Error Estimates for Dissipative Evolution Problem

Giuseppe Savaré

Abstract. We present a quick overview of the general problem to find optimal *a priori* and *a posteriori* error estimates for the approximation of dissipative evolution equations in Hilbert and metric spaces by means of a variational formulation of the implicit Euler scheme. We shall discuss what are the intrinsic metric arguments which are involved in the derivation of the estimates and we present an elementary proof in a simplified finite dimensional case. An application to the porous medium equation in the new framework of the Wasserstein distance is briefly sketched.

1. Introduction

In this paper we are interested in optimal *a priori* and *a posteriori* error estimates for the approximation of the solution $u : [0,T] \to \mathscr{S}$ of the gradient flow equation

$$\begin{cases} u'(t) = -\nabla\phi(u(t)) & t \in [0,T], \\ u(0) = u^0, \end{cases} \tag{1}$$

by means of the implicit Euler scheme

$$\frac{U^n_{\boldsymbol\tau} - U^{n-1}_{\boldsymbol\tau}}{\tau_n} = -\nabla\phi(U^n_{\boldsymbol\tau}), \quad n = 1, \cdots, N; \quad U^0_{\boldsymbol\tau} \approx u^0 \text{ given.} \tag{2}$$

Here $\phi : \mathscr{S} \to \mathbb{R}$ is a given functional defined in a suitable ambient space $\mathscr{S}$, $\boldsymbol\tau := \{\tau_n\}_{n=1}^N$ denotes the variable step sizes induced by a partition $\mathcal{P}_{\boldsymbol\tau}$ of the time interval $[0,T]$ into N subintervals $I^n_{\boldsymbol\tau} := (t^{n-1}_{\boldsymbol\tau}, t^n_{\boldsymbol\tau}]$,

$$\mathcal{P}_{\boldsymbol\tau} := \{0 = t^0_{\boldsymbol\tau} < t^1_{\boldsymbol\tau} < \ldots < t^{N-1}_{\boldsymbol\tau} < t^N_{\boldsymbol\tau} = T\}, \quad \tau_n = t^n_{\boldsymbol\tau} - t^{n-1}_{\boldsymbol\tau} = |I^n_{\boldsymbol\tau}|, \tag{3}$$

and the equation (2) should be intended as a recursive algorithm which indicates how to find $U^n_{\boldsymbol\tau}$ once $U^{n-1}_{\boldsymbol\tau}$ is known.

The *discrete solution* $U_{\boldsymbol\tau}(t^n_{\boldsymbol\tau}) := U^n_{\boldsymbol\tau}, n = 1, \cdots, N$, will thus provide an approximation of the exact solution u at the nodes $t^n_{\boldsymbol\tau}$ of the partition $\mathcal{P}_{\boldsymbol\tau}$. We want

This work was partially supported by grants of M.I.U.R. (cofin2002) and of the Institute of Applied Mathematics and Information Technology (I.M.A.T.I.) of the C.N.R., Pavia, Italy.

to find estimates of the *error* $E_{\boldsymbol\tau}$ between $U_{\boldsymbol\tau}$ and the exact solution u

$$E_{\boldsymbol\tau} := \max_{t\in\mathcal{P}_{\boldsymbol\tau}} d(U_{\boldsymbol\tau}(t), u(t)), \quad d \text{ being a suitable } \textit{distance} \text{ defined on } \mathscr{S},$$

addressing the following fundamental issues.

- The estimates are *a posteriori*, i.e. they are expressed in terms of computable quantities which depend solely on time-steps, discrete solution, and data.
- The *a posteriori* bounds should converge to zero as the maximum of the time steps
$$|\boldsymbol\tau| := \max_{1\le n\le N} \tau_n \tag{4}$$
goes to 0 with an *optimal rate* w.r.t. the regularity of the solution u.
- The estimates should depend on the structural properties (*dissipativity*) of $\mathscr{S}$ and ϕ which ensure the well posedness of (1): in particular, they should not depend on the space dimension of $\mathscr{S}$ and they are explicitly determined without need of solving any auxiliary problem. No extra regularity assumptions should be assumed.
- No *a priori* constraints should be imposed between consecutive time-steps, which could just be tailored to the *a posteriori* error estimators alone.

Four possible frameworks

We will focus on frameworks where (1) is, in some sense to be better specified, *dissipative* and its dependence with respect to (w.r.t.) perturbations of the initial datum can be explicitly controlled in terms of the distance d.

a) The simplest case: $\mathscr{S} := \mathbb{R}^m$, d is the usual Euclidean distance, ϕ is a (quadratic perturbation of a) C^1 convex function: (1) is a system of ODE's.

