LECTURE NOTES ON GRADIENT FLOWS AND OPTIMAL TRANSPORT

SARA DANERI AND GIUSEPPE SAVARÉ

ABSTRACT. We present a short overview on the strongest variational formulation for gradient flows of geodesically λ-convex functionals in metric spaces, with applications to diffusion equations in Wasserstein spaces of probability measures. These notes are based on a series of lectures given by the second author for the Summer School "Optimal transportation: Theory and applications" in Grenoble during the week of June 22-26, 2009.

INTRODUCTION

These notes are based on a series of lectures given by the second author for the Summer School "Optimal transportation: Theory and applications" in Grenoble during the week of June 22-26, 2009.

We try to summarize some of the main results concerning gradient flows of geodesically λ-convex functionals in metric spaces and applications to diffusion PDE's in the Wasserstein space of probability measures. Due to obvious space constraints, the theory and the references presented here are largely incomplete and should be intended as an over-simplified presentation of a quickly evolving subject. We refer to the books [3, 68] for a detailed account of the large literature available on these topics.

In the first section we collect some elementary and well known results concerning gradient flows of smooth convex functions in $\mathbb{R}^d$. We selected just a few topics, which are well suited for a "metric" formulation and provide a useful guide for the more abstract developments.

In the second section we present the main (and strongest) notion of gradient flow in metric spaces characterized by the solution of a metric *evolution variational inequality*: the aim here is to show the consequence of this definition, without any assumptions on the space and on the functional (except completeness and lower semicontinuity): we shall see that solutions to evolution variational inequalities enjoy nice stability, asymptotic, and regularization properties. We also investigate the relationships with two different approaches, *curves of maximal slope* and *minimizing movements*, and we discuss a first stability result with respect to perturbations of the generating functional with respect to Γ-convergence.

The third section is devoted to some fundamental generation results for gradient flows of geodesically λ-convex functionals: here we adopt the method of *minimizing movement* to construct suitable families of discrete approximating solutions and we show three basic convergence results.

Apart from Sections 2.6 (stability of gradient flows with respect to Γ-convergence of the functionals) and 3.1 (existence of curves of maximal slope), we made a substantial effort to avoid any compactness argument in the theory, which is mainly focused on purely metric arguments. So we will present a slightly relaxed version of the minimizing movement scheme, which is always solvable by invoking Ekeland's variational principle, and the main existence and generation results for λ-gradient flows rely on refined Cauchy estimates and crucial geometric assumptions on the distance of the metric space.

The last section is devoted to applications of the metric theory to evolution equations in the so called "Wasserstein spaces" $\mathscr{P}_2(X)$ of probability measures. We recall a few basic facts about such spaces, the characterization of geodesics and absolutely continuous curves, and some geometric properties of the Wasserstein distance. Three basic examples of (or, better, displacement-) λ-convex functionals in $\mathscr{P}_2(\mathbb{R}^d)$ are presented, together with the evolutionary PDE's they are associated with. A short account of possible extensions of the theory to measure-metric spaces concludes the notes.

1. Gradient flows for smooth λ-convex functions in the Euclidean space

In this section we recall some simple properties of the gradient flow of a C^2 function $\varphi : \mathbb{R}^d \to \mathbb{R}$ satisfying the global lower bound $D^2\varphi \geq \lambda I$ for some $\lambda \in \mathbb{R}$. We will focus on those aspects which rely just on the "metric" structure of $\mathbb{R}^d$ and therefore could make sense in more general metric spaces. We denote by $\mathsf{d}(u,v) = |u-v|$ the Euclidean distance on $\mathbb{R}^d$ induced by the scalar product $\langle \cdot, \cdot \rangle$.

Remark 1.1 (A few basic facts about λ-convex functions). We will extensively use the following well known equivalent characterizations of a λ-convex function $\varphi : \mathbb{R}^d \to \mathbb{R}$ (here x, x_0, x_1 are arbitrary points in $\mathbb{R}^d$)

Hessian inequality:

$$D^2\varphi(x) \geq \lambda I \quad \text{i.e.} \quad \langle D^2\varphi(x)\xi, \xi \rangle \geq \lambda |\xi|^2 \quad \text{for every } \xi \in \mathbb{R}^d. \tag{1.1a}$$

λ-monotonicity of $\nabla\varphi$:

$$\langle \nabla\varphi(x_0) - \nabla\varphi(x_1), x_0 - x_1 \rangle \geq \lambda |x_0 - x_1|^2. \tag{1.1b}$$

λ-convexity inequality:

$$\varphi(x_\theta) \leq (1-\theta)\varphi(x_0) + \theta\varphi(x_1) - \frac{\lambda}{2}\theta(1-\theta)|x_0 - x_1|^2 \quad x_\theta := (1-\theta)x_0 + \theta x_1, \, \theta \in [0,1]. \tag{1.1c}$$

Sub-gradient inequality:

$$\langle \nabla\varphi(x_1), x_1 - x_0 \rangle - \frac{\lambda}{2}|x_1 - x_0|^2 \geq \varphi(x_1) - \varphi(x_0) \geq \langle \nabla\varphi(x_0), x_1 - x_0 \rangle + \frac{\lambda}{2}|x_1 - x_0|^2. \tag{1.1d}$$

Notice that

$$\varphi \text{ is } \lambda\text{-convex if and only if } \tilde{\varphi}(x) := \varphi(x) - \frac{\lambda}{2}|x|^2 \text{ is convex.} \tag{1.1e}$$

In particular, there exist constants $a \in \mathbb{R}$, $\boldsymbol{b} \in \mathbb{R}^d$ such that

$$\varphi(x) \geq a + \langle \boldsymbol{b}, x \rangle + \frac{\lambda}{2}|x|^2. \tag{1.1f}$$

Definition 1.2 (Gradient flow). *The gradient flow of φ is the family of maps*

$$\mathsf{S}_t : \mathbb{R}^d \to \mathbb{R}^d, \quad t \in [0, +\infty),$$

characterized by the following property: for every $u_0 \in \mathbb{R}^d$, $\mathsf{S}_0(u_0) := u_0$ and the curve $u_t := \mathsf{S}_t(u_0)$, $t \in (0, +\infty)$, is the unique C^1 solution of the Cauchy problem

$$\frac{\mathrm{d}}{\mathrm{d}t}u_t = -\nabla\varphi(u_t) \quad \text{in } (0, +\infty), \quad \lim_{t\downarrow 0} u_t = u_0. \tag{1.2}$$

By the standard Cauchy-Lipschitz theory and the *a priori* estimates we will show in the next theorem, for every initial datum $u_0 \in \mathbb{R}^d$ equation (1.2) admits a unique global solution so that the family S_t, $t \in [0, +\infty)$, is a continuous semigroup of Lipschitz maps, thus satisfying

$$\mathsf{S}_{t+h}(u_0) = \mathsf{S}_t(\mathsf{S}_h(u_0)), \quad \lim_{t\downarrow 0} \mathsf{S}_t(u_0) = \mathsf{S}_0(u_0) = u_0 \quad \text{for every } u_0 \in \mathbb{R}^d. \tag{1.3}$$

1.1. Basic estimates.

Theorem 1.3 (Basic differential estimates). *Let us assume that $\varphi \in C^2(\mathbb{R}^d)$ is λ-convex; if $u : [0, +\infty) \to \mathbb{R}^d$ is a solution of (1.2) then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}|u_t - v|^2 + \frac{\lambda}{2}|u_t - v|^2 = \mathrm{e}^{-\lambda t}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{\lambda t}\frac{1}{2}|u_t - v|^2\right) \leq \varphi(v) - \varphi(u_t) \quad \text{for every } v \in \mathbb{R}^d, \quad \text{(EVI}_\lambda\text{)}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(u_t) = -|u_t'|^2 = -|\nabla\varphi(u_t)|^2 \leq 0, \tag{EI}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{2\lambda t}|\nabla\varphi(u_t)|^2\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{2\lambda t}|u_t'|^2\right) \leq 0; \tag{SI$_\lambda$}$$

moreover, if v is another solution to (1.2) *then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{\lambda t}\,|u_t - v_t|\right) \le 0. \tag{Cont$_\lambda$}$$

Proof. We sketch here the easy calculations.

For the *evolution variational inequality* (EVI$_\lambda$):

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}|u_t - v|^2 = \langle u_t', u_t - v\rangle \stackrel{(1.2)}{=} \langle \nabla\varphi(u_t), v - u_t\rangle \stackrel{(1.1\mathrm{d})}{\le} \varphi(v) - \varphi(u_t) - \frac{\lambda}{2}|u_t - v|^2.$$

The *energy identity* (EI):

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(u_t) = \langle \nabla\varphi(u_t), u_t'\rangle \stackrel{(1.2)}{=} -|\nabla\varphi(u_t)|^2 \stackrel{(1.2)}{=} -|u_t'|^2.$$

The *slope inequality* (SI$_\lambda$):

$$\frac{\mathrm{d}}{\mathrm{d}t}|\nabla\varphi(u_t)|^2 = 2\langle \mathrm{D}^2\varphi(u_t)\nabla\varphi(u_t), u_t'\rangle \stackrel{(1.2)}{=} -2\langle \mathrm{D}^2\varphi(u_t)\nabla\varphi(u_t), \nabla\varphi(u_t)\rangle \stackrel{(1.1\mathrm{a})}{\le} -2\lambda|\nabla\varphi(u_t)|^2.$$

The *λ-contraction* property (Cont$_\lambda$):

$$\frac{\mathrm{d}}{\mathrm{d}t}|u_t - v_t|^2 = 2\langle u_t' - v_t', u_t - v_t\rangle \stackrel{(1.2)}{=} -2\langle \nabla\varphi(u_t) - \nabla\varphi(v_t), u_t - v_t\rangle \stackrel{(1.1\mathrm{b})}{\le} -2\lambda|u_t - v_t|^2. \qquad \square$$

In order to write in a simple way suitable integrated versions of the previous inequalities, we set

$$\mathsf{E}_\lambda(t) := \int_0^t \mathrm{e}^{\lambda r}\,\mathrm{d}r = \begin{cases} \frac{\mathrm{e}^{\lambda t}-1}{\lambda} & \text{if } \lambda \ne 0, \\ t & \text{if } \lambda = 0. \end{cases} \tag{1.4}$$

Corollary 1.4 (Pointwise and integral inequalities)**.** *If $u : [0, +\infty) \to \mathbb{R}^d$ is a solution to* (1.2) *then*

$$\frac{\mathrm{e}^{\lambda t}}{2}|u_t - v|^2 + \mathsf{E}_\lambda(t)\big(\varphi(u_t) - \varphi(v)\big) + \frac{\big(\mathsf{E}_\lambda(t)\big)^2}{2}|\nabla\varphi(u_t)|^2 \le \frac{1}{2}|u_0 - v|^2 \quad \text{for every } v \in \mathbb{R}^d, \tag{1.5}$$

$$\varphi(u_t) + \frac{1}{2}\int_0^t \left(|u_r'|^2 + |\nabla\varphi(u_r)|^2\right)\mathrm{d}r = \varphi(u_0), \tag{1.6}$$

$$|\nabla\varphi(u_t)| \le \mathrm{e}^{-\lambda t}|\nabla\varphi(u_0)|; \tag{1.7}$$

moreover, if v is another solution to (1.2) *then*

$$|u_t - v_t| \le \mathrm{e}^{-\lambda t}|u_0 - v_0|. \tag{1.8}$$

In particular, when $\lambda > 0$, φ admits a unique minimum point $\bar{u}$ and

$$\frac{\lambda}{2}|u_t - \bar{u}|^2 \le \varphi(u_t) - \varphi(\bar{u}) \le \frac{1}{2\lambda}|\nabla\varphi(u_t)|^2 \tag{1.9}$$

$$|u_t - \bar{u}| \le \mathrm{e}^{-\lambda t}|u_0 - \bar{u}|, \quad \varphi(u_t) - \varphi(\bar{u}) \le \mathrm{e}^{-2\lambda t}\big(\varphi(u_0) - \varphi(\bar{u})\big). \tag{1.10}$$

Proof. We have just to check (1.5): if A_t denotes the quantity in the left-hand side, we show that A_t is nonincreasing. A differentiation in time yields

$$\frac{\mathrm{d}}{\mathrm{d}t}A_t = \mathrm{e}^{\lambda t}\left(\frac{\lambda}{2}|u_t - v|^2 + \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}|u_t - v|^2 + \varphi(u_t) - \varphi(v) + \mathsf{E}_\lambda(t)|\nabla\varphi(u_t)|^2\right)$$

$$+ \mathsf{E}_\lambda(t)\frac{\mathrm{d}}{\mathrm{d}t}\varphi(u_t) + \frac{\big(\mathsf{E}_\lambda(t)\big)^2}{2}\frac{\mathrm{d}}{\mathrm{d}t}|\nabla\varphi(u_t)|^2\Big)$$

$$\stackrel{\text{(EVI}_\lambda)}{\le} \mathsf{E}_\lambda(t)\left(\mathrm{e}^{\lambda t}|\nabla\varphi(u_t)|^2 + \frac{\mathrm{d}}{\mathrm{d}t}\varphi(u_t) + \frac{\mathsf{E}_\lambda(t)}{2}\frac{\mathrm{d}}{\mathrm{d}t}|\nabla\varphi(u_t)|^2\right)$$

$$\stackrel{\text{(EI)}}{=} \mathsf{E}_\lambda(t)\left(\big(\mathrm{e}^{\lambda t} - 1\big)|\nabla\varphi(u_t)|^2 + \frac{\mathsf{E}_\lambda(t)}{2}\frac{\mathrm{d}}{\mathrm{d}t}|\nabla\varphi(u_t)|^2\right)$$

$$\stackrel{\text{(SI}_\lambda)}{\le} \mathsf{E}_\lambda(t)\left(\big(\mathrm{e}^{\lambda t} - 1 - \lambda\mathsf{E}_\lambda(t)\big)|\nabla\varphi(u_t)|^2\right) \stackrel{(1.4)}{=} 0. \qquad \square$$

In terms of the maps S_t, (1.8) yields the λ-contraction estimate

$$\mathsf{d}(\mathsf{S}_t(u_0), \mathsf{S}_t(v_0)) \leq \mathrm{e}^{-\lambda t}\mathsf{d}(u_0, v_0) \quad \text{for every } u_0, v_0 \in \mathbb{R}^d,\ t \geq 0, \tag{1.11}$$

thus showing the Lipschitz property of S_t and the uniqueness and continuous dependence w.r.t. the initial data of the solutions of (1.2).

1.2. **Approximation by the Implicit Euler scheme.** One of the simplest but very useful ways to construct discrete approximations of the solution to (1.2) (and to show its existence by a limiting process) is given by the *implicit Euler scheme.*

For a given time step $\tau > 0$ we consider the associated uniform partition of $[0, +\infty)$

$$\mathcal{P}_\tau := \{0 = t_\tau^0 < t_\tau^1 < ... < t_\tau^n < ...\}, \quad t_\tau^n := n\tau, \tag{1.12}$$

and we look for a discrete sequence $(U_\tau^n)_{n \in \mathbb{N}}$ whose value U_τ^n should provide an effective approximation of $u(t_\tau^n)$. U_τ^n are defined recursively, starting from a suitable choice of $U_\tau^0 \approx u_0$, by solving at each step the equation of the Euler scheme

$$\frac{U_\tau^n - U_\tau^{n-1}}{\tau} = -\nabla\varphi(U_\tau^n) \quad n = 1, 2, \cdots, \tag{1.13}$$

or, equivalently,

$$U_\tau^n = J_\tau(U_\tau^{n-1}), \quad J_\tau := (I + \tau\nabla\varphi)^{-1}. \tag{1.14}$$

Existence of a discrete approximating solution can be easily obtained by looking for the minimizers of the function

$$U \mapsto \Phi(\tau, U_\tau^{n-1}; U) := \frac{1}{2\tau}\left|U - U_\tau^{n-1}\right|^2 + \varphi(U). \tag{1.15}$$

In fact, it is immediate to check that any minimizer U_τ^n of (1.15) solves (1.13); moreover, the function defined by (1.15) is $(\tau^{-1}+\lambda)$-convex and therefore it admits a unique minimizer whenever $\tau^{-1} > -\lambda$.

Denoting by $U_\tau : [0, +\infty) \to \mathbb{R}^d$ the piecewise linear interpolant of the discrete values $(U_\tau^n)_{n \in \mathbb{N}}$ on the grid $\mathcal{P}_\tau$, defined by

$$U_\tau(t) := \frac{t - t_\tau^{n-1}}{\tau}U_\tau^{n-1} + \frac{t_\tau^n - t}{\tau}U_\tau^n \quad \text{if } t \in [t_\tau^{n-1}, t_\tau^n], \tag{1.16}$$

one expects that $U_\tau(t)$ converges to the solution u_t to (1.2) as $\tau \downarrow 0$.

