

# Hamilton-Jacobi equations in the Wasserstein space

W. Gangbo\*, T. Nguyen† and A. Tudorascu‡

## Abstract

We introduce a concept of viscosity solutions for Hamilton-Jacobi equations (HJE) in the Wasserstein space. We prove existence of solutions for the Cauchy problem for certain Hamiltonians defined on the Wasserstein space over the real line. In order to illustrate the link between HJE in the Wasserstein space and Fluid Mechanics, in the last part of the paper we focus on a special Hamiltonian. The characteristics for these HJE are solutions of physical systems in finite dimensional spaces.

## 1 Introduction

In this paper we introduce a concept of viscosity solutions for Hamilton-Jacobi equations in the Wasserstein space  $\mathcal{P}_2(\mathbb{R}^d)$ . Throughout this work,  $\mathcal{P}_2(\mathbb{R}^d)$  denotes the set of Borel probability measures on  $\mathbb{R}^d$  with finite second moments. Let us temporarily set  $\mathcal{M} := \mathcal{P}_2(\mathbb{R}^d)$ . It is well-known that this is a metric space with a weak analogue of a differential and Riemannian structure (cf. e.g. [3], [11]). The tangent space at  $\mu \in \mathcal{M}$  is  $\mathcal{T}_\mu\mathcal{M}$ , the closure of  $\{\nabla\varphi : \varphi \in C_c^\infty(\mathbb{R}^d)\}$  in  $L^2(\mu)$ . Here,  $L^2(\mu)$  is the set of vector fields of  $\mathbb{R}^d$  onto itself, which are  $\mu$ -square integrable. The tangent bundle  $\mathcal{TM}$  is the union of the sets  $\{\mu\} \times \mathcal{T}_\mu\mathcal{M}$ .

We consider “mechanical Lagrangians”  $L : \mathcal{TM} \rightarrow \mathbb{R}$  of the form

$$L(\mu, \mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|_\mu^2 - \mathcal{V}(\mu), \quad (\mu, \mathbf{w}) \in \mathcal{TM}.$$

Here,  $\mathcal{V} : \mathcal{M} \rightarrow \mathbb{R}$ . The Hamiltonian associated to  $L$  is then

$$H(\mu, \zeta) := \sup_{\xi \in \mathcal{T}_\mu\mathcal{M}} \{ \langle \zeta, \xi \rangle_\mu - L(\mu, \xi) \},$$

which is defined for  $(\mu, \zeta) \in \mathcal{TM}$ . Since the tangent space  $\mathcal{T}_\mu\mathcal{M}$  is a Hilbert space, we have identified it with its dual. We use this convention in the rest of the paper.

We give sufficient conditions for the value function

$$U(\mu, t) := \min_{\sigma \in \mathcal{P}_t(\cdot, \mu)} \left\{ \int_0^t L(\sigma(\tau), \dot{\sigma}(\tau)) d\tau + U_0(\sigma(0)) \right\} \quad (1)$$

---

\*Georgia Institute of Technology, gangbo@math.gatech.edu; WG gratefully acknowledges the support provided by NSF grants DMS-03-54729 and DMS-06-00791

†University of Akron, tnguyen@uakron.edu; TN gratefully acknowledges the postdoctoral support provided by NSF grants DMS-03-54729 while in the School of Mathematics at Georgia Institute of Technology

‡Georgia Institute of Technology, adriant@math.gatech.edu; AT gratefully acknowledges the support provided by the School of Mathematics. Key words: Hamilton-Jacobi equations in infinite dimension, viscosity solutions, mass transfer, Wasserstein metric. AMS code: 49J40, 82C40 and 47J25.

to be a viscosity subsolution for the Hamilton-Jacobi equation

$$\frac{\partial U}{\partial t}(\mu, t) + H(\mu, \nabla_{\mu} U(\mu, t)) = 0, \quad U(\mu, 0) = U_0(\mu). \quad (2)$$

Here,  $\nabla_{\mu} U$  is the Wasserstein gradient of the function  $U(\cdot, t) : \mathcal{M} \rightarrow \mathbb{R}$ , as defined in subsection 3.1. We have denoted by  $P_t(\cdot, \mu)$  the set of paths  $\sigma : [0, t] \rightarrow \mathcal{M}$  which are 2-absolutely continuous and such that  $\sigma_t = \mu$ . We refer the reader to subsection 2.1 where we recall the definition of 2-absolutely continuous paths and their properties. Appropriate choices of  $\mathcal{V}, U_0 : \mathcal{M} \rightarrow \mathbb{R}$  reveal that the characteristics for (2) are paths in  $AC^2(0, T; \mathcal{P}_2(\mathbb{R}))$  satisfying specific systems of partial differential equations. These systems of equations appear in Fluid Mechanics and the one-dimensional Euler-Poisson system (see theorem 3.9 and also [9]) is included. Proving that the function in (1) is a viscosity supersolution seems to be a harder task which we are able to complete only when  $d = 1$ . To achieve that goal, we impose that  $m \rightarrow \mathcal{V}(m_{\#}\nu_0) =: \bar{\mathcal{V}}(m)$  is continuously differentiable. In the last section, motivated by the study of the Euler-Poisson system on the real line, we consider the case where  $\mathcal{V}(\mu)$  is the difference of one half of the second moment of  $\mu$  and one fourth of  $\int_{\mathbb{R} \times \mathbb{R}} |z - \bar{z}| d\mu(z) d\mu(\bar{z})$ . Despite the fact that  $\bar{\mathcal{V}}$  fails to be differentiable, we still show existence of viscosity solutions there.

When  $d = 1$ , we establish existence of a minimizing path in (1) by exploiting the fact that,  $\mathcal{M}$  can be identified with a subset of  $L^2(\nu_0)$ . Here,  $\nu_0$  is the Lebesgue measure restricted to the unit interval  $(-1/2, 1/2)$ . That identification is used to show that minimizers of (1) satisfy a stronger time regularity property. We use this to prove that the value function  $U$  is also a supersolution.

Before concluding this introduction let us point out that the theory of Hamilton-Jacobi equation in infinite dimensional spaces has been studied in flat spaces such as Hilbert spaces (cf, e.g. [4], [5]) and in Banach spaces (e.g. [6], [7]). The theory developed in these works cannot be applied directly to the metric space  $\mathcal{M}$ .

## 1.1 Notation and Definitions

- We suppose that  $T > 0$  is a constant. We sometimes give it a specific value, such as  $T = 1$ .
- $|\cdot|$  is the euclidean norm on  $\mathbb{R}^d$  and  $\langle \cdot, \cdot \rangle$  is the standard inner product.
- $C_c^\infty(\mathbb{R}^d)$  is the set of functions on  $\mathbb{R}^d$  which are infinitely differentiable and of compact support.
- If  $\psi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\pm\infty\}$ ,  $\psi^*$  is its Legendre transform.
- **id** denotes the identity map on  $\mathbb{R}^d$  for  $d \geq 1$ .
- As usual, we denote by  $\mathcal{L}^d$  the Lebesgue measure on  $\mathbb{R}^d$ .
- $X$  denotes the unit cube in  $\mathbb{R}^d$ , centered at the origin. In particular, if  $d = 1$ , then  $X = (-1/2, 1/2)$ . We set  $X_T := X \times (0, T)$ . Similarly,  $X_T^2 := X \times X \times (0, T)$ . The measure  $\nu_0$  is the restriction of  $\mathcal{L}^d$  to  $X$  and so it is a Borel probability measure. We write  $\nu_0 = \mathcal{L}^d|_X$ . The product measure of  $\nu_0$  by  $\mathcal{L}^1|_{(0, T)}$  is the measure on  $X_T$  denoted by  $\nu_T$ . We do not display explicitly its dependence on  $T$  since this does not create any confusion in this manuscript.
- $\mathcal{P}_2(\mathbb{R}^d)$  stands for the set of Borel probability measures  $\mu$  on  $\mathbb{R}^d$  with finite second moments:

$$\int_{\mathbb{R}^d} |y|^2 d\mu(y) < \infty.$$

- Given  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ ,  $\Gamma(\mu, \nu)$  is the set of Borel probability measures on  $\mathbb{R}^d \times \mathbb{R}^d$  which have

$\mu$  and  $\nu$  as their marginals. The Wasserstein distance  $W_2$  between  $\mu$  and  $\nu$  is defined by

$$W_2^2(\mu, \nu) = \min_{\gamma \in \Gamma(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma(x, y).$$

The set of  $\gamma$  where the minimum is achieved is nonempty and is denoted by  $\Gamma_o(\mu, \nu)$ . We refer the reader to [3] chapter 7 for the properties of  $W_2$  and  $\Gamma_o(\mu, \nu)$ .  $(\mathcal{P}_2(\mathbb{R}^d), W_2)$  is a metric space which is complete and separable.

- If  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ ,  $L^2(\mu)$  is the set of functions  $\xi : \mathbb{R}^d \rightarrow \mathbb{R}^d$  which are  $\mu$  measurable and such that  $\int_{\mathbb{R}^d} |\xi|^2 d\mu$  is finite. This is a separable Hilbert space for the inner product  $\langle \xi, \bar{\xi} \rangle_\mu = \int_{\mathbb{R}^d} \langle \xi, \bar{\xi} \rangle d\mu$ . We denote the associated norm by  $\|\cdot\|_\mu$ . When  $m = \mathcal{L}^1|_{(0,T)}$ , we write  $\|\cdot\|_{L^2(0,T)}$  for  $\|\cdot\|_m$  to distinguish between the space and time variables.

- If  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ , we denote by  $\mathcal{T}_\mu \mathcal{P}_2(\mathbb{R}^d)$  the closure of  $\{\nabla \varphi : \varphi \in C_c^\infty(\mathbb{R}^d)\}$  in  $L^2(\mu)$ . We refer to  $\mathcal{T}_\mu \mathcal{P}_2(\mathbb{R}^d)$  as the tangent space to  $\mathcal{P}_2(\mathbb{R}^d)$  at  $\mu$  (see section 8.5 of [3]). When  $d = 1$  it is easy to check that  $\mathcal{T}_\mu \mathcal{P}_2(\mathbb{R}) = L^2(\mu)$ .

- If  $(Z, |\cdot|)$  is a norm space,  $L^2(0, T; Z)$  is the set of Borel functions  $M : (0, T) \rightarrow Z$  such that  $\int_0^T |M_t|_Z^2 dt < \infty$ . Here and throughout this work, we write  $M_t$  in place of  $M(t)$ . When  $\mu$  is a Borel probability measure on  $\mathbb{R}^d$  and  $Z = L^2(\mu)$ , we identify  $L^2(0, T; L^2(\mu))$  with  $L^2(\mu \times \mathcal{L}^1|_{(0,T)})$ .

- If  $U : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\pm\infty\}$ , we denote by  $\partial U(\mu)$  the subdifferential of  $U$  at  $\mu$ , as introduced in [2] (see definition 3.1). As shown in [2], for  $\lambda$ -convex functionals, this definition coincides with the one in (10.3.12) of [3]. Since  $\|\cdot\|_\mu$  is uniform and  $\partial U(\mu)$  is a closed convex subset of  $L^2(\mu)$ , it admits a unique element of minimal norm. As it is customary in subdifferential analysis, we denote that element by  $\nabla_\mu U(\mu)$ . We refer to it as the gradient of  $U$  with respect to the Wasserstein distance  $W_2$ . The superdifferential of  $U$  at  $\mu$  is denoted by  $\partial^+ U(\mu)$  and consists of all  $\xi$  such that  $-\xi$  belongs to  $\partial(-U)(\mu)$ . We say that  $U$  is differentiable at  $\mu$  if  $\partial U(\mu)$  and  $\partial^+ U(\mu)$  are nonempty. In that case, both sets coincide and contain a unique element. Observe that there is no confusion denoting the element of minimal norm of  $\partial U(\mu)$  by  $\nabla_\mu U(\mu)$ .

- We also recall that if  $M : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a Borel map and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  then  $M_\# \mu$  is the Borel measure defined by

$$M_\# \mu[C] = \mu[M^{-1}(C)] \quad \text{for all Borel sets } C \subset \mathbb{R}^d.$$

- If  $\mu, \nu$  are Borel probability measures on the real line and  $\mu$  is atom-free, then it is known that there exists a unique (up to a set of  $\mu$ -zero measure) optimal map pushing forward  $\mu$  to  $\nu$ . It is called *the monotone rearrangement* and is obtained as  $G^{-1} \circ F$ , where  $F, G$  are the cumulative distribution functions of  $\mu$  and  $\nu$ . We have  $G(y) = \nu(-\infty, y]$  and  $G^{-1}(x) = \inf\{y \in \mathbb{R} : G(y) > x\}$ . Note that  $G^{-1}$  is the *generalized inverse* of  $G$  [12]. In this work, optimal map on the real line always means *left continuous* optimal map.

- We denote by  $\mathcal{M}on$  the set of monotone nondecreasing, left-continuous functions

$M : (-1/2, 1/2) \rightarrow \mathbb{R}$  which are in  $L^2(\nu_0)$ .

- Suppose  $(\mathcal{M}, dist)$  is a complete metric space and  $\sigma : (0, T) \rightarrow \mathcal{M}$ . We write  $\sigma_t$  for the value of  $\sigma$  at  $t$ :  $\sigma_t := \sigma(t)$ . If there exists  $\beta \in L^2(0, T)$  such that

$$dist(\sigma_t, \sigma_s) \leq \int_s^t \beta(\tau) d\tau \tag{3}$$

for every  $s < t$  in  $(0, T)$ , we say that  $\sigma$  is 2-absolutely continuous. We denote by  $AC^2(0, T; \mathcal{M})$  the set of  $\sigma : (0, T) \rightarrow \mathcal{M}$  that are 2-absolutely continuous.

- Suppose  $\sigma \in AC^2(0, T; \mathcal{M})$ . Since  $\mathcal{M}$  is complete,  $\lim_{t \rightarrow 0^+} \sigma_t$  exists and will be denoted by  $\sigma_0$ . Similarly,  $\sigma_T$  is well-defined. For  $\mathcal{L}^1$ -almost every  $t \in (0, T)$ ,

$$|\sigma'| (t) := \lim_{h \rightarrow 0} \frac{\text{dist}(\sigma_{t+h}, \sigma_t)}{|h|}$$

exists. If the above limit exists at  $t$ , we say that  $|\sigma'|$  exists at  $t$ . We have  $|\sigma'| \leq \beta$  for every  $\beta$  satisfying (3) and

$$\text{dist}(\sigma_t, \sigma_s) \leq \int_s^t |\sigma'|(\tau) d\tau \quad \text{for every } s < t \text{ in } (0, T). \quad (4)$$

The function  $|\sigma'|$  is referred to as the metric derivative of  $\sigma$ . For more details, we refer the reader to section 1.1 of [3]. We denote the  $L^2$ -norm of  $|\sigma'|$  on  $(0, T)$  by  $\|\sigma'\|_{\text{metric}, T}$ . In case there is no confusion about the time interval on which we integrate, we simply write  $\|\sigma'\|_{\text{metric}}$ .

- Suppose  $s, \bar{s} \in \mathcal{M}$ . We denote by  $P_T(s, \bar{s})$  the set of curves  $\sigma \in AC^2(0, T; \mathcal{M})$  such that  $\sigma_0 = s$  and  $\sigma_T = \bar{s}$ . Similarly,  $P_T(\cdot, \bar{s})$  denotes the set of curves  $\sigma \in AC^2(0, T; \mathcal{M})$  such that  $\sigma_T = \bar{s}$ .

- If  $n$  is a integer,  $\mathcal{P}_n$  is the set of  $n$  averages of  $n$  Dirac masses in  $\mathbb{R}$ . When  $d = 1$ , we divide  $X = (-1/2, 1/2)$  into  $n$  intervals of equal length. Recall that  $\nu_0$  is the restriction to  $X$  of the one-dimensional Lebesgue measure. Suppose  $N, \bar{N} \in L^2(\nu_0)$  are constant on each of these subintervals. We denote by  $P_T^n(N, \bar{N})$  the set of  $M$  in  $P_T(N, \bar{N})$  such that for each  $t \in (0, T)$ ,  $M_t$  is constant on each of these subintervals.

## 2 Preliminaries

### 2.1 Absolutely continuous curves on metric spaces

As in [9], throughout this subsection  $(\mathcal{M}, \text{dist})$  is a complete metric space. We assume the existence of a Hausdorff topology  $\tau$  on  $\mathcal{M}$ , weaker than the metric topology. Also, suppose there exists a distance  $\text{dist}_\tau$  such that on bounded subsets of  $(\mathcal{M}, \text{dist})$ , the topology  $\tau$  coincides with the distance topology  $\text{dist}_\tau$ . We assume that closed balls of  $(\mathcal{M}, \text{dist})$  are compact for  $\tau$  and that  $\text{dist}$  is  $\tau$ -sequentially lower semicontinuous on  $B \times B$  whenever  $B$  is a closed ball in  $(\mathcal{M}, \text{dist})$ . For instance, when  $(\mathcal{M}, \text{dist}) = (L^2(\nu_0), \|\cdot\|_{\nu_0})$ , we choose  $\tau$  to be the weak topology. When  $\mathcal{M} = \mathcal{P}_2(\mathbb{R}^d)$ ,  $\text{dist}$  is the Wasserstein distance, we choose  $\tau$  to be the narrow convergence topology (see [3] remark 5.1.1).

The following proposition appears in [9].

