

# EXAMPLE OF A FIRST ORDER DISPLACEMENT CONVEX FUNCTIONAL

JOSÉ A. CARRILLO<sup>†</sup> AND DEJAN SLEPČEV<sup>‡</sup>

ABSTRACT. We present a family of first-order functionals which are displacement convex, that is convex along the geodesics induced by the quadratic transportation distance on the circle. The displacement convexity implies the existence and uniqueness of gradient flows of the given functionals. More precisely, we show the existence and uniqueness of gradient-flow solutions of a class of fourth-order degenerate parabolic equations with periodic boundary data. Moreover, positivity of the absolutely continuous part of the solutions is preserved along the flow.

## 1. INTRODUCTION

The notion of geodesic convexity of functionals on the space of probability measures endowed with transportation distances was introduced by McCann [13] who named it displacement convexity. In the same paper McCann discovered the three main classes of displacement convex functionals on Euclidean space: the internal energy, the external potential energy, and the interaction energy:

$$E_1 = \int_{\mathbb{R}^d} U(\rho) dx, \quad E_2 = \int_{\mathbb{R}^d} V(x) d\rho, \quad \text{and} \quad E_3 = \int_{\mathbb{R}^d} W * \rho d\rho.$$

Here  $U : \mathbb{R}^+ \rightarrow \mathbb{R}$  is the density of the internal energy,  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  is an external potential and  $W : \mathbb{R}^d \rightarrow \mathbb{R}$  is an interaction potential. The functionals are displacement convex under appropriate conditions on  $U$ ,  $V$ , and  $W$ .

Since McCann's pioneering paper, an intriguing question was: Are there other displacement-convex functionals, in particular are there ones that involve derivatives? Versions of this question were asked by Villani [22, Open Problem 5.17] and [23, pg. 892]. Here we present first order functionals which are displacement convex in periodic setting in one dimension. To the best of our knowledge, these are the

---

*Date:* October 14, 2008.

first displacement convex functionals that involve derivatives of the measure density. While the question on  $\mathbb{R}^d$  remains open for any dimension, our examples show that there are no fundamental, structural obstructions to existence of displacement convex functionals that involve derivatives.

We show that first-order functionals of the form

$$(1) \quad E(\rho) = \int_0^1 \left[ \left( \frac{1}{\rho^\beta} \right)_x \right]^2 dx, \quad \beta \neq 0$$

lead to well-defined displacement convex functionals on the set of probability measures on the circle,  $\mathcal{P}(S^1)$ , endowed with the quadratic Wasserstein distance,  $d_W$  if and only if  $\beta \in \left[1, \frac{3}{2}\right]$ . It is worth noting that there are several important functionals that can be written in the above form. Namely the value  $\beta = -1$  corresponds to the Dirichlet energy, while  $\beta = -\frac{1}{2}$  corresponds to the Fisher information. However these functionals are not displacement convex.

Geodesic convexity has important consequences for the existence and uniqueness of gradient flows in the space of probability measures as recognized by Otto [16] and developed by Ambrosio, Gigli, and Savaré [2], and others. Furthermore, uniform geodesic convexity provides quantitative information on asymptotics of gradient flows [16]. Moreover it implies a number of important functional inequalities, see [22, Ch. 9] and references therein.

The existence of gradient flows in Wasserstein metric is based on variational schemes or minimizing movements [9, 1, 2], first considered in this setting by Jordan, Kinderlehrer and Otto. The application of the scheme is greatly simplified when the functional considered is displacement convex. More importantly, the geodesic convexity induces quantitative contraction estimates between solutions of the gradient flows [7, 2] and it implies the uniqueness of these solutions. These results have found applications to nonlinear diffusion equations with/without nonlocal interactions; see for instance [6, 2] and the references therein.

In recent years there have been a number of developments regarding displacement convexity. Carlen et al. [4], have studied displacement convexity of the internal and the interaction energy in the setting of interfacial problems, where the configurations in the state space have infinite mass. Displacement convexity on Riemannian manifolds for internal energy (under Ricci curvature bounds) was predicted by

Otto and Villani [17] and proved by Cordero-Erausquin, McCann, and Schmuckenschläger in [5]. Recently Otto and Westdickenberg [18] have introduced techniques that were further developed by Daneri and Savare [8] to show geodesic convexity of functionals on manifolds using a purely local, Eulerian framework. The connection between bounds on displacement convexity and bounds on Ricci curvature has found important applications, see papers by Lott and Villani [12], Sturm and von Renesse [21], Sturm [20, 19] and references therein. Let us also mention important applications to geometric flows, where convexity along optimal transportation paths is considered on manifolds whose metric is being deformed by a geometric flow, for example the flow by Ricci curvature. The reader can see works of Lott [11], McCann and Topping [15], and references therein. Finally Lisini, Savaré and the authors [3] have considered an adapted notion of displacement convexity to problems where the metric is a generalization of the quadratic transportation distance with a nonlinear weight.

The geodesic convexity of functionals introduced by McCann, has lead to understanding of gradient flows given by the equations of the form

$$(2) \quad \frac{\partial \rho}{\partial t} = \nabla \cdot [\rho \nabla (U'(\rho) + V + W * \rho)],$$

where the unknown  $\rho(t, \cdot)$  is a time-dependent probability measure on  $\mathbb{R}^d$  ( $d \geq 1$ ). Using the techniques of [2] we show existence and uniqueness of the gradient flow of (1). The gradient flow solutions are weak (variational) solutions of the fourth-order diffusion equation:

$$(3) \quad \frac{\partial \rho}{\partial t} = \beta^2 (\rho [-\theta \rho^{\theta-1} \rho_x^2 - 2\rho^\theta \rho_{xx}]_x)_x$$

with periodic boundary conditions, where  $\theta = -2(\beta + 1)$ .

The plan of this paper is as follows. We quickly review and discuss the optimal transport on the circle in Section 2, and present a convenient form of the displacement interpolation. We show the geodesic convexity of the functional (1) in Section 3. The statement is given in Theorem 1, while the proof is divided in a sequence of lemmas. In the Subsection 3.2 we show that when  $\beta \notin [1, \frac{3}{2}]$  the functional (1) is not displacement convex on  $S^1$ . We also show that on  $\mathbb{R}$  the functional is either trivial or not convex. Section 4 is devoted to applying the theory developed in [2] to the gradient flow of (1).

2. OPTIMAL TRANSPORT ON  $S^1$ 

We consider the points on  $S^1$  as equivalence classes of real numbers:  $S^1 = \mathbb{R}/\mathbb{Z}$ . We identify functions on  $S^1$  with functions on  $\mathbb{R}$  of period one. However, when convenient, we also identify the points on the circle with the interval  $[0, 1)$ . The transportation cost we consider is the square of the geodesic distance on  $S^1$ . Let  $\mu_0$  and  $\mu_1$  be probability measures on  $S^1$ , which are absolutely continuous with respect to the Lebesgue measure,  $\mathcal{L}$ . Let  $\rho_0$  and  $\rho_1$  their densities. By results of McCann [14] there exists an optimal transportation map  $\Psi : S^1 \rightarrow S^1$ ,  $\Psi_{\#}\mu_0 = \mu_1$ .

