DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **34**, Number **4**, April **2014**

doi:10.3934/dcds.2014.34.1397

pp. **1397–1441**

OPTIMAL TRANSPORT AND LARGE NUMBER OF PARTICLES

Wilfrid Gangbo and Andrzej Święch

School of Mathematics
Georgia Institute of Technology
686 Cherry Street
Atlanta, GA 30332-0160, USA

Abstract. We present an approach for proving uniqueness of ODEs in the Wasserstein space. We give an overview of basic tools needed to deal with Hamiltonian ODE in the Wasserstein space and show various continuity results for value functions. We discuss a concept of viscosity solutions of Hamilton-Jacobi equations in metric spaces and in some cases relate it to viscosity solutions in the sense of differentials in the Wasserstein space.

1. **Introduction.** We consider infinite dimensional Hamiltonian systems in the Wasserstein space which arise in the study of limits of physical systems of indistinguishable particles in motion when the number of particles tends to infinity, and the associated Hamilton-Jacobi equations. Such systems appear in many interesting cases, for instance in the theory of Mean Field Games pioneered by J-M. Lasry and P-L. Lions [57, 58, 59, 60], which has become a fast growing area during the past few years [1, 2, 21, 45, 49, 50, 51, 52, 56]. The study of Hamilton–Jacobi equations in the Wasserstein space $\mathcal{P}_2(M)$ and in more general metric spaces is an important problem of its own. Here, $M = \mathbb{R}^D$ or $M = \mathbb{T}^D$ and $\mathcal{P}_2(M)$ is the set of Borel measures on M with finite second moments. The theory of Mean Field Games when $M = \mathbb{R}^d$, leads to the investigation of equation

$$\partial_t \mathcal{U}(t,x,\mu) + \frac{1}{2}|\nabla_x \mathcal{U}(t,x,\mu)|^2 + \mathcal{F}(\mu) - \int_{\mathbb{R}^d} \langle \nabla_x \mathcal{U}(t,q,\mu), \nabla_\mu \mathcal{U}(t,q,\mu)\rangle \mu(dq) = 0, \tag{1.1}$$

which is related to the so called mean-field equations in [60]. Here, the variables are $t > 0$, $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. The rigorous treatment of (1.1) is open to our knowledge. A model equation for us will be the Hamilton-Jacobi equation

$$\partial_t \mathcal{U}(t,\mu) + \mathcal{H}(\mu, \nabla_\mu \mathcal{U}(t,\mu)) = 0 \quad \text{on} \quad (0,T) \times \mathcal{P}_2(M), \tag{1.2}$$

where

$$\mathcal{H}(\mu,\xi) := \frac{1}{2}||\xi||^2_\mu + \mathcal{F}(\mu), \qquad (\mu,\xi) \in \mathcal{TP}_2(M). \tag{1.3}$$

Here, $\mathcal{TP}_2(M)$ is the union of the sets $\{\mu\} \times L^2(\mu)$ where $\mu \in \mathcal{P}_2(M)$ and $L^2(\mu)$ stands for the set of Borel maps $\xi : M \to \mathbb{R}^D$ such that $\int_M |\xi|^2 d\mu < \infty$. There is an embedding of $\mathcal{TP}_2(M)$ into $\mathcal{P}_2(M \times \mathbb{R}^D)$ given by $(\mu,\xi) \to (\mathbf{i}d \times \xi)_\# \mu$ and so, $\mathcal{TP}_2(M)$ can be viewed as a subspace of $\mathcal{P}_2(M \times \mathbb{R}^D)$. Here # is the push forward operator (cf. e.g. [7]).

2010 *Mathematics Subject Classification.* Primary: 35F21, 37K05, 49L25; Secondary: 76A99.
Key words and phrases. Optimal transport, conservative systems, infinite dimensional Hamiltonian systems, Hamilton-Jacobi equations.

Hamilton–Jacobi equations in the Wasserstein and related spaces also appear in the study of large deviations of empirical measures for stochastic particle systems, statistical mechanics, fluid mechanics, and many other areas [11, 12, 13, 14, 15, 16, 41, 32, 34, 35, 36]. In this article we give an overview of basic tools needed to deal with Hamiltonian ODE in the Wasserstein space, show various continuity results for value functions, and discuss viscosity solutions of Hamilton-Jacobi equations in the Wasserstein and metric spaces.

In Section 3, inspired by the work of Loeper [61, 62] and Yudovich [68], we present tools for proving uniqueness of solutions $\sigma \in AC_2(0,T;\mathcal{P}_2(M))$ of ordinary differential equations in the Wasserstein space; our study covers the case where $\sigma(t)$ may not be absolutely continuous with respect to the Lebesgue measure. Applying these tools for proving uniqueness of characteristics in Equation (1.2) remains however a challenge because of the lack of regularity of $\mu \to \nabla_\mu \mathcal{U}(t,\cdot)$. The result obtained in Theorem 5.2 (iv) would be, in finite dimension, equivalent to the fact that $\mathcal{U}(t,\cdot)$ is semiconvex and semiconcave and so its gradient is Lipschitz.

In Section 4, we study non–autonomous Hamiltonian equations for a one particle system and link them to systems of infinitely many particles. The idea there is that in order to study infinite dimensional ordinary differential equations of the form

$$\partial_t(\sigma \mathbf{v}) + \nabla \cdot (\sigma \mathbf{v} \otimes \mathbf{v}) = -\sigma \nabla_\mu \mathcal{F}(\sigma)$$

on $(0,T) \times \mathcal{P}_2(M)$, one needs to understand the one particle non–autonomous ordinary differential equations

$$\ddot{q} = -\sigma \nabla_\mu \mathcal{F}(\sigma(t))(q).$$

Making this statement rigorous requires proving some estimates which we establish in Section 5. For simplicity, in Sections 4 – 6 we keep our focus on Hamiltonians of the form

$$\mathcal{F}(\mu) = \int_M (V + W * \mu) d\mu. \tag{1.4}$$

The main result of Section 5 is Theorem 5.2 (iv) which states that the value function provided by the Hopf–Lax formula is differentiable along special paths (cf. also Remark 7 (i)).