Theorem 1.5. *If* $\lim_{\tau \downarrow 0} U_\tau^0 = u_0$ *then the family of piecewise linear interpolants* $(U_\tau)_{\tau>0}$ *satisfies the Cauchy condition as* $\tau \downarrow 0$ *with respect to the uniform convergence on each compact interval* $[0, T]$, $T > 0$; *its unique limit is the solution* u_t *of* (1.2). *Moreover, for every* $T > 0$ *there exists a universal constant* $C(\lambda, T)$ *such that*

$$\sup_{t \in [0,T]} |u_t - U_\tau(t)| \leq |u_0 - U_\tau^0| + C(\lambda, T)|\nabla\varphi(u_0)|\,\tau. \tag{1.17}$$

In particular, when $\lambda = 0$ *we can choose* $C = \frac{1}{\sqrt{2}}$, *independent of* T.

Remarks about the proof. In the present finite dimensional smooth setting, the proof of the convergence of U_τ is not difficult: considering e.g. the case $\lambda = 0$, we can apply the contraction property of the map J_τ defined by (1.14)

$$|J_\tau(x) - J_\tau(y)| \leq |x - y| \quad \text{for every } x, y \in \mathbb{R}^d, \tag{1.18}$$

to obtain the uniform bound

$$\tau^{-1}|U_\tau^n - U_\tau^{n-1}| = |\nabla\varphi(U_\tau^n)| \leq |\nabla\varphi(U_\tau^{n-1})| \quad \text{for every } n \geq 1, \tag{1.19}$$

so that

$$|U_\tau'(t)| \leq \sup_{n \in \mathbb{N}} \tau^{-1}|U_\tau^n - U_\tau^{n-1}| = \tau^{-1}|U_\tau^1 - U_\tau^0| \leq |\nabla\varphi(U_\tau^0)| \quad \text{for every } t \in [0, +\infty) \setminus \mathcal{P}_\tau. \tag{1.20}$$

Since $\lim_{\tau \downarrow 0} |\nabla\varphi(U_\tau^0)| = |\nabla\varphi(u_0)|$ it follows that $(U_\tau)_{\tau>0}$ satisfies a uniform Lipschitz condition and therefore it admits a suitable subsequence uniformly converging to a Lipschitz curve u in each compact interval $[0, T]$. Denoting by $\bar{U}_\tau(t)$ the piecewise constant interpolant

$$\bar{U}_\tau(t) := U_\tau^n \quad \text{if } t \in (t_\tau^{n-1}, t^n], \tag{1.21}$$

the same estimate (1.20) shows that

$$\sup_{t\in(0,+\infty)} |U_\tau(t) - \bar{U}_\tau(t)| \le \tau|\nabla\varphi(U_\tau^0)|, \tag{1.22}$$

so that $\bar{U}_\tau$ has the same limit points than U_τ. On the other hand, (1.13) yields

$$U_\tau'(t) = -\nabla\varphi(\bar{U}_\tau(t)) \quad \text{in } [0, +\infty) \setminus \mathcal{P}_\tau, \tag{1.23}$$

and we can pass to the limit in an integrated form of (1.23) thus showing that u solves (1.2).

The uniform error estimate (1.17) is subtler: a simple derivation in the case $\lambda = 0$ can be found in [51], see also [59, 61]. Its main functional interest relies on the fact that it involves just the lower bound on the Hessian of φ but not its upper bound (and therefore, it does not require a uniform Lipschitz condition on $\nabla\varphi$). $\qquad\square$

1.3. Metric characterization of Gradient flows in $\mathbb{R}^d$.
The energy identity (EI) (with his integrated version (1.6)) and the evolution variational inequality (EVI$_\lambda$) not only provide important estimates on the solution to (1.2) but can also be used to characterize it.

Concerning (EI) we can even relax the identity, as the following proposition shows.

Proposition 1.6 (Curves of maximal slope). *A C^1 curve $u : [0, +\infty) \to \mathbb{R}^d$ is a solution to (1.2) if and only if it satisfies the* Energy Dissipation Inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(u_t) \le -\frac{1}{2}|u_t'|^2 - \frac{1}{2}|\nabla\varphi(u_t)|^2 \quad \text{in } (0, +\infty) \tag{EDI}$$

or its weaker integrated form

$$\varphi(u_t) + \frac{1}{2}\int_0^t \left(|u_r'|^2 + |\nabla\varphi(u_r)|^2\right) \mathrm{d}r \le \varphi(u_0) \quad \text{for every } t \in (0, +\infty). \tag{EDI'}$$

Proof. Il u is a C^1 curve the chain rule yields

$$\varphi(u_t) = \varphi(u_0) + \int_0^t \langle\nabla\varphi(u_r), u_r'\rangle \mathrm{d}r, \tag{1.24}$$

so that (EDI') yields

$$\frac{1}{2}\int_0^t \left|u_r' + \nabla\varphi(u_r)\right|^2 \mathrm{d}r = \frac{1}{2}\int_0^t \left(|u_r'|^2 + |\nabla\varphi(u_r)|^2\right) \mathrm{d}r + \int_0^t \langle\nabla\varphi(u_r), u_r'\rangle \mathrm{d}r \le 0,$$

and therefore $u_r' = -\nabla\varphi(u_r)$ for $\mathscr{L}^1$-a.e. $r \in (0, t)$. Since t is arbitrary and $u \in C^1$, u solves (1.2). $\qquad\square$

Notice that in the previous formulation we did not use the λ-convexity of φ: the argument only relies on the chain rule.

In the following proposition we show that also the evolution variational inequality (EVI$_\lambda$) characterizes a solution of (1.2). In fact, if (EVI$_\lambda$) admits a solution *for every initial datum u_0*, then φ is λ-convex.

Proposition 1.7 (Characterization of Gradient Flows through the EVI). *If $u : [0, +\infty) \to \mathbb{R}^d$ is a C^1 curve solving (EVI$_\lambda$) then u is a solution to (1.2).*

Proof. Applying the chain rule for the squared distance function $\frac{1}{2}|\cdot -v|^2$ we easily have

$$\langle u_t', u_t - v\rangle \le \varphi(v) - \varphi(u_t) - \frac{\lambda}{2}|u_t - v|^2 \quad \text{for every } v \in \mathbb{R}^d, \ t > 0. \tag{1.25}$$

Choosing $v := u_t + \varepsilon\xi$, for $\varepsilon > 0$ and $\xi \in \mathbb{R}^d$ and dividing by ε we obtain

$$-\langle u_t', \xi\rangle \le \varepsilon^{-1}\left(\varphi(u_t + \varepsilon\xi) - \varphi(u_t)\right) - \frac{\lambda\varepsilon}{2}|\xi|^2 \quad \text{for every } \xi \in \mathbb{R}^d.$$

Passing to the limit as $\varepsilon \downarrow 0$ we eventually get

$$-\langle u_t', \xi \rangle \leq \langle \nabla\varphi(u_t), \xi \rangle \quad \text{for every } \xi \in \mathbb{R}^d,$$

so that $-u_t' = \nabla\varphi(u_t)$. $\hfill\square$

Proposition 1.8. *Let us suppose that there exists a C^1 semigroup $\tilde{\mathsf{S}}_t : \mathbb{R}^d \to \mathbb{R}^d$, $t \geq 0$, of smooth maps such that for every $u_0 \in \mathbb{R}^d$ the curve $u_t := \tilde{\mathsf{S}}_t(u_0)$ satisfies* (EVI$_\lambda$). *Then φ is λ-convex.*

Proof. We consider for simplicity the case $\lambda = 0$; for arbitrary $u^0, u^1 \in \mathbb{R}^d$ we set

$$u^s := (1-s)u^0 + su^1, \quad u_t^s := \tilde{\mathsf{S}}_t(u^s)$$

and we want to show that

$$\frac{\mathrm{d}}{\mathrm{d}s}\varphi(u^s)_{|s=0} \leq \frac{\mathrm{d}}{\mathrm{d}s}\varphi(u^s)_{|s=1}.$$

We get

$$\frac{\mathrm{d}}{\mathrm{d}s}\varphi(u^s)_{|s=0} = \langle \nabla\varphi(u^0), u^1 - u^0 \rangle \overset{(1.2)}{=} -\langle \frac{\mathrm{d}}{\mathrm{d}t}u_t^0_{|t=0}, u^1 - u^0 \rangle = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}|u_t^0 - u^1|^2\Big)_{|t=0}$$

$$\overset{(\text{EVI}_\lambda)}{\leq} \varphi(u^1) - \varphi(u^0) \overset{(\text{EVI}_\lambda)}{\leq} -\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}|u^0 - u_t^1|^2\Big)_{|t=0} \overset{(1.2)}{\leq} \langle \nabla\varphi(u^1), u^1 - u^0 \rangle = \frac{\mathrm{d}}{\mathrm{d}s}\varphi(u^s)_{|s=1} \;\square$$

1.4. **Extensions to more general functional settings.** The simple finite dimensional theory for smooth functionals has been extended in various directions; without claiming any completeness, we quote here four different points of view:

The theory of differential inclusions and maximal monotone operators in Hilbert spaces, developed in the seventies by KOMURA [39], CRANDALL-PAZY [25], CRANDALL-LIGGETT [24], BRÉZIS [13], BÉNILAN [10], J.L. LIONS [40]: we refer to the monographs [14, 8, 40]. In this framework one considers the gradient flow generated by a proper lower semicontinuous λ-convex functional $\phi : H \to (-\infty, +\infty]$, where H is a separable Hilbert space. By using tools of convex analysis, clever regularization techniques, and replacing $\nabla\varphi$ with the multivalued subdifferential operator $\partial\phi$, one can basically reproduce all the estimates and results we briefly discussed in the finite dimensional setting which just depend on the lower bound of the Hessian of φ, avoiding any strong compactness assumptions.

In this framework, the resolvent operator $J_\tau := (I + \tau\partial\phi)^{-1}$ is single-valued and non-expansive, i.e.

$$\mathsf{d}(J_\tau[u], J_\tau[v]) \leq \mathsf{d}(u, v) \quad \text{for every } u, v \in H, \quad \tau > 0. \tag{1.26}$$

This property is the key ingredient to prove, as in the CRANDALL-LIGGETT generation theorem [24], uniform convergence of the exponential formula

$$u_t = \lim_{n \to +\infty}(J_{t/n})^n[u_0], \quad \mathsf{d}(u_t, (J_{t/n})^n[u_0]) \leq \frac{2|\partial\phi|(u_0)t}{\sqrt{n}} \tag{1.27}$$

and therefore to define a contraction semigroup on $\overline{D(\phi)}$.

Being generated by a convex functional, this semigroup exhibits a nice regularization effect, since $u_t \in D(\partial\phi)$ even if $u_0 \in \overline{D(\phi)}$. Moreover, the curve u_t can be characterized as the unique solution of the evolution variational inequality (EVI$_\lambda$), whose formulation goes back to [41]. Optimal error estimates for the implicit Euler discretization in the spirit of (1.17) have been obtained by [7, 59, 61, 51].

The theory of the curves of maximal slope in metric spaces, developed in the eighties by DE GIORGI, DEGIOVANNI, MARINO, TOSQUES in a series of papers originating from [30, 29], and culminating in [31, 44] (but see also the more recent [16] and the presentation of [2, 3]). Here $\phi : X \to (-\infty, +\infty]$ is a proper and lower semicontinuous functional defined in the complete metric space X and one looks for absolutely continuous curves satisfying a suitable form of the Energy dissipation inequality (EDI), where $|u'|$ should be interpreted as the *metric velocity of the curve u* and $|\nabla\varphi(u)|$ should be replaced by the *metric slope of ϕ*. The theory is usually based on local compactness of the sublevels of ϕ and various kind of assumptions on its slope, yielding in particular its lower semicontinuity and the possibility to write a weak form of the chain rule. The

advantage of this approach relies on its flexibility, but in general metric spaces uniqueness and stability properties of curves of maximal slope are not known.

Limits of discrete solutions, generalized minimizing movements. This is the weakest approach, which has been clarified in [28] and independently applied to different kind of problems (see e.g. [43], [35], [36], [48]). It just provides a general approximating scheme which is quite useful to construct some limit curves by compactness arguments, but one can hardly deduce refined properties of these curves from general metric results and each example deserves a careful ad hoc investigation.

Evolution variational inequalities in metric spaces: this is the strongest point of view, which is related to the metric evolution variational inequality (EVI$_\lambda$) and goes back to BÉNILAN [10] notion of integral solutions to evolution equations in Banach spaces. Its application to gradient flows in metric spaces has been developed in [3] and it will be adopted in these notes.

2. GRADIENT FLOWS AND EVOLUTION VARIATIONAL INEQUALITIES IN METRIC SPACES

The aim of this section is to study the metric notion of gradient flows associated to the (metric formulation of the) evolution variational inequality (EVI$_\lambda$).

Throughout the rest of these notes, (X, d) will be a *complete and separable metric space* and $\phi : X \to (-\infty, +\infty]$ a proper and l.s.c. functional on X with non empty domain $D(\phi) = \{v \in X : \phi(v) < +\infty\}$. We will look for curves $u : [0, +\infty) \to X$ which satisfy properties that depend only on the metric structure of X and that in the case of a smooth function $\phi = \varphi$ on $X = \mathbb{R}^d$ satisfy the ODE (1.2).

2.1. A few metric concepts.
Let us first recall the notion of *metric velocity* and *metric slope* (see e.g. [3]).

Definition 2.1 (Absolutely continuous curves). *We say that a curve $v : (a, b) \subset \mathbb{R} \to X$ belongs to $AC^p_{(\mathrm{loc})}(a, b; X)$ for some $p \in [1, +\infty]$ if there exists $m \in L^p_{(\mathrm{loc})}(a, b)$ such that*

$$\mathsf{d}(v_s, v_t) \leq \int_s^t m(r) \, dr \qquad \text{for every } a < s \leq t < b. \tag{2.1}$$

If $p = 1$ we say that v is a (locally) absolutely continuous curve.

Theorem 2.2 (Metric derivative). *If $v : (a, b) \to X$ is an absolutely continuous curve then the limit*

$$|v'|(t) = \lim_{s \to t} \frac{\mathsf{d}(v_s, v_t)}{|t - s|} \tag{2.2}$$

exists for $\mathscr{L}^1$-a.e. $t \in (a, b)$ and it is called metric derivative *of v at the point t. Moreover, the function $t \mapsto |v'|(t)$ belongs to $L^1(a, b)$, it is an admissible integrand for the right hand side of (2.1), and it is minimal in the following sense:*

$$|v'|(t) \leq m(t) \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (a, b), \text{ for each function } m \text{ satisfying } (2.1).$$

Definition 2.3 (Metric Slope). *The metric slope of ϕ at a point $v \in X$ is given by*

$$|\partial\phi|(v) = \begin{cases} +\infty & \text{if } v \notin D(\phi), \\ 0 & \text{if } v \in D(\phi) \text{ is isolated}, \\ \limsup_{w \to v} \dfrac{\left(\phi(v) - \phi(w)\right)^+}{\mathsf{d}(v, w)} & \text{otherwise}. \end{cases} \tag{2.3}$$

2.2. Structural properties of solutions to Evolution Variational Inequalities.
The next (quite restrictive) definition is modeled on the case of λ-convex functionals in Euclidean-like spaces and has been introduced and discussed in [3, Chap. 4].

Definition 2.4 (EVI and Gradient flow). *A solution of the evolution variational inequality* EVI$_\lambda(X, \mathsf{d}, \phi)$, $\lambda \in \mathbb{R}$, *is a locally absolutely continuous curve $u : t \in (0, +\infty) \mapsto u_t \in D(\phi)$ such that*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{d}^2(u_t, v) + \frac{\lambda}{2}\mathsf{d}^2(u_t, v) \leq \phi(v) - \phi(u_t) \quad \mathscr{L}^1\text{-a.e. in } (0, +\infty), \quad \text{for every } v \in D(\phi). \quad \text{(EVI}_\lambda)$$

A λ-gradient flow of ϕ is a family of continuous maps $\mathsf{S}_t : \overline{D(\phi)} \to D(\phi)$, $t > 0$, such that for every $u \in \overline{D(\phi)}$

$$\lim_{t \downarrow 0} \mathsf{S}_t(u) = u =: \mathsf{S}_0(u), \qquad \mathsf{S}_{t+h}(u) = \mathsf{S}_h(\mathsf{S}_t(u)) \quad \textit{for every } t, h \geq 0, \tag{2.4a}$$

$$\textit{the curve } t \mapsto \mathsf{S}_t(u) \quad \textit{is a solution of } \mathrm{EVI}_\lambda(X, \mathsf{d}, \phi). \tag{2.4b}$$

The next result shows that (EVI_λ) can be formulated avoiding differentiation and without assuming the absolute continuity of u (see [60] for the proof).