b) Infinite dimensional Hilbert spaces $\mathscr{S} := \mathscr{H}$, the distance being induced by the scalar product $\langle\cdot,\cdot\rangle$. The most interesting case in view of applications to nonlinear evolutionary PDE's [18] is provided by lower semicontinuous (l.s.c.) convex functionals ϕ (continuous quadratic perturbations are still allowed), which take their values in the *extended* real line $(-\infty,+\infty]$. A crucial feature of the convexity of ϕ is the variational formulation of both (1) and (2) by means of evolution inequalities. In fact, the solution of (1) can also be characterized by means of the system

$$\langle u'(t), u(t)-v\rangle + \phi(u(t)) \le \phi(v) \quad \forall\, v\in\mathscr{H}, \tag{5}$$

which, at the discrete level, corresponds to

$$\langle \tau_n^{-1}(U^n_{\boldsymbol\tau} - U^{n-1}_{\boldsymbol\tau}), U^n_{\boldsymbol\tau} - V\rangle + \phi(U^n_{\boldsymbol\tau}) \le \phi(V) \quad \forall\, V\in\mathscr{H}; \tag{6}$$

they both lead to the nonexpansivity property of the continuous (and discrete) trajectories starting from two initial values u^0, v^0 (resp. $U^0_{\boldsymbol\tau}, V^0_{\boldsymbol\tau}$)

$$d(u(t),v(t)) \le d(u^0,v^0), \quad d(U^n_{\boldsymbol\tau}, V^n_{\boldsymbol\tau}) \le d(U^0_{\boldsymbol\tau}, V^0_{\boldsymbol\tau}) \quad t\in[0,T],\ n=1,\cdots,N. \tag{7}$$

In the framework of this well developed theory (see e.g. [4]) the same approximation results of case *a)* can be reproduced. We refer to [14] for a detailed account of the various contributions and for a presentation of the estimates and of the related applications; here we simply quote the optimal *a priori* estimates of [3], [15], and the linear theory of [8] (the last of an important series of papers on this subject).

c) $\mathscr{S}$ is a Riemannian manifold, endowed with its Riemannian distance d. In this case equation (1) should be imposed in the tangent space of $\mathscr{S}$ at the point $u(t)$ and we could interpret the notion of convexity as *convexity along geodesics.*

Conversely, (2) requires more attention, since now the difference $U^n_{\boldsymbol{\tau}} - U^{n-1}_{\boldsymbol{\tau}}$ is no more defined. One possible way to circumvent this difficulty is to rewrite (2) in a variational form, which is already suggested in the Euclidean case. We simply notice that among the solutions of (2) we can select the minima of the functionals

$$V \mapsto \Phi^n_{\boldsymbol{\tau}}(V) := \tfrac{1}{2\tau_n} d^2(V, U^{n-1}_{\boldsymbol{\tau}}) + \phi(V), \tag{8a}$$

i.e. $U^n_{\boldsymbol{\tau}},\ n = 1, \cdots, N$, solve recursively the family of minimum problems

$$\Phi^n_{\boldsymbol{\tau}}(U^n_{\boldsymbol{\tau}}) = \min_{V \in \mathscr{S}} \Phi^n_{\boldsymbol{\tau}}(V) \quad n = 1, \cdots, N. \tag{8b}$$

(8a,b) still make sense in a Riemannian manifold and they can be used to approximate (1). Of course, we expect that the error estimates are affected by the curvature of $\mathscr{S}$; nevertheless, it is interesting to note that, at least in the case of a simply connected manifold with *nonpositive* sectional curvature, we can reproduce the same result of the Euclidean framework. It is one of the most interesting contribution of [11], [2] to show that, as in the Euclidean case, these estimates do not directly rely on the finite dimension or on the local differentiable structure of the manifold, but on simple geometric properties of its Riemannian distance.

d) $(\mathscr{S}, d)$ is a metric space. In this case is no more clear what is the meaning of (1), whereas the recursive variational scheme (8a,b) still applies and provides a family of (hopefully) approximate solutions to (1). This lead many authors to use this approximation procedure to define a solution of (1), when the differential problem does not exhibit a useful linear structure. In an abstract metric framework the contributions of E. De Giorgi and his collaborators [7], [6], [1], show how to give a meaning to (1) and provide general conditions for the convergence (up to an extraction of a subsequence) of the variational scheme with uniform step sizes.

Before discussing the delicate question of convergence estimate in this abstract setting, let us briefly recall a particular example, which was also one of the main motivation of the present investigation: it arises from the recent papers of F. Otto [9, 16] on a new variational interpretation of the Fokker-Planck and the porous medium equations as gradient flows w.r.t. suitable functionals in the space of probability measures metrized by the so called Wasserstein distance. For the sake of simplicity, here we will consider only measures which are absolutely continuous w.r.t. the Lebesgue one, and we will identify them with their densities.

An example of the metric setting: the porous medium equation and the Wasserstein distance.

