Journal of Convex Analysis
Volume 16 (2009), No. 1, 71–87

The Discrete Brezis-Ekeland Principle

Ulisse Stefanelli*
IMATI - CNR, v. Ferrata 1, 27100 Pavia, Italy
ulisse.stefanelli@imati.cnr.it, http://www.imati.cnr.it/ulisse/

Received: July 17, 2007
Revised manuscript received: March 14, 2008

We discuss a global-in-time variational approach to the time-discretization of gradient flows of convex functionals in Hilbert spaces. In particular, a discrete version of the celebrated Brezis-Ekeland variational principle is considered. The variational principle consists in the minimization of a functional on entire time-discrete trajectories. The latter functional admits a unique minimizer which solves the classical backward Euler scheme. This variational characterization is exploited in order to re-obtain in a variational fashion and partly extend the known convergence analysis for the Euler method. The relation between this variational technique and a posteriori error control and space approximation is also discussed.

Keywords: Gradient flow, Euler method, Brezis-Ekeland principle, convergence, error control

2000 Mathematics Subject Classification: 35K55

1. Introduction

This note is concerned with the classical *gradient flow*

$$u' + \partial\phi(u) \ni f \quad \text{a.e. in } (0,T), \; u(0) = u_0, \tag{1}$$

where the trajectory $t \in [0,T] \mapsto u(t)$ takes values in the Hilbert space H, the functional

$$\phi : H \to (-\infty, \infty] \text{ is convex, proper, and lower semicontinuous,}$$

the prime stands for time differentiation, and $\partial\phi$ is the subdifferential of ϕ in the sense of Convex Analysis. Moreover, we let $f \in L^2(0,T;H)$ and $u_0 \in D(\phi) := \{v \in H \;:\; \phi(v) \neq \infty\}$ (effective domain). Consequently, solutions u of (1) in $H^1(0,T;H)$ uniquely exist.

Gradient flows can be regarded as the paradigm of parabolic evolution. They arise almost ubiquitously in connection with applications to dissipative systems and have hence attracted constant attention during the last four decades. It is beyond our purposes to even attempt a review of the existing literature. Let us however mention that, even restricting to the present quite classical setting [5], the gradient flow (1) stems in a variety of different applications such as heat conduction, the Stefan problem, the Hele-Shaw cell,

*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.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

porous media, parabolic variational inequalities, some classes of ODEs with obstacles, some degenerate parabolic PDEs, the mean curvature flow, among others [26].

The gradient flow (1) is classically time-discretized by means of the so-called *backward Euler scheme*. Let $\{0 = t^0 < \cdots < t^N = T\}$ be a partition of the interval $[0, T]$ and $(f^1, \dots, f^N) \in H^N$ be a suitable approximation of f (local means on the partition, for instance), we say that the vector $(u^0, \dots, u^N) \in H^{N+1}$ solves the Euler scheme if

$$\delta u^i + \partial\phi(u^i) \ni f^i \quad \text{for } i = 1, \dots, N, \; u^0 = u_0, \tag{2}$$

where $\delta u^i = (u^i - u^{i-1})/\tau^i$ denotes the i-th difference quotient and $\tau^i = t^i - t^{i-1}$ is the i-th time-step.

We are interested in a global-in-time variational characterization of the solution of the Euler scheme (2). To this end, we shall introduce the functional $K_N : H^{N+1} \to [0, \infty]$ as

$$\begin{aligned} K_N(u^0, \dots, u^N) = {} & \left(\sum_{i=1}^N \tau^i \left(|\delta u^i|^2 - \left(f^i, \delta u^i\right) \right) + \phi(u^N) - \phi(u_0) \right)^+ \\ & + \sum_{i=1}^N \tau^i \left(\phi(u^i) + \phi^*\left(f^i - \delta u^i\right) - (f^i, u^i) + \frac{\tau^i}{2}|\delta u^i|^2 \right) \\ & + \frac{1}{2}|u^N|^2 - \frac{1}{2}|u^0|^2 + \frac{1}{2}|u^0 - u_0|^2. \end{aligned}$$

Here $(\cdot, \cdot)$ is the scalar product in H, $|\cdot|$ is the corresponding norm, and we have denoted by ϕ^* the conjugate of ϕ, i.e. $\phi^*(w) := \sup\{(w, u) - \phi(u), \; u \in H\}$ for all $w \in H$ and by $(\cdot)^+$ the standard positive part. Note that K_N is convex and coercive with respect to the weak topology in H^{N+1} (hence lower semicontinuous).

The starting point of our analysis relies on the observation that K_N admits a unique minimizer which in turn solves the Euler scheme (see Lemma 2.2).

(Characterization) $(u^0, \dots, u^N)$ solves the Euler scheme (2) if and only if $K_N(u^0, \dots, u^N) = \min K_N = 0$. (3)

The aim of this note is to recast the convergence analysis for the Euler scheme via this variational characterization. This procedure leads to new variational proofs of known facts as well as some interesting novel result. We shall remark that characterization (3) holds indeed in a quantitative form as the value of the functional $K_N(u^0, \dots, u^N)$ is controlling the distance of the vector $(u^0, \dots, u^N)$ from the solution of the Euler scheme (2).

An even simpler functional $J_N : H^{N+1} \to [0, \infty]$ featuring the same characterization above is

$$\begin{aligned} J_N(u^0, \dots, u^N) = {} & \sum_{i=1}^N \tau^i \left(\phi(u^i) + \phi^*\left(f^i - \delta u^i\right) - (f^i, u^i) + \frac{\tau^i}{2}|\delta u^i|^2 \right) \\ & + \frac{1}{2}|u^N|^2 - \frac{1}{2}|u^0|^2 + \frac{1}{2}|u^0 - u_0|^2. \end{aligned}$$

In particular, the unique minimizer of J_N attains the value 0 and solves the Euler scheme. The characterization of solution to the Euler scheme as the minimizer of J_N was already exploited by *Lemaire* [15] for discussing the asymptotic behavior for large times $T \to \infty$ of the iterative solution of (2) (namely, τ^i = constant and $N \to \infty$). We shall focus here on the convergence of the discrete scheme to the continuous problem by considering $\tau \to 0$ instead. To this aim, some uniform-in-N coercivity will be needed. More precisely, by letting $\widehat{u}_N$ denote the piecewise affine interpolant of the vector $(u_N^0, \ldots, u_N^N)$ on the partition, we will ask that minimizing sequences as $N \to \infty$ are such that the corresponding interpolants $\widehat{u}_N$ are weakly pre-compact in $H^1(0,T;H)$. Now, this is not generally the case for J_N unless ϕ is quadratically bounded (namely, $\phi(u) \leq c|u|^2 + c$ for some $c > 0$), a quite restrictive assumption with respect to applications (especially to PDEs). On the other hand, the latter equi-coercivity for $\widehat{u}_N$ is fulfilled in the case of K_N due to the presence of the positive part term.

