

# A VARIATIONAL PRINCIPLE FOR HARDENING ELASTOPLASTICITY\*

ULISSE STEFANELLI<sup>†</sup>

**Abstract.** We present a variational principle governing the quasistatic evolution of a linearized elastoplastic material. In case of linear hardening, the novel characterization allows to recover and partly extend some known results and proves itself to be especially well-suited for discussing general approximation and convergence issues. In particular, the variational principle is exploited in order to prove in a novel setting the convergence of time and space-time discretizations as well as to provide some possible a posteriori error control.

**Key words.** Variational principle, Elastoplasticity, Approximation.

**AMS subject classifications.** 35K55, 49S05, 74C05

**1. Introduction.** The primal initial-boundary value problem of elastoplasticity consists in determining the generalized deformation state of a material subject to external mechanical actions. In particular, starting from some initial state and for a given load and traction, one shall determine the displacement  $u$  of the body from the reference configuration, the inelastic (plastic) part  $p$  of its strain, and, possibly, a vector of internal hardening variables  $\xi$ . In the small deformation regime and within the frame of associative elastoplasticity, the problem is classically formulated in a variational form as that of finding the absolutely continuous trajectory  $t \in [0, T] \mapsto y(t) \in Y$  ( $Y$  is a Banach space) such that

$$\partial\psi(\dot{y}) + Ay \ni \ell \quad \text{a.e. in } (0, T), \quad y(0) = y_0, \quad (1.1)$$

where  $y = (u, p, \xi)$  stands for the vector of unknown fields,  $A : Y \rightarrow Y^*$  (dual) is linear, continuous, and symmetric, and  $\psi : Y \rightarrow [0, \infty]$  is the positively 1-homogeneous and convex dissipation potential ( $\partial$  is the classical subdifferential in the sense of Convex Analysis, see below). Moreover,  $\ell : [0, T] \rightarrow Y^*$  is a given and suitably smooth generalized load (possibly including surface tractions) and  $y_0$  represents the initial state. The reader is referred to Section 2 for some brief mechanical motivation as well as to the classical monographs by DUVAUT & LIONS [7], HAN & REDDY [13], LEMAITRE & CHABOCHE [18], and SIMO & HUGHES [42] for a comprehensive collection of results.

The aim of this paper is that of investigating a global-in-time variational formulation of problem (1.1). In particular, we shall introduce the functional  $\mathcal{F} : W^{1,1}(0, T; Y) \rightarrow [0, \infty]$  on trajectories as

$$\mathcal{F}(y) = \int_0^T (\psi(\dot{y}) + \psi^*(\ell - Ay) - \langle \ell - Ay, \dot{y} \rangle),$$

where  $\psi^*$  stands for the conjugate  $\psi^*(w) = \sup_{v \in Y} (\langle w, v \rangle - \psi(v))$  of  $\psi$  and  $\langle \cdot, \cdot \rangle$  denotes the duality pairing between  $Y^*$  and  $Y$ . The starting point of this analysis

---

\*This research was performed during a visit to the Seminar for Applied Mathematics at ETH Zürich and the Institute of Mathematics of the University of Zürich under the sponsorship of the STM CNR 2006 Program and the Swiss National Science Foundation. The support and hospitality of both institutions is gratefully acknowledged.

<sup>†</sup>IMATI - CNR, v. Ferrata 1, I-27100, Pavia, Italy ([ulisse.stefanelli@imati.cnr.it](mailto:ulisse.stefanelli@imati.cnr.it)).

relies on the fact that solutions of (1.1) and minimizers of  $\mathcal{F}$  fulfilling the initial condition and making  $\mathcal{F}$  zero coincide, namely (see Theorem 3.1)

$$y \text{ solves (1.1) iff } \mathcal{F}(y) = \min \mathcal{F} = 0 \text{ and } y(0) = y_0. \quad (1.2)$$

This variational characterization has a clear mechanical interpretation. Indeed, since  $\psi$  is positively 1-homogeneous, its conjugate  $\psi^*$  turns out to be the indicator function of the convex set  $\partial\psi(0)$ . Hence,  $\min \mathcal{F}$  is actually a constrained minimization problem and we have

$$\mathcal{F}(y) = 0 \text{ iff } \begin{cases} \ell - Ay \in \partial\psi(0) \text{ a.e. in } (0, T), \\ \varphi(T, y(T)) + \int_0^T \psi(\dot{y}) = \varphi(0, y(0)) - \int_0^T \langle \dot{\ell}, y \rangle, \end{cases} \quad (1.3)$$

where we have used the notation  $(t, y) \mapsto \varphi(t, y) = \frac{1}{2}\langle Ay, y \rangle - \langle \ell(t), y \rangle$ . The first relation above expresses the so-called *local stability* [26] of the trajectory whereas the second is nothing but the energy balance at time  $T$ . More precisely,  $\varphi(t, y)$  denotes the complementary energy at time  $t$  for the state  $y$ ,  $\int_0^T \psi(\dot{y})$  represents the dissipation of the system on  $[0, T]$ , and  $-\int_0^T \langle \dot{\ell}, y \rangle$  is related to external actions on  $[0, T]$ . Hence, minimizing  $\mathcal{F}$  consists in selecting the (only) stable trajectory which conserves the energy. In this regards, the reader is referred to the pioneering papers by MOREAU [33, 34, 35].

The interest of variational characterization (1.2) of the differential problem (1.1) relies on the possibility of exploiting the general tools from the Calculus of Variations. Some care is however required. Indeed, although  $\mathcal{F}$  is convex and lower semicontinuous with respect to the weak topology of  $W^{1,1}(0, T; Y)$ , the functional generally fails to be coercive. Moreover, one is not just asked to minimize  $\mathcal{F}$  but also to prove that the minimum is 0. This considerations suggest that the Direct Method is hardly applicable in order to get solutions to (1.1) via the characterization in (1.2).

The first issue of this paper is that of exploiting the variational principle in (1.2) in order to address general approximation procedures. Since solutions and minimizers coincide, a natural tool in order to frame an abstract approach to limiting procedures within (1.1) is that of considering the corresponding minimum problems via  $\Gamma$ -convergence [11]. As the value of the functional is directly quantified to be 0 on the minimizers, what is actually needed here for passing to limits are so-called  $\Gamma$ -lim inf inequalities only and the latter are generally easily available. We shall specifically focus on the case of linear hardening elastoplasticity and apply the above-mentioned perspective in order to recover in a unified and more transparent frame and partly generalize some convergence results for conformal finite elements (Thm. 5.3), time-discretizations (Thm. 6.5), and fully-discrete space-time approximations (Thm. 7.1). In particular, for time-discretization we develop a discrete version of the variational principle (1.2) in the same spirit of the theory of *variational integrators* [25] (see Subsection 6.1). This connection entails also some generalized view at the classical discrete time-schemes (see Subsection 6.5).

A second novel point of the present variational approach consists in the possibility of exploiting  $\mathcal{F}$  in order to estimate a posteriori some approximation error. By letting  $\mathcal{F}(y) = 0$ , we will check that (Cor. 4.5)

$$\max_{[0, T]} \frac{1}{2} \langle A(y - v), y - v \rangle \leq \mathcal{F}(v) \quad \forall v \in W^{1,1}(0, T; Y), \quad v(0) = y_0.$$

If  $A$  shows some coercivity (which is precisely the case of linearized hardening, see Subsection 2.4), and  $v$  is the outcome of some approximation procedure, the estimate above may serve as the basis for some a posteriori estimation procedure, possibly headed to adaptivity (see Subsection 6.7). Let us stress that the latter and (1.3) entail that the distance of a (stable) trajectory from the solution to (1.1) can be uniformly estimated by means of its energy production along the path.

The variational characterization in (1.2) is strictly linked to the celebrated principle by BREZIS, ELEKAND, & NAYROLES [3, 4, 38, 39] for the gradient flow of a convex functional  $j : Y \rightarrow (-\infty, \infty]$  in a Hilbert space  $Y$ , namely

$$\dot{y} + \partial j(y) \ni \ell \quad \text{a.e. in } (0, T). \quad (1.4)$$

In particular, by using the convexity of  $j$  (see details in Section 3.1), this inclusion may be equivalently rewritten as the scalar relation

$$j(y) + j^*(\ell - \dot{y}) - \langle \ell - \dot{y}, y \rangle = 0 \quad \text{a.e. in } (0, T) \quad (1.5)$$

where  $j^*$  is the conjugate of  $j$ . As the above left-hand-side is almost everywhere non-negative for all trajectories  $y$  and it is 0 iff  $y$  solves (1.4), the latter is equivalent to minimize the global functional

$$\mathcal{I}(y) = \int_0^T \left( j(y) + j^*(\ell - \dot{y}) - \langle \ell - \dot{y}, y \rangle \right)$$

and check that  $\mathcal{I}(y) = 0$ . Since its introduction, the latter principle has continuously attracted attention. In particular, it has been exploited in the direction of proving existence [40, 2, 41, 10, 9] (note that the above-mentioned obstructions to the application of the Direct Methods again appear) and the description of long-time dynamics [17]. Moreover, the Brezis-Ekeland-Nayroles approach has been adapted to the case of second order [21, 22] and doubly nonlinear equations [45], and to the study of variable time step discretizations of (1.4) [43] as well.

The variational characterization (1.2) stems basically from the same idea as in (1.5). Namely, the differential inclusion in (1.1) is equivalently rewritten by convexity of  $\psi$  as

$$\psi(\dot{y}) + \psi^*(\ell - Ay) - \langle \ell - Ay, \dot{y} \rangle = 0 \quad \text{a.e. in } (0, T),$$

and (1.2) follows by noting that the above left-hand-side is almost everywhere non-negative for all trajectories  $y$  and it is 0 iff  $y$  is fulfilling the inclusion in (1.1). On the other hand, apart from this conceptual analogy, one has indeed to mention that the characterization (1.2) has little in common with the original Brezis-Ekeland-Nayroles principle. In particular, the two principles turn out to be different even in the case of a quadratic functional  $\psi$  on a Hilbert space.

One has to mention that, of course, (1.2) is not the only possible global-in-time variational characterization of (1.1). Besides minimizing the  $L^2$  space-time norm of the residual (which might be little interesting since the order of the problem is doubled), one has at least to mention VISINTIN [48], where generalized solutions are obtained as minimal elements of a certain partial-order relation on the trajectories, and the recent contribution by MIELKE & ORTIZ [27] where the functional

$$y \mapsto e^{-T/\varepsilon} \varphi(T, y(T)) + \int_0^T e^{-t/\varepsilon} \left( \psi(\dot{y}) + \frac{1}{\varepsilon} \varphi(t, y) \right) \quad (1.6)$$

is minimized among trajectories with  $y(0) = y_0$ . Under extra-smoothness conditions on  $\psi$  (not fulfilled in the current frame), the Euler-Lagrange equations of the latter functional are

$$\begin{aligned} -\varepsilon D^2 \psi(\dot{y}) \ddot{y} + D\psi(\dot{y}) + Ay &= \ell \\ y(0) = y_0, \quad D\psi(\dot{y}(T)) + Ay(T) &= \ell(T). \end{aligned}$$

In particular, minimizing the functional in (1.6) consists in performing a suitable elliptic (in time) regularization of the problem. In the specific case of  $\psi$  positively 1-homogeneous, the limit  $\varepsilon \rightarrow 0$  can be carried out and the minimizers of the functional in (1.6) are proved to converge to the solution of (1.1). The latter approach is quite different from that of (1.2). On the one hand, it is much more general as it naturally applies to the non-smooth case as well (no derivatives of  $\psi$  and  $\phi$  are involved). The results of [27] have been recently extended in the direction of discretizations and relaxation in [31].

**2. Mechanical model.** Let us provide the reader with a brief introduction to the mechanical setting under consideration. Our aim is just that of recalling some essential features of the models as well as their variational formulation. In particular, we restrain from reporting here an extensive discussion on associative elastoplasticity as the latter can be easily recovered from the many contributions on the subject. The reader is particularly referred to the mentioned monographs for some comprehensive presentation.

**2.1. Preliminaries.** We will denote by  $\mathbb{R}_{\text{sym}}^{3 \times 3}$  the space of symmetric  $3 \times 3$  tensors endowed with the natural scalar product  $a : b := \text{tr}(ab) = a_{ij}b_{ij}$  (summation convention). The space  $\mathbb{R}_{\text{sym}}^{3 \times 3}$  is orthogonally decomposed as  $\mathbb{R}_{\text{sym}}^{3 \times 3} = \mathbb{R}_{\text{dev}}^{3 \times 3} \oplus \mathbb{R}1_2$ , where  $\mathbb{R}1_2$  is the subspace spanned by the identity 2-tensor  $1_2$  and  $\mathbb{R}_{\text{dev}}^{3 \times 3}$  is the subspace of deviatoric symmetric  $3 \times 3$  tensors. In particular, for all  $a \in \mathbb{R}_{\text{sym}}^{3 \times 3}$ , we have that  $a = a_{\text{dev}} + \text{tr}(a)1_2/3$ .

We shall assume the reference configuration  $\Omega$  to be a non-empty, bounded, and connected open set in  $\mathbb{R}^3$  with Lipschitz continuous boundary. The space dimension 3 plays essentially no role throughout the analysis and we would be in the position of reformulating our results in  $\mathbb{R}^d$  with no particular intricacy. Our unknown variables are the displacement of the body  $u \in \mathbb{R}^3$ , the plastic strain  $p \in \mathbb{R}_{\text{dev}}^{3 \times 3}$ , and a vector of internal variables  $\xi \in \mathbb{R}^m$  ( $m \in \mathbb{N}$ ) which will describe the hardening of the material. We will denote by  $\varepsilon(u)$  the standard symmetric gradient.