Theorem 2.5 (Derivative free characterization of solutions to (EVI_λ))**.** *A curve $u : (0, +\infty) \to \overline{D(\phi)}$ is a solution of $\mathrm{EVI}_\lambda(X, \mathsf{d}, \phi)$ according to Definition 2.4 if and only if for every $s, t \in (0, +\infty)$ with $s < t$ and $v \in D(\phi)$*

$$\frac{\mathrm{e}^{\lambda(t-s)}}{2} \mathsf{d}^2(u_t, v) - \frac{1}{2} \mathsf{d}^2(u_s, v) \leq \mathsf{E}_\lambda(t - s)\Big(\phi(v) - \phi(u_t)\Big). \tag{EVI'_λ}$$

Notice that (EVI'_λ) yields the pointwise right-upper differential inequality

$$\frac{1}{2} \frac{\mathrm{d}^+}{\mathrm{d}t} \mathsf{d}^2(u_t, v) + \frac{\lambda}{2} \mathsf{d}^2(u_t, v) \leq \phi(v) - \phi(u_t) \quad \textit{for every } v \in D(\phi), \tag{2.5}$$

at *every* time $t > 0$: here $\frac{\mathrm{d}^+}{\mathrm{d}t} \zeta$ denotes the right-upper Dini derivative $\limsup_{h \downarrow 0} h^{-1}(\zeta(t+h) - \zeta(t))$.

The next result collects many useful properties of solutions to $\mathrm{EVI}_\lambda(X, \mathsf{d}, \phi)$ (see [60] and an analogous result of [4] in the Wasserstein framework): they reproduce in the metric framework the estimates of the previous section and show that (EVI_λ) contains all the information concerning the gradient flow of ϕ.

Theorem 2.6 (Properties of solutions to (EVI_λ))**.** *Let $u, u^1, u^2 : [0, +\infty) \to X$ be solutions of $\mathrm{EVI}_\lambda(X, \mathsf{d}, \phi)$.*

λ-contraction and uniqueness:

$$\mathsf{d}(u_t^1, u_t^2) \leq \mathrm{e}^{-\lambda(t-s)} \mathsf{d}(u_s^1, u_s^2) \quad \textit{for every } 0 \leq s < t < +\infty. \tag{2.6}$$

In particular, for every $u_0 \in \overline{D(\phi)}$ there is at most one solution u of $\mathrm{EVI}_\lambda(X, \mathsf{d}, \phi)$ satisfying the initial condition $\lim_{t \downarrow 0} u_t = u_0$.

Regularizing effects: *u is locally Lipschitz continuous in $(0, +\infty)$ and $u_t \in D(|\partial\phi|) \subset D(\phi)$ for every $t > 0$. Moreover in the time interval $[0, +\infty)$*

$$\textit{the map } t \mapsto \phi(u_t) \textit{ is non-increasing and (locally semi-, if } \lambda < 0) \textit{ convex,} \tag{2.7}$$

$$\textit{the map } t \mapsto \mathrm{e}^{\lambda t} |\partial\phi|(u_t) \textit{ is non-increasing and right continuous,} \tag{2.8}$$

the following regularization/a priori estimate holds

$$\frac{\mathrm{e}^{\lambda t}}{2} \mathsf{d}^2(u_t, v) + \mathsf{E}_\lambda(t)\Big(\phi(u_t) - \phi(v)\Big) + \frac{\big(\mathsf{E}_\lambda(t)\big)^2}{2} |\partial\phi|^2(u_t) \leq \frac{1}{2} \mathsf{d}^2(u_0, v) \tag{2.9}$$

for every $v \in D(\phi)$; in particular

$$\phi(u_t) \leq \phi(v) + \frac{1}{2\mathsf{E}_\lambda(t)} \mathsf{d}^2(u_0, v), \tag{2.10}$$

$$|\partial\phi|^2(u_t) \leq \frac{1}{2\mathrm{e}^{\lambda t} - 1} |\partial\phi|^2(v) + \frac{1}{(\mathsf{E}_\lambda(t))^2} \mathsf{d}^2(u_0, v) \quad \textit{if } -\lambda t < \log 2. \tag{2.11}$$

Asymptotic expansion for $t \downarrow 0$: *If $u_0 \in D(|\partial\phi|)$ and $\lambda \leq 0$ then for every $v \in D(\phi)$ and $t \geq 0$*

$$\frac{\mathrm{e}^{2\lambda t}}{2} \mathsf{d}^2(u_t, v) - \frac{1}{2} \mathsf{d}^2(u_0, v) \leq \mathsf{E}_{2\lambda}(t)\big(\phi(v) - \phi(u_0)\big) + \frac{t^2}{2} |\partial\phi|^2(u_0). \tag{2.12}$$

Right and left limits, energy identity: *For every $t > 0$ the* right *limits*

$$|\dot{u}_{t+}| := \lim_{h \downarrow 0} \frac{d(u_t, u_{t+h})}{h}, \quad \frac{d}{dt}\phi(u_{t+}) := \lim_{h \downarrow 0} \frac{\phi(u_{t+h}) - \phi(u_t)}{h} \tag{2.13}$$

exist, they satisfy

$$\frac{d}{dt}\phi(u_{t+}) = -|\dot{u}_{t+}|^2 = -|\partial\phi|^2(u_t), \tag{2.14}$$

and they define a right continuous map. (2.13) and (2.14) hold at $t = 0$ iff $u_0 \in D(|\partial\phi|)$. Moreover, there exists an at most countable set $\mathcal{C} \subset (0, +\infty)$ such that the analogous identities for the left *limits hold for every $t \in (0, +\infty) \setminus \mathcal{C}$.*

Asymptotic behavior: *If $\lambda > 0$, then ϕ admits a unique minimum point $\bar{u}$ and for every $t \geq t_0 \geq 0$ we have*

$$\frac{\lambda}{2}d^2(u_t, \bar{u}) \leq \phi(u_t) - \phi(\bar{u}) \leq \frac{1}{2\lambda}|\partial\phi|^2(u_t), \tag{2.15a}$$

$$d^2(u_t, \bar{u}) \leq d^2(u_{t_0}, \bar{u})e^{-\lambda(t-t_0)}, \tag{2.15b}$$

$$\phi(u_t) - \phi(\bar{u}) \leq \Big(\phi(u_{t_0}) - \phi(\bar{u})\Big)e^{-2\lambda(t-t_0)}, \quad \phi(u_t) - \phi(\bar{u}) \leq \frac{1}{2}e^{\lambda(t-t_0)}d^2(u_{t_0}, \bar{u}), \tag{2.15c}$$

$$|\partial\phi|(u_t) \leq |\partial\phi|(u_{t_0})e^{-\lambda(t-t_0)}, \quad |\partial\phi|(u_t) \leq \frac{1}{e^{\lambda(t-t_0)}d(u_{t_0}, \bar{u})}. \tag{2.15d}$$

If $\lambda = 0$ and $\bar{u}$ is any minimum point of ϕ then we have

$$|\partial\phi|(u_t) \leq \frac{d^2(u_0, \bar{u})}{t}, \quad \phi(u_t) - \phi(\bar{u}) \leq \frac{d^2(u_0, \bar{u})}{2t}, \tag{2.16}$$

$$\text{the map} \quad t \mapsto d^2(u_t, \bar{u}) \quad \text{is not increasing.}$$

Continuity of the energy and the slope: *If $u^n \in C^0([0, +\infty); X)$ are solutions of $\mathrm{EVI}_\lambda(X, d, \phi)$ such that $\lim_{n \uparrow +\infty} u_0^n = u_0$, then*

$$\lim_{n \uparrow +\infty} \phi(u_t^n) = \phi(u_t) \qquad \text{for every } t > 0, \tag{2.17}$$

$$\lim_{n \uparrow +\infty} |\partial\phi|(u_t^n) = |\partial\phi|(u_t) \qquad \text{for every } t \in (0, +\infty) \setminus \mathcal{C}. \tag{2.18}$$

We just sketch the proof of the contraction property (2.6). *For a fixed $s \in (0, +\infty)$ we have that*

$$\frac{\partial}{\partial t}\frac{1}{2}d^2(u_t^1, u_s^2) + \frac{\lambda}{2}d^2(u_t^1, u_s^2) \leq \phi(u_s^2) - \phi(u_t^1) \quad \text{for every } t \in (0, +\infty), \tag{2.19}$$

while for a fixed $t \in (0, +\infty)$

$$\frac{\partial}{\partial s}\frac{1}{2}d^2(u_t^1, u_s^2) + \frac{\lambda}{2}d^2(u_t^1, u_s^2) \leq \phi(u_t^1) - \phi(u_s^2) \quad \text{for every } s \in (0, +\infty). \tag{2.20}$$

Adding (2.19) and (2.20) we get

$$\frac{\partial}{\partial t}\frac{1}{2}d^2(u_t^1, u_s^2) + \frac{\partial}{\partial s}\frac{1}{2}d^2(u_t^1, u_s^2) + \lambda d^2(u_t^1, u_s^2) \leq 0;$$

Applying [3, Lemma 4.3.4] we obtain

$$\frac{d}{dt}d^2(u_t^1, u_t^2) \leq -2\lambda d^2(u_t^1, u_t^2) \quad \mathcal{L}^1\text{-a.e. in } (0, +\infty)$$

and therefore we obtain (2.6). $\qquad\square$

Theorem 2.6 concerns each single solution to (EVI_λ); when the λ-gradient flow S_t of ϕ exists we have further interesting properties, showing that the formulation by (EVI_λ) is really stronger than all the other metric approaches.

2.3. **λ-Gradient flows and λ-convexity along geodesics.** Let us first recall the notion of (minimal, constant speed) geodesics in a metric space X and the related convexity.

Definition 2.7 (Constant speed geodesics). *A curve $\gamma : [0,1] \to X$ is a* constant speed geodesic *(or simply* geodesic*) if*

$$\mathsf{d}(\gamma_s, \gamma_t) = |t - s|\, \mathsf{d}(\gamma_0, \gamma_1) \quad \text{for every } 0 \le s \le t \le 1. \tag{2.21}$$

A set $D \subset X$ is geodesically convex *if every couple of points $x_0, x_1 \in D$ can be connected by a geodesic γ contained in D.*

Definition 2.8 (λ-convexity along curves and geodesically λ-convex functionals). *We say that $\phi : X \to (-\infty, +\infty]$ is* λ-convex *along a curve $\gamma : [0,1] \to X$ if*

$$\phi(\gamma_s) \le (1-s)\phi(\gamma_0) + s\phi(\gamma_1) - \frac{\lambda}{2}s(1-s)\mathsf{d}^2(\gamma_0, \gamma_1) \quad \text{for every } s \in [0,1]. \tag{2.22}$$

We say that ϕ is geodesically λ-convex *if every couple of points $x_0, x_1 \in D(\phi)$ can be connected by a geodesic γ along which ϕ is λ-convex. If ϕ is geodesically convex and it is λ-convex along every geodesic connecting $x_0, x_1 \in D(\phi)$ in $\overline{D(\phi)}$ then we say that ϕ is* strongly *geodesically λ-convex.*

Theorem 2.9 ([27]). *If the λ-gradient flow S_t of ϕ exists then ϕ is λ-convex along any geodesic in $\overline{D(\phi)}$. In particular, if $\overline{D(\phi)}$ is geodesically convex, then ϕ is strongly geodesically λ-convex.*

Proof. Let $\gamma : s \in [0,1] \mapsto \gamma^s \in \overline{D(\phi)}$ be a geodesic with $\gamma^0, \gamma^1 \in D(\phi)$ and let us set $\gamma_t^s := \mathsf{S}_t(\gamma^s)$.

Applying (EVI'_λ) we have for every $s \in [0,1]$ and $t > 0$

$$\frac{1}{2}\mathrm{e}^{\lambda t}\mathsf{d}^2(\gamma_t^s, \gamma^0) - \frac{1}{2}\mathsf{d}^2(\gamma^s, \gamma^0) \le \mathsf{E}_\lambda(t)\big(\phi(\gamma^0) - \phi(\gamma_t^s)\big), \tag{2.23}$$

$$\frac{1}{2}\mathrm{e}^{\lambda t}\mathsf{d}^2(\gamma_t^s, \gamma^1) - \frac{1}{2}\mathsf{d}^2(\gamma^s, \gamma^1) \le \mathsf{E}_\lambda(t)\big(\phi(\gamma^1) - \phi(\gamma_t^s)\big). \tag{2.24}$$

Multiplying (2.23) by $(1-s)$ and (2.24) by s and adding the two inequalities we get

$$\frac{\mathrm{e}^{\lambda t}}{2}\big((1-s)\mathsf{d}^2(\gamma_t^s, \gamma^0) + s\mathsf{d}^2(\gamma_t^s, \gamma^1)\big) - \frac{1}{2}\big((1-s)\mathsf{d}^2(\gamma^s, \gamma^0) + s\mathsf{d}^2(\gamma^s, \gamma^1)\big)$$
$$\le \mathsf{E}_\lambda(t)\big((1-s)\phi(\gamma^0) + s\phi(\gamma^1) - \phi(\gamma_t^s)\big). \tag{2.25}$$

We now observe that the elementary inequality

$$(1-s)a^2 + sb^2 \ge s(1-s)(a+b)^2 \quad \text{for every } a, b \in \mathbb{R}, \quad s \in [0,1], \tag{2.26}$$

and the triangular inequality yield

$$(1-s)\mathsf{d}^2(\gamma_t^s, \gamma^0) + s\mathsf{d}^2(\gamma_t^s, \gamma^1) \overset{(2.26)}{\ge} s(1-s)\big(\mathsf{d}(\gamma_t^s, \gamma^0) + \mathsf{d}(\gamma_t^s, \gamma^1)\big)^2$$
$$\ge s(1-s)\mathsf{d}^2(\gamma^0, \gamma^1). \tag{2.27}$$

On the other hand, since γ is a geodesic we have

$$(1-s)\mathsf{d}^2(\gamma^s, \gamma^0) + s\mathsf{d}^2(\gamma^s, \gamma^1) = s(1-s)\mathsf{d}^2(\gamma^0, \gamma^1). \tag{2.28}$$

Inserting (2.27) and (2.28) in (2.25) we get

$$\frac{\mathrm{e}^{\lambda t} - 1}{2}s(1-s)\mathsf{d}^2(\gamma^0, \gamma^1) \le \mathsf{E}_\lambda(t)\big((1-s)\phi(\gamma^0) + s\phi(\gamma^1) - \phi(\gamma_t^s)\big). \tag{2.29}$$

Dividing then both sides of (2.29) by $\mathsf{E}_\lambda(t)$ and passing to the limit as $t \downarrow 0$ we obtain

$$\phi(\gamma^s) \le (1-s)\phi(\gamma^0) + s\phi(\gamma^1) - \frac{\lambda}{2}s(1-s)\mathsf{d}^2(\gamma^0, \gamma^1) \quad \text{for every } s \in [0,1].$$

$\square$

2.4. λ-gradient flows and curves of maximal slope.

Definition 2.10 (Curves of maximal slope). *We say that a curve $u \in AC_{\mathrm{loc}}^2(0, +\infty; X)$ is a curve of maximal slope for the functional ϕ if the energy dissipation inequality*

$$\frac{1}{2} \int_s^t |u'|^2(r)\, dr + \frac{1}{2} \int_s^t |\partial\phi|^2(u_r)\, dr \le \phi(u_s) - \phi(u_t) \tag{2.30}$$

holds for all $0 < s \le t < +\infty$.

The notion of curve of maximal slope has been first introduced (in a slightly different form) by De Giorgi and provides a weak notion of gradient flow for nonsmooth functionals, also nonconvex. If ϕ admits a λ-gradient flow according to Definition 2.4, then these two definitions coincide.

Theorem 2.11. *Let us assume that the λ-gradient flow S of ϕ exists and let $u \in AC_{\mathrm{loc}}^2(0, +\infty; X)$ be satisfying (2.30) with $\lim_{t\downarrow 0} u_t = u_0 \in D(\phi)$. Then $u_t = \mathsf{S}_t(u_0)$ for every $t \ge 0$ and (2.30) is in fact an identity for every $0 \le s < t < +\infty$.*

2.5. λ-gradient flows and the minimizing movements variational scheme.

A general variational method to approximate gradient flows (and often to prove their existence) is provided by the so-called *minimizing movements* variational scheme. In his original formulation (see e.g. [28]), the method consists in finding a discrete approximation $\overline{U}_\tau$ of the continuous gradient flow u by solving a recursive variational scheme, which is the natural generalization of (1.15) to a metric-space setting. If $\tau > 0$ denotes the step size of the uniform partition $\mathcal{P}_\tau$ (1.12), starting from a suitable approximation U_τ^0 of u_0 one looks at each step $((n-1)\tau, n\tau]$ for the minimizers of the functional

$$U \mapsto \Phi(\tau, U_\tau^{n-1}; U) := \frac{1}{2\tau} \mathsf{d}^2(U, U_\tau^{n-1}) + \phi(U). \tag{2.31}$$

$\overline{U}_\tau$ thus takes a value $U_\tau^n \in \operatorname{argmin} \Phi(\tau, U^{n-1}; \cdot)$ on each interval $((n-1)\tau, n\tau]$.

