergence and show that $\mathcal{S}(t)$ is weakly closed.

Theorem 6.3 *Assume that E, Δ and I satisfy the above assumption. Take a hierarchical sequence of discretizations $P^{(1)} \subset P^{(2)} \subset \ldots$ with $\delta(P^{(j)}) \to 0$ and define $\widehat{z}^{(l)} = Z_{P^{(j_l)}}$ and z^∞ as in Proposition 6.2. Then $z = (z^\infty)_\pm$ is a solution of (GF) if there exists a dense subset $\mathcal{T}$ of $[0,T]$ with $\mathcal{T} \subset \bigcup_{j=1}^\infty P^{(j)}$ such that one of the following two conditions holds*
(1) $\widehat{z}^{(l)}(t) \to z^\infty(t)$ for all $t \in \mathcal{T}$;
(2) $\mathcal{S}(t)$ is weakly closed for all $t \in \mathcal{T}$.

Proof. The two conditions are such that we can conclude $z(t) \in \mathcal{S}(t)$ for all $t \in \mathcal{T}$ by using the stability of $z^{(j_l)}(t)$ at time t for all sufficiently large l and the closedness of $\mathcal{S}(t)$ in the suitable topology.

Lemma 5.1(b) and the density of $\mathcal{T}$ then imply that z is stable for all $t \in [0,T]$ and the theorem is established. ■

Remark 6.4 *The above theorem was derived under the assumption of reflexivity of X. However, we only used that closed, bounded subsets of E are sequentially weakly compact. Interesting applications involve the choice $X = \mathrm{L}^1(\Omega)$ and $E = \{\, z \,:\, \|z\|_\infty \le 1 \,\}$, cf. [MTL00].*

Actually in case (1) the assumption of the weak continuity of $\partial_t I(t,\cdot)$ is unnecessary, strong continuity would suffice. A simple example to which the second part of our theorem applies is the following. Let $E = X$ where X is an infinite dimensional Hilbert space. Let $\Delta(z) = \|z\|$ and

$$I(t,z) = \frac{\alpha}{2}\|z\|^2 + \|z\|^4 - \langle g(t), z\rangle$$

for a small $\alpha > 0$ and a suitable smooth function $g : [0,T] \to X$. By Example 5.5 we know that the stable sets $\mathcal{S}(t)$ are not weakly closed for an open set of funtions g.

Following the spirit of the existence theorem we can also deduce a uniqueness result which requires that the stable sets are convex.

Theorem 6.5 *If in addition to the above assumptions the function I has the form $I(t,z) = J(z) - \langle g^*(t), z\rangle$ where J is strictly convex and the stable sets $\mathcal{S}(t)$ are convex for all $t \in [0,T]$, then there is a unique solution for each initial condition $z_0 \in E$.*

Proof. The existence of a solution is a consequence of Theorem 6.3, since convexity of $\mathcal{S}(t)$ implies weak closedness via the strong closedness, cf. Lemma 5.1.

For any two solutions $z_j : [0,T] \to E$ with $z_j(0) = z_0 \in \mathcal{S}(0)$ define $\widetilde{z}(t) = \frac{1}{2}(z_0(t)+z_1(t))$. By convexity of the stable sets we know that $\widetilde{z}(t)$ is stable for all t. Now assume $z_0(t) \neq z_1(t)$ for some $t > 0$. Using strict convexity of J, the energy identity (3.4) and the linearity of $\partial_t I$ we obtain

$$\begin{aligned} I(t,\widetilde{z}(t)) + \int_{[0,t]} \Delta(\mathrm{d}\widetilde{z}) &< \tfrac{1}{2}[I(t,z_0(t)) + I(t,z_1(t))] + \int_{[0,t]} \tfrac{1}{2}[\Delta(\mathrm{d}z_0)+\Delta(\mathrm{d}z_2)] \\ &= \tfrac{1}{2}[I(t,z_0(0)) + I(0,z_1(0))] - \int_0^t \tfrac{1}{2}[\langle \dot{g}^*, z_0\rangle + \langle \dot{g}^*, z_1\rangle]\mathrm{d}s \\ &= I(0,z_0) + \int_0^t \partial_t I(s,\widetilde{z}(s))\mathrm{d}s. \end{aligned}$$

Thus, $\widetilde{z}$ also satisfies the global energy condition (E) and is a solution as well. However, by Lemma 3.7 every solution satisfies the energy equality (3.4) which is not the case here due to the strict convexity. Hence, we conclude $z_0 = z_1$ on $[0,T]$. ■

In [CoV90] uniqueness for our situation is obtained only if $E = X$ and I is a quadratic functional, i.e., $I(t,z) = \langle Az - g^*(t), z\rangle$ where $A : X \to X^*$ is symmetric and positive definite. Clearly, this is a special case of the above result.

7 The good case: I uniformly convex and $X = E$

In this section we establish the well posedness of the time–continuous evolution problem in the convex case. By Theorem 3.4.c and Theorem 3.5.c we see that all our formulations are equivalent. It is obvious that one can only expect that the solutions are unique if the I is strictly convex. In the degenerate cases it is easy to construct examples for nonuniqueness. If we additionally assume that I satisfies a smoothness condition we get a pretty complete picture. For every initial value $z(0) = z_0 \in X$ there exists a unique continuous process z and $z(t)$ depends continuously on z_0. Furthermore the solutions to the incremental problem converge to z as the fineness of the discretization tends to 0.

This improves the result in [HaR95] where the existence of a solution $z \in W^{2,1}([0,T],X)$ (resp. with $\dot{z} \in \mathrm{BV}([0,T],X)$, cf. Section 7.2) has to be assumed before showing convergence of the incremental method. At the end of the section we state the optimal regularity in time of the solutions. This requires that the boundary of F^* is smooth.

Unfortunately our smoothness assumption on $I(t,\cdot)$ rules out the case $E \neq X$ and we are unable to generalize the methods of this chapter to cases where E has a boundary.

Theorem 7.1 (Well posedness in the convex, smooth case) *Assume $E = X$ and $I(t,\cdot) \in \mathrm{C}^3(X)$ α–uniformly convex in z. Then for every z_0 in X there exists a unique solution $z \in W^{1,\infty}([0,T],X)$ of (GF) (or alternatively of (LF), (VI) or (SF)). For each $t \in [0,T]$ the state $z(t)$ depends Lipschitz continuously on the initial value z_0. Furthermore there exists a constant $C > 0$ so that*

$$\|z - Z_P\|_{L^\infty([0,T],X)} \leq C\sqrt{\delta}$$

with Z_P from (4.2) and $\delta = \min\{ t_i - t_{i-1} \ : \ i=1,\dots,N \}$ is the fineness of the time discretization.

Proof. The equivalence of (GF), (LF), (VI) and (SF) is established in the Theorems 3.4 and 3.5. The proof of existence of the solutions to the Cauchy problem is given in Theorem 7.3. There also the convergence of the solutions of the incremental problems to the time continuous solution is established. The uniqueness and the continuous dependence on the initial value is established in Theorem 7.4 ■

7.1 Strong convergence

We first prove a stability result for the incremental problem when I is perturbed. This will then be used to compare the incremental solutions for two different discretizations.

Proposition 7.2 *Let $E = X$, $j \in \{1,2\}$, $I^j \in \mathrm{C}^3(X,\mathbb{R})$, I^j α–uniformly convex. Then, there exists a constant $C > 0$ such that for all discretizations $P = \{0, t_1, \dots, t_{N-1}, T\}$ the solutions $(z_k^j)_{k=0,\dots N}$, $j = 1,2$, of the incremental problem satisfy*

$$\|z_k^1 - z_k^2\| \le C\rho^{1/2}$$

where $\rho = \sup_{z\in X} \|\mathrm{D}I^1(\cdot,z) - \mathrm{D}I^2(\cdot,z)\|_{L^\infty([0,T],X^*)}$.

