A variational principle for gradient flows in metric spaces

Riccarda Rossi [a], Giuseppe Savaré [b], Antonio Segatti [b], Ulisse Stefanelli [c]

[a] *Dipartimento di Matematica, Università di Brescia, v. Valotti 9 I-25133 Brescia, Italy*

[b] *Dipartimento di Matematica, Università di Pavia, v. Ferrata 1, I-27100 Pavia, Italy*

[c] *IMATI - CNR, v. Ferrata 1, I-27100 Pavia, Italy and WIAS, Mohrenstr. 39, D-10117 Berlin, Germany*

Received *****; accepted after revision +++++

Presented by

Abstract

We present a novel variational approach to gradient-flow evolution in metric spaces. In particular, we advance a functional defined on entire trajectories, whose minimizers converge to curves of maximal slope for geodesically convex energies. The crucial step of the argument is the reformulation of the variational approach in terms of a dynamic programming principle, and the use of the corresponding Hamilton-Jacobi equation. The result is applicable to a large class of nonlinear evolution PDEs including nonlinear drift-diffusion, Fokker-Planck, and heat flows on metric-measure spaces.

To cite this article: R. Rossi, G. Savaré, A. Segatti, and U. Stefanelli ...

Résumé

Un principe variationnel pour les équations de flot gradient dans les espaces métriques. Nous présentons une nouvelle approche variationnelle pour l'étude d'évolution de *flot gradient* dans des espaces métriques. En particulier, nous proposons une fonctionnelle définie sur des trajectoires entières. Nous démontrons que les minimums de cette fonctionnelle convergent vers des courbes de descente maximale dans le cas d'une énergie géodésiquement convexe. Le point crucial de l'argument est la reformulation de l'approche variationnelle en terms du principe de la programmation dynamique. Ce resultat peut s'appliquer à une large classe d'évolution nonlineaires qui peuvent être reformulées comme des flots gradient dans des espaces métriques de Wasserstein.

Pour citer cet article : R. Rossi, G. Savaré, A. Segatti, and U. Stefanelli ...

1. The variational principle

Gradient flows are the paradigm of parabolic evolution and, as such, have received constant attention since the end of the 60s. They arise almost ubiquitously in connection with applications such as heat conduction, the Stefan problem, the Hele-Shaw cell, porous media, parabolic variational inequalities, some classes of ODEs with obstacles, degenerate parabolic PDEs, and the mean curvature flow for Cartesian graphs, just to mention a few. More recently, following the pioneering work by Otto [10], an even larger class of PDE problems including transport, nonlinear drift-diffusion, and Fokker-Planck equations have been translated into gradient flows, in the framework of probability spaces endowed with the Wasserstein metric. In this connection, the reader is referred to the monograph by Ambrosio, Gigli, & Savaré [2] for a collection of results.

The aim of this note is to illustrate a novel variational view at gradient flows in metric spaces. Let (X,d) be a complete and separable metric space and the functional

$$\phi : X \to (-\infty,\infty] \text{ be proper, geodesically convex, and with compact sublevels.} \tag{1}$$

Geodesic convexity means that every couple of points in $D(\phi) := \{\phi < \infty\}$ can be connected by a minimal and constant speed geodesic $\gamma : [0,1] \to X$ (thus satisfying $d(\gamma(s),\gamma(t)) = (t-s)d(\gamma(0),\gamma(1))$), such that $\phi(\gamma(t)) \le (1-t)\phi(\gamma(0))+t\phi(\gamma(1))$ for all $t \in [0,1]$. Note that most of the assumptions in (1) are here chosen for the sake of presentation simplicity and may be relaxed. In particular, the geodesic convexity requirement in (1) can be weakened, and geodesically λ-convex functional can be considered, see [9].

We shall consider the *global-in-time* functionals $I^\varepsilon : AC^2(\mathbb{R}_+;X) \to (-\infty,\infty]$ given, for $\varepsilon > 0$, by

$$I^\varepsilon(u) = \int_0^\infty \mathrm{e}^{-t/\varepsilon} \left(\frac{1}{2}|u'|^2(t)+\frac{1}{\varepsilon}\phi(u(t))\right) \mathrm{d}t. \tag{2}$$

Here, $\mathbb{R}_+ := [0,\infty)$, and $AC^2(\mathbb{R}_+;X)$ is the set of absolutely continuous curves $t \in \mathbb{R}_+ \mapsto u(t) \in X$, for which the *metric derivative* $t \mapsto |u'|(t) := \lim_{s\to t} d(u(t),u(s))/|t-s|$ exists a.e. and belongs to $L^2(\mathbb{R}_+)$ [2].