Definition 2.12 (The minimizing movement variational scheme). *Let us consider a time step $\tau > 0$ and a discrete initial datum $U_\tau^0 \in D(\phi)$. A τ-discrete minimizing movement starting from U_τ^0 is any sequence $(U_\tau^n)_{n\in\mathbb{N}}$ in $D(\phi)$ which satisfies*

$$\Phi(\tau, U_\tau^{n-1}; U_\tau^n) \le \Phi(\tau, U_\tau^{n-1}; V) \quad \text{for every } V \in X, \ n \in \mathbb{N}. \tag{2.32}$$

A discrete solution $\overline{U}_\tau$ is any piecewise constant interpolant of a τ-discrete minimizing movement on the grid $\mathcal{P}_\tau$ defined by

$$\overline{U}_\tau(0) = U_\tau^0, \quad \overline{U}_\tau(t) \equiv U_\tau^n \quad \text{if } t \in (t_\tau^{n-1}, t_\tau^n], \ n \ge 1. \tag{2.33}$$

The existence of a minimizing sequence $\{U_\tau^n\}_{n\in\mathbb{N}}$ is usually obtained by invoking the direct method of the Calculus of Variations, thus requiring that the functional (2.31) has compact sublevels with respect to some Hausdorff topology σ on X (see e.g. the setting of [3, § 2.1]). In the next section we will discuss another possibility, still considered in [3], when the functional (2.31) satisfies a strong convexity assumption.

In a general setting it is also possible to avoid these restrictions by applying the Ekeland's Variational Principle to the functional (2.31): this approach only requires the completeness of the metric space.

Definition 2.13 (A relaxed minimizing movement variational scheme). *Let us consider a time step $\tau > 0$, a relaxation parameter $\eta \ge 0$, and a discrete initial datum $U_{\tau,\eta}^0 \in D(\phi)$. A (τ, η)-discrete minimizing movement starting from $U_{\tau,\eta}^0$ is any sequence $(U_{\tau,\eta}^n)_{n\in\mathbb{N}}$ in $D(\phi)$ which satisfies*

$$\Phi(\tau, U_{\tau,\eta}^{n-1}; U_{\tau,\eta}^n) \le \Phi(\tau, U_{\tau,\eta}^{n-1}; V) + \frac{\eta}{2} \mathsf{d}(U_{\tau,\eta}^n, U_{\tau,\eta}^{n-1})\, \mathsf{d}(V, U_{\tau,\eta}^n) \quad \text{for every } V \in D(\phi), \tag{2.34a}$$

and the further condition

$$\Phi(\tau, U_{\tau,\eta}^{n-1}; U_{\tau,\eta}^n) = \frac{1}{2\tau} \mathsf{d}^2(U_{\tau,\eta}^n, U_{\tau,\eta}^{n-1}) + \phi(U_{\tau,\eta}^n) \le \phi(U_{\tau,\eta}^{n-1}), \tag{2.34b}$$

for every $n \in \mathbb{N}$. A (τ, η)-discrete solution $\overline{U}_{\tau,\eta}$ is any piecewise constant interpolant of a (τ, η)-discrete minimizing movement on the grid $\mathcal{P}_\tau$, as in (2.33).

Notice that when $\eta = 0$ a solution to (2.34a) is a minimizer of (2.31) (and in particular satisfies (2.34b)), so that the usual discrete solutions arising from the minimizing movement scheme are included in this more general relaxed framework. The next result [60], which follows directly from Ekeland's variational principle, shows that the previous scheme is always solvable when $\eta > 0$.

Theorem 2.14. *Let us assume that X is complete and ϕ is quadratically bounded from below, i.e. for some $\kappa_o, \phi_o \in \mathbb{R}$, $o \in X$*

$$\phi(x) + \frac{\kappa_o}{2}\mathsf{d}^2(x, o) \geq \phi_o \quad \text{for every } x \in X. \tag{2.35}$$

Then for every $\eta > 0$, $\tau > 0$ with $\tau^{-1} > -\kappa_0$, and $U^0_{\tau,\eta} \in D(\phi)$, the relaxed minimizing movement scheme admits at least a (τ, η)-discrete solution $(U^n_{\tau,\eta})_{n \in \mathbb{N}}$.

Since under the general assumptions of Theorem 2.14 the relaxed minimizing movement scheme admits a (τ, η)-discrete solution $\overline{U}_{\tau,\eta}$ for fixed $\eta > 0$ and arbitrarily small step size τ, it is natural to ask what its limit as $\tau \downarrow 0$. A first result in this direction is provided by the next theorem, which shows that the minimizing movement scheme is *consistent* with the definition of λ-gradient flow 2.4. Notice that in Theorem 2.15 we will assume *a priori* that the λ-gradient flow of ϕ exists to get the convergence of $\overline{U}_{\tau,\eta}$; in Section 3 we will discuss how to remove this strong assumption. Still it is sometimes useful to know that any λ-gradient flow, no matter how it has been constructed, admits a uniformly converging discrete approximation, which exhibits nice variational properties.

Theorem 2.15. *Let us assume that there exists the λ-gradient flow S_t of ϕ according to Definition 2.4 and that $\overline{D(\phi)}$ is geodesically convex. Let $\tau > 0, \eta \geq 0$ satisfy $\eta - \lambda < \frac{1}{2\tau}$, and let the sequence $(U^n_{\tau,\eta})_{n \in \mathbb{N}} \subset D(\phi)$ be a (τ, η)-discrete minimizing movement with $U^0_{\tau,\eta} \in D(|\partial\phi|)$. Setting $\alpha = \alpha_{\tau,\eta} := \frac{1}{2\tau}\log(1 + 2(\lambda - \eta)\tau)$ we have the* a priori *error estimate*

$$\mathsf{d}(\mathsf{S}_t(u_0), \overline{U}_{\tau,\eta}(t)) \leq \mathsf{d}(u_0, U^0_{\tau,\eta}) + \mathrm{e}^{-\alpha T}\sqrt{T\tau}|\partial\phi|(U^0_{\tau,\eta}) \quad \text{for every } t \in [0, T]. \tag{2.36}$$

In particular if for some $\eta \geq 0$ and every $\tau \in (0, \tau_0)$ $\overline{U}_{\tau,\eta}$ is a family of (τ, η)-discrete solutions with $U^0_{\tau,\eta} = u_0 \in D(|\partial\phi|)$, then $\lim_{\tau \downarrow 0} \overline{U}_{\tau,\eta}(t) = \mathsf{S}_t(u_0)$ uniformly on every compact interval.

Let us remark that η has been kept fixed in the previous convergence result, so that the coefficients $\alpha_{\tau,\eta} = \frac{1}{2\tau}\log(1 + 2(\lambda - \eta)\tau)$ in the estimate (2.36) are uniformly bounded from below as $\tau \downarrow 0$.

2.6. **Stability of λ-gradient flows under Γ-convergence.** We conclude this section by showing a simple stability property of Gradient Flows with respect to perturbations of the generating functional ϕ. Here we consider a coercive family of Γ-converging functionals $\phi^h : X \to (-\infty, +\infty]$, $h \in \bar{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$, which are quadratically bounded from below, uniformly w.r.t. h: for some $o \in X$, $\phi_o, \kappa_o \in \mathbb{R}$ they satisfy

$$\phi^h(x) + \frac{\kappa_o}{2}\mathsf{d}^2(x, o) \geq \phi_o \quad \text{for every } x \in X, \ h \in \mathbb{N}. \tag{2.37}$$

In the next definition we jointly recall the (sequential) notions of Γ-convergence and of coercivity [26, Def. 1.12]. For notational convenience, we will identify monotone subsequences $(h_n)_{n \in \mathbb{N}}$ with their unbounded image $H = \{h_n : n \in \mathbb{N}\} \subset \mathbb{N}$; expressions like $\lim_{h \in H}$, $\liminf_{h \in H}$ have an obvious meaning as limits for $h \uparrow +\infty, h \in H$.

Definition 2.16 (Sequential $\Gamma(X, \mathsf{d})$-convergence of coercive functionals). *We say that $(\phi^h)_{h \in \mathbb{N}}$ is a coercive family of functionals $\Gamma(X, \mathsf{d})$-sequentially converging to a proper functional $\phi^\infty : X \to (-\infty, +\infty]$ if the following two conditions are satisfied:*

(1) For every infinite subset $H \subset \mathbb{N}$ and every bounded sequence $(x^h)_{h \in H}$ with $\sup_{h \in H} \phi^h(x^h) < +\infty$, there exists an infinite subsequence $H' \subset H$ such that $\lim_{h \in H'} x^h = x^\infty \in D(\phi^\infty)$ and

$$\liminf_{h \in H'} \phi^h(x^h) \geq \phi^\infty(x^\infty). \tag{2.38}$$

(2) For every $\bar{x}^\infty \in D(\phi^\infty)$ there exists a sequence $(\bar{x}^h)_{h \in \mathbb{N}}$ such that

$$\lim_{h \uparrow +\infty} \mathsf{d}(\bar{x}^h, \bar{x}^\infty) = 0, \quad \lim_{h \uparrow +\infty} \phi^h(\bar{x}^h) = \phi^\infty(\bar{x}^\infty). \tag{2.39}$$

It is possible to prove that a coercive family of λ-convex functionals $(\phi^h)_{h \in \mathbb{N}}$ $\Gamma(X, \mathsf{d})$-converging to ϕ^∞ always satisfies the uniform lower bound (2.37).

Let us now state our first convergence result [60].

Theorem 2.17. *Let $(\phi^h)_{h \in \mathbb{N}}$ be a coercive family of functionals $\Gamma(X, \mathsf{d})$-converging to ϕ^∞ and let us assume that the λ-gradient flows S^h exist for every $h \in \mathbb{N}$. Then the functional ϕ^∞ admits a λ-gradient flow S^∞ and for every sequence $u_0^h \in \overline{D(\phi^h)}$ converging to $u_0^\infty \in \overline{D(\phi^\infty)}$ we have*

$$\lim_{h \uparrow +\infty} \mathsf{S}_t^h(u_0^h) = \mathsf{S}_t^\infty(u_0^\infty), \quad \lim_{h \uparrow +\infty} \phi^h(\mathsf{S}_t^h(u_0^h)) = \phi^\infty(\mathsf{S}_t^\infty(u_0^\infty)) \quad \text{for every } t > 0, \tag{2.40}$$

locally uniformly on $(0, +\infty)$.

Proof. Here we consider the simpler case when (2.37) holds for $\kappa_o = 0$; it is not restrictive to assume $\lambda \leq 0$ and $\phi_o \geq 0$.

Step 1: uniform bounds. We set $u_t^h := \mathsf{S}_t^h(u_0^h)$ and we fix a compact time interval $[0, T]$, $T > 0$, a point $o^\infty \in D(\phi^\infty)$ and a corresponding sequence o^h as in (2.39). (EVI$'_\lambda$) yields

$$\mathsf{d}^2(u_t^h, o^h) \leq \left(\mathsf{d}^2(u_0{}^h, o^h) + 2\mathsf{E}_\lambda(t)\, \phi^h(o^h) \right) e^{-\lambda t}; \tag{2.41}$$

and therefore there exists a constant $C_1(T)$ independent of h such that

$$\mathsf{d}(u_t^h, o^h) \leq C_1(T) \quad \text{for every } t \in [0, T],\ h \in \mathbb{N}. \tag{2.42}$$

The regularizing estimate (2.9) yields

$$\frac{e^{\lambda t}}{2} \mathsf{d}^2(u_t^h, o^h) + \mathsf{E}_\lambda(t)\phi^h(u_t^h) + \frac{\left(\mathsf{E}_\lambda(t)\right)^2}{2}|\partial\phi^h|^2(u_t^h) \leq \frac{1}{2}\mathsf{d}^2(u_0^h, o^h) + \mathsf{E}_\lambda(t)\phi^h(o^h) \leq C_2(T) \tag{2.43}$$

if $t \in (0, T]$, for a suitable constant $C_2(T)$ independent of h.

In particular for every $0 < S < T$ there exists a constant $C(S, T)$ such that

$$\phi^h(u_t^h) \leq C(S, T), \quad |\partial\phi^h|(u_t^h) = |\dot{u}_{t+}^h| \leq C(S, T) \quad \text{for every } t \in [S, T]. \tag{2.44}$$

Step 2: compactness. By the estimates of the previous point, the sequence $(u^h)_{h \in \mathbb{N}}$ is uniformly Lipschitz in each bounded interval $[S, T]$ of $(0, +\infty)$ and for every fixed t $\{u_t^h\}_{h \in \mathbb{N}}$ satisfies the assumptions of Definition 2.16, so that $(u_t^h)_{h \in \mathbb{N}}$ is relatively compact in X. Applying Ascoli-Arzelà theorem we can find a subsequence $H = (h_n)_{n \in \mathbb{N}}$ such that u^{h_n} converge locally uniformly in time to a locally Lipschitz curve u^∞ in $(0, +\infty)$.

Step 3: characterization of the limit. Let us now fix an arbitrary point $v^\infty \in D(\phi^\infty)$ and a corresponding approximating sequence $v^h \in D(\phi^h)$ as in (2.39). By (EVI$'_\lambda$) of Theorem 2.5 we know that

$$\frac{e^{\lambda(t-s)}}{2}\mathsf{d}^2(u_t^h, v^h) - \frac{1}{2}\mathsf{d}^2(u_s^h, v^h) \leq \mathsf{E}_\lambda(t-s)\left(\phi^h(v^h) - \phi^h(u_t^h) \right); \tag{2.45}$$

We then pass to the limit in (2.45) as $h \uparrow +\infty, h \in H$, using the facts that u_t^h converges pointwise to u_t in X and applying (2.38) for u_t^h and (2.39) for v^h; we obtain

$$\frac{e^{\lambda(t-s)}}{2}\mathsf{d}^2(u_t^\infty, v^\infty) - \frac{1}{2}\mathsf{d}^2(u_s^\infty, v^\infty) \leq \mathsf{E}_\lambda(t-s)\left(\phi(v^\infty) - \phi(u_t^\infty) \right) \tag{2.46}$$

for every $v^\infty \in D(\phi^\infty)$, $0 \leq s < t$. A further application of Theorem 2.5 shows that u^∞ solves EVI$_\lambda(X, \mathsf{d}, \phi^\infty)$.

In order to check that $\lim_{t \downarrow 0} u_t^\infty = u_0^\infty$ we use (2.46) at $s = 0$ and the lower semicontinuity of ϕ^∞, which yields

$$\limsup_{t \downarrow 0} \mathsf{d}^2(u_t^\infty, v^\infty) \leq \mathsf{d}^2(u_0^\infty, v^\infty) \quad \text{for every } v^\infty \in D(\phi^\infty); \tag{2.47}$$

since $u_0^\infty \in \overline{D(\phi^\infty)}$ we conclude that $\lim_{t \downarrow 0} \mathsf{d}(u_t^\infty, u_0^\infty) = 0$.

Since the limit is the unique solution of EVI$_\lambda(X, \mathsf{d}, \phi^\infty)$ starting from u_0^∞, we conclude that the whole sequence u^h converge to u^∞.

Step 4: convergence of energy. We argue as in the proof of (2.17) and (2.18): for a fixed $t > 0$ and applying (2.39) to u_t^∞ we find a sequence $(\bar{u}_t^h)_{h \in \mathbb{N}}$ converging to u_t^∞ with $\lim_{h\uparrow+\infty} \phi^h(\bar{u}_t^h) = \phi^\infty(u_t^\infty)$. By estimate (2.43), the slope $|\partial\phi^h|(u_t^h)$ is uniformly bounded by a constant M_t so that

$$\phi^h(\bar{u}_t^h) \geq \phi^h(u_t^h) - M_t \mathsf{d}(\bar{u}_t^h, u_t^h) - \frac{\lambda}{2}\mathsf{d}^2(\bar{u}_t^h, u_t^h).$$

Passing to the limit as $h \uparrow +\infty$ we get $\limsup_{h\uparrow+\infty} \phi^h(u_t^h) \leq \phi^\infty(u_t^\infty)$, which combined with (2.38) yields the second identity of (2.40). □

3. Convergence of the minimizing movement method and generation results

We have seen in Theorem 2.15 a first convergence theorem for the relaxed minimizing movement method: it basically says that *if ϕ admits a λ-gradient flow* according to Definition 2.4 then any family of discrete solutions converges to the unique continuous solution of (EVI$_\lambda$) as the time step converges to 0.

In this section we revert this point of view and we try to prove *the existence of the λ-gradient flow* when ϕ is geodesically λ-convex by studying the convergence of the (relaxed) minimizing movement Method.

In the following we present three different results in this direction:

(1) A simpler convergence result when *the sublevels of ϕ are locally compact*: in this case we avoid any geometric restriction on the distance d of X and we do not need any Cauchy estimate. On the other hand, the (not necessarily unique) limit points of the discrete solutions are just curves of Maximal Slope, according to Definition 2.10: in general it is not possible to prove that they solve (EVI$_\lambda$).

(2) A first generation result for λ-gradient flows, by assuming that the minimizing movement generating functional $\Phi(\tau, U; V)$ defined by (2.31) satisfies a suitable convexity property (which results from the combination of the convexity of d^2 and of ϕ).