Let us consider continuous positive densities  $\rho_0$  and  $\rho_1$ :

$$\rho_0 > 0 \text{ and } \rho_1 > 0,$$

then the map  $\Psi$  is continuous, see [10] and the references therein. We can consider the mapping  $\Psi$  as a mapping from  $[0, 1]$  to  $\mathbb{R}/\mathbb{Z}$ .

Consider the mapping  $\tilde{\Psi} : [0, 1] \rightarrow [-\frac{1}{2}, \frac{3}{2}]$  such that  $\tilde{\Psi}(x)$  is the smallest (often unique) element of  $\Psi(x)$  for which  $|\tilde{\Psi}(x) - x| \leq \frac{1}{2}$ . The Euclidean distance between  $x$  and  $\tilde{\Psi}(x)$  is always less than or equal to  $\frac{1}{2}$  and thus coincides with the geodesic distance on  $S^1$ . Let  $\tilde{\mu}_1 = \tilde{\Psi}_{\#}\mu_0$ . Note that  $\tilde{\Psi}$  is the optimal transport with Euclidean quadratic cost between  $\mu_0$  and  $\tilde{\mu}_1$ . Thus  $\tilde{\Psi}$  is a monotone mapping. Let  $a = \tilde{\Psi}(0)$ . Then  $a + 1 = \tilde{\Psi}(1)$ . Monotonicity implies that  $\text{Range}(\tilde{\Psi}) = [a, a + 1]$ . Furthermore regularity of optimal transportation maps on  $\mathbb{R}$  implies that  $\tilde{\Psi}$  is smooth. Since we could have considered an interval  $[b, b + 1]$  instead of  $[0, 1]$  to define  $\tilde{\Psi}$ , it follows that  $\Psi$  is smooth.

Note that displacement interpolation is represented by the diffeomorphisms  $\tilde{\Psi}_s : [0, 1] \rightarrow [as, as + 1]$  defined by

$$\tilde{\Psi}_s(x) = (1 - s)x + s\tilde{\Psi}(x) = x + s(\tilde{\Psi}(x) - x).$$

The displacement interpolant is given by

$$(4) \quad \mu_s = (\tilde{\Psi}_s)_{\#}\mu_0.$$

Let

$$f(x) = \tilde{\Psi}(x) - x.$$

If  $\rho_0$  and  $\rho_1$  are smooth then so is  $\tilde{\Psi}$  and consequently so are  $\tilde{\Psi}_s$  and  $\rho_s$ , the density of  $\mu_s$ . Then

$$(5) \quad \rho_s(y) = \frac{\rho_0(x)}{\tilde{\Psi}'_s(x)} = \frac{\rho_0(x)}{1 + sf'(x)} \quad \text{where } y = \tilde{\Psi}_s(x).$$

## 3. DISPLACEMENT CONVEXITY

Let us first consider the energy  $\bar{E} : \mathcal{P}(S^1) \rightarrow [0, \infty]$  defined as follows: For  $\beta \neq 0$ , if  $\mu$  is a probability measure with density  $\rho$  then

$$(6) \quad \bar{E}(\mu) = \int_0^1 \left[ \left( \frac{1}{\rho^\beta} \right)_x \right]^2 dx.$$

where the derivative considered is the weak derivative. In the case that  $\mu$  is not absolutely continuous with respect to the Lebesgue measure or if the above expression is not defined we define  $\bar{E}(\mu) = \infty$ . For convenience we will often write  $\bar{E}(\rho)$  instead of  $\bar{E}(\mu)$ .

The energy  $E$  is defined to be the lower-semicontinuous envelope of  $\bar{E}$  with respect to the Wasserstein distance:

$$(7) \quad E(\mu) = \inf \{ \liminf_{n \rightarrow \infty} \bar{E}(\mu_n) \mid \lim_{n \rightarrow \infty} d_W(\mu_n, \mu) = 0 \}.$$

An important property of  $E$  is that if  $\bar{E}(\mu) < \infty$  then  $E(\mu) = \bar{E}(\mu)$ . It follows from Lemmas 5 and 6.

The following examples further illuminate the properties of the energy:

**Example 1.** For  $\beta > \frac{1}{2}$  there are measures with finite energy with densities which are not bounded. In particular let  $\rho : [0, 1] \rightarrow [0, \infty]$  be symmetric about  $x = \frac{1}{2}$  with

$$\rho(x) = cx^\gamma \quad \text{on} \quad \left(0, \frac{1}{2}\right]$$

where  $\gamma \in (-1, -\frac{1}{2\beta})$  and  $c$  is chosen so that  $\int_0^1 \rho dx = 1$ . Then

$$\bar{E}(\rho) \lesssim \int_0^{1/2} x^{-2-2\beta\gamma} dx < \infty$$

since  $\gamma < -\frac{1}{2\beta}$ .

**Example 2.** For  $\beta > \frac{1}{2}$  there are measures which are not in  $\mathcal{P}_{ac}(S^1)$  with finite energies.

Using the construction of the previous example there exists a function  $f \in C^\infty((0, \frac{1}{4}), (0, \infty))$  of the form  $f(x) = ax^\gamma + b$  such that  $\gamma \in (-1, -\frac{1}{2\beta})$ ,  $\int_0^{1/4} f(x) dx = \frac{1}{3}$  and  $f(\frac{1}{4}) = \frac{1}{6}$ . Let

$$\mu = \frac{1}{4}\delta_0 + \rho\mathcal{L}.$$

where  $\rho$  is symmetric about  $\frac{1}{2}$  and

$$\rho(x) = \begin{cases} f(x) & \text{if } x \in (0, \frac{1}{4}) \\ \frac{1}{6} & \text{if } x \in (\frac{1}{4}, \frac{1}{2}). \end{cases}$$

To show that  $E(\mu) < \infty$  it is enough to construct a sequence  $\mu_n \rightarrow \mu$  as  $n \rightarrow \infty$  in Wasserstein distance with  $\overline{E}(\mu_n) \leq 2 \int_0^{\frac{1}{4}} (f^{-\beta})_x^2 dx < \infty$  for all  $n \geq 2$ . Here is the required construction: For  $n \geq 2$  let  $\tilde{\rho}_n$  be symmetric about  $\frac{1}{2}$  with

$$\tilde{\rho}_n(x) = \begin{cases} \frac{n}{8} & \text{if } x \in (0, \frac{1}{n}) \\ \rho(x - \frac{1}{n} + f^{-1}(\frac{n}{8})) & \text{if } x \in (\frac{1}{n}, \frac{1}{2}). \end{cases}$$

Note that  $\int_0^{1/2} \tilde{\rho}_n dx < \frac{1}{2}$ . Thus there exists  $m < 1$  such that for  $\rho_n = \tilde{\rho}_n \vee m$  we have  $\int_0^{1/2} \rho_n dx = \frac{1}{2}$ . Furthermore note that  $\rho_n \in H^1(S^1)$ ,  $E(\rho_n) \leq 2 \int_0^{\frac{1}{4}} (f^{-\beta})_x^2 dx$ , and  $\rho_n$  converges to  $\rho$  in Wasserstein distance as  $n \rightarrow \infty$ .

**Example 3.** Corollary 10 implies that if  $\mu = \alpha\delta_a + \rho\mathcal{L}$ , where  $\rho$  is continuous on  $S^1 \setminus \{a\}$ , has finite energy,  $E(\mu) < \infty$ , then

$$\limsup_{x \rightarrow a^-} \rho(x) = \limsup_{x \rightarrow a^+} \rho(x) = \infty.$$

This in particular implies  $E(\delta_a) = \infty$  and  $E(\frac{1}{2}\delta_a + \frac{1}{2}\mathcal{L}) = \infty$ .