**Proposition 2.1.** *Suppose that  $\sigma$  belongs to  $AC^2(0, T; \mathcal{M})$  and  $s_0 \in \mathcal{M}$ . Then*

$$\pi \|\text{dist}(\sigma(\cdot), s_0)\|_{L^2(0, T)} \leq 2T \|\sigma'\|_{\text{metric}} + \pi \sqrt{T} \text{dist}(\sigma_T, s_0). \quad (5)$$

**Remark 2.2.** *Let  $\mathcal{M} := \mathcal{P}_2(\mathbb{R}^d)$  and let  $\text{dist}$  be the Wasserstein distance:  $\text{dist} := W_2$ . By theorem 8.3.1 in [3],  $\sigma \in AC^2(0, T; \mathcal{M})$  if and only if there exists a Borel map  $(y, t) \rightarrow \mathbf{v}_t(y) \in \mathbb{R}^d$  such that  $t \rightarrow \|\mathbf{v}_t\|_{\sigma_t} \in L^2(0, T)$  and the continuity equation*

$$\partial_t \sigma_t + \nabla_y \cdot (\sigma_t \mathbf{v}_t) = 0 \quad \text{in } \mathbb{R}^d \times (0, T)$$

holds in the sense of distributions. We call  $\mathbf{v}$  a velocity associated to the path  $\sigma$ . One can choose a velocity associated to the path  $\sigma$  so that it is uniquely determined by the following properties: if  $\mathbf{w}$  is another velocity associated to the path, then  $\|\mathbf{v}_t\|_{L^2(\sigma_t)} \leq \|\mathbf{w}_t\|_{L^2(\sigma_t)}$  and  $\mathbf{v}_t \in \mathcal{T}_{\sigma_t} \mathcal{P}_2(\mathbb{R}^d)$  for  $\mathcal{L}^1$ -almost every  $t \in (0, T)$ . We refer to  $\mathbf{v}_t$  as the tangent velocity field at  $\sigma_t$ , or the velocity of minimal norm associated to  $\sigma$ . For the velocity of minimal norm we have  $\|\mathbf{v}_t\|_{L^2(\sigma_t)} = |\sigma'(t)|$  for  $\mathcal{L}^1$ -almost every  $t \in (0, T)$ .

## 2.2 Value functions and the Lax–Oleinik semigroup

Throughout this subsection we assume that  $\mathcal{M}$ ,  $\text{dist}$  and  $\text{dist}_\tau$  are as in subsection 2.1. Suppose  $\mathcal{V} : \mathcal{M} \rightarrow \mathbb{R}$  is a Borel map and there exist  $\mathbf{s}_0 \in \mathcal{M}$  and  $e_0, e_1 \in \mathbb{R}$  such that

$$\mathcal{V}(\mathbf{s}) \leq e_0 \text{dist}^2(\mathbf{s}, \mathbf{s}_0) + e_1 \quad \text{for all } \mathbf{s} \in \mathcal{M}. \quad (6)$$

If  $\sigma \in AC^2(0, T; \mathcal{M})$ , then  $\sigma \in C([0, T]; \mathcal{M})$  by (4) and so  $t \rightarrow \mathcal{V}(\sigma_t)$  is a Borel function as a composition of two Borel functions. From (5) and (6) to conclude that for every  $\epsilon > 0$ , we have

$$\mathcal{A}_T(\sigma) \geq \frac{1}{2\pi^2} [\pi^2 - 8(1 + \epsilon)e_0 T^2] \int_0^T |\sigma'|^2(t) dt - T [e_0(1 + 1/\epsilon) \text{dist}^2(\sigma_T, \mathbf{s}_0) + e_1], \quad (7)$$

where  $\mathcal{A}_T(\sigma)$  is the action defined by

$$\mathcal{A}_T := \mathcal{A}_0^T \quad \text{and} \quad \mathcal{A}_s^t(\sigma) := \int_s^t \left( \frac{1}{2} |\sigma'|^2(\tau) - \mathcal{V}(\sigma_\tau) \right) d\tau.$$

For  $U_0 : \mathcal{M} \rightarrow \mathbb{R}$  we define the value function

$$U(\mathbf{s}, t) := \inf_{\sigma} \{ \mathcal{A}_t(\sigma) + U_0(\sigma(0)) : \sigma(t) = \mathbf{s}, \sigma \in AC^2(0, t; \mathcal{M}) \}. \quad (8)$$

**Remark 2.3.** Suppose  $U_0$  is bounded below. Let  $\bar{\sigma}(\tau) = \mathbf{s}$  for all  $\tau \in [0, t]$ , then we have

$$U(\mathbf{s}, t) \leq \mathcal{A}_t(\bar{\sigma}) + U_0(\bar{\sigma}(0)) = -t\mathcal{V}(\mathbf{s}) + U_0(\mathbf{s}) < \infty.$$

(i) If in addition  $8e_0 T^2 < \pi^2$ , then we may choose  $\epsilon > 0$  such that  $\pi^2 - 8(1 + \epsilon)e_0 T^2 > 0$ . Thus, we conclude from (7) that  $U(\mathbf{s}, t) > -\infty$  for  $t \in [0, T]$  and  $\mathbf{s} \in \mathcal{M}$ . Consequently,  $U : \mathcal{M} \times [0, T] \rightarrow \mathbb{R}$ .

(ii) If  $\mathcal{V}(\mathbf{s}) \leq \bar{e}_0 \text{dist}^p(\mathbf{s}, \mathbf{s}_0) + \bar{e}_1$  for some  $p \in (0, 2)$ , then for every  $e_0 > 0$  there exists  $e_1 > 0$  such that (6) holds. Choosing  $e_0$  arbitrarily small, we have  $8e_0 T^2 < \pi^2$  holds for  $T > 0$  large. Hence all conditions in (i) are satisfied, so in this case we also have  $U : \mathcal{M} \times [0, T] \rightarrow \mathbb{R}$ .

**Lemma 2.4.** Suppose  $U_0 : \mathcal{M} \rightarrow \mathbb{R}$  is bounded below,  $T > 0$  and  $8e_0 T^2 < \pi^2$ . Then the value function  $U$  defined in (8) satisfies

$$U(\mathbf{s}, t) = \inf_{\sigma} \{ \mathcal{A}_s^t(\sigma) + U(\sigma_s, s) : \sigma_t = \mathbf{s}, \sigma \in AC^2(s, t; \mathcal{M}) \} \quad (9)$$

for  $0 \leq s < t \leq T$ .

**Proof:** The method of proof of this lemma is standard. Fix  $\mathbf{s} \in \mathcal{M}$  and  $s, t$  such that  $0 \leq s < t \leq T$ . Denote by  $V(\mathbf{s}, t)$  the expression at the right handside of (9) and let  $\epsilon > 0$ . There exists  $\sigma^\epsilon \in AC^2(0, t, \mathcal{M})$  such that  $\sigma_t^\epsilon = \mathbf{s}$  and

$$U(\mathbf{s}, t) \geq -\epsilon + \mathcal{A}_0^t(\sigma^\epsilon) + U_0(\sigma^\epsilon(0)).$$

Consequently,  $\tau \rightarrow \frac{1}{2}|(\sigma^\epsilon)'|^2(\tau) - \mathcal{V}(\sigma_\tau^\epsilon) \in L^1(0, t)$  and hence we can write  $\mathcal{A}_0^t(\sigma^\epsilon) = \mathcal{A}_0^s(\sigma^\epsilon) + \mathcal{A}_s^t(\sigma^\epsilon)$ . Thus, we obtain from the above inequality that

$$U(\mathbf{s}, t) \geq -\epsilon + \mathcal{A}_0^s(\sigma^\epsilon) + \mathcal{A}_s^t(\sigma^\epsilon) + U_0(\sigma^\epsilon(0)) \geq -\epsilon + U(\sigma_s^\epsilon, s) + \mathcal{A}_s^t(\sigma^\epsilon) \geq -\epsilon + V(\mathbf{s}, t).$$

Since  $\epsilon > 0$  is arbitrary, we conclude that  $U(\mathbf{s}, t) \geq V(\mathbf{s}, t)$ . To obtain the reverse inequality, it suffices to show that for any  $\epsilon > 0$ ,

$$U(\mathbf{s}, t) \leq \epsilon + \mathcal{A}_s^t(\sigma) + U(\sigma_s, s)$$

for all  $\sigma \in AC^2(s, t; \mathcal{M})$  satisfying  $\sigma_t = \mathbf{s}$  and  $\mathcal{A}_s^t(\sigma) < +\infty$ . Let  $\sigma$  be such a path and choose  $\sigma^\epsilon \in AC^2(0, s; \mathcal{M})$  be such that  $\sigma_s^\epsilon = \sigma_s$  and

$$U(\sigma_s, s) \geq -\epsilon + \mathcal{A}_0^s(\sigma^\epsilon) + U_0(\sigma_0^\epsilon). \quad (10)$$

We extend  $\sigma^\epsilon$  to  $(s, t]$  by setting

$$\sigma^\epsilon(\tau) = \sigma(\tau) \quad \text{for } \tau \in (s, t].$$

The fact that  $\sigma^\epsilon$  is continuous at  $s$ , ensures it belongs to  $AC^2(0, t; \mathcal{M})$ . Because  $\sigma_t^\epsilon = \mathbf{s}$  we have

$$U(\mathbf{s}, t) \leq \mathcal{A}_0^t(\sigma^\epsilon) + U_0(\sigma_0^\epsilon) = \mathcal{A}_0^s(\sigma^\epsilon) + \mathcal{A}_s^t(\sigma) + U_0(\sigma_0^\epsilon),$$

where we have used the fact that  $\mathcal{A}_0^s(\sigma^\epsilon)$  and  $\mathcal{A}_s^t(\sigma)$  are finite to obtain the last equality. This together with (10) yields

$$U(\mathbf{s}, t) \leq \epsilon + U(\sigma_s, s) + \mathcal{A}_s^t(\sigma).$$

QED.

### 3 Hamilton-Jacobi equations on Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$

In this subsection,  $d \geq 1$  is an integer and we apply the results of the previous section to the space  $\mathcal{M} := \mathcal{P}_2(\mathbb{R}^d)$  endowed with the Wasserstein metric  $dist = W_2$ . We assume that  $\mathcal{V} : \mathcal{M} \rightarrow \mathbb{R}$  is Borel and is bounded below on bounded subsets of  $\mathcal{M}$ . We introduce the Lagrangian

$$L(\mu, \xi) = \frac{1}{2} \|\xi\|_\mu^2 - \mathcal{V}(\mu), \quad (\mu, \xi) \in \mathcal{T}\mathcal{M}.$$

The Hamiltonian associated to  $L$  is then

$$H(\mu, \zeta) := \sup \{ \langle \zeta, \xi \rangle_\mu - L(\mu, \xi) : \xi \in \mathcal{T}_\mu \mathcal{M} \} = \frac{1}{2} \|\zeta\|_\mu^2 + \mathcal{V}(\mu), \quad (\mu, \zeta) \in \mathcal{T}\mathcal{M}.$$

If  $\sigma \in AC^2(0, T; \mathcal{M})$  we denote by  $\mathbf{v}$  the velocity of minimal norm associated to  $\sigma$ . We introduce the action

$$\mathcal{A}(\sigma) := \int_0^T L(\sigma_t, \mathbf{v}_t) dt$$

defined on  $AC^2(0, T; \mathcal{M})$ . Recall that if  $\mu, \nu \in \mathcal{M}$ , then  $P_T(\mu, \nu)$  denotes the set of paths  $\sigma$  in  $AC^2(0, T; \mathcal{M})$  such that  $\sigma_0 = \mu$  and  $\sigma_T = \nu$ .

### 3.1 Definition of viscosity solution on Wasserstein space

The notion of viscosity sub and super solutions can be defined in terms of the subdifferentials and superdifferentials of functions on  $\mathcal{M}$ . The definitions of sub (super) differential we are going to give here coincide with that of [3] for  $\lambda$ -convex function. Otherwise, in general, they differ.

**Definition 3.1.** Fix  $\mu \in \mathcal{M}$  and  $\xi \in \mathcal{T}_\mu \mathcal{M}$ . Let  $\mathcal{W} : \mathcal{M} \rightarrow \mathbb{R} \cup \{\pm\infty\}$  be a proper functional.

(i) We say that  $\xi$  belongs to the subdifferential of  $\mathcal{W}$  at  $\mu$  and we write  $\xi \in \partial \mathcal{W}(\mu)$  if

$$\mathcal{W}(\nu) - \mathcal{W}(\mu) \geq \sup_{\gamma \in \Gamma_o(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \xi(\bar{y})(y - \bar{y}) d\gamma(\bar{y}, y) + o(W_2(\mu, \nu)) \quad \forall \nu \in \mathcal{M}.$$

Here  $\Gamma_o(\mu, \nu)$  is the set of optimal plans between  $\mu$  and  $\nu$ .

(ii) We say that  $\xi$  belongs to the superdifferential of  $\mathcal{W}$  at  $\mu$  and we write  $\xi \in \partial^+ \mathcal{W}(\mu)$  if  $-\xi \in \partial \mathcal{W}(\mu)$ .

**Remark 3.2.** As expected, when the sets  $\partial \mathcal{W}(\mu)$  and  $\partial^+ \mathcal{W}(\mu)$  are both nonempty, then they coincide and consist of a single element. Indeed, suppose  $\xi_1 \in \partial \mathcal{W}(\mu)$  and  $\xi_2 \in \partial^+ \mathcal{W}(\mu)$ . Let  $\varphi \in C_c^2(\mathbb{R})$  which we assume to be distinct from the null function to avoid trivialities. The map  $M^\varepsilon := \mathbf{id} + \varepsilon \nabla \varphi$  is the gradient of a convex function for  $|\varepsilon| \ll 1$ . Thus, setting  $\mu_\varepsilon := M^\varepsilon \# \mu$ , we get  $\gamma^\varepsilon := (\mathbf{id} \times M^\varepsilon) \# \mu \in \Gamma_o(\mu, \mu_\varepsilon)$ . We have

$$0 \geq \lim_{\varepsilon \rightarrow 0} \frac{\int_{\mathbb{R}^d \times \mathbb{R}^d} (\xi_1(y) - \xi_2(y))(\bar{y} - y) d\gamma^\varepsilon(y, \bar{y})}{W_2(\mu, \mu^\varepsilon)} = \frac{1}{\|\nabla \varphi\|_\mu} \int_{\mathbb{R}^d} \langle \xi_1 - \xi_2, \nabla \varphi \rangle d\mu.$$

Since  $\varphi$  is arbitrary and  $\xi_1, \xi_2 \in \mathcal{T}_\mu \mathcal{M}$  we obtain that  $\xi_1 \equiv \xi_2$   $\mu$ -almost everywhere.

**Definition 3.3.** In case the convex set  $\partial \mathcal{W}(\mu)$  is nonempty, as it is standard in convex analysis, its unique element of minimal  $\|\cdot\|_\mu$ -norm is denoted by  $\nabla_\mu \mathcal{W}(\mu)$ . It is called the gradient of  $\mathcal{W}$  with respect to the Wasserstein distance at  $\mu$ . In particular, if  $\partial \mathcal{W}(\mu_0) \cap \partial^+ \mathcal{W}(\mu) \neq \emptyset$  then  $\nabla_\mu \mathcal{W}(\mu)$  is the unique element of the intersection.

**Remark 3.4.** Suppose that  $\partial^+ \mathcal{W}(\mu)$  is nonempty and denote its unique element of minimal  $\|\cdot\|_\mu$ -norm by  $\nabla_\mu^+ \mathcal{W}(\mu)$ . If  $\partial \mathcal{W}(\mu)$  is also nonempty then  $\nabla_\mu \mathcal{W}(\mu) = \nabla_\mu^+ \mathcal{W}(\mu)$ . Thus, there is no confusion denoting both elements of minimal norm by  $\nabla_\mu \mathcal{W}(\mu)$ .

In the sequel, if  $T > 0$  we set

$$Q_T := \mathcal{M} \times (0, T).$$

By analogy with the standard finite-dimensional theory, we have the following definitions:

**Definition 3.5.** Let  $U : Q_T \rightarrow \mathbb{R}$ .

(i) We say that  $U$  is a viscosity subsolution for (2) if  $U$  is upper semicontinuous on  $Q_T$ ,

$$U(\cdot, 0) \leq U_0, \text{ and } \theta + H(\mu, \zeta) \leq 0 \text{ for all } (\mu, t) \in Q_T \text{ and all } (\zeta, \theta) \in \partial U(\mu, t). \quad (11)$$

(ii) We say that  $U$  is a viscosity supersolution for (2) if  $U$  is lower semicontinuous on  $Q_T$ ,

$$U(\cdot, 0) \geq U_0, \text{ and } \theta + H(\mu, \zeta) \geq 0 \text{ for all } (\mu, t) \in Q_T \text{ and all } (\zeta, \theta) \in \partial U(\mu, t). \quad (12)$$

**Remark 3.6.** If  $U$  is a viscosity solution, then, in view of remark 3.2, we deduce that (2) is satisfied at all points  $(\mu, t) \in Q_T$  where  $\partial U(\mu, t) \cap \partial^+ U(\mu, t) \neq \emptyset$ , which are precisely the points where  $U$  is differentiable.