Proof. We generalize the idea of the proof of Theorem 2.3 in [HaR95]. We introduce the notation $\sigma^j(t,z) = \mathrm{D}I^j(t,z), e_k = z_k^1 - z_k^2$ and the difference operator $\tau_k\zeta = \zeta_k - \zeta_{k-1}$ where ζ stands for $t, z^j, \sigma^j(t, z_k^l)$ or e.

Convexity, $E = X$ and (4.3) give $\langle \sigma^j(t_k, z_k^j), w - \tau_k z^j\rangle + \Delta(w) - \Delta(\tau_k z^j) \ge 0$ for all $w \in X$. Inserting $w = \tau_k z^{3-j}$ and adding the equations for $j = 1$ and 2 gives

$$\langle \sigma^1(t_k, z_k^1) - \sigma^2(t_k, z_k^2), \tau_k e\rangle \le 0. \tag{7.1}$$

The final estimate is derived using the quantity

$$\gamma_k \stackrel{\mathrm{def}}{=} \langle \sigma^1(t_k, z_k^1) - \sigma^1(t_k, z_k^2), e_k\rangle = \langle \mathrm{D}I^1(t_k, z_k^1) - \mathrm{D}I^1(t_k, z_k^2), z_k^1 - z_k^2\rangle$$

which by uniform convexity controls the error e_k via $\alpha\|e_k\|^2 \le \gamma_k$. The increment $\tau_k\gamma = \gamma_k - \gamma_{k-1}$ can be estimated via (7.1) as follows

$$\begin{aligned} \tau_k\gamma &= \langle \sigma^1(t_k, z_k^1) - \sigma^1(t_k, z_k^2), \tau_k e\rangle + \langle \tau_k(\sigma^1(t_k, z_k^1) - \sigma^1(t_k, z_k^2)), e_{k-1}\rangle \\ &= 2\langle \sigma^1(t_k, z_k^1) - \sigma^2(t_k, z_k^2), \tau_k e\rangle + \beta_k \end{aligned}$$

where $\beta_k = \langle \tau_k(\sigma^1(t_k, z_k^1) - \sigma^1(t_k, z_k^2)), e_{k-1}\rangle - \langle \sigma^1(t_k, z_k^1) - \sigma^1(t_k, z_k^2), \tau_k e\rangle + 2\langle \sigma^1(t_k, z_k^2) - \sigma^2(t_k, z_k^2), \tau_k e\rangle$ takes the form

$$\begin{aligned} \beta_k &= \langle A_k e_k - A_{k-1}e_{k-1}, e_{k-1}\rangle - \langle A_k e_k, \tau_k e\rangle + 2\langle \sigma^1(t_k, z_k^2) - \sigma^2(t_k, z_k^2), \tau_k e\rangle \\ &= -\langle A_k \tau_k e, \tau_k e\rangle + \langle (A_k - A_{k-1})e_{k-1}, e_{k-1}\rangle + 2\langle \sigma^1(t_k, z_k^2) - \sigma^2(t_k, z_k^2), \tau_k e\rangle. \end{aligned}$$

The symmetric operators $A_k \in \mathrm{Lin}(X, X^*)$ are defined via $A_k = \int_{[0,1]} \mathrm{D}^2 I^1(t_k, z_k^2 + \theta e_k)\,\mathrm{d}\theta$ and satisfy $A_k e_k = \sigma^1(t_k, z_k^1) - \sigma^1(t_k, z_k^2)$. By convexity and three–times differentiability we obtain

$$\langle A_k y, y\rangle \ge 0 \text{ and } \|A_k - A_{k-1}\| \le C_3\left(\|\tau_k z^1\| + \|\tau_k z^2\|\right).$$

Together with $\|\tau_k e\| \leq \|\tau_k z^1\| + \|\tau_k z^2\|$ we find

$$\tau_k \gamma \leq \left[2G_k + \frac{C_3}{\alpha}\gamma_{k-1} \right] \left(\|\tau_k z^1\| + \|\tau_k z^2\| \right),$$

where $G_k = \sup_{z\in X} \|\sigma^1(t_k, z) - \sigma^2(t_k, z)\|_{X^*}$. The Lipschitz continuity of z^j (see Theorem 4.4) gives the existence of a constant C which is independent of the discretization P such that

$$\gamma_k \leq \gamma_{k-1} + (t_k - t_{k-1})\widehat{C}[G_k + \gamma_{k-1}] \text{ for } k = 1, \ldots, N,$$

where $\widehat{C} = C \sup_{z\in X}[\|\partial_t \sigma^1(\cdot, z)\|_\infty + \|\partial_t \sigma^2(\cdot, z)\|_\infty]$. With $\gamma_0 = 0$ and $G_k \leq \rho$ we obtain

$$\gamma_k \leq \widehat{C}\rho \sum_{n=1}^{k} (t_n - t_{n-1}) \prod_{j=n+1}^{k} [1 + \widehat{C}(t_j - t_{j-1})] \leq \widehat{C}\rho \mathrm{e}^{\widehat{C}T} T.$$

Together with $\|z_k^1 - z_k^2\|^2 \leq \frac{1}{\alpha}\gamma_k$ this is the desired result. ■

Theorem 7.3 *Under the assumptions of Proposition 7.2 there is a unique solution $z \in W^{1,\infty}([0,T], X)$ and for each discretization P the incremental solution $z_k = Z_P(t_k)$ satisfies the error estimate*

$$\|z_k - z(t_k)\| \leq C\rho(P)^{1/2} \text{ for } t_k \in P$$

where C is independent of P.

Proof. Uniqueness and Lipschitz continuity are shown in Theorems 7.4 and 7.5 below. The existence part is based on the global definition of solutions via stability (S) and energy inequality (E) and strong convergence. We use the discretizations $P^{(j)} = \{kT2^{-j} : k = 0, \ldots, 2^j\}$ with the associated incremental solutions $z^{(j)} = Z_{P^{(j)}}$.

The idea is to consider $z^{(j-1)}$ as an incremental solution on $P^{(j)}$ of a slightly modified problem. Then $z^{(j-1)}$ and $z^{(j)}$ can be compared using the previous proposition. We let $I^{(j)}(kT2^{-j}, \cdot) = I(kT2^{-j}, \cdot)$ for even k and $I^{(j)}(kT2^{-j}, \cdot) = I((k+1)T2^{-j}, \cdot)$ for odd k. Between the points in $P^{(j)}$ the potential $I^{(j)}$ is assumed to be linear in t. Thus, we have

$$\|\mathrm{D}I^{(j)} - \mathrm{D}I\|_\infty \leq \frac{T}{2^{j-1}} \|\partial_t \mathrm{D}I\|_\infty \text{ and } \|\partial_t \mathrm{D}I^{(j)}\|_\infty \leq 2\|\partial_t \mathrm{D}I\|_\infty,$$

where $\|\mathrm{D}I\|_\infty = \sup\{\|\mathrm{D}I(t, z)\|_{X^*} : t \in [0, T], z \in X\}$.

Since $z^{(j-1)}$ is the incremental solution on $P^{(j)}$ with we can apply Proposition 7.2 and obtain

$$\|z^{(j)}(t) - z^{(j-1)}(t)\| \leq C2^{-j/2} \text{ for } t \in P^{(j)}. \tag{7.2}$$

Thus, keeping $t \in \mathcal{T}^n = \{kT2^{-n} : k \in \mathbb{N}, k \leq 2^n\}$ fixed, the sequence $(z^{(j)}(t))_{j\in\mathbb{N}}$ is a Cauchy sequence. Its limit $z^\infty(t)$ provides a Lipschitz continuous solution by Theorem 6.3 (the incremental solutions satisfy a uniform Lipschitz bound, cf. Theorem 4.4).