Let us introduce the set $\mathscr{S} = \mathscr{P}^2_a(\mathbb{R}^k)$

$$\mathscr{P}^2_a(\mathbb{R}^k) := \Big\{ u \in L^1(\mathbb{R}^k) : u \geq 0, \int_{\mathbb{R}^k} u(x)\,dx = 1, \int_{\mathbb{R}^k} |x|^2 u(x)\,dx < +\infty \Big\}, \tag{9a}$$

endowed with the *Wasserstein distance* (we refer to the recent book [17] for a comprehensive and up-to-date overview of this and related topics)

$$\begin{aligned} W^2(v,w) := \min \Big\{ \int_{\mathbb{R}^k} |\boldsymbol{r}(x) - x|^2\, v(x)\,dx : \quad & \boldsymbol{r} : \mathbb{R}^k \to \mathbb{R}^k \text{ is a Borel map}, \\ \int_{\mathbb{R}^k} \zeta(y) w(y)\,dy = \int_{\mathbb{R}^k} \zeta(\boldsymbol{r}(x)) v(x)\,dx \quad & \text{for all functions } \zeta \in C^0_c(\mathbb{R}^k) \Big\}. \end{aligned} \tag{9b}$$

The map $\boldsymbol{r}$ that realizes the minimum in (9b) is usually called *the optimal transportation* between v and w. For $m > 1$ we also introduce the functional on $\mathscr{P}^2_a(\mathbb{R}^k)$

$$\varphi(u) := \frac{1}{m-1} \int_{\mathbb{R}^k} u^m(x)\,dx. \tag{9c}$$

In [16] F. Otto showed that the nonlinear (*porous medium*) diffusion equation

$$\begin{cases} \partial_t u - \Delta(u^m) = 0 & \text{in } \mathbb{R}^k \times (0,T), \\ u(x,0) = u^0(x) & \text{in } \mathbb{R}^k, \quad u^0 \in \mathscr{P}^2_a(\mathbb{R}^k), \end{cases} \tag{9d}$$

can be considered as the *gradient flow* of the functional φ in $\mathscr{P}^2_a(\mathbb{R}^k)$ w.r.t. the distance W: here the meaning of "*gradient flow*" has to be intended exactly as the limit of the discrete solutions $U_{\boldsymbol{\tau}}$ obtained by the recursive minimization algorithm (8a,b) with the choice $d := W$ and $\phi := \varphi$ corresponding to (9a,b).

Even if $\mathscr{P}^2_a(\mathbb{R}^k)$ is a convex subset of $L^1(\mathbb{R}^k)$, the segments do not provide the shortest path (w.r.t. W) connecting two elements, since they do not have even finite length. However, it is worth mentioning that the functional φ defined by (9c) is still convex along the (constant speed) geodesics of $\mathscr{P}^2_a(\mathbb{R}^k)$, as showed by [12].

Geodesic convexity is not enough for error estimates in metric spaces.

Let us recall that the well posedness (and therefore also the problem to find optimal error estimates) for gradient flows of convex functionals in general *Banach* spaces (even finite dimensional) is still completely open. This remark indicates, at least euristically, that severe restrictions should be imposed to the general metric setting, and that the geometric properties of the distance should play a crucial role.

A first step in this direction has been obtained by U. Mayer [11], who considered gradient flows of geodesically convex functionals on *nonpositively curved metric spaces*: these are length spaces where the distance from any given point $w \in \mathscr{S}$ satisfies the inequality

$$d^2(v_t, w) \leq (1-t) d^2(v_0, w) + t d^2(v_1, w) - t(1-t) d^2(v_0, v_1) \quad t \in [0,1], \tag{10}$$

along any constant speed geodesic curve $\{v_t\}_{t\in[0,1]}$ connecting two arbitrary points $v_0, v_1 \in \mathscr{S}$. This property, which was introduced by Aleksandrov on the basis of the analogous inequality satisfied in Euclidean spaces (in fact an equality) and in Riemannian manifolds of nonpositive sectional curvature [10, §2.3], allows to prove (7), and to obtain the convergence of the approximation scheme (and a suboptimal *a priori* error estimate) by following the same ideas of the celebrated Crandall-Liggett generation Theorem [5]. It is interesting to note that if d is induced by a norm in a vector space and it satisfies (10), then the norm obeys the parallelogram rule (simply choose $t = 1/2$ in (10)) and therefore it is induced by a scalar product.

These assumptions, however, do not cover the case of the Wasserstein distance (9a,b). In fact, it is possible to prove (cf. [2]) that the distance of this space satisfies the opposite inequality of (10), thus providing a *positively* curved space, as formally suggested also by [16].

A new geometric condition

The main idea to overcome this difficulty and to derive better error estimates is to allow more flexibility to the choice of the connecting curves, which do not need to be geodesics but could depend on the distance and on the functional. This approach relies on a careful analysis of the techniques of [14] and is completely developed in an intrinsic metric framework in [2], where a particular attention to the case of the Wasserstein distance and its applications is devoted. It is surprising that all the estimates are based only on the following geometrical property:

for every triple $w, v_0, v_1 \in \mathscr{S}$ there exists a curve $\{v_t\}_{t\in[0,1]}$ connecting v_0 to v_1 which satisfies

$$\phi(v_t) \le (1-t)\phi(v_0) + t\phi(v_1) \quad t \in [0,1], \tag{11a}$$

$$d^2(w, v_t) \le (1-t)d^2(w, v_0) + td^2(w, v_1) - t(1-t)d^2(v_0, v_1). \tag{11b}$$

In the next section, following the ideas of [2], we will present a new elementary derivation of the estimates of [14] in the easiest finite dimensional framework (thus avoiding any technical tools related to the theory of maximal monotone operators), trying to use only the *metric properties* of the Euclidean spaces related to (11a,b). We shall obtain estimates independent of the dimension and of the vectorial structure, which therefore will be immediately appliable to general Hilbert and metric spaces. An application to the porous medium equation will be briefly sketched at the end of the paper.