Our convergence result reads as follows.

> **(Convergence)** Let $K_N(u_N^0, \ldots, u_N^N) \to 0$ as $N \to \infty$.
> Then, the corresponding piecewise interpolants $\widehat{u}_N$ converge strongly in $H^1(0,T;H)$ to the solution u of the gradient flow (1). (4)

Notice that, by choosing $K_N(u_N) = 0$, we re-obtain the convergence for the solutions of Euler scheme.

A first novelty of this variational approach consists in the fact that the convergence result (4) is not restricted to the special case of the Euler scheme (2) but can be applied to some more general situations as well. Although minimizers of (3) and solutions to (2) coincide, minimizing sequences as $N \to \infty$ need not solve the Euler scheme. Hence, the variational formulation via K_N may be regarded as a tool for discussing convergence for generic approximating time-discrete trajectories, possibly not related to the Euler scheme (or where the relations in (2) are not solved exactly but rather approximately).

A second interesting feature of the functional K_N is that, as already mentioned, it controls the uniform distance from the solution to the Euler scheme. Hence, K_N may be used in order to implement some a posteriori error control procedure, possibly aimed at adaptivity (see Subsection 5.4).

Let us however state from the very beginning that our interest for the minimization problem in (3) is purely theoretical. Indeed, apart from the already mentioned specific points above, we do not believe that minimizing K_N could represent a computationally valid alternative to the sequential solution of the Euler scheme. First of all, the Euler scheme consists in solving N nonlinear equations in one unknown while checking stationarity for K_N implies the solution of a system of $N+1$ nonlinear equations with up to three unknowns each. Secondly, the formulation of K_N requires the knowledge of ϕ^* whereas the Euler scheme (2) and its sequential variational formulation (7) below are formulated in terms of ϕ and $\partial\phi$ only.

2. The discrete Brezis-Ekeland principle

The characterization (3) corresponds to a discrete version of the variational principle formulated by *Brezis & Ekeland* [7, 6] and *Nayroles* [23, 24] (see also [3, Sect. 3.4]). The latter characterizes solutions u of the gradient flow (1) as minimizers of the functional $J : H^1(0,T;H) \to [0,\infty]$ defined as

$$J(u) = \int_0^T \big(\phi(u) + \phi^*(f - u') - (f,u)\big) + \frac{1}{2}|u(T)|^2 - \frac{1}{2}|u_0|^2 + \frac{1}{2}|u(0) - u_0|^2.$$

Let us stress that $\phi(u) + \phi^*(w) \geq (w,u)$ for all u, $w \in H$, and that the equality holds iff $w \in \partial\phi(u)$. In particular, one readily checks that $J(v) \geq 0$ for all $v \in H^1(0,T;H)$ and that $J(u) = \min J = 0$ iff u solves (1).

The functional J is convex and lower semicontinuous with respect to the weak topology of $H^1(0,T;H)$. It may however fail to be coercive with respect to the latter. More specifically, minimizing sequences of J need not be weakly pre-compact in $H^1(0,T;H)$ (take $H = \mathbb{R}$, $\phi^*(\cdot) = |\cdot|$, $f = 0$, $u_0 = 0$, and $u_n(t) = \min\{n^{2/3}t, n^{-1/3}\}$). On the other hand, note that J is coercive if ϕ^* has at least a quadratic growth, namely if ϕ is quadratically bounded.

In order to gain coercivity it is worth introducing the functional $K : H^1(0,T;H) \to [0,\infty]$ given by

$$K(u) = \left(\int_0^T \big(|u'|^2 - (f,u')\big) + \phi(u(T)) - \phi(u_0)\right)^+ + J(u).$$

Indeed, the latter is coercive with respect to the weak topology of $H^1(0,T;H)$, convex, and lower semicontinuous. Moreover, if u solves (1) then the chain rule [5, Lemme 3.3, p. 73] entails that

$$\phi(u(T)) - \phi(u_0) = \int_0^T \frac{\mathrm{d}}{\mathrm{dt}}\phi(u) = \int_0^T (f - u', u').$$

In particular, we have that $K(u) = 0$ so that the Brezis-Ekeland characterization can be restated as follows.

Theorem 2.1 (Brezis-Ekeland, 1976). *u solves* (1) *iff* $K(u) = \min K = 0$.

The formulation of the functional K_N is directly tailored for a discrete version of the latter variational principle. The key point is that, letting $(u^0,\dots,u^N)$ be given and the functionals $L^i : H \times H \to [0,\infty]$ be defined as

$$L^i(u,v) = \phi(u) + \phi^*\left(f^i - \frac{u-v}{\tau^i}\right) - \left(f^i - \frac{u-v}{\tau^i}, u\right) \quad \text{for } i = 1,\dots,N,$$

the Euler scheme (2) can be rewritten as

$$L^i(u^i, u^{i-1}) = \min L^i = 0 \quad \text{for } i = 1,\dots,N, \ u^0 = u_0. \tag{5}$$

Moreover, we readily check that

$$J_N(u^0,\dots,u^N) = \sum_{i=1}^N \tau^i L^i(u^i,u^{i-1}) + \frac{1}{2}|u^0 - u_0|^2, \tag{6}$$

so that clearly $J_N \geq 0$ and $J_N(u^0,\dots,u^N) = 0$ iff $(u^0,\dots,u^N)$ solves Euler. A fortiori, since $K_N \geq J_N$, if $K_N(u^0,\dots,u^N) = 0$, then $(u^0,\dots,u^N)$ solves Euler. Moreover, the converse implication holds as well (see Subsection 3.1 below) and we have the following.

Lemma 2.2. $(u^0,\dots,u^N)$ *solves the Euler scheme* (2) *iff*

$$K_N(u^0,\dots,u^N) = \min K_N = 0.$$

The minimum problems of (5) have to be compared to the variational formulation of the Euler scheme, namely

$$u^0 = u_0 \quad \text{and} \quad G^i(u^i,u^{i-1}) = \min G^i(\cdot,u^{i-1}) \ \text{ for } i=1,\dots,N,$$
$$\text{where } G^i(u,v) = \frac{|u-v|^2}{2\tau^i} + \phi(u). \tag{7}$$

Although L^i and G^i admit the same minimizer, the advantage in dealing with L^i is patent since we also have the property $\min L^i = 0$. This is really the key point for the construction of a global minimization formulation of Euler (see (6)). Note that no such construction is currently available starting from G^i. A global minimization procedure based of the functionals G^i would require the use of weights encoding the fact that the minimization of G^1 has to be accorded a disproportional higher priority with respect to the minimization of G^2, the minimization of G^2 with respect to that of G^3 and so on. The reader is referred to *Mielke & Ortiz* [18] and *Mielke & Stefanelli* [19, 20] for some detail in this direction.