The estimate for general discretizations $P = P^{(1)}$ is obtained by successive bisection of the previous discretization $P^{(j)}$ such that $\delta(P^{(j+1)}) = \delta(P^{(j)})/2$. Doing the same trick as above we again obtain $\|\mathrm{D}I^{(j)} - \mathrm{D}I\|_\infty \leq 2\delta(P^{(j)})$ and $\|\partial_t \mathrm{D}I^{(j)}\|_\infty \leq 2\|\partial_t \mathrm{D}I\|_\infty$. Estimate (7.2) is replaced by

$$\|z^{(j-1)}(t) - z^{(j)}(t)\| \leq C\delta(P^{(j)})^{1/2} \text{ for } t \in P^{(j)}.$$

Restricting to $t \in P = P^{(1)}$ and adding all these estimates we find

$$\|z^{(1)}(t) - z^{\infty}(t)\| \le \sum_{j=1}^{\infty} C\left(\delta(P)2^{-j+1}\right)^{1/2} = \tilde{C}\delta(P)^{1/2}.$$

Since $z^{(1)}(t_k) = z_k$ is the original incremental solution and $z^{\infty}(t) = z(t)$ the time–continuous one the result is established. ∎

7.2 Uniqueness results

It is easy to construct examples with $I(t,\cdot)$ (not strictly) convex such that the solution is not unique, see e.g. Example 7.6. A first uniqueness result was obtained in Theorem 6.5.

Theorem 7.4 *Assume that $E = X$, that I is α–uniformly convex and that $I(t,\cdot) \in \mathrm{C}^3(X,\mathbb{R})$. Then the solutions are unique and depend Lipschitz continuously on the initial value.*

Proof. Let z_1 and z_0 be two solutions. Define

$$\gamma(t) = \langle \sigma_1^* - \sigma_0^*, z_1(t) - z_0(t)\rangle \quad \text{with } \sigma_j^* = \mathrm{D}I(t, z_j(t)),$$

then $\|z_1(t) - z_0(t)\|^2 \le \gamma(t)/\alpha$ by α–uniform convexity. Moreover, by Theorem 7.5 below we know that $\dot{z}_j = (1+\|\dot{z}_j\|)\mathrm{rd}(z_j)$ exists a.e. in $[0,T]$ and satisfies $\|\dot{z}_j(t)\| \le C_1/\alpha$. Thus, we have

$$\dot{\gamma}(t) = \langle \partial_t \mathrm{D}I(t,z_1) - \partial_t \mathrm{D}I(t,z_0), z_1 - z_0\rangle + \langle r_1^*, \dot{z}_1\rangle + \langle r_0^*, \dot{z}_0\rangle$$

where $r_j^* = \mathrm{D}^2 I(t,z_j)[z_j - z_{1-j}] - \sigma_{1-j}^* + \sigma_j^* = 2(\sigma_j^* - \sigma_{1-j}^*) + b_j^*$. Using the estimates

$$\|b_j^*\| = \|\mathrm{D}I(t,z_{1-j}) - \mathrm{D}I(t,z_j) - \mathrm{D}^2 I(t,z_j)[z_{1-j} - z_j]\| \le C_3 \|z_1 - z_0\|^2$$

and $\|\partial_t \mathrm{D}I(t,z_1) - \partial_t \mathrm{D}I(t,z_0)\| \le C_2 \|z_1 - z_0\|$ we find

$$\dot{\gamma} \le C_2 \|z_1 - z_0\|^2 + 2C_3 \|z_1 - z_0\|^2 \frac{C_1}{\alpha} + 2\langle \sigma_0^*, \dot{z}_0 - \dot{z}_1\rangle + 2\langle \sigma_1^*, \dot{z}_1 - \dot{z}_0\rangle. \tag{7.3}$$

To estimate the sum of the last two terms we use the variational inequality (VI) for the solutions $z = z_j$ where, by homogeneity, we can replace $\mathrm{rd}(z_j)$ by $\dot{z}_j$. We insert the test functions $w = \dot{z}_{1-j}$ and subtraction of the two equations yields

$$\langle \sigma_1^* - \sigma_0^*, \dot{z}_1 - \dot{z}_0\rangle \le 0. \tag{7.4}$$

Here the characteristic functions vanish as $\mathrm{T}_{z_j(t)}E = X$ (recall $E = X$) and the terms involving Δ annihilate.

Thus, (7.3) gives $\dot{\gamma} \le C_5 \gamma$ with $C_5 = (C_2 + 2C_1C_3/\alpha)/\alpha$ and, hence, $\|z_1(t) - z_0(t)\|^2 \le \mathrm{e}^{C_5 t}\gamma(0)/\alpha$ which implies uniqueness and Lipschitz continuity. ∎

Theorem 7.4 can be generalized when we find a replacement for (7.4). Such a generalization is given via the *structure condition* (C.1) in Appendix C.

7.3 Temporal regularity

In the general case solutions $z : [0,T] \to E$ will not be continuous but just lie in $\mathrm{BV}([0,T],X)$. This includes the case of I being convex in z, when it is not strictly convex. We obtain continuity if $I(t,\cdot)$ is strictly convex and Lipschitz continuity if $I(t,\cdot)$ is uniformly convex. Even in the simplest case $\dot z$ will have jumps due to the intrinsic nondifferentiability of $\Delta : X \to \mathbb{R}$. Under suitable strong assumptions we will show $\dot z \in \mathrm{BV}([0,T],X)$.

Theorem 7.5 *(a) If $I : [0,T]\times E \to \mathbb{R}$ is continuous and if $I(t,\cdot) : E \to \mathbb{R}$ is strictly convex for all $t \in [0,T]$, then any solution $z : [0,T] \to E$ is continuous.*
(b) If additionally $I(t,\cdot)$ is α–uniformly convex, then $z : [0,T] \to E$ is Lipschitz continuous with

$$\|z(t)-z(s)\| \le \frac{C_2}{\alpha}|t-s|$$

where C_2 is defined in assumption (2.6)(b).

Proof. (a) Take $t \in [0,T]$, then the left and right limits z_- and z_+ exist and satisfy energy identity (2.7). For $\theta \in (0,1)$ let $z_\theta = \theta z_+ + (1-\theta)z_-$ and assume $z_+ \neq z_-$, then strict convexity gives

$$I(t,z_\theta) + \Delta(z_\theta - z_-) < \theta I(t,z_+) + (1-\theta)I(t,z_-) + \theta\Delta(z_+ - z_-) = I(t,z_-).$$

This contradicts the stability of z_- and we conclude the desired continuity $z_+ = z_-$.

(b) Using uniform convexity, for $t > s$, we obtain

$$\begin{aligned}
\alpha\|z(t)-z(s)\|^2 &\le I(s,z(t)) - I(s,z(s)) - \langle \mathrm{D}I(s,z(s)), z(t)-z(s)\rangle \\
&\le -\textstyle\int_{[s,t)} \partial_t I(\tau,z(t))\,\mathrm{d}\tau + I(t,z(t)) - I(s,z(s)) + \Delta(z(t)-z(s)) \\
&= \textstyle\int_{[s,t)} [\partial_t I(\tau,z(\tau)) - \partial_t I(\tau,z(t))]\,\mathrm{d}\tau - \int_{[s,t)} \Delta(\mathrm{d}z) + \Delta(z(t)-z(s)) \\
&\le \textstyle\int_{[s,t)} |\partial_t I(\tau,z(\tau)) - \partial_t I(\tau,z(t))|\,\mathrm{d}\tau \;\le\; \int_{[s,t)} C_2\|z(\tau)-z(t)\|\,\mathrm{d}s.
\end{aligned}$$

Here we have used ($\mathrm{S}_{\mathrm{loc}}$) at s for the second estimate, the energy balance (2.6) on $[s,t]$, $\int_{[s,t)} \Delta(\dot z(\tau))\,\mathrm{d}\tau \ge \Delta(z(t)-z(s))$ for the fourth estimate and finally assumption (2.6).