In Section 6 we consider functions more general than those appearing in (1.4) and prove that the value function provided by the Hopf–Lax formula is Lipschitz. Most of the techniques used there mimic those used in the finite dimensional setting. The new ingredient is Lemma 8.3 which says that any 2–absolutely continuous curve in the Wasserstein space can be in some sense translated in any prescribed direction while its velocities are controlled. Moreover there is a difficulty which one encounters when trying to show that the value function is semi–concave. Given a curve $\sigma \in AC_2(0,T;\mathcal{P}(M))$, $\nu \in \mathcal{P}_2(M)$ and $t \in (0,T)$, one can consider the path σ^ν which coincides with σ on $[0,t]$ and extend it to the geodesic which connects σ_t to ν. In the Hilbert space setting, the analogue of the path σ^ν is used to prove that the value function given by the Hopf–Lax formula is λ–concave if $\mathcal{F}$ is λ–convex. Making that proof work in the Wasserstein setting is a harder task which we could complete only under some restrictive smoothness assumptions on the initial value function $\mathcal{U}_0$ (cf. Theorem 5.2). In a Hilbert space, one can translate any curve with its tangents in any given direction whereas we are lacking of ways of performing the analogue operation in the Wasserstein space. This substantially complicates the proof of the fact that the value function is differentiable along characteristics unless one imposes that the initial value function is of class C^3 in a sense to be specified.

In a Hilbert space, there is a natural Poisson structure and the study of Hamilton-Jacobi equations has a long history (see next paragraphs). The characteristics exist and are unique when the initial function is smooth (cf. the recent study [47]). In the Wasserstein space there are major difficulties one has to face. Indeed, one can show the existence of a Hamiltonian flow $\Psi : [0,\infty) \times \mathcal{P}_2(M \times \mathbb{R}^D) \to \mathcal{P}_2(M \times \mathbb{R}^D)$ (cf. [6]) for the Hamiltonian

$$\breve{\mathcal{H}}(\gamma) := \frac{1}{2} \int_{M \times \mathbb{R}^D} |p|^2 \gamma(dq, dp) + \mathcal{F}(\mu), \qquad \mu = \pi_{M\#}\gamma \tag{1.5}$$

which extends $\mathcal{H}$ from $\mathcal{TP}_2(M)$ to $\mathcal{P}_2(M \times \mathbb{R}^D)$. However, if we choose $(\mu, \xi) \in \mathcal{TP}_2(M)$ and identify it with $\gamma = (\mathbf{i}d \times \xi)_{\#}\mu \in \mathcal{P}_2(M \times \mathbb{R}^D)$, $\Psi(t, \gamma)$ may escape $\mathcal{TP}_2(M)$ and so, there is no known Hamiltonian flow for $\mathcal{H}$. An existence or uniqueness theory remains open in $\mathcal{TP}_2(M)$. We refer the reader to [20] where a Hamiltonian flow for $\mathcal{H}$ and $M = \mathbb{R}$ was proposed via some selection criteria.

The terminology of Hamiltonian systems in the Wasserstein space which we use throughout this manuscript is justified by the fact that there exists a Poisson structure on $\mathcal{P}(\mathbb{T}^d \times \mathbb{R}^d)$ as well as on $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ (cf. [38] and [55], [63]). In some cases, for instance when $M = \mathbb{T}^1$, one can exploit the well–developed theory of Hamiltonian systems on the Hilbert space $L^2(M)$ to study Hamiltonian systems on the Wasserstein space $\mathcal{P}_2(M)$ (cf. e.g. [40, 41, 42, 46]). A direct approach on the infinite dimensional torus $\mathcal{P}(\mathbb{T}^d)$ appeared for the first time in [43].

Hamilton-Jacobi-Bellman equations in infinite dimensional spaces have a long history. Earlier results in Hilbert spaces can be found in [9]. The theory of viscosity solutions in Hilbert spaces started with papers of M. Crandall and P. L. Lions [24]-[30]. Other notions of viscosity solution were also introduced (see e.g. [54, 65, 66]) and there is by now a huge literature on the subject, including a theory of second order equations in Hilbert spaces. As regards equations in spaces of probability measures and more general metric spaces, several approaches have been introduced. In [34] a very general theory of viscosity solutions in metric spaces was proposed. The main motivation of [34] was to apply it to equations coming from large deviation problems for particle systems. More concrete problems in the Wasserstein space have been studied in [33, 35, 36]. The definitions of viscosity solutions there were based on the use of special test functions related to the problems that reduced the state space to measures absolutely continuous with respect to Lebesgue measure and guaranteed coercivity estimates. P. L. Lions in [60] proposed an approach in which an equation in the Wasserstein space is pulled to an equation in a Hilbert space L^2 where measures are replaced by random variables in L^2 having given laws. The definition of viscosity solution for equations in the Wasserstein space given in [41] is based on the notions of sub- and superdifferentials of functions in the Wasserstein space. In [44] a notion of metric viscosity solution was introduced. It looks at the behavior of functions along curves and it is substantially based on the sub- and super-optimality inequalities of dynamic programming. Another paper that studies a special Hamilton-Jacobi-Isaacs equation in the space of measures associated with a zero-sum differential game with imperfect information is [22]. Finally we mention the papers [8, 48, 64] where it was proved that Hopf-Lax formulas satisfy certain differential inequalities and equalities involving local slopes (see (7.2)) for the associated Hamilton-Jacobi equation.

In this paper we discuss the notion of viscosity solution in the Wasserstein space using the notion of sub- and superdifferentials and a notion of viscosity solution in a geodesic metric space. In Section 6 we show that in the Wasserstein space, the Hopf-Lax formula provides a subsolution in the viscosity sense in terms of the subdifferential of a value function. The Hopf-Lax formula is not known to provide a supersolution in the viscosity sense in terms of the superdifferential of the value function except in some simple cases [53]. In Section 7 we discuss a notion of viscosity solution in a geodesic metric space for Hamilton-Jacobi equations whose gradient variable only depends on its "length". We prove a general comparison result, show that a viscosity solution can be obtained by Perron's method, and prove in a model case that the function given by the Hopf-Lax formula is a viscosity solution.

This manuscript relies on the material developed by Ambrosio, Gigli and Savaré [7] which contains the classical theory the mass transport is built upon. We also refer the reader to [67] for an alternative presentation of the mass transport theory.

2. **Preliminaries.**

Throughout this manuscript M is either $\mathbb{R}^d$, $\mathbb{T}^d$, $\mathbb{R}^d \times \mathbb{R}^d$ or $\mathbb{T}^d \times \mathbb{R}^d$ and $\mathbf{i}d : M \to M$ is the identity map on M. If $x, y \in M$ then $|x - y|$ is the natural distance between x and y. We write $M \subset \mathbb{R}^D$ having in mind that either $D = d$ or $D = 2d$.