Before closing this section, let us mention that 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 [14] and the survey [17], we shall restrain here from giving a detailed presentation of the subject and limit ourselves in mentioning that some Γ-convergence techniques have been recently exploited in the (finite-dimensional) frame of variational integrators by *Müller & Ortiz* [22] (see also [16]).

2.1. Existence via the Brezis-Ekeland principle

The functional K is coercive and lower semicontinuous with respect to the weak topology of $H^1(0,T;H)$. Hence, one is tempted to exploit the Brezis-Ekeland characterization of Theorem 2.1 in order to obtain solutions to (1) via the Direct Method. This strategy has however proved itself to be much more involved with respect to classical maximal monotone techniques (see [5]). The difficulty arises from the fact that one is not just asked to minimize K but also to prove that the minimum is 0.

Conditional existence results for (1) by means of the Direct Method have been firstly obtained by *Rios* [28, 31] (see also [29, 30]). Later on, *Auchmuty* [4] proved that, in the controlled-growth case, the minimum problem can be reformulated as a saddle point problem for which the minimax value 0 is achieved (see also [2]). Again in the controlled-growth case, *Roubíček* [32] checked that the optimality conditions imply (1) (see also the recent monograph [33, Sec. 8.10]). Finally, the full extent of maximal monotone methods has been recovered via the Brezis-Ekeland approach by *Ghoussoub & Tzou* [13]. In the latter paper, the authors eventually overcome the controlled-growth assumption within the far-reaching theory of (anti)selfdual Lagrangians by *Ghoussoub* (see the recent monograph [10]). We shall mention some further results by *Ghoussoub & McCann* [12] for quadratic perturbations of convex functionals, the analysis of the long-time dynamics of autonomous gradient flows by *Lemaire* [15], and the extension of the Brezis-Ekeland principle to doubly nonlinear [38] and rate-independent situations [36, 37, 39] by this author.

3. Characterization and well-posedness

3.1. The discrete variational principle

Let us provide a proof of Lemma 2.2. For the sake of later purposes we define the functionals $L^i : H \times H \to [0, \infty]$ and $Z^i : D(\phi) \times H \to [-\infty, \infty)$ as

$$L^i(u, v) = \phi(u) + \phi^* \left(f^i - \frac{u - v}{\tau^i} \right) - \left(f^i - \frac{u - v}{\tau^i}, u \right) \quad i = 1, \ldots, N,$$
$$Z^i(u, v) = \left(\frac{u - v}{\tau^i} - f^i, \frac{u - v}{\tau^i} \right) + \frac{\phi(u) - \phi(v)}{\tau^i} \quad i = 2, \ldots, N,$$
$$Z^1(u, v) = \left(\frac{u - v}{\tau^1} - f^1, \frac{u - v}{\tau^1} \right) + \frac{\phi(u) - \phi(u_0)}{\tau^1}.$$

We start by recalling that the Euler scheme (2) can be rewritten as

$$L^i(u^i, u^{i-1}) = \min L^i(\cdot, u^{i-1}) = 0 \;\; \text{for } i = 1, \ldots, N, \;\; u^0 = u_0. \tag{8}$$

We have that

$$K_N(u^0, \ldots, u^N) = \left(\sum_{i=1}^N \tau^i Z^i(u^i, u^{i-1}) \right)^+ + \sum_{i=1}^N \tau^i L^i(u^i, u^{i-1}) + \frac{1}{2}|u^0 - u_0|^2, \tag{9}$$

Owing to (9), we readily check that $K_N(u^0, \ldots, u^N) = 0$ yields that $(u^0, \ldots, u^N)$ solves (8) (hence Euler).

Viceversa, let $(u^0, \ldots, u^N)$ solve (8). In particular $u^0 = u_0$. By exploiting Fenchel's duality, we readily get that, $\forall u, v \in D(\phi)$, $i = 2, \ldots, N$,

$$\begin{aligned} \tau^i Z^i(u, v) = & - \left(f^i - \frac{u - v}{\tau^i}, u \right) + \left(f^i - \frac{u - v}{\tau^i}, v \right) + \phi(u) - \phi(v) \\ \leq & - \left(f^i - \frac{u - v}{\tau^i}, u \right) + \phi^* \left(f^i - \frac{u - v}{\tau^i} \right) + \phi(v) + \phi(u) - \phi(v) \\ = & \; L^i(u, v). \end{aligned} \tag{10}$$

A similar calculation yields that

$$\tau^1 Z^1(u, v) \leq L^1(u, v) + \phi(v) - \phi(u_0).$$

Owing to (8) and (9) the assertion of Lemma 2.2 follows.

3.2. Well-posedness by a purely variational method

Relying on the well-posedness of the Euler scheme (2), Lemma 2.2 ensures that the minimum problem in (3) is also well-posed and that the corresponding value of the functional is 0. We shall stress that these facts can be checked by purely variational methods, that is, without exploiting the well-posedness theory for the Euler scheme (2).

First of all, the functional K_N is lower semicontinuous on H^{N+1} and has bounded sublevels (the latter follows at once by arguing iteratively for the component u^i as $i = 0, 1, \ldots, N$ increases). Hence, the minimum problem (3) has a solution.