**2.2. Constitutive relation.** Moving within the small-strain regime, we additively decompose the linearized deformation  $\varepsilon(u)$  into the elastic strain  $e$  and the inelastic (or plastic) strain  $p$  as

$$\varepsilon(u) = e + p.$$

Let  $\mathbb{C}$  be the elasticity tensor. By regarding the latter as a symmetric positive definite linear map  $\mathbb{C} : \mathbb{R}_{\text{sym}}^{3 \times 3} \rightarrow \mathbb{R}_{\text{sym}}^{3 \times 3}$ , we shall assume that the orthogonal subspaces  $\mathbb{R}_{\text{dev}}^{3 \times 3}$  and  $\mathbb{R}1_2$  are invariant under  $\mathbb{C}$ . This amounts to say that indeed

$$\mathbb{C}a = \mathbb{C}_{\text{dev}} a_{\text{dev}} + \kappa \text{tr}(a)1_2,$$

for a given  $\mathbb{C}_{\text{dev}} : \mathbb{R}_{\text{dev}}^{3 \times 3} \rightarrow \mathbb{R}_{\text{dev}}^{3 \times 3}$  and a constant  $\kappa$ , and all  $a \in \mathbb{R}_{\text{sym}}^{3 \times 3}$ . The case of isotropic materials is given by  $\mathbb{C}_{\text{dev}} = 2G(1_4 - 1_2 \otimes 1_2/3)$  and  $G$  and  $\kappa$  are respectively

the shear and the bulk moduli. The latter decomposition is not exploited in our analysis but it is clearly suggested by the mechanical application. Moreover, we shall introduce two linear symmetric positive semi-definite hardening moduli  $\mathbb{H}_p : \mathbb{R}_{\text{dev}}^{3 \times 3} \rightarrow \mathbb{R}_{\text{dev}}^{3 \times 3}$  and  $\mathbb{H}_\xi : \mathbb{R}^m \rightarrow \mathbb{R}^m$  (to be identified with a fourth order tensor and a matrix, respectively) and define the Helmholtz free energy  $W : \mathbb{R}_{\text{sym}}^{3 \times 3} \times \mathbb{R}_{\text{dev}}^{3 \times 3} \times \mathbb{R}^m \rightarrow [0, \infty)$  of the material as

$$W(\varepsilon(u), p, \xi) := \frac{1}{2}(\varepsilon(u) - p) : \mathbb{C}(\varepsilon(u) - p) + \frac{1}{2}p : \mathbb{H}_p p + \frac{1}{2}\xi^T \cdot \mathbb{H}_\xi \xi.$$

The generalized stresses  $(\sigma, \eta)$  are conjugate to the above-defined generalized strains  $(e, \xi)$  via the energy  $W$ . In particular, the material is classically assumed to show elastic response,

$$\sigma = \frac{\partial W}{\partial e} = \mathbb{C}e = \mathbb{C}(\varepsilon(u) - p). \quad (2.1)$$

and the thermodynamic force  $\eta$  driving the evolution of the internal variables  $\xi$  is defined as

$$\eta = -\frac{\partial W}{\partial \xi} = -\mathbb{H}_\xi \xi. \quad (2.2)$$

Moreover, moving within the frame of associative elastoplasticity, we assume the existence of a function  $R : \mathbb{R}_{\text{dev}}^{3 \times 3} \times \mathbb{R}^m \rightarrow [0, \infty]$  convex, positively 1-homogeneous, and lower semicontinuous such that

$$\partial R(\dot{p}, \dot{\xi}) \ni \begin{pmatrix} \sigma - \mathbb{H}_p p \\ \eta \end{pmatrix}. \quad (2.3)$$

In particular,  $R$  is asked to be the support function of a convex set  $C^* \in \mathbb{R}_{\text{dev}}^{3 \times 3} \times \mathbb{R}^m$ , i.e.  $R(p) = \sup_{q \in C^*} q : p$ . We will indicate with  $R^*$  its conjugate, namely the indicator function of  $C^*$  given by  $R^*(q) = 0$  if  $q \in C^*$  and  $R^*(q) = \infty$  otherwise. Moreover, we let  $C$  be the domain of  $R$ , namely  $C = D(R) = \{(p, \xi) \in \mathbb{R}_{\text{dev}}^{3 \times 3} \times \mathbb{R}^m : R(p, \xi) < \infty\}$ .

Finally, the above material relations (2.1)-(2.3) can be condensed as the following constitutive material law

$$\partial R(\dot{p}, \dot{\xi}) + \begin{pmatrix} (\mathbb{C} + \mathbb{H}_p)p \\ \mathbb{H}_\xi \xi \end{pmatrix} \ni \begin{pmatrix} \mathbb{C}\varepsilon(u) \\ 0 \end{pmatrix}, \quad (2.4)$$

which in turn can be rephrased in the form of (1.1) by letting

$$y = (p, \xi), \quad Y = \mathbb{R}_{\text{dev}}^{3 \times 3} \times \mathbb{R}^m, \quad \psi = R, \\ A(p, \xi) = ((\mathbb{C} + \mathbb{H}_p)p, \mathbb{H}_\xi \xi), \quad \ell = (\mathbb{C}\varepsilon(u), 0). \quad (2.5)$$

Let us close this subsection by explicitly mentioning three classical linear hardening models [13, Ex. 4.8, p.88]

**Linear kinematic hardening** choose  $\mathbb{H}_p = h_p 1_4$  where  $h_p > 0$ , and  $\mathbb{H}_\xi = 0$ . In this case the internal variable  $\xi$  is not evolving and shall be removed from the set of unknowns.

**Linear isotropic hardening** choose  $\mathbb{H}_p = 0$ ,  $m = 1$ , and  $\mathbb{H}_\xi = h_\xi > 0$ . Moreover, let  $D(R) = \{(p, \xi) \in \mathbb{R}_{\text{dev}}^{3 \times 3} \times \mathbb{R} : |p| \leq \xi\}$ .

**Linear combined kinematic-isotropic hardening** let  $\mathbb{H}_p = h_p 1_4$ ,  $m = 1$ , and  $\mathbb{H}_\xi = h_\xi$  where  $h_p, h_\xi > 0$ . Moreover, let  $D(R) = \{(p, \xi) \in \mathbb{R}_{\text{dev}}^{3 \times 3} \times \mathbb{R} : |p| \leq \xi\}$ .

It is beyond the purpose if this introduction to discuss and justify the above-mentioned material models. The reader shall check the cited references for comments on their relevance within applications and some mechanical motivation.

**2.3. Variational formulation of the quasistatic evolution.** Let us now move to the consideration of the full equilibrium problem. To this aim, we assume that the boundary  $\partial\Omega$  is partitioned in two disjoint open sets  $\Gamma_{\text{tr}}$  and  $\Gamma_{\text{Dir}}$  with  $\partial\Gamma_{\text{tr}} = \partial\Gamma_{\text{Dir}}$  (in  $\partial\Omega$ ). We ask  $\Gamma_{\text{Dir}}$  to be such that there exists a positive constant  $c_{\text{Korn}}$  depending on  $\Gamma_{\text{Dir}}$  and  $\Omega$  such that the Korn inequality

$$c_{\text{Korn}} \|u\|_{H^1(\Omega; \mathbb{R}^3)}^2 \leq \|u\|_{L^2(\Gamma_{\text{Dir}}; \mathbb{R}^3)}^2 + \|\varepsilon(u)\|_{L^2(\Omega; \mathbb{R}_{\text{sym}}^{3 \times 3})}^2, \quad (2.6)$$

holds true for all  $u \in H^1(\Omega; \mathbb{R}^3)$ . It would indeed suffice to impose  $\Gamma_{\text{Dir}}$  to have a positive surface measure (see, e.g., [7, Thm. 3.1, p. 110]).

For the sake of simplicity, we will prescribe homogeneous Dirichlet boundary conditions on  $\Gamma_{\text{Dir}}$  (our analysis extends with little notational intricacy to the case of non-homogeneous Dirichlet boundary conditions as well). On  $\Gamma_{\text{tr}}$  some time-dependent traction will be prescribed instead.

As for the full quasistatic evolution of the material we shall couple the constitutive relation (2.4) with the equilibrium equation

$$\operatorname{div} \sigma + f = 0 \quad \text{in } \Omega. \quad (2.7)$$

Here, we assume to be given the body force  $f : [0, T] \rightarrow L^2(\Omega; \mathbb{R}^3)$  and a surface traction  $g : [0, T] \rightarrow L^2(\Gamma_{\text{tr}}; \mathbb{R}^3)$ .

Then, one can rephrase the problem into the form of (1.1) by choosing

$$y = (u, p, \xi), \quad (2.8)$$

$$Y = \left\{ (u, p, \xi) \in H^1(\Omega; \mathbb{R}^3) \times L^2(\Omega; \mathbb{R}_{\text{dev}}^{3 \times 3}) \times L^2(\Omega; \mathbb{R}^m) \right. \\ \left. \text{such that } u = 0 \text{ on } \Gamma_{\text{Dir}} \right\}, \quad (2.9)$$

$$\langle A(u, p, \xi), (v, q, z) \rangle = \int_{\Omega} \left( (\varepsilon(u) - p) : \mathbb{C}(\varepsilon(v) - q) + p : \mathbb{H}_p q + \xi^T \cdot \mathbb{H}_{\xi} z \right) \\ \forall (v, q, z) \in Y, \quad (2.10)$$

$$\psi(u, p, \xi) = \int_{\Omega} R(p, \xi), \quad (2.11)$$

and defining the total load  $\ell : [0, T] \rightarrow Y^*$  as

$$\langle \ell(t), (u, p, \xi) \rangle = \int_{\Omega} f \cdot u + \int_{\Gamma_{\text{tr}}} g \cdot u \, d\mathcal{H}^2 \quad \forall u \in H^1(\Omega; \mathbb{R}^3), \quad t \in [0, T],$$

where  $\mathcal{H}^2$  is the 2-dimensional Hausdorff measure.

**2.4. The coercivity of  $A$ .** Let us close this introductory discussion by explicitly commenting on the coercivity of the bilinear form induced by  $A$ . We shall recall some sufficient conditions on  $\mathbb{H}_p$ ,  $\mathbb{H}_{\xi}$ , and  $R$  in such a way that there exists a constant  $\alpha > 0$  such that

$$\langle Ay, y \rangle \geq \alpha |y|^2 \quad \forall y \in D(\psi) \quad (2.12)$$

where  $|\cdot|$  is the norm in  $Y$ . This issue is fairly classical [13, Sec. 7.3, p. 167] and we discuss it here for the sake of completeness only.

Of course (2.12) holds (and even for all  $y \in Y$ ) whenever  $\mathbb{H}_p$  and  $\mathbb{H}_{\xi}$  are positive definite (this is the case of the above-mentioned linear combined kinematic-isotropic hardening).

As we have already observed, in case  $\mathbb{H}_\xi = 0$ , the problem naturally reduces to the pair  $(u, p)$  only. Up to this reduction, (2.12) holds (again for all  $y \in Y$ ) when  $\mathbb{H}_p$  is positive definite. This is exactly the case of linear kinematic hardening.

On the other hand, in case  $\mathbb{H}_p = 0$ , the plastic strain will still evolve and one has (2.12) if  $D(R)$  is bounded in the  $p$ -direction for all  $\xi$ , namely if [13, (7.51)]

$$D(R) \subset \{(p, \xi) \in \mathbb{R}_{\text{dev}}^{3 \times 3} \times \mathbb{R}^m : \beta |p|^2 \leq \xi^T \cdot \mathbb{H}_\xi \xi \text{ for some constant } \beta > 0\}, \quad (2.13)$$

which is clearly the case for linear isotropic hardening.

Some generalization of the latter condition could in principle be considered for the case when  $\mathbb{H}_p$  and  $\mathbb{H}_\xi$  are only semi-definite. In particular, (2.12) holds if one assumes (2.13) and

$$\xi \neq 0 \text{ and } \xi^T \cdot \mathbb{H}_\xi \xi = 0 \Rightarrow R(p, \xi) = \infty \quad \forall p \in \mathbb{R}_{\text{dev}}^{3 \times 3}.$$

Let us mention that the most critical case in the class of (2.4) is  $\mathbb{H}_p = 0$ ,  $\mathbb{H}_\xi = 0$  where actually no hardening takes place. This is the situation *perfect plasticity* for which the Sobolev space framework above is not appropriate and one would consider the space  $BD(\Omega)$  of functions of bounded deformations instead [6]. We shall make clear that, even if our variational characterization covers the case of perfect plasticity, the subsequent approximation results apply to the linear hardening situation only.

### 3. Characterization.

**3.1. General assumptions.** Let us start by recalling notation and enlisting the basic assumptions for the following. First of all, we will ask that

$$Y \text{ is a separable and reflexive Banach space.} \quad (3.1)$$

We will use the symbols  $|\cdot|$  for the norm of  $Y$  and  $\langle \cdot, \cdot \rangle$  for the duality pairing between  $Y^*$  (dual) and  $Y$ . The norm in  $Y^*$  will be denoted by  $|\cdot|_*$  instead.

We introduce the functional

$$\begin{aligned} \psi : Y \rightarrow [0, \infty] \quad & \text{proper, convex, lower semicontinuous,} \\ & \text{and positively 1-homogeneous.} \end{aligned} \quad (3.2)$$

Equivalently,  $\psi$  is required to be the support function of a convex and closed set  $C^* \subset Y^*$  containing 0, namely

$$\psi(y) = \sup\{\langle y^*, y \rangle : y^* \in C^*\}. \quad (3.3)$$

We shall define  $C = D(\psi)$ . Hence, the conjugate  $\psi^* : Y^* \rightarrow [0, \infty]$ , which is classically defined as  $\psi^*(y^*) = \sup_{y \in Y} (\langle y^*, y \rangle - \psi(y))$ , is the indicator function of the convex set  $C^*$ , namely  $\psi^*(y^*) = 0$  if  $y^* \in C^*$  and  $\psi^*(y^*) = \infty$  otherwise. Let us remark that  $\psi$  fulfills the triangle inequality  $\psi(a) \leq \psi(b) + \psi(c)$  whenever  $a = b + c$ .

We shall use the symbol  $\partial$  in order to denote the usual subdifferential in the sense of Convex Analysis, namely

$$y^* \in \partial\psi(y) \quad \text{iff} \quad y \in D(\psi) \quad \text{and} \quad \langle y^*, w - y \rangle \leq \psi(w) - \psi(y) \quad \forall w \in Y.$$

Similarly, we define

$$\begin{aligned} y \in \partial\psi^*(y^*) \quad & \text{iff} \quad y^* \in D(\psi^*) \quad \text{and} \quad \langle w^* - y^*, y \rangle \leq \psi^*(w^*) - \psi^*(y^*) \quad \forall w^* \in Y^* \\ & \text{iff} \quad y^* \in C^* \quad \text{and} \quad \langle w^* - y^*, y \rangle \leq 0 \quad \forall w^* \in C^*. \end{aligned}$$

Finally, we recall Fenchel's inequality

$$\psi(y) + \psi^*(y^*) \geq \langle y^*, y \rangle \quad \forall y \in Y, y^* \in Y^*,$$

and remark that equality holds iff  $y^* \in \partial\psi(y)$  (or, equivalently,  $y \in \partial^*\psi^*(y^*)$ ).

As for the operator  $A$  we require

$$A : Y \rightarrow Y^* \quad \text{linear, continuous, and symmetric,} \quad (3.4)$$

and define the function

$$y \rightarrow \phi(y) = \frac{1}{2} \langle Ay, y \rangle,$$

so that  $A = D\phi$ . Moreover, we will ask  $\phi$  to be coercive on  $C = D(\psi)$ , namely we assume that there exists a positive constant  $\alpha$  such that

$$\phi(y) \geq \frac{\alpha}{2} |y|^2 \quad \forall y \in C. \quad (3.5)$$

As we have already commented in Subsection 2.4, the latter coercivity is fulfilled in the situation of elastoplastic evolution with linear kinematic, isotropic, or combined kinematic-isotropic hardening and will turn out to be sufficient for both the forthcoming characterization results.

On the other hand, the following uniqueness-type results will be checked under some stronger coercivity frame and we will ask for

$$\phi(y) \geq \frac{\alpha}{2} |y|^2 \quad \forall y \in C - C. \quad (3.6)$$

Clearly, condition (3.6) is fulfilled when  $\phi$  happens to be coercive on the whole space  $Y$ . The latter applies in particular to the case of linear kinematic and combined kinematic-isotropic hardening elastoplasticity. In this case,  $\phi$  defines an equivalent (squared) norm in  $Y$ .