(3) A second generation result when $\mathsf{d}^2(\cdot, v)$ is semiconcave along geodesics and the metric space satisfies a *local angle condition* between triple of geodesics emanating from the same point.

Differently from the first approach, the last ones provide explicit Cauchy estimates ensuring the convergence of the method and do not require any local compactness of the sublevels of ϕ.

3.1. **Convergence of the variational scheme in the locally compact case.** Let us first consider the case when ϕ is geodesically λ-convex and its sublevels are locally compact, i.e. $\exists\, o \in X$ s.t.

$$\left\{x \in X : \phi(x) \leq R \text{ and } \mathsf{d}(x, o) \leq R\right\} \quad \text{are compact in } X \text{ for every } R > 0. \tag{3.1}$$

Combining [3, Proposition 2.2.3, Corollary 2.4.11] we get

Theorem 3.1 (Limits of discrete minimizing movements are curves of maximal slope)**.** *If ϕ is geodesically λ-convex and satisfies (3.1) then for every $\tau > 0$ satisfying $\tau^{-1} > -\lambda$ and $U_\tau^0 \in D(\phi)$ the minimizing movement variational scheme admits at least one solution $(U_\tau^n)_{n\in\mathbb{N}}$. If moreover*

$$\lim_{\tau\downarrow 0} U_\tau^0 = u_0, \quad \lim_{\tau\downarrow 0} \phi(U_\tau^0) = \phi(u_0), \tag{3.2}$$

and $\overline{U}_\tau$ is a family of discrete solutions, any infinitesimal sequence of time steps $\tau_n \downarrow 0$ admits a convergent subsequence (still denoted by τ_n) and a limit curve $u \in AC_{\text{loc}}^2([0, +\infty); X)$ such that

$$\lim_{n\uparrow+\infty} \overline{U}_{\tau_n}(t) = u_t, \quad \lim_{n\uparrow+\infty} \phi(\overline{U}_{\tau_n}(t)) = \phi(u_t) \quad \text{for every } t \geq 0 \tag{3.3}$$

uniformly in each compact interval $[0, T]$. u is a curve of maximal slope (see Definition 2.10), satisfying the energy identity

$$\frac{1}{2}\int_s^t |u'|^2(r)\, dr + \frac{1}{2}\int_s^t |\partial\phi|^2(u_r)\, dr = \phi(u_s) - \phi(u_t) \quad \text{for all } 0 < s \leq t < +\infty. \tag{3.4}$$

Corollary 3.2 (Existence of Curves of Maximal Slope)**.** *Under the same assumptions of the previous theorem, for every $u_0 \in D(\phi)$ there exists a curve of maximal slope $u \in AC^2_{\mathrm{loc}}([0, +\infty); X)$ starting from u_0 and satisfying (3.4).*

3.2. Generation of λ-gradient flows by strong convexity of Φ.

In the case when X is an Hilbert space and ϕ is a l.s.c. λ-convex functional, it is well known that the minimizing movement variational scheme admits a unique solution and the corresponding discrete solution U_τ converges to the solution of (EVI_λ). Applying this approximation scheme, it is then possible to show the existence of the λ-gradient flow of ϕ according to Definition 2.4.

Similar results for minimizing movements of convex functionals in Banach spaces do not always hold: indeed, the characterization of gradient flows through the EVI depends not only on the convexity of ϕ but also on structural properties of the distance d.

One fundamental property is the 1-convexity of the function $v \mapsto \frac{1}{2}\mathsf{d}(v, w)^2$, i.e.

$$\mathsf{d}^2(v_s, w) \leq (1 - s)\mathsf{d}^2(v_0, w) + s\mathsf{d}^2(v_1, w) - s(1 - s)\mathsf{d}^2(v_0, v_1)$$

$$\text{for every } v_0, v_1, \text{ for every } [0, 1] \ni s \mapsto v_s \text{ geodesic between } v_0 \text{ and } v_1, \tag{3.5}$$

which in Banach spaces is equivalent to the fact that d is induced by a scalar product.

(3.5) is satisfied by the geodesic distance on Riemannian manifolds of non-positive sectional curvature and characterizes the Aleksandrov *non-positively curved* (NPC) length spaces, see e.g. [37, 15].

Actually, using (3.5) and adapting a Crandall-Liggett argument, MAYER [46] was able to prove (1.26) and then (1.27) also for geodesically convex functionals on NPC spaces.

A crucial consequence of (3.5) and the λ-convexity of ϕ is that the generating functional $\Phi(\tau, V; U)$ of the minimizing movement scheme (2.31)

$$\Phi(\tau, V; U) := \frac{1}{2\tau}\mathsf{d}^2(U, V) + \phi(U) \quad \tau > 0, \ U, V \in X, \tag{3.6}$$

satisfies the $\tau^{-1} + \lambda$-convexity condition along geodesics, i.e.

$$\text{the map } U \mapsto \Phi(\tau, V; U) \text{ is geodesically } (\tau^{-1} + \lambda)\text{-convex for every } V \in X. \tag{3.7}$$

One of the main contributions of [3, Chapter 4] is to show that (3.7) can be relaxed, by assuming the $(\tau^{-1} + \lambda)$-convexity of $\Phi(\tau, V; \cdot)$ along more general families of curves in X connecting two arbitrary points in $D(\phi)$.

This improvement has been essential to apply the generation result in Wasserstein spaces, which do not satisfy (3.5) except for the 1-dimensional case.

Theorem 3.3 (Convergence of the minimizing movement scheme and generation result [3])**.** *Let us assume that the functional Φ defined in (3.6) satisfies the following property: for every $V, U_0, U_1 \in D(\phi)$ there exists a curve $\gamma_s : [0, 1] \to X$ with $\gamma_0 = U_0$ and $\gamma_1 = U_1$, such that*

$$U \mapsto \Phi(\tau, V; U) \text{ is } \left(\frac{1}{\tau} + \lambda\right)\text{-convex on } \gamma \text{ for each } 0 < \tau < \frac{1}{\lambda^-}, \tag{3.8}$$

i.e.

$$\Phi(\tau, V; \gamma_s) \leq (1 - s)\Phi(\tau, V; U_0) + s\Phi(\tau, V; U_1) - \frac{1 + \lambda\tau}{2\tau}s(1 - s)\mathsf{d}^2(U_0, U_1). \tag{3.9}$$

(1) *For every $U_\tau^0 = u_0 \in \overline{D(\phi)}$ and $\tau > 0$ with $1 + \tau\lambda > 0$ the minimizing movement method 2.12 admits a unique solution $(U_\tau^n)_{n \in \mathbb{N}} \subset D(\phi)$*

(2) *The corresponding discrete solutions $\bar{U}_\tau$ converge to u as $\tau \downarrow 0$ uniformly on compact intervals.*

(3) *The limit u is the unique solution of (EVI_λ). In particular, ϕ admits a λ-gradient flow according to Definition 2.4, thus satisfying all the properties stated in Theorem 2.6.*

(4) *There exist universal constants $C_{\lambda,T}$ such that if $u_0 \in D(|\partial\phi|)$ the optimal error estimate holds:*

$$\mathsf{d}(u(t), \bar{U}_\tau(t)) \leq C_{\lambda,T}|\partial\phi|(u_0)\tau \quad \text{for every } t \in [0, T]. \tag{3.10}$$

The main arguments of the proof of Theorem 3.3 in a simplified setting can be found in [62]. Sub-optimal convergence estimates, inspired by the Crandall-Ligget approach, have also been obtained in a different way in [5] and in [23].

3.3. Generation results for geodesically convex functionals in spaces with a semiconcave squared distance.
In this subsection we consider a geodesically λ-convex functional in a complete metric space (X, d) whose squared distance satisfies a *semi-concavity* condition.

Definition 3.4 (Semi-concavity of the squared distance function). *We say that $D(\phi) \subset X$ is a K-SC (Semi-Concave) space if for every geodesic $[0,1] \ni s \mapsto v_s \in D(\phi)$ and for every $w \in D(\phi)$ we have*

$$\mathsf{d}^2(v_s, w) \geq (1-s)\mathsf{d}^2(v_0, w) + s\mathsf{d}^2(v_1, w) - \mathrm{K}s(1-s)\mathsf{d}^2(v_0, v_1) \quad \textit{for every } s \in [0,1]. \qquad (3.11)$$

Examples of K-SC spaces

-PC SPACES: X is positively curved (PC) in the Aleksandrov sense if and only if X is K-SC with $\mathrm{K} = 1$.

-ALEKSANDROV SPACES: if X is an Aleksandrov space whose curvature is bounded from below by a negative constant $-k$ and $D = \mathrm{diam}(X) < +\infty$, then X is a K-SC space with $\mathrm{K} = \frac{D\sqrt{k}}{\tanh(D\sqrt{k})}$. This class includes all Riemannian manifolds whose sectional curvature is bounded from below.

-PRODUCT AND L^2-SPACES: if $\{(X_i, \mathsf{d}_i)\}_{i \in \mathbb{N}}$ is a countable collection of K-SC spaces, then the product $\prod\limits_{i \in \mathbb{N}} X_i$ with the usual product distance is a K-SC space. If μ is a finite measure on some separable measure space Ω, then $\mathscr{X} := L^2_\mu(\Omega; X) = \{f : \Omega \to X : \int_\Omega \mathsf{d}^2(f(\omega), x_0) \, \mathrm{d}\mu(\omega) < +\infty \text{ for some } x_0 \in X\}$ endowed with the distance $d^2_{\mathscr{X}}(f, g) = \int_\Omega \mathsf{d}^2(f(\omega), g(\omega)) \, \mathrm{d}\mu(\omega)$ is K-SC whenever X is K-SC.

-WASSERSTEIN SPACE: $(\mathscr{P}_2(X), W_2)$ is K-SC if and only if X is K-SC (see the next section).

We will also assume that the (upper) angle between couple of geodesics emanating from the same point satisfies a suitable condition.

Definition 3.5 (Upper angles). *Let x^1, x^2 be two geodesics emanating from the same initial point $x_0 := x_0^1 = x_0^2$. Their upper angle $\sphericalangle_u(x^1, x^2) \in [0, \pi]$ is defined by*

$$\cos(\sphericalangle_u(x^1, x^2)) := \liminf_{s,t \downarrow 0} \frac{\mathsf{d}^2(x_0, x_s^1) + \mathsf{d}^2(x_0, x_t^2) - \mathsf{d}^2(x_s^1, x_t^2)}{2\mathsf{d}(x_0, x_s^1)\mathsf{d}(x_0, x_t^2)}$$

Definition 3.6 (Local angle condition (LAC)). *We say that $D(\phi) \subset X$ satisfies the local angle condition (LAC) if for any triple of geodesics $x^i : [0,1] \to D(\phi)$, $i = 1,2,3$, emanating from the same initial point x_0 the corresponding angles $\theta^{ij} := \sphericalangle_u(x^i, x^j)$ satisfy one of the following equivalent conditions:*
1. $\theta^{12} + \theta^{23} + \theta^{31} \leq 2\pi$.
2. There exist a Hilbert space H and vectors $w^i \in H$ such that $\langle w^i, w^j \rangle_H = \cos(\theta^{ij})$ for $1 \leq i, j \leq 3$.
3. For any choice of $\xi^1, \xi^2, \xi^3 \geq 0$ one has that $\sum\limits_{i,j=1}^{3} \cos(\theta^{ij})\xi^i\xi^j \geq 0$.

Examples of (LAC) spaces

-A BANACH SPACE X satisfies (LAC) if and only if X is a HILBERT space.

-RIEMANNIAN MANIFOLDS AND ALEKSANDROV SPACES with curvature bounded from below satisfy (LAC). In particular if (3.11) holds with $\mathrm{K} = 1$ then X satisfies (LAC).

-PRODUCT AND L^2-SPACES: $\prod\limits_{i \in \mathbb{N}} X_i$ satisfies (LAC) if and only if each (X_i, d_i) does; $L^2_\mu(\Omega; X)$ satisfies (LAC) if and only if X satisfies it.

-WASSERSTEIN SPACE: $\mathscr{P}_2(X)$ satisfies (LAC) if and only if X does.

Theorem 3.7 (Generation theorem for geodesically λ-convex functionals in K-SC and (LAC) spaces). *Let (X, d) be a complete metric space and let $\phi : (-\infty, +\infty]$ be a proper, l.s.c. and λ-geodesically convex functional.*

(1) For every $\tau, \eta > 0$ with $1 + \tau\lambda > 0$ and $U_{\tau,\eta}^0 = u_0 \in D(\phi)$ the relaxed minimizing movement scheme (2.34a,b) admits at least one solution $(U_{\tau,\eta}^n)_{n \in \mathbb{N}}$.

(2) If $D(\phi)$ is a K-SC space and satisfies the (LAC) then the discrete solution $\bar{U}_{\tau,\eta}$ converges to u as $\tau \downarrow 0$ uniformly in each compact interval.

(3) The limit u is the unique solution of (EVI_λ). In particular, ϕ admits a λ-gradient flow according to Definition 2.4, thus satisfying all the properties stated in Theorem 2.6.

4. Wasserstein spaces and diffusion equations

4.1. **The Wasserstein space.** Here, just to set the notation, we collect some basic definitions and properties of the Wasserstein space which will be used in the sequel. For a more detailed overview on this topic we refer to [67, 68], [3] and [4].

Transport maps and couplings. We denote by X_i, for some $i \in \mathbb{N}$, a separable and complete metric space. $\mathscr{P}(X)$ is the space of Borel probability measures on X.

If $\mu \in \mathscr{P}(X_1)$ and $\boldsymbol{t} : X_1 \to X_2$ is a Borel map, we denote by $\boldsymbol{t}_{\#}\mu \in \mathscr{P}(X_2)$ the *push-forward of μ through $\boldsymbol{t}$*, defined by

$$\boldsymbol{t}_{\#}\mu(B) := \mu(\boldsymbol{t}^{-1}(B)) \quad \text{for every } B \in \mathscr{B}(X_2). \tag{4.1}$$

We denote by π^i, for $i = 1, ..., n$, the canonical projection operator from a product space $X_1 \times ... \times X_n$ into X_i, defined by

$$\pi^i(x_1, ..., x_n) := x_i.$$

Given $\mu_1 \in \mathscr{P}(X_1)$ and $\mu_2 \in \mathscr{P}(X_2)$, the class $\Gamma(\mu_1, \mu_2)$ of *transport plans* or *couplings* between μ_1 and μ_2 is defined by

$$\Gamma(\mu_1, \mu_2) := \left\{ \boldsymbol{\gamma} \in \mathscr{P}(X_1 \times X_2) : \pi^1_{\#}\boldsymbol{\gamma} = \mu_1, \ \pi^2_{\#}\boldsymbol{\gamma} = \mu_2 \right\}.$$

To each couple of measures $\mu_1 \in \mathscr{P}(X_1)$, $\mu_2 = \boldsymbol{t}_{\#}\mu_1 \in \mathscr{P}(X_2)$ linked by a Borel map $\boldsymbol{t} : X_1 \to X_2$ we can associate the coupling

$$\boldsymbol{\mu} := (\boldsymbol{i}_{X_1} \times \boldsymbol{t})_{\#}\mu_1 \in \mathscr{P}(X_1 \times X_2), \quad \boldsymbol{i}_{X_1} \text{ being the identity map on } X_1. \tag{4.2}$$

If $\boldsymbol{\mu}$ is representable as in (4.2) we say that $\boldsymbol{\mu}$ is induced by $\boldsymbol{t}$ and $\boldsymbol{t}$ is a *transport map* between μ_1 and μ_2. Each coupling $\boldsymbol{\mu} \in \Gamma(\mu_1, \mu_2)$ concentrated on a $\boldsymbol{\mu}$-measurable graph in $X_1 \times X_2$ admits the representation (4.2) for some μ_1-measurable map $\boldsymbol{t}$, which therefore transports μ_1 into μ_2.

Wasserstein distance. Given a complete and separable metric space (X, d) we denote by $\mathscr{P}_2(X)$ the space of Borel probability measures with finite quadratic moment: $\mu \in \mathscr{P}(X)$ belongs to $\mathscr{P}_2(X)$ iff

$$\int_X \mathsf{d}^2(x, x_o) \, \mathrm{d}\mu(x) < +\infty \quad \text{for some (and thus any) point } x_o \in X. \tag{4.3}$$

For every couple of measures $\mu, \nu \in \mathscr{P}_2(X)$ we consider the Kantorovich problem for the cost d^2

$$W_2^2(\mu, \nu) := \min\left\{ \int_{X \times X} \mathsf{d}^2(x, y) \, \mathrm{d}\boldsymbol{\gamma}(x, y) : \boldsymbol{\gamma} \in \Gamma(\mu, \nu) \right\}. \tag{4.4}$$

It is not difficult to check, by the direct method of calculus of variations, that the minimum problem (4.4) admits at least a solution. The subset of $\Gamma(\mu, \nu)$ given by the *optimal transport plans* for (4.4) will be denoted by $\Gamma_{\text{opt}}(\mu, \nu)$. Notice that if there exists $\boldsymbol{\gamma} = (\boldsymbol{i}_{X_1} \times \boldsymbol{t})_{\#}\mu \in \Gamma(\mu, \nu)$, we have

$$\int_{X \times X} \mathsf{d}^2(x, y) \, \mathrm{d}\boldsymbol{\gamma}(x, y) = \int_X \mathsf{d}^2(x, \boldsymbol{t}(x)) \, \mathrm{d}\mu(x).$$

The quantity $W_2(\mu, \nu)$ defined by (4.4) is a distance between the measures $\mu, \nu \in \mathscr{P}_2(X)$ which enjoys remarkable properties.