### 3.2 Existence of viscosity subsolutions in $\mathcal{P}_2(\mathbb{R}^d)$

We suppose there exist  $e_0, e_1 \in \mathbb{R}$  such that

$$\mathcal{V}(\mu) \leq e_0 \int_{\mathbb{R}^d} |y|^2 d\mu(y) + e_1 \quad \forall \mu \in \mathcal{M}. \quad (13)$$

Given a path  $\sigma \in AC^2(0, T; \mathcal{M})$ , we always denote by  $\mathbf{v}$  its velocity of minimal norm. With this convention, by remark 2.2, the value function  $U$  defined in (8) is

$$U(\mu, t) := \inf_{\sigma} \left\{ \int_0^t L(\sigma_\tau, \mathbf{v}_\tau) d\tau + U_0(\sigma(0)) : \sigma(t) = \mu, \sigma \in AC^2(0, t; \mathcal{M}) \right\} \quad \text{for } \mu \in Q_T. \quad (14)$$

For  $\mu, \nu \in \mathcal{M}$  we define

$$\mathcal{C}_t(\nu, \mu) = \inf_{\sigma} \left\{ \int_0^t L(\sigma_\tau, \mathbf{v}_\tau) d\tau : \sigma \in P_t(\nu, \mu) \right\}.$$

Note that

$$U(\mu, t) = \inf_{\nu \in \mathcal{M}} \{ \mathcal{C}_t(\nu, \mu) + U_0(\nu) \}.$$

**Lemma 3.7.** *Suppose that  $\mathcal{V} : \mathcal{M} \rightarrow \mathbb{R}$  is Borel and bounded below on bounded subsets of  $\mathcal{M}$ . Then  $U$  is upper semicontinuous on  $Q_T$  and  $(\nu, \mu, t) \rightarrow \mathcal{C}_t(\nu, \mu)$  is upper semicontinuous on  $\mathcal{M} \times \mathcal{M} \times (0, T)$ .*

**Proof:** Let  $(\mu_n, t_n) \rightarrow (\mu, t)$  and suppose  $\varepsilon > 0$ . Then there exists  $\sigma \in P_t(\cdot, \mu)$  such that

$$\int_0^t L(\sigma_\tau, \mathbf{v}_\tau) d\tau + U_0(\sigma(0)) \leq U(\mu, t) + \varepsilon. \quad (15)$$

Let  $\bar{\sigma}^n \in P_1(\mu, \mu_n)$  be a geodesic connecting  $\mu$  and  $\mu_n$ . Define  $\sigma^n(\tau) = \sigma(nt\tau/(n-1)t_n)$  for  $\tau \in [0, (n-1)t_n/n]$  and  $\sigma^n(\tau) = \bar{\sigma}^n((n\tau/t_n) + (1-n))$  for  $\tau \in [(n-1)t_n/n, t_n]$ . Then  $\sigma^n \in P_{t_n}(\cdot, \mu_n)$ . By changing variables and using the properties of geodesics, we have

$$\begin{aligned} \int_0^{t_n} L(\sigma_\tau^n, \mathbf{v}_\tau^n) d\tau &= \int_0^{\frac{(n-1)t_n}{n}} L(\sigma_\tau^n, \mathbf{v}_\tau^n) d\tau + \int_{\frac{(n-1)t_n}{n}}^{t_n} L(\sigma_\tau^n, \mathbf{v}_\tau^n) d\tau \\ &= \frac{(n-1)t_n}{nt} \int_0^t L(\sigma_\tau, \mathbf{v}_\tau) d\tau + \frac{t_n}{n} \int_0^1 L(\bar{\sigma}_\tau^n, \bar{\mathbf{v}}_\tau^n) d\tau \\ &= \frac{(n-1)t_n}{nt} \int_0^t L(\sigma_\tau, \mathbf{v}_\tau) d\tau + \frac{t_n}{2n} W_2^2(\mu, \mu_n) - \frac{t_n}{n} \int_0^1 \mathcal{V}(\bar{\sigma}_\tau^n) d\tau. \end{aligned}$$

Moreover, the set  $\{\bar{\sigma}^n(\tau) : \tau \in [0, 1]\}$  is bounded uniformly in  $n$  as  $W_2(\bar{\sigma}_\tau^n, \mu) \leq W_2(\mu, \mu_n)$ . Therefore, as  $\mathcal{V}$  is bounded below on bounded subsets, we obtain

$$\limsup_{n \rightarrow \infty} \int_0^{t_n} L(\sigma_\tau^n, \mathbf{v}_\tau^n) d\tau \leq \int_0^t L(\sigma_\tau, \mathbf{v}_\tau) d\tau.$$

This together with the facts that  $\sigma^n \in P_{t_n}(\cdot, \mu_n)$  and  $\bar{\sigma}^n(0) = \sigma(0)$  yields

$$\limsup_{n \rightarrow \infty} U(\mu_n, t_n) \leq \limsup_{n \rightarrow \infty} \int_0^{t_n} L(\sigma_\tau^n, \mathbf{v}_\tau^n) d\tau + U_0(\sigma^n(0)) \leq \int_0^t L(\sigma_\tau, \mathbf{v}_\tau) d\tau + U_0(\sigma(0)).$$

It then follows from (15) that  $\limsup_{n \rightarrow \infty} U(\mu_n, t_n) \leq U(\mu, t) + \varepsilon$  yielding that  $U$  is upper semicontinuous on  $Q_T$ . The proof of the upper semicontinuity of  $\mathcal{C}$  on  $\mathcal{M} \times \mathcal{M} \times (0, T)$  is similar. QED.

We soon give a sufficient condition which ensures existence of viscosity subsolutions for the infinite-dimensional Hamilton-Jacobi equation

$$\frac{\partial U}{\partial t}(\mu, t) + \frac{1}{2} \int_{\mathbb{R}} |\nabla_{\mu} U(\mu, t)(x)|^2 d\mu(x) + \mathcal{V}(\mu) = 0, \quad U(\mu, 0) = U_0(\mu) \quad (16)$$

on  $Q_T$ . When appropriate conditions are imposed on  $\mathcal{V}$ , then a minimizer  $\sigma$  exists in (14). It satisfies the system of partial differential equations

$$\begin{cases} \partial_t \sigma_t + \operatorname{div}(\sigma_t \mathbf{v}_t) = 0 & \text{in } \mathbb{R}^d \times (0, T) \\ \partial_t(\sigma_t \mathbf{v}_t) + \operatorname{div}(\sigma_t \mathbf{v}_t \otimes \mathbf{v}_t) = -\sigma_t \nabla_{\mu} \mathcal{V}(\sigma_t) & \text{in } \mathbb{R}^d \times (0, T) \\ \sigma_T = \mu, \quad \mathbf{v}_0 = \nabla_{\mu} U_0(\sigma_0). \end{cases} \quad (17)$$

This result, which can be found in theorem 3.9, gives the clear link between infinite-dimensional Hamilton-Jacobi equations on  $\mathcal{M}$  and Fluids Mechanics. The system of equations (17) (i) has to be understood in the following sense:

$$\int_0^T dt \int_{\mathbb{R}^d} (\partial_t \phi_t + \langle \nabla \phi_t, \mathbf{v}_t \rangle) d\sigma_t + \int_{\mathbb{R}^d} \phi_0 d\sigma_0 = 0, \quad (18)$$

whereas (ii) amounts to

$$\int_0^T dt \int_{\mathbb{R}^d} \langle \partial_t \Phi_t + \nabla_x \Phi_t \mathbf{v}_t, \mathbf{v}_t \rangle d\sigma_t + \int_{\mathbb{R}^d} \langle \Phi_0, \nabla_{\mu} U_0(\sigma_0) \rangle d\sigma_0 = \int_0^T dt \int_{\mathbb{R}^d} \langle \Phi_t, \nabla_{\mu} \mathcal{V}(\sigma_t) \rangle d\sigma_t \quad (19)$$

for all  $\Phi = \nabla \phi$  where  $\phi \in C_c^1([0, T] \times \mathbb{R}^d)$ . Here  $\phi \in C_c^1([0, T] \times \mathbb{R}^d)$  is the set of  $C^1$  functions on  $[0, T] \times \mathbb{R}^d$  which are the restriction of compactly supported  $C^1$  functions on  $(-1, T) \times \mathbb{R}^d$ .

**Definition 3.8.** *Suppose  $\sigma \in AC^2(0, T; \mathcal{M})$  and  $\mathbf{v}$  is its velocity of minimal norm. We say that  $(\sigma, \mathbf{v})$  satisfies (17) if it satisfies (18) and (19).*

In the second part of the next theorem, we assume that  $\mathcal{V}$  satisfies the following condition:  $\partial \mathcal{V}(\mu)$  is nonempty for every  $\mu \in \mathcal{M}$ , and for any  $C > 0$  and any compact set  $K \subset \mathcal{M}$ , there exists a constant  $C^* > 0$  such that

$$\sup_{\mu, \Psi} \left\{ \frac{\mathcal{V}((\mathbf{id} + \varepsilon \Psi) \# \mu) - \mathcal{V}(\mu)}{\varepsilon} - \langle \nabla_{\mu} \mathcal{V}(\mu), \Psi \rangle_{\mu} : \mu \in K, \Psi \in C_c^{\infty}(\mathbb{R}^d), \|\Psi\|_{\infty} \leq C \right\} \geq -C^* \quad (20)$$

for all  $\varepsilon > 0$  sufficiently small.

**Theorem 3.9.** *Suppose that  $\mathcal{V}$  satisfies (13),  $8\varepsilon_0 T^2 < \pi^2$  and  $\mathcal{V}$  is lower semicontinuous with respect to  $W_2$ . Also, suppose that  $U_0 : \mathcal{M} \rightarrow \mathbb{R}$  is bounded below and  $\mathcal{V}$  is bounded below on bounded subsets of  $\mathcal{M}$ . Then  $U$  from (14) satisfies:*

- (i)  $U$  is a viscosity subsolution of (16) over  $Q_T$ .
- (ii) Assume in addition that  $\mu \rightarrow \nabla_{\mu} \mathcal{V}(\mu)$  is a Borel map and  $\partial U_0(\nu)$  is nonempty for all  $\nu \in \mathcal{M}$  and (20) holds. Then if  $\sigma$  is a minimizer of (14) for  $t = T$ , we have  $(\sigma, \mathbf{v})$  satisfies the system of equations (17).

**Proof:** Clearly, (14) implies  $U(\mu, 0) = U_0(\mu)$  for all  $\mu \in \mathcal{M}$ , so the first inequalities in (11) and (12) hold trivially. By lemma 3.7,  $U$  is upper semicontinuous.

**Step 1:** We prove (i). Let  $(\mu, t) \in Q_T$  and assume that  $(\zeta, \theta) \in \partial U(\mu, t)$ . Take an arbitrary  $\varphi \in C_c^\infty(\mathbb{R})$  and  $\varepsilon > 0$  small enough such that  $x \rightarrow |x|^2/2 + \lambda\varphi(x)$  is a strictly convex function for all  $\lambda \in [0, \varepsilon]$ . Let  $\phi_\tau^\varepsilon(x) := |x|^2/2 + (t - \tau)\varphi(x)$ , which is strictly convex for  $\tau \in (t - \varepsilon, t)$ . Set

$$M^\varepsilon := \mathbf{id} + \varepsilon \nabla \varphi, \quad \mu^\varepsilon := M^\varepsilon \# \mu, \quad \sigma_\tau^\varepsilon := (\nabla \phi_\tau^\varepsilon) \# \mu \quad \tau \in [t - \varepsilon, t].$$

Note that  $\sigma^\varepsilon$  is the geodesic connecting  $\mu^\varepsilon$  at time  $t - \varepsilon$  to  $\mu$  at time  $t$ . Also,

$$W_2(\mu, \mu^\varepsilon) = \varepsilon \|\nabla \varphi\|_\mu, \quad \gamma^\varepsilon := (\mathbf{id} \times M^\varepsilon) \# \mu \in \Gamma_o(\mu, \mu^\varepsilon). \quad (21)$$

We extend  $\sigma^\varepsilon$  to  $[0, t - \varepsilon)$  by choosing  $\sigma^\varepsilon \in P_{t-\varepsilon}(\cdot, \mu^\varepsilon)$  so that

$$U(\mu^\varepsilon, t - \varepsilon) \geq -\varepsilon^2 + \mathcal{A}_0^{t-\varepsilon}(\sigma^\varepsilon) + U_0(\sigma^\varepsilon(0)).$$

We have  $\sigma^\varepsilon \in P_t(\cdot, \mu)$  and so,

$$\begin{aligned} U(\mu, t) &\leq \mathcal{A}_0^t(\sigma^\varepsilon) + U_0(\sigma^\varepsilon(0)) = \mathcal{A}_0^{t-\varepsilon}(\sigma^\varepsilon) + \mathcal{A}_{t-\varepsilon}^t(\sigma^\varepsilon) + U_0(\sigma^\varepsilon(0)) \\ &\leq U(\mu^\varepsilon, t - \varepsilon) + \varepsilon^2 + \mathcal{A}_{t-\varepsilon}^t(\sigma^\varepsilon). \end{aligned} \quad (22)$$

The velocity of minimal norm associated to  $\sigma_\varepsilon$  on  $(t - \varepsilon, t)$  is  $\mathbf{v}^\varepsilon$  defined by

$$\mathbf{v}_\tau^\varepsilon = \frac{\nabla(\phi_\tau^\varepsilon)^* - \mathbf{id}}{t - \tau}.$$

For  $\tau \in (t - \varepsilon, t)$ , we have

$$\mathbf{v}_\tau^\varepsilon \circ \nabla \phi_\tau^\varepsilon = -\nabla \varphi, \quad W_2(\sigma_\tau^\varepsilon, \mu) = \|\nabla \phi_\tau^\varepsilon - \mathbf{id}\|_\mu = |t - \tau| \|\nabla \varphi\|_\mu.$$

Hence,

$$\mathcal{A}_{t-\varepsilon}^t(\sigma^\varepsilon) = \int_{t-\varepsilon}^t L(\sigma_\tau^\varepsilon, \mathbf{v}_\tau^\varepsilon) d\tau = \frac{\varepsilon}{2} \|\nabla \varphi\|_\mu^2 - \int_{t-\varepsilon}^t \mathcal{V}((\mathbf{id} + (t - \tau)\nabla \varphi) \# \mu) d\tau.$$

It follows from this, (22) and the assumption  $\mathcal{V}$  is lower semicontinuous that

$$\limsup_{\varepsilon \downarrow 0} \frac{U(\mu, t) - U(\mu^\varepsilon, t - \varepsilon)}{\varepsilon} \leq \frac{1}{2} \int_{\mathbb{R}^d} |\nabla \varphi|^2 d\mu - \mathcal{V}(\mu). \quad (23)$$

We use the fact that  $(\zeta, \theta) \in \partial U(\mu, t)$  and (21) to obtain

$$\begin{aligned} U(\mu^\varepsilon, t - \varepsilon) - U(\mu, t) &\leq \iint_{\mathbb{R}^d \times \mathbb{R}^d} \langle \zeta(y), \bar{y} - y \rangle d\gamma^\varepsilon(y, \bar{y}) + (t - \varepsilon - t)\theta + o(W_2(\mu, \mu^\varepsilon)) + o(\varepsilon) \\ &= \varepsilon \int_{\mathbb{R}^d} \langle \zeta, \nabla \varphi \rangle d\mu - \varepsilon\theta + o(\varepsilon). \end{aligned}$$

Thus,

$$\liminf_{\varepsilon \downarrow 0} \frac{U(\mu, t) - U(\mu^\varepsilon, t - \varepsilon)}{\varepsilon} \geq \theta - \int_{\mathbb{R}^d} \langle \zeta, \nabla \varphi \rangle d\mu. \quad (24)$$

Combine (23) and (24) to deduce  $\theta + H(\mu, \zeta) \leq 1/2 \int_{\mathbb{R}^d} |\zeta + \nabla\varphi|^2 d\mu$ . Since  $\varphi \in C_c^\infty(\mathbb{R}^d)$  is arbitrary and  $\{\nabla\phi : \phi \in C_c^\infty(\mathbb{R}^d)\}$  is dense in  $\mathcal{T}_\mu\mathcal{M}$  which contains  $\zeta$ , we conclude that  $\theta + H(\mu, \zeta) \leq 0$ . This proves (11) and hence  $U$  is a viscosity subsolution.