Let t be fixed and define $\gamma(\tilde s) = \max\{\|z(s)-z(t)\| : s \in [\tilde s, t]\}$, then we have shown $\alpha\|z(t)-z(s)\|^2 \le C_2|t-s|\gamma(s) \le C_2|t-\tilde s|\gamma(\tilde s)$ for all $\tilde s \le s \le t$. Thus, we conclude $\alpha\gamma(\tilde s) \le C_2|t-\tilde s|$ which is the desired result. ■

After having established the well–posedness of the evolution problem (i.e. existence and uniqueness) in certain cases it is desirable to check whether any of the assumptions can be dropped. Unfortunately we do not have an example for nonexistence, except in pathological cases. We can, however, give an example which illustrates that the assumption of uniform convexity in Theorem 7.4 can not be dropped.

Example 7.6 *Let $X = \mathbb{R}$, $E = [-1,1]$, $\Delta(v) = |v|$ and $I(z) = \frac{\alpha}{2}z^2 - g\cdot z$. The existence of solutions is clear by trivial compactness in finite dimensions. Since $E \subsetneq X$ we can not apply Theorem 7.4 directly but one can easily show that for E satisfies the structure condition (C.1). Therefore, by Proposition C.1 we have uniqueness if $\alpha > 0$. For every $\alpha \in \mathbb{R}$ the solutions are monotone in time as long as g is monotone, see Fig. 2 on the left side. If $\alpha \le 0$ we lose uniqueness and Lipschitz continuity. For $\alpha < 0$ every solution z takes only values within $\{-1,1\}$, see Fig. 2 on the right side.*

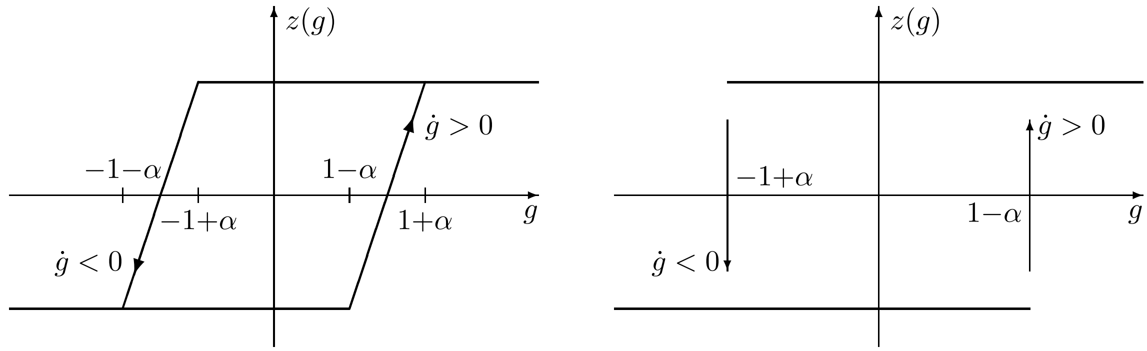

Figure 2: Solutions in Example 7.6 are unique and Lipschitz if $\alpha > 0$ (to the left) and discontinuous and nonunique if $\alpha < 0$ (to the right).

In [HaR95] a convergence result for the incremental problem is given under the assumption that the time continuous solution lies in $z \in W^{2,1}([0,T],X)$, that is $\ddot z = \frac{\mathrm{d}^2}{\mathrm{d}t^2} z \in L^1([0,T],X)$. In fact, their error estimate is

$$\|z_k - z(t_k)\|^2 \le C \max_{j=1,\dots,k}\{t_j - t_{j-1}\} \sum_{j=1}^{k} \|\dot z(t_j) - \tfrac{1}{t_j - t_{j-1}}(z(t_j) - z(t_{j-1}))\|, \tag{7.5}$$

where the constant C does not depend on k and the discretization $0 = t_0 < t_1 < \ldots < t_k$.

The last sum can be estimated by $\int_{[0,t_k]} \|\mathrm{d}\dot z\|$, thus we only need $\dot z \in \mathrm{BV}([0,T],X)$ and not $\dot z \in W^{1,1}([0,T],X)$ as stated in [HaR95]. This difference is crucial since the latter inclusion is false for typical cases whereas the former can be shown under natural additional assumptions.

Example 7.7 *We consider a very simple example with $E = X = \mathbb{R}$, $\Delta(v) = |v|$, $I(t,z) = \frac{1}{2}z^2 - (4t - t^2)z$ and the initial condition $z(0) = 0$. A simple calculation gives the unique solution*

$$z(t) = \begin{cases} 0 & \text{for } t \le 2-\sqrt{3}, \\ 4t - t^2 - 1 & \text{for } t \in [2-\sqrt{3}, 2], \\ 3 & \text{for } t \in [2, 2+\sqrt{2}], \\ 4t - t^2 + 1 & \text{for } t \ge 2+\sqrt{2}. \end{cases}$$

The boundary of the set $F^* = [-1,1]$ is $\{-1,1\}$, and $\dot z$ jumps upon hitting it ($t = 2-\sqrt{3}$ or $2+\sqrt{2}$) but not upon leaving it ($t = 2$).

Theorem 7.8 *Assume $E = X = X^{**}$ and that the boundary of $F^* \subset X^*$ is of class C^2. Let the functional $I(t,z)$ be uniformly convex and C^3. Then, the unique solution z satisfies $\dot z \in \mathrm{BV}([0,T],X)$.*

Proof. According to Theorems 7.5, 7.4 and 7.3 we know that the solution $z : [0,T] \to X$ exists, is unique and Lipschitz continuous. Thus, we have $\dot z \in L^\infty([0,T],X)$.

Since the boundary ∂F^* is smooth there is for each $z^* \in \partial F^*$ a unique outward unit normal $N(z^*) \in X$. The mapping $N : \partial F^* \to X; z^* \to N(z^*)$ is Lipschitz continuous and $\partial\chi_{F^*}(z^*) = \{\lambda N(z^*) : \lambda \in [0,\infty)\}$.

For the given solution z let $\sigma^*(t) = \mathrm{D}I(t, z(t)) \in F^*$. Since z is Lipschitz and I is smooth we know that $\sigma^* : [0, T] \to X^*$ is Lipschitz continuous. Moreover ($\mathrm{E}_{\mathrm{loc}}$) and ($\mathrm{S}_{\mathrm{loc}}$) imply

$$\dot{z}(t) = \begin{cases} 0 & \text{if } -\sigma^*(t) \in \mathrm{int}(F^*), \\ \lambda(t)N(-\sigma^*(t)) & \text{for a.a. } t \in \mathcal{T} \end{cases} \tag{7.6}$$

where $\mathcal{T} = \{t \in [0, T] : -\sigma^*(t) \in \partial F^*\}$ is closed. Now take any $t \in \mathcal{T}$ with $\lambda(t) > 0$ in (7.6). Since $\dot{\sigma}^*(t) = A(t)\dot{z}(t)+\partial_t \mathrm{D}I(t, z(t))$ is perpendicular to $N(-\sigma^*(t))$, we find $\lambda(t) = \frac{-\langle \partial_t \mathrm{D}I(t,z(t)), N(-\sigma^*(t))\rangle}{\langle A(t)N(-\sigma^*(t)), N(-\sigma^*(t))\rangle}$, where $A(t) = \mathrm{D}^2 I(t, z(t)) \in \mathrm{Lin}(X, X^*)$. By the implicit function theorem and the uniqueness (forward in time) it can be shown that there exists $\varepsilon > 0$ such that $z : [t, t + \varepsilon] \to X$ satisfies the differential equation

$$\dot{z} = F(t, z) := \frac{\langle -\partial_t \mathrm{D}I(t, z), \tilde{N}(t, z)\rangle}{\langle \mathrm{D}^2 I(t, z)\tilde{N}(t, z), \tilde{N}(t, z)\rangle}\tilde{N}(t, z) \tag{7.7}$$

where $\tilde{N}(t, z) = N(-\mathrm{D}I(t, z))$. By our smoothness assumptions $F(t, z)$ is continuous in t and Lipschitz in $z \in \mathcal{Z}(t) = \{\, z \in X \,:\, \mathrm{D}I(t, z) \in \partial F^* \,\}$ which is a smooth manifold.