2. Main result

In this section we shall deal with the simplest situtation, which is provided by a

convex C^1 functional ϕ defined in a finite dimensional space $\mathscr{S} := \mathbb{R}^m$,

endowed with the usual euclidean distance $d(u, v) := |u - v|$.

Since the elementary proof of this result will be based only on the fact that for each $w, v_0, v_1 \in \mathscr{S}$

the segment $v_t := (1-t)v_0 + tv_1$, $t \in [0,1]$, satisfies (11a,b),

we will keep the notation $d(u, v)$ instead of $|u - v|$ for the distance.

Theorem 2.1 (Optimal error estimate)**.** *Let u be the solution of* (1)*, let $U_{\boldsymbol{\tau}}$ be the discrete solution obtained by solving* (2) *or, equivalently,* (8a,b)*. Then*

$$E_{\boldsymbol{\tau}}^2 = \max_{t \in \mathcal{P}_{\boldsymbol{\tau}}} d^2(U_{\boldsymbol{\tau}}(t), u(t)) \leq d^2(u^0, U_{\boldsymbol{\tau}}^0) + \sum_{n=1}^N \tau_n^2 \mathscr{E}_{\boldsymbol{\tau}}^n, \tag{12}$$

where $\mathscr{E}_{\boldsymbol{\tau}}^n$ is the a posteriori error estimator *defined by*

$$\mathscr{E}_{\boldsymbol{\tau}}^n := \frac{\phi(U_{\boldsymbol{\tau}}^{n-1}) - \phi(U_{\boldsymbol{\tau}}^n)}{\tau_n} - \frac{d^2(U_{\boldsymbol{\tau}}^{n-1}, U_{\boldsymbol{\tau}}^n)}{\tau_n^2} \geq 0. \tag{13}$$

Moreover, if ϕ is bounded from below, then

$$\sum_{n=1}^N \tau_n^2 \mathscr{E}_{\boldsymbol{\tau}}^n \leq |\boldsymbol{\tau}| \big(\phi(U_{\boldsymbol{\tau}}^0) - \inf_{\mathscr{S}} \phi\big). \tag{14}$$

Finally, the following optimal apriori *bound in terms of the gradient of ϕ holds*

$$\sum_{n=1}^N \tau_n^2 \mathscr{E}_{\boldsymbol{\tau}}^n \leq \tfrac{1}{2} |\boldsymbol{\tau}|^2 \, |\nabla \phi(U_{\boldsymbol{\tau}}^0)|^2. \tag{15}$$

We will divide the proof in some steps.

A "metric variational inequality" for u. Our first remark is that (5) can also be written as

$$\tfrac{d}{dt} \tfrac{1}{2} d^2(u(t), v) + \phi(u(t)) \leq \phi(v) \quad \forall\, v \in \mathscr{S},\ t \in [0,T]. \tag{16}$$

Observe that, being u regular, (16) contains all the information of (1); in particular it is easy to derive from (16) the contraction property for continuous trajectories (7). In fact, if w is another solution of (16), which we rewrite by introducing another variable s

$$\tfrac{d}{ds} \tfrac{1}{2} d^2(w(s), v) + \phi(w(s)) \leq \phi(v) \quad \forall\, v \in \mathscr{S},\ t \in [0,T], \tag{17}$$

then choosing $v := w(s)$ in (16), $v := u(t)$ in (17), and summing up the two inequalities, we find

$$\tfrac{\partial}{\partial t} \tfrac{1}{2} d^2(u(t), w(s)) + \tfrac{\partial}{\partial s} \tfrac{1}{2} d^2(u(t), w(s)) \leq 0 \quad s, t \in [0,T]. \tag{18}$$

Evaluating this relation for $s = t$ we obtain

$$\tfrac{d}{dt} \tfrac{1}{2} d^2(u(t), w(t)) \leq 0 \quad \text{i.e.} \quad d(u(t), w(t)) \leq d(u^0, w^0) \quad t \in [0,T]. \tag{19}$$

A "metric variational inequality" for $U_{\boldsymbol{\tau}}^n$. The second step consists in writing an analogous variational inequality for the discrete solution: here we will use in a crucial way (11a) and (11b). In fact, choosing $w := U_{\boldsymbol{\tau}}^{n-1}$, it is easy to see that they yield the following strong convexity property for the functionals $\Phi_{\boldsymbol{\tau}}^n(\cdot)$:

$$\Phi_{\boldsymbol{\tau}}^n(v_t) \leq (1-t)\Phi_{\boldsymbol{\tau}}^n(v_0) + t\Phi_{\boldsymbol{\tau}}^n(v_1) - \tfrac{1}{2\tau_n} t(1-t) d^2(v_0, v_1). \tag{20}$$