We shall now exploit [9, Thm. 4.1] in order to have that

$$\forall v \in H \quad \min L^i(\cdot, v) = 0. \tag{11}$$

This is of course equivalent to the existence of a solution of the Euler scheme. Nevertheless, the proof of [9, Thm. 4.1] relies on duality theory instead of on that of maximal monotone operators (the two theories are actually one, see [11]). In particular, relation (6) and (11) entail that

$$\min J_N \leq \min \frac{1}{2}|u^0 - u_0|^2 = 0.$$

The same holds for K_N. By exploiting (10) for $i = 2, \ldots, N$ and (11) we get that

$$\min K_N \leq \min \left((\phi(u^0) - \phi(u_0))^+ + \frac{1}{2}|u^0 - u_0|^2 \right) = 0.$$

3.3. Uniqueness at the discrete level

We shall now conclude this discussion by checking uniqueness. To this aim, we perform the change of variables

$$(u^0, u^1, \ldots, u^N) \to (u^0, \delta u^1, \ldots, \delta u^N), \tag{12}$$

and compute that

$$\begin{aligned} & J_N(u^0, \delta u^1, \ldots, \delta u^N) \\ = & \sum_{i=1}^{N} \tau^i \left(\phi \left(u^0 + \sum_{j=1}^{i} \tau^j \delta u^j \right) + \phi^*(f^i - \delta u^i) - \left(f^i, u^0 + \sum_{j=1}^{i} \tau^j \delta u^j \right) \right) \\ & \quad + \frac{1}{2} \left| u^0 + \sum_{i=1}^{N} \tau^i \delta u^i \right|^2 + \frac{1}{2} \sum_{i=1}^{N} (\tau^i)^2 |\delta u^i|^2 - \frac{1}{2}|u^0|^2 + \frac{1}{2}|u^0 - u_0|^2. \end{aligned}$$

Hence, in these new variables the functional J_N (and K_N as well) is uniformly convex and the minimizer is unique. Let us recall that, given a normed space E, a functional $\psi : E \to (-\infty, \infty]$ is said to be *uniformly convex* if it exists $\lambda > 0$ such that

$$\psi(\theta u + (1-\theta)v) \leq \theta\phi(u) + (1-\theta)\phi(v) - \frac{\lambda}{2}\theta(1-\theta)\|u - v\|_E^2 \quad \forall u, \, v \in E, \ \theta \in [0, 1].$$

In particular, in case ψ is two-times differentiable, we have $\mathrm{D}^2\psi(u)[v, v] \geq \lambda \|v\|_E^2$, with obvious notation.

4. Convergence

We shall now turn to the convergence analysis. In the following we make an extensive use of the following notation: letting $v = (v^0, \dots, v^N)$ be a vector, we will denote by $\widehat{v}$ and $\overline{v}$ two functions of the time interval $[0, T]$ which interpolate the values of the vector v piecewise linearly and backward constantly on the partition, respectively. Namely

$$\widehat{v}(0) = v^0, \quad \widehat{v}(t) = \gamma^i(t) v^i + \big(1 - \gamma^i(t)\big) v^{i-1},$$
$$\overline{v}_n(0) = v^0, \quad \overline{v}_n(t) = v^i, \quad \text{for } t \in (t^{i-1}, t^i], \ i = 1, \dots, N$$

where

$$\gamma^i(t) = (t - t^{i-1})/\tau^i \ \text{ for } t \in (t^{i-1}, t^i], \ i = 1, \dots, N.$$

Moreover, we recall that $\delta v^i = (v^i - v^{i-1})/\tau^i$ for $i = 1, \dots, N$ (so that $\widehat{v}' = \overline{\delta v}$).

Assume now we are given the partitions $\{0 = t_n^0 < t_n^1 < \cdots < t^{N_n} = T\}$ and denote by $\tau_n^i = t_n^i - t_n^{i-1}$ the i-th time-step and by $\tau_n = \max_{1 \le i \le N_n} \tau_n^i$ the diameter of the n-th partition. No constraints are imposed on the possible choice of the time-steps throughout this analysis besides $\tau_n \to 0$ as $n \to \infty$. Moreover, let some approximations $(f_n^1, \dots, f_n^{N_n})$ of the function f be given. The corresponding Euler schemes read

$$\frac{u_n^i - u_n^{i-1}}{\tau_n^i} + \partial\phi(u_n^i) \ni f_n^i \ \text{ for } i = 1, \dots, N_n, \ u_n^0 = u_0, \tag{13}$$

The functionals $K_n : H^{N_n+1} \to [0, \infty]$ are defined as

$$\begin{aligned} K_n(u^0, \dots, u^{N_n}) = &\left(\sum_{i=1}^{N_n} \tau_n^i \left(|\delta u^i|^2 - \big(f_n^i, \delta u^i\big) \right) + \phi(u^{N_n}) - \phi(u_0) \right)^+ \\ &+ \sum_{i=1}^{N_n} \tau_n^i \left(\phi(u^i) + \phi^* \left(f_n^i - \delta u^i\right) - (f_n^i, u^i) + \frac{\tau_n^i}{2} |\delta u^i|^2 \right) \\ &+ \frac{1}{2} |u^{N_n}|^2 - \frac{1}{2} |u^0|^2 + \frac{1}{2} |u^0 - u_0|^2. \end{aligned}$$

We shall prove that minimizing sequences of K_n are converging to solutions of the gradient flow (1) as $n \to \infty$.

Theorem 4.1 (Convergence). *Assume that $\overline{f}_n \to f$ strongly in $L^2(0, T; H)$ and $K_n(u_n) \to 0$. Then, $\widehat{u}_n \to u$ strongly in $H^1(0, T; H)$ where u solves the gradient flow* (1).

We shall explicitly mention that the above convergence requirement on $\overline{f}_n$ is fairly standard and is met, for instance, by choosing

$$f_n^i = \frac{1}{\tau_n^i} \int_{t_n^{i-1}}^{t_n^i} f \ \text{ for } i = 1, \dots, N_n. \tag{14}$$

Proof of Theorem 4.1. By exploiting the above introduced notation, we have that

$$K_n(u_n) \geq \left(\int_0^T \left(|\widehat{u}_n'|^2 - (\overline{f}_n, \widehat{u}_n') \right) + \phi(\widehat{u}_n(T)) - \phi(u_0) \right)^+$$
$$+ \int_0^T \left(\phi(\overline{u}_n) + \phi^*(\overline{f}_n - \widehat{u}_n') - (\overline{f}_n, \overline{u}_n) \right)$$
$$+ \frac{1}{2}|\widehat{u}_n(T)|^2 + \frac{1}{2}|\widehat{u}_n(0)|^2 + \frac{1}{2}|u_0|^2 - (\widehat{u}_n(0), u_0). \tag{15}$$

Hence, if $\sup_n K_n(u_n) < \infty$ then $\widehat{u}_n$ is bounded in $H^1(0,T;H)$. By suitably extracting some (not relabeled) subsequence we have that $\widehat{u}_n \to u$ weakly in $H^1(0,T;H)$. In particular $\widehat{u}_n(t) \to u(t)$ weakly for all $t \in [0,T]$. Moreover $\overline{u}_n \to u$ weakly in $L^2(0,T;H)$ as well since

$$\|\widehat{u}_n - \overline{u}_n\|^2_{L^2(0,T;H)} = \sum_{i=1}^{N_n} \int_{t_n^{i-1}}^{t_n^i} (1 - \gamma_n^i(t))^2 \|u_n^i - u_n^{i-1}\|^2 \leq \frac{\tau_n^2}{3} \|\widehat{u}_n'\|^2_{L^2(0,T;H)}.$$

Finally, we can pass to the lim inf in (15) by lower semicontinuity and check that

$$K(u) \leq \liminf_{n\to\infty} K_n(u_n) = 0.$$

Namely, owing to the characterization of Theorem 2.1, u solves the gradient flow (1). Hence, it is unique and the whole sequence $\widehat{u}_n$ converges.