We shall make use of the following notation

$$\chi(y) = \phi(y) + |y|^2 \quad \forall y \in Y$$

Indeed the latter choice is just motivated by simplicity and could be replaced as well by any other  $\chi : Y \rightarrow [0, \infty)$  such that  $\chi(y) = 0$  iff  $y = 0$  and that  $y \mapsto \chi(y) - \phi(y)$  is lower semicontinuous.

Finally, we shall fix data such that

$$\ell \in L^\infty(0, T; Y^*), \quad y_0 \in C. \quad (3.7)$$

The restriction on the choice of the initial datum in  $C$  is motivated by the coercivity assumption on  $\phi$  in (3.5). On the other hand, we shall explicitly mention that the usual choice for  $y_0$  in elastoplasticity is  $y_0 = 0$ .

In the forthcoming of the paper the above assumptions (3.1)-(3.5) and (3.7) will be tacitly assumed (unless explicitly stated). It should be however clear that the above choice is motivated by the sake of simplicity. Indeed, most of the following results still hold under suitably weaker assumptions, as we shall comment.

**3.2. The functional.** Let the *Lagrangian*  $L : (0, T) \times Y \times Y \rightarrow [0, \infty]$  be defined as

$$\begin{aligned} L(t, y, p) &= \psi(p) + \psi^*(\ell(t) - Ay) - \langle \ell(t) - Ay, p \rangle \\ &\text{for a.e. } t \in (0, T), \forall y, p \in Y, \end{aligned} \quad (3.8)$$

and the functional  $F : W^{1,1}(0, T; Y) \rightarrow [0, \infty]$  as

$$F(y) = \int_0^T L(t, y(t), \dot{y}(t)) \, dt + \chi(y(0) - y_0). \quad (3.9)$$

Now, by simply using the chain rule, we obtain that

$$F(y) = \int_0^T \left( \psi(\dot{y}) + \psi^*(\ell - Ay) - \langle \ell, \dot{y} \rangle \right) + \phi(y(T)) - \phi(y(0)) + \chi(y(0) - y_0).$$

A first remark is that, by exploiting the particular form of  $\chi$ ,

$$\begin{aligned} F(y) &= \int_0^T \left( \psi(\dot{y}) + \psi^*(\ell - Ay) - \langle \ell, \dot{y} \rangle \right) \\ &\quad + \phi(y(T)) + \phi(y_0) - \langle Ay(0), y_0 \rangle + |y(0) - y_0|^2. \end{aligned} \quad (3.10)$$

In particular,  $F$  is clearly convex.

**3.3. The characterization.** Let us state here our variational principle.

**THEOREM 3.1** (Variational principle).  $y \in W^{1,1}(0, T; Y)$  solves (1.1) iff

$$F(y) = 0 = \min F.$$

*Proof.* Owing to Fenchel's inequality we have that

$$L(t, y, p) = 0 \quad \text{iff } \ell(t) - Ay \in \partial\psi(p)$$

and, clearly,  $\chi(y(0) - y_0) = 0$  iff  $y(0) = y_0$ . Hence, all solutions  $y$  of (1.1) are such that  $F(y) = 0$  and vice-versa.  $\square$

Let us remark that the latter variational characterization result holds in much greater generality. The proof made no use of the separability and reflexivity of  $Y$  nor of the linearity of  $A$  (besides its being single-valued and such that  $t \mapsto Ay(t)$  is measurable). Moreover, the positive 1-homogeneity of  $\psi$  is unessential [44]. In particular, the variational approach of Theorem 3.1 can be directly extended to a variety of different dissipative systems possibly including viscous evolution as well. We shall address this perspective in a forthcoming contribution.

We have already observed that  $F$  is convex. Moreover,  $F$  is lower semicontinuous with respect to the weak topology of  $W^{1,1}(0, T; Y)$  since all weakly convergent sequences in  $W^{1,1}(0, T; Y)$  are pointwise weakly convergent as well. Hence, one could be tempted to use the Direct Method in order to get the existence of minimizers, i.e. solutions to equation (1.1). As we commented in the Introduction, this seems to be no trivial task.

First of all, the functional  $F$  need not be coercive with respect to the weak topology of  $W^{1,1}(0, T; Y)$ . Indeed, the functional  $\psi$  may degenerate and hence not control the norm of its argument. Moreover, even in the case when  $\psi$  is non-degenerate, the homogeneity assumption just entails that the sublevels of  $F$  are bounded in  $W^{1,1}(0, T; Y)$  and no weak compactness follows.

Secondly, even assuming coercivity in the weak topology of  $W^{1,1}(0, T; Y)$ , one would still need to prove that the minimum 0 is attained.

**3.4. The variational principle for hardening elastoplasticity.** By referring to the notations of Section 2, let us now present the actual form of the functional  $F$  for the case of the constitutive relation for linearized elastoplastic materials with linear hardening (see (2.5)). In this case the functional reads

$$\begin{aligned} F(p, \xi) &= \int_0^T \left( R(\dot{p}, \dot{\xi}) + R^*(\mathbb{C}(\varepsilon(u) - p) - \mathbb{H}_p p, -\mathbb{H}_\xi \xi) \right) \\ &\quad - \int_0^T \left( (\mathbb{C}(\varepsilon(u) - p) - \mathbb{H}_p p) : \dot{p} - \xi^T \cdot \mathbb{H}_\xi \dot{\xi} \right) \\ &\quad + \frac{1}{2} (p(0) - p_0) : (\mathbb{C} + \mathbb{H}_p)(p(0) - p_0) + \frac{1}{2} (\xi(0) - \xi_0)^T \cdot \mathbb{H}_\xi (\xi(0) - \xi_0) \\ &\quad + |(p(0), \xi(0)) - (p_0, \xi_0)|^2, \end{aligned}$$

for  $(p_0, \xi_0) \in \mathbb{R}_{\text{dev}}^{3 \times 3} \times \mathbb{R}^m$  and  $\varepsilon(u) \in L^\infty(0, T; \mathbb{R}_{\text{sym}}^{3 \times 3})$  some given initial datum

In the situation of the quasistatic evolution, for some given initial datum  $(u_0, p_0, \xi_0) \in Y$ , a load  $f \in L^\infty(0, T; L^2(\Omega; \mathbb{R}^3))$ , and a traction  $g \in L^\infty(0, T; L^2(\Gamma_{\text{tr}}; \mathbb{R}^3))$ , the functional reads (see (2.8)-(2.11))

$$\begin{aligned} F(u, p, \xi) &= \int_0^T \int_\Omega \left( R(\dot{p}, \dot{\xi}) + R^*(\mathbb{C}(\varepsilon(u) - p) - \mathbb{H}_p p, -\mathbb{H}_\xi \xi) \right) \\ &\quad - \int_0^T \int_\Omega f \cdot \dot{u} - \int_0^T \int_{\Gamma_{\text{tr}}} g \cdot \dot{u} \, d\mathcal{H}^2 \\ &\quad + \int_0^T \int_\Omega (\varepsilon(u) - p) : \mathbb{C}(\varepsilon(u) - p) + p : \mathbb{H}_p \dot{p} + \xi^T \cdot \mathbb{H}_\xi \dot{\xi} \\ &\quad + \frac{1}{2} \int_\Omega (\varepsilon(u(0)) - u_0) - (p(0) - p_0) : \mathbb{C}(\varepsilon(u(0)) - u_0) - (p(0) - p_0) \\ &\quad + \frac{1}{2} \int_\Omega \left( (p(0) - p_0) : \mathbb{H}_p (p(0) - p_0) + (\xi(0) - \xi_0)^T \cdot \mathbb{H}_\xi (\xi(0) - \xi_0) \right) \\ &\quad + \frac{1}{2} \int_\Omega |(u(0), p(0), \xi(0)) - (u_0, p_0, \xi_0)|^2, \end{aligned}$$

for all points  $(u, p, \xi) \in Y$  such that

$$\begin{aligned} \int_\Omega (\varepsilon(u) - p) : \mathbb{C} \varepsilon(v) &= \int_\Omega f \cdot v - \int_{\Gamma_{\text{tr}}} g \cdot v \, d\mathcal{H}^2 \\ \forall v \in H^1(\Omega; \mathbb{R}^3) \text{ with } v &= 0 \text{ on } \Gamma_{\text{Dir}}, \text{ a.e. in } (0, T), \end{aligned}$$

and  $F(u, p, \xi) = \infty$  otherwise.

**4. Properties of the minimizers.** For the sake of illustrating the variational principle of Theorem 3.1, we shall collect here some properties of the trajectories belonging to the domain of the functional  $F$  and, in particular, of the minimizers.

**4.1. Trajectories are in  $C$ .** LEMMA 4.1. *Let  $F(y) < \infty$ . Then  $y(t) \in C$  for all  $t \in [0, T]$ .*

*Proof.* Since  $F(y) < \infty$  we have that  $\dot{y} \in C$  almost everywhere in  $(0, T)$ . Hence, for all  $t \in [0, T]$ , we have that  $\int_0^t \dot{y} \in C$  by Jensen's inequality. On the other hand  $y_0 \in C$  and  $y(t) = y_0 + \int_0^t \dot{y}$ . The assertion follows by recalling that  $C$  is a cone.  $\square$

**4.2. Stability at regular points of  $\ell$ .** Assume  $\ell : [0, T] \rightarrow Y^*$  is given and let the set of *stable* states  $S(t) \subset Y$  for  $t \in [0, T]$  be defined as

$$S(t) = \{y \in Y : \phi(y) - \langle \ell(t), y \rangle \leq \phi(w) - \langle \ell(t), w \rangle + \psi(w - y) \quad \forall w \in Y\}.$$

In particular, one has that  $y \in S(t)$  iff it minimizes the convex functional  $G : w \mapsto \phi(w) - \langle \ell(t), w \rangle + \psi(w - y)$ , namely

$$y \in S(t) \quad \text{iff} \quad 0 \in \partial G(y) = Ay - \ell(t) + \partial\psi(0).$$

Let us now check that indeed  $\partial\psi(0) = C^*$ . To this end, owing to (3.3), given  $y^* \in C^*$  one has that  $\langle y^*, w \rangle \leq \psi(w)$  for all  $w \in Y$ . Namely  $C^* \subset \partial\psi(0)$ . On the other hand, assume that there exists  $\xi \in \partial\psi(0) \setminus C^*$ . Hence, by the Hahn-Banach Theorem one finds  $y \in Y$ ,  $\alpha \in \mathbb{R}$ , and  $\varepsilon > 0$  such that

$$\psi(y) \stackrel{\xi \in \partial\psi(0)}{\geq} \langle \xi, y \rangle \geq \alpha + \varepsilon > \alpha - \varepsilon \geq \langle y^*, y \rangle \quad \forall y^* \in C^*$$

and, by passing to the supremum as  $y^* \in C^*$ , we obtain a contradiction. We may summarize this discussion as follows

$$y \in S(t) \quad \text{iff} \quad \ell(t) - Ay \in C^*.$$

LEMMA 4.2 (Stability of the minimizers). *Let  $\ell$  be either left- or right-weakly continuous at some point  $t \in [0, T]$  and  $F(y) < \infty$ . Then  $y(t) \in S(t)$ .*

*Proof.* Since  $F(y) < \infty$ , we have that  $\ell(t) - Ay(t) \in C^*$  for all  $t \in (0, T) \setminus N$ , where  $|N| = 0$ . Choose a sequence  $t_k \in (0, T) \setminus N$  such that  $t_k \rightarrow t$  (from the left or from the right) and  $\ell(t_k) \rightarrow \ell(t)$  weakly in  $Y^*$ . Hence,  $\ell(t_k) - Ay(t_k) \rightarrow \ell(t) - Ay(t)$  weakly in  $Y^*$  and  $\ell(t) - Ay(t) \in C^*$ .  $\square$

In particular, if  $\ell$  happens to be right-continuous at 0, the functional  $F$  will not attain the minimum value 0 unless the initial datum  $y_0$  is stable, namely  $y_0 \in S(0)$ .

**4.3. Equivalent formulations.** Letting now  $\ell \in W^{1,1}(0, T; Y^*)$ , problem (1.1) admits some alternative equivalent formulations [26, Sec. 2.1]. We explicitly mention that  $y \in W^{1,1}(0, T; Y)$  is said to be an *energetic solution* if it solves the *energetic formulation* [32] of (1.1), namely

$$y(t) \in S(t) \quad \forall t \in [0, T], \tag{4.1}$$

$$\phi(y(t)) - \langle \ell(t), y(t) \rangle + \int_0^t \psi(\dot{y}) = \phi(y(0)) - \langle \ell(0), y(0) \rangle - \int_0^t \langle \dot{\ell}, y \rangle$$

$$\forall t \in [0, T], \tag{4.2}$$

$$y(0) = y_0. \tag{4.3}$$

MIELKE & THEIL [32] proved that the latter is equivalent to (1.1) and hence, owing to the characterization of Theorem 3.1, to  $F(y) = 0 = \min F$  (note that the analysis in [32] is much more general and is in particular allowing discontinuous in time evolutions by introducing the above notion of energetic solutions in the frame of functions of bounded variation). For the aim of pointing out some features of our variational approach, we shall present here a direct proof of this fact.

LEMMA 4.3 (Equivalence with the energetic formulation). *Let  $\ell \in W^{1,1}(0, T; Y^*)$ . Then,  $F(y) = 0 = \min F$  iff  $y$  fulfills (4.1)-(4.3).*

*Proof.* Owing to Lemma 4.2, we readily have that the stability condition (4.1) holds iff  $\psi^*(\ell - Ay) = 0$  almost everywhere.

Let  $y$  be such that  $F(y) = 0$ . Then (4.1) and (4.3) hold and  $L(t, y(t), \dot{y}(t)) = 0$  for a.e.  $t \in (0, T)$ . In particular, for all  $t \in [0, T]$ ,

$$\begin{aligned} 0 &= \int_0^t L(s, y(s), \dot{y}(s)) \, ds = \int_0^t \left( \psi(\dot{y}) - \langle \ell, \dot{y} \rangle \right) + \phi(y) \Big|_0^t \\ &= (\phi(y) - \langle \ell, y \rangle) \Big|_0^t + \int_0^t \psi(\dot{y}) + \int_0^t \langle \dot{\ell}, y \rangle, \end{aligned}$$

so that the energy equality (4.2) holds for all  $t \in [0, T]$ .

On the contrary, let  $y \in W^{1,1}(0, T; Y)$  fulfill (4.1)-(4.3). Then  $\chi(y(0) - y_0) = 0$  and  $\psi^*(\ell - Ay) = 0$  almost everywhere (see above). Hence  $F(y) = 0$  follows from the energy equality (4.2) at time  $T$  and an integration by parts.  $\square$

Let us mention that the last lemma proves in particular that the energy equality (4.2) could be equivalently enforced *at the final time*  $T$  only. Moreover, it proves that, as already commented in the Introduction, all *stable trajectories*  $t \mapsto y(t)$  (i.e. trajectories such that  $y(t) \in S(t)$  for all  $t \in [0, T]$ ) are such that the following energy inequality holds

$$\phi(y(t)) - \langle \ell(t), y(t) \rangle + \int_0^t \psi(\dot{y}) \geq \phi(y(0)) - \langle \ell(0), y(0) \rangle - \int_0^t \langle \dot{\ell}, y \rangle \quad \forall t \in [0, T].$$

Hence, we have provided a proof to [26, Prop. 5.7] (note that the referred result is however more general as it is concerned with the *BV* situation, the energy is implicitly depending on time, and no linear structure on  $Y$  is required).