The main result of this paper is the following theorem:

**Theorem 1.** *Consider  $\beta \in [1, \frac{3}{2}]$ . The functional  $E$  is displacement convex on  $\mathcal{P}(S^1)$ .*

*Proof.* The proof relies on facts that we establish in a series of lemmas. In particular the proof uses that:

- i) If  $E(\mu) < \infty$  then there exists a sequence of probability measures,  $\{\mu_n\}_{n=1,2,\dots}$  with smooth densities such that  $\mu_n$  converges to  $\mu$  in Wasserstein distance and  $E(\mu_n)$  converges to  $E(\mu)$  as  $n \rightarrow \infty$ .
- ii) Convexity of  $E$  holds along geodesics connecting probability measures with smooth densities.

The Claim i) is included in Lemma 7 and it follows from Lemmas 4, 5, and 6. The Claim ii) is proved in Lemma 8 for  $\overline{E}$ . It follows for  $E$  since for smooth measures  $E(\mu) = \overline{E}(\mu)$ , by Lemma 6.

The geodesic convexity of  $E$  then follows by a standard approximation argument. In particular let  $\mu_0$  and  $\mu_1$  be probability measures with finite energy:  $E(\mu_0) < \infty$  and  $E(\mu_1) < \infty$ . Then there exist approximation sequences  $\{\mu_{0,n}\}_{n=1,2,\dots}$  and  $\{\mu_{1,n}\}_{n=1,2,\dots}$  of measures with smooth densities such that

$$d_W(\mu_{i,n}, \mu_i) \rightarrow 0 \quad \text{and} \quad \overline{E}(\mu_{i,n}) \rightarrow E(\mu_i) \quad \text{as } n \rightarrow \infty \quad \text{for } i = 0, 1.$$

Let  $\pi_n$  be the optimal transportation plan between  $\mu_{0,n}$  and  $\mu_{1,n}$ . Stability of optimal transportation plans (see [23, Theorem 5.20]) then implies that along a subsequence, which can we assume to be the whole sequence,

$$\pi_n \rightarrow \pi \quad \text{as } n \rightarrow \infty \quad \text{in the weak topology of measures}$$

where  $\pi$  is an optimal transportation plan between  $\mu_0$  and  $\mu_1$ . Let  $\Pi_1$  and  $\Pi_2$  be the projections  $S^1 \times S^1 \rightarrow S^1$  to the first and the second coordinate respectively. Then the displacement interpolant  $\mu_{s,n} = [\mu_{0,n}, \mu_{1,n}]_s$  is given by

$$\mu_{s,n} = ((1-s)\Pi_1 + s\Pi_2)_\# \pi_n.$$

Therefore, for  $\mu_s = [\mu_0, \mu_1]_s$ ,

$$\mu_{s,n} \rightarrow \mu_s \quad \text{as } n \rightarrow \infty \quad \text{in the weak topology of measures.}$$

Using the properties of  $E$  and the convexity of  $\overline{E}$

$$\begin{aligned} E(\mu_s) &\leq \liminf_{n \rightarrow \infty} \overline{E}(\mu_{s,n}) \\ &\leq \liminf_{n \rightarrow \infty} (1-s)\overline{E}(\mu_{0,n}) + s\overline{E}(\mu_{1,n}) \\ &= (1-s)E(\mu_0) + sE(\mu_1). \end{aligned}$$

□

The following lemma is a local estimate on the modulus of continuity.

**Lemma 2.** *Let  $\beta > 0$ . Consider a probability measure  $\mu$  with density  $\rho$  such that  $E_o = \overline{E}(\rho) < \infty$ . Then  $v = \rho^{-\beta}$  is a continuous function*

and if  $\rho(x_0) < \infty$  then

(8)

$$\rho(x) \leq \left( \frac{1}{\rho(x_0)^\beta} - \sqrt{E_o d(x, x_0)} \right)^{-\frac{1}{\beta}} \quad \text{if } d(x, x_0) < \frac{1}{E_o} \rho(x_0)^{-2\beta}$$

(9)

$$\rho(x) \geq \left( \frac{1}{\rho(x_0)^\beta} + \sqrt{E_o d(x, x_0)} \right)^{-\frac{1}{\beta}} \quad \text{for all } x \in S^1.$$

*Proof.* Since  $\int_{S^1} v_x^2 dx = E(\rho) < \infty$ ,  $v$  is a continuous function due to the Sobolev embeddings. Assume  $v(x_0) > 0$  and  $x \in [0, 1]$ . By reparameterizing the circle (translation and reflection) we can assume that  $0 < x_0 < x \leq x_0 + \frac{1}{2} < 1$ . Other cases are analogous. Using that line segment is the minimizer of Dirichlet energy for fixed boundary data one obtains:

$$E_o \geq \int_{x_0}^x v_x^2(s) ds \geq \left( \frac{v(x) - v(x_0)}{x - x_0} \right)^2 (x - x_0).$$

Therefore

$$\left( \frac{1}{\rho(x)^\beta} - \frac{1}{\rho(x_0)^\beta} \right)^2 \leq E_o (x - x_0).$$

Hence

$$\frac{1}{\rho(x_0)^\beta} - \sqrt{E_o (x - x_0)} \leq \frac{1}{\rho(x)^\beta} \leq \frac{1}{\rho(x_0)^\beta} + \sqrt{E_o (x - x_0)}$$

which implies the desired inequalities.  $\square$

As a corollary we obtain the following positive lower bound:

**Corollary 3.** *Assume  $\beta > 0$  and  $E_o = \overline{E}(\rho) < \infty$ . Then*

$$(10) \quad \rho \geq \left( \frac{1}{1 + \sqrt{E_o}} \right)^{\frac{1}{\beta}} \quad \text{on } S^1.$$

*Proof.* Since  $\int_{S^1} \rho dx = 1$  there exists  $x_0$  such that  $\rho(x_0) \geq 1$ . The desired inequality follows by applying the estimate (9) to arbitrary  $x \in S^1$  and using that  $d(x, x_0) < 1$ .  $\square$

Next two lemmas provide tools for approximating measures with finite energy  $\overline{E}$ , by smooth measures:

**Lemma 4.** Assume  $\beta > 0$ . Consider a probability measure  $\mu$  with density  $\rho$  such that  $\overline{E}(\rho) < \infty$ . For every  $\varepsilon > 0$  there exists a probability measure with density  $\tilde{\rho} \in H^1(S^1)$  such that

$$\overline{E}(\tilde{\rho}) \leq \overline{E}(\rho) \quad \text{and} \quad d_W^2(\rho, \tilde{\rho}) < \varepsilon.$$

*Proof.* Let  $\varepsilon > 0$ . Given  $m \geq 0$  let  $\sigma_-(m) = \int_0^1 (m - \rho(x))_+ dx$ . Note that  $\sigma_-$  is a nondecreasing Lipschitz continuous function with  $\sigma_-(m_2) - \sigma_-(m_1) \leq m_2 - m_1$  for  $m_2 > m_1$ . Consider also  $\sigma_+(M) = \int_0^1 (\rho(x) - M)_+ dx$ , whose properties mirror those of  $\sigma_-$ .