Let us now move to the proof of the strong convergence. Since

$$\int_0^T \left(|\widehat{u}_n'|^2 - (\overline{f}_n, \widehat{u}_n') \right) + \phi(\widehat{u}_n(T)) - \phi(u_0) \leq K_n(u_n),$$

by passing to the lim sup as $n \to \infty$ we have that

$$\limsup_{n\to\infty} \int_0^T |\widehat{u}_n'|^2 \leq \limsup_{n\to\infty} K_n(u_n) - \liminf_{n\to\infty} \phi(\widehat{u}_n(T)) + \phi(u_0) + \limsup_{n\to\infty} \int_0^T (\overline{f}_n, \widehat{u}_n')$$
$$\leq -\phi(u(T)) + \phi(u_0) + \int_0^T (f, u') = \int_0^T |u'|^2.$$

Hence, the strong convergence $\widehat{u}_n' \to u'$ in $L^2(0,T;H)$ follows. □

As already mentioned, the weak pre-compactness in $H^1(0,T;H)$ of minimizing sequences for J_n does not generally hold (unless $J_n(u_n) = 0$, of course). By explicitly requiring it, we can however rephrase the above convergence result as follows.

Theorem 4.2 (Convergence for J_n). *Assume that $\overline{f}_n \to f$ strongly in $L^2(0,T;H)$, $J_n(u_n) \to 0$, and $\widehat{u}_n$ are weakly pre-compact in $H^1(0,T;H)$. Then, $\widehat{u}_n \to u$ weakly in $H^1(0,T;H)$ where u solves the gradient flow* (1).

Note that, differently from the case of Theorem 4.1, no strong convergence is here inferred (but see Subsection 5.2 below).

4.1. Existence via the Brezis-Ekeland principle

We shall specifically remark that, by exploiting the above-introduced time-discretization, we are in position of re-obtaining the existence of solutions to the gradient flow (1) by a variational technique. Let us state this fact here for the sake of clarity.

Lemma 4.3. *Under the above assumptions there exist a (unique) solution of the gradient flow* (1)*. The latter is the weak limit in* $H^1(0,T;H)$ *of minimizing sequences of* K_n *as* $n \to \infty$.

Since we have already checked in Subsequence 3.2 that a minimizing sequence of K_n necessarily exists, the lemma follows directly from Theorem 4.1.

5. Error control and strong convergence

5.1. Error with respect to Euler

The functional K_N can be used in order to estimate the distance of vectors. Let us start with the following.

Lemma 5.1 (J_N controls the distance). *Let* u, $v \in H^{N+1}$. *Then*

$$\frac{\eta(1-\eta)}{2} \max_{i=1,\ldots,N} |u^i - v^i|^2 \leq \eta J_N(u) + (1-\eta) J_N(v) \quad \forall \eta \in [0,1]. \tag{16}$$

Proof. Define, $J^i : H^{N+1} \to [0,\infty]$ as

$$J^i(u) = \sum_{j=1}^{i} \tau^j L^j(u^j, u^{j-1}) + \frac{1}{2}|u^0 - u_0|^2$$

for $i = 0, 1, \ldots, N$. It is straightforward to check that, for all i,

$$u \mapsto J^i(u) = C^i(u) + \frac{1}{2}|u^i|^2$$

where C^i is convex. Hence, for all u, $v \in H^{N+1}$ and $\eta \in [0,1]$, one has that

$$0 \leq J^i(\eta u + (1-\eta)v) \leq \eta J^i(u) + (1-\eta) J^i(v) - \frac{\eta(1-\eta)}{2}|u^i - v^i|^2$$

and the assertion follows. □

In particular, letting $e = (e^0, \ldots, e^N) \in H^{N+1}$ be the solution to (2) and choosing $v = e$, we have following.

Corollary 5.2 (Uniform error with respect to Euler).

$$\frac{1}{2} \max_{i=1,\ldots,N} |u^i - e^i|^2 \leq J_N(u) \quad \forall u \in H^{N+1}. \tag{17}$$

Namely, we have obtained that

$$\|\widehat{u} - \widehat{e}\|_{C([0,T];H)} \le (2J_N(u))^{1/2}. \tag{18}$$

Besides the latter error control with respect to Euler, a global estimate can be obtained as well by exploiting the (global) uniform convexity of J_N with respect to the variables $(u^0, \delta u^1, \dots, \delta u^N)$ (see (12)). In particular, as $J_N(e) = 0$, we obtain that

$$\frac{1}{2}|u^0 - u_0|^2 + \sum_{i=1}^{N} \frac{(\tau^i)^2}{2} |\delta u^i - \delta e^i|^2 \le J_N(u). \tag{19}$$

The latter can be exploited in order to re-obtain the uniform error control of (18). Moreover, it yields

$$\|\widehat{u}' - \widehat{e}'\|^2_{L^2(0,T;H)} \le \frac{2J_N(u)}{\min \tau^i}. \tag{20}$$

as well.

5.2. Strong convergence

The estimates (18) and (20) entail the possibility to sharpen the convergence statement of Theorem 4.1. In particular, since we already know that the solution to the Euler scheme converges to the solution of the gradient flow strongly in $H^1(0,T;H)$ (see Theorem 4.1), we have the following.

Lemma 5.3 (Strong convergence). *Under the same assumptions of Theorem* 4.2*, the convergence $u_n \to u$ is strong in $C([0,T];H)$. Moreover, if*

$$J_n(u_n)/\min \tau_n^i \to 0,$$

then the convergence is strong in $H^1(0,T;H)$.