Before closing this subsection, let us explicitly remark that the above inferred equivalence between formulations has been obtained for the absolutely continuous case  $\ell \in W^{1,1}(0, T; Y^*)$  only, whereas the characterization of Theorem 3.1 holds more generally for bounded  $\ell$ .

**4.4. The functional controls the uniform distance: uniqueness.** So far, we have simply reformulated known results in a variational fashion. Here, we present some novel results instead.

LEMMA 4.4 (Uniform distance control via  $F$ ). *We have that*

$$\begin{aligned} \eta(1 - \eta) \max_{t \in [0, T]} \phi(u(t) - v(t)) &\leq \eta F(u) + (1 - \eta) F(v) \\ \forall u, v \in W^{1,1}(0, T; Y), \quad \eta \in [0, 1]. \end{aligned} \tag{4.4}$$

*Proof.* The statement follows from the quadratic character of  $\phi$ . Fix  $t \in [0, T]$  and define  $F^t : W^{1,1}(0, t; Y) \rightarrow [0, \infty]$  as

$$\begin{aligned} F^t(y) &= \int_0^t L(s, y(s), \dot{y}(s)) \, ds + \chi(y(0) - y_0) \\ &= \int_0^t \left( \psi(\dot{y}) + \psi^*(\ell - Ay) - \langle \ell, \dot{y} \rangle \right) \\ &\quad + \phi(y(t)) + \phi(y_0) - \langle Ay(0), y_0 \rangle + |y(0) - y_0|^2. \end{aligned}$$

Then, clearly  $y \mapsto G^t(y) = F^t(y) - \phi(y(t))$  is convex. Hence, letting  $w = \eta u + (1 - \eta)v$  we have that

$$\begin{aligned} 0 &\leq F^t(w) \\ &\leq \eta(G^t(u) + \phi(u(t))) + (1 - \eta)(G^t(v) + \phi(v(t))) - \eta(1 - \eta)\phi(u(t) - v(t)), \end{aligned}$$

whence the assertion follows.  $\square$

The latter lemma exploits the quadratic character of  $\phi$  only. In particular, no coercivity for  $\phi$  is assumed. It should however be clear that its application (as the title of the lemma indeed suggests) will always be referred to the situation where the stronger (3.6) is required, namely when the left-hand side of (4.4) controls

$$\eta(1 - \eta) \frac{\alpha}{2} \max_{[0, T]} |u - v|^2.$$

We now present two immediate corollaries of Lemma 4.4.

**COROLLARY 4.5** (Uniform distance from the minimizer). *Let  $F(y) = 0$ . Then*

$$\max_{t \in [0, T]} \phi(y(t) - v(t)) \leq F(v).$$

This corollary encodes an interesting novel feature of our variational approach, for it provides possible a posteriori error estimator to be used within approximation procedures. It is interesting to remark that the uniform distance of any stable trajectory from the minimizer is controlled by means of its energy production along the path only. Again, although Corollary 4.5 holds under no coercivity assumptions of  $\phi$ , let us mention that its application will be restricted to the frame of (3.6). Finally, we have uniqueness of the minimizers of  $F$  attaining the value 0.

**COROLLARY 4.6** (Uniqueness). *Assume (3.6). Then, there exists at most one trajectory  $y$  such that  $F(y) = 0$ .*

**4.5. Lipschitz bound.** From this point on and throughout the remainder of the paper we shall tacitly assume

$$\ell \in W^{1, \infty}(0, T; Y^*), \quad y_0 \in S(0). \quad (4.5)$$

As already commented after Lemma 4.2, the above restriction on the initial datum is mandatory whenever  $\ell$  admits a weak-right-limit in 0.

As for  $\ell$ , the extra Lipschitz continuity assumption is motivated by the rate-independence of the problem (every absolutely continuous datum can be time-rescaled to a Lipschitz continuous datum) and the following well-known result.

**LEMMA 4.7** (Lipschitz bound). *Assume (3.6) and let  $\ell \in W^{1, \infty}(0, T; Y^*)$ , and  $F(y) = 0$ . Then*

$$\|\dot{y}\|_{L^\infty(0, T; Y)} \leq \frac{1}{\alpha} \|\dot{\ell}\|_{L^\infty(0, T; Y^*)} \quad \text{a.e. in } (0, T). \quad (4.6)$$

The proof of the lemma is exactly the classical one [32, Thm. 7.5], but formulated by means of our variational arguments. We provide it for the sake of completeness.

*Proof.* Let  $0 \leq s < t \leq T$  be fixed. Since  $L(y, \dot{y}) = 0$  almost everywhere we have that

$$\int_s^t \left( \psi(\dot{y}) + \langle \dot{\ell}, y \rangle \right) + (\phi(y) - \langle \ell, y \rangle) \Big|_s^t = 0.$$

On the other hand, owing to the strong monotonicity of  $A$  and the fact that  $y(s) \in S(s)$  (see Lemma 4.2) one obtains

$$\begin{aligned} \phi(y(t) - y(s)) &\leq \phi(y(t)) - \langle \ell(s), y(t) \rangle + \psi(y(t) - y(s)) - \phi(y(s)) - \langle \ell(s), y(s) \rangle \\ &= (\phi(y) - \langle \ell, y \rangle) \Big|_s^t + \int_s^t \langle \dot{\ell}(r), y(t) \rangle \, dr + \psi(y(t) - y(s)). \end{aligned}$$

By taking the sum of these two relations and recalling that, by Jensen,

$$\psi(y(t) - y(s)) = \psi\left(\int_s^t \dot{y}\right) \leq \int_s^t \psi(\dot{y})$$

we get that

$$\phi(y(t) - y(s)) \leq \int_s^t \langle \dot{\ell}(r), y(t) - y(r) \rangle dr.$$

Finally, an application of some extended Gronwall lemma (see [26, Thm. 3.4]) entails that

$$\frac{\alpha}{2} |y(t) - y(s)|^2 \leq \frac{1}{2} \|\dot{\ell}\|_{L^\infty(0, T; Y^*)} |y(t) - y(s)| (t - s)$$

and the assertion follows.  $\square$

**5. Space approximation and stability under data perturbation.** We now apply the characterization results of Theorem 3.1 to the approximation of solutions of (1.1). As already commented in the Introduction, we shall proceed via  $\Gamma$ -convergence [11]. The reader is referred to the monographs by ATTOUCH [1] and DAL MASO [5] for some comprehensive discussion on this topic. Indeed, since Theorem 3.1 directly quantifies the value of the minimum to be 0, what is actually needed for passing to limits are  $\Gamma$ -lim inf inequalities only. We shall illustrate this fact by discussing the simple case of stability under data perturbation first.

**LEMMA 5.1** (Stability under data perturbation). *Assume (3.6), let  $\ell_h \rightarrow \ell$  strongly in  $L^1(0, T; Y^*)$  being uniformly Lipschitz continuous, and  $y_{0,h} \rightarrow y_0$ . Moreover, let  $F_h : W^{1,1}(0, T; Y) \rightarrow [0, \infty]$  be defined as*

$$F_h(y) = \int_0^T \left( \psi(\dot{y}) + \psi^*(\ell_h - Ay) - \langle \ell_h - Ay, \dot{y} \rangle \right) + \chi(y(0) - y_{0,h}),$$

and let  $F_h(y_h) = 0$ . Then  $y_h \rightarrow y$  weakly star in  $W^{1,\infty}(0, T; Y)$  and  $F(y) = 0$ .

*Proof.* Owing to Lemma 4.7, we find a (not relabeled) subsequence  $y_h$  such that  $y_h \rightarrow y$  weakly star in  $W^{1,\infty}(0, T; Y)$ . Hence, we have by lower semicontinuity

$$\begin{aligned} 0 \leq F(y) &\leq \liminf_{h \rightarrow 0} \left( \int_0^T \left( \psi(\dot{y}_h) + \psi^*(\ell_h - Ay_h) - \langle \ell_h, \dot{y}_h \rangle \right) \right. \\ &\quad \left. + \phi(y_h(T)) + \phi(y_{0,h}) - \langle Ay_h(0), y_{0,h} \rangle + |y_h(0) - y_{0,h}|^2 \right) \\ &= \liminf_{h \rightarrow 0} F_h(y_h) = 0. \end{aligned}$$

Hence,  $F(y) = 0$ ,  $y$  is unique, and the assertion follows from the fact that the whole sequence converges.  $\square$

**5.1. Preliminaries on functional convergence.** In order to move to more general approximation situations, we are forced to discuss a suitable functional convergence notion. We limit ourselves in introducing the relevant definitions, referring to the mentioned monographs for all the necessary details.

Recall that  $Y$  is a real reflexive Banach space. Letting  $f_n, f : Y \rightarrow (-\infty, \infty]$  be convex, proper, and lower semicontinuous, we say that  $f_n \rightarrow f$  in the *Mosco sense in*  $Y$  [1, 36] iff, for all  $y \in Y$ ,

$$\begin{aligned} f(y) &\leq \liminf_{n \rightarrow \infty} f_n(y_n) \quad \forall y_n \rightarrow y \text{ weakly in } Y, \\ \exists y_n \rightarrow y \text{ strongly in } Y &\text{ such that } f(y) = \limsup_{n \rightarrow \infty} f_n(y_n). \end{aligned}$$

In particular,  $f_n \rightarrow f$  in the Mosco sense iff  $f_n \rightarrow f$  in the sense of  $\Gamma$ -convergence with respect to both the weak and the strong topology in  $Y$ .

We will consider the situation of approximating functionals  $\psi_h$ . By [1, Thm. 3.18, p. 295], we have that  $\psi_h \rightarrow \psi$  in the Mosco sense in  $Y$  iff  $\psi_h^* \rightarrow \psi^*$  in the Mosco sense in  $Y^*$ . By assuming the functionals  $\psi_h$  to be positively 1-homogeneous, it turns out that the Mosco convergence  $\psi_h \rightarrow \psi$  in  $Y$  is equivalent to the Mosco convergence of sets  $C_h^* \rightarrow C^*$  in  $Y^*$  which reads,

$$\begin{aligned} C_n^* \ni y_n^* \rightarrow y^* \text{ weakly in } Y^* &\Rightarrow y^* \in C^*, \\ \forall y^* \in C^*, \exists y_n^* \in C_n^* : y_n^* \rightarrow y^* &\text{ strongly in } Y^*. \end{aligned}$$

Finally, we repeatedly use a lemma from [45] which we report it here, for the sake of completeness.

LEMMA 5.2 (Cor. 4.4, [45]). *Let  $p \in [1, \infty]$  and  $f_h, f : Y \rightarrow (-\infty, \infty]$  be convex, proper, and lower semicontinuous such that*

$$f(y) \leq \inf \left\{ \liminf_{h \rightarrow 0} f_h(y_h) : y_h \rightarrow y \text{ weakly in } Y \right\} \quad \forall y \in Y.$$

Moreover, let  $y_h \rightarrow y$  weakly in  $W^{1,p}(0, T; Y)$  (weakly star if  $p = \infty$ ), Then, we have that

$$\int_0^T f(y(t)) \, dt \leq \liminf_{h \rightarrow 0} \int_0^T f_h(y_h(t)) \, dt.$$

**5.2. Space approximations.** We now move to the analysis of some space approximation situation, indeed specifically tailored for the case of conformal finite elements. Let us enlist here our assumptions for the sake of later referencing.

We assume to be given

$$\begin{aligned} Y_h \subset Y \text{ closed subspaces such that } h \mapsto Y_h &\text{ increases} \\ \text{and } \cup_{h>0} Y_h &\text{ is dense in } Y, \end{aligned} \tag{5.1}$$

$$\phi_h(y) = \phi(y) \text{ if } y \in Y_h \text{ and } \phi_h(y) = \infty \text{ otherwise.} \tag{5.2}$$

$$\psi_h : Y \rightarrow (-\infty, \infty] \text{ convex, proper, and lower semicontinuous,} \tag{5.3}$$

$$\psi_h \text{ positively 1-homogeneous,} \tag{5.4}$$

$$\psi_h \rightarrow \psi \text{ in the Mosco sense in } Y, \tag{5.5}$$

$$\phi(y) \geq \frac{\alpha}{2} |y|^2 \quad \forall y \in C_h - C_h \text{ where } C_h = D(\psi_h), \tag{5.6}$$

$$\ell_h \rightarrow \ell \text{ pointwise strongly in } Y^*, \tag{5.7}$$

$$\ell_h \text{ uniformly Lipschitz continuous,} \tag{5.8}$$

$$y_{0,h} \in Y_h, \quad y_{0,h} \rightarrow y_0. \tag{5.9}$$

We shall mention that within the frame of conformal finite elements methods the subspaces  $Y_h$  are obviously taken to be finite-dimensional and that the approximating functionals  $\phi_h$  and  $\psi_h$  are usually the restrictions of the functionals  $\phi$  and  $\psi$  on the subspace  $Y_h$ . This is exactly our choice here for  $\phi_h$ . In particular, one shall observe that  $\phi_h \rightarrow \phi$  in the Mosco sense in  $Y$ ,  $D(\partial\phi_h) = Y_h$ , and that

$$A_h y = \partial\phi_h(y) = \partial\phi(y) = D\phi(y) = Ay \quad \forall y \in Y_h. \quad (5.10)$$

As for  $\psi_h$  we are allowing some extra freedom (let us however remark that (5.6) follows from (3.6) as soon as  $\psi_h$  is the restriction of  $\psi$  to  $Y_h$  since, in this case,  $C_h = C \cap Y_h$ ). On the other hand, we are asking  $\psi_h$  to be positively 1-homogeneous, namely we are considering the case of some rate-independent approximation of (1.1) only. The reader is referred instead to EFENDIEV & MIELKE [8], MIELKE, ROSSI, & SAVARÉ [29, 28], TOADER & ZANINI [47], and ZANINI [49] for some results in the direction of rate-dependent approximation of rate-independent processes.

Finally, we shall (re)define the approximating functionals as  $F_h : W^{1,1}(0, T; Y) \rightarrow [0, \infty]$  as

$$F_h(y) = \int_0^T \left( \psi_h(\dot{y}) + \psi_h^*(\ell_h - A_h y) - \langle \ell_h - A_h y, \dot{y} \rangle \right) + \chi_h(y(0) - y_{0,h}),$$

where  $A_h = \partial\phi_h$  and  $\chi_h(\cdot) = \phi_h(\cdot) + |\cdot|^2$ . We have the following.