Note that  $\sigma_-(1) = \sigma_+(1)$ . Let  $\varepsilon_0 = \sigma_-(1)$ . If  $\varepsilon_0 = 0$  then  $\rho \equiv 1$ , so we can consider the case  $\varepsilon_0 > 0$ . We can assume  $\varepsilon < \varepsilon_0$ . Since  $\sigma_-(0) = 0$  and  $\sigma_-(1) > \varepsilon$  there exists  $m_\varepsilon < 1$  such that  $\sigma_-(m_\varepsilon) = \varepsilon$ . Analogously there exists  $M_\varepsilon > 1$  such that  $\sigma_+(M_\varepsilon) = \varepsilon$ . Now define

$$\tilde{\rho} = (\rho \vee m_\varepsilon) \wedge M_\varepsilon.$$

Construction ensures  $\int_0^1 \tilde{\rho} = 1$  and  $d_W^2(\rho, \tilde{\rho}) < \varepsilon$ .

Let  $\tilde{v}(x) = \frac{1}{\tilde{\rho}^\beta}$ . Since  $v \in H^1(S^1)$  and

$$\tilde{v} = (v \vee M_\varepsilon^{-\beta}) \wedge m_\varepsilon^{-\beta}$$

then  $\tilde{v} \in H^1(S^1)$ , as a composition of an  $H^1$  and a Lipschitz function. Note that  $\int_0^1 (\tilde{v}_x)^2 dx \leq \int_0^1 v_x^2 dx$  and that  $\tilde{v}$  is bounded by positive numbers from below and above. Since  $\tilde{\rho} = (\tilde{v})^{-1/\beta}$  and the function  $s \rightarrow s^{-1/\beta}$  is Lipschitz on the range of  $\tilde{v}$  the function  $\tilde{\rho}$  is in  $H^1(S^1)$  by the chain rule.  $\square$

**Lemma 5.** Assume  $\beta > 0$ . Consider a probability density  $\rho \in H^1(S^1)$  for which  $\overline{E}(\rho) < \infty$ . Then

$$\overline{E}(\rho) = \beta^2 \int_0^1 \rho^\theta \rho_x^2 dx \quad \text{where } \theta = -2(\beta + 1).$$

Moreover, there exists a sequence of probability measures with smooth densities  $\rho_n$  such that

$$d_W(\rho_n, \rho) \rightarrow 0 \quad \text{and} \quad \overline{E}(\rho_n) \rightarrow \overline{E}(\rho) \quad \text{as } n \rightarrow \infty.$$

*Proof.* Continuity of  $\rho$  and Corollary 3 imply that  $\rho$  is bounded by positive numbers from below and above:

$$0 < m \leq \rho \leq M < \infty$$

Thus  $v = \rho^{-\beta}$  is also bounded from below and above. Since  $\int_0^1 v_x^2 = \overline{E}(\rho) < \infty$ , we conclude that  $v \in H^1(S^1)$ . Given that  $v = \rho^{-\beta}$  and

that  $z \rightarrow z^{-\beta}$  is Lipschitz on range of  $\rho$  we can use the chain rule to obtain that  $v_x = -\beta \frac{\rho_x}{\rho^{\beta+1}}$ .

Let  $\eta$  be a smooth compactly supported mollifier. Let  $\eta_\delta = \frac{1}{\delta} \eta(\frac{\cdot}{\delta})$  and  $\rho_n = \eta_{\frac{1}{n}} * \rho$  where convolution is computed on  $S^1$ . Note that  $d_W(\rho_n, \rho) \rightarrow 0$  as  $n \rightarrow \infty$ , since support of  $\eta_{\frac{1}{n}}$  is shrinking to a point. Furthermore,

$$\begin{aligned} |\overline{E}(\rho_n) - \overline{E}(\rho)| &= \beta^2 \left| \int_0^1 \rho_n^\theta (\rho_n)_x^2 - \rho_n^\theta \rho_x^2 + \rho_n^\theta \rho_x^2 - \rho^\theta \rho_x^2 dx \right| \\ &\leq m^\theta \beta^2 \int_0^1 |(\rho_n)_x^2 - \rho_x^2| dx + m^{2\theta} \beta^2 \int_0^1 |\rho_n^{-\theta} - \rho^{-\theta}| \rho_x^2 dx, \end{aligned}$$

which converges to 0 since  $\rho_n$  converges to  $\rho$  in  $H^1$  and thus in  $L^\infty$ .  $\square$

We now prove the lower-semicontinuity of  $\overline{E}$  in the following sense:

**Lemma 6.** *Assume  $\beta > 0$ . Let  $\{\mu_n\}_{n=1,2,\dots}$  be a sequence of probability measures with densities in  $H^1(S^1)$ , converging in the weak topology to a probability measure  $\mu$ :*

$$d_W(\mu_n, \mu) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

*Assume  $\overline{E}(\mu) < \infty$  then  $\overline{E}(\mu) \leq \liminf_{n \rightarrow \infty} \overline{E}(\mu_n)$ .*

*Proof.* Consider  $\mu$  and  $\{\mu_n\}_{n=1,2,\dots}$  satisfying the assumptions of the lemma. Since  $\overline{E}(\mu)$  and  $\overline{E}(\mu_n)$  are finite  $\mu, \mu_n \in \mathcal{P}_{ac}(S^1)$  and thus we can consider their densities  $\rho, \rho_n$ . Let  $E_o = \liminf_{n \rightarrow \infty} \overline{E}(\mu_n)$ . We can assume without a loss of generality that  $E_o < \infty$  and that  $\lim_{n \rightarrow \infty} \overline{E}(\mu_n) = E_o$ . We can furthermore assume that  $\overline{E}(\mu_n)$  is bounded from above by  $E_o + 1$ . Let  $v_n = \rho_n^{-\beta}$ . Boundedness of  $\overline{E}(\rho_n)$  implies that  $\int_0^1 (v_n)_x^2 dx$  is bounded. Corollary 3 implies that  $v_n$  are uniformly bounded in  $L^\infty$  and consequently in  $H^1$ . Therefore there exists a subsequence, which we can assume to be the whole sequence, such that as  $n \rightarrow \infty$

$$v_n \rightharpoonup v \text{ in } H^1, \quad v_n \rightarrow v \text{ a.e.}, \quad \text{and } v_n \rightarrow v \text{ in } L^\infty.$$

We claim that  $|\{v = 0\}| = 0$ . To show this note that  $v_n \geq 0$  and  $\int_0^1 v_n^{-1/\beta} dx = 1$  imply

$$|\{v_n < \varepsilon^\beta\}| = |\{\varepsilon v_n^{-1/\beta} > 1\}| \leq \varepsilon \int_0^1 v_n^{-1/\beta} dx = \varepsilon$$

for all  $n$  and all  $\varepsilon > 0$ . Therefore  $|\{|v_n - v| \geq \varepsilon^\beta\}| \geq |\{v = 0\}| - \varepsilon$ . Since  $v_n \rightarrow v$  in  $L^\infty$ , taking  $n \rightarrow \infty$  and then  $\varepsilon \rightarrow 0$ , implies  $|\{v = 0\}| = 0$ . Therefore  $\rho_n$  converge almost everywhere to  $\tilde{\rho} = v^{-1/\beta}$ :

$$\rho_n \rightarrow \tilde{\rho} \text{ a.e. as } n \rightarrow \infty.$$

If  $\tilde{\rho} = \rho$  then the weak lower semicontinuity in  $H^1$  of the Dirichlet energy yields that

$$\overline{E}(\rho) = \int_0^1 v_x^2 dx \leq \lim_{n \rightarrow \infty} \int_0^1 (v_n)_x^2 dx = \lim_{n \rightarrow \infty} \overline{E}(\rho_n).$$