5.3. A posteriori error estimates

By relying on the theory of a posteriori error control for the Euler scheme by *Nochetto, Savaré & Verdi* [25, 26] (see also the extensions in [35, 27]), the above estimates can be exploited in order to provide an explicit a posteriori control of the discretization error. Within this subsection, u is the solution of the gradient flow (1), $\widehat{e}$ is the piecewise interpolant of the solution of the Euler scheme with $\widehat{e}(0) = u(0)$, and $\widehat{u}$ is the piecewise interpolant of the vector $(u^0, \dots, u^N)$. The result in [26, Thm. 3.2] entails that

$$\|u - \widehat{e}\|_{C([0,T];H)} \le \left(-\sum_{i=1}^{N} (\tau^i)^2 Z^i(e^i, e^{i-1}) \right)^{1/2} + \|f - \overline{f}\|_{L^1(0,T;H)}.$$

Note that, for all $i = 1, \dots, N$, one has that $0 = -L^i(e^i, e^{i-1})/\tau^i \le -Z^i(e^i, e^{i-1})$. By taking into account (18), we obtain the a posteriori error estimate

$$\|u - \widehat{u}\|_{C([0,T];H)} \le \left(-\sum_{i=1}^{N} (\tau^i)^2 Z^i(e^i, e^{i-1}) \right)^{1/2} + \|f - \overline{f}\|_{L^1(0,T;H)} + (2J_N(u))^{1/2}. \tag{21}$$

The latter can be used in order to provide explicit convergence rates. Henceforth, the symbol $\lesssim$ will denote an inequality up to a constant depending on data only.

In case $\phi \geq 0$ and f_n^i are defined by local means as in (14) one has that [26, Thm. 3.16]

$$\|u - \widehat{u}\|_{C([0,T];H)} \leq \tau^{1/2} \left(3\phi(u_0) + 3\|f\|^2_{L^2(0,T;H)}\right)^{1/2} + (2J_N(u))^{1/2}$$
$$\lesssim \tau^{1/2} + (J_N(u))^{1/2}. \tag{22}$$

In case $u_0 \in D(\partial\phi) = \{u \in H : \partial\phi(u) \neq \emptyset\}$, $f \in BV(0,T;H)$, and $f_n^i = \lim_{t \downarrow t_n^i} f(t)$, we have [26, Thm. 3.16]

$$\|u - \widehat{u}\|_{C([0,T];H)} \leq \tau \left(\frac{1}{\sqrt{2}}|(f_+(0) - \partial\phi(u_0))^\circ| + 2\mathrm{Var}\, f\right) + (2J_N(u))^{1/2}$$
$$\lesssim \tau + (J_N(u))^{1/2}, \tag{23}$$

where $(\cdot)^\circ$ denotes the element of minimal norm and

$$\mathrm{Var}\, f = \sup\left\{\sum_{j=1}^{M} |f(s^j) - f(s^{j-1})| \ : \ \{0 = s^0 < \cdots < s^M = T\}\right\}.$$

An alternative a posteriori control strategy relies on the direct use of the continuous functional J. Indeed, we readily check the following.

Lemma 5.4 (J controls the uniform distance from the solution). *Let $J(u) = 0$ and $v \in H^1(0,T;H)$. Then,*

$$\frac{1}{2}\|u - v\|^2_{C^0([0,T];H)} \leq J(u). \tag{24}$$

The proof is nothing but a continuous version of the argument of Lemma 5.1. Estimate (24) entails an alternative bound of the error between the continuous solution u and its approximation $\widehat{u}$ which is obtained by evaluating the continuous functional on the approximate solution.

5.4. Adaptivity

The above-introduced a posteriori error estimates can of course be exploited in order to develop an adaptive strategy. The error control in the uniform norm up to a given tolerance $\theta > 0$

$$\|u - \widehat{u}\|_{C^0([0,T];H)} \leq \theta \tag{25}$$

for some piecewise approximation $\widehat{u}$ with $|\widehat{u}(0) - u_0|^2 \leq \theta^2/2$, can be enforced, for instance, by choosing the partition $\{0 = t^0 < \cdots < t^N = T\}$ in such a way that

$$\int_{t^{i-1}}^{t^i} \left(\phi(\widehat{u}) + \phi^*(f - \widehat{u}') - (f - \widehat{u}', \widehat{u})\right) \leq \frac{\theta^2}{4N}$$

for $i = 1, \ldots, N$. Namely, by uniformly distributing the error along the partition.

Moreover, by exploiting the computable quantities $L^i(u^i, u^{i-1})$ and $Z^i(u^i, u^{i-1})$ and the theory by *Nochetto, Savaré & Verdi* [25, 26], the bound (25) can be achieved by asking for

$$-\tau^i Z^i(u, u^{i-1}) \le \frac{\theta^2}{16\,T} \quad \text{and} \quad L^i(u, u^{i-1}) \le \frac{\theta^2}{32\,T},$$

where we are considering the case $f = 0$ for the sake of simplicity.

5.5. The uniformly convex case

In case the functionals ϕ or ϕ^* are uniformly convex of constant $\gamma > 0$, the results of Subsections 5.1–5.2 are sharper. Let us collect here some remarks.

ϕ uniformly convex. In this case, the functionals J_n turn out to be uniformly convex as well. Referring to the notations of Subsection 5.1 and owing to [26, Thm. 3.14], in the uniformly convex case we gain an exponential decay of the error $|u - \widehat{e}_n|$. We can combine the above referred result and estimate (18) in order to sharpen the posteriori error control (21) at time T as follows

$$|u(T) - \widehat{u}_n(T)| \le \left(-\sum_{i=1}^{N_n} (\tau_n^i)^2 \exp(-\gamma(T - t_n^i)) Z^i(e_n^i, e_n^{i-1}) \right)^{1/2} + \int_0^T \exp(-\gamma(T-t)/2)(f - \overline{f}_n)(t)\ \mathrm{d}t + (2J_n(u_n))^{1/2}. \qquad (26)$$

ϕ^* uniformly convex. In this case, the global estimate (19) can be refined to

$$\frac{1}{2}|u_n^0 - u_0|^2 + \frac{\gamma}{2}\sum_{i=1}^{N_n} \tau_n^i |\delta u_n^i - \delta e_n^i|^2 \le J_n(u_n), \qquad (27)$$

which in particular entails that

$$\|\widehat{u}_n' - \widehat{e}_n'\|^2_{L^2(0,T;H)} \le \frac{2}{\gamma} J_n(u_n),$$

and we have the following sharper version of Lemma 5.3.