**THEOREM 5.3** (Convergence of space approximations). *Assume (5.1)-(5.9) and let  $F_h(y_h) = 0$ . Then  $y_h \rightarrow y$  weakly star in  $W^{1,\infty}(0, T; Y)$  and  $F(y) = 0$ .*

*Proof.* By Lemma 4.7, we find a (not relabeled) subsequence  $y_n \rightarrow y$  weakly star in  $W^{1,\infty}(0, T; Y)$  and weakly pointwise. Since  $F_h(y_h) = 0$  we readily check that  $y(t) \in Y_h$  for all  $t \in [0, T]$ . In particular,  $A_h y_h = Ay_h$  for all  $t \in [0, T]$  owing to (5.10). Hence, by lower semicontinuity,

$$\begin{aligned} 0 \leq F(y) &\leq \liminf_{h \rightarrow 0} \left( \int_0^T \left( \psi_h(\dot{y}_h) + \psi_h^*(\ell_h - Ay_h) - \langle \ell_h, \dot{y}_h \rangle \right) \right. \\ &\quad \left. + \phi(y_h(T)) + \phi(y_{0,h}) - \langle Ay_h(0), y_{0,h} \rangle + |y_h(0) - y_{0,h}|^2 \right) \\ &= \liminf_{h \rightarrow 0} \left( \int_0^T \left( \psi_h(\dot{y}_h) + \psi_h^*(\ell_h - A_h y_h) - \langle \ell_h, \dot{y}_h \rangle \right) \right. \\ &\quad \left. + \phi_h(y_h(T)) - \phi_h(y_h(0)) \right) \\ &= \liminf_{h \rightarrow 0} F_h(y_h) = 0, \end{aligned}$$

Note that the integral terms containing  $\psi$  and  $\psi^*$  pass to the  $\liminf$  by means of Lemma 5.2.  $\square$

By inspecting the proof of Theorem 5.3 (which of course generalizes Lemma 5.1), one realizes that, whenever the weak-star precompactness in  $W^{1,\infty}(0, T; Y)$  of the sequence  $y_h$  is assumed, the convergence statement holds more generally in the case  $F_h(y_h) \rightarrow 0$ . Namely, by directly asking for the above-mentioned compactness, one could consider the convergence of some approximated solutions  $y_h$  such that, possibly,  $F_h(y_h) > 0$ . We rephrase this fact in the following statement.

LEMMA 5.4 ( $\Gamma$ -lim inf inequality for  $F_h$ ). *Assume (5.1)-(5.3), (5.5)-(5.7), and (5.9). Then,*

$$F(u) \leq \inf \left\{ \liminf_{h \rightarrow 0} F_h(y_h) : y_h \rightarrow y \text{ weakly star in } W^{1,\infty}(0, T; Y) \right\}.$$

Note that the homogeneity of  $\psi_h$ , the uniform convexity of  $\phi_h$ , and the Lipschitz continuity of  $\ell_h$  play no role here.

Finally, again by looking carefully to the proof of Theorem 5.3 one could wonder if the requirement on the Mosco convergence of  $\psi_h$  could be weakened. Indeed, what we are actually using is only that

$$\psi \leq \Gamma\text{-}\liminf_{h \rightarrow 0} \psi_h \quad \text{and} \quad \psi^* \leq \Gamma\text{-}\liminf_{h \rightarrow 0} \psi_h^* \quad (5.11)$$

with respect to the weak topologies of  $Y$  and  $Y^*$ , respectively. On the other hand, in our specific situation, [45, Lemma 4.1] entails that (5.11) and the fact that  $\psi_h \rightarrow \psi$  in the Mosco sense in  $Y$  are equivalent.

This observation motivates once again the belief that Mosco convergence is the right frame in order to pass to limits within rate-independent problems. For the sake of completeness, let us recall that a first result in the direction of the approximation of the play operator ( $Y$  Hilbert and  $A$  coercive on  $Y$ ) under the Hausdorff convergence of the characteristic sets  $C_h^* = D(\psi_h^*)$  is contained in [16, Thm. 3.12, p. 34] whereas the extension of this result to the more general situation of Mosco converging sets as well as some application to parabolic PDEs with hysteresis is discussed in [46]. More recently, MIELKE, ROUBÍČEK, & STEFANELLI [30] addressed in full generality the issue of  $\Gamma$ -convergence and relaxation for the energetic solutions of rate-independent processes. An alternative convergence proof in the specific case of convex energies is obtained by means of the Brezis-Ekeland-Nayroles approach in [45].

**6. Time-discretization.** Assume now we are given the partitions  $P_n = \{0 = t_n^0 < t_n^1 < \dots < t_n^{N_n} = T\}$  and denote by  $\tau_n^i = t_n^i - t_n^{i-1}$  the  $i$ -th time-step and by  $\tau_n = \max_{1 \leq i \leq N_n} \tau_n^i$  the diameter of the  $n$ -th partition. No constraints are imposed on the possible choice of the time-steps throughout this analysis besides  $\tau_n \rightarrow 0$  as  $n \rightarrow \infty$ . Moreover, let the parameter  $\theta \in [1/2, 1]$  be given.

In the following we will make an extensive use of the following notation: letting  $v = (v^0, \dots, v^{N_n})$  be a vector, we will denote by  $\hat{v}_n$  and  $\bar{v}_n$  two functions of the time-interval  $[0, T]$  which interpolate the values of the vector  $v$  piecewise linearly and backward constantly on the partition  $P_n$ , respectively. Namely

$$\begin{aligned} \hat{v}_n(0) &= v^0, \quad \hat{v}_n(t) = \gamma_n^i(t)v^i + (1 - \gamma_n^i(t))v^{i-1}, \\ \bar{v}_n(0) &= v^0, \quad \bar{v}_n(t) = v^i, \quad \text{for } t \in (t_n^{i-1}, t_n^i], \quad i = 1, \dots, N_n \end{aligned}$$

where

$$\gamma_n^i(t) = (t - t_n^{i-1})/\tau_n^i \quad \text{for } t \in (t_n^{i-1}, t_n^i], \quad i = 1, \dots, N_n.$$

Moreover, we let  $\delta v^i = (v^i - v^{i-1})/\tau_n^i$  for  $i = 1, \dots, N_n$  (so that  $\hat{v}_n = \overline{\delta v_n}$ ) and denote by  $v_\theta$  the vector with components  $v_\theta^i = \theta v^i + (1 - \theta)v^{i-1}$ .

Recall that  $\ell \in W^{1,\infty}(0, T; Y^*)$  and  $y_0 \in S(0)$ . We shall be concerned with the so-called  $\theta$ -scheme for problem (1.1):

$$\partial\psi\left(\frac{y_n^i - y_n^{i-1}}{\tau_n^i}\right) + A(\theta y_n^i + (1-\theta)y_n^{i-1}) \ni \ell(\theta t_n^i + (1-\theta)t_n^{i-1})$$

$$\text{for } i = 1, \dots, N_n, \quad (6.1)$$

$$y_n^0 = y_0. \quad (6.2)$$

One usually refers to the latter as backward or implicit Euler scheme for the choice  $\theta = 1$  and as Crank-Nicolson scheme for  $\theta = 1/2$ .

Owing to the above-introduced notation, the latter scheme can be equivalently rewritten as

$$\partial\psi(y_n^i - y_n^{i-1}) + Ay_{n,\theta}^i \ni \ell(t_{n,\theta}^i) \quad \text{for } i = 1, \dots, N_n, \quad y_n^0 = y_0. \quad (6.3)$$

Clearly, the  $\theta$ -scheme (6.3) is rate-independent. Namely, no time-step appears in (6.3) and the choice of the partition affects the solution via the values of the load  $\ell$  only. In this concern, our focus on variable time-steps partition could be simplified by considering proper rescaled loads  $\ell$  instead. We shall however keep up with it, especially in order to underline the possibility of adapting the partition according to some a posteriori analysis (see Subsection 6.7).

Before moving on, let us comment that, for all  $n$ , the latter scheme as a unique solution. Indeed, given  $y_n^{i-1} \in C$ , it suffices to (uniquely) solve iteratively the incremental problem

$$y_n^i \in \text{Arg min}_{y \in Y} \left( \theta\phi(y) - \langle \ell(t_{n,\theta}^i) - (1-\theta)Ay_n^{i-1}, y \rangle + \psi(y - y_n^{i-1}) \right). \quad (6.4)$$

Note that, since  $y_n^{i-1} \in C$ , the functional under minimization turns out to be uniformly convex. Hence, by (3.5), the minimum problem has a unique solution. In particular, exactly as in Lemma 4.1 we have the following.

LEMMA 6.1.  $y_n^i \in C$  for all  $i = 0, 1, \dots, N_n$ .

A crucial observation is that, as in the continuous case, the discrete trajectories show some sort of stability as well.

LEMMA 6.2 (Stability of the discrete trajectories). *We have that*

$$y_n^i \in \text{Arg min}_{y \in Y} \left( \theta\phi(y) - \langle \ell(t_{n,\theta}^i) - (1-\theta)Ay_n^{i-1}, y \rangle + \psi(y - y_n^i) \right)$$

$$\text{for } i = 1, \dots, N_n. \quad (6.5)$$

In particular, if  $\theta = 1$ , then  $y_n^i \in S(t_n^i)$ .

*Proof.* From the incremental formulation (6.4) and the triangle inequality for  $\psi$  we get that, for all  $y \in Y$ ,

$$\begin{aligned} & \theta\phi(y_n^i) - \langle \ell(t_{n,\theta}^i) - (1-\theta)Ay_n^{i-1}, y_n^i \rangle + \psi(y_n^i - y_n^{i-1}) \\ & \leq \theta\phi(y) - \langle \ell(t_{n,\theta}^i) - (1-\theta)Ay_n^{i-1}, y \rangle + \psi(y - y_n^{i-1}) \\ & \leq \theta\phi(y) - \langle \ell(t_{n,\theta}^i) - (1-\theta)Ay_n^{i-1}, y \rangle + \psi(y - y_n^i) + \psi(y_n^i - y_n^{i-1}), \end{aligned}$$

whence the assertion follows.  $\square$

Again as in the continuous case, we readily check that

$$\text{relation (6.5) holds iff } \ell(t_{n,\theta}^i) - Ay_{n,\theta}^i \in C^*. \quad (6.6)$$

**6.1. The discrete variational principle.** We shall now present a discrete version of the variational principle of Theorem 3.1.

We define  $L_n^{\theta,i}(y, z) : Y \times Y \rightarrow [0, \infty]$  as

$$\begin{aligned} L_n^{\theta,i}(y, z) = & \psi\left(\frac{y-z}{\tau_n^i}\right) + \psi^*\left(\ell(t_{n,\theta}^i) - A(\theta y + (1-\theta)z)\right) \\ & - \langle \ell(t_{n,\theta}^i) - A(\theta y + (1-\theta)z), \frac{y-z}{\tau_n^i} \rangle, \end{aligned}$$

and the functionals  $F_n^\theta : Y^{N_n+1} \rightarrow [0, \infty]$  as

$$F_n^\theta(y_n^0, \dots, y_n^{N_n}) = \sum_{i=1}^{N_n} \tau_n^i L_n^{\theta,i}(y_n^i, y_n^{i-1}) + \chi(y_n^0 - y_0).$$

LEMMA 6.3 (Discrete variational principle).  $(y_n^0, \dots, y_n^{N_n})$  solves (6.3) iff  $F_n^\theta(y_n^0, \dots, y_n^{N_n}) = 0 = \min F_n^\theta$ .

*Proof.* Analogously to the continuous case, we have that, for all  $i = 1, \dots, N_n$ ,

$$\partial\psi(\delta y_n^i) + A y_{n,\theta}^i \ni \ell(t_{n,\theta}^i) \quad \text{iff} \quad L_n^{\theta,i}(y_n^i, y_n^{i-1}) = 0,$$

and  $y_n^0 = y_0$  iff  $\chi(y_n^0 - y_0) = 0$ .  $\square$

Let us observe that the functional  $F_n^\theta$  is convex and lower-semicontinuous. Moreover, by the homogeneity of  $\psi$  (see (3.2)),  $F_n^\theta$  is actually independent of the time steps. In fact, we have

$$\begin{aligned} F_n^\theta(y_n^0, \dots, y_n^{N_n}) = & \sum_{i=1}^{N_n} \left( \psi(y_n^i - y_n^{i-1}) + \psi^*(\ell(t_{n,\theta}^i) - A y_{n,\theta}^i) \right) \\ & - \langle \ell(t_{n,\theta}^i) - A y_{n,\theta}^i, y_n^i - y_n^{i-1} \rangle + \chi(y_n^0 - y_0). \end{aligned}$$

The idea of dealing with time-discretizations via a discrete variational principle closely relates our analysis to the theory of so-called *variational integrators*. The latter are numerical schemes stemming from the approximation of the action functional in Lagrangian Mechanics. By referring the reader to the monograph [12] and the survey [25], we shall restrain here from giving a detailed presentation of the subject and limit ourselves to some (necessarily sketchy) considerations. Letting  $(t, y, p) \in [0, T] \times \mathbb{R}^m \times \mathbb{R}^m \mapsto \mathcal{L}(t, y, p)$  denote the Lagrangian of a (finite dimensional, for simplicity) system, the Hamilton principle asserts that the actual trajectory  $t \mapsto y(t)$  of the system minimizes the action functional

$$y \mapsto \int_0^T \mathcal{L}(t, y(t), \dot{y}(t)) \, dt$$

among all curves with prescribed endpoints, thus solving the Lagrange equations

$$\partial_{y_i} \mathcal{L} - \frac{d}{dt} \partial_{p_i} \mathcal{L} = 0 \quad \text{for } i = 1, \dots, m. \quad (6.7)$$

Hence, a natural idea is that of deriving numerical schemes for Lagrangian mechanics by applying some quadrature procedure to the action functional, i.e. discretizing

Hamilton's principle. The resulting discrete schemes show comparable performance with respect to other methods but generally enjoy some interesting extra (and often crucial) properties such as the conservation of suitable quantities [20]. Variational integrators have been intensively applied in finite-dimensional contexts and, more recently, to the situation of nonlinear wave equations [24] and non-equilibrium elasticity [19].

The present analysis may bear some resemblance to the above-mentioned theory. Indeed, the formulation of the  $\theta$ -scheme in the case  $\theta = 1/2$  stems exactly from the midpoint quadrature of the functional  $F$  as

$$\begin{aligned} & \int_{t_n^{i-1}}^{t_n^i} L(t, \hat{y}(t), \hat{y}'(t)) \, dt \\ &= \tau_n^i L(t_{n,1/2}^i, \hat{y}(t_{n,1/2}^i), \hat{y}'(t_{n,1/2}^i)) \\ &= \psi(y^i - y^{i-1}) + \psi^* \left( \ell(t_{n,1/2}^i) - A \left( \frac{y^i + y^{i-1}}{2} \right) \right) \\ &\quad - \langle \ell(t_{n,1/2}^i) - A \left( \frac{y^i + y^{i-1}}{2} \right), y_n^i - y_n^{i-1} \rangle, \end{aligned}$$

where  $\hat{y}$  is taken to be piecewise affine on the partition  $P_n$ .

On the other hand, our focus here is quite different. First of all, we are not dealing with the Hamilton principle (endpoints are not fixed) as we are not aimed at solving the Euler-Lagrange equations for  $F$  (i.e., solve (6.7)). Secondly, we are specifically interested at infinite-dimensional situations, namely PDEs. Finally, the only choice of  $\theta$  which is directly related with a quadrature of  $F$  is  $\theta = 1/2$  and we are not considering higher order schemes.

Before closing this discussion, let us mention that some  $\Gamma$ -convergence techniques have been recently exploited in the (finite-dimensional) frame of variational integrators by MÜLLER & ORTIZ [37] (see also [23]).