Theorem 4.1. *Let* (X, d) *be a complete and separable metric space. Then,* W_2 *defines a distance on* $\mathscr{P}_2(X)$ *and* $(\mathscr{P}_2(X), W_2)$ *is a complete and separable metric space. Moreover, for a given sequence* $\{\mu_k\}_{k \in \mathbb{N}} \subset \mathscr{P}_2(X)$ *we have*

$$\lim_{k \to +\infty} W_2(\mu_k, \mu) = 0 \quad \Leftrightarrow \quad \begin{cases} \displaystyle\int_X f \, \mathrm{d}\mu_k \to \int_X f \, \mathrm{d}\mu & \text{for every } f \in C_b^0(X) \\[2mm] \displaystyle\lim_{R \uparrow +\infty} \int_{X \setminus B_R(x_0)} \mathsf{d}^2(x, x_0) \, \mathrm{d}\mu_k(x) = 0 & \text{uniformly w.r.t. } k \in \mathbb{N}. \end{cases}$$

The metric space $(\mathscr{P}_2(X), W_2)$ is called the $(L^2$-$)$ *Wasserstein space* on X. When $X = \mathbb{R}^d$ we denote by $\mathscr{P}_2^a(\mathbb{R}^d)$ the subset of $\mathscr{P}_2(\mathbb{R}^d)$ defined by

$$\mathscr{P}_2^a(\mathbb{R}^d) := \{\mu \in \mathscr{P}_2(\mathbb{R}^d) : \mu \ll \mathscr{L}^d\}. \tag{4.5}$$

Here we recall the following basic result on the existence and uniqueness of optimal transport plans induced by maps (which are then called *optimal transport maps*) in the case in which the initial measure μ belongs to $\mathscr{P}_2^a(\mathbb{R}^d)$.

Theorem 4.2 (Existence and uniqueness of optimal transport maps, [38, 12])**.** *For any* $\mu \in \mathscr{P}_2^a(\mathbb{R}^d)$ *and* $\nu \in \mathscr{P}_2(\mathbb{R}^d)$*, Kantorovich's optimal transport problem* (4.4) *has a unique solution* $\boldsymbol{\gamma}$*, which is concentrated on the graph of a transport map* $\boldsymbol{t}$*.* $\boldsymbol{t}$ *is the unique minimizer of Monge's optimal transport problem on* $\mathbb{R}^d$ *for the Euclidean distance*

$$\min\left\{ \int_{\mathbb{R}^d} |x - \boldsymbol{r}(x)|^2 \, \mathrm{d}\mu(x) : \boldsymbol{r}_{\#}\mu = \nu \right\}.$$

The map $\boldsymbol{t}$ *is cyclically monotone and there exists a convex open set* $\Omega \subset \mathbb{R}^d$ *with* $\mu(\mathbb{R}^d \setminus \Omega) = 0$ *and a convex function* $\phi : \Omega \to \mathbb{R}$ *such that* $\boldsymbol{t}(x) = \nabla\phi(x)$ *for* μ-*a.e.* $x \in \Omega$.

Geodesics and curvature properties of $(\mathscr{P}_2(\mathbb{R}^d), W_2)$.

Theorem 4.3 (Geodesics in the Wasserstein space)**.** *Given* $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$ *and* $\boldsymbol{\gamma} \in \Gamma_{\mathrm{opt}}(\mu, \nu)$, *the curve*

$$[0, 1] \ni s \mapsto \mu_s = \left((1 - s)\pi^1 + s\pi^2\right)_{\#} \boldsymbol{\gamma}.$$

is a constant speed geodesic between μ *and* ν*, i.e. it satisfies*

$$W_2(\mu_s, \mu_t) = |s - t| W_2(\mu_0, \mu_1) \quad \text{for every } s, t \in [0, 1].$$

Vice versa, any constant speed geodesic between μ *and* ν *can be built in this way.*
If $\boldsymbol{\gamma} = (\boldsymbol{i} \times \boldsymbol{t})_{\#}\mu$*, then*

$$\mu_s = \left((1 - s)\boldsymbol{i} + s\boldsymbol{t}\right)_{\#}\mu, \quad s \in [0, 1].$$

In particular, $(\mathscr{P}_2(\mathbb{R}^d), W_2)$ *is a geodesic space.*

In view of the application to the Wasserstein framework of the theory of gradient flows in metric spaces developed in the previous section, we recall the following theorem (see Theorem 7.3.2 and Example 7.3.3 of [3])

Theorem 4.4 (($\mathscr{P}_2(\mathbb{R}^d), W_2$) is a PC-space)**.** *For any* $\mu_0, \mu_1, \mu_2 \in \mathscr{P}_2(\mathbb{R}^d)$ *we have*

$$W_2^2(\mu_s, \mu_2) \geq (1 - s)W_2^2(\mu_0, \mu_2) + sW_2^2(\mu_1, \mu_2) - s(1 - s)W_2^2(\mu_0, \mu_1) \quad \text{for every } s \in [0, 1], \tag{4.6}$$

where μ_s *is any constant speed geodesic between* μ_0 *and* μ_1.
Moreover, when $d \geq 2$ *there is no constant* $\lambda \in \mathbb{R}$ *such that* $W_2^2(\cdot, \mu_2)$ *is* λ-*convex along geodesics.*

According to Aleksandrov's notion of curvature for metric spaces, (4.6) can be interpreted by saying that the Wasserstein space is a *positively curved* metric space (or PC-space).
Then, the square of the Wasserstein distance along geodesics does not satisfy the 1-convexity assumption (3.5), which would be the most natural to prove the generation Theorem 3.3 for the gradient flows of λ-convex functionals.
However, the theory developed in the previous section allows for a great flexibility in the choice of the connecting curves. In particular, for the Wasserstein space on $\mathbb{R}^d$ the 1-convexity property (3.5) is satisfied along the following class of curves:

Definition 4.5 (Generalized geodesics). *A generalized geodesic joining μ_2 to μ_3 (with base point μ_1) is a curve of the type*

$$[0,1] \ni s \mapsto \mu_s^{2\to3} := ((1-s)\pi^2 + s\pi^3)_{\#}\boldsymbol{\mu}, \tag{4.7}$$

where

$$\boldsymbol{\mu} \in \Gamma(\mu_1,\mu_2,\mu_3) \quad and \quad \pi^{1,2}_{\ \#}\boldsymbol{\mu} \in \Gamma_{\mathrm{opt}}(\mu_1,\mu_2), \quad \pi^{1,3}_{\ \#}\boldsymbol{\mu} \in \Gamma_{\mathrm{opt}}(\mu_1,\mu_3). \tag{4.8}$$

Here $\Gamma(\mu_1,\mu_2,\mu_3) := \{\boldsymbol{\gamma} \in \mathscr{P}(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d) : \pi^i_{\#}\boldsymbol{\gamma} = \mu^i, i = 1,2,3\}$ and $\pi^{i,j} : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ is the projection on the i-th and j-th coordinate.

Proposition 4.6 (1-convexity of the Wasserstein distance along generalized geodesics). *Let $\mu_1,\mu_2,\mu_3 \in \mathscr{P}_2(\mathbb{R}^d)$ and let $\boldsymbol{\mu} \in \Gamma(\mu_1,\mu_2,\mu_3)$ such that $\pi^{1,i}_{\#}\boldsymbol{\mu} \in \Gamma_{\mathrm{opt}}(\mu_1,\mu_i)$, for $i = 2,3$. Then,*

$$W_2^2\big(\mu_s^{2\to3},\mu_1\big) \le (1-s)W_2^2(\mu_1,\mu_2) + sW_2^2(\mu_1,\mu_3) - s(1-s)W_2^2(\mu_2,\mu_3) \quad for\ every\ s \in [0,1].$$

In particular, the function $\frac{1}{2}W_2^2(\mu_1,\cdot)$ is 1-convex along generalized geodesics.

4.2. **Absolutely continuous curves in** $(\mathscr{P}_2(\mathbb{R}^d), W_2)$. We recall here some basic properties of absolutely continuous curves in the Wasserstein space, which are related to the "dynamic interpretation" by BENAMOU-BRENIER [9]. The main result is the following [3, Theorem 8.3.1]:

Theorem 4.7 (Absolutely continuous curves and the continuity equation). *Let $\mu_t : (0,+\infty) \to \mathscr{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve and let $|\mu'| \in L^1(0,+\infty)$ be its metric derivative. Then there exists a Borel vector field $\boldsymbol{v} : (x,t) \mapsto \boldsymbol{v}_t(x)$ such that*

$$\boldsymbol{v}_t \in L^2(\mu_t;\mathbb{R}^d), \quad \|\boldsymbol{v}_t\|_{L^2(\mu_t;\mathbb{R}^d)} \le |\mu'|(t) \quad for\ \mathscr{L}^1\text{-}a.e.\ t \in (0,+\infty) \tag{4.9}$$

and the continuity equation

$$\frac{\partial}{\partial t}\mu_t + \nabla \cdot (\boldsymbol{v}_t\mu_t) = 0 \quad in\ \mathbb{R}^d \times (0,+\infty) \tag{4.10}$$

holds in the sense of distributions.
Moreover,

$$\boldsymbol{v}_t \in \overline{\{\nabla\varphi : \varphi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu_t;\mathbb{R}^d)} \quad for\ \mathscr{L}^1\text{-}a.e.\ t \in (0,+\infty). \tag{4.11}$$

Conversely, if a curve $(0,+\infty) \ni t \mapsto \mu_t \in \mathscr{P}_2(\mathbb{R}^d)$ is continuous w.r.t. the weak topology on $\mathscr{P}(\mathbb{R}^d)$ and it satisfies the continuity equation (4.10) for some Borel vector field $\boldsymbol{v}_t$ with

$$\int_0^{+\infty} \|\boldsymbol{v}_t\|_{L^2(\mu_t;\mathbb{R}^d)} < +\infty,$$

then μ_t is an absolutely continuous curve and $|\mu'|(t) \le \|\boldsymbol{v}_t\|_{L^2(\mu_t;\mathbb{R}^d)}$ for $\mathscr{L}^1$-a.e. $t \in (0,+\infty)$.

Then, the minimal norm for the vector fields $\boldsymbol{v}_t$ satisfying (4.10) for an absolutely continuous curve μ_t is given by its metric derivative. Furthermore, such "minimal" vector fields satisfy (4.11). This fact suggests the following definition (we refer to [3, Chap. 8.4]).

Definition 4.8 (Tangent space). *Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. We define the tangent space to $\mathscr{P}_2(\mathbb{R}^d)$ at the point μ as*

$$\mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d) := \overline{\{\nabla\varphi : \varphi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu_t;\mathbb{R}^d)}. \tag{4.12}$$

Proposition 4.9 (Tangent vectors to absolutely continuous curves). *Let $\mu_t : (0,+\infty) \to \mathscr{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve and let $\boldsymbol{v}_t \in L^2(\mu_t;\mathbb{R}^d)$ be a Borel vector field such that (4.10) holds. Then $\boldsymbol{v}_t$ satisfies (4.9) if and only if $\boldsymbol{v}_t \in \mathrm{Tan}_{\mu_t}\mathscr{P}_2(\mathbb{R}^d)$ for $\mathscr{L}^1$-a.e. $t \in (0,+\infty)$. The vector $\boldsymbol{v}_t$ is uniquely determined $\mathscr{L}^1$-a.e. by (4.9) and (4.10).*

Tangent vector fields are also strictly related to the first order infinitesimal behavior of the Wasserstein distance along absolutely continuous curves.

Proposition 4.10. *Let $\mu_t : (0, +\infty) \to \mathscr{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve and let $\boldsymbol{v}_t \in \text{Tan}_{\mu_t} \mathscr{P}_2(\mathbb{R}^d)$ be the tangent vector characterized by Proposition 4.9. Then, for $\mathscr{L}^1$-a.e. $t \in (0, +\infty)$ the following property holds:*

$$\lim_{h \to 0} \frac{W_2(\mu_{t+h}, (\boldsymbol{i} + h\boldsymbol{v}_t)_\# \mu_t)}{|h|} = 0. \tag{4.13}$$

Then, if μ_t and μ_{t+h} are linked by an optimal transport map $\boldsymbol{t}_{\mu_t}^{\mu_{t+h}}$, we have

$$\lim_{h \to 0} \frac{\boldsymbol{t}_{\mu_t}^{\mu_{t+h}} - \boldsymbol{i}}{h} = \boldsymbol{v}_t \quad \text{in } L^2(\mu_t; \mathbb{R}^d).$$

As an application of (4.13) we are able to show the $\mathscr{L}^1$-a.e. differentiability of $t \mapsto W_2(\mu_t, \sigma)$ along absolutely continuous curves μ_t in terms of tangent vectors and optimal transport plans; this provides a useful formula for the left hand side of the (EVI$_\lambda$)

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} W_2^2(\mu_t, \sigma) \le \phi(\sigma) - \phi(\mu_t) \quad \text{for every } \sigma \in D(\phi)$$

Theorem 4.11. *Let $\mu_t : (0, +\infty) \to \mathscr{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve, let $\boldsymbol{v}_t \in \text{Tan}_{\mu_t} \mathscr{P}_2(\mathbb{R}^d)$ be the tangent vector characterized by Proposition 4.9 and let $\sigma \in \mathscr{P}_2(X)$. Then*

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} W_2^2(\mu_t, \sigma) = \int_{\mathbb{R}^d \times \mathbb{R}^d} \langle x - y, \boldsymbol{v}_t(x) \rangle \, \mathrm{d}\boldsymbol{\gamma}(x, y) \quad \text{for every } \boldsymbol{\gamma} \in \Gamma_{\text{opt}}(\mu_t, \sigma), \tag{4.14}$$

for $\mathscr{L}^1$-a.e $t \in (0, +\infty)$. In particular, if $\mu_t \in \mathscr{P}_2^a(\mathbb{R}^d)$,

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} W_2^2(\mu_t, \sigma) = \int_{\mathbb{R}^d} \langle x - \boldsymbol{t}_{\mu_t}^\sigma(x), \boldsymbol{v}_t(x) \rangle \, \mathrm{d}\mu_t(x), \tag{4.15}$$

where $\boldsymbol{t}_{\mu_t}^\sigma$ is the unique optimal transport map between μ_t and σ.

4.3. Geodesically λ-convex functionals in $\mathscr{P}_2(\mathbb{R}^d)$.
In this section we introduce the three main classes of λ-geodesically convex functionals on the Wasserstein space $(\mathscr{P}_2(\mathbb{R}^d), W_2)$ introduced by MCCANN [47] (for the proofs of the main results we refer to Chapter 9 of [3]).

Example 4.12 (Potential energy). Let $V : \mathbb{R}^d \to \mathbb{R}$ be a λ_V-convex function for some $\lambda_V \in \mathbb{R}$ and let us define the potential energy

$$\mathscr{V}(\mu) := \int_{\mathbb{R}^d} V(x) \, \mathrm{d}\mu(x) \quad \text{for every } \mu \in \mathscr{P}_2(\mathbb{R}^d). \tag{4.16}$$

For every $\mu_1, \mu_2 \in D(\mathscr{V})$ and $\boldsymbol{\mu} \in \Gamma(\mu_1, \mu_2)$ we have

$$\mathscr{V}\left(\left[(1-s)\pi^1 + s\pi^2\right]_\# \boldsymbol{\mu}\right) \le (1-s)\mathscr{V}(\mu_1) + s\mathscr{V}(\mu_2) - \frac{\lambda_V}{2} s(1-s) \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \, \mathrm{d}\boldsymbol{\mu}(x, y). \tag{4.17}$$

In particular, $\mathscr{V}$ is geodesically λ_V-convex on $\mathscr{P}_2(\mathbb{R}^d)$.

Example 4.13 (Interaction energy). Let $\lambda_W \le 0$ and let $W : \mathbb{R}^d \to \mathbb{R}$ be a λ_W-convex function with $W(-x) = W(x)$ for every $x \in \mathbb{R}^d$, and let us set

$$\mathscr{W}(\mu) := \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} W(x - y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y) \quad \text{for every } \mu \in \mathscr{P}_2(\mathbb{R}^d). \tag{4.18}$$

For every $\mu_1, \mu_2 \in D(\mathscr{W})$ and $\boldsymbol{\mu} \in \Gamma(\mu_1, \mu_2)$ we have

$$\mathscr{W}\left(\left[(1-s)\pi^1 + s\pi^2\right]_\# \boldsymbol{\mu}\right) \le (1-s)\mathscr{W}(\mu_1) + s\mathscr{W}(\mu_2) - \frac{\lambda_W}{2} s(1-s) \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \, \mathrm{d}\boldsymbol{\mu}(x, y). \tag{4.19}$$

In particular, $\mathscr{W}$ is geodesically λ_W-convex on $\mathscr{P}_2(\mathbb{R}^d)$.