Lemma 5.5 (Improved convergence, ϕ^* uniformly convex). *Let ϕ^* be uniformly convex. Under the assumptions of Theorem 4.2, $\widehat{u}_n \to u$ strongly in $H^1(0,T;H)$ where u solves the gradient flow (1).*

6. Data approximations, space-time discretization

The data f and u_0 and the functional ϕ have been kept fixed throughout the above analysis. On the other hand, the variational method is such that approximations in f, u_0, and ϕ can be easily considered as well. Although limits in the diameter if the partition and the data could be considered in any order, we shall discuss here the situation of the joint limit

$$(f_h, u_{0,h}, \phi_h) \to (f, u_0, \phi) \quad \text{as } h \to 0$$

which encodes the case of full space-time approximations (see below). The above functional convergence $\phi_h \to \phi$ is intended in the *Mosco sense* [1, 21] namely, for all $u \in H$,

$$\phi(u) \le \liminf_{h\to 0} \phi_h(u_h) \quad \forall u_h \to u \text{ weakly in } H,$$
$$\exists u_h \to u \ \text{ strongly in } H \ \text{ such that } \phi(u) = \limsup_{h\to 0} \phi_h(u_h).$$

In particular, $\phi_h \to \phi$ in the Mosco sense iff $\phi_h \to \phi$ in the sense of Γ-convergence with respect to both the weak and the strong topology in H [8].

The corresponding Euler schemes read

$$\frac{u^i_{n,h} - u^{i-1}_{n,h}}{\tau^i_n} + \partial\phi_h(u^i_{n,h}) \ni f^i_{n,h} \ \text{ for } i = 1,\dots,N_n,\ u^0_n = u_{0,h}, \tag{28}$$

Let us now define $K_{n,h} : H^{N_n+1} \to [0,\infty]$ as

$$\begin{aligned} K_{n,h}(u^0,\dots,u^{N_n}) = &\left(\sum_{i=1}^{N_n} \tau^i_n \left(|\delta u^i|^2 - \left(f^i_{n,h}, \delta u^i\right)\right) + \phi_h(u^{N_n}) - \phi_h(u_{0,h})\right)^+ \\ &+ \sum_{i=1}^{N_n} \tau^i_n \left(\phi_h(u^i) + \phi^*_h\left(f^i_{n,h} - \delta u^i\right) - (f^i_{n,h}, u^i) + \frac{\tau^i_n}{2}|\delta u^i|^2\right) \\ &+ \frac{1}{2}|u^{N_n}|^2 - \frac{1}{2}|u^0|^2 + \frac{1}{2}|u^0 - u_{0,h}|^2. \end{aligned}$$

The convergence result in this case reads as follows.

Theorem 6.1 (Data approximations). *Assume that H is separable, $(f_h, u_{0,h}, \phi_h)$ fulfill the general assumptions of Section 1, and ϕ_h are uniformly bounded below. Moreover, let $\overline{f}_{n,h} \to f$ strongly in $L^2(0,T;H)$, $u_{0,h} \to u_0$ strongly, and $\phi_h \to \phi$ in the Mosco sense as $(n,h) \to (\infty,0)$. If $K_{n,h}(u_{n,h}) \to 0$ then $\widehat{u}_{n,h} \to u$ weakly in $H^1(0,T;H)$ where u solves the gradient flow (1).*

Proof. After extracting not relabeled subsequences, one can pass to the lim inf as $(n,h) \to (\infty,0)$ in

$$\begin{aligned} K_{n,h}(u_{n,h}) \ge &\left(\int_0^T \left(|\widehat{u}'_{n,h}|^2 - (\overline{f}_{n,h}, \widehat{u}'_{n,h})\right) + \phi_h(\widehat{u}_{n,h}(T)) - \phi_h(u_{0,h})\right)^+ \\ &+ \int_0^T \left(\phi_h(\overline{u}_{n,h}) + \phi^*_h(\overline{f}_{n,h} - \widehat{u}'_{n,h}) - (\overline{f}_{n,h}, \overline{u}_{n,h})\right) \\ &+ \frac{1}{2}|\widehat{u}_{n,h}(T)|^2 + \frac{1}{2}|\widehat{u}_{n,h}(0)|^2 + \frac{1}{2}|u_{0,h}|^2 - (\widehat{u}_{n,h}(0), u_{0,h}). \end{aligned}$$

In order to do so, the only terms that need some comment are those containing the approximating functionals ϕ_h and ϕ^*_h. To this aim, we shall exploit a technical tool from [38, Cor. 4.4] (but see also [34]) in order to ensure that

$$\int_0^T \phi(u) \le \liminf_{(n,h)\to(\infty,0)} \int_0^T \phi_h(\overline{u}_{n,h}),$$
$$\int_0^T \phi^*(f - u') \le \liminf_{(n,h)\to(\infty,0)} \int_0^T \phi^*_h(\overline{f}_{n,h} - \widehat{u}'_{n,h}).$$

In particular, one has that

$$K(u) \le \liminf_{(n,h)\to(\infty,0)} K_{n,h}(u_{n,h}),$$

and the assertion follows. □

Remark 6.2. There is no particular intricacy in considering the case of the functional J as well (apart from the fact that weak pre-compactness of the minimizing sequence has to be explicitly assumed).

6.1. Space-time discretization

A direct application of Theorem 6.1 entails the convergence of a combined time-discretization and conformal finite element approximation of the gradient flow (1). Let (f, u_0, ϕ) fulfill the general assumptions of Section 1. Moreover, we will ask ϕ to be continuous and bounded below and assume to be given a nested sequence of finite dimensional spaces $H_h \subset H$ such that $\bigcup_{h>0} H_h$ is dense in H. We define the approximating functionals as

$$\phi_h(u) = \phi(u) \quad \text{if } u \in H_h \quad \text{and} \quad \phi_h(u) = \infty \quad \text{otherwise.}$$

By letting $\pi_h : H \to H_h$ denote the standard projection in H, we let

$$f_h = \pi_h f, \quad f^i_{n,h} = \frac{1}{\tau^i_n}\int_{t^{i-1}_n}^{t^i_n} f_h \quad \text{for } i = 1, \dots, N_n, \quad \text{and} \quad u_{0,h} = \pi_h u_0.$$

It is a standard matter to check that the assumptions of Theorem 6.1 are fulfilled. In particular,

$$\begin{aligned}\|f - \overline{f}_{n,h}\|_{L^2(0,T;H)} &\le \|f - \pi_h f\|_{L^2(0,T;H)} + \|\pi_h f - \overline{f}_{n,h}\|_{L^2(0,T;H)} \\ &\le \|f - \pi_h f\|_{L^2(0,T;H)} + \|f - \overline{f}_n\|_{L^2(0,T;H)}.\end{aligned}$$

Moreover, $\phi_h \to \phi$ in the Mosco sense. By applying Theorem 6.1 we readily get that the space-time discretized solutions $\widehat{u}_{n,h}$ such that $K_{n,h}(u_{n,h}) \to 0$ are weakly converging in the weak topology of $H^1(0,T;H)$ to the solution of the corresponding continuous problem.

References

[1] H. Attouch: Variational Convergence for Functions and Operators, Pitman, Boston (1984).