To show that  $\tilde{\rho} = \rho$  consider a point  $x_o$  such that  $\rho_n(x_o) \rightarrow \tilde{\rho}(x_o) < \infty$  as  $n \rightarrow \infty$ . It suffices to show that  $\tilde{\rho} = \rho$  in some neighborhood of  $x_o$ . We can assume that  $\rho_n(x_o) < \tilde{\rho}(x_o) + 1 = M$  for all  $n$ . By estimate (8) of Lemma 2, for  $\delta^{-1} = 2eM^{2\beta} > 0$  for all  $n$

$$\rho_n(x) \leq 2M \quad \text{on } (x_o - \delta, x_o + \delta).$$

Therefore  $\rho_n$  is bounded in  $L^\infty$  on the interval  $(x_o - \delta, x_o + \delta)$ . Thus, it has a subsequence,  $\{\rho_{n_k}\}_{k=1,2,\dots}$  converging weakly in  $L^1$ .

$$\rho_{n_k} \rightharpoonup \tilde{\rho} \text{ in } L^1((x_o - \delta, x_o + \delta)) \quad \text{as } k \rightarrow \infty.$$

Since weak  $L^1$  convergence implies convergence in the weak topology of measures, and since  $\rho_n$  converges to  $\rho$  in the weak topology of measures, we conclude that  $\tilde{\rho} = \rho$  on  $(x_o - \delta, x_o + \delta)$ .  $\square$

Let us remark that Lemmas 5 and 6 imply that  $E(\mu) = \overline{E}(\mu)$  whenever  $\overline{E}(\mu) < \infty$ .

**Lemma 7.** *assume  $\beta > 0$ . If  $E(\mu) < \infty$  then there exists a sequence of probability measures,  $\{\mu_n\}_{n=1,2,\dots}$  with smooth densities such that  $\mu_n$  converges to  $\mu$  in Wasserstein distance and  $\overline{E}(\mu_n)$  converges to  $E(\mu)$  as  $n \rightarrow \infty$ .*

*Proof.* We first assume that  $\overline{E}(\mu) < \infty$  and thus  $\mu = \rho \mathcal{L}$ . Due to Lemma 5 we can construct a sequence of densities  $\rho_n \in H^1(S^1)$  such that

$$\overline{E}(\rho_n) \leq \overline{E}(\rho), \text{ and } d_W(\rho, \rho_n) < \frac{1}{n}$$

for all  $n$ . Therefore, we have a sequence of measures  $\mu_n = \rho_n \mathcal{L}$  with  $H^1(S^1)$  densities converging to  $\mu$  in  $d_W$  sense, and thus Lemma 6 implies that

$$\limsup_{n \rightarrow \infty} \overline{E}(\mu_n) \leq \overline{E}(\mu) \leq \liminf_{n \rightarrow \infty} \overline{E}(\mu_n),$$

from which we deduce there exists a sequence of probability measures  $\mu_n$  with densities in  $H^1(S^1)$  satisfying the desired properties. Now, the approximation Lemma 5 allows one to replace densities in  $H^1(S^1)$  by smooth densities in the above statement.

Finally, let us assume that  $E(\mu) < \infty$ , then by the definition of the lower semicontinuous envelope of  $\bar{E}$ , there exists  $\{\mu_n\}_{n=1,2,\dots}$  with  $\bar{E}(\mu_n) < \infty$  such that  $E(\mu) = \lim_{n \rightarrow \infty} \bar{E}(\mu_n)$  and  $d_W(\mu, \mu_n) \rightarrow 0$ . Taking into account the first part of this proof, for any  $\varepsilon > 0$ , we can find a probability density with smooth density  $\tilde{\mu}_n = \tilde{\rho}_n \mathcal{L}$  for all  $n$  such that

$$|\bar{E}(\mu_n) - \bar{E}(\tilde{\mu}_n)| < \varepsilon \quad \text{and} \quad d_W(\mu_n, \tilde{\mu}_n) < \varepsilon.$$

Thus, the desired result follows.  $\square$

We end the sequence of lemmas needed to prove Theorem 1 by showing the displacement convexity of  $\bar{E}$  functional in the smooth setting.

**Lemma 8.** *Let  $\beta \in [1, \frac{3}{2}]$ . Consider probability measures  $\mu_0$  and  $\mu_1$  with smooth and positive densities  $\rho_0$  and  $\rho_1$ . Let  $\mu_s$  be the displacement interpolant defined by (4). Then the function*

$$s \rightarrow \bar{E}(\mu_s)$$

*is convex on  $[0, 1]$ .*

*Proof.* Let  $\mathcal{E}(s) = \bar{E}(\mu_s)$ . It is enough to show that  $\mathcal{E}''(s) \geq 0$  for all  $s$ . Since  $\mu_s$  has smooth density for all  $s \in [0, 1]$  and since interpolation between  $\mu_s$  and  $\mu_1$  is given by a linear reparameterization of  $\{\mu_r\}_{r \in [s, 1]}$ , it suffices to show that  $\mathcal{E}''(0) \geq 0$ . Since  $\rho_s$  is smooth

$$\mathcal{E}(s) = \int_0^1 \left[ \left( \frac{1}{\rho_s} \right)_y \right]^2 dy = \beta^2 \int_{as}^{as+1} \rho_s^\theta (\rho_s)_y^2 dy$$

where  $\theta = -2(\beta + 1)$ . Let  $\rho = \rho_0$ . Using the change of variables  $y = x + sf(x)$  we obtain

$$\mathcal{E}(s) = \beta^2 \int_0^1 \rho^\theta(x) \left( \rho_x - \frac{s\rho(x)f''(x)}{1 + sf'(x)} \right)^2 \frac{1}{(1 + sf'(x))^{3+\theta}} dx.$$

Differentiating twice gives

$$\mathcal{E}''(0) = \beta^2 \int_0^1 \rho^\theta [2A^2 + 4(4 + \theta)AB + (4 + \theta)(3 + \theta)B^2] dx$$

where

$$A = \rho f'' \text{ and } B = \rho' f'.$$

The quadratic form in  $A$  and  $B$  is nonnegative if

$$2(4 + \theta)(3 + \theta) \geq \frac{1}{4} 4^2 (4 + \theta)^2$$

that is if

$$(\theta + 4)(\theta + 5) \leq 0$$

which holds precisely when

$$\theta \in [-5, -4]$$

which is the case when

$$\beta \in \left[1, \frac{3}{2}\right].$$

□

Let us complement the previous result by several corollaries and remarks.