**Step 2:** We prove (ii). In order to do that, let us introduce the augmented action

$$\mathcal{B}(\sigma) := \mathcal{A}_0^T(\sigma) + U_0(\sigma_0).$$

Fix  $\sigma \in AC^2(0, T; \mathcal{M})$ , let

$$\varphi \in C_c^\infty([0, T] \times \mathbb{R}^d), \quad \Phi := \nabla\varphi \quad \text{and} \quad \sigma_t^s = (\mathbf{id} + s\Phi(t, \cdot))_{\#}\sigma_t.$$

Let  $\mathbf{v}^s$  (resp.  $\mathbf{v}$ ) be the velocities of minimal norm associated to  $\sigma^s$  (resp.  $\sigma$ ). By definition, the continuity equation (17) (i) holds. Lemma 8 in [9] ensures

$$\int_0^T \|\mathbf{v}_t^s\|_{\sigma_t^s}^2 dt - \int_0^T \|\mathbf{v}_t\|_{\sigma_t}^2 dt \leq \int_0^T \left( C_t^2 s^2 + 2s \langle \mathbf{v}_t, \partial_t \Phi_t + \nabla_x \Phi_t \mathbf{v}_t \rangle_{\sigma_t} \right) dt, \quad (25)$$

where  $C_t = \|\partial_t \Phi + \nabla \Phi_t \mathbf{v}_t\|_{L^2(\sigma_t)}$ . Observe that

$$[\mathbf{id} \times (\mathbf{id} + s\Phi_t)]_{\#}\sigma_t \in \Gamma_o(\sigma_t, \sigma_t^s), \quad \text{for } 0 \leq s \ll 1.$$

Hence if  $\zeta_0 \in \partial U_0(\sigma_0)$ , then

$$U_0(\sigma_0^s) - U_0(\sigma_0) \leq s \langle \zeta_0, \Phi_0 \rangle_{\sigma_0} + o(s).$$

This together with (25) yields

$$\begin{aligned} \mathcal{B}(\sigma^s) - \mathcal{B}(\sigma) &\leq \int_0^T \left( C_t^2 \frac{s^2}{2} + s \langle \mathbf{v}_t, \partial_t \Phi_t + \nabla_x \Phi_t \mathbf{v}_t \rangle_{\sigma_t} \right) dt \\ &\quad - \int_0^T [\mathcal{V}(\sigma_t^s) - \mathcal{V}(\sigma_t)] dt + s \langle \zeta_0, \Phi_0 \rangle_{\sigma_0} + o(s). \end{aligned} \quad (26)$$

In the sequel, we assume that  $\sigma$  is a minimizer of (14) for  $t = T$ . We then have  $\mathcal{B}(\sigma) = U(\mu, T) \leq \mathcal{B}(\sigma^s)$ . Using this, we obtain from (26) that

$$0 \leq \int_0^T \langle \mathbf{v}_t, \partial_t \Phi_t + \nabla_x \Phi_t \mathbf{v}_t \rangle_{\sigma_t} dt - \limsup_{s \downarrow 0} \int_0^T \frac{\mathcal{V}(\sigma_t^s) - \mathcal{V}(\sigma_t)}{s} dt + \langle \zeta_0, \Phi_0 \rangle_{\sigma_0}. \quad (27)$$

Recall that (4) ensures that  $\sigma$  is continuous and so,  $\sigma([0, T])$  is a compact subset of  $\mathcal{M}$ . Thus,  $\sigma([0, T])$ ,  $\sigma^s([0, T])$  and  $\bar{\sigma}^s([0, T])$  are bounded sets in  $\mathcal{M}$ , where  $\bar{\sigma}_t^s := (\mathbf{id} - s\Phi(t, \cdot))_{\#}\sigma_t$ . Moreover, it follows from (20) that

$$\frac{\mathcal{V}(\sigma_t^s) - \mathcal{V}(\sigma_t)}{s} + C^* \geq \langle \nabla_\mu \mathcal{V}(\sigma_t), \Phi_t \rangle_{\sigma_t} \geq - \frac{\mathcal{V}(\bar{\sigma}_t^s) - \mathcal{V}(\sigma_t)}{s} - C^*.$$

Therefore, we conclude in particular that  $t \rightarrow \langle \nabla_\mu \mathcal{V}(\sigma_t), \Phi_t \rangle_{\sigma_t} \in L^1(0, T)$  since the function  $\mathcal{V}$  is bounded on bounded subsets. Furthermore, by Fatou's lemma and the fact  $\partial \mathcal{V}(\sigma_t)$  is nonempty, we obtain

$$\liminf_{s \downarrow 0} \int_0^T \frac{\mathcal{V}(\sigma_t^s) - \mathcal{V}(\sigma_t)}{s} dt \geq \int_0^T \liminf_{s \downarrow 0} \frac{\mathcal{V}(\sigma_t^s) - \mathcal{V}(\sigma_t)}{s} dt \geq \int_0^T \langle \nabla_\mu \mathcal{V}(\sigma_t), \Phi_t \rangle_{\sigma_t} dt.$$

Combining this with (27) , we get

$$0 \leq \int_0^T \left( \langle \mathbf{v}_t, \partial_t \Phi_t + \nabla_x \Phi_t \mathbf{v}_t \rangle_{\sigma_t} - \langle \nabla_\mu \mathcal{V}(\sigma_t), \Phi_t \rangle_{\sigma_t} \right) dt + \langle \zeta_0, \Phi_0 \rangle_{\sigma_0}. \quad (28)$$

We substitute  $\Phi$  by  $-\Phi$  in (28) to conclude that

$$0 = \int_0^T \left( \langle \mathbf{v}_t, \partial_t \Phi_t + \nabla_x \Phi_t \mathbf{v}_t \rangle_{\sigma_t} - \langle \nabla_\mu \mathcal{V}(\sigma_t), \Phi_t \rangle_{\sigma_t} \right) dt + \langle \zeta_0, \Phi_0 \rangle_{\sigma_0}.$$

Hence (17) (ii) holds.

QED.

**Remark 3.10.** We remark that condition (20) is satisfied if  $\mathcal{V}(\mu) = W_2^2(\mu, \nu_0)$ .

### 3.3 Conservation of the Hamiltonian along minimizing paths

In the next proposition, we prove conservation of the Hamiltonian along paths minimizing  $\mathcal{A}$  over  $P_T(\mu, \nu)$ . We use a technique which is standard in the finite dimensional setting. This technique has been applied to the Wasserstein setting by Ambrosio and Figalli [1] when  $\mathcal{V} \equiv 0$ .

**Proposition 3.11.** *Assume that  $\mathcal{V} : \mathcal{M} \rightarrow \mathbb{R}$  is Borel and bounded on compact subsets of  $\mathcal{M}$ . Suppose  $\sigma$  minimizes  $\mathcal{A}_T$  over  $P_T(\mu, \nu)$  and let  $\mathbf{v}$  be the velocity of minimal norm associated to  $\sigma$ . Then there exist  $\mathcal{N} \subset [0, T]$  a set of zero measure and  $\mathbf{v}^* : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that  $\mathbf{v}_t = \mathbf{v}_t^*$  for all  $t \in [0, T] \setminus \mathcal{N}$  and  $H(\sigma_t, \mathbf{v}_t^*)$  is independent of  $t$ .*

**Proof:** We have  $\mathcal{A}_T(\bar{\sigma}) < \infty$ , where  $\bar{\sigma}$  is the geodesic connecting  $\mu$  to  $\nu$ . Thus,  $\mathcal{A}_T(\sigma) < \infty$ . Because  $\mathcal{V}$  is bounded on compact subsets of  $\mathcal{M}$  and  $\sigma([0, T])$  is a compact set, we conclude that  $\mathcal{V} \circ \sigma \in L^\infty(0, T)$ . Thus

$$t \rightarrow \|\mathbf{v}_t\|_{\sigma_t}^2 \in L^1(0, T) \quad \text{and} \quad t \rightarrow H(\sigma_t, \mathbf{v}_t) \in L^1(0, T). \quad (29)$$

Let  $\varphi \in C_c^\infty(0, T)$ . We consider the functions  $T_\varepsilon : [0, T] \rightarrow [0, T]$  defined by

$$T_\varepsilon(s) = s + \varepsilon\varphi(s).$$

These are increasing diffeomorphisms for  $|\varepsilon| \ll 1$ . We denote by  $S_\varepsilon$  the inverse of  $T_\varepsilon$ . It satisfies

$$S_\varepsilon(t) = t - \varepsilon\varphi(t) + o(\varepsilon), \quad \dot{T}_\varepsilon(S_\varepsilon(t)) = 1 - \varepsilon\dot{\varphi}(t) + o(\varepsilon). \quad (30)$$

Here,  $\dot{\varphi}(t)$  is the derivative of  $\varphi$  with respect to  $t$ . We define

$$\sigma_s^\varepsilon := \sigma_{T_\varepsilon(s)} \quad \text{and} \quad \mathbf{v}_s^\varepsilon := \dot{T}_\varepsilon(s) \mathbf{v}_{T_\varepsilon(s)} \quad \text{for } s \in [0, T],$$

where  $\dot{T}_\varepsilon(s)$  stands for the derivative of  $T_\varepsilon(s)$  with respect to  $s$ . One readily checks that  $\mathbf{v}^\varepsilon$  is the velocity of minimal norm associated to  $\sigma^\varepsilon$ . Thus,

$$\begin{aligned}
\mathcal{A}_T(\sigma^\varepsilon) &= \int_0^T \left( \frac{1}{2} \dot{T}_\varepsilon^2(s) \|\mathbf{v}_{T_\varepsilon(s)}\|_{\sigma_{T_\varepsilon(s)}}^2 - \mathcal{V}(\sigma_{T_\varepsilon(s)}) \right) ds \\
&= \int_0^T \left( \frac{1}{2} \dot{T}_\varepsilon(S_\varepsilon(t)) \|\mathbf{v}_t\|_{\sigma_t}^2 - \frac{1}{\dot{T}_\varepsilon(S_\varepsilon(t))} \mathcal{V}(\sigma_t) \right) dt \tag{31}
\end{aligned}$$

$$= \int_0^T \left( L(\sigma_t, \mathbf{v}_t) - \varepsilon \dot{\varphi}(t) H(\sigma_t, \mathbf{v}_t) + o(\varepsilon) \right) dt \tag{32}$$

$$= \mathcal{A}_T(\sigma) - \varepsilon \int_0^T \dot{\varphi}(t) H(\sigma_t, \mathbf{v}_t) dt + o(\varepsilon). \tag{33}$$

The equality (31) is obtained by a change of variables whereas (32) is a consequence of (30). To obtain (33) we have used the fact that, by (29),

$$t \rightarrow L(\sigma_t, \mathbf{v}_t) \in L^1(0, T) \quad \text{and} \quad t \rightarrow \dot{\varphi}(t) H(\sigma_t, \mathbf{v}_t) \in L^1(0, T).$$

Since  $\sigma^\varepsilon \in P_T(\mu, \nu)$  and  $\sigma$  minimizes  $\mathcal{A}_T$  over  $P_T(\mu, \nu)$ , it follows from (33) that

$$\int_0^T \dot{\varphi}(t) H(\sigma_t, \mathbf{v}_t) dt = 0.$$

As  $\varphi \in C_c^\infty(0, T)$  is arbitrary, this concludes the proof. QED.

**Remark 3.12.** *Let  $\sigma$  and  $\mathbf{v}^*$  be as in proposition 3.11. The proposition implies that  $t \rightarrow \|\mathbf{v}_t^*\|_{\sigma_t} \in L^\infty(0, T)$ . If, in addition  $\mathcal{V}$  is continuous, then as  $\mathcal{V} \circ \sigma$  is also continuous, we in fact have  $t \rightarrow \|\mathbf{v}_t^*\|_{\sigma_t} \in C[0, T]$ .*

## 4 Hamilton-Jacobi equations on $\mathcal{P}_2(\mathbb{R})$ : Hamiltonians with smooth potentials

In this section we study Hamilton-Jacobi equations in the metric space  $\mathcal{M} := \mathcal{P}_2(\mathbb{R})$ . We recall that

$$X = (-1/2, 1/2) \text{ and } \nu_0 = \mathcal{L}_X^1.$$

Unlike the case  $d > 1$ , we can still prove existence of a minimizer of the action  $\mathcal{A}_T$  over the set  $P_T(\nu, \mu)$  while allowing the potential  $\mathcal{V}(\mu)$  to contain a term such as  $\mathcal{N}_a(\mu) = a \int_{\mathbb{R}} y^2 d\mu(y)$ . For  $a > 0$ , the functional  $\mathcal{N}_a$  is not upper semicontinuous on bounded subsets of  $\mathcal{M}$  for the narrow convergence. This is a source of additional difficulty which we overcome when  $d = 1$ .

For the sake of illustration, let us describe first some of the results we obtain in a simple case. For parameters  $a \in \mathbb{R}$  we define the Lagrangians:

$$L^a(\mu, w) = \frac{1}{2} (\|w\|_\mu^2 - a M_2(\mu)) \quad \text{for } \mu \in \mathcal{M} \text{ and } w \in \mathcal{T}_\mu \mathcal{M},$$

where  $M_2(\mu) := \int_{\mathbb{R}^d} |y|^2 d\mu(y)$  is the second moment of  $\mu$ . The associated action is

$$\mathcal{A}_T(\sigma) := \int_0^T L^a(\sigma_t, v_t) dt, \quad \sigma \in AC^2(0, T; \mathcal{M}).$$

Here  $v$  is the velocity of minimal norm for  $\sigma$ . Suppose  $\{\sigma^n\}_{n \in \mathbf{N}} \subset P_T(\mu, \nu)$  is such that

$$\sup_{n \in \mathbf{N}} \int_0^T \|v^n\|_{\sigma_t^n}^2 dt < \infty \quad \text{and} \quad \sup_{n \in \mathbf{N}} \int_0^T M_2(\sigma_t^n) dt < \infty. \quad (34)$$

Here,  $v^n$  is the velocity of minimal norm for  $\sigma^n$ . Suppose in addition that for each  $t$ ,  $\{\sigma_t^n\}_{n \in \mathbf{N}}$  converges narrowly to  $\sigma_t$  as  $n$  tends to  $\infty$ . In general, we have

$$\liminf_{n \rightarrow \infty} \int_0^T \|v_t^n\|_{\sigma_t^n}^2 dt \geq \int_0^T \|v_t\|_{\sigma_t}^2 dt, \quad \text{and} \quad \liminf_{n \rightarrow \infty} \int_0^T M_2(\sigma_t^n) dt \geq \int_0^T M_2(\sigma_t) dt. \quad (35)$$

Therefore, only if  $a \leq 0$  does (35) yield in general that

$$\liminf_{n \rightarrow \infty} \mathcal{A}_T(\sigma^n) \geq \mathcal{A}_T(\sigma). \quad (36)$$

The fact that we do not know whether (36) holds for  $a > 0$  is an obstruction to proving existence of a minimizer of  $\mathcal{A}_T$  over  $P_T(\nu, \mu)$  in that case when  $d \geq 2$ . However, when  $d = 1$ , it is shown in [9] that one can weaken the condition on  $a$  by only imposing that  $T^2 a < \pi^2$  and still conclude that (36) holds. We will use this to prove existence of a path minimizing  $\mathcal{A}_T(\sigma) + U_0(\sigma_0)$  over  $P_T(\cdot, \mu)$ . Since  $M_2$  is continuous and  $\sigma$  is continuous, the conservation of the Hamiltonian obtained in proposition 3.11 gives that if we modify  $v$  on a set of zero measure then  $t \rightarrow \|v_t\|_{\sigma_t}$  is continuous. Only in one dimension are we able to obtain the stronger conclusion that for every  $t$ ,  $\|v_t\|_{\sigma_t}$  equals the metric derivative  $|\sigma'(t)|$ . That property shall be instrumental for proving that  $U$  defined in (14) is a viscosity solution of (16).

In the sequel, we consider a class of Lagrangians which are more general than  $L^a$ . These are

$$L(\mu, w) = \frac{1}{2} (\|w\|_{\mu}^2 - aM_2(\mu)) - \mathcal{W}(\mu) \quad \text{for } (\mu, w) \in \mathcal{TM},$$

where  $a \in \mathbb{R}$ . The Hamiltonian associated to  $L$  is

$$H(\mu, w) = \frac{1}{2} (\|w\|_{\mu}^2 + aM_2(\mu)) + \mathcal{W}(\mu) \quad \text{for } (\mu, w) \in \mathcal{TM}.$$

Similarly, we define

$$\bar{L}(M, \mathbf{n}) = \frac{1}{2} (\|\mathbf{n}\|_{\nu_0}^2 - a\|M\|_{\nu_0}^2) - \bar{\mathcal{W}}(M) \quad \text{for } (M, \mathbf{n}) \in L^2(\nu_0) \times L^2(\nu_0),$$

where  $a \in \mathbb{R}$  and the function  $\bar{\mathcal{W}} : L^2(\nu_0) \rightarrow \mathbb{R}$  is defined by

$$\bar{\mathcal{W}}(m) := \mathcal{W}(m \# \nu_0) \quad \text{for } m \in L^2(\nu_0). \quad (37)$$

The Hamiltonian associated to  $\bar{L}$  is

$$\bar{H}(M, \mathbf{n}) = \frac{1}{2} (\|\mathbf{n}\|_{\nu_0}^2 + a\|M\|_{\nu_0}^2) + \bar{\mathcal{W}}(M) \quad \text{for } (M, \mathbf{n}) \in L^2(\nu_0) \times L^2(\nu_0).$$

**Remark 4.1.** Let  $\text{Mon}$  be the set of  $m \in L^2(\nu_0)$  which are monotone nondecreasing. Here  $\nu_0 := \mathcal{L}_{|X}^1$  and  $X = (-1/2, 1/2)$ . It is well-known that  $I : \text{Mon} \rightarrow \mathcal{P}_2(\mathbb{R})$  defined by  $I(m) = m \# \nu_0$  is an isomorphism. It is shown in [9] that this induces a map  $\mathcal{I} : M \in H^1(0, T; \text{Mon}) \rightarrow$

$AC^2(0, T; \mathcal{P}_2(\mathbb{R}))$  defined by  $\mathcal{I}(M)_t = I(M_t) =: \sigma_t$ . Furthermore, if  $v$  is the velocity of minimal norm associated to  $\sigma$  then  $v_t \circ M_t = M'_t$  for almost every  $t \in (0, T)$ . Here,  $M'$  is the functional time derivative of  $M$ . Hence,  $\|v_t\|_{\sigma_t} = \|M'_t\|_{\nu_0}$  for almost every  $t \in (0, T)$ . Furthermore, if  $\bar{M} \in H^1(0, T; L^2(\nu_0))$  is such that  $\bar{M}_t \# \nu_0 = M_t \# \nu_0$ , then  $\|M'_t\|_{\nu_0} \leq \|\bar{M}'_t\|_{\nu_0}$  for almost every  $t \in (0, T)$ .

Given  $U_0$  an initial real valued function defined on  $\mathcal{M}$ , we introduce the augmented actions

$$\mathbf{A}_t(\sigma) := \int_0^t L(\sigma_\tau, v_\tau) d\tau + U_0(\sigma_0) \quad \text{for } \sigma \in AC^2(0, t; \mathcal{M}).$$

Here  $v$  is the velocity of minimal norm associated to  $\sigma$ . For these  $\sigma$ , let  $M \in H^1(0, t; L^2(\nu_0))$  be uniquely determined by the fact that  $M_t$  pushes  $\nu_0$  forward to  $\sigma_t$  and  $M_t$  is monotone nondecreasing. It follows from remark 4.1 that  $\mathbf{A}_t(\sigma) = \mathbf{B}_t(M)$ , where

$$\mathbf{B}_t(M) := \int_0^t \bar{L}(M_\tau, M'_\tau) d\tau + \bar{U}_0(M_0). \quad (38)$$

Here,  $M'$  stands for the functional derivative of  $M$  with respect to time, and the function  $\bar{U}_0 : L^2(\nu_0) \rightarrow \mathbb{R}$  is defined by

$$\bar{U}_0(m) := U_0(m \# \nu_0) \quad \text{for } m \in L^2(\nu_0). \quad (39)$$

We impose that there exist  $1 < p < 2$ ,  $\bar{e}_0 \geq 0$  and  $\bar{e}_1 \in \mathbb{R}$  such that

$$\mathcal{W}(\mu) \leq \bar{e}_0 \int_{\mathbb{R}} |y|^p d\mu(y) + \bar{e}_1 \quad \text{for } \mu \in \mathcal{M}. \quad (40)$$

Note that this is equivalent to

$$\bar{\mathcal{W}}(m) \leq \bar{e}_0 \int_X |m|^p dx + \bar{e}_1 \quad \text{for } m \in L^2(\nu_0). \quad (41)$$

We also assume that  $\bar{U}_0$  is  $\lambda$ -convex and the subdifferential  $\partial \bar{U}_0(m)$  is nonempty for every  $m \in L^2(\nu_0)$ . We denote the element of minimal norm of that convex set by  $\nabla_{L^2} \bar{U}_0(m)$ .