Example 4.14 (Internal energy). Let $U : [0, +\infty) \to \mathbb{R}$ be a convex function such that

$$U(0) = 0, \qquad \liminf_{s \downarrow 0} \frac{U(s)}{s^\alpha} > -\infty \quad \text{for some } \alpha > \frac{d}{d+2}, \qquad \lim_{s \to +\infty} \frac{U(s)}{s} = +\infty \tag{4.20}$$

the map $s \mapsto s^d U(s^{-d})$ is convex and non-increasing on $(0, +\infty)$. \tag{4.21}

The internal energy functional

$$\mathscr{U}(\mu) := \begin{cases} \displaystyle\int_{\mathbb{R}^d} U(\rho(x))\,\mathrm{d}\mathscr{L}^d(x) & \text{if } \mu \ll \mathscr{L}^d,\ \rho = \dfrac{d\mu}{d\mathscr{L}^d}, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.22}$$

is geodesically convex and lower semicontinuous in $\mathscr{P}_2(\mathbb{R}^d)$. Among the functionals $\mathscr{U}$ with U satisfying (4.20) and (4.21) we have

$$\text{the entropy functional } U(s) = s \log s \tag{4.23}$$

$$\text{the power functional } U(s) = \frac{s^m}{m-1}, \quad m > 1. \tag{4.24}$$

Property (4.21) is also satisfied in the range $1 - \frac{1}{d} < m < 1$; however, since in this case U is not superlinear at infinity (the third condition of (4.20)), we have to consider the relaxed lower semicontinuous functional

$$\mathscr{U}^*(\mu) = \int_{\mathbb{R}^d} \frac{\rho^m}{m-1}\,\mathrm{d}\mathscr{L}^d, \quad \text{if } \mu = \rho\mathscr{L}^d + \mu^\perp,\ \mu^\perp \perp \mathscr{L}^d.$$

Example 4.15 (Relative entropy). Let μ, γ be two measures in $\mathscr{P}(\mathbb{R}^d)$. Then, the relative entropy of μ w.r.t. γ is the functional defined by

$$\text{Ent}_\gamma(\mu) := \begin{cases} \displaystyle\int \frac{d\mu}{d\gamma} \log \frac{d\mu}{d\gamma}\,\mathrm{d}\gamma, & \text{if } \mu \ll \gamma, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.25}$$

We note that (4.25) corresponds to the internal energy associated to the function (4.23) when $\gamma = \mathscr{L}^d$ (which nevertheless is not in $\mathscr{P}(\mathbb{R}^d)$).

Proposition 4.16. *The relative entropy* Ent_γ *is geodesically convex in* $\mathscr{P}_2(\mathbb{R}^d)$ *if and only if one of the following conditions holds:*

1. $\gamma = \mathrm{e}^{-V}\mathscr{L}^d$ *for some convex function* $V : \mathbb{R}^d \to \mathbb{R}$; $\hspace{2cm}$ (4.26)

2. γ *is log-concave, i.e. for every couple of open sets* $A, B \subset \mathbb{R}^d$, $t \in [0,1]$

$$\log \gamma((1-t)A + tB) \geq (1-t)\log \gamma(A) + t\log \gamma(B). \tag{4.27}$$

Now we introduce the notion of convexity which will be crucial to apply the metric theory of gradient flows developed in Section 3 to the main examples 4.12, 4.13, 4.14, 4.15 of geodesically λ-convex functionals in the Wasserstein space.

Definition 4.17 (Convexity along generalized geodesics). *Given* $\lambda \in \mathbb{R}$, *we say that* $\phi : \mathscr{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ *is* λ-*convex along generalized geodesics if for any* $\mu_1, \mu_2, \mu_3 \in D(\phi)$ *there exists a generalized geodesic* $[0,1] \ni s \mapsto \mu_s^{2\to3}$ *joining* μ_2 *to* μ_3 *induced by a plan* $\boldsymbol{\mu}$ *satisfying (4.8) such that*

$$\phi(\mu_s^{2\to3}) \leq (1-s)\phi(\mu_2) + s\phi(\mu_3) - \frac{\lambda}{2}s(1-s)\int |x_2 - x_3|^2\,\mathrm{d}\boldsymbol{\mu}(x_1, x_2, x_3) \quad \text{for every } s \in [0,1]. \tag{4.28}$$

Lemma 4.18. $[(1/\tau + \lambda)$-convexity of $\Phi(\tau, \mu_1; \cdot)]$ *Let* $\phi : \mathscr{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ *be a proper functional which is* λ-*convex along generalized geodesics for some* $\lambda \in \mathbb{R}$. *Then, for each* $\mu_1 \in D(\phi)$ *and for each* $0 < \tau < \frac{1}{\lambda^-}$ *the functional*

$$\Phi(\tau, \mu_1; \mu) := \frac{1}{2\tau}W_2^2(\mu_1, \mu) + \phi(\mu) \text{ satisfies the convexity assumption (3.8).}$$

By Lemma 4.18, whenever ϕ is proper, l.s.c. and λ-convex along generalized geodesics in $\mathscr{P}_2(\mathbb{R}^d)$ we can apply Theorem 3.3 and get the existence, uniqueness and regularizing estimates for the solutions of the EVI_λ.

The examples of geodesically convex functionals in $\mathscr{P}_2(\mathbb{R}^d)$ which have been introduced in this section are also convex along the generalized geodesics.

Theorem 4.19. *The functionals on $\mathscr{P}_2(\mathbb{R}^d)$ considered by the examples 4.12, 4.13 (with $\lambda \leq 0$), 4.14 (under condition (4.21)), and 4.15 (under condition (4.27)) are λ-convex along generalized geodesics.*

4.4. **Gradient flows in $(\mathscr{P}_2(\mathbb{R}^d), W_2)$ and evolutionary PDE's.** In this section we show some applications of the generation result 3.3 to the existence, well-posedness, and asymptotic behavior of nonnegative solutions $\rho : \mathbb{R}^d \times (0, +\infty) \to \mathbb{R}$ of evolutionary PDE's of the type

$$\frac{\partial}{\partial t}\rho - \nabla \cdot \left(\rho \nabla \frac{\delta\phi}{\delta\rho} \right) = 0 \quad \text{in } \mathbb{R}^d \times (0, +\infty), \tag{4.29}$$

where $\frac{\delta\phi(\rho)}{\delta\rho}$ is the first variation of a suitable integral functional; here we consider the case of functionals which are a positive linear combination of the three kinds of contributions considered in the Examples 4.12, 4.13, and 4.14, i.e. $\phi(\rho) := \alpha_1 \mathscr{U}(\rho) + \alpha_2 \mathscr{V}(\rho) + \alpha_3 \mathscr{W}(\rho)$ where $\alpha_i \geq 0$ and

$$\mathscr{U}(\rho) := \int_{\mathbb{R}^d} U(\rho(x))\,\mathrm{d}x,$$

$$\mathscr{V}(\rho) := \int_{\mathbb{R}^d} V(x)\rho(x)\,\mathrm{d}x, \tag{4.30}$$

$$\mathscr{W}(\rho) := \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} W(x-y)\rho(x)\rho(y)\,\mathrm{d}x\mathrm{d}y,$$

so that

$$\frac{\delta\phi(\rho)}{\delta\rho} = \alpha_1 U'(\rho) + \alpha_2 V + \alpha_3 W * \rho. \tag{4.31}$$

In the particular cases of the Fokker-Planck equation ($\phi = \mathscr{U} + \mathscr{V}$ and $U(r) = r \log r$)

$$\frac{\partial}{\partial t}\rho - \nabla \cdot (\nabla\rho + \rho\nabla V) = 0 \quad \text{in } \mathbb{R}^d \times (0, +\infty), \tag{4.32}$$

and of the nonlinear diffusion equations ($\phi = \mathscr{U}$, $U(r) = \frac{1}{m-1}r^m$)

$$\frac{\partial}{\partial t}\rho - \Delta\rho^m = 0, \quad m \geq 1 - \frac{1}{d}, \tag{4.33}$$

the Wasserstein approach has been introduced by the remarkable papers of JORDAN-KINDERLEHRER-OTTO [36] and OTTO [56] and then extended in many interesting directions, covering a wide range of applications: see e.g. [54, 1, 17, 18, 21, 33, 57, 22, 20, 6, 11, 34, 5, 19, 50, 45, 49].

The results presented here are just examples of the transport approach.

Theorem 4.20. *Let V, W, U be as in the examples 4.12, 4.13 and 4.14, let $\mathscr{V}, \mathscr{W}, \mathscr{U}$ be defined as in (4.30), and let $\phi := \alpha_1 \mathscr{U} + \alpha_2 \mathscr{V} + \alpha_3 \mathscr{W}$. For every $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$ there exists a unique solution $\mu_t \in \mathrm{Lip}_{\mathrm{loc}}(0, +\infty; \mathscr{P}_2(\mathbb{R}^d))$ satisfying $\mathrm{EVI}_\lambda(\mathscr{P}_2(\mathbb{R}^d), W_2, \phi)$, $\lambda := \alpha_2 \lambda_V + \alpha_3 \lambda_W$,*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}W_2^2(\mu_t, \sigma) \leq \phi(\sigma) - \phi(\mu_t) - \frac{\lambda}{2}W_2^2(\mu_t, \sigma) \quad \text{for every } \sigma \in D(\phi) \tag{4.34}$$

with $\lim_{t\downarrow 0} \mu_t = \mu_0$ in $\mathscr{P}_2(\mathbb{R}^d)$; the curve μ satisfies all the properties stated in Theorem 2.6, the continuity equation

$$\frac{\partial}{\partial t}\mu_t + \nabla \cdot (\mu_t \boldsymbol{v}_t) = 0 \quad \text{in } \mathbb{R}^d \times (0, +\infty), \quad \text{with} \tag{4.35}$$

$$\boldsymbol{v}_t \in \mathrm{Tan}_{\mu_t}\mathscr{P}_2(\mathbb{R}^d) \quad \mathscr{L}^1\text{-a.e. in } (0, +\infty) \text{ and } \quad t \mapsto \int_{\mathbb{R}^d} |\boldsymbol{v}_t|^2\,\mathrm{d}\mu_t = |\mu_t'|^2 \in L^\infty_{\mathrm{loc}}(0, +\infty), \tag{4.36}$$

and for $\mathscr{L}^1$-a.e. $t \in (0, +\infty)$ the velocity vector $\boldsymbol{v}_t \in \mathrm{Tan}_{\mu_t}\mathscr{P}_2(\mathbb{R}^d)$ satisfies the "subdifferential inequality"

$$\int_{\mathbb{R}^d} \langle \boldsymbol{v}_t(x), x - y \rangle + \frac{\lambda}{2}|y - x|^2\,\mathrm{d}\boldsymbol{\gamma}_t(x, y) \leq \phi(\sigma) - \phi(\mu_t) \quad \text{for every } \boldsymbol{\gamma}_t \in \Gamma_{\mathrm{opt}}(\mu_t, \sigma). \tag{4.37}$$

We can give an explicit characterization of the system (4.35), (4.37). Here we consider the simpler case when U, V, W are differentiable and satisfy a *doubling condition*: for a function $f : \mathbb{R}^h \to \mathbb{R}$ it means that there exists a constant $C > 0$ such that

$$f(x+y) \leq C(1 + f(x) + f(y)) \quad \text{for every } x, y \in \mathbb{R}^h. \tag{4.38}$$

We also set

$$L_U(r) := rU'(r) - U(r) \quad \text{if } r > 0, \quad L_U(0) = 0. \tag{4.39}$$

Theorem 4.21. *Under the same assumptions of the previous theorem, let us also suppose that U, V, W are differentiable and satisfy the doubling condition* (4.38). *The locally Lipschitz curve μ characterized by* (4.34) *(or by* (4.35), (4.37)*) solves the following evolutionary PDE's in $\mathbb{R}^d \times (0, +\infty)$*

Transport equation, $\phi = \mathscr{V}$, $\boldsymbol{v}_t = -\nabla V$:

$$\frac{\partial}{\partial_t}\mu_t - \nabla \cdot (\mu_t \nabla V) = 0. \tag{4.40}$$

Nonlocal interaction equation, $\phi = \mathscr{W}$, $\boldsymbol{v}_t = -(\nabla W) * \mu_t$

$$\frac{\partial}{\partial_t}\mu_t - \nabla \cdot (\mu_t(\nabla W * \mu_t)) = 0. \tag{4.41}$$

Fokker-Planck equation, $\phi = \mathscr{U} + \mathscr{V}$, $U(r) = r \log r$, $-\mu_t \boldsymbol{v}_t = \nabla \mu_t + \mu_t \nabla V$

$$\frac{\partial}{\partial_t}\mu_t - \nabla \cdot (\nabla \mu_t + \mu_t \nabla V) = 0. \tag{4.42}$$

In this case, $\mu_t = \rho_t \mathscr{L}^d$ *with* $\rho_t \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$ *for $\mathscr{L}^1$-a.e. $t \in (0, +\infty)$.*
Nonlinear diffusion equation, $\phi = \mathscr{U}$, $\mu_t \boldsymbol{v}_t = -\nabla L_U(\rho_t)$ *where* $\mu_t = \rho_t \mathscr{L}^d \ll \mathscr{L}^d$,

$$\frac{\partial}{\partial_t}\mu_t - \Delta(L_U(\rho_t)) = 0, \tag{4.43}$$

with $L_U(\rho_t) \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$ *for $\mathscr{L}^1$-a.e. $t \in (0, +\infty)$.*
Drift-diffusion with non local interactions, $\phi = \mathscr{U} + \mathscr{V} + \mathscr{W}$, $-\mu_t \boldsymbol{v}_t = \nabla L_U(\rho_t) + \mu_t \nabla V + \mu_t((\nabla W) * \mu_t)$, $\mu_t = \rho_t \mathscr{L}^d \ll \mathscr{L}^d$,

$$\frac{\partial}{\partial_t}\mu_t - \nabla \cdot \big(\nabla L_U(\rho_t) + \mu_t \nabla V + \mu_t((\nabla W) * \mu_t)\big) = 0 \tag{4.44}$$

We refer to [3, Chap. 11] for the proofs and for more general and detailed results; here we just give a sketch of the argument showing that (4.35), (4.37) yield (4.42) when $\phi = \mathscr{U} + \mathscr{V}$ in the case of $U(r) = r \log r$.

Let us fix a time $t > 0$ where (4.37) holds, a smooth test function $\zeta \in C^\infty_c(\mathbb{R}^d)$, and $\boldsymbol{t}_\varepsilon := \boldsymbol{i} + \varepsilon \nabla \zeta$. If $|\varepsilon| \max_{\mathbb{R}^d} \|\mathrm{D}^2 \zeta\| < 1$ the coupling $\boldsymbol{\gamma}_\varepsilon := (\boldsymbol{i}, \boldsymbol{t}_\varepsilon)_\# \mu_t$ is optimal between μ_t and $(\boldsymbol{t}_\varepsilon)_\# \mu_t$ so that (4.37) yields

$$-\varepsilon \int_{\mathbb{R}^d} \langle \boldsymbol{v}_t(x), \nabla \zeta(x) \rangle \, \mathrm{d}\mu_t(x) \leq \phi((\boldsymbol{t}_\varepsilon)_\# \mu_t) - \phi(\mu_t).$$

Setting

$$\rho_t := \frac{\mathrm{d}\mu_t}{\mathrm{d}\mathscr{L}^d}, \quad \rho_t^\varepsilon := \frac{\mathrm{d}(\boldsymbol{t}_\varepsilon)_\# \mu_t}{\mathrm{d}\mathscr{L}^d}$$

we get

$$-\varepsilon \int_{\mathbb{R}^d} \langle \boldsymbol{v}_t, \nabla \zeta \rangle \, \mathrm{d}\mu_t \leq \int_{\mathbb{R}^d} \rho_t^\varepsilon \log \rho_t^\varepsilon \, \mathrm{d}\mathscr{L}^d - \int_{\mathbb{R}^d} \rho_t \log \rho_t \, \mathrm{d}\mathscr{L}^d + \int_{\mathbb{R}^d} \big(V(\boldsymbol{t}_\varepsilon(x)) - V(x)\big) \, \mathrm{d}\mu_t(x)$$

Applying the change of variables formula

$$\rho_t^\varepsilon(\boldsymbol{t}_\varepsilon(x)) \det[\boldsymbol{i} + \varepsilon \mathrm{D}^2 \zeta(x)] = \rho_t(x),$$

we obtain

$$-\varepsilon \int_{\mathbb{R}^d} \langle \boldsymbol{v}_t, \nabla \zeta \rangle \, \mathrm{d}\mu_t \leq -\int_{\mathbb{R}^d} \rho(x) \log\big(\det[\boldsymbol{i} + \varepsilon \mathrm{D}^2 \zeta(x)]\big) \, \mathrm{d}\mathscr{L}^d + \int_{\mathbb{R}^d} \big(V(\boldsymbol{t}_\varepsilon(x)) - V(x)\big) \, \mathrm{d}\mu_t(x) \tag{4.45}$$

Finally, dividing by ε and taking the limit of (4.45) as ε tends to 0 we get

$$- \int_{\mathbb{R}^d} \langle \boldsymbol{v}_t, \nabla\zeta \rangle \, \mathrm{d}\mu_t = \int_{\mathbb{R}^d} \Big(- \Delta\zeta(x) + \nabla V(x) \cdot \nabla\zeta \Big) \, \mathrm{d}\mu_t \quad \text{for every } \zeta \in C_c^\infty(\mathbb{R}^d),$$

so that μ satisfies the distributional formulation of (4.42).