Define the functions

$$h(t) = \begin{cases} \lambda(t) & \text{for } t \in \mathcal{T}, \\ 0 & \text{else;} \end{cases} \quad \text{and } \widehat{N}(t) = \begin{cases} N(-\sigma^*(t)) & \text{for } t \in \mathcal{T}, \\ \text{linear interpolant} & \text{else.} \end{cases}$$

Then $\widehat{N} : [0, T] \to X$ is Lipschitz. The argument with (7.7) shows that $h(t)$ is Lipschitz continuous from the right with a fixed Lipschitz constant independent of t and ε. In particular, h can only be discontinuous at points t^* where $\lim_{t \nearrow t^*} h(t) = 0$ and $\lim_{t \searrow t^*} h(t) > 0$. Between such points h is Lipschitz with a fixed constant. Hence, h can be written as a sum of a Lipschitz function $\tilde{h}$ and a piecewise constant function with at most countably many positive jumps $j_l > 0$ at times t^*_l.

By Theorem 7.5 $\|\dot{z}(t)\|$ and hence $h(t)$ is bounded by C_2/α, which gives

$$\sum_{l=1}^{\infty} j_l + h(0) - T\mathrm{Lip}(\tilde{h}) \leq \frac{C_2}{\alpha}.$$

This implies that h is of bounded variation:

$$\mathrm{Var}(h; [0, T]) \leq T\mathrm{Lip}(\tilde{h}) + \sum_{l=1}^{\infty} j_l \leq 2T\mathrm{Lip}(\tilde{h}) - h(0) + \frac{C_2}{\alpha}.$$

Now the formula $\dot{z}(t) = h(t)\widehat{N}(t)$ gives the desired result, since

$$\begin{aligned} \mathrm{Var}(h(\cdot)\widehat{N}(\cdot); [0, T]) &\leq \mathrm{Var}(h; [0, T])\|\widehat{N}(\cdot)\|_\infty + \|h(\cdot)\|_\infty \mathrm{Var}(\widehat{N}; [0, T]) \\ &\leq \mathrm{Var}(h; [0, T]) + \frac{C_2}{\alpha} T\mathrm{Lip}(\widehat{N}). \end{aligned}$$

■

A The reduced derivative

As above we consider $z \in \mathrm{BV}_-([0,T],X)$. Our aim is to define a substitute for the derivative which works well for rate–independent processes. It will be called reduced derivative and its properties are (i) it is a multiple of the derivative if it exists and (ii) at jump points it is a multiple of the jump vector.

For $z \in \mathrm{BV}_-([0,T],X)$ we define $\widehat{\tau} : [0,T] \to [0,\infty)$ via

$$\widehat{\tau}(t) = t + \int_{[0,t]} \|\mathrm{d}z\| = t + \mathrm{Var}(z,[0,t]).$$

Then, $\widehat{\tau} = \widehat{\tau}_-$ (left–sided limit) and $\widehat{\tau}_+$ (right–sided limit) are strictly increasing and coincide except at the (at most countable) jump points of z. With $\widehat{T} = \widehat{\tau}(T)$ we define the continuous inverse

$$\widehat{t} : \begin{cases} [0,\widehat{T}] & \to \quad [0,T], \\ \qquad \tau & \mapsto \quad \max\{t \in [0,T] : \widehat{\tau}(t) \le \tau\} \end{cases}$$

and the *stretched function* $\widehat{z} \in \mathrm{C}^0([0,\widehat{T}],X)$ via

$$\widehat{z}(\tau) = (1-\theta)z_-(t) + \theta z_+(t) \text{ for } \tau = (1-\theta)\widehat{\tau}_-(t) + \theta\widehat{\tau}_+(t). \tag{A.1}$$

Thus, we have $z(t) = \widehat{z}(\widehat{\tau}(t))$ and $\widehat{z}$ is linearly interpolated at jump points, i.e. at points where $\widehat{\tau}_+(t) > \widehat{\tau}_-(t)$.

By construction we have, for $0 \le \tau_1 < \tau_2 \le \widehat{T}$,

$$\|\widehat{z}(\tau_2)-\widehat{z}(\tau_1)\| \le \tau_2-\tau_1 - (t_2-t_1) \le \tau_2 - \tau_1,$$

where $t_1 \le t_2$ satisfies $\tau_j = (1-\theta_j)\widehat{\tau}_-(t_j) + \theta_j\widehat{\tau}_+(t_j)$. Thus,

$$\widehat{z} \in \mathrm{W}^{1,\infty}([0,\widehat{T}],X), \quad \|\frac{\mathrm{d}}{\mathrm{d}\tau}\widehat{z}\|_{\mathrm{L}^\infty([0,\widehat{T}],X)} \le 1.$$

The derivative $\widehat{v}_z(\tau) = \frac{\mathrm{d}}{\mathrm{d}\tau}\widehat{z}(\tau)$ is defined almost everywhere in the sense of Lebesgue measure ($\lambda-$a.e.) and the reduced derivative is the pullback of $\widehat{v}_z$ via $\widehat{\tau} : [0,T] \to [0,\widehat{T}]$, more precisely

$$\mathrm{rd}(z)(t) \stackrel{\text{def}}{=} \widehat{v}_z(\frac{1}{2}[\widehat{\tau}_-(t)+\widehat{\tau}_+(t)]) \text{ for } t \in [0,T]. \tag{A.2}$$

Associated with this pullback of the derivative is the pullback μ_z of the Lebesgue measure λ on $[0,\widehat{T}]$:

$$\mu_z([t_1,t_2)) \stackrel{\text{def}}{=} \lambda([\widehat{\tau}(t_1),\widehat{\tau}(t_2))) = \widehat{\tau}(t_2) - \widehat{\tau}(t_1) = t_2 - t_1 + \int_{[t_1,t_2]} \|\mathrm{d}z\|.$$

By the general construction of pullbacks the reduced derivative $\mathrm{rd}(z)$ is defined μ_z-a.e. in $[0,T]$. In particular, if t is a jump point, we have $\mu_z(\{t\}) = \|z_+(t)-z_-(t)\|$ and $\mathrm{rd}(z)(t) = \mathrm{Sign}(z_+(t)-z_-(t)) \in \partial B_1^X(0)$. If z has a derivative $\dot{z}$, then $\mathrm{rd}(z)(t) = (1+\|\dot{z}(t)\|)^{-1}\dot{z}(t)$ and $\mu_z((t_1,t_2)) = \int_{t_1}^{t_2}[1+\|\dot{z}(s)\|]\mathrm{d}s$.

We will need the following results.