#### 4.1 Basic properties of augmented actions

The main results will be obtained under the following assumptions: we suppose that  $\bar{U}_0$  is Frechet differentiable,  $\lambda$ -convex for some real number  $\lambda$  satisfying

$$(H1) \quad T\lambda^- < 1 - \frac{4a^+ T^2}{\pi^2}.$$

We suppose that  $\mathcal{W}$  is upper semicontinuous for the narrow convergence on bounded subsets of  $\mathcal{M}$ :

$$(H2) \quad \limsup_{n \rightarrow \infty} \mathcal{W}(\mu^n) \leq \mathcal{W}(\mu) \quad \text{whenever } \{\mu^n\}_{n \in \mathbb{N}} \text{ is bounded in } \mathcal{M} \text{ and converges narrowly to } \mu.$$

Occasionally, we will assume that the superdifferential  $\partial^* \bar{\mathcal{W}}(m)$  is nonempty at every  $m \in L^2(\nu_0)$  and so, its element of minimal norm  $\nabla_{L^2} \bar{\mathcal{W}}(m)$  exists. In that case, we assume that

$$(H3) \quad m \longrightarrow \nabla_{L^2} \bar{\mathcal{W}}(m) \text{ is Borel and bounded on compact subsets of } L^2(\nu_0).$$

We introduce the bilinear and quadratic functionals on  $H^1(0, T; L^2(\nu_0))$  as follows

$$\mathcal{B}_T(M, N) := \int_0^T \int_{\mathbb{R}} (M'_\tau N'_\tau - aM_\tau N_\tau) d\nu_0 d\tau, \quad 2\mathcal{Q}_T(M) := \mathcal{B}_T(M, M).$$

We first point out that elementary computations yield, if  $R \in H^1(0, T; L^2(\nu_0))$ ,

$$\|R_t\|_{\nu_0} \leq \|R_T\|_{\nu_0} + \sqrt{T-t} \left( \int_t^T \|R'_\tau\|_{\nu_0}^2 d\tau \right)^{1/2} \quad \text{for } t \in [0, T]. \quad (42)$$

If, in addition  $R_T = 0$ , we recall a sharper inequality, the Poincaré-Wirtinger inequality given in proposition 2.1:

$$\pi^2 \int_0^T \|R_\tau\|_{\nu_0}^2 d\tau \leq 4T^2 \int_0^T \|R'_\tau\|_{\nu_0}^2 d\tau. \quad (43)$$

Hence,

$$2\mathcal{Q}_T(R) \geq \left(1 - 4\frac{a+T^2}{\pi^2}\right) \int_0^T \|R'_\tau\|_{\nu_0}^2 d\tau. \quad (44)$$

**Remark 4.2.** Suppose  $\bar{U}_0$  is  $\lambda$ -convex and its subdifferential  $\partial\bar{U}_0(m)$  is nonempty for every  $m \in L^2(\nu_0)$ . Let  $M, N \in AC^2(0, T; L^2(\nu_0))$  be such that  $M_T = N_T$ . Then

$$\begin{aligned} \mathcal{Q}_T(N) + \bar{U}_0(N_0) &\geq \mathcal{Q}_T(M) + \bar{U}_0(M_0) + \mathcal{B}_T(M, N - M) + \langle \nabla_{L^2} \bar{U}_0(M_0), N_0 - M_0 \rangle_{\nu_0} \\ &+ \frac{1}{2} \left(1 - 4a^+ \frac{T^2}{\pi^2} - T\lambda^-\right) \int_0^T \|N'_\tau - M'_\tau\|_{\nu_0}^2 d\tau. \end{aligned} \quad (45)$$

**Proof:** Setting  $R = M - N$  and  $t = 0$  in (42) we have

$$\|M_0 - N_0\|_{\nu_0}^2 \leq T \int_0^T \|M'_\tau - N'_\tau\|_{\nu_0}^2 d\tau. \quad (46)$$

It is straightforward to obtain

$$\begin{aligned} \mathcal{Q}_T(N) + \bar{U}_0(N_0) &\geq \mathcal{Q}_T(M) + \bar{U}_0(M_0) + \mathcal{B}_T(M, N - M) + \langle \nabla_{L^2} \bar{U}_0(M_0), N_0 - M_0 \rangle_{\nu_0} \\ &+ \mathcal{Q}_T(N - M) + \frac{\lambda}{2} \|M_0 - N_0\|_{\nu_0}^2. \end{aligned} \quad (47)$$

We use (44), (46) and (47) to obtain (45). QED.

**Corollary 4.3.** Suppose  $\bar{U}_0$  is  $\lambda$ -convex, its subdifferential  $\partial\bar{U}_0(m)$  is nonempty for every  $m \in L^2(\nu_0)$  and  $\pi^2 - 4a^+T^2 \geq \pi^2T\lambda^-$ . Let  $\{M^n\}_{n \in \mathbf{N}}$  be a sequence in  $AC^2(0, T; L^2(\nu_0))$  such that for each  $t \in [0, T]$ ,  $\{M_t^n\}_n$  converges weakly to  $M_t$  in  $L^2(\nu_0)$ . Then

$$\liminf_{n \rightarrow \infty} \{ \mathcal{Q}_T(M^n) + \bar{U}_0(M_0^n) \} \geq \mathcal{Q}_T(M) + \bar{U}_0(M_0).$$

**Proof:** By lemma 3 in [9],  $\{M^n\}_n$  converges weakly to  $M$  in  $L^2(X \times (0, T))$  and  $\{(M^n)'\}_n$  converges weakly to  $M'$  in  $L^2(X \times (0, T))$ . Thus,

$$\lim_{n \rightarrow \infty} \mathcal{B}_T(M, M^n - M) = 0.$$

This, together with remark 4.2 and the fact that

$$\lim_{n \rightarrow \infty} \langle \nabla_{L^2} \bar{U}_0(M_0), M_0^n - M_0 \rangle_{\nu_0} = 0,$$

yields the proof. To apply remark 4.2, we have used that  $\pi^2 - 4a^+T^2 \geq \pi^2T\lambda^-$ . QED.

We now show a coercivity property of the action  $\mathbf{B}_T$  defined in (38).

**Lemma 4.4** (Coercivity). *Suppose  $\bar{U}_0$  is  $\lambda$ -convex and its subdifferential  $\partial \bar{U}_0(m)$  is nonempty for every  $m \in L^2(\nu_0)$ . Suppose  $\bar{W}$  satisfies (41). Let  $M, N \in AC^2(0, T; L^2(\nu_0))$  be such that  $M_T = N_T$ . Then, for each  $\delta > 0$ , we have*

$$\begin{aligned} \mathbf{B}_T(N) &\geq \mathcal{Q}_T(M) + \bar{U}_0(M_0) + \frac{1}{2} \left( 1 - 4a^+ \frac{T^2}{\pi^2} - 2b_\delta - T\lambda^- - \delta \right) \int_0^T \|N'_\tau - M'_\tau\|_{\nu_0}^2 d\tau \\ &\quad - c_\delta - \frac{1}{2\delta} \left( \sqrt{T} \|\nabla_{L^2} \bar{U}_0(M_0)\|_{\nu_0} + \sqrt{\int_0^T \|M'_\tau\|_{\nu_0}^2 d\tau} + \frac{2|a|T}{\pi} \sqrt{\int_0^T \|M_\tau\|_{\nu_0}^2 d\tau} \right)^2, \end{aligned}$$

where the constants  $b_\delta$  and  $c_\delta$  are defined by

$$\pi^2 b_\delta := 4\bar{e}_0 p T^2 \delta^{\frac{2}{p}} \quad \text{and} \quad 2c_\delta := 2\bar{e}_1 T + 2\bar{e}_0 p \delta^{\frac{2}{p}} \int_0^T \|M_\tau\|_{\nu_0}^2 d\tau + \bar{e}_0 (2-p) T \delta^{\frac{2}{p-2}}.$$

**Proof:** Young's inequality yields that for every  $\delta > 0$ ,

$$|M|^p \leq \frac{p\delta^{\frac{2}{p}}}{2} |M|^2 + \frac{2-p}{2} \delta^{\frac{2}{p-2}}.$$

We use this, together with (41) and (43) in which we set  $R = M - N$  to obtain

$$\int_0^T \bar{W}(N_\tau) d\tau \leq c_\delta + b_\delta \int_0^T \|N'_\tau - M'_\tau\|_{\nu_0}^2 d\tau. \quad (48)$$

We combine (45) and (48) to obtain

$$\begin{aligned} \mathbf{B}_T(N) &\geq \mathcal{Q}_T(M) + \bar{U}_0(M_0) + \mathcal{B}_T(M, N - M) + \langle \nabla_{L^2} \bar{U}_0(M_0), N_0 - M_0 \rangle_{\nu_0} \\ &\quad + \frac{1}{2} \left( 1 - 4a^+ \frac{T^2}{\pi^2} - T\lambda^- - 2b_\delta \right) \int_0^T \|N'_\tau - M'_\tau\|_{\nu_0}^2 d\tau - c_\delta. \end{aligned} \quad (49)$$

Note that Hölder's inequality and (42) with  $R = N - M$  yield

$$|\langle \nabla_{L^2} \bar{U}_0(M_0), N_0 - M_0 \rangle_{\nu_0}| \leq \sqrt{T} \|\nabla_{L^2} \bar{U}_0(M_0)\|_{\nu_0} \left( \int_0^T \|N'_\tau - M'_\tau\|_{\nu_0}^2 d\tau \right)^{\frac{1}{2}} \quad (50)$$

and

$$\begin{aligned}
|\mathcal{B}_T(M, N - M)| &\leq \sqrt{\int_0^T \|M'_\tau\|_{\nu_0}^2 d\tau} \sqrt{\int_0^T \|N'_\tau - M'_\tau\|_{\nu_0}^2 d\tau} \\
&+ |a| \sqrt{\int_0^T \|M_\tau\|_{\nu_0}^2 d\tau} \sqrt{\int_0^T \|N_\tau - M_\tau\|_{\nu_0}^2 d\tau} \\
&\leq \sqrt{\int_0^T \|M'_\tau\|_{\nu_0}^2 d\tau} \sqrt{\int_0^T \|N'_\tau - M'_\tau\|_{\nu_0}^2 d\tau} \\
&+ |a| \sqrt{\int_0^T \|M_\tau\|_{\nu_0}^2 d\tau} \sqrt{\frac{4T^2}{\pi^2} \int_0^T \|N'_\tau - M'_\tau\|_{\nu_0}^2 d\tau}. \tag{51}
\end{aligned}$$

We combine (49)–(51) to get

$$\begin{aligned}
\mathbf{B}_T(N) &\geq \mathcal{Q}_T(M) + \bar{U}_0(M_0) \\
&- \sqrt{\int_0^T \|N'_\tau - M'_\tau\|_{\nu_0}^2 d\tau} \left( \sqrt{T} \|\nabla_{L^2} \bar{U}_0(M_0)\|_{\nu_0} + \sqrt{\int_0^T \|M'_\tau\|_{\nu_0}^2 d\tau} + \frac{2|a|T}{\pi} \sqrt{\int_0^T \|M_\tau\|_{\nu_0}^2 d\tau} \right) \\
&+ \frac{1}{2} \left( 1 - 4a^+ \frac{T^2}{\pi^2} - 2b_\delta - T\lambda^- \right) \int_0^T \|N'_\tau - M'_\tau\|_{\nu_0}^2 d\tau - c_\delta.
\end{aligned}$$

Using Young's inequality to estimate the expression in the second line above, we conclude the proof. QED.

**Remark 4.5.** Let  $\{m^n\}_{n \in \mathbf{N}} \subset L^2(\nu_0)$  be a sequence of monotone nondecreasing maps converging weakly to  $m$  in  $L^2(\nu_0)$ . Since  $\{m^n\}_{n \in \mathbf{N}}$  is bounded in  $L^2(\nu_0)$ , by remark 6 of [9], it is bounded in  $BV_{loc}(X)$ . The Sobolev Embedding theorem ensures that  $\{m^n\}_{n \in \mathbf{N}}$  converges to  $m$  in  $L^1_{loc}(X)$ . In fact, because  $X$  is a bounded set,  $\{m^n\}_{n \in \mathbf{N}}$  converges to  $m$  in  $L^p(\nu_0)$  for all  $1 \leq p < 2$ .

**Corollary 4.6.** Suppose  $\mathcal{W}$  satisfies (H2) and the growth condition (40). Let  $\{M^n\}_{n \in \mathbf{N}}$  be a sequence in  $L^2(0, T; L^2(\nu_0))$  such that  $M_t^n$  is monotone nondecreasing for every  $t$  in  $[0, T]$ . Suppose that for each  $t \in [0, T]$ ,  $\{M_t^n\}_n$  converges weakly to  $M_t$  in  $L^2(\nu_0)$ . Then

$$\limsup_{n \rightarrow \infty} \int_0^T \bar{\mathcal{W}}(M_t^n) dt \leq \int_0^T \bar{\mathcal{W}}(M_t) dt.$$

**Proof:** For  $t \in [0, T]$  we define

$$\sigma_t^n := (M_t^n)_{\#} \nu_0 \quad \text{and} \quad \sigma_t := M_t_{\#} \nu_0.$$

By remark 4.5,  $\{M_t^n\}_{n \in \mathbf{N}}$  converges to  $M_t$  in  $L^p(\nu_0)$  for any  $1 \leq p < 2$ . Hence,

$$\lim_{n \rightarrow \infty} \bar{e}_0 \int_X |M_t^n|^p dx + \bar{e}_1 = \bar{e}_0 \int_X |M_t|^p dx + \bar{e}_1 \tag{52}$$

and  $\{\sigma_t^n\}_{n \in \mathbf{N}}$  converges narrowly to  $\sigma_t$ . Thus,

$$\liminf_{n \rightarrow \infty} -\bar{\mathcal{W}}(M_t^n) = \liminf_{n \rightarrow \infty} -\mathcal{W}(\sigma_t^n) \geq -\mathcal{W}(\sigma_t) = -\bar{\mathcal{W}}(M_t).$$

This, together with (52), (41) and Fatou's lemma, yields

$$\begin{aligned}
\int_0^T dt \int_X \bar{e}_0 |M_t|^p dx + \liminf_{n \rightarrow \infty} \int_0^T -\bar{\mathcal{W}}(M_t^n) dt &= \liminf_{n \rightarrow \infty} \int_0^T \left( \int_X \bar{e}_0 |M_t^n|^p - \bar{\mathcal{W}}(M_t^n) \right) dt \\
&\geq \int_0^T \liminf_{n \rightarrow \infty} \left( \int_X \bar{e}_0 |M_t^n|^p - \bar{\mathcal{W}}(M_t^n) \right) dt \\
&\geq \int_0^T \left( \int_X \bar{e}_0 |M_t|^p - \bar{\mathcal{W}}(M_t) \right) dt.
\end{aligned}$$

This proves the corollary. QED.

**Remark 4.7** (Strict convexity). *Suppose  $\bar{U}_0$  is  $\lambda$ -convex,  $-\bar{\mathcal{W}}$  is  $l$ -convex and the subdifferentials  $\partial \bar{U}_0(m)$ ,  $\partial(-\bar{\mathcal{W}})(m)$  are nonempty for every  $m \in L^2(\nu_0)$ . Assume also that  $\bar{\mathcal{W}}$  satisfies (H3) and the growth condition (40). If  $M, N \in AC^2(0, T; L^2(\nu_0))$  are such that  $M_T = N_T$  then*

$$\begin{aligned}
\mathbf{B}_T(N) &\geq \mathbf{B}_T(M) + \mathcal{B}_T(M, N - M) + \langle \nabla_{L^2} \bar{U}_0(M_0), N_0 - M_0 \rangle_{\nu_0} \\
&\quad - \int_0^T \langle \nabla_{L^2} \bar{\mathcal{W}}(M_\tau), N_\tau - M_\tau \rangle_{\nu_0} d\tau + c(T, l, \lambda, a) \int_0^T \|N'_\tau - M'_\tau\|_{\nu_0}^2 d\tau. \quad (53)
\end{aligned}$$

Here,  $2c(T, l, \lambda, a) = 1 - 4(a^+ + l^-) \frac{T^2}{\pi^2} - T\lambda^-$ . Hence, if  $c(T, l, \lambda, a) > 0$ , then  $\mathbf{B}_T$  is strictly convex on  $P_T(\cdot, m)$  for every  $m \in L^2(\nu_0)$ .