Recall that $\mathcal{P}_2(M)$, the set of probability measures on M with bounded second moments, is endowed with the Wasserstein metric W_2, which makes it a geodesic space. Given $\mu, \nu \in \mathcal{P}_2(M)$, we denote by $\Gamma(\mu, \nu)$ the set of Borel measures γ on $M \times M$ which have μ as the first marginal and ν as the second marginal. We denote by $\Gamma_o(\mu, \nu)$ the subset of $\Gamma(\mu, \nu)$ which consists of measures γ such that

$$W_2^2(\mu, \nu) = \int_{M \times M} |x - y|^2 \gamma(dx, dy).$$

When M is a bounded set then $\mathcal{P}_2(M)$ coincides with $\mathcal{P}(M)$, the set of Borel probability measures on M. If $\mu \in \mathcal{P}_2(M)$ and $\xi : M \to \mathbb{R}^D$ is a Borel vector field such that $||\xi||_\mu^2 := \int_M |\xi|^2 \mu(dq) < \infty$, we write $\xi \in L^2(\mu)$. We denote by $T_\mu \mathcal{P}_2(M)$ the closure of $\nabla C_c^\infty(M)$ in $L^2(\mu)$, and denote by $T\mathcal{P}_2(M)$ the set of (μ, ξ) such that $\mu \in \mathcal{P}_2(M)$ and $\xi \in T_\mu \mathcal{P}_2(M)$. If n is a positive integer, $\mathcal{P}^n(M)$ is the set of discrete measures of the form

$$\mu^{\mathbf{x}} := \frac{1}{n} \sum_{i=1}^n \delta_{x^i}, \quad \text{where} \quad x^1, \cdots, x^n \in M, \quad \mathbf{x} := (x^1, \cdots, x^n).$$

Given a metric space $\mathbb{S}$ and time dependent function $f : [0, T] \to \mathbb{S}$, throughout this manuscript, we write f_t in place of $f(t)$. For instance if $X : [0, T] \times M \to \mathbb{R}^D$, we write $X_t(q)$ instead of $X(t, q)$. If $\sigma \in AC_2(0, T, \mathcal{P}_2(M))$ we write σ_t instead of $\sigma(t)$.

Theorem 2.1, stated below, is a fundamental theorem of the Monge–Kantorovich mass transport theory which was first due to Brenier [19] and was later refined by Gangbo–McCann [39].

Theorem 2.1. *Assume $M = \mathbb{R}^d$ or $M = \mathbb{T}^d$, $\mu, \nu \in \mathcal{P}_2(M)$ and μ vanishes on $(d-1)$–rectifiable sets. Then there exists a unique $\gamma \in \Gamma_o(\mu, \nu)$. Furthermore, there exists a Borel map $T : M \to M$ such that $\gamma = (\mathbf{i}d \times T)_\# \mu$ and so, $T_\# \mu = \nu$.*

The following stability result on optimal couplings can be found in Proposition 7.1.3 [7].

Theorem 2.2. *Assume* $\{\mu^n\}_n, \{\nu_n\}_n \subset \mathcal{P}_2(M)$ *converge narrowly to* μ, ν *respectively and* $\gamma_n \in \Gamma_o(\mu_n, \nu_n)$. *Then,* $\{\gamma_n\}_n$ *is narrowly relatively compact in* $\mathcal{P}_2(M \times M)$ *and any narrow limit point belongs to* $\Gamma_o(\mu, \nu)$.

3. Uniqueness of ODEs in the Wasserstein space.

3.1. Properties of curves in the Wasserstein space.

Definition 3.1. Let $(\mathbb{S}, \mathrm{dist})$ be a metric space. A curve $t \in (a, b) \mapsto \sigma_t \in \mathbb{S}$ is *2–absolutely continuous* if there exists $\beta \in L^2(a, b)$ such that $\mathrm{dist}(\sigma_t, \sigma_s) \leq \int_s^t \beta(\tau)d\tau$ for all $a < s < t < b$. We then write $\sigma \in AC_2(a, b; \mathbb{S})$. For such curves the limit $|\sigma'|(t) := \lim_{s \to t} \mathrm{dist}(\sigma_t, \sigma_s)/|t-s|$ exists for $\mathcal{L}^1$–almost every $t \in (a, b)$. We call this limit the *metric derivative* of σ at t. It satisfies $|\sigma'| \leq \beta$ $\mathcal{L}^1$–almost everywhere (cf. e.g. [7]).

Remark 1. (i) If $\sigma \in AC_2(a, b; \mathbb{S})$, since $|\sigma'| \in L^2(a, b)$ and $\mathrm{dist}(\sigma_s, \sigma_t) \leq \int_s^t |\sigma'|(\tau)d\tau$ for $a < s < t < b$, we can apply Hölder's inequality to conclude that $\mathrm{dist}^2(\sigma_s, \sigma_t) \leq c|t-s|$ where $c = \int_a^b |\sigma'|^2(\tau)d\tau$.

(ii) It follows from (i) that σ is continuous and so, since $[0, T]$ is a compact set, so is $\{\sigma_t |\ t \in [a, b]\}$, the range of σ. In particular the range of σ is a bounded set and if $\mathbf{s} \in \mathbb{S}$, by the triangle inequality $\mathrm{dist}(\sigma_s, \mathbf{s}) \leq \sqrt{c|s-a|} + \mathrm{dist}(\sigma_a, \mathbf{s})$.

The proof of the following proposition can be found in [7] (cf. [43] when $M = \mathbb{T}^d$).

Proposition 1. *If* $\sigma \in AC_2(a, b; M)$ *then there exists a Borel map* $\mathbf{v} : (a, b) \times M \to \mathbb{R}^D$ *such that* $\mathbf{v}_t \in L^2(\sigma_t)$ *for* $\mathcal{L}^1$*–almost every* $t \in (a, b)$, $t \to ||\mathbf{v}_t||_{\sigma_t}$ *belongs to* $L^2(a, b)$ *and*

$$\partial_t \sigma + \nabla \cdot (\sigma \mathbf{v}) = 0$$

in the sense of distributions: for all $\phi \in C_c^\infty((a, b) \times M)$,

$$\int_a^b \int_M \Big(\partial_t \phi + \langle \nabla \phi, \mathbf{v} \rangle\Big)\, d\sigma_t\, dt = 0. \tag{3.1}$$

We refer to $\mathbf{v}$ *as a velocity for* σ.

Furthermore, one can choose $\mathbf{v}$ *such that* $\mathbf{v}_t \in T_{\sigma_t}\mathcal{P}_2(M)$ *and* $||\mathbf{v}_t||_{\sigma_t} = |\sigma'|(t)$ *for* $\mathcal{L}^1$*–almost every* $t \in (a, b)$. *In that case, for* $\mathcal{L}^1$*–almost every* $t \in (a, b)$, $\mathbf{v}_t$ *is uniquely determined. We denote this velocity* $\dot{\sigma}$ *and refer to it as the velocity of minimal norm, since if* $\mathbf{w}$ *is any other velocity for* σ *then* $||\dot{\sigma}_t||_{\sigma_t} \leq ||\mathbf{w}_t||_{\sigma_t}$ *for* $\mathcal{L}^1$*–almost every* $t \in (a, b)$.