[2] J.-P. Aubin: Variational principles for differential equations of elliptic, parabolic and hyperbolic type, in: Mathematical Techniques of Optimization, Control and Decision, J.-P. Aubin et al. (ed.), Birkhäuser, Boston (1981) 31–45.

[3] J.-P. Aubin. A. Cellina: Differential Inclusions, Grundlehren der Mathematischen Wissenschaften 264, Springer, Berlin (1984).

[4] G. Auchmuty: Saddle-points and existence-uniqueness for evolution equations, Differ. Integral Equ. 6(5) (1993) 1161–1171.

[5] H. Brezis: Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert, North-Holland Math. Studies 5, North-Holland, Amsterdam (1973).

[6] H. Brezis, I. Ekeland: Un principe variationnel associé à certaines équations paraboliques. Le cas dépendant du temps, C. R. Acad. Sci., Paris, Sér. A 282 (1976) 971–974.

[7] H. Brezis, I. Ekeland: Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps, C. R. Acad. Sci., Paris, Sér. A 282 (1976) 1197–1198.

[8] G. Dal Maso: An Introduction to Γ-convergence, Progress in Nonlinear Differential Equations and their Applications 8, Birkhäuser, Basel (1993).

[9] N. Ghoussoub: Anti-selfdual Lagrangians: variational resolutions of non self-adjoint equations and dissipative evolutions, Ann. Inst. H. Poincaré, Anal. Non Linéaire 24(2) (2007) 171–205.

[10] N. Ghoussoub: Selfdual Partial Differential Systems and their Variational Principles, Springer, Berlin (2009).

[11] N. Ghoussoub: Maximal monotone operators are sefldual vector fields and vice-versa, preprint (2006), available at: http://www.pims.math.ca/∼nassif/ .

[12] N. Ghoussoub, R. J. McCann: A least action principle for steepest descent in a non-convex landscape, in: Partial Differential Equations and Inverse Problems, C. Conca et al. (ed.), Contemp. Math. 362, Amer. Math. Soc., Providence (2004) 177–187.

[13] N. Ghoussoub, L. Tzou: A variational principle for gradient flows, Math. Ann. 330(3) (2004) 519–549.

[14] E. Hairer, Ch. Lubich, G. Wanner: Geometric Numerical Integration. Structure-preserving Algorithms for Ordinary Differential Equations, 2nd Ed., Springer Series in Computational Mathematics 31, Springer, Berlin (2006).

[15] B. Lemaire: An asymptotical variational principle associated with the steepest descent method for a convex function, J. Convex Analysis 3(1) (1996) 63–70.

[16] F. Maggi, M. Morini: A Γ-convergence result for variational integrators of Lagrangians with quadratic growth, ESAIM, Control Optim. Calc. Var. 10(4) (2004) 656–665.

[17] J. E. Marsden, M. West: Discrete mechanics and variational integrators, Acta Numerica 10 (2001) 357–514.

[18] A. Mielke, M. Ortiz: A class of minimum principles for characterizing the trajectories and the relaxation of dissipative systems, ESAIM, Control Optim. Calc. Var. 14(3) (2008) 494–516.

[19] A. Mielke, U. Stefanelli: A discrete variational principle for rate-independent evolution, Adv. Calc. Var. 1(4) (2008) 399–431.

[20] A. Mielke, U. Stefanelli: Weighted energy-dissipation functionals for gradient flows, submitted (2009).

[21] U. Mosco: Convergence of convex sets and of solutions of variational inequalities, Adv. Math. 3 (1969) 510–585.

[22] S. Müller, M. Ortiz: On the Γ-convergence of discrete dynamics and variational integrators, J. Nonlinear Sci. 14(3) (2004) 279–296.

[23] B. Nayroles: Deux théorèmes de minimum pour certains systèmes dissipatifs, C. R. Acad. Sci., Paris, Sér. A 282(17) (1976) 1035–1038.

[24] B. Nayroles: Un théorème de minimum pour certains systèmes dissipatifs, variante hilbertienne, Trav. Sémin. d'Anal. Convexe 6, Exp. 2 (1976).

[25] R. Nochetto, G. Savaré, C. Verdi: Error control of nonlinear evolution equations, C. R. Acad. Sci., Paris, Sér. I, Math. 326(12) (1998) 1437–1442.

[26] R. Nochetto, G. Savaré, C. Verdi: A posteriori error estimates for variable time-step discretization of nonlinear evolution equations, Commun. Pure Appl. Math. 53(5) (2000) 525–589.

[27] R. H. Nochetto, G. Savaré: Nonlinear evolution governed by accretive operators in Banach spaces: error control and applications, Math. Models Methods Appl. Sci. 16(3) (2006) 439–477.

[28] 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 283(3) (1976) 83–86.

[29] H. Rios: La question d'existence de solutions pour certaines équations à opérateurs monotones vue comme problème de minimum ou comme problème de point-selle, Trav. Sémin. d'Anal. Convexe 6, Exp. 13 (1976).

[30] H. Rios: Étude de certains problèmes paraboliques: existence et approximation des solutions, Trav. Sémin. d'Anal. Convexe 8(1), Exp. 1 (1978).

[31] H. Rios: Une étude d'existence sur certains problèmes paraboliques, Ann. Fac. Sci. Toulouse, V. Ser., Math. 1(3) (1979) 235–255.

[32] T. Roubíček: Direct method for parabolic problems, Adv. Math. Sci. Appl. 10(1) (2000) 57–65.

[33] T. Roubíček: Nonlinear Partial Differential Equations with Applications, Int. Ser. Numer. Math. 153, Birkhäuser, Basel (2005).

[34] A. Salvadori: On the M-convergence for integral functionals on L^p_X, Atti Sem. Mat. Fis. Univ. Modena 33(1) (1984) 137–154.

[35] G. Savaré: Error estimates for dissipative evolution problems, in: Free Boundary Problems: Theory and Applications (Trento, 2002), P. Colli et al. (ed.), Int. Ser. Numer. Math. 147, Birkhäuser, Basel (2003) 281–291.

[36] U. Stefanelli: A variational principle in non-smooth mechanics, in: MFO Workshop: Analysis and numerics for rate-independent processes, Oberwolfach Reports 4(1) (2007) 622–624.

[37] U. Stefanelli: A variational characterization of rate-independent evolution, submitted (2008).

[38] U. Stefanelli: The Brezis-Ekeland principle for doubly nonlinear equations, SIAM J. Control Optim. 47(3) (2008) 1615–1642.

[39] U. Stefanelli: A variational principle for hardening elasto-plasticity, SIAM J. Math. Anal. 40(2) (2008) 623–652.