Theorem A.1 *Let $z \in \mathrm{BV}_-([0,T],X)$ with $\mu_z \in \mathcal{M}([0,T])$ and $\mathrm{rd}(z) \in \mathrm{L}^\infty([0,T],\mu_z)$ as above.*

(a) Let $f : [0,T] \times X \to \mathbb{R}$ be continuous and $f(t,\cdot) : X \to \mathbb{R}$ homogeneous of degree 1, then

$$\int_0^T f(\cdot, \mathrm{d}z) = \int_0^T f(t, \mathrm{rd}(z)(t))\mu_z(\mathrm{d}t).$$

(b) For any $A \in \mathrm{C}^0([0,T],X^*)$ we have

$$\int_0^T \langle A(\cdot), \mathrm{d}z\rangle = \int_0^T \langle A(t), \mathrm{rd}(z)(t)\rangle \mu_z(\mathrm{d}t).$$

(c) For any $0 \le s < t \le T$ and any $g \in \mathrm{C}^1([0,T] \times X, \mathbb{R})$ we have

$$g(t,z(t)) - g(s,z(s)) = \int_s^t \partial_t g(r,z(r))\mathrm{d}r + \int_{[s,t)} \langle G(r), \mathrm{rd}(z)(r)\rangle \mu_z(\mathrm{d}r)$$

where $G(r) = \int_{\theta=0}^1 \mathrm{D}g(r,(1-\theta)z_-(r) + \theta z_+(r))\mathrm{d}\theta$.

Proof. ad (a): Let $t_{j,n} = jT/n$, $\tau_{j,n} = \widehat{\tau}(t_{j,n})$ and $\delta_{j,n} = \tau_{j,n} - \tau_{j-1,n}$. Then

$$\begin{aligned}\int_0^T f(\cdot, \mathrm{d}z) &= \lim_{n\to\infty} \sum_{j=1}^n f(t_{j,n}, z(t_{j,n}) - z(t_{j-1,n})) \\ &= \lim_{n\to\infty} \sum_{j=1}^n f(\widehat{t}(\tau_{j,n}), \frac{1}{\delta_{j,n}}[\widehat{z}(\tau_{j,n}) - \widehat{z}(\tau_{j-1,n})])\delta_{j,n} \\ &= \int_0^{\widehat{T}} f(\widehat{t}(\tau), \frac{\mathrm{d}}{\mathrm{d}\tau}\widehat{z}(\tau))\mathrm{d}\tau.\end{aligned}$$

The latter convergence holds, since we may introduce further points on jump intervals where $\widehat{\tau}_+(t) > \widehat{\tau}_-(t)$ without changing the sum (use $\widehat{t} = t$ on this interval and homogeneity of $f(t,\cdot)$). By pullback (transformation formula) we obtain the desired result.

ad (b). For general $\widehat{A} \in \mathrm{C}^0([0,\widehat{T}],X^*)$ we obtain via pullback

$$\int_0^{\widehat{T}} \langle \widehat{A}(\tau), \frac{\mathrm{d}}{\mathrm{d}\tau}\widehat{z}(\tau)\rangle \mathrm{d}\tau = \int_0^T \langle B(t), \mathrm{rd}(z)(t)\rangle \mu_z(\mathrm{d}t)$$

with $B(f) = \int_{\theta=0}^1 \widehat{A}((1-\theta)\widehat{\tau}_-(t) + \theta\widehat{\tau}_+(t))\mathrm{d}\theta$.

As in the proof of (a) we obtain $\widehat{A}(\tau) = A(\widehat{t}(\tau))$ which gives $B(t) = A(t)$ and the assertion follows.

ad (c). We replace $z \in \mathrm{BV}_-([0,T],X)$ by the stretched function $\widehat{z} \in \mathrm{W}^{1,\infty}((0,T),X)$ and let $\sigma = \widehat{\tau}(s)$ and $\tau = \widehat{\tau}(t)$. We find

$$\begin{aligned}g(t,z(t)) - g(s,z(s)) &= g(\widehat{t}(\tau), \widehat{z}(\tau)) - g(\widehat{t}(\sigma), \widehat{z}(\sigma)) \\ &= \int_\sigma^\tau \frac{\mathrm{d}}{\mathrm{d}\rho}[g(\widehat{t}(\rho), \widehat{z}(\rho)]\mathrm{d}\rho \\ &= \int_\sigma^\tau \partial_t g(\widehat{t}(\rho), \widehat{z}(\rho))\widehat{t}'(\rho)\mathrm{d}\rho + \int_\sigma^\tau \langle \mathrm{D}g(\widehat{t}(\rho), \widehat{z}(\rho)), \widehat{z}'(\rho)\rangle \mathrm{d}\rho \\ &= \int_s^t \partial_t g(r,z(r))\mathrm{d}r + \int_{[s,t)} \langle G(r), \mathrm{rd}(z)(r)\rangle \mu_z(\mathrm{d}r),\end{aligned}$$

where we have used (b) for the last step. ■

A simple consequence of (a) is

$$\int_{[s,t)} \mathrm{r d}(z)(r)\mu_z(\mathrm{d}r) = \int_{[s,t)} \mathrm{d}z = z(t) - z(s). \tag{A.3}$$

B Duality and Cones

For a function $f : X \to (-\infty, \infty]$ we define the Legendre–Fenchel transform $\mathcal{L}f$ via

$$\mathcal{L}f : X^* \to (-\infty, \infty]; v^* \mapsto \sup\{\langle v^*, v\rangle - f(v) : v \in X\}.$$

For a function $f^* : X^* \to (-\infty, \infty]$ we define the inverse Legendre–Fenchel transform $\mathcal{L}^* f^*$ via

$$\mathcal{L}^* f^* : X \to (-\infty, \infty]; v \mapsto \sup\{\langle v^*, v\rangle - f^*(v^*) : v^* \in X^*\}.$$

The functions $\mathcal{L}f$ and $\mathcal{L}^* f^*$ are always lower semicontinuous and convex. If f was lower semicontinuous and convex from the beginning then $\mathcal{L}^*\mathcal{L}f = f$. For these results and the proof of the following theorem we refer to [EkT76].

For convex functions f the subdifferential ∂f is defined via

$$\partial f(z) = \{\, v^* \in X^* \ : \ \forall\, w \in W : f(w) \geq f(z) + \langle v^*, w{-}z\rangle \,\}.$$

Theorem B.1 *Let $f : X \to \mathbb{R} \cup \{\infty\}$ be convex and lower semicontinuous. Set $f^* = \mathcal{L}f$. Then the following statements are equivalent:*
1. $v^* \in \partial f(v)$
2. $f(v{+}w) \geq f(v) + \langle v^*, w\rangle$ for all $w \in X$
3. $v \in \operatorname{argmax}\{\langle v^*, w\rangle - f(w) : w \in X\}$
4. $\langle v^*, v\rangle = f(v) + f^*(v^*)$
5. $v^* \in \operatorname{argmax}\{\langle w^*, v\rangle - f^*(w^*) : w^* \in X^*\}$
6. $f^*(v^*{+}w^*) \geq f^*(v^*) + \langle w^*, v\rangle$ for all $w^* \in X^*$
7. $v \in \partial f^*(v^*)$.

Let C be a closed convex cone in X, i.e. $v \in C$ implies $\alpha v \in C$ for all $\alpha \in [0, \infty)$. The dual cone C^* is defined as

$$C^* = \{v^* \in X^* : \langle v^*, v\rangle \leq 0 \text{ for all } v \in C\}.$$

This duality can also be expressed by the characteristic functions of the cones, namely $\mathcal{L}\chi_C = \chi_{C^*}$. The sum $C_1 + C_2$ and the intersection $C_1 \cap C_2$ of convex cones are again convex cones. Moreover, $(C_1{+}C_2)^* = C_1^* \cap C_2^*$ and $(C_1 \cap C_2)^* = C_1^*{+}C_2^*$.

In particular, for convex sets $E \subset X$ the (inward) tangent cone $\mathrm{T}_z E$ (see (2.1)) is closed and convex and its dual cone is the normal cone $\mathrm{N}_z^* E = \partial\chi_E(z)$. The following lemma is used in Section 3.