**6.2. Stability of the  $\theta$ -scheme.** It is known since HAN & REDDY [15, 13] that the choice  $\theta < 1/2$  in (6.3) leads to an unconditionally unstable scheme and that, on the contrary, for  $\theta \in [1/2, 1]$  the  $\theta$ -scheme is stable in  $H^1(0, T; Y)$  when  $Y$  is a Hilbert space and the partitions are chosen to be uniform.

Here we shall provide an alternative stability proof by taking into account the Banach-space frame.

LEMMA 6.4 (Stability). *Assume (3.6) and let  $\theta \in [1/2, 1]$ . Then, the solution to the  $\theta$ -scheme (6.3) fulfills*

$$\|\hat{y}_{n,\theta}\|_{L^\infty(0,T;Y)} \leq \frac{1}{\alpha} \|\dot{\ell}\|_{L^\infty(0,T;Y^*)} \quad \text{if } \theta = 1 \text{ or } \theta = \frac{1}{2}. \quad (6.8)$$

Moreover, for constant time-steps,

$$\|\hat{y}_{n,\theta}\|_{L^\infty(0,T;Y)} \leq \frac{1}{\alpha(2\theta - 1)} \|\dot{\ell}\|_{L^\infty(0,T;Y^*)} \quad \text{if } \frac{1}{2} < \theta < 1. \quad (6.9)$$

Our argument coincides with that of [32, Thm. 4.4] in the case of the implicit Euler scheme, i.e.  $\theta = 1$  and it is an extension of the latter for the case  $1/2 < \theta < 1$ . Here, we do not play with the variational inequality by choosing suitable tests but use the scalar relations  $L_n^{\theta,i}(y_n^i, y_n^{i-1}) = 0$  instead (this makes however no substantial

difference since the latter scalar relations are exactly the outcome of the test on the variational inequality in [32, Thm. 4.4]).

The stability proof for the Crank-Nicolson scheme  $\theta = 1/2$  is quite different from former arguments and stems as a direct outcome of our variational approach. In both cases  $\theta = 1$  and  $\theta = 1/2$ , the stability constant  $1/\alpha$  is sharp (see Lemma 4.7).

We complement this analysis by providing the stability for the  $\theta$ -scheme for  $1/2 < \theta < 1$  in the case of constant time-steps (likely with a non-optimal, although explicit, stability constant).

Note that the stability estimates (6.8)-(6.9) do not hold in the classical parabolic situation (i.e.  $\psi$  quadratic on a Hilbert space) unless the very restrictive compatibility assumption  $Ay_0 = \ell(0)$  is made (and in this case, the proof below just goes through). Evidence of this failure is available even in the simplest scalar situation ( $Y = \mathbb{R}$ ) of problem  $\dot{y}(t) + y(t) = 0$  for  $t > 0$  and  $y(0) = y_0$ . Indeed, the corresponding  $\theta$ -scheme fulfills (6.8)-(6.9) iff  $y_0 = 0$  (and hence  $y \equiv 0$ ). The current rate-independent situation turns out to be better behaved since the restriction  $Ay_0 = \ell(0)$  of the parabolic case is replaced by the weaker  $y_0 \in S(0)$  (see (4.5)).

*Proof.* Let us prove the stability of the Crank-Nicolson scheme  $\theta = 1/2$  first. For this aim, it suffices to recall that

$$\begin{aligned} 0 &= F_n^{1/2}(y_n^0, \dots, y_n^{N_n}) \\ &= \int_0^T \left( \psi(\dot{\hat{y}}_n) + \psi^*(\hat{\ell}_n - A\hat{y}_n) - \langle \hat{\ell}_n - A\hat{y}_n, \dot{\hat{y}}_n \rangle \right) + \chi(\hat{y}_n(0) - y_0). \end{aligned}$$

Hence  $\hat{y}_n$  minimizes the functional  $F$  where  $\ell$  is replaced by  $\hat{\ell}_n$ . The stability estimate follows from Lemma 4.7.

Let us now move to the case  $1/2 < \theta \leq 1$ . Relation (6.5) applied at level  $i - 1$  for some  $i = 2, \dots, N_n$  along with the choice  $y = y_n^i$  entails that

$$\begin{aligned} &\theta\phi(y_n^i - y_n^{i-1}) + \theta\phi(y_n^{i-1}) - \langle \ell(t_{n,\theta}^{i-1}) - (1-\theta)Ay_n^{i-2}, y_n^{i-1} \rangle \\ &\leq \theta\phi(y_n^i) - \langle \ell(t_{n,\theta}^{i-1}) - (1-\theta)Ay_n^{i-2}, y_n^i \rangle + \psi(y_n^i - y_n^{i-1}), \end{aligned}$$

where the extra-term  $\theta\phi(y_n^i - y_n^{i-1})$  is obtained from the fact that  $\phi$  is quadratic. Hence, we have that

$$\begin{aligned} &\theta\phi(y_n^i - y_n^{i-1}) + \theta\phi(y_n^{i-1}) - \theta\phi(y_n^i) \\ &\leq \langle \ell(t_{n,\theta}^{i-1}), y_n^{i-1} - y_n^i \rangle + (1-\theta)\langle A(y_n^{i-2} - y_n^{i-1}), y_n^i - y_n^{i-1} \rangle + \psi(y_n^i - y_n^{i-1}) \\ &\quad + (1-\theta)\langle Ay_n^{i-1}, y_n^i - y_n^{i-1} \rangle \\ &= \langle \ell(t_{n,\theta}^{i-1}), y_n^{i-1} - y_n^i \rangle + (1-\theta)\langle A(y_n^{i-2} - y_n^{i-1}), y_n^i - y_n^{i-1} \rangle + \psi(y_n^i - y_n^{i-1}) \\ &\quad - (1-\theta)\left(\phi(y_n^{i-1}) + \phi(y_n^i - y_n^{i-1}) - \phi(y_n^i)\right), \end{aligned}$$

so that

$$\phi(e_n^i) + \phi(y_n^{i-1}) - \phi(y_n^i) \leq -\langle \ell(t_{n,\theta}^{i-1}), e_n^i \rangle + (\theta - 1)\langle Ae_n^{i-1}, e_n^i \rangle + \psi(e_n^i), \quad (6.10)$$

where we have used  $e_n^i = y_n^i - y_n^{i-1}$  in order to shorten notations.

Next, from  $L_n^{\theta,i}(y_n^i, y_n^{i-1}) = 0$  for  $i = 1, \dots, N_n$ , we obtain that

$$\begin{aligned} 0 &= \psi(e_n^i) - \langle \ell(t_{n,\theta}^i) - Ay_n^i, e_n^i \rangle \\ &= \psi(e_n^i) - \langle \ell(t_{n,\theta}^i), e_n^i \rangle + \theta\left(\phi(y_n^i) + \phi(e_n^i) - \phi(y_n^{i-1})\right) \\ &\quad - (1-\theta)\left(\phi(y_n^{i-1}) + \phi(e_n^i) - \phi(y_n^i)\right). \end{aligned}$$

In particular, we have checked that

$$\psi(e_n^i) + \phi(y_n^i) - \phi(y_n^{i-1}) + (2\theta - 1)\phi(e_n^i) = \langle \ell(t_{n,\theta}^i), e_n^i \rangle. \quad (6.11)$$

We take the sum between the latter and (6.10) and get that

$$2\theta\phi(e_n^i) \leq \langle \ell(t_{n,\theta}^i) - \ell(t_{n,\theta}^{i-1}), e_n^i \rangle + (\theta - 1)\langle Ae_n^{i-1}, e_n^i \rangle,$$

or, equivalently,

$$\langle Ae_{n,\theta}^i, e_n^i \rangle \leq \langle \ell(t_{n,\theta}^i) - \ell(t_{n,\theta}^{i-1}), e_n^i \rangle. \quad (6.12)$$

Now, if  $\theta = 1$ , we conclude that

$$|e_n^i| \leq \frac{1}{\alpha} |\ell(t_{n,\theta}^i) - \ell(t_{n,\theta}^{i-1})|_*,$$

and the assertion follows.

In case  $1/2 \leq \theta < 1$  and for a constant time-step partition, one proceeds from (6.12) by computing

$$\begin{aligned} & \tau_n \|\hat{\ell}_n\|_{L^\infty(0,T;Y^*)} \sqrt{\frac{2}{\alpha}} \sqrt{\phi(e_n^i)} \\ & \geq \langle \ell(t_{n,\theta}^i) - \ell(t_{n,\theta}^{i-1}), e_n^i \rangle \\ & \geq \theta \langle Ae_n^i, e_n^i \rangle + (1 - \theta) \langle Ae_n^{i-1}, e_n^i \rangle \\ & = (2\theta - 1) \langle Ae_n^i, e_n^i \rangle + (1 - \theta) \langle A(e_n^i + e_n^{i-1}), e_n^i \rangle \\ & \geq 2(2\theta - 1)\phi(e_n^i) + (1 - \theta) (\phi(e_n^i) - \phi(e_n^{i-1})) \end{aligned}$$

Note that the coefficient  $(2\theta - 1)$  is strictly positive as  $\theta > 1/2$ . By using the fact that  $y_n^0 = y_0 \in S(0)$  (recall (4.5)), we readily check that

$$\phi(e_n^1) + \phi(y_n^0) - \langle \ell(0), y_n^0 \rangle \leq \phi(y_n^1) - \langle \ell(0), y_n^1 \rangle + \psi(e_n^1) \quad (6.13)$$

and, by adding the latter to (6.11) for  $i = 1$  we have

$$2\theta\phi(e_n^1) \leq \langle \ell(t_{n,\theta}^1) - \ell(0), e_n^1 \rangle \leq \tau_n \|\hat{\ell}_n\|_{L^\infty(0,T;Y^*)} \sqrt{\frac{2}{\alpha}} \sqrt{\phi(e_n^1)}. \quad (6.14)$$

Let us define

$$a_i^2 = \phi\left(\frac{y_n^i - y_n^{i-1}}{\tau_n}\right) = \phi(e_n^i)/\tau_n^2,$$

$$C_0 = \frac{2(2\theta - 1)}{1 - \theta}, \quad C_1 = \frac{1}{1 - \theta} \sqrt{\frac{2}{\alpha}} \|\hat{\ell}_n\|_{L^\infty(0,T;Y^*)}, \quad C_2 = \frac{C_1}{C_0},$$

so that, owing to (6.13), (6.14), and using the fact that  $2(2\theta - 1) < 2\theta$ ,

$$(C_0 + 1)a_i^2 - a_{i-1}^2 \leq C_1 a_i \quad \text{for } i = 2, \dots, N_n,$$

$$\begin{aligned} a_1 & \leq \frac{1}{2\theta} \sqrt{\frac{2}{\alpha}} \|\hat{\ell}_n\|_{L^\infty(0,T;Y^*)} \leq \frac{1}{2(2\theta - 1)} \sqrt{\frac{2}{\alpha}} \|\hat{\ell}_n\|_{L^\infty(0,T;Y^*)} \\ & = \frac{1}{1 - \theta} \sqrt{\frac{2}{\alpha}} \|\hat{\ell}_n\|_{L^\infty(0,T;Y^*)} \frac{1 - \theta}{2(2\theta - 1)} \\ & = \frac{C_1}{C_0} = C_2. \end{aligned}$$

Now, since  $(C_0 + 1)C_2^2 - C_1C_2 = C_2^2$ , we easily prove by induction that  $a_i \leq C_2$  and the assertion follows.  $\square$

**6.3. Convergence.** We shall prove the weak-star  $W^{1,\infty}(0, T; Y)$  convergence for the  $\theta$ -method. This result has to be compared with that of Han & Reddy [14, Thm. 3.4] where the uniform convergence of the backward constant interpolations is obtained. Our result is weaker than that of [14, Thm. 3.4] since we are not providing strong convergence. On the other hand, we believe our half-page proof to be possibly more transparent than the long argument developed in [14]. Let us moreover mention that in the Hilbertian case and for  $A$  coercive on  $Y$ , the strong convergence in  $W^{1,p}(0, T; Y)$  for all  $p < \infty$  of the Euler method  $\theta = 1$  has been proved in [16, Prop. 3.9, p. 33].

**THEOREM 6.5** (Convergence for the  $\theta$ -method). *Assume (3.6) and let  $F_n^\theta(y_n^0, \dots, y_n^{N_n}) = 0$ . Then,  $\hat{y}_n \rightarrow y$  weakly star in  $W^{1,\infty}(0, T; Y)$  where  $F(y) = 0 = \min F$ .*

*Proof.* Owing to Lemma 6.4, we can extract a (not relabeled) subsequence such that  $\hat{y}_n \rightarrow y$  weakly star in  $W^{1,\infty}(0, T; Y)$ , hence weakly pointwise in  $Y$ . Moreover, we clearly have that both  $\bar{y}_n$  and  $\bar{y}_{n,\theta}$  converge at the same limit weakly star in  $L^\infty(0, T; Y)$ . Finally, we directly check that  $\bar{\ell}_{n,\theta} \rightarrow \ell$  strongly in  $L^\infty(0, T; Y^*)$ . By observing that, since  $\theta \geq 1/2$ ,

$$\begin{aligned} \tau_n^i \langle A(\theta y_n^i + (1-\theta)y_n^{i-1}), \delta y_n^i \rangle &= \phi(y_n^i) + (2\theta - 1)\phi(y_n^i - y_n^{i-1}) - \phi(y_n^{i-1}) \\ &\geq \phi(y_n^i) - \phi(y_n^{i-1}), \end{aligned}$$

we compute that

$$\begin{aligned} 0 &= F_n^\theta(y_n^0, \dots, y_n^{N_n}) \\ &\geq \int_0^T \left( \psi(\hat{y}_n) + \psi^*(\bar{\ell}_{n,\theta} - A\bar{y}_{n,\theta}) - \langle \bar{\ell}_{n,\theta}, \hat{y}_n \rangle \right) \\ &\quad + \phi(\hat{y}_n(T)) - \phi(\hat{y}_n(0)) + \chi(\hat{y}_n(0) - y_0) \\ &= \int_0^T \left( \psi(\hat{y}_n) + \psi^*(\bar{\ell}_{n,\theta} - A\bar{y}_{n,\theta}) - \langle \bar{\ell}_{n,\theta}, \hat{y}_n \rangle \right) \\ &\quad + \phi(\hat{y}_n(T)) + \phi(y_0) - \langle A\hat{y}_n(0), y_0 \rangle + |\hat{y}_n(0) - y_0|^2. \end{aligned}$$

Finally, it suffices to pass to the  $\liminf$  above as  $n \rightarrow \infty$  and exploit lower semi-continuity and the stated convergences in order to obtain that  $F(y) \leq 0$ . Hence, by Theorem 3.1 and Corollary 4.6,  $y$  is the only solution to (1.1) and the whole sequence  $\hat{y}_n$  converges.  $\square$

**6.4. The functional controls the uniform distance.** We shall reproduce at the discrete level the results of Subsection 4.6. We begin by showing how to possibly control the uniform distance of two vectors by means of the discrete functional  $F_n^\theta$ .