One can check that, for all $\varepsilon > 0$ and $\bar u \in D(\phi)$ the problem

$$\min_{u\in K(\bar u)} I^\varepsilon(u) \quad \text{with } K(\bar u) := \{u \in AC^2(\mathbb{R}_+;X) \, : \, u(0) = \bar u\}, \tag{3}$$

admits a solution, see [9]. Our aim is to show the connection between this minimization problem and *curves of maximal slope (for the functional ϕ, with respect to the upper gradient $|\partial\phi|$ and originating from $\bar u$)*. The latter are trajectories $u \in AC^2(\mathbb{R}_+;X)$ such that $u(0) = \bar u$ and

$$\phi(u(t)) + \frac{1}{2}\int_0^t |u'|^2(t)\mathrm{d}t + \frac{1}{2}\int_0^t |\partial\phi|^2(u(t))\mathrm{d}t = \phi(\bar u) \quad \text{for all } t \ge 0. \tag{4}$$

In (4), the symbol $|\partial\phi|(u) := \limsup_{v\to u} (\phi(u)-\phi(v))^+/d(u,v)$, for $u \in D(\phi)$, stands for the *local descending slope* of ϕ at u [2]. This concept is the natural analogue of the *gradient of the energy* in a metric setting, where, in absence of a linear structure, one has to resort to suitable surrogates of gradients and time derivatives. In particular, under assumptions (1), if X is a Hilbert space endowed with its strong metric, curves of maximal slopes coincide with classical gradient flows, i.e. solutions of the differential inclusion $u'(t) \in -\partial\phi(u(t))$ [8]. Our main result reads as follows.

Theorem 1.1 (Variational principle) *As $\varepsilon \downarrow 0$ the minimizers in (3) admit a subsequence which locally uniformly converges to a curve of maximal slope.*

This convergence result entails the possibility of reformulating the differential problem (4) as a (limit of a class of) minimization problem(s). In particular, it paves the way to the application of the specific tools of the Calculus of Variations to (4), especially the Direct Method, relaxation, and Γ-convergence. As a by-product of Theorem 1.1, we have an alternative existence proof for curves of maximal slope (see [2]).

We recall that the variational approach to gradient flows via the minimization of I_ε has been firstly applied to mean curvature evolution by Ilmanen [5]. Then, two examples of relaxation of gradient flows via I^ε are provided by Conti & Ortiz [4] in the context of microstructure evolution. In the Hilbert case, Theorem 1.1, along with a number of related results, has been obtained [8]. This variational approach has been applied to rate-independent evolution by Mielke & Ortiz [6] and further detailed in [7], whereas the doubly nonlinear case is addressed in [1].

We shift here the attention to the metric case. As already mentioned, our interest for gradient flows in metric spaces is not at all academical, but rather motivated by possible applications to evolution PDEs with nonnegative solutions $u : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}_+$, in the general form

$$\partial_t u - \nabla \cdot \big(u\nabla(\delta_u\phi(u))\big) = 0 \quad \text{in } \mathbb{R}^d \times \mathbb{R}_+. \tag{5}$$

where $\delta_u\phi(u)$ is the suitably defined first variation of an integral functional, resulting from the linear combination of the terms $\mathcal{U}(u) = \int_{\mathbb{R}^d} U(u(x))\,\mathrm{d}x$, $\mathcal{V}(u) = \int_{\mathbb{R}^d} V(x)u(x)\,\mathrm{d}x$, $\mathcal{W}(u) = \int_{\mathbb{R}^d\times\mathbb{R}^d} W(x-y)u(x)u(y)\,\mathrm{d}x\,\mathrm{d}y$ under qualified smoothness and growth assumptions. The functionals $\mathcal{U}$, $\mathcal{V}$, and $\mathcal{W}$ are generally referred to as the *internal*, the *potential*, and the *interaction* energies, respectively. In particular, the choices $F = \mathcal{U}+\mathcal{V}$, $U(r) = r\log r$, and $F = \mathcal{U}$, $U(r) = r^m/(m-1)$, $m \ge 1-1/d$, respectively yield the Fokker-Planck and the nonlinear diffusion equations. Transport and nonlinear drift-diffusion equations (with or without non-local interactions) can be considered as well. The aforementioned PDE is by now classically reformulated as a gradient flow equation, in the metric space $\mathcal{P}_2(\mathbb{R}^d)$ of probability measures with finite second moment, endowed with the Wasserstein 2-metric. Another possible application concerns the heat flow in a Polish metric-measure space (M,d,m) satisfying the Lott-Sturm-Villani condition $CD(K,\infty)$ [3]: in this case $X = \mathcal{P}_2(M)$, $\phi(\mu) = \mathrm{Ent}_m(\mu)$ is the relative entropy functional, and the family of minimizers u_ε converge to the unique solution $\mu = \rho m$, $\partial_t\rho - \Delta_{m,d}\rho = 0$ (see [3] for the definition of the operator $\Delta_{m,d}$).