**Corollary 4.8.** *Suppose  $\bar{U}_0$  is  $\lambda$ -convex,  $-\bar{\mathcal{W}}$  is  $l$ -convex and their subdifferentials  $\partial \bar{U}_0(m)$ ,  $\partial(-\bar{\mathcal{W}})(m)$  are nonempty for every  $m \in L^2(\nu_0)$ . Assume also (H3) and (40) are satisfied and  $1 - 4(a^+ + l^-)T^2/\pi^2 \geq T\lambda^-$ . Let  $\{M^n\}_{n \in \mathbf{N}}$  be a sequence in  $AC^2(0, T; L^2(\nu_0))$  such that for each  $t \in [0, T]$ ,  $\{M_t^n\}_n$  converges weakly to  $M_t$  in  $L^2(\nu_0)$ . Then*

$$\liminf_{n \rightarrow \infty} \mathbf{B}_T(M^n) \geq \mathbf{B}_T(M).$$

**Proof:** Use remark 4.7 in place of remark 4.2 in the proof of corollary 4.3. QED.

## 4.2 Existence of viscosity solutions for the Cauchy problem

We will show that the value function  $U$  defined in (14) is a minimum which satisfies the Hamilton-Jacobi equation (16). Our study includes “linear” functionals of the form

$$U_0(\mu) = \int_{\mathbb{R}} u_0 d\mu,$$

where  $u_0 \in C^1(\mathbb{R})$  is  $\lambda$ -convex as a function defined on  $\mathbb{R}$  and  $-C \leq u_0(t) \leq C(|t|^2 + 1)$  for some  $C > 0$ . In that case  $\bar{U}_0(M) = \int_X u_0 \circ M d\nu_0$ .

**Theorem 4.9.** (i) *Suppose (H1), (H2) and (40) hold. Assume also  $\bar{U}_0$  is bounded below and the subdifferential  $\partial \bar{U}_0(m)$  is nonempty for every  $m \in L^2(\nu_0)$ . Then for any monotone nondecreasing function  $m \in L^2(\nu_0)$ , we have  $\mathbf{B}_T$  admits a minimizer  $M^o$  over  $P_T(\cdot, m)$ .*

(ii) *Suppose in addition (H3) holds,  $-\bar{\mathcal{W}}$  is  $l$ -convex,  $\partial(-\bar{\mathcal{W}})(m)$  is nonempty for every  $m$  and*

$T\lambda^- + 4(a^+ + l^-)T^2/\pi^2 < 1$ . Then the minimizer of  $\mathbf{B}_T$  over  $P_T(\cdot, m)$  is unique.

(iii) In addition to the assumptions in (i), suppose that  $\bar{\mathcal{W}}, \bar{U}_0$  are differentiable and  $\nabla_{L^2}\bar{\mathcal{W}}$  is continuous. Then  $M^o \in H^2(0, T; L^2(\nu_0))$  and satisfies the Euler-Lagrange equation

$$(M_t^o)'' + aM_t^o = -\nabla_{L^2}\bar{\mathcal{W}}(M_t^o), \quad M_T^o = m, \quad (M^o)'_0 = \nabla_{L^2}\bar{U}_0(M_0^o). \quad (54)$$

(iv) If the assumptions in (ii) holds then any element of  $P_T(\cdot, m)$  which is a solution for (54) minimizes  $\mathbf{B}_T$  over  $P_T(\cdot, m)$ .

**Proof:** Suppose the assumptions in (i) hold. Let  $\bar{M}_t \equiv m$  for all  $t \in [0, T]$ . Since  $\mathbf{B}_T(\bar{M})$  is finite, we conclude that  $\mathbf{B}_T$  is not identically  $+\infty$  on  $P_T(\cdot, m)$ . This, together with the facts  $\bar{U}_0$  is bounded below and  $\mathcal{W}$  satisfies (40), implies that  $\inf \{\mathbf{B}_T(M) : M \in P_T(\cdot, m)\}$  is finite. Because of (H1) we may choose  $\delta > 0$  small enough so that  $0 < 1 - 4a^+ \frac{T^2}{\pi^2} - 2b_\delta - T\lambda^- - \delta$ , where  $b_\delta$  is the constant given in lemma 4.4. Lemma 4.4 and the Poincaré-Wirtinger inequality (43) yield that the sublevel sets of  $\mathbf{B}_T$  over  $P_T(\cdot, m)$  are bounded subsets of  $H^1(0, T; L^2(\nu_0))$ . Hence, every minimizing sequence  $\{M^n\}_{n \in \mathbb{N}}$  of  $\mathbf{B}_T$  over  $P_T(\cdot, m)$  is a bounded subset of  $H^1(0, T; L^2(\nu_0))$ . By remark 5 of [9] and the fact that  $m$  is monotone nondecreasing, we may assume without loss of generality that for each  $t \in [0, T]$ ,  $M_t^n$  is monotone nondecreasing. We apply proposition 4 of [9] to obtain a subsequence of  $\{M^n\}_{n \in \mathbb{N}}$  and  $M^o \in H^1(0, T; L^2(\nu_0))$  such that for each  $t \in [0, T]$ ,  $\{M_t^n\}_{n \in \mathbb{N}}$  converges weakly in  $L^2(\nu_0)$  to  $M_t^o$ . It then follows from remark 4.5 that for each  $t \in [0, T]$ ,  $\{M_t^n\}_{n \in \mathbb{N}}$  converges strongly to  $M_t^o$  in  $L^p(\nu_0)$  for all  $1 \leq p < 2$ . Hence  $M_t^o$  is monotone nondecreasing and  $M^o \in P_T(\cdot, m)$ . We apply corollary 4.3 and corollary 4.6 to conclude that

$$\liminf_{n \rightarrow \infty} \mathbf{B}_T(M^n) \geq \mathbf{B}_T(M^o).$$

This proves that  $M^o$  minimizes  $\mathbf{B}_T$  over  $P_T(\cdot, m)$ . Under the assumptions in (ii), remark 4.7 gives that  $\mathbf{B}_T$  is strictly convex on the convex set  $P_T(\cdot, m)$ . This ensures uniqueness of the minimizer.

Suppose that the assumptions in (iii) hold. Let  $M \in H^1(0, T; L^2(\nu_0))$  be such that  $M_T \equiv 0$ . Because  $M^o + \varepsilon M \in P_T(\cdot, m)$  for every  $\varepsilon \in \mathbb{R}$  we have

$$\begin{aligned} 0 &= \frac{d}{d\varepsilon} \mathbf{B}_T(M^o + \varepsilon M)|_{\varepsilon=0} \\ &= \mathcal{B}_T(M^o, M) + \langle \nabla_{L^2}\bar{U}_0(M_0^o), M_0 \rangle_{\nu_0} - \int_0^T \langle \nabla_{L^2}\bar{\mathcal{W}}(M_\tau^o), M_\tau \rangle_{\nu_0} d\tau \\ &= \int_0^T \left( \langle (M_\tau^o)', M_\tau' \rangle_{\nu_0} - \langle aM_\tau^o + \nabla_{L^2}\bar{\mathcal{W}}(M_\tau^o), M_\tau \rangle_{\nu_0} \right) d\tau + \langle \nabla_{L^2}\bar{U}_0(M_0^o), M_0 \rangle_{\nu_0}. \end{aligned} \quad (55)$$

This proves that  $M^o \in H^2(0, T; L^2(\nu_0))$  and (54) holds. To obtain (55), we have used that the function  $f : [0, T] \times [0, 1] \rightarrow L^2(\nu_0)$  defined by  $f(t, s) = \nabla_{L^2}\bar{\mathcal{W}}(M_t^o + sM_t)$  is continuous and so, it is bounded on the compact set  $[0, T] \times [0, 1]$ . Hence,

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \int_0^T \frac{\bar{\mathcal{W}}(M_t^o + \varepsilon M_t) - \bar{\mathcal{W}}(M_t^o)}{\varepsilon} dt &= \lim_{\varepsilon \rightarrow 0} \int_0^T dt \int_0^1 \langle \nabla_{L^2}\bar{\mathcal{W}}(M_t^o + \varepsilon s M_t), M_t \rangle_{\nu_0} ds \\ &= \int_0^T \langle \nabla_{L^2}\bar{\mathcal{W}}(M_t^o), M_t \rangle_{\nu_0} dt. \end{aligned}$$

Suppose now that the assumptions in (iv) hold. Let  $\bar{M} \in P_T(\cdot, m)$  be a solution for (54). Then for any  $M \in H^1(0, T; L^2(\nu_0))$  satisfying  $M_T \equiv 0$ , we have

$$\mathbf{B}_T(\bar{M}, M) + \langle \nabla_{L^2} \bar{U}_0(\bar{M}_0), M_0 \rangle_{\nu_0} - \int_0^T \langle \nabla_{L^2} \bar{W}(\bar{M}_\tau), M_\tau \rangle_{\nu_0} d\tau = 0. \quad (56)$$

Now for any  $N \in P_T(\cdot, m)$ , by setting  $M := N - \bar{M}$  in (56) and using remark 4.7 we conclude that  $\mathbf{B}_T(N) > \mathbf{B}_T(\bar{M})$  unless  $N = \bar{M}$ . In particular, we have proved that  $\mathbf{B}_T(N) \geq \mathbf{B}_T(\bar{M})$  for all  $N \in P_T(\cdot, m)$  which means  $\bar{M}$  is a minimizer of  $\mathbf{B}_T$  over  $P_T(\cdot, m)$ . QED.

**Corollary 4.10.** *Suppose as in theorem 4.9 that (H1), (H2), (40) hold,  $\bar{U}_0$  is bounded below and  $\partial \bar{U}_0(m)$  is nonempty for every  $m$ . Then for any  $\mu \in \mathcal{M}$ , we have  $\mathbf{A}_T$  admits a minimizer  $\sigma^o$  over  $P_T(\cdot, \mu)$ . If in addition (H3) holds,  $-\bar{W}$  is  $l$ -convex,  $\partial(-\bar{W})(m)$  is nonempty for every  $m$  and  $T\lambda^- + 4(a^+ + l^-)T^2/\pi^2 < 1$ , then  $\sigma^o$  is unique.*

**Proof:** This is a direct consequence of remark 4.1 and theorem 4.9. Indeed, suppose  $\mu \in \mathcal{M}$  and let  $m \in L^2(\nu_0)$  be the monotone nondecreasing map satisfying  $m_{\#}\nu_0 = \mu$ . It then follows from remark 4.1 that  $M^o$  minimizes  $\mathbf{B}_T$  over  $P_T(\cdot, m)$  if and only if  $\mathcal{I}(M^o)$  minimizes  $\mathbf{A}_T$  over  $P_T(\cdot, \mu)$ . Here,  $\mathcal{I}$  is the map defined in remark 4.1. We now use theorem 4.9 to conclude the proof of the corollary. QED.

**Remark 4.11.** *Suppose as in theorem 4.9 that (H1), (H2), (H3), (40) hold and  $\bar{U}_0$  is bounded below. Assume also  $-\bar{W}$  is  $l$ -convex,  $\partial(-\bar{W})(m)$  and  $\partial \bar{U}_0(m)$  are nonempty for every  $m$ , and  $T\lambda^- + 4(a^+ + l^-)T^2/\pi^2 < 1$ . For each  $m \in L^2(\nu_0)$  we define by  $m^*$  the monotone rearrangement of  $m$  characterized by the facts that  $m^*$  is monotone nondecreasing and  $m^*_{\#}\nu_0 = m_{\#}\nu_0$ . Suppose  $E \subset L^2(\nu_0)$  is a convex set such that for any  $m \in E$  we have  $m^* \in E$ . Then one can readily check that the arguments used in the proof of theorem 4.9 can be applied to prove that if  $m \in E$  then  $\mathbf{B}_T$  admits a unique minimizer over the set of  $E_T$  of all  $M \in P_T(\cdot, m)$  such that  $M_t \in E$  for all  $t \in [0, T]$ .*

The following corollary ensures the continuity of the value function  $U$ .

**Corollary 4.12.** *Suppose (H1), (H2) and (40) hold, and  $\mathcal{W}$  is bounded below on bounded subsets of  $\mathcal{M}$ . Assume also  $\bar{U}_0$  is bounded below and the subdifferential  $\partial \bar{U}_0(m)$  is nonempty for every  $m \in L^2(\nu_0)$ . Then the value function  $U$  is continuous on  $Q_T$ .*

**Proof:** We know that  $U : Q_T \rightarrow \mathbb{R}$ . Since  $U$  is upper semicontinuous by lemma 3.7, it suffices to prove that  $U$  is lower semicontinuous. Suppose  $(\mu_n, t_n) \rightarrow (\mu, t)$  in  $Q_T$  and let  $\varepsilon > 0$ . Let  $m_n$  and  $m$  be the monotone nondecreasing maps satisfying  $m_{n\#}\nu_0 = \mu_n$  and  $m_{\#}\nu_0 = \mu$ . For each  $n$ , we can find a path  $\sigma^n \in P_{t_n}(\cdot, \mu_n)$  such that

$$U(\mu_n, t_n) + \varepsilon \geq \int_0^{t_n} L(\sigma_\tau^n, \mathbf{v}_\tau^n) d\tau + U_0(\sigma^n(0)).$$

Let  $\bar{\sigma}^n(\tau) := \sigma^n(\frac{t_n\tau}{t})$ , for  $\tau \in [0, t]$ . Also, for each  $n$  and  $\tau$ , let  $M_\tau^n$  be the monotone nondecreasing map that pushes  $\nu_0$  forward to  $\bar{\sigma}_\tau^n$ . Then we have  $M^n \in P_t(\cdot, m_n)$  and the above inequality can be rewritten as

$$U(\mu_n, t_n) + \varepsilon \geq \frac{t_n}{t} \int_0^t L(\bar{\sigma}_\tau^n, \bar{\mathbf{v}}_\tau^n) d\tau + U_0(\bar{\sigma}^n(0)) = \frac{t_n}{t} \int_0^t \bar{L}(M_\tau^n, (M^n)'_\tau) d\tau + \bar{U}_0(M_0^n). \quad (57)$$

Since  $\limsup_{n \rightarrow \infty} U(\mu_n, t_n) \leq U(\mu, t) < +\infty$ , by arguing similarly as in the proof of theorem 4.9 (i) we see that there exist a subsequence of  $\{M^n\}_n$  and  $M \in P_t(\cdot, m)$  such that for each  $\tau \in [0, t]$ ,  $\{M_\tau^n\}_n$  converges weakly in  $L^2(\nu_0)$  to  $M_\tau$ . Furthermore,

$$\liminf_{n \rightarrow \infty} \left\{ \frac{t_n}{t} \int_0^t \bar{L}(M_\tau^n, (M^n)'_\tau) d\tau + \bar{U}_0(M_0^n) \right\} \geq \int_0^t \bar{L}(M_\tau, M'_\tau) d\tau + \bar{U}_0(M_0). \quad (58)$$

We note that to obtain the analogous results of lemma 4.4, corollary 4.3 and corollary 4.6, we need to use the following Poincaré-Wirtinger inequality (see proposition 2.1): for any  $N \in P_t(\cdot, m)$ ,

$$\left( \int_0^t \|M_\tau^n - N_\tau\|_{\nu_0}^2 d\tau \right)^{1/2} \leq \frac{2t}{\pi} \left( \int_0^t \|(M^n)'_\tau - N'_\tau\|_{\nu_0}^2 d\tau \right)^{1/2} + \sqrt{t} \|m_n - m\|_{\nu_0}.$$

Now let  $\sigma_\tau := M_{\tau\#}\nu_0$  for  $\tau \in [0, t]$ . Then  $\sigma \in P_t(\cdot, \mu)$  and hence by using (57), (58) and the fact  $M_\tau$  is monotone nondecreasing for every  $\tau \in [0, t]$ , we obtain

$$\liminf_{n \rightarrow \infty} U(\mu_n, t_n) + \varepsilon \geq \int_0^t \bar{L}(M_\tau, M'_\tau) d\tau + \bar{U}_0(M_0) = \int_0^t L(\sigma_\tau, \mathbf{v}_\tau) d\tau + U_0(\sigma(0)) \geq U(\mu, t).$$

This concludes the proof. QED.

**Theorem 4.13.** *Suppose (H1), (H2), (40) hold and  $\bar{W}$  is bounded below on bounded subsets of  $L^2(\nu_0)$ . Assume  $\bar{W}$  is differentiable and  $\nabla_{L^2}\bar{W}$  is continuous. Suppose also that  $\bar{U}_0$  is differentiable and bounded below. Then the value function  $U$  defined in (14) is a viscosity solution for (16).*

**Proof:** We have from corollary 4.12 that  $U$  is continuous on  $Q_T$ . Also by theorem 3.9 (i),  $U$  is a viscosity subsolution for (16). It remains to show that it is a viscosity supersolution. For that fix  $(\mu, t) \in Q_T$  and let  $(\zeta, \theta) \in \partial U(\mu, t)$ . Corollary 4.10 ensures existence of a minimizer  $\sigma$  in (14). For  $0 \leq s \leq t$  we denote by  $M_s$  the unique monotone nondecreasing map pushing  $\nu_0$  forward to  $\sigma_s$ . It follows that  $M$  minimize  $\mathbf{B}_t$  over  $P_t(\cdot, M_t)$  and hence theorem 4.9 (iii) yields  $M \in H^2(0, t; L^2(\nu_0))$ . Fix  $0 < \varepsilon < t$ . By the semigroup property of lemma 2.4 and remark 4.1,

$$U(\mu, t) - U(\sigma_{t-\varepsilon}, t - \varepsilon) = \int_{t-\varepsilon}^t L(\sigma_\tau, v_\tau) d\tau = \int_{t-\varepsilon}^t \bar{L}(M_\tau, M'_\tau) d\tau. \quad (59)$$

Let  $\gamma_\varepsilon := (M_t \times M_{t-\varepsilon})\# \nu_0 \in \Gamma_o(\sigma_t, \sigma_{t-\varepsilon})$ . We use (59) to obtain

$$\begin{aligned} \frac{1}{\varepsilon} \int_{t-\varepsilon}^t \bar{L}(M_\tau, M'_\tau) d\tau &= \frac{U(\mu, t) - U(\sigma_{t-\varepsilon}, t - \varepsilon)}{\varepsilon} \\ &\leq \int_{\mathbb{R} \times \mathbb{R}} \langle \zeta(x), \frac{x-y}{\varepsilon} \rangle d\gamma_\varepsilon(x, y) + \theta + \frac{1}{\varepsilon} o(W_2(\sigma_{t-\varepsilon}, \sigma_t)) + \frac{o(\varepsilon)}{\varepsilon} \\ &= \int_X \langle \zeta(M_t x), \frac{M_t x - M_{t-\varepsilon} x}{\varepsilon} \rangle dx + \theta + \frac{o(W_2(\sigma_{t-\varepsilon}, \sigma_t))}{W_2(\sigma_{t-\varepsilon}, \sigma_t)} \frac{\|M_t - M_{t-\varepsilon}\|_{\nu_0}}{\varepsilon} + \frac{o(\varepsilon)}{\varepsilon}. \end{aligned} \quad (60)$$

Letting  $\varepsilon$  tend to  $0^+$  in (60) and using the fact  $M \in H^2(0, t; L^2(\nu_0))$ , we conclude that

$$-\theta \leq \int_X \langle \zeta(M_t x), M'_t x \rangle dx - \bar{L}(M_t, M'_t) \leq \bar{H}(M_t, \zeta \circ M_t) = H(\mu, \zeta).$$

Therefore, the second inequality in (12) holds and we conclude the proof of the theorem. QED.