Lemma B.2 *Let $E \subset X$ be closed and convex and take $z_0, z_1 \in E$. For $\theta \in (0, 1)$ let $z_\theta = (1{-}\theta)z_0 + \theta z_1$; then $\mathrm{T}_{z_\theta} E = \mathrm{T}_{z_0} E + \mathrm{T}_{z_1} E$ for all $\theta \in (0, 1)$.*

Proof. We set $C_\theta = \mathrm{T}_{z_\theta}E$ and show $C_j \subset C_\theta$ for $j=0$ and 1. By convexity of C_θ this implies $C_0 + C_1 \subset C_\theta$. To show $C_0 \subset C_\theta$ we consider $r_n > 0$ and $v_n \in X$ such that $w_n = z_0 + r_n v_n \in E$ and $v_n \to v \in C_0$. Then $z_\theta + (1-\theta) r_n v_n = (1-\theta) w_n + \theta z_1 \in E$ and $v \in C_\theta$ follows.

The opposite inclusion follows by duality from showing $C_0^* \cap C_1^* \subset C_\theta$. Choose $\nu^* \in C_0^* \cap C_1^*$ and consider the hyperplanes $H_j \subset X$ which have the same (not opposite) normal ν^* and contain z_j. As the set E touches in H_j in z_j and lies on one side of it, we conclude that $H_0 = H_1$. In particular, $z_\theta \in H_0$ which implies $\nu^* \in C_\theta^*$. ■

Proposition B.3 *Let $F^* \subset X^*$ be closed convex set, $C \subset X$ a closed convex cone and $f(v) = (\mathcal{L}^* \chi_{F^*})(v) + \chi_C(v)$. Then, $\mathcal{L} f = \chi_{F^* + C^*}$.*

Proof. Each $v^* \in F^* + C^*$ has the form $v^* = w^* + y^*$ with $w^* \in F^*$ and $y^* \in C^*$. Then,

$$\begin{aligned}(\mathcal{L} f)(v^*) &= \sup\{\langle w^*, v\rangle - (\mathcal{L}^* \chi_{F^*}(v) + \langle y^*, v\rangle - \chi_C(v) : v \in X\} \\ &\le \mathcal{L}(\mathcal{L}^* \chi_{F^*})(w^*) + \mathcal{L}(\chi_C)(y^*) = \chi_{F^*}(w^*) + \chi_{C^*}(y^*) = 0 + 0 = 0.\end{aligned}$$

Now assume $v^* \notin F^* + C^*$. By Mazur's Lemma there exists $\alpha \in \mathbb{R}$ and $v \in X$ such that $\langle v^*, v\rangle > \alpha \ge \langle z^*, v\rangle$ for all $z^* \in F^* + C^*$. This implies

$$\begin{aligned}(\mathcal{L} f)(v^*) &\ge \langle v^*, v\rangle - (\mathcal{L}^* \chi_{F^*})(v) - \mathcal{L}^* \chi_{C^*}(v) \\ &= \langle v^*, v\rangle - \sup_{w^* \in F^*} \langle w^*, v\rangle - \sup_{y^* \in C^*} \langle y^*, v\rangle \\ &= \langle v^*, v\rangle - \sup_{w^* \in F^*, y^* \in C^*} \langle w^* + y^*, v\rangle \ge \langle v^*, v\rangle - \alpha > 0.\end{aligned}$$

Testing with γv rather than v we conclude $(\mathcal{L}^* f)(v^*) \ge \gamma[\langle v^*, v\rangle - \alpha]$ for all $\gamma > 0$ and obtain $(\mathcal{L}^* f)(v^*) = \infty$. This concludes the proof. ■

C Structure condition

Many of the above results need the assumption $E = X$. This is due to the fact that the stability condition ($\mathrm{S}_{\mathrm{loc}}$) changes discontinuously near or on the boundary ∂E via the tangent cone $\mathrm{T}_z E$. If we want to compare two different solutions z_1 and z_2, then the local formulation (LF) involves two different tangent cones $\mathrm{T}_{z_j(t)} E$ such that we cannot guarantee $\dot z_1(t) \in \mathrm{T}_{z_2(t)} E$ and vice versa. However, this feature was essential in the uniqueness and convergence result, see (7.4) and (7.1).

We propose a new structure condition. For $z_j \in E$ we define $\sigma_j^* = \mathrm{D}_z I(t, z_j)$ which lies in the set $-(F^* + \mathrm{N}_{z_j}^* E)$ whenever z_j is stable at time t. Our *structure condition* involves E, F^* and $I(t,z)$ simultaneously and reads as follows.

(SC) [Structure Condition] *For all $R > 0$ there exists a constant $C_{\mathrm{struc}} > 0$ such that for all $t \in [0,T]$, all $z_1, z_2 \in \mathcal{S}(t) \cap \{\|z\|_X \le R\}$ and all $v_j \in V_{z_j}$ the estimate*

$$\langle \sigma_1^* - \sigma_2^*, v_1 - v_2\rangle \le C_{\mathrm{struc}} \|z_1 - z_2\|^2 (\|v_1\| + \|v_2\|) \tag{C.1}$$

holds, where $V_z = \mathrm{N}_{(-\sigma^*)}(F^* + \mathrm{N}_z^* E)$ is the set of possible velocities.

Note that we always have $\mathrm{N}_{(-\sigma_j^*)}(F^*+\mathrm{N}_{z_j}^* E) \subset \mathrm{T}_{z_j}E$ (cf. also the sweeping process formulation (SP) in (3.8)). The intuition behind the structure condition is that, for uniqueness, it is sufficient that the terms which involve the differentials of I and the velocities have a sign.

Proposition C.1 *Let $I \in \mathrm{C}^3(E\times[0,T])$ and let I be α–uniformly convex. If additionally (SC) is satisfied, then every solution of (GF) depends continuously on the initial data and, in particular, uniqueness holds.*

Proof. The claim follows by repeating the proof of Theorem 7.4 where (SC) takes the role of the estimate (7.4) which is just (SC) with $C_{\text{struc}} = 0$. ∎

We illustrate now one case, where (SC) holds and one case where (SC) is violated.

Example C.2 *Let $X = \mathbb{R}^2$, $E = [0,\infty)\times\mathbb{R}$, $I(t,z) = \frac{1}{2}\|z\|_2^2 + \langle \binom{t}{0}, z\rangle$ and $\Delta(v) = |v_1|+|v_2|$ or equivalently $F^* = [-1,1]^2$. Then, the structure condition (SC) holds with $C_{\text{struc}} = 0$.*

We show this by direct calculation. The problem is especially simply as X and X^* can be identified by $\mathrm{D}^2 I(t,z) = \mathrm{id}_{\mathbb{R}^2}$. The stable sets are convex and given by

$$\mathcal{S}(t) = [0, \max\{0, 1-t\}]\times[-1,1] = \{\, z\in E \,:\, \mathrm{D}I(t,z)\in F^*\,\}.$$

The sets $V_z = \mathrm{N}_{-\sigma^*}(F^*+\mathrm{N}_z E)$ of possible velocities satisfies $V_z = \{0\}$ in the interior of $\mathcal{S}(t)$ as well as on the (open) line $\{0\}\times(-1,1)\subset\partial E$. For the other boundary points of $\partial\mathcal{S}(t)$ we find $V_z = -\mathrm{N}_z\mathcal{S}(t)$. In particular,

$$V_z = \begin{cases} \{\, \binom{0}{-\mu} \,:\, \mu\geq 0\,\} & \text{for } z = \binom{\alpha}{1} \text{ with } \alpha\in[0,1-t);\\ \{\, \binom{-\lambda}{-\mu} \,:\, \lambda,\mu\geq 0\,\} & \text{for } z = \binom{1-t}{1};\\ \{\, \binom{-\lambda}{0} \,:\, \lambda\geq 0\,\} & \text{for } z = \binom{1-t}{\beta} \text{ with } |\beta|<1;\\ \{\, \binom{-\lambda}{\mu} \,:\, \lambda,\mu\geq 0\,\} & \text{for } z_1 = \binom{1-t}{-1};\\ \{\, \binom{0}{\mu} \,:\, \mu\geq 0\,\} & \text{for } z = \binom{\alpha}{-1} \text{ with } \alpha\in[0,1-t). \end{cases}$$

Using $\sigma^* = \mathrm{D}I(t,z) = z + \binom{t}{0}$ the structure condition with constant $C_{\text{struc}} = 0$ reduces to $\langle z_2-z_1, v_2-v_1\rangle \leq 0$ for all $v_j\in V_{z_j}$. However, by convexity of $\mathcal{S}(t)$ we have $z_{3-j}-z_j\in\mathrm{T}_{z_j}\mathcal{S}(t)$ whereas $-v_j\in\mathrm{N}_{z_j}\mathcal{S}(t)$, see Fig. 3. This implies $\langle z_j-z_{3-j}, v_j\rangle\leq 0$ and adding these two relations gives the desired result.