Assuming $\mathbf{v}$ *is the velocity of minimal norm for* σ *then for* $\mathcal{L}^1$*–almost every* $t \in (a, b)$

$$\lim_{h \to 0} \frac{W_2\big(\sigma_{t+h}, (\mathbf{i}d + h\mathbf{v}_t)_\# \sigma_t\big)}{|h|} = 0. \tag{3.2}$$

Remark 2. Suppose $\{\sigma^n\}_n \subset AC_2(a, b; M)$. By definition of $|(\sigma^n)'|$

$$\sup_{n \in \mathbb{N}} \int_a^b |(\sigma^n)'|^2(t)dt < \infty$$

if and only if there are velocities $\mathbf{v}^n : (a, b) \times M \to \mathbb{R}^D$ for σ^n such that

$$\sup_{n \in \mathbb{N}} \int_a^b ||\mathbf{v}^n||^2_{\sigma_t^n} dt < \infty.$$

The following proposition is a consequence of Propositions 3 and 4 of [40].

Proposition 2. *Suppose $\nu \in \mathcal{P}_2(M)$, $\{\sigma^n\}_n \subset AC_2(a,b;M)$ and*

$$\sup_{n\in\mathbb{N}} W_2(\sigma_0^n,\nu), \quad \sup_{n\in\mathbb{N}} \int_a^b |(\sigma^n)'|^2(t)dt < \infty.$$

Then there exist $\sigma \in AC_2(a,b;M)$ and an increasing sequence of integers $\{n_k\}_k$ such that for all $t \in [a,b]$, $\{\sigma_t^{n_k}\}_k$ converges narrowly to σ_t. Furthermore, we have

$$\liminf_{k\to\infty} \int_a^b |(\sigma^{n_k})'|^2(t)dt \geq \int_a^b |\sigma'|^2(t)dt. \tag{3.3}$$

Lemma 3.2. *(i) Let $\sigma^i \in AC_2(a,b;\mathcal{P}_2(M))$ and let $\mathbf{v}^i$ be velocities for σ^i $(i=1,2)$. Set*

$$g(t) = W_2(\sigma_t^1,\sigma_t^2).$$

Then $g \in W^{1,2}(a,b)$ and the distributional derivative of $1/2g^2$ satisfies almost everywhere

$$\frac{1}{2}(g^2)'(t) \leq \int_{M\times M} \langle \mathbf{v}_t^1(x) - \mathbf{v}_t^2(y), x-y\rangle \gamma_t(dx,dy) \tag{3.4}$$

for any $\gamma_t \in \Gamma_o(\sigma_t^1,\sigma_t^2)$.

Proof. let $\bar{\mathbf{v}}^i$ be velocities of minimal norm for σ^i $(i=1,2)$. Since

$$||\bar{\mathbf{v}}_t^1||_{\sigma_t^1} + ||\bar{\mathbf{v}}_t^2||_{\sigma_t^2} \leq ||\mathbf{v}_t^1||_{\sigma_t^1} + ||\mathbf{v}_t^2||_{\sigma_t^2} =: \beta(t),$$

if $a \leq t_1 \leq t_2 \leq b$, by the triangle inequality and Remark 1 (i)

$$g(t_1) \leq W_2(\sigma_{t_1}^1,\sigma_{t_2}^1) + g(t_2) + W_2(\sigma_{t_2}^2,\sigma_{t_1}^2) \leq g(t_2) + \int_{t_1}^{t_2} \beta(t)dt.$$

Interchanging the role of t_1 and t_2 we conclude that $|g(t_2)-g(t_1)| \leq \int_{t_1}^{t_2}\beta(t)dt$. This proves that $g \in W^{1,2}(a,b)$. Hence, $1/2g^2$ is $W^{1,1}(a,b)$, its pointwise derivative exists and coincides almost everywhere with its distributional derivative.

Recall that the set $\mathcal{N}$ of $t \in (a,b)$ such that Equation (3.2) fails to hold for either $(\sigma^1,\bar{\mathbf{v}}^1)$ or $(\sigma^2,\bar{\mathbf{v}}^2)$ is of null measure. Let $t \in (a,b)\setminus\mathcal{N}$. For $|h|$ small enough, by the triangle inequality

$$g(t+h) \leq W_2(\sigma_{t+h}^1, (\mathbf{i}d + h\bar{\mathbf{v}}_t^1)_\#\sigma_t^1) + \bar{g}(h) + W_2((\mathbf{i}d+h\bar{\mathbf{v}}_t^2)_\#\sigma_t^2,\sigma_{t+h}^2) = \bar{g}(h) + o(h),$$

where

$$\bar{g}(h) = W_2((\mathbf{i}d + h\bar{\mathbf{v}}_t^1)_\#\sigma_t^1, (\mathbf{i}d+h\bar{\mathbf{v}}_t^2)_\#\sigma_t^2).$$

Hence,

$$g^2(t+h) \leq \bar{g}^2(h) + o(h). \tag{3.5}$$

Let $\gamma \in \Gamma_o(\sigma_t^1,\sigma_t^2)$ and define the Borel measure

$$\gamma^h = \Big((\mathbf{i}d + h\bar{\mathbf{v}}_t^1)\times(\mathbf{i}d+h\bar{\mathbf{v}}_t^2)\Big)_\#\gamma.$$

We have

$$\gamma^h \in \Gamma\big((\mathbf{i}d+h\bar{\mathbf{v}}_t^1)_\#\sigma_t^1, (\mathbf{i}d+h\bar{\mathbf{v}}_t^2)_\#\sigma_t^2\big)$$

and so,

$$\begin{aligned}\bar{g}^2(h) &\leq \int_{M\times M} |w-z|^2\gamma^h(dw,dz)\\ &= \int_{M\times M} |x+h\bar{\mathbf{v}}_t^1(x)-y-t\bar{\mathbf{v}}_t^2(y)|^2\gamma(dx,dy)\\ &= g^2(t)+2h\int_{M\times M}\langle x-y,\bar{\mathbf{v}}_t^1(x)-\bar{\mathbf{v}}_t^2(y)\rangle\gamma(dx,dy)+O(h^2). \quad (3.6)\end{aligned}$$

If $t\in(a,b)\setminus\mathcal{N}$ and g^2 is differentiable at t, Equations (3.5) and (3.6) imply

$$\frac{1}{2}(g^2)'(t)\leq\int_{M\times M}\langle\bar{\mathbf{v}}_t^1(x)-\bar{\mathbf{v}}_t^2(y),x-y\rangle\gamma_t(dx,dy). \quad (3.7)$$

Since $\nabla\cdot\big(\sigma_t^i(\mathbf{v}_t^i-\bar{\mathbf{v}}_t^i)\big)=0$ we combine Proposition 8.5.4 of [7] and (3.7) to conclude that (3.4) holds. □

3.2. **Uniqueness of solutions of ODEs driven by vector fields on $\mathcal{P}_2(M)$.** Let $\mathcal{O}$ be a subset of $\mathcal{P}_2(M)$ and let X be a vector field on $\mathcal{O}$ in the sense that for each $\mu\in\mathcal{O}$, $X(\mu)\in L^2(\mu)$. We assume that X is continuous in the sense that for each $Y\in C_b(M,\mathbb{R}^D)$

$$\mu\to\Lambda_\mu(Y)=\langle X(\mu),Y\rangle_\mu \quad \text{is contintuous.} \quad (3.8)$$

We further assume that

$$m:=\sup_{\mu\in\mathcal{O}}||X(\mu)||_{L^1(\mu)}<\infty. \quad (3.9)$$

Remark 3. Suppose Equation (3.8) holds and let $\sigma\in AC_2(0,T;\mathcal{P}_2(M))$.