2. The metric inner variation equation

In order to gain some insight into the convergence result of Theorem 1.1, let us present the specific form of the Euler-Lagrange equation for the minimum problem (3).

Lemma 2.1 (Metric inner variations) *Let u_ε minimize I^ε on $K(\bar u)$. Then, u_ε fulfills*

$$|u_\varepsilon'|^2(t) + \frac{\mathrm{d}}{\mathrm{d}t}\Big(\phi(u_\varepsilon(t)) - \frac{\varepsilon}{2}|u_\varepsilon'|^2(t)\Big) = 0 \quad \textit{for a.a. } t>0. \tag{6}$$

In the limiting case $\varepsilon = 0$, relation (6) represents the balance between the variation in energy (density) $(\mathrm{d}/\mathrm{d}t)\phi(u_\varepsilon(t))$ and the dissipated energy (density) $-|u_\varepsilon'|^2(t)$. For $\varepsilon > 0$, the extra term $(\mathrm{d}/\mathrm{d}t)(\varepsilon/2)|u_\varepsilon'|^2(t)$ corresponds to a *nonlocal-in-time* correction. If X is a Hilbert space, the Euler-Lagrange equation for I^ε reads [8]

$$-\varepsilon u_\varepsilon''(t) + u_\varepsilon'(t) + \partial\phi(u_\varepsilon(t)) \ni 0 \quad \text{for a.a. } t>0,$$

so that relation (6) stems by testing the latter by u_ε'. Eventually, minimizing I^ε basically corresponds in addressing an *elliptic-in-time* regularization of the original gradient flow evolution.

3. The Dynamic Programming principle

Let us introduce the *value function* $\bar u \in D(\phi) \mapsto V^\varepsilon(\bar u)$ defined by

$$V^\varepsilon(\bar u) := \min_{u\in K(\bar u)} I^\varepsilon(u) = I^\varepsilon(u_\varepsilon),$$

for any minimizer u_ε of problem (3). We have (cf. [9]) that V^ε is l.s.c., bounded from below, and that it monotonically converges to ϕ everywhere. In particular, $V^\varepsilon(\bar u) \le \phi(\bar u)$ for all $\bar u \in X$. Moreover, we have that

$$\big(u_\varepsilon \to u,\ \sup_\varepsilon \phi(u_\varepsilon) < \infty\big) \ \Rightarrow\ \phi(u) \le \liminf_{\varepsilon\searrow 0} V^\varepsilon(u_\varepsilon) \ \text{ and } \ \frac{1}{2}|\partial\phi|^2(u) \le \liminf_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\big(\phi(u_\varepsilon) - V^\varepsilon(u_\varepsilon)\big). \tag{7}$$

The crucial tool towards Theorem 1.1 is a metric version of the classical *Dynamic Programming principle.*

Proposition 3.1 (Dynamic Programming principle) *For every $T > 0$ there holds*

$$V^\varepsilon(\bar u) = \min_{u\in K(\bar u)}\left(\int_0^T \mathrm{e}^{-t/\varepsilon}\left(\frac{1}{2}|u'|^2(t) + \frac{1}{\varepsilon}\phi(u(t))\right)\mathrm{d}t + V^\varepsilon(u(T))e^{-T/\varepsilon}\right). \tag{8}$$

The core of the proof of Theorem 1.1 consists in working out the relations between the value function V^ε and the energy ϕ, relying on (8). In particular, one can prove that V^ε is continuous on the sublevels of ϕ, that the map $t \mapsto V^\varepsilon(u(t))$ is absolutely continuous, and we have the crucial relation [9, Prop. 2.8]

$$-\frac{\mathrm{d}}{\mathrm{d}t}V^\varepsilon(u_\varepsilon(t)) = \frac{1}{2}|u_\varepsilon'|^2(t) + \frac{1}{\varepsilon}\phi(u_\varepsilon(t)) - \frac{1}{\varepsilon}V^\varepsilon(u_\varepsilon(t)) \quad \text{for a.a. } t>0. \tag{9}$$