**3.1. Bounds from below for densities.** We can show the following generalizations of Corollary 3:

**Corollary 9.** *Assume  $E_o = E(\mu) < \infty$ . Let  $\rho_{ac}$  the the absolutely continuous part of the measure  $\mu$ . Then*

$$(11) \quad \rho_{ac} \geq \left( \frac{1}{1 + \sqrt{E_o}} \right)^{\frac{1}{\beta}} \quad \text{a.e. on } S^1.$$

*Proof.* Lemma 7 implies that there exists a sequence of absolutely continuous measures  $\mu_n$ , with continuous densities  $\rho_n$  such that:

$$d_W(\mu_n, \mu) \rightarrow 0 \quad \text{and} \quad \overline{E}(\rho_n) \rightarrow E(\mu) \quad \text{as } n \rightarrow \infty.$$

Using the Corollary 3, for every  $\varepsilon > 0$ ,

$$\rho_n \geq \left( \frac{1}{1 + \sqrt{E_o + \varepsilon}} \right)^{\frac{1}{\beta}} \quad \text{on } S^1$$

for all  $n$  large enough. Since  $\mu_n \rightarrow \mu$  in the weak topology of measures

$$\mu \geq \left( \frac{1}{1 + \sqrt{E_o + \varepsilon}} \right)^{\frac{1}{\beta}} \quad \text{on } S^1.$$

Taking  $\varepsilon \rightarrow 0$  and considering the absolutely continuous parts of the left and the right-hand side implies the conclusion of the corollary. □

**Corollary 10.** *Consider a probability measure  $\mu$  on  $S^1$  that has an atom at  $x = a$ . Let  $\rho_{ac}$  be the absolutely continuous part of the measure  $\mu$ . If  $E_o = E(\mu) < \infty$  then*

$$\rho_{ac}(x) \geq (E_o d(x, a))^{-\frac{1}{2\beta}}.$$

*Proof.* Definition of the energy (7) and Lemmas 4 and 5 imply that there exists a sequence of absolutely continuous measures  $\mu_n$ , with continuous densities  $\rho_n$  such that:

$$d_W(\mu_n, \mu) \rightarrow 0 \quad \text{and} \quad \overline{E}(\rho_n) \rightarrow E(\mu) \quad \text{as } n \rightarrow \infty.$$

Since  $\mu \geq \alpha \delta_a$  for some  $\alpha > 0$ , for every  $\varepsilon > 0$  there exists  $n_0$  such that for all  $n \geq n_0$  there exists  $x_n \in B(a, \varepsilon)$  such that  $\rho_n(x_n) > \frac{1}{\varepsilon}$ . Then Lemma 2 implies

$$\rho_n(x) \geq \left( \varepsilon^\beta + \sqrt{E_o(d(x, a) + \varepsilon)} \right)^{-\frac{1}{\beta}}.$$

Taking  $n \rightarrow \infty$  implies that

$$\mu \geq \left( \varepsilon^\beta + \sqrt{E_o(d(x, a) + \varepsilon)} \right)^{-\frac{1}{\beta}}.$$

The conclusion of the corollary now follows by considering the absolutely continuous parts of the measures and taking  $\varepsilon \rightarrow 0$ .  $\square$

**3.2. Lack of displacement convexity.** We first present an example that shows that for  $\beta \notin [1, \frac{3}{2}]$ ,  $\beta \neq 0$  the functional (6) is not displacement convex even if only geodesics between smooth densities are considered. For  $\beta = -1$  the functional corresponds to the Dirichlet energy. The lack of displacement convexity of the Dirichlet energy was known in the community, but we are not aware of a published reference.

**Example 1.** Assume  $\beta \notin [1, \frac{3}{2}]$  and  $\beta \neq 0$ . We use computations of Lemma 8 to build our example. In particular it suffices to find  $\rho$  and  $f$  sufficiently smooth and such that

$$\mathcal{E}''(0) = \beta^2 \int_0^1 \rho^\theta [2(\rho f'')^2 + 4(4+\theta)\rho \rho' f' f'' - (4+\theta)(3+\theta)(\rho' f')^2] dx < 0$$

Here  $\theta = -2(\beta + 1) \notin [-5, -4]$ . Let  $h = f'$ . then

(12)

$$\mathcal{E}''(0) = \beta^2 \int_0^1 \rho^\theta \left[ 2(\rho h' + (4+\theta)\rho' h)^2 - (4+\theta)(5+\theta)(\rho' h)^2 \right] dx.$$

Let us denote by  $\Phi(x)$  the integrand in the above integral. It suffices to find  $\rho$  and  $h$  which are periodic, Lipschitz and such that  $\int_{S^1} h = 0$  (so that  $f$  is periodic) for which  $\mathcal{E}''(0) < 0$ . The existence of a smooth example follows via mollification. We first define auxiliary functions  $\tilde{\rho}$  and  $\tilde{h}$  on the interval  $[0, 2M]$  for some  $M > 1$ . Let  $c = e^{-(4+\theta)}$ . We require  $\tilde{\rho}$  and  $\tilde{h}$  to be symmetric with respect to  $x = M$ , so we only need to define them on  $[0, M]$ . Let

$$(13) \quad \tilde{\rho}(x) = \begin{cases} e^x & \text{for } x \in [0, 1] \\ e & \text{for } x \in [1, M], \end{cases}$$

$$(14) \quad \tilde{h}(x) = \begin{cases} e^{-(4+\theta)x} & \text{for } x \in [0, 1] \\ -(x - l - 1)\frac{c}{l} & \text{for } x \in [1, 1 + 2l] \\ -c & \text{for } x \in [1 + 2l, 1 + 2l + L]. \end{cases}$$

where  $M = 1 + 2l + L$  and  $l$  and  $L$  are to be determined. The functions  $\tilde{\rho}$  and  $\tilde{h}$  on  $[0, M]$  are represented graphically on Figure 3.2.

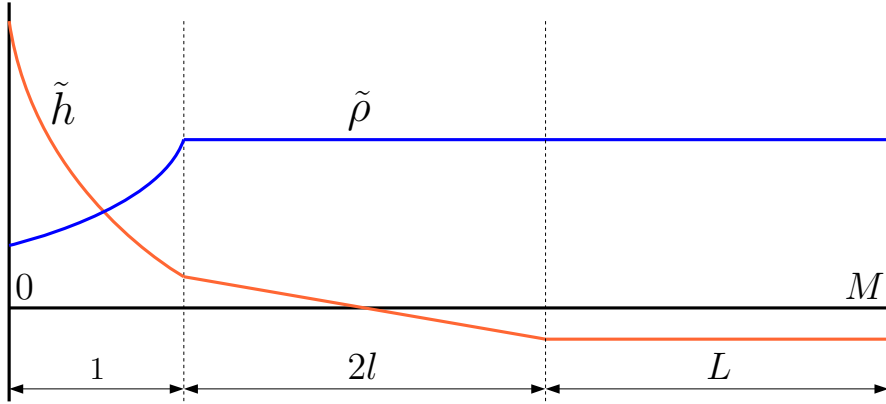


FIGURE 1. Schematic illustration of the example of lack of convexity (for  $\theta > -4$ )

First note that on  $[0, 1]$

$$\tilde{\rho}\tilde{h}' + (4 + \theta)\tilde{\rho}'\tilde{h} = 0$$

and thus, since  $(4 + \theta)(5 + \theta) > 0$ ,

$$I = \int_0^1 \tilde{\Phi}(x) dx < 0$$

where  $\tilde{\Phi}$  is the integrand in (12) with  $\rho$  and  $h$  replaced by  $\tilde{\rho}$  and  $\tilde{h}$ . Note that on  $[1, 1 + 2l]$

$$II = \int_1^{1+2l} \tilde{\Phi}(x) dx = \int_1^{1+2l} 2e^{2+\theta} \frac{c^2}{l^2} dx \lesssim \frac{1}{l}.$$

Choose  $l$  large enough to ensure  $I + II < 0$ . Choose  $L$  so that  $\int_0^M \tilde{h}(x) dx = 0$ . Note that

$$III = \int_{1+2l}^{1+2l+L} \tilde{\Phi}(x) dx = 0.$$

Therefore

$$\int_0^M \tilde{\Phi}(x) dx < 0.$$

Now define  $\rho$  and  $h$  on  $S^1$ :

$$\rho(x) = \lambda \tilde{\rho}(2Mx), \quad h = \tilde{h}(2Mx)$$

where  $\lambda > 0$  is chosen so that  $\int_0^1 \rho(x) = 1$ . Due to homogeneity of the terms in  $\Phi$  we have that  $\mathcal{E}''(0) < 0$ .