(i) We have

$$m:=\sup_{t\in[0,T]}||X(\sigma_t)||_{L^1(\sigma_t)}<\infty.$$

(ii) If $Z\in C^r(0,T;C_b(M,\mathbb{R}^D))$ then $A:t\to A(t)=\Lambda_{\sigma_t}(Z(t,\cdot))$ is continuous.

Proof. (i) Consider the linear maps

$$\lambda_t:C_b(M,\mathbb{R}^D)\to\mathbb{R},\quad Y\to\Lambda_{\sigma_t}(Y).$$

By Remark 1 and Equation (3.8) $t\to\lambda_t(Y)$ is continuous as the composition of two continuous functions and so it is bounded on the compact set $[0,T]$. By the uniform boundedness principle

$$\infty>\sup_{t\in[0,T]}||\lambda_t||_{L(C_b(M,\mathbb{R}^D))}=\sup_{t\in[0,T]}||X(\sigma_t)||_{L^1(\sigma_t)}.$$

(ii) Assume that $Z\in C^r(0,T;C_b(M,\mathbb{R}^D))$. We will only show that A is continuous at every $t\in(0,T)$ since the proof of that case can easily be adapted to the cases $t=0$ or $t=T$. For $|h|$ small enough, we have

$$A(t+h)=\lambda_{t+h}(Z(t,\cdot))+\lambda_{t+h}(Z(t+h,\cdot)-Z(t,\cdot)), \quad (3.10)$$

$$\limsup_{h\to0}|\lambda_{t+h}(Z(t+h,\cdot)-Z(t,\cdot))|\leq\limsup_{h\to0}m||Z(t+h,\cdot)-Z(t,\cdot)||_{C_b(M,\mathbb{R}^D)}=0 \quad (3.11)$$

and

$$\lim_{h\to0}\lambda_{t+h}(Z(t,\cdot))=\lambda_t(Z(t,\cdot)). \quad (3.12)$$

We combine (3.10–3.12) to conclude the proof of (ii). □

Remark 4. Let ω_* be a real valued Borel function on $[0,\infty)$, let $\mu^1,\mu^2 \in O$ and let $\gamma \in \Gamma_o(\mu^1,\mu^2)$. Thanks to the Cauchy–Schwarz inequality, a sufficient condition to have

$$\int_{M\times M} \langle X(\mu^1)(x) - X(\mu^2)(y), x-y\rangle \gamma(dx,dy) \le \omega_*(W_2(\mu^1,\mu^2))W_2(\mu^1,\mu^2) \quad (3.13)$$

is

$$\int_{M\times M} |X(\mu^1)(x) - X(\mu^2)(y)|^2 \gamma(dx,dy) \le \omega_*^2(W_2(\mu^1,\mu^2)).$$

Theorem 3.3. *Let ω_* be a real valued Borel function on $[0,\infty)$ such that $\omega_*(y) > \omega_*(0) = 0$ for all $y \in [0,\infty)$ and for some $a \in (0,\infty)$*

$$\int_0^a \frac{dy}{\omega_*(y)} = \infty. \quad (3.14)$$

Assume that for every $\mu^1,\mu^2 \in O$ there exists $\gamma \in \Gamma_o(\mu^1,\mu^2)$ such that (3.13) holds. If $\sigma^i \in AC_2(0,T;\mathcal{O})$ and $X(\sigma^i_t) : (0,T)\times M \to \mathbb{R}^D$ are velocities for σ^i $(i=1,2)$, then $\sigma^1 = \sigma^2$ on $[0,T]$ provided that $\sigma^1(0) = \sigma^2(0)$.

Proof. We are to prove that $G \equiv 0$, where

$$G(t) = W_2^2(\sigma^1_t, \sigma^2_t).$$

By Lemma 3.2, $G \in W^{1,2}(a,b)$ and its distributional derivative, which coincides almost everywhere with its pointwise derivative satisfies

$$\begin{aligned} \dot G(t) &\le 2\int_{M\times M} \langle X(\sigma^1_t)(x) - X(\sigma^2_t)(y), x-y\rangle \gamma_t(dx,dy) \\ &\le 2\omega_*\Big(\sqrt{G(t)}\Big)\Big(\sqrt{G(t)}\Big) \\ &= \omega\big(G(t)\big), \end{aligned} \quad (3.15)$$

for any $\gamma_t \in \Gamma_o(\sigma^1_t,\sigma^2_t)$. Here we have set $\omega(y) = 2\sqrt{y}\omega_*(\sqrt{y})$ for $y \ge 0$. Observe that ω is a nonnegative Borel function on $[0,\infty)$ such that $\omega(y) > \omega(0) = 0$ for $y > 0$. Thanks to Lemma 8.1, Equations (3.14) and (3.15), together with the fact that $G(0) = 0$, imply $G \equiv 0$ on $[0,T]$. □

Remark 5. In fact one can reach the conclusions of Theorem 3.3 under weaker assumptions. More precisely, let ω_* be a real valued Borel function on $[0,\infty)$ such that $\omega_*(y) > \omega_*(0) = 0$ for all $y \in [0,\infty)$ and for some $a \in (0,\infty)$ Equation (3.14) holds. Suppose $\sigma^i \in AC_2(0,T;\mathcal{O})$ and $X(\sigma^i_t) : (0,T)\times M \to \mathbb{R}^D$ are velocities for σ^i $(i=1,2)$. Suppose there exists $\bar a > 0$ such that (3.13) holds for all $\mu^1,\mu^2 \in O$ satisfying

$$W_2(\mu^1,\sigma^1_0),\ W_2(\mu^2,\sigma^2_0) \le \bar a,$$

and some $\gamma \in \Gamma_o(\mu^1,\mu^2)$. If $\sigma^1_0 = \sigma^2_0$ then $\sigma^1 = \sigma^2$ on $[0,T]$.