We end this section by showing that the value function  $U$  enjoys a semiconcavity property. We first give the definition of semiconcave functions on  $\mathcal{P}_2(\mathbb{R})$ .

**Definition 4.14.** *We say that a function  $v : \mathcal{P}_2(\mathbb{R}) \rightarrow \mathbb{R}$  is semiconcave if there exists a constant  $C$  such that for every  $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R})$ , we have*

$$(1 - \alpha)v(\mu_1) + \alpha v(\mu_2) - v(((1 - \alpha)T_1 + \alpha T_2)_\# \nu_0) \leq \alpha(1 - \alpha)CW_2^2(\mu_1, \mu_2) \quad \forall \alpha \in [0, 1],$$

where  $T_1$  and  $T_2$  are the monotone nondecreasing maps satisfying  $T_1\#\nu_0 = \mu_1$  and  $T_2\#\nu_0 = \mu_2$ .

In the following result,  $U_0$  is only required to be bounded below.

**Proposition 4.15.** *Suppose  $\mathcal{W}$  satisfies (40) and  $\bar{W}$  is  $\kappa$ -convex on  $L^2(\nu_0)$ . Suppose also  $U_0$  is bounded below. Then the value function  $U(\cdot, t)$  is semiconcave for any  $t > 0$  satisfying  $4at^2 < \pi^2$ . More precisely, for every  $\mu_1, \mu_2 \in \mathcal{M}$  and  $\alpha \in [0, 1]$ , we have*

$$(1 - \alpha)U(\mu_1, t) + \alpha U(\mu_2, t) - U(((1 - \alpha)T_1 + \alpha T_2)_\# \nu_0, t) \leq \frac{\alpha(1 - \alpha)}{2} \left( \frac{1}{t} - \frac{(a + \kappa)t}{3} \right) W_2^2(\mu_1, \mu_2).$$

Here  $T_1$  and  $T_2$  are the monotone nondecreasing maps satisfying  $T_1\#\nu_0 = \mu_1$  and  $T_2\#\nu_0 = \mu_2$ .

**Proof:** Notice that due to (40),  $4at^2 < \pi^2$  and  $U_0$  is bounded below, we obtain from the Poincaré-Wirtinger inequality that  $U(\mu, t) > -\infty$  for all  $\mu \in \mathcal{P}_2(\mathbb{R})$ . Therefore, if we let  $\varepsilon > 0$  then there exists  $\sigma \in P_t(\cdot, ((1 - \alpha)T_1 + \alpha T_2)_\# \nu_0)$  such that

$$U(((1 - \alpha)T_1 + \alpha T_2)_\# \nu_0, t) \geq \int_0^t L(\sigma_s, \dot{\sigma}_s) ds + U_0(\sigma(0)) - \varepsilon. \quad (61)$$

For each  $s \in [0, t]$ , let  $M_s$  be the monotone nondecreasing map pushing  $\nu_0$  to  $\sigma_s$ . As  $\sigma_t = ((1 - \alpha)T_1 + \alpha T_2)_\# \nu_0$ , we have  $M_t = (1 - \alpha)T_1 + \alpha T_2$ . For  $s \in [0, t]$ , define

$$\sigma_s^1 := \left( M_s + \frac{s}{t} \alpha (T_1 - T_2) \right)_\# \nu_0 \quad \text{and} \quad \sigma_s^2 := \left( M_s - \frac{s}{t} (1 - \alpha) (T_1 - T_2) \right)_\# \nu_0.$$

Then we have  $\sigma_t^1 = T_{1\#}\nu_0$  and  $\sigma_t^2 = T_{2\#}\nu_0$ . This together with (61) yields

$$\begin{aligned}
& (1-\alpha)U(T_{1\#}\nu_0, t) + \alpha U(T_{2\#}\nu_0, t) - U(((1-\alpha)T_1 + \alpha T_2)\#\nu_0, t) \\
& \leq (1-\alpha) \int_0^t L(\sigma_s^1, \dot{\sigma}_s^1) ds + \alpha \int_0^t L(\sigma_s^2, \dot{\sigma}_s^2) ds - \int_0^t L(\sigma_s, \dot{\sigma}_s) ds + \varepsilon \\
& = \frac{1}{2} \int_0^t \left[ (1-\alpha) \int_{\mathbb{R}} |\dot{\sigma}_s^1(x)|^2 d\sigma_s^1(x) + \alpha \int_{\mathbb{R}} |\dot{\sigma}_s^2(x)|^2 d\sigma_s^2(x) - \int_{\mathbb{R}} |\dot{\sigma}_s(x)|^2 d\sigma_s(x) \right] ds \\
& \quad - \frac{a}{2} \int_0^t \left[ (1-\alpha)M_2(\sigma_s^1) + \alpha M_2(\sigma_s^2) - M_2(\sigma_s) \right] ds \\
& \quad - \int_0^t \left[ (1-\alpha)\mathcal{W}(\sigma_s^1) + \alpha\mathcal{W}(\sigma_s^2) - \mathcal{W}(\sigma_s) \right] ds + \varepsilon \\
& \leq \frac{1}{2} \int_0^t \int_{\mathbb{R}} \left[ (1-\alpha)|\dot{M}_s + \frac{\alpha}{t}(T_1 - T_2)|^2 + \alpha|\dot{M}_s - \frac{(1-\alpha)}{t}(T_1 - T_2)|^2 - |\dot{M}_s|^2 \right] d\nu_0(x) ds \\
& \quad - \frac{a}{2} \int_0^t \int_{\mathbb{R}} \left[ (1-\alpha)|M_s + \frac{s}{t}\alpha(T_1 - T_2)|^2 + \alpha|M_s - \frac{s}{t}(1-\alpha)(T_1 - T_2)|^2 - |M_s|^2 \right] d\nu_0(x) ds \\
& \quad - \int_0^t \left[ (1-\alpha)\bar{\mathcal{W}}(M_s + \frac{s}{t}\alpha(T_1 - T_2)) + \alpha\bar{\mathcal{W}}(M_s - \frac{s}{t}(1-\alpha)(T_1 - T_2)) - \bar{\mathcal{W}}(M_s) \right] ds + \varepsilon.
\end{aligned}$$

It follows from this and the assumption  $\bar{\mathcal{W}}$  is  $\kappa$ -convex on  $L^2(\nu_0)$  that

$$\begin{aligned}
& (1-\alpha)U(T_{1\#}\nu_0, t) + \alpha U(T_{2\#}\nu_0, t) - U(((1-\alpha)T_1 + \alpha T_2)\#\nu_0, t) \\
& \leq \frac{\alpha(1-\alpha)}{2} \left[ \frac{1}{t} - \frac{at}{3} - \frac{\kappa t}{3} \right] \|T_1 - T_2\|_{\nu_0}^2 + \varepsilon.
\end{aligned}$$

Since  $\|T_1 - T_2\|_{\nu_0} = W_2(\mu_1, \mu_2)$  and  $\varepsilon > 0$  is arbitrary, we conclude the proof. QED.

## 5 An example of non smooth potential: the 1-d Euler Poisson system

In this section, we assume that we are given a function  $\bar{U}_0$  which is differentiable such that

(H4)  $\{\nabla_{L^2}\bar{U}_0(m_n)\}_{n=1}^{\infty}$  converges to  $\nabla_{L^2}\bar{U}_0(m)$  whenever  $\{m_n\}_{n=1}^{\infty}$  converges to  $m$  in  $L^2(\nu_0)$ .

In the previous section we dealt with smooth potentials. From the point of view of Fluid Mechanics, it is interesting to extend the study to include cases of non smooth potentials. We consider the potential

$$\mathcal{W}(\mu) = -1/4 \int_{\mathbb{R} \times \mathbb{R}} |x - \bar{x}| d\mu(x) d\mu(\bar{x}). \tag{62}$$

Up to a multiplicative constant, this is the Green function for the Poisson equation. That potential was considered in [9] and [10] in the study of some versions of the Euler-Poisson system.

**Remark 5.1.** *Writing*

$$W_2^2(\mu, \nu_0) = \int_{\mathbb{R}} y^2 d\mu(y) + 2\mathcal{W}(\mu) + \frac{1}{12}$$

and knowing that  $W_2^2(\cdot, \nu_0)$  is differentiable on  $\mathcal{M}$ , (cf. e.g. [3]), we obtain that  $\mathcal{W}$  is differentiable everywhere on  $\mathcal{M}$ . However,  $\bar{\mathcal{W}}(m) := \mathcal{W}(m_{\#}\nu_0)$  is not differentiable everywhere on  $L^2(\nu_0)$ .

In the sequel, we work with the Lagrangian

$$L(\mu, w) = \frac{1}{2}\|w\|_{\mu}^2 - \frac{1}{2}\int_{\mathbb{R}} y^2 d\mu(y) + \frac{1}{4}\int_{\mathbb{R}\times\mathbb{R}} |y - \bar{y}| d\mu(y) d\mu(\bar{y}).$$

## 5.1 Existence of viscosity solutions

**Theorem 5.2.** *Suppose both (H1) (with  $a = 1$ ) and (H4) hold. Assume also  $\bar{U}_0$  is differentiable and bounded below. Let  $m \in L^2(\nu_0)$  be monotone nondecreasing. Then, there is a unique minimizer of  $\mathbf{B}_T$  over  $P_T(\cdot, m)$ . Also,  $M$  is a minimizer of  $\mathbf{B}_T$  over  $P_T(\cdot, m)$  if and only if it satisfies the Euler-Lagrange equation*

$$M_t'' x + M_t x = \frac{1}{2}\int_X \omega_t(x, \bar{x}) dx, \quad M_T = m, \quad M_0' = \nabla_{L^2} \bar{U}_0(M_0). \quad (63)$$

for some  $\omega : X \times X \times (0, T) \rightarrow \mathbb{R}$  such that  $\omega_t(x, \bar{x}) = -\omega_t(\bar{x}, x)$  and  $\omega_t(x, \bar{x}) \in \partial|\cdot|(M_t x - M_t \bar{x})$ .

**Proof:** Let us define

$$\bar{\mathcal{C}}_T(m_0, m_1) := \inf \left\{ \int_0^T \bar{L}(M_t, M_t') dt : M \in P_T(m_0, m_1) \right\}. \quad (64)$$

By theorem 4.9,  $\mathbf{B}_T$  admits a unique minimizer  $M^o$  over  $P_T(\cdot, m)$ . If we set  $m_0 = M_0^o$  and  $m_1 = M_T^o$  then  $M^o$  is the unique minimizer in (64). By proposition 7 [9], the differential equation in (63) holds.

For  $\varepsilon > 0$  let  $W^\varepsilon \in C^2(\mathbb{R})$  be an even, convex function such that

$$0 = W^\varepsilon(0) = \min W^\varepsilon, \quad \|W^\varepsilon - W\|_\infty \leq \varepsilon, \quad \|(W^\varepsilon)'\|_\infty \leq 1 + \varepsilon, \quad \|(W^\varepsilon)''\|_\infty < \infty.$$

Here,  $W(\omega) := |\omega|$ . We introduce the Lagrangian

$$\bar{L}^\varepsilon(m, \mathbf{n}) = \frac{1}{2}(\|\mathbf{n}\|_{\nu_0}^2 - \|m\|_{\nu_0}^2) - \bar{\mathcal{W}}^\varepsilon(m), \quad \bar{\mathcal{W}}^\varepsilon(m) := -\frac{1}{4}\int_{X \times X} W^\varepsilon(m(x) - m(\bar{x})) dx d\bar{x}.$$

The associated Hamiltonian is

$$\bar{H}^\varepsilon(m, \mathbf{n}) = \frac{1}{2}(\|\mathbf{n}\|_{\nu_0}^2 + \|m\|_{\nu_0}^2) + \bar{\mathcal{W}}^\varepsilon(m).$$

We introduce the augmented action as

$$\bar{\mathbf{B}}_T^\varepsilon(M) = \int_0^T \bar{L}^\varepsilon(M_\tau, M_\tau') d\tau + U_o(M_0).$$

Note that

$$|\bar{\mathbf{B}}_T^\varepsilon - \bar{\mathbf{B}}_T| \leq \frac{\varepsilon}{4}. \quad (65)$$

By theorem 4.9,  $\mathbf{B}_T^\varepsilon$  admits a unique minimizer over  $P_T(\cdot, m)$ , namely  $M^\varepsilon \in H^2(0, T; L^2(\nu_0))$  satisfying

$$(M^\varepsilon)_t'' + M^\varepsilon = -\nabla_{L^2} \mathcal{W}^\varepsilon(M_t^\varepsilon) \quad \text{and} \quad (M^\varepsilon)'_0 = \nabla_{L^2} \bar{U}_0(M_0^\varepsilon). \quad (66)$$

By remark 4.1 and the fact that  $m$  is monotone nondecreasing and  $M^\varepsilon$  is the unique minimizer, we have that  $M_t^\varepsilon$  is monotone nondecreasing for each  $t \in [0, T]$ . By (65)

$$\bar{\mathbf{B}}_T(M^\varepsilon) \leq \frac{\varepsilon}{4} + \bar{\mathbf{B}}_T^\varepsilon(M^\varepsilon) \leq \frac{\varepsilon}{4} + \bar{\mathbf{B}}_T^\varepsilon(M^o) \leq \frac{\varepsilon}{2} + \bar{\mathbf{B}}_T(M^o). \quad (67)$$

This, together with lemma 4.4 and again (65) gives that

$$\sup_{\varepsilon \in (0,1)} \int_0^T \|(M^\varepsilon)'_t\|_{\nu_0}^2 dt < \infty. \quad (68)$$

Since  $\|(W^\varepsilon)'\|_\infty \leq 1 + \varepsilon$ , we have that

$$\|\nabla_{L^2} \bar{\mathcal{W}}^\varepsilon\|_{\nu_0} \leq (1 + \varepsilon)/2 \leq 1. \quad (69)$$

We use (68) and the fact that

$$\|M_t^\varepsilon\|_{\nu_0}^2 \leq 2\|m\|_{\nu_0}^2 + 2T \int_0^T \|(M^{\varepsilon_k})'_\tau\|_{\nu_0}^2 d\tau$$

to conclude that

$$\sup_{\varepsilon, t} \{\|M_t^\varepsilon\|_{\nu_0}^2 : t \in [0, T], \varepsilon \in (0, 1)\} < \infty. \quad (70)$$

We combine (66) and (68) to conclude that

$$\sup_{\varepsilon, t} \{\|(M^\varepsilon)_t''\|_{\nu_0}^2 : t \in [0, T], \varepsilon \in (0, 1)\} < \infty. \quad (71)$$

We use (68), (70) and (71) to obtain a sequence  $\varepsilon_k \subset (0, 1)$  converging to 0 as  $k$  tends to  $\infty$  and  $M^* \in H^2(0, T; L^2(\nu_0))$  such that respectively  $\{(M^{\varepsilon_k})''\}_{k \in \mathbf{N}}$ ,  $\{(M^{\varepsilon_k})'\}_{k \in \mathbf{N}}$  and  $\{(M^{\varepsilon_k})\}_{k \in \mathbf{N}}$  converge respectively weakly in  $L^2(X \times (0, T))$  to  $(M^*)''$ ,  $(M^*)'$  and  $M^*$ . Furthermore, as in the proof of theorem 4.9 that there exists a subsequence of  $\varepsilon_k \subset (0, 1)$  we don't relabel such that for each  $t \in [0, T]$ ,  $\{(M^{\varepsilon_k})'_t\}_{k \in \mathbf{N}}$  converges weakly to  $M_t^*$  in  $L^2(\nu_0)$  and  $\{M_t^{\varepsilon_k}\}_{k \in \mathbf{N}}$  converges strongly to  $M_t^*$  in  $L^p(\nu_0)$  for  $1 \leq p < 2$ . Naturally,  $M^* \in P_T(\cdot, m)$ .