The role of the value function V^ε is further illustrated by observing that the minimizer u_ε is itself a curve of maximal slope for V^ε, with respect to the *conditioned local slope* $|\tilde\partial V^\varepsilon|$, which is defined at $u \in D(\phi)$ by

$$|\tilde\partial V^\varepsilon|(u) := \limsup_{v\to u,\,\phi(v)\to\phi(u)} \frac{(V^\varepsilon(u) - V^\varepsilon(v))^+}{d(u,v)} \le |\partial V^\varepsilon|(u)$$

(note that the lim sup is taken along sequences also "converging in energy"). Namely, there holds

$$V^\varepsilon(u_\varepsilon(t)) + \frac{1}{2}\int_0^t |u_\varepsilon'|^2(t)\,\mathrm{d}t + \frac{1}{2}\int_0^t |\tilde\partial V^\varepsilon|^2(u_\varepsilon(t))\,\mathrm{d}t = V^\varepsilon(\bar u) \quad \text{for all } t>0.$$

This is indeed a consequence of (9), combined with the following relation (to be interpreted as a metric version of the *Hamilton-Jacobi equation* for V^ε):

$$2V^\varepsilon(u_\varepsilon(t)) + \varepsilon|\tilde\partial V^\varepsilon|^2(u_\varepsilon(t)) = 2\phi(u_\varepsilon(t)) \quad \text{for all } t>0.$$

4. Convergence proof

We comment here the line of the proof of Theorem 1.1, referring to [9] for all details. Let u_ε be a minimizer of I^ε on $K(\bar u)$. An important fact is that the function $t \mapsto \phi(u_\varepsilon(t))$ is (convex and) nonincreasing in $\mathbb{R}_+$. The convergence proof follows by passing to the limit as $\varepsilon \to 0$ in the integral of (9), namely

$$V^\varepsilon(u_\varepsilon(t)) + \frac{1}{2}\int_0^t |u_\varepsilon'|^2(s)\mathrm{d}s + \int_0^t \frac{1}{\varepsilon}\big(\phi(u_\varepsilon(s)) - V^\varepsilon(u_\varepsilon(s))\big)\mathrm{d}s = V^\varepsilon(\bar u) \quad \text{for all } t>0. \tag{10}$$

In particular, the latter and the fact that $V^\varepsilon(\bar u) \le \phi(\bar u)$ entail that $|u_\varepsilon'|$ is uniformly bounded in $L^2(\mathbb{R}_+)$ and nonincreasing in $\mathbb{R}_+$. Hence, by the compactness of the sublevels of ϕ (see (1)), one extracts a not relabeled subsequence such that $u_\varepsilon(t) \to u(t)$ in X for all $t \ge 0$ and $|u_\varepsilon'| \rightharpoonup m$ weakly in $L^2(\mathbb{R}_+)$, with $m \ge |u'|$ a.e. in $(0,+\infty)$. In order to pass to the lim inf in (10) we use the second of (7) and the fact that $V^\varepsilon(\bar u) \nearrow \phi(\bar u)$. Hence, the '$\le$' inequality in (4) is established. The missing '$\ge$' inequality follows directly from the fact that, under assumptions (1), the local slope $|\partial\phi|$ is a *strong upper gradient*, see [2, Def. 1.2.1].

Acknowledgements

R.R. and G.S. are partially supported by MIUR-PRIN 2008 grant *Optimal mass transportation, geometric and functional inequalities and applications*. U.S. is partly supported by FP7-IDEAS-ERC-StG grant 200947 *BioSMA*, CNR-AVČR Grant *SmartMath*, and the Alexander von Humboldt Foundation.

References

[1] G. Akagi and U. Stefanelli, Weighted energy-dissipation functionals for doubly nonlinear evolution, *J. Funct. Anal.*, , 260 (2011), 9:2541–2578.

[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008. Second edition.

[3] L. Ambrosio, N. Gigli, and G. Savaré, Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. Preprint arXiv:1106.2090 (2011).

[4] S. Conti and M. Ortiz, Minimum principles for the trajectories of systems governed by rate problems, *J. Mech. Phys. Solids*, 56 (2008), 1885–1904.

[5] T. Ilmanen, Elliptic regularization and partial regularity for motion by mean curvature, *Mem. Amer. Math. Soc.*, 108 (1994), 520:x+90.

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

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

[8] A. Mielke and U. Stefanelli. Weighted energy-dissipation functionals for gradient flows, *ESAIM Control Optim. Calc. Var.*, 17 (2011), 52-85.

[9] R. Rossi, G. Savaré, A. Segatti and U. Stefanelli. Elliptic Regularization for gradient flows in metric spaces. In preparation, 2011.

[10] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26 (2001), 101–174.