3.3. **Examples.** We consider sets $\mathcal{O} \subset \mathcal{P}(M)$ when $M = \mathbb{T}^d$ or $M = \mathbb{R}^d$, and vector fields X defined on $\mathcal{O} \subset \mathcal{P}_2(M)$. We give examples of existence and uniqueness of solutions for some well–known initial value problem problems of the form

$$\sigma_0 = \mu, \qquad \partial_t\sigma + \nabla\cdot(\sigma X(\sigma)) = 0 \quad \text{in} \quad \mathcal{D}'\Big((0,T)\times\mathbb{T}^d\Big). \quad (3.16)$$

Example 3.4 (A trivial example)**.** When

$$X(\mu) = \int_M \nabla(V + W * \mu) d\mu.$$

and $\nabla V, \nabla W : M \to \mathbb{R}^d$ are Lipschitz functions then X satisfies (3.8), (3.9) and (3.13) with $\mathcal{O} = \mathcal{P}_2(M)$.

The following result is a well-kown one, due to V. Yudovich [68], but we present a proof based on Theorem 3.3.

Example 3.5 (2-d Euler incompressible systems in terms of vorticities)**.** Let m be a positive real number and let $\mathcal{O}$ be the set of probability measures μ on $\mathbb{T}^2$ such that $\mu = \varrho \mathcal{L}^2$, and $-m \leq \varrho - 1 \leq m$. Define ϕ^μ such that $\triangle \phi^\mu = \varrho - 1$ on $\mathbb{T}^d$ so that $\nabla \phi^\mu$ is uniquely determined. Set

$$X(\mu) = (\nabla \phi^\mu)^\perp.$$

For all $T > 0$ (3.16) admits a unique solution $t \to \sigma_t \in \mathcal{O}$.

Proof. Here, we only deal with the issue of uniqueness. For a constant C_m which depends only on m (cf. e.g. [62] Proposition 5.2)

$$||X(\mu_1) - X(\mu_2)||_{L^2(\mathbb{T}^2)} \leq C_m W_2(\mu_1, \mu_2). \tag{3.17}$$

Therefore, since $W_2 \leq 1/\sqrt{2}$, Equation (3.9) holds.

Set

$$H(y) = y \ln^2 y.$$

Note that

$$H \quad \text{increases on } [0, e^{-2}], \tag{3.18}$$

H is concave on $[0, e^{-1}]$, $H(0) = 0$ and so, if $a \in [0, e^{-1}]$ and $\lambda \in [0,1]$, then

$$\lambda H(a) = \lambda H(a) + (1-\lambda) H(0) \leq H\big(\lambda a + (1-\lambda)0\big) = H(\lambda a). \tag{3.19}$$

Choose $0 < \alpha < e^{-2}$. For instance, we can choose $\alpha = e^{-2}/2$. Let $\mu_1, \mu_2 \in \mathcal{O}$ and let $\gamma \in \Gamma_0(\mu_1, \mu_2)$. We have

$$\begin{aligned}\int_{\mathbb{T}^2 \times \mathbb{T}^2} |X(\mu_2)(q) - X(\mu_1)(q)|^2 \gamma(dq, dr) =& ||X(\mu_2) - X(\mu_1)||^2_{\mu_1} \\ \leq& m ||X(\mu_2) - X(\mu_1)||^2_{\mathbb{T}^2}.\end{aligned} \tag{3.20}$$

Thanks to Remark 5 we may assume without loss of generality that

$$\frac{1}{-\ln W_2(\mu_1, \mu_2)} < \alpha < e^{-2}. \tag{3.21}$$

Note that the diameter of $\mathbb{T}^2$ is $1/\sqrt{2}$ and so, $W_2(\mu_1, \mu_2) < 1$. Hence, $-\ln W_2(\mu_1, \mu_2) > 0$. Increasing the value of C_m if necessary, we have (cf. e.g. [17] chapter 8)

$$|X(\mu)(q) - X(\mu)(r)| \leq C_m |q - r| \ln \frac{1}{|q-r|^2}.$$

Thus,

$$\int_{\mathbb{T}^2 \times \mathbb{T}^2} |X(\mu_2)(q) - X(\mu_2)(r)|^2 \gamma(dq, dr) \leq C_m^2 \int_{\mathbb{T}^2 \times \mathbb{T}^2} H(|q-r|^2) \gamma(dq, dr). \tag{3.22}$$

Set

$$A := \{(q, r) \in \mathbb{T}^2 \times \mathbb{T}^2, \; | \; |q - r| < \alpha\}, \quad B := \mathbb{T}^2 \times \mathbb{T}^2 \setminus A.$$

We have

$$W_2^2(\mu_1,\mu_2) = \int_{\mathbb{T}^2\times\mathbb{T}^2} |q-r|^2\gamma(q,r) \geq \alpha^2\gamma(B)$$

and so,

$$\gamma(B) \leq \frac{1}{\alpha^2} W_2^2(\mu_1,\mu_2) \leq W_2^2(\mu_1,\mu_2)\ln^2 W_2^2(\mu_1,\mu_2). \tag{3.23}$$

Set

$$D_m := C_m^2 \sup_{l\in[0,1/\sqrt{2}]} l^2\ln^2 l^2.$$

Since the diameter of $\mathbb{T}^2$ is $1/\sqrt{2}$, we have $|q-r|\leq 1/\sqrt{2}$ and thus by (3.23),

$$C_m^2\int_B H(|q-r|^2)\gamma(dq,dr) \leq D_m\gamma(B) \leq D_m W_2^2(\mu_1,\mu_2)\ln^2 W_2^2(\mu_1,\mu_2). \tag{3.24}$$

Write

$$\int_A H(|q-r|^2)\gamma(dq,dr) = \gamma(A)\int_A H\big(|q-r|^2\big)\tilde{\gamma}(dq,dr),$$

where $\tilde{\gamma}=\gamma/\gamma(A)$ is a probability measure. Since H is a concave function, we apply Jensen's inequality to conclude that

$$\int_A H(|q-r|^2)\gamma(dq,dr) \leq \gamma(A) H\Big(\int_A |q-r|^2\tilde{\gamma}(dq,dr)\Big).$$

Thanks to (3.19) we conclude that

$$\int_A H(|q-r|^2)\gamma(dq,dr) \leq H\Big(\int_A |q-r|^2\gamma(dq,dr)\Big).$$

We use that H increases on $[0,e^{-2}]$ (cf. (3.18)) and that by Equation (3.21), $W_2 \leq e^{-e^2} \leq e^{-2}$ to conclude that

$$\int_A H(|q-r|^2)\gamma(dq,dr) \leq H\Big(\int_{\mathbb{T}^2\times\mathbb{T}^2} |q-r|^2\gamma(dq,dr)\Big) = H\Big(W_2^2(\mu_1,\mu_2)\Big). \tag{3.25}$$