Starting from the minimum property (8b) and applying (20) with $v_0 := U_{\boldsymbol{\tau}}^n$, $v_1 := V$, we get

$$\Phi_{\boldsymbol{\tau}}^n(U_{\boldsymbol{\tau}}^n) \leq \Phi_{\boldsymbol{\tau}}^n(v_t) \leq (1-t)\Phi_{\boldsymbol{\tau}}^n(U_{\boldsymbol{\tau}}^n) + t\Phi_{\boldsymbol{\tau}}^n(V) - \tfrac{1}{2\tau_n} t(1-t) d^2(U_{\boldsymbol{\tau}}^n, V) \quad \forall\, V \in \mathscr{S}.$$

The minimum condition says that the right derivative at $t = 0$ of the right hand side is nonnegative; thus we find

$$\Phi^n_{\boldsymbol{\tau}}(V) - \Phi^n_{\boldsymbol{\tau}}(U^n_{\boldsymbol{\tau}}) - \tfrac{1}{2\tau} d^2(U^n_{\boldsymbol{\tau}}, V) \geq 0 \quad \forall V \in \mathscr{S}, \tag{21}$$

which can also be written as

$$\tfrac{1}{\tau_n}\Big(\tfrac{1}{2} d^2(U^n_{\boldsymbol{\tau}}, V) - \tfrac{1}{2} d^2(U^{n-1}_{\boldsymbol{\tau}}, V)\Big) + \tfrac{1}{2\tau_n} d^2(U^n_{\boldsymbol{\tau}}, U^{n-1}_{\boldsymbol{\tau}}) + \phi(U^n_{\boldsymbol{\tau}}) \leq \phi(V). \tag{22}$$

A continuous formulation of (22). We want to write (22) as a true differential evolution inequality for the discrete solution $U_{\boldsymbol{\tau}}$, in order to compare it with the continuous one (16) and to try to reproduce the same comparison argument which we used for (18) and (19). A first idea, which was exploited in [14], is to use a linear interpolation between the consecutive values of the discrete solution. But an even simpler choice is possible (as in [13]): we interpolate the values of the functions on $U^n_{\boldsymbol{\tau}}$ instead of interpolating the arguments, since this does not even require an underlying linear structure. Therefore, we set

$$\phi_{\boldsymbol{\tau}}(s) := \text{“the linear interpolant of } \phi(U^{n-1}_{\boldsymbol{\tau}}) \text{ and } \phi(U^n_{\boldsymbol{\tau}})\text{” if } s \in (t^{n-1}_{\boldsymbol{\tau}}, t^n_{\boldsymbol{\tau}}], \tag{23}$$

i.e.

$$\phi_{\boldsymbol{\tau}}(s) := \frac{t^n_{\boldsymbol{\tau}} - s}{\tau_n}\phi(U^{n-1}_{\boldsymbol{\tau}}) + \frac{s - t^{n-1}_{\boldsymbol{\tau}}}{\tau_n}\phi(U^n_{\boldsymbol{\tau}}) \quad s \in (t^{n-1}_{\boldsymbol{\tau}}, t^n_{\boldsymbol{\tau}}]. \tag{24}$$

Analogously, for any $V \in \mathscr{S}$ we set

$$d^2_{\boldsymbol{\tau}}(s; V) := \frac{t^n_{\boldsymbol{\tau}} - s}{\tau_n} d^2(U^{n-1}_{\boldsymbol{\tau}}, V) + \frac{s - t^{n-1}_{\boldsymbol{\tau}}}{\tau_n} d^2(U^n_{\boldsymbol{\tau}}, V) \quad s \in (t^{n-1}_{\boldsymbol{\tau}}, t^n_{\boldsymbol{\tau}}]. \tag{25}$$

Since

$$\tfrac{d}{ds} d^2_{\boldsymbol{\tau}}(s; V) = \tfrac{1}{\tau_n}\Big(d^2(U^n_{\boldsymbol{\tau}}, V) - d^2(U^{n-1}_{\boldsymbol{\tau}}, V)\Big) \quad s \in (t^{n-1}_{\boldsymbol{\tau}}, t^n_{\boldsymbol{\tau}}],$$

(22) becomes

$$\tfrac{d}{ds}\tfrac{1}{2} d^2_{\boldsymbol{\tau}}(s; V) + \phi_{\boldsymbol{\tau}}(s) \leq \phi(V) + \tfrac{1}{2}\mathscr{R}_{\boldsymbol{\tau}}(s) \quad \forall s \in (0, T) \setminus \mathcal{P}_{\boldsymbol{\tau}}, \tag{26}$$

where we set for $s \in (t^{n-1}_{\boldsymbol{\tau}}, t^n_{\boldsymbol{\tau}}]$, $t^{n-1/2}_{\boldsymbol{\tau}} := \frac{1}{2}(t^n_{\boldsymbol{\tau}} + t^{n-1}_{\boldsymbol{\tau}})$,