**LEMMA 6.6** (Uniform distance control via  $F_n^\theta$ ). *Let the vectors  $u = (u^0, \dots, u^{N_n})$ ,  $v = (v^0, \dots, v^{N_n}) \in Y^{N_n+1}$  be given. Then,*

$$\begin{aligned} &\eta(1-\eta) \max_{1 \leq i \leq N_n} \phi(u^i - v^i) \\ &\leq \eta F_n^\theta(u^0, \dots, u^{N_n}) + (1-\eta) F_n^\theta(v^0, \dots, v^{N_n}) \quad \forall \eta \in [0, 1]. \end{aligned}$$

*Proof.* This proof follows the same lines of that of Corollary 4.4. Let  $1 \leq i \leq N_n$  be fixed and define  $F_n^{\theta,i} : Y^{i+1} \rightarrow [0, \infty]$  as

$$\begin{aligned} F_n^{\theta,i}(y^0, \dots, y^i) &= \sum_{j=1}^i \tau_n^j L_n^{\theta,j}(y^j, y^{j-1}) + \chi(y^0 - y_0) \\ &= \sum_{j=1}^i \psi(y^j - y^{j-1}) + \psi^*(\ell(t_{n,\theta}^j) - Ay_\theta^j) - \langle \ell(t_{n,\theta}^j), y^j - y^{j-1} \rangle \\ &\quad + \phi(y^i) + (2\theta - 1) \sum_{j=1}^i \phi(y^j - y^{j-1}) + \phi(y_0) - \langle Ay^0, y_0 \rangle + |y^0 - y_0|^2. \end{aligned}$$

Then, clearly  $y = (y^0, \dots, y^i) \mapsto G_n^{\theta,i}(y) = F_n^{\theta,i}(y) - \phi(y^i)$  is convex. Hence, letting  $w = \eta u + (1 - \eta)v$  for some  $\eta \in [0, 1]$  we have that

$$\begin{aligned} 0 &\leq F_n^{\theta,i}(w) \\ &\leq \eta(G_n^{\theta,i}(u) + \phi(u^i)) + (1 - \eta)(G_n^{\theta,i}(v) + \phi(v^i)) - \eta(1 - \eta)\phi(u^i - v^i), \end{aligned}$$

whence the assertion follows.  $\square$

Again, note that the latter lemma controls the uniform norm of the distance only if the stronger (3.6) is required. The following corollary of Lemma 6.6 will be the starting point for some possible a posteriori error control procedure (see Subsection 6.6).

**COROLLARY 6.7** (Uniform distance from the minimizer). *Let  $F_n^\theta(y^0, \dots, y^{N_n}) = 0$ . Then*

$$\max_{1 \leq i \leq N_n} \phi(y^i - v^i) \leq F_n^\theta(v^0, \dots, v^{N_n}) \quad \forall (v^0, \dots, v^{N_n}) \in Y^{N_n+1}.$$

Moreover, we re-obtain a proof of the uniqueness of the solution of the  $\theta$ -method.

**COROLLARY 6.8** (Uniqueness of the minimizer). *Assume (3.6). Then, there exists at most one  $y = (y^0, \dots, y^{N_n})$  such that  $F_n^\theta(y) = 0$ .*

**6.5. The generalized  $\theta$ -method.** Although minimizers of  $F_n^\theta$  and solutions of the  $\theta$ -scheme (6.3) coincide, minimizing sequences of  $F_n^\theta$  need not solve (6.3). This extra freedom allows the minimization formulation to capture the convergence of some generalized  $\theta$ -method, where the relations in (6.3) are not solved exactly but rather approximated. Namely, we shall look for vectors  $u_n = (u_n^0, \dots, u_n^{N_n})$  such that

$$F_n^\theta(u_n) \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

instead of  $F_n^\theta(u_n) = 0$  for all  $n \in \mathbb{N}$ .

From the computational viewpoint, note that the  $\theta$ -scheme consists in solving  $N_n$  nonlinear equations in one unknown each while checking for stationarity for  $F_n^\theta$  implies the solution of a tridiagonal system of  $N_n + 1$  nonlinear equations with (up to) three unknowns each. This entails in particular that minimizing  $F_n^\theta$  instead of solving (6.3) could be of a scarce interest if one is merely concerned in reproducing the  $\theta$ -scheme with no error. On the other hand, the issue of solving up to some tolerance turns out to be particularly relevant whenever one is aimed at implementing

an optimization procedure for the solution of (6.3). Indeed, one should be prepared to run the algorithm (some descent method, say) until some given tolerance is reached.

Our starting point for a possible convergence analysis of the generalized  $\theta$ -method is the following classical error control result.

**THEOREM 6.9** (Mielke & Theil [32]). *Assume (3.6). Then,  $\hat{y}_n \rightarrow y$  uniformly and  $F(y) = 0$ . In particular,*

$$\max_{t \in [0, T]} |(\hat{y}_n - y)(t)| \leq C_e \sqrt{\tau_n} \quad (6.15)$$

where  $C_e$  depends only on data and is independent of  $n$ .

More precisely, in [32] solely the case of the Euler scheme  $\theta = 1$  is discussed. However, an easy adaptation of the argument entails the result for  $\theta \in [1/2, 1)$  as well.

By explicitly comparing the minimizing sequence  $u_n = (u_n^0, \dots, u_n^{N_n})$  with the corresponding solution  $(y_n^0, \dots, y_n^{N_n})$  of the  $\theta$ -method, we have the following.

**THEOREM 6.10** (Convergence for the generalized  $\theta$ -method). *Assume (3.6) and let  $F_n^\theta(u_n^0, \dots, u_n^{N_n}) \rightarrow 0$ . Then,  $\hat{u}_n \rightarrow y$  uniformly, where  $F(y) = 0$ . In particular,*

$$\max_{t \in [0, T]} |(\hat{u}_n - y)(t)| \leq C_e \sqrt{\tau_n} + \left( \frac{2}{\alpha} F_n^\theta(u_n^0, \dots, u_n^{N_n}) \right)^{1/2}. \quad (6.16)$$

*Proof.* We have that

$$\begin{aligned} \max_{t \in [0, T]} |(y - \hat{u}_n)(t)| &\leq \max_{t \in [0, T]} |(y - \hat{y}_n)(t)| + \max_{t \in [0, T]} |(\hat{y}_n - \hat{u}_n)(t)| \\ &\leq C_e \sqrt{\tau_n} + \max_{1 \leq i \leq N_n} |u_n^i - y_n^i| \\ &\leq C_e \sqrt{\tau_n} + \left( \frac{2}{\alpha} \max_{1 \leq i \leq N_n} \phi(u_n^i - y_n^i) \right)^{1/2} \end{aligned}$$

and we conclude by applying Corollary 6.7.  $\square$

**6.6. A posteriori error control.** Let us now exploit both Corollary 4.5 and Theorem 6.10 in order to provide some possible a posteriori estimates of the approximation error by means of solutions  $u_n$  of the generalized  $\theta$ -method described above.

**LEMMA 6.11** (A posteriori error control via  $F_n^\theta$ ). *Assume (3.6) and let  $F_n^\theta(u_n^0, \dots, u_n^{N_n}) \sim \tau_n^s$  for some  $s > 0$  and  $F(y) = 0$ . Then,*

$$\max_{t \in [0, T]} |(\hat{u}_n - y)(t)| \sim \tau_n^r \quad \text{where } 2r = \max\{1, s\}.$$

**LEMMA 6.12** (A posteriori error control via  $F$ ). *Assume (3.6) and let  $F(\hat{u}_n) \sim \tau_n^s$  for some  $s > 0$  and  $F(y) = 0$ . Then,  $\max_{t \in [0, T]} |(\hat{u}_n - y)(t)| \sim \tau_n^{s/2}$ .*

We are also in the position of proving the weak-star convergence of the time derivatives of solutions  $u_n$  of the generalized  $\theta$ -method by comparing them with the corresponding derivatives of the exact solution of the  $\theta$ -method.

**LEMMA 6.13** (Improved convergence for the generalized  $\theta$ -method). *Assume (3.6) and let  $F_n^\theta(u_n^0, \dots, u_n^{N_n}) \sim \tau_n^2$ . Then,  $\hat{u}_n$  is equibounded in  $W^{1, \infty}(0, T; Y)$ . In particular,  $\hat{u}_n \rightarrow y$  weakly star in  $W^{1, \infty}(0, T; Y)$ .*

*Proof.* Let  $(y_n^0, \dots, y_n^{N_n})$  be the solution of the  $\theta$ -scheme. By exploiting Lemma 6.7, we check that

$$\begin{aligned} |u_n^i - u_n^{i-1}| &\leq |u_n^i - y_n^i| + |y_n^i - y_n^{i-1}| + |y_n^{i-1} - u_n^{i-1}| \\ &\leq \tau_n^i \|\hat{y}_n\|_{L^\infty(0,T;Y)} + 2 \left( \frac{2}{\alpha} F_n^\theta(u_n^0, \dots, u_n^{N_n}) \right)^{1/2}. \end{aligned}$$

The uniform bound on  $\|\hat{u}_n\|_{W^{1,\infty}(0,T;Y)}$  follows by dividing the latter by  $\tau_n^i$ , taking the maximum as  $1 \leq i \leq N_n$ , and recalling Lemma 6.4.  $\square$

**6.7. Adaptivity.** Assuming (3.6), the above introduced a posteriori error estimators can be exploited in order to develop an adaptive strategy. In particular, the error control in the uniform norm up to a given tolerance  $\text{tol} > 0$

$$\max_{t \in [0,T]} |(y - \hat{y}_n)(t)| \leq \text{tol}$$

for some piecewise approximation  $\hat{y}_n$  with  $\chi(\hat{y}_n(0) - y_0) \leq \alpha \text{tol}^2/4$  can be inferred, for instance, by choosing time steps in such a way that

$$\int_{t_{n-1}^{i-1}}^{t_n^i} L(t, \hat{y}_n(t), \dot{\hat{y}}_n(t)) \leq \frac{\alpha \text{tol}^2}{4N_n}.$$

Namely, by uniformly distributing the error along the partition.

Alternatively, one could develop an adaptive strategy by considering just computed quantities at the discrete level by asking for

$$\tau_n^i L_n^{\theta,i}(y_n^i, y_n^{i-1}) \leq \frac{\alpha \text{tol}^2}{32N_n} \quad \text{for} \quad \tau_n \leq \frac{\text{tol}^2}{16C_e^2}$$

and exploiting Theorem 6.10.

**7. Space-time approximations.** Let us combine the results of the previous sections (and use the corresponding notation) in order to state and prove a result on the convergence of full space-time approximations. Our results have to be compared with the former convergence analysis by HAN & REDDY [13]. Our approach leads to a convergence proof with respect to a weaker topology. However, it is on the one hand slightly more general (some assumptions on the spaces and the functionals, see (H1)-(H2) [13, p. 264], are not required) and on the other hand has a much simpler proof.

**THEOREM 7.1** (Convergence of space-time approximations). *Assume (5.1)-(5.9),  $\theta \in [1/2, 1]$ , define  $L_{n,h}^{\theta,i}(y, z) : Y \times Y \rightarrow [0, \infty]$  as*

$$\begin{aligned} L_{n,h}^{\theta,i}(y, z) &= \psi_h \left( \frac{y-z}{\tau_n^i} \right) + \psi_h^* (\ell_h(t_{n,\theta}^i) - A_h(\theta y + (1-\theta)z)) \\ &\quad - \langle \ell_h(t_{n,\theta}^i) - A_h(\theta y + (1-\theta)z), \frac{y-z}{\tau_n^i} \rangle, \end{aligned}$$

where  $A_h = \partial\phi_h$ , and let the functionals  $F_{n,h}^\theta : Y^{N_n+1} \rightarrow [0, \infty]$  be defined as

$$F_{n,h}^\theta(y^0, \dots, y^{N_n}) = \sum_{i=1}^{N_n} \tau_n^i L_{n,h}^{\theta,i}(y^i, y^{i-1}) + \chi_h(y^0 - y_0),$$

where  $\chi_h(\cdot) = \phi_h(\cdot) + |\cdot|^2$  (note that  $D(F_{n,h}^\theta) \subset Y_h^{N_n+1}$ ). Finally, let  $F_{n,h}(y_h^0, \dots, y_h^{N_n}) = 0$ . We have:

- (a)  $\hat{y}_{n,h} \rightarrow y_h$  weakly star in  $W^{1,\infty}(0, T; Y)$  as  $(n, h) \rightarrow (\infty, h)$  and  $F_h(y_h) = 0$ .
- (b)  $\hat{y}_{n,h} \rightarrow \hat{y}_n$  weakly star in  $W^{1,\infty}(0, T; Y)$  as  $(n, h) \rightarrow (n, 0)$  and  $F_n^\theta(y_n) = 0$ .
- (c)  $\hat{y}_n \rightarrow y$  weakly star in  $W^{1,\infty}(0, T; Y)$  as  $(n, 0) \rightarrow (\infty, 0)$  and  $F(y) = 0$ .
- (d)  $y_h \rightarrow y$  weakly star in  $W^{1,\infty}(0, T; Y)$  as  $(\infty, h) \rightarrow (\infty, 0)$  and  $F(y) = 0$ .
- (e)  $\hat{y}_{n,h} \rightarrow y$  weakly star in  $W^{1,\infty}(0, T; Y)$  as  $(n, h) \rightarrow (\infty, 0)$  and  $F(y) = 0$ .

The thesis of the Theorem is illustrated in Figure 1. In particular, we aim at showing that the space (or data) and time-limit can be taken in any order. Note that Limit (c) has been already checked in Theorem 6.5 and that the very same argument yields Limit (a) as well (recall that  $Y_h$  is closed). Moreover, Limit (d) is discussed in Theorem 5.3. So what we are actually left to check are Limits (b) and (e) only.

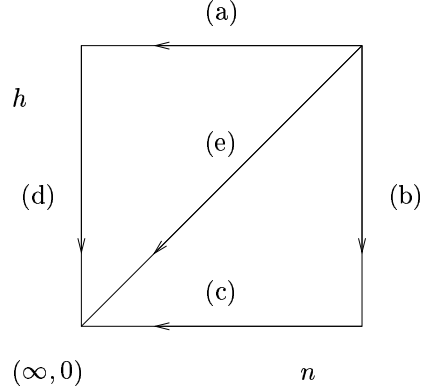


Fig. 7.1: Convergences for space-time approximations (see Theorem 7.1)

*Proof.* Limit (b). The assertion follows once we check that, for all  $i = 1, \dots, N_n$ , if  $y_{n,h}^{i-1} \rightarrow y_n^{i-1}$  weakly in  $Y$ , one has the weak convergence  $y_{n,h}^i \rightarrow y_n^i$  as well. Recall that

$$\begin{aligned} y_{n,h}^i &\in \operatorname{Argmin}_{y \in Y} \left( \theta \phi_h(y) - \langle \ell_h(t_{n,\theta}^i) - (1-\theta)A_h y_{n,h}^{i-1}, y \rangle + \psi_h(y - y_{n,h}^{i-1}) \right) \\ &= \operatorname{Argmin}_{y \in Y_h} \left( \theta \phi(y) - \langle \ell_h(t_{n,\theta}^i) - (1-\theta)A y_{n,h}^{i-1}, y \rangle + \psi_h(y - y_{n,h}^{i-1}) \right). \end{aligned}$$

Hence, since we have (5.6), the sequence  $y_{n,h}^i$  is weakly precompact and, up to the extraction of a (not relabeled) subsequence,  $y_{n,h}^i \rightarrow \tilde{y}$  weakly in  $Y$ . Let us prove that  $\tilde{y}$  solves the incremental problem (6.4). Indeed, we have that

$$\begin{aligned} 0 &\leq L_n^{\theta,i}(\tilde{y}, y_n^{i-1}) \\ &\leq \liminf_{h \rightarrow 0} \left( \psi_h(y_{n,h}^i - y_{n,h}^{i-1}) + \psi_h^*(\ell_h(t_{n,\theta}^i) - A y_{n,h,\theta}^i) \right. \\ &\quad \left. - \langle \ell_h(t_{n,\theta}^i) - A y_{n,h,\theta}^i, y_{n,h}^i - y_{n,h}^{i-1} \rangle \right) \\ &= \liminf_{h \rightarrow 0} L_{n,h}^{\theta,i}(y_{n,h}^i, y_{n,h}^{i-1}) = 0 \end{aligned}$$

where we have used the Mosco convergence in (5.5) and the pointwise convergence of  $\ell_h$  (5.7). Since the only solution of (6.4) is  $y_n^i$ , we have that  $\tilde{y} = y_n^i$  and the whole sequence converges.