By (3.24) and (3.25)

$$\begin{aligned}
&\int_{\mathbb{T}^2\times\mathbb{T}^2} |X(\mu_2)(q)-X(\mu_2)(r)|^2\gamma(dq,dr)\\
\leq\;& D_m W_2^2(\mu_1,\mu_2)\ln^2 W_2^2(\mu_1,\mu_2)\\
+\;& C_m^2 H\big(W_2^2(\mu_1,\mu_2)\big)\\
=\;& \big(\frac{D_m}{4}+C_m^2\Big) H\big(W_2^2(\mu_1,\mu_2)\big).
\end{aligned} \tag{3.26}$$

By (3.17), (3.20) and (3.26)

$$\begin{aligned}
\int_{\mathbb{T}^2\times\mathbb{T}^2} |X(\mu_2)(r)-X(\mu_1)(q)|^2\gamma(dq,dr) \;&\leq\; 2\big(\frac{D_m}{4}+C_m^2\big) H\big(W_2^2(\mu_1,\mu_2)\\
&+\; 2mC_m^2 W_2^2(\mu_1,\mu_2),
\end{aligned} \tag{3.27}$$

which yields (3.8). Thanks to (3.27), Remark 4 yields (3.13) if we set

$$\omega_*(t) = t\sqrt{2}\sqrt{mC_m^2+\Big(\frac{D_m}{4}+C_m^2\Big)\ln^2 t}.$$

By Remark 5, Equation (3.16) admits at most one solution $t\to\sigma_t\in\mathcal{O}$. $\square$

4. **One particle Hamiltonian systems.** Throughout this section we suppose that $\mathcal{U}_0 : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ and:

(U1) $\mathcal{U}_0$ is differentiable on $\mathcal{P}(\mathbb{T}^d)$ (cf. Definition 6.2) and

$$\sup_{\mu \in \mathcal{P}(\mathbb{T}^d)} ||\nabla_\mu \mathcal{U}_0(\mu)||_{C^2(\mathbb{T}^d)} < +\infty.$$

(U2) If $\{\mu_k\}_k \subset \mathcal{P}(\mathbb{T}^d)$ converges narrowly to μ, then $\{\nabla_\mu \mathcal{U}_0(\mu_k)\}$ converges uniformly to $\nabla_\mu \mathcal{U}_0(\mu)$ on $\mathbb{T}^d$.

Examples include

$$\mathcal{U}_0(\mu) = \int_{\mathbb{T}^d} (v_0 + w_0 * \mu) d\mu,$$

where $v_0, w_0 \in C^3(\mathbb{T}^d)$. Using the terminology of [43], (U1–U2) imply that $\mathcal{U}_0 \in C^1(\mathcal{P}(\mathbb{T}^d))$.

We assume that $V, W \in C^3(\mathbb{T}^d)$, W is even and $C_* > 0$ satisfies

$$||V||_{C^3(\mathbb{T}^d)},\ ||W||_{C^3(\mathbb{T}^d)},\ \sup_{\mu \in \mathcal{P}(\mathbb{T}^d)} ||\nabla_\mu \mathcal{U}_0(\mu)||_{C^2(\mathbb{T}^d)} \le C_*. \tag{4.1}$$

We will denote by $C_{V,W}$ a generic constant depending only on V and W. We denote by $\pi_{\mathbb{T}^d} : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{T}^d$ and $\pi_{\mathbb{R}^d} : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}^d$ the maps

$$\pi_{\mathbb{T}^d}(q,p) = q, \quad \pi_{\mathbb{R}^d}(q,p) = p \quad (q,p) \in \mathbb{T}^d \times \mathbb{R}^d.$$

Given $T > 0$ and $\sigma \in AC_2\big(0, T, \mathcal{P}(\mathbb{T}^d)\big)$ we define the one particle Hamiltonian

$$H^\sigma(t,q,p) = \frac{|p|^2}{2} + V(q) + W * \sigma_t(q)$$

and consider the Hamiltonian vector field

$$X_{H^\sigma}(t,q,p) = \big(p, -\nabla\big(V + W * \sigma_t\big)(q)\big).$$

We have

$$X_{H^\sigma} \in C\big([0,T]; C^2(\mathbb{T}^d \times \mathbb{R}^d)\big),$$

$$||\nabla_{(q,p)} X_{H^\sigma}||_\infty^2 \le d + \big(||\nabla V||_\infty + ||\nabla W||_\infty\big)^2 =: s_\infty \tag{4.2}$$

and

$$\sup_{\sigma, T}\Big\{||\nabla^2_{(q,p)} X_{H^\sigma}||_\infty \mid \sigma \in AC_2(0,T,\mathcal{P}(\mathbb{T}^d)), T \in (0,1]\Big\} < +\infty. \tag{4.3}$$

Consider the flow, which may be defined globally in time, as the solution of the initial value problem

$$\dot\Phi^\sigma_t = X_{H^\sigma}(t, \Phi^\sigma_t), \qquad \Phi^\sigma_0(q,p) = (q,p).$$

4.1. **Compactness properties of Hamiltonian flows.** Thanks to Equations (4.2–4.3) the standard theory of Hamiltonian systems ensures the existence of constants C_0 and C_1 independent of $T \in (0,1]$ such that

$$||\nabla \Phi^\sigma_t||_\infty,\ ||\nabla \dot\Phi^\sigma_t||_\infty,\ ||\nabla^2 \Phi^\sigma_t||_\infty \le C_0 \exp(C_1 t). \tag{4.4}$$

Furthermore,

$$|\dot\Phi^\sigma_t|^2 = |X_{H^\sigma}(\cdot, \Phi^\sigma_t)|^2 \le s_\infty. \tag{4.5}$$

The diameter of $\mathbb{T}^d$ being smaller than $\sqrt{d}/2$, integrating, we have

$$|\Phi^\sigma(t,q,p)|^2 \le \frac{d}{4} + |p|^2 + T s_\infty. \tag{4.6}$$

Lemma 4.1. *Suppose* $\mathbf{v}$ *is a velocity for* σ*, and*

$$c = \sup_{t\in[0,T]} ||\mathbf{v}_t||_{\sigma_t} < \infty. \tag{4.7}$$

Then

$$||\nabla_{(t,q,p)} X_{H^\sigma}||^2_\infty \leq s_\infty + c^2 ||\nabla^2 W||_\infty. \tag{4.8}$$

Proof. We have the distributional derivatives

$$X_{H^\sigma}(t,q,p) = \Big(p, -\nabla(V + W * \sigma_t)\Big)$$

and so, since $\mathbf{v}$ is a velocity for σ we have

$$\partial_t X_{H^\sigma}(t,q,p) = \Big(0, \int_{\mathbb{T}^d} \nabla^2 W(q-y)\mathbf{v}_t(y)\sigma_t(dy)\Big).$$