$$\begin{aligned} \tfrac{1}{2}\mathscr{R}_{\boldsymbol{\tau}}(s) &:= \phi_{\boldsymbol{\tau}}(s) - \phi(U^N_{\boldsymbol{\tau}}) - \tfrac{1}{2\tau_n} d^2(U^n_{\boldsymbol{\tau}}, U^{n-1}_{\boldsymbol{\tau}}) \\ &= \frac{t^n_{\boldsymbol{\tau}} - s}{\tau_n}\Big(\phi(U^{n-1}_{\boldsymbol{\tau}}) - \phi(U^n_{\boldsymbol{\tau}})\Big) - \frac{d^2(U^n_{\boldsymbol{\tau}}, U^{n-1}_{\boldsymbol{\tau}})}{2\tau_n} \\ &= (t^n_{\boldsymbol{\tau}} - s)\mathscr{E}^n_{\boldsymbol{\tau}} + (t^{n-1/2}_{\boldsymbol{\tau}} - s)\frac{d^2(U^n_{\boldsymbol{\tau}}, U^{n-1}_{\boldsymbol{\tau}})}{\tau_n^2}. \end{aligned} \tag{27}$$

The comparison argument. Taking a convex combination w.r.t. the variable $s \in I^n_{\boldsymbol{\tau}}$ of (16) written for $v := U^{n-1}_{\boldsymbol{\tau}}$ and $v := U^n_{\boldsymbol{\tau}}$, we easily get

$$\tfrac{\partial}{\partial t}\tfrac{1}{2} d^2_{\boldsymbol{\tau}}(s, u(t)) + \phi(u(t)) \leq \phi_{\boldsymbol{\tau}}(s) \quad t, s \in [0, T], \tag{28}$$

whereas choosing $V := u(t)$ in (26) we find

$$\tfrac{\partial}{\partial s}\tfrac{1}{2} d^2_{\boldsymbol{\tau}}(s; u(t)) + \phi_{\boldsymbol{\tau}}(s) \leq \phi(u(t)) + \tfrac{1}{2}\mathscr{R}_{\boldsymbol{\tau}}(s) \quad t \in [0, T],\ s \in [0, T] \setminus \mathcal{P}_{\boldsymbol{\tau}}. \tag{29}$$

Summing up (28) and (29) we end up with

$$\frac{\partial}{\partial t} d^2_{\boldsymbol{\tau}}(s,u(t)) + \frac{\partial}{\partial s} d^2_{\boldsymbol{\tau}}(s;u(t)) \le \mathscr{R}_{\boldsymbol{\tau}}(s) \quad t \in [0,T],\ s \in [0,T] \setminus \mathcal{P}_{\boldsymbol{\tau}}. \tag{30}$$

Choosing $s = t$ we eventually find

$$\frac{d}{dt} d^2_{\boldsymbol{\tau}}(t,u(t)) \le \mathscr{R}_{\boldsymbol{\tau}}(t) \quad t \in [0,T] \setminus \mathcal{P}_{\boldsymbol{\tau}}, \tag{31}$$

and therefore, being $t \mapsto d_{\boldsymbol{\tau}}(t,u(t))$ continuous,

$$d^2_{\boldsymbol{\tau}}(T,u(T)) \le d^2_{\boldsymbol{\tau}}(0,u^0) + \int_0^T \mathscr{R}_{\boldsymbol{\tau}}(t)\,dt. \tag{32}$$

Since $d_{\boldsymbol{\tau}}(t,V) = d(U_{\boldsymbol{\tau}}(t),V)$ if $t \in \mathcal{P}_{\boldsymbol{\tau}}$, a simple evaluation through (27) of the contribution of each subinterval $I^n_{\boldsymbol{\tau}}$ to the integral gives

$$d^2(U_{\boldsymbol{\tau}}(T),u(T)) \le d^2(U^0_{\boldsymbol{\tau}},u^0) + \sum_{n=1}^N \tau_n^2 \mathscr{E}^n_{\boldsymbol{\tau}}. \tag{33}$$

Cancellation effects and rate of convergence. Choosing $V := U^{n-1}_{\boldsymbol{\tau}}$ in (22) we find

$$\frac{1}{\tau_n} d^2(U^n_{\boldsymbol{\tau}},U^{n-1}_{\boldsymbol{\tau}}) + \phi(U^n_{\boldsymbol{\tau}}) \le \phi(U^{n-1}_{\boldsymbol{\tau}}); \tag{34}$$

this shows that $\mathscr{E}^n_{\boldsymbol{\tau}} \ge 0$ and that $n \mapsto \phi(U^n_{\boldsymbol{\tau}})$ is decreasing; we get

$$\sum_{n=1}^N \tau_n^2 \mathscr{E}^n_{\boldsymbol{\tau}} \le \sum_{n=1}^n \tau_n \Big(\phi(U^{n-1}_{\boldsymbol{\tau}}) - \phi(U^n_{\boldsymbol{\tau}})\Big) \le |\boldsymbol{\tau}| \Big(\phi(u^0_{\boldsymbol{\tau}}) - \phi(U^N_{\boldsymbol{\tau}})\Big) \tag{35}$$