Let us mention that, if the functionals  $\psi_h$  are uniformly linearly bounded (which is quite common in practice), then one could prove the latter convergence to be actually

strong: namely  $y_{n,h}^{i-1} \rightarrow y_n^{i-1}$  strongly in  $Y$  implies the strong convergence  $y_{n,h}^i \rightarrow y_n^i$ . Indeed, let  $w_h$  and  $\tilde{w}_h$  be such that  $w_h - y_{n,h}^i \rightarrow 0$  strongly in  $Y$ ,  $\psi_h(w_h - y_{n,h}^i) \rightarrow 0$ ,  $\tilde{w}_h \in Y_h$  and  $\tilde{w}_h - w_h \rightarrow 0$  strongly in  $Y$ . Then

$$\begin{aligned} & \theta \phi(y_{n,h}^i) - \langle \ell(t_{n,\theta}^i) - (1-\theta)A_h y_{n,h}^{i-1}, y_{n,h}^i \rangle + \psi_h(y_{n,h}^i - y_{n,h}^{i-1}) \\ & \leq \theta \phi_h(\tilde{w}_h) - \langle \ell(t_{n,\theta}^i) - (1-\theta)A_h y_{n,h}^{i-1}, \tilde{w}_h \rangle + \psi_h(\tilde{w}_h - w_h) + \psi_h(w_h - y_{n,h}^i). \end{aligned}$$

If  $\psi_h$  are uniformly linearly bounded above, then  $\psi_h(\tilde{w}_h - w_h) \rightarrow 0$  with  $h \rightarrow 0$ . Whence, by passing to the lim sup in the latter, we check that  $\limsup_{h \rightarrow 0} \phi(y_{n,h}^i) \leq \phi(y_n^i)$ , which, together with lower semicontinuity gives  $\phi(y_{n,h}^i) \rightarrow \phi(y_n^i)$  and the strong convergence follows from the reflexivity of  $Y$ .

Limit (e). Lemma 6.4, the uniform Lipschitz continuity of  $\ell_h$  (5.8), and the initial datum convergence (5.9), entail that  $\hat{y}_{n,h}$  are uniformly Lipschitz continuous as well. Hence, by extracting a (not relabeled) subsequence,  $\hat{y}_{n,h} \rightarrow y$  weakly star in  $W^{1,\infty}(0,T;Y)$ . In order to check that  $y$  solves (1.1), let us remark that, being  $\ell_{n,h,\theta}^i = \ell_h(t_{n,\theta}^i)$ ,

$$\bar{\ell}_{n,h,\theta} \rightarrow \ell \quad \text{strongly in } L^1(0,T;Y^*)$$

and that, by [45, Cor. 4.4]

$$\begin{aligned} \int_0^T \psi(y) & \leq \liminf_{h \rightarrow 0} \int_0^T \psi_h(\hat{y}_{n,h}) \\ \int_0^T \psi^*(\ell - Ay) & \leq \liminf_{h \rightarrow 0} \int_0^T \psi_h^*(\bar{\ell}_{n,h,\theta} - A\bar{y}_{n,h,\theta}), \end{aligned}$$

and compute that

$$\begin{aligned} 0 & \leq F(y) \\ & \leq \liminf_{h \rightarrow 0} \left( \int_0^T (\psi_h(\hat{y}_{n,h}) + \psi_h^*(\bar{\ell}_{n,h,\theta} - A\bar{y}_{n,h,\theta}) - \langle \bar{\ell}_{n,h,\theta}, \hat{y}_{n,h} \rangle) \right. \\ & \quad \left. + \phi(\hat{y}_{n,h}(T)) + \phi(y_{0,h}) - \langle A\hat{y}_{n,h}(0), y_{0,h} \rangle + |\hat{y}_{n,h}(0) - y_{0,h}|^2 \right) \\ & \leq \liminf_{h \rightarrow 0} F_{n,h}^\theta(y_{n,h}^0, \dots, y_{n,h}^{N_n}) = 0, \end{aligned}$$

and we have that  $F(y) = 0$ .  $\square$

We shall conclude by briefly mention some further results which can be obtained by suitably adapting to the current fully-discretized situation the arguments developed above for time-discretizations. Firstly, in the same spirit of Lemma 6.6, one could consider the possibility of estimating the distance of a vector from the minimizer of  $F_{n,h}^\theta$  by means of the functional itself. Secondly, the use of Corollary 4.5 would entail the possibility of an a posteriori error control an some adaptive strategy along the lines of Subsection 6.7 could be considered. Finally, by relying on the known convergence estimates for full space-time discretized problems [13] one could obtain a convergence and an a posteriori error control result for some generalized space-time approximated problem where  $F_{n,h}^\theta$  are not exactly minimized and one considers minimizing sequences instead (see Subsection 6.5). We shall develop these considerations elsewhere.

## REFERENCES

- [1] H. ATTOUCH, *Variational convergence for functions and operators*, Pitman (Advanced Publishing Program), Boston, MA, 1984.
- [2] G. AUCHMUTY, *Saddle-points and existence-uniqueness for evolution equations*, *Differential Integral Equations*, 6 (1993), pp. 1161–1171.
- [3] H. BREZIS AND I. EKELAND, *Un principe variationnel associé à certaines équations paraboliques. Le cas dépendant du temps*, *C. R. Acad. Sci. Paris Sér. A-B*, 282 (1976), pp. Ai, A1197–A1198.
- [4] ———, *Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps*, *C. R. Acad. Sci. Paris Sér. A-B*, 282 (1976), pp. Aii, A971–A974.
- [5] G. DAL MASO, *An introduction to  $\Gamma$ -convergence*, *Progress in Nonlinear Differential Equations and their Applications*, 8, Birkhäuser Boston Inc., Boston, MA, 1993.
- [6] G. DAL MASO, A. DESIMONE, AND M. G. MORA, *Quasistatic evolution problems for linearly elastic-perfectly plastic materials*, *Arch. Ration. Mech. Anal.*, 180 (2006), pp. 237–291.
- [7] G. DUVAUT AND J.-L. LIONS, *Inequalities in Mechanics and Physics*, Springer, Berlin, 1976.
- [8] M. A. EFENDIEV AND A. MIELKE, *On the rate-independent limit of systems with dry friction and small viscosity*, *J. Convex Anal.*, 13 (2006), pp. 151–167.
- [9] N. GHOUSOUB, *Selfdual partial differential systems and their variational principles*, Universitext, Springer-Verlag, In press, 2008.
- [10] N. GHOUSOUB AND L. TZOU, *A variational principle for gradient flows*, *Math. Ann.*, 330 (2004), pp. 519–549.
- [11] E. DE GIORGI AND T. FRANZONI, *Su un tipo di convergenza variazionale*, *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* (8), 58 (1975), pp. 842–850.
- [12] E. HAIRER, CH. LUBICH, AND G. WANNER, *Geometric numerical integration*, vol. 31 of *Springer Series in Computational Mathematics*, Springer-Verlag, Berlin, second ed., 2006. Structure-preserving algorithms for ordinary differential equations.
- [13] W. HAN AND B.D. REDDY, *Plasticity, Mathematical theory and numerical analysis*, Springer-Verlag, New York, 1999.
- [14] ———, *Convergence of approximations to the primal problem in plasticity under conditions of minimal regularity*, *Numer. Math.*, 87 (2000), pp. 283–315.
- [15] W. HAN AND B. D. REDDY, *Computational plasticity: the variational basis and numerical analysis*, *Comput. Mech. Adv.*, 2 (1995), pp. 283–400.
- [16] P. KREJČÍ, *Hysteresis, Convexity and Dissipation in Hyperbolic Equations*, vol. 8 of *GAKUTO Int. Series Math. Sci. Appl.*, Gakkotosho, Tokyo, 1996.
- [17] B. LEMAIRE, *An asymptotical variational principle associated with the steepest descent method for a convex function*, *J. Convex Anal.*, 3 (1996), pp. 63–70.
- [18] J. LEMAITRE AND J.-L. CHABOCHE, *Mechanics of solid materials*, Cambridge University Press, 1990.
- [19] A. LEW, J. E. MARSDEN, M. ORTIZ, AND M. WEST, *Asynchronous variational integrators*, *Arch. Ration. Mech. Anal.*, 167 (2003), pp. 85–146.
- [20] ———, *Variational time integrators*, *Internat. J. Numer. Methods Engrg.*, 60 (2004), pp. 153–212.
- [21] M. MABROUK, *Un principe variationnel pour une équation non linéaire du second ordre en temps*, *C. R. Acad. Sci. Paris Sér. I Math.*, 332 (2001), pp. 381–386.
- [22] ———, *A variational principle for a nonlinear differential equation of second order*, *Adv. in Appl. Math.*, 31 (2003), pp. 388–419.
- [23] F. MAGGI AND M. MORINI, *A  $\Gamma$ -convergence result for variational integrators of Lagrangians with quadratic growth*, *ESAIM Control Optim. Calc. Var.*, 10 (2004), pp. 656–665 (electronic).
- [24] J. E. MARSDEN, G. W. PATRICK, AND S. SHKOLLER, *Multisymplectic geometry, variational integrators, and nonlinear PDEs*, *Comm. Math. Phys.*, 199 (1998), pp. 351–395.
- [25] J. E. MARSDEN AND M. WEST, *Discrete mechanics and variational integrators*, *Acta Numer.*, 10 (2001), pp. 357–514.
- [26] A. MIELKE, *Evolution of rate-independent systems*, in *Handbook of Differential Equations, evolutionary equations*, C. Dafermos and E. Feireisl, eds., vol. 2, Elsevier, 2005, pp. 461–559.
- [27] A. MIELKE AND M. ORTIZ, *A class of minimum principles for characterizing the trajectories and the relaxation of dissipative systems*, *ESAIM Control Optim. Calc. Var.*, (to appear, 2008). Preprint available at: <http://www.wias-berlin.de/main/publications/wias-publ/index.cgi.en>.
- [28] A. MIELKE, R. ROSSI, AND G. SAVARÉ, *In preparation*, (2008).

- [29] ———, *A metric approach to a class of doubly nonlinear evolution equations and application*, Ann. Sc. Norm. Super. Pisa Cl. Sci., (to appear, 2008). Preprint available at: <http://www.wias-berlin.de/main/publications/wias-publ/index.cgi.en>.
- [30] A. MIELKE, T. ROUBÍČEK, AND U. STEFANELLI,  *$\Gamma$ -limits and relaxations for rate-independent evolutionary problems*, Calc. Var. Partial Differential Equations, (to appear, 2008).
- [31] A. MIELKE AND U. STEFANELLI, *A discrete variational principle for rate-independent evolution*, In preparation, (2008).
- [32] A. MIELKE AND F. THEIL, *On rate-independent hysteresis models*, NoDEA, Nonlinear Diff. Equations Applications, 11 (2004), pp. 151–189.
- [33] J.-J. MOREAU, *La notion de sur-potentiel et les liaisons unilatérales en élastostatique*, C. R. Acad. Sci. Paris Sér. A-B, 267 (1968), pp. A954–A957.
- [34] J.-J. MOREAU, *Sur les lois de frottement, de viscosité et plasticité*, C. R. Acad. Sci. Paris Sér. II Méc. Phys. Chim. Sci. Univers Sci. Terre, 271 (1970), pp. 608–611.
- [35] ———, *Sur l'évolution d'un système élasto-visco-plastique*, C. R. Acad. Sci. Paris Sér. A-B, 273 (1971), pp. A118–A121.
- [36] U. MOSCO, *Convergence of convex sets and of solutions of variational inequalities*, Advances in Math., 3 (1969), pp. 510–585.
- [37] S. MÜLLER AND M. ORTIZ, *On the  $\Gamma$ -convergence of discrete dynamics and variational integrators*, J. Nonlinear Sci., 14 (2004), pp. 279–296.
- [38] B. NAYROLES, *Deux théorèmes de minimum pour certains systèmes dissipatifs*, C. R. Acad. Sci. Paris Sér. A-B, 282 (1976), pp. Aiv, A1035–A1038.
- [39] ———, *Un théorème de minimum pour certains systèmes dissipatifs. Variante hilbertienne*, Travaux Sémin. Anal. Convexe, 6 (1976), p. 22.
- [40] H. RIOS, *Étude de la question d'existence pour certains problèmes d'évolution par minimisation d'une fonctionnelle convexe*, C. R. Acad. Sci. Paris Sér. A-B, 283 (1976), pp. Ai, A83–A86.
- [41] T. ROUBÍČEK, *Direct method for parabolic problems*, Adv. Math. Sci. Appl., 10 (2000), pp. 57–65.
- [42] J.C. SIMO AND T.J.R. HUGHES, *Computational inelasticity*, vol. 7 of Interdisciplinary Applied Mathematics, Springer-Verlag, New York, 1998.
- [43] U. STEFANELLI, *The discrete Brezis-Ekeland principle*, Preprint IMATI - CNR, 14PV07/14/0 (2007).
- [44] ———, *A variational principle in non-smooth mechanics*, in MFO Workshop: Analysis and numerics for rate-independent processes, vol. 4 of Oberwolfach Reports, 2007, pp. 1:622–624.
- [45] ———, *The Brezis-Ekeland principle for doubly nonlinear equations*, SIAM J. Control Optim., (to appear, 2008). Preprint available at: <http://www.imati.cnr.it/ulisse/pubbl.html>.
- [46] ———, *Some remarks on convergence and approximation for a class of hysteresis problems*, Istit. Lombardo Accad. Sci. Lett. Rend. A, (to appear, 2008). Preprint available at: <http://www.imati.cnr.it/ulisse/pubbl.html>.
- [47] R. TOADER AND C. ZANINI, *An artificial viscosity approach to quasistatic crack growth*, Submitted, (2007). Preprint available at <http://cvgmt.sns.it/>.
- [48] A. VISINTIN, *A new approach to evolution*, C. R. Acad. Sci. Paris Sér. I Math., 332 (2001), pp. 233–238.
- [49] C. ZANINI, *Singular perturbations of finite dimensional gradient flows*, Discrete Contin. Dyn. Syst., 18 (2007), pp. 657–675.