Example C.3 *We take the same X, I and Δ as in the previous example, but now $E = \{\, z\in\mathbb{R}^2 \,:\, \langle z, \binom{1}{1}\rangle \leq 0\,\}$. Now, the stable sets may be nonconvex:*

$$\mathcal{S}(t) = \Big(\big[\tbinom{t}{0}+F^*\big]\cap E\Big)\cup\{\, z = \tfrac{\theta}{2}\tbinom{1}{-1} \,:\, \theta\in[-2-t, 2-t]\,\}.$$

Now choose $z_1 = \binom{-1}{1-\varepsilon}$ and $z_2 = \binom{-1-\varepsilon}{1+\varepsilon}$ with a small positive ε. Then, we find $V_{z_1}\{\, \binom{\mu}{0} \,:\, \mu\geq 0\,\}$ and $V_{z_2} = \{0\}$. With $v_1 = \binom{\mu}{0}$ and $v_2 = 0$ we obtain $\langle\sigma_1^*-\sigma_2^*, v_1-v_2\rangle = \varepsilon\mu = 5^{-1/2}\|z_1-z_2\|(\|v_1\|-\|v_2\|)$, see Fig. 3. For $\varepsilon\to 0$ we see that (SC) cannot hold for any finite C_{struc}.

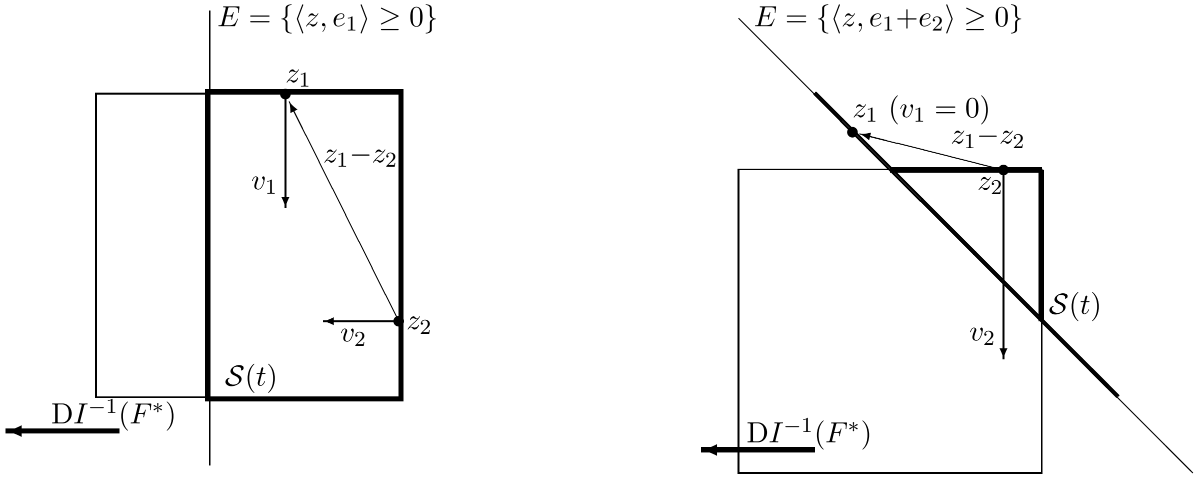

Figure 3: Illustration for the structure condition in Examples C.2 (to the left) and Example C.3 (to the right). In both cases we have $v_j \in V_{z_j}$ and the structure condition reduces to $\langle z_1 - z_2, v_1 - v_2\rangle \leq 0$. In the situation on the left–hand side it is satisfied but for the right–hand side the structure condition can't be satisfied.

Acknowledgements

F.T. thanks the *Deutscher Akademischer Austauschdienst* for generous support. The research was partially supported by DFG through the SFB 404 *Multifield Problems*.

References

[BaP86] V. Barbu & T. Precupanu. *Convexity and optimization in Banach spaces*, D. Reidel Pub. Co, Dordrecht, 2nd edition, 1986.

[BrS96] M. Brokate, J. Sprekels. *Hysteresis and Phase Transitions*, Applied Math. Sciences Vol. 121, Springer–Verlag, New York, 1996.

[CHM01] C. Carstensen, K. Hackl, A. Mielke. Nonconvex potentials and microstructures in finite–strain plasticity. *Proc. London Royal Soc. A* 2001, in press.

[CoV90] P. Colli, A. Visintin. On a class of doubly nonlinear evolution equations. *Comm. Partial Diff. Eqns.* **15** (1990) 737–756

[EkT76] I. Ekeland, R. Temam. *Convex Analysis and Variational Problems.* Horth–Holland & Elsevier Publ. Co., 1976.

[GoM00] S. Govindjee, A. Mielke, G.J. Hall. The free-energy of mixing for n–variant martensitic phase transformations using quasi–convex analysis. *J. Mech. Physics Solids*, submitted May 2001.

[HaR95] W. Han & B.D. Reddy. Computational plasticity: the variational basis and numerical analysis. *Comp. Mech. Advances* **2** (1995) 283–400.

[KuM00] M.S. Kuczma, A. Mielke. Influence of hardening and inhomogeneity on internal loops in pseudoelasticity. *Zeits. Angew. Math. Mechanik* **80** (2000) 291–306.

[KuM97] M. Kunze & M.D.P. Monteiro Marquez. Existence of solutions for degenerate sweeping processes. *J. Convex Analysis* **4** (1997) 165–176.

[KuM98] M. Kunze & M.D.P. Monteiro Marquez. On the discretization of degenerate sweeping processes. *Portugalia Mathematica* **55** (1998) 219–232.

[Mie00] A. Mielke. Finite plasticity, Lie groups and geodesics on SL(d). In "*Geometry, Dynamics, and Mechanics*, P. Newton, A. Weinstein, P. Holmes (eds), Springer–Verlag, to appear. SFB 404 Preprint 26, Stuttgart, 2000.

[Mie01] A. Mielke. Energetic formulation of multiplicative elastoplasticity using the dissipation metric. In preparation, 2001.

[Mon93] M.D.P. Monteiro Marques. Differential Inclusions in Nonsmooth Mechanical Problems. Birkhäuser 1993.

[Mor77] J.J. Moreau. Evolution problem associated with a moving convex set in a Hilbert space. *J. Diff. Eqns.* **26** (1977) 347–374.

[MTL00] A. Mielke, F. Theil, V.I. Levitas. A variational formulation of rate–independent phase transformations using an extremum principle. Preprint 2000–15, SFB 404, Universität Stuttgart, June 2000.

[Vis94] A. Visintin. *Differential Models of Hysteresis.* Applied Math. Sciences Vol. 111, Springer–Verlag, Berlin, 1994.

[Vis01] A. Visintin. A new approach to evolution. *C.R.A.S. Paris* **t. 332**, 233–238, 2001.