which proves (14). (15) requires a more refined argument. We start by writing (22) at the time step $n-1$ with the choice $V := U^n_{\boldsymbol{\tau}}$

$$\begin{aligned} &\phi(U^{n-1}_{\boldsymbol{\tau}}) - \phi(U^n_{\boldsymbol{\tau}}) \\ &\qquad \le \frac{1}{2\tau_{n-1}} \Big(d^2(U^{n-2}_{\boldsymbol{\tau}},U^n_{\boldsymbol{\tau}}) - d^2(U^{n-1}_{\boldsymbol{\tau}},U^n_{\boldsymbol{\tau}}) - d^2(U^{n-1}_{\boldsymbol{\tau}},U^{n-2}_{\boldsymbol{\tau}})\Big). \end{aligned} \tag{36}$$

Now we observe that the triangle and the Young inequalities yield

$$\begin{aligned} d^2(U^{n-2}_{\boldsymbol{\tau}},U^n_{\boldsymbol{\tau}}) &\le \Big(d(U^{n-2}_{\boldsymbol{\tau}},U^{n-1}_{\boldsymbol{\tau}}) + d(U^{n-1}_{\boldsymbol{\tau}},U^n_{\boldsymbol{\tau}})\Big)^2 \\ &\le \big(1 + \tfrac{\tau_n}{\tau_{n-1}}\big) d^2(U^{n-2}_{\boldsymbol{\tau}},U^{n-1}_{\boldsymbol{\tau}}) + \big(1 + \tfrac{\tau_{n-1}}{\tau_n}\big) d^2(U^{n-1}_{\boldsymbol{\tau}},U^n_{\boldsymbol{\tau}}). \end{aligned} \tag{37}$$

By substituting this inequality in (36), we obtain after a division by τ_n

$$\frac{\phi(U^{n-1}_{\boldsymbol{\tau}}) - \phi(U^n_{\boldsymbol{\tau}})}{\tau_n} \le \frac{d^2(U^{n-1}_{\boldsymbol{\tau}},U^n_{\boldsymbol{\tau}})}{2\tau_n^2} + \frac{d^2(U^{n-1}_{\boldsymbol{\tau}},U^{n-2}_{\boldsymbol{\tau}})}{2\tau_{n-1}^2}. \tag{38}$$

If we insert (38) into the expression of $\mathscr{E}^n_{\boldsymbol{\tau}}$ we get

$$\mathscr{E}^n_{\boldsymbol{\tau}} \le \frac{d^2(U^{n-1}_{\boldsymbol{\tau}},U^{n-2}_{\boldsymbol{\tau}})}{2\tau_{n-1}^2} - \frac{d^2(U^{n-1}_{\boldsymbol{\tau}},U^n_{\boldsymbol{\tau}})}{2\tau_n^2} \tag{39}$$

and therefore, by (13) and the positivity of $\mathscr{E}^n_{\boldsymbol{\tau}}$, we eventually find

$$\sum_{n=1}^N \tau_n^2 \mathscr{E}^n_{\boldsymbol{\tau}} \le |\boldsymbol{\tau}|^2 \sum_{n=1}^N \mathscr{E}^n_{\boldsymbol{\tau}} \le |\boldsymbol{\tau}|^2 \Big(\frac{\phi(U^0_{\boldsymbol{\tau}}) - \phi(U^1_{\boldsymbol{\tau}})}{\tau_1} - \frac{d^2(U^1_{\boldsymbol{\tau}},U^0_{\boldsymbol{\tau}})}{2\tau_1^2}\Big). \tag{40}$$

The modulus of the gradient of a convex function. For convex functions the modulus of the gradient can also be characterized as

$$|\nabla\phi(w)| = \sup\Big\{\frac{\phi(w)-\phi(v)}{d(w,v)} : v \in \mathscr{S}, v \neq w, \phi(v) \leq \phi(w)\Big\}; \tag{41}$$

thus, since $\phi(U^1_{\boldsymbol\tau}) \leq \phi(U^0_{\boldsymbol\tau})$, we get

$$\tfrac{\phi(U^0_{\boldsymbol\tau})-\phi(U^1_{\boldsymbol\tau})}{\tau_1} - \tfrac{d^2(U^1_{\boldsymbol\tau},U^0_{\boldsymbol\tau})}{2\tau_1^2} \leq |\nabla\phi(U^0_{\boldsymbol\tau})|\tfrac{d(U^1_{\boldsymbol\tau},U^0_{\boldsymbol\tau})}{\tau_1} - \tfrac{d^2(U^1_{\boldsymbol\tau},U^0_{\boldsymbol\tau})}{2\tau_1^2} \leq \tfrac{1}{2}|\nabla\phi(U^0_{\boldsymbol\tau})|^2.$$

General metric space and the application to the porous medium equation

It should be clear from the above proof that

Corollary 2.2. *The statement of Theorem* 2.1 *holds even if* $(\mathscr{S}, d)$ *is a metric space,* ϕ *is a l.s.c. (extended) real functional defined on* $\mathscr{S}$ *whose gradient norm is defined by* (41)*,* u *is a solution of the "metric evolution inequality"* (20)*,* $U_{\boldsymbol\tau}$ *solves the variational algorithm* (8a,b)*, and* (11a,b) *hold.*