We use (67) and then apply corollary 4.3 and corollary 4.6 to conclude that

$$\mathbf{B}_T(M^o) \geq \limsup_{k \rightarrow \infty} \mathbf{B}_T(M^{\varepsilon_k}) \geq \liminf_{k \rightarrow \infty} \mathbf{B}_T(M^{\varepsilon_k}) \geq \mathbf{B}_T(M^*). \quad (72)$$

Since  $M_t^* \in P_T(\cdot, m)$  and  $M^o$  is the unique minimizer of  $\mathbf{B}_T$  over  $P_T(\cdot, m)$ , (67) yields  $M^* = M^o$  and so,

$$\mathbf{B}_T(M^o) = \lim_{k \rightarrow \infty} \mathbf{B}_T(M^{\varepsilon_k}) = \mathbf{B}_T(M^*). \quad (73)$$

By remark 4.7, there exists a constant  $c > 0$  such that

$$\begin{aligned} \mathbf{B}_T^\varepsilon(M^\varepsilon) &\geq \mathbf{B}_T^\varepsilon(M^o) + \mathcal{B}_T(M^o, M^\varepsilon - M^o) + \langle \nabla_{L^2} \bar{U}_0(M_0), M_0^\varepsilon - M_0^o \rangle_{\nu_0} \\ &\quad - \int_0^T \langle \nabla_{L^2} \bar{\mathcal{W}}^\varepsilon(M_\tau^o), M_\tau^\varepsilon - M_\tau^o \rangle_{\nu_0} d\tau + c \int_0^T \|(M^\varepsilon)'_\tau - (M^o)'_\tau\|_{\nu_0}^2 d\tau. \end{aligned} \quad (74)$$

By (65), (69) and (74)

$$\begin{aligned} \frac{\varepsilon}{2} + \mathbf{B}_T(M^\varepsilon) &\geq -\frac{\varepsilon}{2} + \mathbf{B}_T(M^o) + \mathcal{B}_T(M^o, M^\varepsilon - M^o) + \langle \nabla_{L^2} \bar{U}_0(M_0), M_0^\varepsilon - M_0^o \rangle_{\nu_0} \\ &\quad - \int_0^T \|M_\tau^\varepsilon - M_\tau^o\|_{L^1(\nu_0)} d\tau + c \int_0^T \|(M^\varepsilon)'_\tau - (M^o)'_\tau\|_{\nu_0}^2 d\tau. \end{aligned} \quad (75)$$

We use (73), (75), the fact that  $\{M^{\varepsilon_k}\}_{k \in \mathbf{N}}$  and  $\{(M^{\varepsilon_k})'\}_{k \in \mathbf{N}}$  converge respectively weakly to  $M$  and  $M'$  in  $L^2(X \times (0, T))$  to conclude that

$$\mathbf{B}_T(M^o) \geq \mathbf{B}_T(M^o) + c \lim_{k \rightarrow \infty} \int_0^T \|(M^{\varepsilon_k})'_\tau - (M^o)'_\tau\|_{\nu_0}^2 d\tau.$$

Thus,

$$\lim_{k \rightarrow \infty} \int_0^T \|(M^{\varepsilon_k})'_\tau - (M^o)'_\tau\|_{\nu_0}^2 d\tau = 0$$

and so,

$$\limsup_{k \rightarrow \infty} \|M_0^{\varepsilon_k} - M_0^o\|_{\nu_0}^2 \leq \lim_{k \rightarrow \infty} T \int_0^T \|(M^{\varepsilon_k})'_\tau - (M^o)'_\tau\|_{\nu_0}^2 d\tau = 0. \quad (76)$$

To obtain the previous inequality we have used the fact that  $M_T^{\varepsilon_k} - M_T^o = 0$  and a simple computation. We now use (76), the second identity in (66), the fact that  $\{(M^{\varepsilon_k})'_0\}_{k \in \mathbf{N}}$  converges weakly to  $(M^*)'_0$  in  $L^2(\nu_0)$  and the assumption (H4) to conclude that  $(M^o)'_0 = \nabla_{L^2} \bar{U}_0(M_0^o)$ .

It remains to prove that every solution of the Euler-Lagrange equations (63) is a minimizer of  $\mathbf{B}_T$  over  $P_T(\cdot, m)$ . We skip the proof as it is similar to that of theorem 4.9. We also refer the reader to proposition 7 [9]. QED.

**Corollary 5.3.** *Suppose both (H1) (with  $a = 1$ ) and (H4) hold. Assume also  $\bar{U}_0$  is differentiable and bounded below. Then the value function  $U$  defined in (1) is a viscosity solution for the Hamilton-Jacobi equation (2).*

**Proof:** By corollary 4.12 and theorem 3.9 (i), we only need to show that  $U$  is a viscosity supersolution. To prove this, we argue as in the proof of theorem 4.13 and use theorem 5.2 instead of theorem 4.9 (iii). QED.

## 5.2 Connection with the finite dimensional case

The following result connects our infinite-dimensional objects with the corresponding finite-dimensional, classical restrictions.

**Proposition 5.4.** *Suppose that  $U_0 : \mathcal{P}_2(\mathbb{R}) \rightarrow \mathbb{R}$  satisfies (H1) (with  $a = 1$ ) and (H4). Assume also  $\bar{U}_0$  is differentiable and bounded below. Then the viscosity solution  $U$  given by (1) for our infinite-dimensional Hamilton-Jacobi with initial data  $U_0$  has the following property: if  $\mu_{\mathbf{x}}^n = (1/n) \sum_{i=1}^n \delta_{x_i}$  for some real  $x_1, x_2, \dots, x_n \in \mathbb{R}$ , then the optimal trajectory for  $U(\mu, t)$  consists of averages of  $n$  point masses as well.*

Before we prove this, let us introduce the finite-dimensional version of (2), namely

$$\begin{cases} \partial_t U^n + H^n(\mathbf{x}, \nabla U^n) = 0 & \text{in } Q_T^n := \mathbb{R}^n \times (0, T), \\ U^n(\cdot, 0) = U_0^n & \text{on } \mathbb{R}^n, \end{cases} \quad (77)$$

where

$$\begin{aligned} H^n(\mathbf{x}, \mathbf{p}) &:= H(\mu_{\mathbf{x}}^n, \mathbf{p}) = \frac{|\mathbf{p}|^2}{2n} + \frac{1}{2} W_2^2(\mu_{\mathbf{x}}^n, \nu_0) \\ &= \frac{|\mathbf{p}|^2}{2n} + \frac{|\mathbf{x}|^2}{2n} - \frac{1}{4n^2} \sum_{i,j=1}^n |x_i - x_j| + \frac{1}{24} \text{ for } \mathbf{x} \in \mathbb{R}^n, \mathbf{p} \in \mathbb{R}^n. \end{aligned}$$

Here we endow  $\mathbb{R}^n$  with the inner product  $\langle \cdot, \cdot \rangle / n$  (inherited from  $L^2(\mu_{\mathbf{x}}^n)$ ) and we use the identification  $\mathbb{R}^n \simeq \mathcal{T}_{\mu_{\mathbf{x}}^n} \mathcal{P}_2(\mathbb{R})$ . The associated Lagrangian becomes

$$\begin{aligned} L^n(\mathbf{x}, \mathbf{v}) &:= L(\mu_{\mathbf{x}}^n, \mathbf{w}) = \frac{|\mathbf{v}|^2}{2n} - \frac{1}{2} W_2^2(\mu_{\mathbf{x}}^n, \nu_0) \\ &= \frac{|\mathbf{v}|^2}{2n} - \frac{|\mathbf{x}|^2}{2n} + \frac{1}{4n^2} \sum_{i,j=1}^n |x_i - x_j| - \frac{1}{24} \text{ for } \mathbf{x}, \mathbf{v} \in \mathbb{R}^n. \end{aligned}$$

Note that, since  $L^n$  is invariant to coordinate permutations, we have that

$$\int_0^t L^n(\sigma^\kappa(\tau), \dot{\sigma}^\kappa(\tau)) d\tau + U_0^n(\sigma^\kappa(0))$$

is independent of the permutation  $\kappa \in \Sigma_n$ , where  $\mathbf{x}^\kappa = (x_{\kappa(1)}, x_{\kappa(2)}, \dots, x_{\kappa(n)})$ . Due to  $U_0^n$  sharing this property, it follows that

$$U^n(\mathbf{x}, t) = U^n(\mathbf{x}^\kappa, t) \text{ for all } \mathbf{x} \in \mathbb{R}^n, \kappa \in \Sigma_n, t \in [0, T]. \quad (78)$$

Here  $U^n$  is the viscosity solution for (77) given by

$$U^n(\mathbf{x}, t) := \min_{\sigma \in AC^2(0,t;\mathbb{R}^n)} \left\{ \int_0^t L^n(\sigma(\tau), \dot{\sigma}(\tau)) d\tau + U_0^n(\sigma_0) : \sigma_t = \mathbf{x} \right\}. \quad (79)$$

The optimal trajectory for  $U^n(\mathbf{x}^\kappa, t)$  is obviously  $\sigma_0^\kappa$ , where  $\sigma_0$  is optimal for  $U^n(\mathbf{x}, t)$ . Note that, according to the proof of Proposition 5.4 (see below), the solution for (2) at  $(\mu_{\mathbf{x}}^n, t)$  is

$$U(\mu_{\mathbf{x}}^n, t) = \min_{\kappa \in \Sigma_n} U^n(\mathbf{x}^\kappa, t). \quad (80)$$

Along with (78), (80) implies the result stated next.

**Theorem 5.5.** *For any integer  $n \geq 1$  one has*

$$U(\mu_{\mathbf{x}}^n, t) = U^n(\mathbf{x}, t) \text{ for all } \mathbf{x} \in \mathbb{R}^n \text{ and all } t \in [0, T],$$

where  $U$  and  $U^n$  are the variational solutions for (2) and (77) respectively.

We shall next give a proof of Proposition 5.4.

**Proof:** (of Proposition 5.4) Let us denote, as before

$$\mathbf{A}_t(\sigma) := \mathcal{A}_t(\sigma) + U_0(\sigma_0), \quad \sigma \in AC^2(0, t; L^2(\nu_0))$$

and consider the associated  $\mathbf{B}_t(M)$ . W.l.o.g., we may assume  $x_1 \leq x_2 \leq \dots \leq x_n$ . As in theorem 5.2, one sees that  $\mathbf{B}_t$  has a unique minimizer  $\tilde{M}$  over  $AC^2(0, t; L_n)$  among the paths of prescribed terminal point  $\tilde{M}_t$ . Here,  $\tilde{M}_t$  is the monotone nondecreasing map such that  $\tilde{M}_t \# \nu_0 = \mu_{\mathbf{x}}^n$ . The symbol  $L_n$  denotes the set of all such maps which are piecewise constant on the uniform  $n$ -division of  $X$ . Since, in particular,  $\tilde{M}$  minimizes  $\mathbf{B}_t$  over the set of paths in  $AC^2(0, t; L_n)$  which have the same initial and terminal point as  $\tilde{M}$ , if we let  $r_i(\tau) := \tilde{M}_\tau(z)$  for  $z \in (c_{i-1}, c_i)$ ,  $i = 1, \dots, n$ , we deduce  $\mathbf{r}(t) = \mathbf{x}$  and  $r_1(\tau) \leq r_2(\tau) \leq \dots \leq r_n(\tau)$  for all  $\tau \in [0, t]$ . Here,  $c_i := -1/2 + i/n$ . Also,  $\mathbf{r} \in H^2(0, t; \mathbb{R}^n)$ . For  $\mathcal{L}^1$ -almost every  $\tau$  point of twice differentiability of  $\mathbf{r}$ ,

$$\tilde{M}_\tau'' x = \dot{r}_i(\tau) \text{ for all } x \in (c_{i-1}, c_i). \quad (81)$$

Consider  $M \in P_t(\cdot, \tilde{M}_t)$  a path such that  $M_\tau$  are monotone nondecreasing maps and let  $M_\tau \# \nu_0 = \mu_\tau$  for  $0 \leq \tau < t$ . The corresponding path of measures is denoted by  $\sigma$ , whereas the path given by  $\tilde{M}$  is denoted by  $\tilde{\sigma}$ . The convexity of  $e \rightarrow G(e) = |e|/2$  implies

$$\begin{aligned} 2\mathbf{A}_t(\sigma) + \frac{1}{12} &= \int_{X_t} (|M'|^2 - |M|^2) d\nu + \frac{1}{2} \int_{X_t^2} |M_\tau x - M_\tau \bar{x}| dx d\bar{x} d\tau + 2\bar{U}_0(M_0) \\ &\geq 2\mathbf{A}_t(\tilde{\sigma}) + \frac{1}{12} + \int_{X_t} (|M' - \tilde{M}'|^2 - |M - \tilde{M}|^2) d\nu + 2[\bar{U}_0(M_0) - \bar{U}_0(\tilde{M}_0)] \\ &\quad + 2 \sum_{i=1}^n \int_{c_{i-1}}^{c_i} d\nu_0 \int_0^t \left( (M'_\tau - \dot{r}_i(\tau)) \dot{r}_i(\tau) - (M_\tau - r_i(\tau)) r_i(\tau) \right) d\tau \end{aligned} \quad (82)$$

$$+ \frac{1}{2} \sum_{j \neq i}^n \int_{X_{ij}} dx d\bar{x} \int_0^t w_{ij}(\tau) \left( (M_\tau x - r_i(\tau)) - (M_\tau \bar{x} - r_j(\tau)) \right) d\tau. \quad (83)$$

where we have set  $X_{ij} = (c_{i-1}, c_i) \times (c_{j-1}, c_j)$  and  $w_{ij}$  are from [9], Theorem 5, equation (51). To be precise,  $w_{ij}(t) \in \partial \cdot | \cdot | (r_i(t) - r_j(t))$  such that  $w_{ji} = -w_{ij}$  and

$$\ddot{r}_i(t) + r_i(t) = \frac{1}{2n} \sum_{j \neq i} w_{ij}(t), \quad i = 1, \dots, n.$$

Due to the minimizing property of  $\tilde{M}$  over  $AC^2(0, t; L_n)$  with terminal endpoint fixed  $\tilde{M}_t$ , one can easily infer (as in the proof of Theorem 5.2) that  $\tilde{M}'_0 = \nabla_{L^2} \bar{U}_0(\tilde{M}_0)$ . Thus, we can write

$$\langle \nabla_{L^2} \bar{U}_0(\tilde{M}_0), M_0 - \tilde{M}_0 \rangle_{L^2(\nu_0)} = \sum_{i=1}^n \int_{c_{i-1}}^{c_i} \dot{r}_i(0) (M_0 - r_i(0)) d\nu_0. \quad (84)$$

Since  $\tilde{M}_t = M_t$ , we use

$$\int_{X_t} (|M' - \tilde{M}'|^2 - |M - \tilde{M}|^2) d\nu \geq \frac{1}{t} \left( 1 - \frac{4t^2}{\pi^2} \right) \|\tilde{M}_0 - M_0\|_{L^2(\nu_0)}^2$$

again, along with (H1), to deduce in light of (84)

$$\int_{X_t} (|M' - \tilde{M}'|^2 - |M - \tilde{M}|^2) d\nu + 2(\bar{U}_0(M_0) - \bar{U}_0(\tilde{M}_0)) \geq 2 \sum_{i=1}^n \int_{c_{i-1}}^{c_i} \dot{r}_i(0)(M_0 - r_i(0)) d\nu_0. \quad (85)$$

We can integrate by parts the expression in (82) and obtain explicit boundary terms that vanish when we combine the result with (83), (85). Finally, we exploit the fact that  $w_{ji} = -w_{ij}$  to obtain

$$\begin{aligned} 2\mathbf{A}_t(\sigma) + \frac{1}{12} &\geq 2\mathbf{A}_t(\tilde{\sigma}) + \frac{1}{12} \\ &\quad - 2 \sum_{i=1}^n \int_{c_{i-1}}^{c_i} d\nu_0 \int_0^t \left( \dot{r}_i(\tau) + r_i(\tau) - \frac{1}{2n} \sum_{j \neq i} w_{ij}(\tau) \right) (M_\tau - r_i(\tau)) d\tau \\ &= 2\mathbf{A}_t(\tilde{\sigma}) + \frac{1}{12}. \end{aligned}$$

QED.

## References

- [1] L. Ambrosio and A. Figalli. Geodesics in the space of measure-preserving maps and plans. *Preprint*
- [2] L. Ambrosio and W. Gangbo. Hamiltonian ODEs in the Wasserstein spaces of probability measures. *Comm. Pure Appl. Math.* **61**, no. 1, 18–53, 2008.
- [3] L. Ambrosio, N. Gigli and G. Savaré. Gradient flows in metric spaces and the Wasserstein spaces of probability measures. Lectures in Mathematics, ETH Zurich, Birkhäuser, 2005.
- [4] V. Barbu and G. Da Prato. Hamilton-Jacobi Equations in Hilbert Spaces. Pitman London, 1983.
- [5] V. Barbu and G. Da Prato. Hamilton-Jacobi Equations in Hilbert Spaces; Variational and Semigroup Approach. *Scuola Normale Superiore report*, Pisa 1984.
- [6] M. Crandall and P. Lions. Hamilton-Jacobi equations in infinite dimensions I. Uniqueness of viscosity solutions. *J. Funct. Anal.* **62**, 379–396, 1985.
- [7] M. Crandall and P. Lions. Hamilton-Jacobi equations in infinite dimensions II. Existence of viscosity solutions. *J. Funct. Anal.* **65**, 368–405, 1986.
- [8] M. Cullen, W. Gangbo and G. Pisante. Semigeostrophic equations discretized in reference and dual variables. *Arch. Ration. Mech. Anal.* **185**, no. 2, 341–363, 2007.
- [9] W. Gangbo, T. Nguyen and A. Tudorascu. Euler-Poisson systems as action minimizing paths in the Wasserstein space. *Arch. Rat. Mech. Anal.* (to appear), 2008.

- [10] T. Nguyen and A. Tudorascu. Euler/Euler-Poisson systems via adhesion dynamics and scalar conservation laws. *SIAM J. Math. Anal.* (to appear), 2008.
- [11] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations* **26**, no. 1–2, 101–174, 2001.
- [12] C. Villani. Topics in optimal transportation. *Graduate Studies in Mathematics* **58**, American Mathematical Society, 2003.