We now show that the energy (6) is not displacement on  $\mathbb{R}$ . We first show that if  $\beta \geq \frac{1}{2}$  then there are no probability densities  $\rho$  with finite energy. Namely Lemma 2 (when modified to apply on  $\mathbb{R}$ ) implies that  $\rho(x) > ax^{\frac{-1}{2\beta}} \geq ax^{-1}$  for some  $a > 0$  and all  $x$  sufficiently large. But such  $\rho$  cannot be integrable. If  $\beta = 0$  then  $\bar{E}$  is identically equal to 0. So it suffices to provide an example of lack of convexity for  $\beta < \frac{1}{2}$ ,  $\beta \neq 0$ .

**Example 2.** We follow the procedure and notation above. Let us choose  $\gamma \in (-\frac{1}{2\beta}, -1)$  if  $0 < \beta < \frac{1}{2}$  and  $\gamma < -1$  if  $\beta < 0$ . We define

$$(15) \quad \tilde{\rho}(x) = \begin{cases} e^x & \text{for } x \in [0, 1] \\ e & \text{for } x \in [0, l] \\ e(x-l)^\gamma & \text{for } x > l+1, \end{cases}$$

$$(16) \quad \tilde{h}(x) = \begin{cases} e^{-(4+\theta)x} & \text{for } x \in [0, 1] \\ -(x-l-1)\frac{c}{l} & \text{for } x \in [1, 1+l] \\ 0 & \text{for } x > l+1. \end{cases}$$

where  $l$  needs to be set. We extend  $\tilde{\rho}$  and  $\tilde{h}$  to  $\mathbb{R}$  to be symmetric about  $x = 1$ . We define  $\rho(x) = \lambda \tilde{\rho}(x)$  and  $h(x) = \tilde{h}(x)$ , with  $\lambda$  chosen so that  $\int_{\mathbb{R}} \rho = 1$ . The choice of  $\gamma$  ensures that  $\rho$  is integrable and

that  $\overline{E}(\rho)$  is finite. Verifying that for  $l$  large the construction gives  $\mathcal{E}''(0) < 0$  is as in Example 1.

#### 4. GRADIENT FLOW

As a consequence of the convexity along geodesics of the functional  $E$ , we can apply the general theory developed in [2] for gradient flows in the space of probability measures. We need to set up some notations to elaborate the consequences of this theory.

Given an initial measure  $\mu_0 \in \mathcal{P}(S^1)$  and time-step  $\tau > 0$ , let us consider the variational scheme or minimizing movement scheme [9, 1, 2] recursively defined by  $\mu_0^\tau = \mu_0$  and

$$(17) \quad \mu_{n+1}^\tau \in \arg \min_{\mu \in \mathcal{P}(S^1)} \left\{ E(\mu) + \frac{1}{2\tau} d_W^2(\mu_n^\tau, \mu) \right\},$$

for all  $n \in \mathbb{N}$ . The scheme is well defined since the functional is non-negative, lower semi-continuous with respect to  $d_W$  convergence, and its sub-level sets are relatively compact with respect to  $d_W$  since we are dealing with probability measures on a compact set.

Moreover, there is a well-defined notion of the subdifferential  $\partial E(\mu)$  of the functional  $E$  for any  $\mu \in \mathcal{P}(S^1)$  as developed in [2, Section 10.3]. Actually, due to geodesic convexity of  $E$ , the minimal slope of the functional is lower semicontinuous with respect to  $d_W$  and thus our functional satisfies the weak lower semicontinuity assumption in [2, Assumption 11.3.1]. Under these conditions, the existence of gradient flows is ensured by [2, Theorem 11.3.2]. More precisely, let us define the interpolating measure curve

$$\mu^\tau(t) := \mu_n^\tau \quad \text{for } t \in ((n-1)\tau, n\tau]$$

for all  $t > 0$  with  $\mu^\tau(0) := \mu_0$ .

**Theorem 11** (Existence of the Gradient Flow). *Given any  $\mu_0 \in \mathcal{P}(S^1)$  such that  $E(\mu_0) < \infty$ , then the sequence of measures  $\mu^\tau(t)$  converges to  $\mu_t$  locally uniformly in  $[0, \infty)$  with  $\mu \in AC_{loc}^2((0, \infty); \mathcal{P}(S^1))$  where  $\mu$  is a solution of the gradient flow equation*

$$v_t = -\partial^o(\mu_t), \quad \|v_t\|_{L^2(\mu_t; S^1)} = |\mu'| (t) \quad \text{a.e. } t > 0$$

with  $\mu_t \rightarrow \mu^0$  as  $t \rightarrow 0$ , where  $v_t$  is the tangent vector to the curve  $\mu_t$ . Moreover, the tangent vectors satisfy the energy identity

$$\int_a^b \int_{S^1} |v_t(x)|^2 d\mu_t(x) dt + E(a) = E(b)$$

for all  $0 \leq a \leq b < \infty$ .

We refer to the book [2] for all the notation and discussion of the ingredients in the theorem above. Let us just mention that  $\partial^\circ$  is the minimal selection in the subdifferential, that is, the vector in the subdifferential with minimal norm and  $|\mu'| (t)$  is the metric derivative of the curve  $\mu_t$ . Moreover, let us recall that  $L^2(\mu_t; S^1)$  is the set of  $L^2$  integrable functions with respect to  $\mu_t$  containing the set of tangent vectors, tangent bundle, to the curve of measures  $\mu_t$  given by

$$Tan_\mu \mathcal{P}(S^1) := \overline{\{\varphi_x : \varphi \in C^\infty(S^1)\}}^{L^2(\mu_t; S^1)}.$$

In fact, let us remember that due to [2, Theorem 8.3.1], we deduce from  $\|v_t\|_{L^2(\mu_t; S^1)} = |\mu'| (t)$  that

$$\frac{\partial \mu_t}{\partial t} + (v_t \mu_t)_x = 0$$

in distributional sense in  $[0, \infty) \times S^1$ . Finally, let us recall that the set  $AC_{loc}^2((0, \infty); \mathcal{P}(S^1))$  consists of absolutely continuous curves  $\mu_t$  in  $\mathcal{P}(S^1)$  such that  $|\mu'| \in L_{loc}^2(0, \infty)$ .

The geodesic convexity of the functional also plays an important role for the uniqueness of solutions of gradient flows. The following result follows readily from [2, Theorem 11.1.4].