If we want to apply the previous result to the example of (9a,b,c,d), for a triple $w, v_0, v_1 \in \mathscr{P}^2_a(\mathbb{R}^k)$ we should exhibit a connecting curve v_t, $t \in [0,1]$, which satisfies (11a,b). Let us denote by $\boldsymbol{r}_0, \boldsymbol{r}_1$ the optimal transportation maps between w and v_0, v_1 respectively (cf. (9b)); we define the curve v_t by duality with continuous functions

$$\int_{\mathbb{R}^k} \zeta(y) v_t(y)\, dy := \int_{\mathbb{R}^k} \zeta((1-t)\boldsymbol{r}_0(x) + t\boldsymbol{r}_1(x)) w(x)\, dx \quad \forall\, \zeta \in C^0_c(\mathbb{R}^k), \tag{42}$$

and it is possible to prove [2] that (11a,b) are satisfied. Thus we obtain

Corollary 2.3. *Let* u *be the solution of the porous medium equation* (9d) *and let* $U_{\boldsymbol\tau}$ *be the discrete solution obtained by solving the recursive minimization scheme* (8a,b) *with the choices of* d, ϕ *corresponding to* (9b,c)*. Then the* a posteriori *estimates of Theorem* 2.1 *hold; in particular, if* $U^0_{\boldsymbol\tau} = u^0$

$$\sup_{t \in \mathcal{P}_{\boldsymbol\tau}} W^2(u(t), U_{\boldsymbol\tau}(t)) \leq |\boldsymbol\tau| \frac{1}{m-1} \int_{\mathbb{R}^k} (u^0(x))^m\, dx, \tag{43}$$

and, if $u^{m-1/2} \in H^1(\mathbb{R}^k)$*,*

$$\sup_{t \in \mathcal{P}_{\boldsymbol\tau}} W^2(u(t), U_{\boldsymbol\tau}(t)) \leq \frac{|\boldsymbol\tau|^2}{2} \frac{m^2}{(m-1/2)^2} \int_{\mathbb{R}^k} \Big|\nabla[(u^0(x))^{m-1/2}]\Big|^2\, dx. \tag{44}$$

References

[1] L. Ambrosio, *Minimizing movements*, Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5) **19** (1995), 191–246.

[2] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows of probability measures*, in preparation (2003).

[3] C. Baiocchi, *Discretization of evolution variational inequalities*, Partial differential equations and the calculus of variations, Vol. I (F. Colombini, A. Marino, L. Modica, and S. Spagnolo, eds.), Birkhäuser Boston, Boston, MA, 1989, pp. 59–92.

[4] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Publishing Co., Amsterdam, 1973, North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).

[5] M. G. Crandall and T. M. Liggett, *Generation of semi-groups of nonlinear transformations on general Banach spaces*, Amer. J. Math. **93** (1971), 265–298.

[6] E. De Giorgi, *New problems on minimizing movements*, Boundary Value Problems for PDE and Applications (C. Baiocchi and J. L. Lions, eds.), Masson, 1993, pp. 81–98.

[7] E. De Giorgi, A. Marino, and M. Tosques, *Problems of evolution in metric spaces and maximal decreasing curve*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8) **68** (1980), no. 3, 180–187.

[8] K. Eriksson, C. Johnson, and S. Larsson, *Adaptive finite element methods for parabolic problems VI: Analytic semigroups*, SIAM J. Numer. Anal. **35** (1998), 1315–1325.

[9] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), no. 1, 1–17 (electronic).

[10] J. Jost, *Nonpositive curvature: geometric and analytic aspects*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1997.

[11] U. F. Mayer, *Gradient flows on nonpositively curved metric spaces and harmonic maps*, Comm. Anal. Geom. **6** (1998), no. 2, 199–253.

[12] R. J. McCann, *A convexity principle for interacting gases*, Adv. Math. **128** (1997), no. 1, 153–179.

[13] R. H. Nochetto and G. Savaré, *Nonlinear evolution governed by accretive operators: error control and applications*, to appear (2003).

[14] R. H. Nochetto, G. Savaré, and C. Verdi, *A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations*, Comm. Pure Appl. Math. **53** (2000), no. 5, 525–589.

[15] J. Rulla, *Error analysis for implicit approximations to solutions to Cauchy problems*, SIAM J. Numer. Anal. **33** (1996), 68–87.

[16] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations **26** (2001), no. 1-2, 101–174.

[17] C. Villani, *Topics in optimal oransportation*, Graduate Studies in Mathematics, **58**, AMS (2003).

[18] A. Visintin, *Models of phase transitions*, Progress in Nonlinear Differential Equations and Their Applications, vol. 28, Birkhäuser, Boston, 1996.

Department of Mathematics,
Università di Pavia,
via Ferrata, 1
I-27100 Pavia, Italy
E-mail address: savare@imati.cnr.it
URL: http://www.imati.cnr.it/~savare