4.5. **The heat flow on Riemannian manifolds and metric-measure spaces.** We conclude these notes by giving a short account of possible applications of the Wasserstein setting to the generation of the heat flow in Riemannian manifolds and metric-measure spaces.

Let us start with a compact and smooth Riemannian manifold (M, g); we denote by d_g its Riemannian distance and by $\gamma = \mathrm{Vol}_g \in \mathscr{P}(M)$ its (normalized) volume measure.

In $\mathscr{P}_2(M)$ we consider the Relative Entropy functional Ent_γ as in (4.25). VON RENESSE-STURM [66] proved

Theorem 4.22. *The Relative Entropy functional* Ent_γ *is geodesically* λ-*convex in* $\mathscr{P}_2(M)$ *if and only if* M *satisfies the lower Ricci curvature bound*

$$\mathrm{Ric}(M) \geq \lambda \quad \textit{i.e.} \quad \mathrm{Ric}_x(v, v) \geq \lambda |v|_g^2 \quad \textit{for all } x \in M \textit{ and } v \in \mathrm{Tan}_x(M). \tag{4.46}$$

In this case, it is possible to show (see [55, 57, 27, 53, 68, 32]) that the Relative Entropy functional Ent_γ generates a λ-gradient flow $\mathsf{S}_t : \mathscr{P}_2(M) \to \mathscr{P}_2(M)$ according to definition 2.4, which coincides with the classical heat flow on M.

Theorem 4.23. *The relative entropy functional* Ent_γ *generates a* λ-*gradient flow* S_t *in* $\mathscr{P}_2(M)$ *according to Definition 2.4 (and thus satisfying all the properties stated in Theorems 2.6 and 2.15). A curve* $\mu_t \in \mathscr{P}_2(M)$ *is a solution of* $\mathrm{EVI}_\lambda(M, \mathsf{d}_g, \mathrm{Ent}_\gamma)$ *if and only if its density* $\rho_t = \mathrm{d}\mu_t / \mathrm{d}\gamma$ *solves the Heat equation*

$$\frac{\partial}{\partial t}\rho_t - \Delta_g \, \rho_t = 0 \quad \textit{in } M \times (0, +\infty),$$

where Δ_g *is the Laplace-Beltrami operator on* M.

The adimensionality of the form of the Entropy functional (4.25) and the purely metric character of the EVI suggest that one can use them to define a heat flow on more general *measure-metric* spaces (X, d, γ), where (X, d) is a complete and separable metric space and $\gamma \in \mathscr{P}(X)$. Indeed, as it has been often pointed out in the previous sections, the EVI formulation gives nice regularity, stability and asymptotic properties for the related flow. We briefly sketch two possible approaches:

Approximation by measured Gromov-Hausdorff convergence. We consider a sequence of smooth and compact Riemannian manifolds $(M^h, \mathsf{d}^h, \mathrm{Vol}^h)$ converging to a limit measure-metric space (X, d, γ) in the measured Gromov-Hausdorff convergence: it means [64] that a sequence $\{\hat{\mathsf{d}}^h\}_{k \in \mathbb{N}}$ of (complete, separable) *coupling semidistances* on the disjoint union $M^h \sqcup X$ exists such that the restriction of $\hat{\mathsf{d}}^h$ on M^h (resp. X) coincides with d^h (resp. d) and

$$\lim_{k\uparrow\infty} \hat{W}_2^h(\mathrm{Vol}^h, \gamma) = 0, \quad \hat{W}_2^h \text{ is the Wasserstein distance on } \mathscr{P}_2(M^h \sqcup X) \text{ induced by } \hat{\mathsf{d}}^h \tag{4.47}$$

A sequence $\mu^h \in \mathscr{P}_2(M^h)$ converges to $\mu \in \mathscr{P}_2(X)$ if $\lim_{k\uparrow+\infty} \hat{W}_2^h(\mu^h, \mu) = 0$. Adapting the arguments of Theorem 2.17 it is possible to prove the following asymptotic result:

Theorem 4.24 ([63]). *Let us assume that the compact Riemannian manifolds* M^h *satisfy the uniform lower bound on the Ricci curvature* $\mathrm{Ric}(M^h) \geq \lambda$ *for some* $\lambda \in \mathbb{R}$ *independent of* k *and converge to* (X, d, γ) *in the measured Gromov-Hausdorff sense. Then the Relative Entropy functional* Ent_γ *admits a* λ-*gradient flow* S_t *on* $\mathscr{P}_2(X)$ *and for every sequence of initial measures* $\mu_0^h \in \mathscr{P}_2(M^h)$ *converging to* $\mu_0 \in \mathscr{P}_2(X)$ *the corresponding solution* μ_t^h *of the Heat flow on* M^h *converges to* $\mathsf{S}_t(\mu_0)$ *in* $\mathscr{P}_2(X)$ *for every* $t > 0$.

Applying Theorem 2.9 one finds in particular that the limit Entropy functional Ent_γ is *strongly* geodesically λ-convex (at least when the support of γ is X), a stability result that has been proved by [65, 42].

Intrinsic costruction. Starting from Theorem 4.22, Sturm [65] and Lott-Villani [42] introduced the concept of metric-measure spaces (X, d, γ) satisfying a lower Ricci curvature bound, by requiring that the relative entropy functional Ent_γ is geodesically λ-convex in $\mathscr{P}_2(X)$.

Definition 4.25 (Lower Ricci curvature bounds for metric-measure spaces)**.** *We say that a metric-measure space (X, d, γ) has Ricci curvature bounded from below by a certain $\lambda \in \mathbb{R}$ (and we write $\mathrm{Ric}(X) \geq \lambda$) if the relative entropy Ent_γ is λ-geodesically convex on X.*

It is then natural to look for other intrinsic properties of X which are sufficient to deduce the existence of the associated EVI semigroup. It is interesting to notice that if the relative entropy functional generates a λ-gradient flow S_t then S_t is a semigroup of *linear* operators [63]. In the case of compact positively curved (PC) Alexandrov spaces the existence of a λ-contracting gradient flow can be deduced by a general unpublished result of [58] and has been recently proved by Ohta [52].

In more general cases, we can apply Theorem 3.7:

Theorem 4.26. *Let us suppose that (X, d, γ) is a complete and separable metric-measure space with Ricci curvature bounded from below, according to Definition 4.25, and measure γ with full support $\mathrm{supp}(\gamma) = X$. If X satisfies the Local Angle Condition 3.6 and it is K-semiconcave as in 3.4, then the relative entropy functional Ent_γ generates a λ-gradient flow on $\mathscr{P}_2(X)$ which can be uniquely extended to a Markov semigroup (i.e. linear, order preserving, strongly continuous, contractive) in every space $L^p(\gamma)$, $p \in [1, +\infty)$.*

References

[1] Martial Agueh. Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory. *Adv. Differential Equations*, 10(3):309–360, 2005.

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

[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[4] Luigi Ambrosio and Giuseppe Savaré. Gradient flows of probability measures. In *Handbook of Evolution Equations (III)*. Elsevier, 2006.

[5] Luigi Ambrosio, Giuseppe Savaré, and Lorenzo Zambotti. Existence and stability for Fokker-Planck equations with log-concave reference measure. *Probab. Theory Relat. Fields*, 145(3-4):517–564, 2009.

[6] Luigi Ambrosio and Sylvia Serfaty. A gradient flow approach to an evolution problem arising in superconductivity. *Comm. Pure Appl. Math.*, 61(11):1495–1539, 2008.

[7] Claudio Baiocchi. Discretization of evolution variational inequalities. In F. Colombini, A. Marino, L. Modica, and S. Spagnolo, editors, *Partial differential equations and the calculus of variations, Vol. I*, pages 59–92. Birkhäuser Boston, Boston, MA, 1989.

[8] Viorel Barbu. *Nonlinear semigroups and differential equations in Banach spaces.* Editura Academiei Republicii Socialiste România, Bucharest, 1976. Translated from the Romanian.

[9] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[10] Philippe Bénilan. Solutions intégrales d'équations d'évolution dans un espace de Banach. *C. R. Acad. Sci. Paris Sér. A-B*, 274:A47–A50, 1972.

[11] A. Blanchet, V. Calvez, and J.A. Carrillo. Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model. *SIAM J. Numer. Anal.*, 46:691–721, 2008.

[12] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.

[13] H. Brézis. Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations. In *Contribution to Nonlinear Functional Analysis, Proc. Sympos. Math. Res. Center, Univ. Wisconsin, Madison, 1971*, pages 101–156. Academic Press, New York, 1971.

[14] 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).

[15] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[16] Tiziana Cardinali, Giovanni Colombo, Francesca Papalini, and Mario Tosques. On a class of evolution equations without convexity. *Nonlinear Anal.*, 28(2):217–234, 1997.

[17] E. A. Carlen and W. Gangbo. Constrained steepest descent in the 2-Wasserstein metric. *Ann. of Math. (2)*, 157(3):807–846, 2003.

[18] E. A. Carlen and W. Gangbo. Solution of a model Boltzmann equation via steepest descent in the 2-Wasserstein metric. *Arch. Ration. Mech. Anal.*, 172(1):21–64, 2004.

[19] J. A. Carrillo, S. Lisini, G. Savaré, and D. Slepcev. Nonlinear mobility continuity equations and generalized displacement convexity. *J. Funct. Anal.*, 258(4):1273–1309, 2010.

[20] José A. Carrillo, Marco Di Francesco, and Corrado Lattanzio. Contractivity of Wasserstein metrics and asymptotic profiles for scalar conservation laws. *J. Differential Equations*, 231(2):425–458, 2006.

[21] José A. Carrillo, Robert J. McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 2003.

[22] José A. Carrillo, Robert J. McCann, and Cédric Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006.

[23] Philippe Clément. Introduction to gradient flows in metric spaces. *Lecture Notes, University of Bielefeld*, 2009. Available online at `https://igk.math.uni-bielefeld.de/study-materials/notes-clement-part2.pdf`.

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

[25] M. G. Crandall and A. Pazy. Semi-groups of nonlinear contractions and dissipative sets. *J. Functional Analysis*, 3:376–418, 1969.

[26] Gianni Dal Maso. *An Introduction to Γ-Convergence*, volume 8 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, Boston, 1993.

[27] Sara Daneri and Giuseppe Savaré. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.*, 40(3):1104–1122, 2008.

[28] Ennio De Giorgi. New problems on minimizing movements. In Claudio Baiocchi and Jacques Louis Lions, editors, *Boundary Value Problems for PDE and Applications*, pages 81–98. Masson, 1993.

[29] Ennio De Giorgi, Marco Degiovanni, Antonio Marino, and Mario Tosques. Evolution equations for a class of nonlinear operators. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 75(1-2):1–8 (1984), 1983.

[30] Ennio De Giorgi, Antonio Marino, and Mario Tosques. Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 68(3):180–187, 1980.

[31] Marco Degiovanni, Antonio Marino, and Mario Tosques. Evolution equations with lack of convexity. *Nonlinear Anal.*, 9(12):1401–1443, 1985.

[32] Matthias Erbar. The heat equation on manifolds as a gradient flow in the Wasserstein space. *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques*, 46(1):1–23, 2010.

[33] L. C. Evans, O. Savin, and W. Gangbo. Diffeomorphisms and nonlinear heat flows. *SIAM J. Math. Anal.*, 37(3):737–751 (electronic), 2005.

[34] S. Fang, J. Shao, and T.K. Sturm. Wasserstein space over the wiener space. *webdoc.sub.gwdg.de*, Jan 2008.

[35] U. Gianazza and G. Savaré. Abstract evolution equations on variable domains: an approach by minimizing movements. *Ann. Sc. Norm. Sup. Pisa Cl. Sci. (4)*, 23:149–178, 1996.

[36] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17 (electronic), 1998.

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

[38] M. Knott and C. S. Smith. On the optimal mapping of distributions. *J. Optim. Theory Appl.*, 43(1):39–49, 1984.

[39] Y. Kōmura. Nonlinear semi-groups in Hilbert space. *J. Math. Soc. Japan*, 19:493–507, 1967.

[40] J.-L. Lions. *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires*. Dunod, Gauthier-Villars, Paris, 1969.

[41] J.-L. Lions and G. Stampacchia. Variational inequalities. *Comm. Pure Appl. Math.*, 20:493–519, 1967.

[42] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.

[43] Stephan Luckhaus. Solutions for the two-phase Stefan problem with the Gibbs-Thomson Law for the melting temperature. *Euro. Jnl. of Applied Mathematics*, 1:101–111, 1990.

[44] Antonio Marino, Claudio Saccon, and Mario Tosques. Curves of maximal slope and parabolic variational inequalities on nonconvex constraints. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 16(2):281–330, 1989.

[45] Daniel Matthes, Robert J. McCann, and Giuseppe Savaré. A family of nonlinear fourth order equations of gradient flow type. *Comm. Partial Differential Equations*, 34(10-12):1352–1397, 2009.

[46] Uwe F. Mayer. Gradient flows on nonpositively curved metric spaces and harmonic maps. *Comm. Anal. Geom.*, 6(2):199–253, 1998.

[47] Robert J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.

[48] Alexander Mielke, Florian Theil, and Valery I. Levitas. A variational formulation of rate-independent phase transformations using an extremum principle. *Arch. Ration. Mech. Anal.*, 162(2):137–177, 2002.

[49] Luca Natile, Mark A. Peletier, and Giuseppe Savaré. Contraction of general transportation costs along solutions to Fokker-Planck equations with monotone drifts. *arXiv:1002.0088v1*, 2010.

[50] Luca Natile and Giuseppe Savarè. A Wasserstein approach to the one-dimensional sticky particle system. *arxiv:0902.4373v2*, 2009.

[51] Ricardo H. Nochetto, Giuseppe Savaré, and Claudio Verdi. A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations. *Comm. Pure Appl. Math.*, 53(5):525–589, 2000.

[52] Shin-ichi Ohta. Gradient flows on wasserstein spaces over compact alexandrov spaces. Technical report, Universität Bonn, 2007.

[53] Shin-ichi Ohta. Gradient flows on Wasserstein spaces over compact Alexandrov spaces. *Amer. J. Math.*, 131(2):475–516, 2009.

[54] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[55] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[56] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[57] Felix Otto and Michael Westdickenberg. Eulerian calculus for the contraction in the Wasserstein distance. *SIAM J. Math. Anal.*, 37(4):1227–1255 (electronic), 2005.

[58] G. Perelman and A. Petrunin. Quasigeodesics and gradient curves in alexandrov spaces. Unpublished preprint, available online at `www.math.psu.edu/petrunin/papers/papers.html`.

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

[60] G. Savaré. Gradient flows and evolution variational inequalities in metric spaces. *In preparation*, 2010.

[61] Giuseppe Savaré. Weak solutions and maximal regularity for abstract evolution inequalities. *Adv. Math. Sci. Appl.*, 6(2):377–418, 1996.

[62] Giuseppe Savaré. Error estimates for dissipative evolution problems. In *Free boundary problems (Trento, 2002)*, volume 147 of *Internat. Ser. Numer. Math.*, pages 281–291. Birkhäuser, Basel, 2004.

[63] Giuseppe Savaré. Gradient flows and diffusion semigroups in metric spaces under lower curvature bounds. *C. R. Math. Acad. Sci. Paris*, 345(3):151–154, 2007.

[64] Karl-Theodor Sturm. Convex functionals of probability measures and nonlinear diffusions on manifolds. *J. Math. Pures Appl. (9)*, 84(2):149–168, 2005.

[65] Karl-Theodor Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.

[66] Karl-Theodor Sturm and Max-K. von Renesse. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.

[67] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[68] Cédric Villani. *Optimal transport. Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2009.

S.I.S.S.A., Via Beirut 2-4, 34151, Trieste, Italy
E-mail address: `daneri@sissa.it`

Dipartimento di Matematica, Università di Pavia. Via Ferrata, 1 – 27100 Pavia, Italy.
E-mail address: `giuseppe.savare@unipv.it`
URL: `http://www.imati.cnr.it/~savare`