**Theorem 12** ( $d_W$ -Contraction). *Given two gradient flow solutions  $\mu_t^1$  and  $\mu_t^2$  in the sense of the theorem above, then*

$$d_W(\mu_t^1, \mu_t^2) \leq d_W(\mu_0^1, \mu_0^2)$$

for all  $t \geq 0$ . In particular, we have a unique gradient flow solution for any given  $\mu_0 \in \mathcal{P}(S^1)$  with  $E(\mu_0) < \infty$ . Moreover, the gradient flow solution is characterized by a system of evolution variational inequalities:

$$\frac{d}{dt} d_W^2(\mu_t, \sigma) \leq E(\sigma) - E(\mu_t) \quad \text{a.e. } t > 0,$$

for all  $\sigma \in \mathcal{P}(S^1)$  such that  $E(\sigma) < \infty$ .

In order to apply this theory to the underlying partial differential equation, we have to characterize the minimal selection element in the subdifferential of the functional  $E$ . This point in our case is more involved than the known zero-order functionals since we do not have a good characterization of the lower semicontinuous envelope of  $E$  which is the case for the internal energy in [2, Subsection 10.4.3]. Let

us mention that in smooth setting the strong subdifferential characterization, and thus the minimal slope, is obtained in [2, Lemma 10.4.1]. More precisely, if we assume that the measure  $\mu$  has a positive smooth density  $\rho$ , i.e.,  $\mu = \rho\mathcal{L}$ , then

$$\partial^\circ(\mu) = \left( \frac{\delta E}{\delta \rho} \right)_x$$

with

$$\frac{\delta E}{\delta \rho} := \beta^2 (\theta \rho^{\theta-1} \rho_x^2 - 2(\rho^\theta \rho_x)_x) = -\beta^2 (\theta \rho^{\theta-1} \rho_x^2 - 2\rho^\theta \rho_{xx}).$$

where  $\theta = -2(\beta + 1)$ . As a consequence, assuming that the solutions of the gradient flow are smooth and positive for all times, then they correspond to solutions of the fourth-order partial differential equation

$$(18) \quad \frac{\partial \rho}{\partial t} = \beta^2 (\rho [-\theta \rho^{\theta-1} \rho_x^2 - 2\rho^\theta \rho_{xx}]_x)_x$$

It remains open to give the full characterization of the minimal selection in the subdifferential. However, we should mention that Theorem 12 is the first result of uniqueness of solutions, in this gradient flow sense, for a fourth-order diffusion equation that we are aware of.

Let us finally mention that, using the bound proved in Corollary 9, we conclude that the solutions of the gradient flow which are positive and have finite energy at  $t = 0$ , remain positive for all times, with a constant-in-time positive bound from below for their absolutely continuous with respect to Lebesgue parts. On the other hand, due to the existence of non absolutely continuous measures for which the functional  $E$  is finite, we cannot exclude the possibility that blowup in  $L^\infty$ -sense or delta mass formation happens starting from smooth data.

**Acknowledgements.** JAC acknowledges support from the project MTM2008-06349-C03-03 DGI-MCI (Spain). DS is grateful to Robert McCann and Felix Otto for stimulating discussions. DS acknowledges NSF support via the grant DMS-0638481. DS is also thankful to the Center for Nonlinear Analysis (NSF grants DMS-0405343 and DMS-0635983) for its support during the preparation of this paper. JAC acknowledges IPAM–UCLA where part of this work was done.

## REFERENCES

- [1] M. AGUEH, *Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory*, Adv. Differential Equations., 10, No 3 (2005) pp. 309-360.
- [2] L.A. AMBROSIO, N. GIGLI AND G. SAVARÉ, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics, Birkhäuser, 2005.
- [3] J.A. CARRILLO, S. LISINI, G.SAVARÉ, AND D. SLEPČEV, *Nonlinear mobility continuity equations and generalized displacement convexity*, preprint.
- [4] E.A. CARLEN, M.C. CARVALHO, R. ESPOSITO, J.L.LEBOWITZ, AND R. MARRA, *Displacement convexity and minimal fronts at phase boundaries*, preprint.
- [5] D. CORDERO-ERAUSQUIN, R.J. MCCANN, AND M. SCHMUCKENSCHLÄGER, *A Riemannian interpolation inequality à la Borell, Brascamp and Lieb*, Invent. Math., 146 No. 2 (2001), pp. 219–257.
- [6] J.A. CARRILLO, R.J. MCCANN AND C. VILLANI, *Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates*, Rev. Matemática Iberoamericana, 19 (2003), pp. 1-48.
- [7] J. A. CARRILLO, R. J. MCCANN AND C. VILLANI, *Contractions in the 2-Wasserstein length space and thermalization of granular media*, Arch. Rat. Mech. Anal., 179 (2006), pp. 217–263.
- [8] S. DANERI AND G. SAVARE, *Eulerian calculus for the displacement convexity in the Wasserstein metric*, preprint.
- [9] R. JORDAN, D. KINDERLEHRER, AND F. OTTO, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal., 29 (1998), pp. 1–17.
- [10] Y.-H. KIM AND R. J. MCCANN, *Continuity, curvature, and the general covariance of optimal transportation*, J. Eur. Math. Soc., to appear.
- [11] J. LOTT, *Optimal transport and Perelman's reduced volume*, preprint.
- [12] J. LOTT AND C. VILLANI, *Ricci curvature for metric-measure spaces via optimal transport* Ann. of Math, to appear.
- [13] R. J. MCCANN, *A convexity principle for interacting gases*, Adv. Math., 128, 1 (1997), pp. 153–179.
- [14] R. J. MCCANN, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal., 11 (2001) pp. 589–608.
- [15] R.J. MCCANN AND P. TOPPING, *Ricci flow, entropy, and optimal transportation*, preprint.
- [16] F. OTTO, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations, 26 (2001) pp. 101–174.
- [17] F. OTTO AND C. VILLANI, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal., 173, No. 2, (2000), pp. 361-400.
- [18] F. OTTO AND M. WESTDICKENBERG, *Eulerian Calculus for the contraction in the Wasserstein distance* SIAM J. Math. Anal. 37 (2005), no. 4, pp. 1227–1255.

- [19] K-T. STURM, *On the geometry of metric measure spaces. II*, Acta Math., 196, No. 1, (2006), pp. 133–177.
- [20] K-T. STURM, *On the geometry of metric measure spaces. I*, Acta Math., 196, No. 1, (2006), pp. 65–131.
- [21] M. VON RENESSE AND K-T. STURM, *Transport inequalities, gradient estimates, entropy, and Ricci curvature*, Comm. Pure Appl. Math., 58, No.7, (2005), pp. 923–940.
- [22] C. VILLANI, *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*, American Mathematical Society, Providence, RI, 2003.
- [23] C. VILLANI, *Optimal transport, old and new*, Lecture Notes for the 2005 Saint-Flour summer school, to appear, Springer 2008.

† ICREA (INSTITUCIÓ CATALANA DE RECERCA I ESTUDIS AVANÇATS) AND DEPARTAMENT DE MATEMÀTIQUES, UNIVERSITAT AUTÒNOMA DE BARCELONA, E-08193 BELLATERRA, SPAIN.

*E-mail address:* carrillo@mat.uab.es

‡ DEPARTMENT OF MATHEMATICAL SCIENCES CARNEGIE MELLON UNIVERSITY PITTSBURGH, USA.

*E-mail address:* slepcev@math.cmu.edu