This, together with (4.2) yields (4.8). $\square$

Corollary 1. *Suppose* $\sigma, \sigma^n \in AC_2(0,T;\mathcal{P}(\mathbb{T}^d))$ *and* $\mathbf{v}^n$ *is a velocity for* σ^n *such that for each* $t \in [0,T]$ $\{\sigma^n_t\}_n$ *converges narrowly to* σ_t*. Suppose*

$$c = \sup_{t,n}\{||\mathbf{v}^n_t||_{\sigma^n_t} \mid n \in \mathbb{N}, t \in [0,T]\} < \infty.$$

Then

(i) $\{X_{H^{\sigma^n}}\}_n$ *converges uniformly to* X_{H^σ} *on* $[0,T]\times\mathbb{T}^d\times\mathbb{R}^d$.
(ii) $\{\Phi^{\sigma^n}\}_n$ *converges locally uniformly to* Φ^σ *on* $[0,T]\times\mathbb{T}^d\times\mathbb{R}^d$.

Proof. (i) Since for each $t\in[0,T]$, $\{\sigma^n_t\}_n$ converges narrowly to σ_t we obtain that $\{\nabla \bar{H}^{\sigma^n}\}_n$ converges pointwise to $\nabla\bar{H}^\sigma$. We apply Lemma 4.1 to $\{X_{\bar{H}^{\sigma^n}}\}_n$ and use the compact embedding of $W^{1,\infty}\big([0,T]\times\mathbb{T}^d\big)$ into $C\big([0,T]\times\mathbb{T}^d\big)$ to conclude that $\{X_{\bar{H}^{\sigma^n}}\}_n$ converges uniformly to $X_{\bar{H}^\sigma}$ on $[0,T]\times\mathbb{T}^d$. This proves (i).

(ii) By (4.4–4.6) if $K\subset\mathbb{R}^d$ then $\{\Phi^{\sigma^n}\}_n$ is precompact for the uniform convergence on $[0,T]\times\mathbb{T}^d\times K$. If a subsequence of $\{\Phi^{\sigma^n}\}_n$ converges uniformly on $[0,T]\times\mathbb{T}^d\times K$ to a function Φ, then by (i) $\Phi = \Phi^\sigma$. $\square$

4.2. **The Hamiltonian flows restricted to subsets of the cotangent bundle.**

For $\nu\in\mathcal{P}(\mathbb{T}^d)$, set

$$S^{\sigma,\nu}_t = \pi_{\mathbb{T}^d}\circ\Phi^\sigma_t(\cdot, \nabla_\mu \mathcal{U}_0(\nu)).$$

Since $\Phi^\sigma_t(q+l,p) = \Phi^\sigma_t(q,p)+l$ for all $q,p\in\mathbb{R}^d$ and all $l\in\mathbb{Z}^d$ we conclude that

$$S^{\sigma,\nu}_t(q+l) = S^{\sigma,\nu}_t(q) + l. \tag{4.9}$$

Hence we can view $S^{\sigma,\nu}_t$ as a map of $\mathbb{T}^d$ into $\mathbb{T}^d$.

Note that

$$\dot{S}^{\sigma,\nu}_t(q) = \nabla_\mu\mathcal{U}_0(\nu) + \int_0^t \nabla(V + W*\sigma_\tau)(S^{\sigma,\nu}_\tau(q))d\tau \tag{4.10}$$

and

$$\ddot{S}^{\sigma,\nu}_t = \nabla(V+W*\sigma_t)\circ S^{\sigma,\nu}_t. \tag{4.11}$$

We use (U1) and Equations (4.4), (4.11) to obtain

$$\sup_{\sigma,\nu,T}\Big\{||\nabla_{t,q}\dot{S}^{\sigma,\nu}||_\infty + ||\nabla^2_{qq}S^{\sigma,\nu}||_\infty \mid T\in(0,1], \sigma\in AC_2(0,T;\mathcal{P}(\mathbb{T}^d)), \nu\in\mathcal{P}(\mathbb{T}^d)\Big\} < \infty. \tag{4.12}$$

Remark 6. If $t\in[0,T]$, $\sigma\in AC_2(0,T;\mathcal{P}(\mathbb{T}^d))$ and $\nu\in\mathcal{P}(\mathbb{T}^d)$ then $S^{\sigma,\nu}_t:\mathbb{R}^d\to\mathbb{R}^d$ is surjective.

Proof. Fix arbitrary $y \in \mathbb{R}^d$. Choose $r > 1$ large enough so that $Ts_\infty < r$ and $2|y| < r$. If $q \in \partial B_{2r}$, the boundary of the closed ball in $\mathbb{R}^d$ of radius $2r$, centered at the origin, then by the Mean Value Theorem there exists $\theta \in (0,t)$ such that $S_t^{\sigma,\nu} q = q + t\dot{S}_\theta^{\sigma,\nu} q$. Hence by Equation (4.5)

$$|S_t^{\sigma,\nu} q| \geq |q| - ts_\infty > 2r - r = r > |y|.$$

This proves that $y \notin S_t^{\sigma,\nu}(\partial B_{2r})$ and so, $f(t) := \deg\,(S_t^{\sigma,\nu}, B_{2r}, y)$, the topological degree of $S_t^{\sigma,\nu}$ on $\bar{B}_{2R}$ at y, is a well–defined continuous function of t. Since $f(t)$ assumes only integer values and $f(0) = 1$, we conclude that $f(t) \equiv 1$. Thus, y belongs to the range of $S_t^{\sigma,\nu}$ (cf. e.g. [37]). □

The identity

$$S_t^{\sigma,\nu} = \mathbf{i}d + \int_0^t \dot{S}_\tau^{\sigma,\nu} d\tau$$

yields

$$\nabla S_t^{\sigma,\nu} = I_d + \int_0^t \nabla(\dot{S}_\tau^{\sigma,\nu}) d\tau. \tag{4.13}$$

We combine (4.12) and (4.13) to obtain a constant $C_{\mathcal{U}_0,V,W}$ such that

$$||\nabla S_t^{\sigma,\nu} - I_d|| \leq tC_{\mathcal{U}_0,V,W}. \tag{4.14}$$

for all $t \in [0,T]$. Hence there is $T_* \in (0,1]$ such that if $T \leq T_*$ then

$$\det \nabla S_t^{\sigma,\nu} \geq \frac{1}{2} \tag{4.15}$$

for all $\sigma \in AC_2(0,T;\mathcal{P}(\mathbb{T}^d))$, $\nu \in \mathcal{P}(\mathbb{T}^d)$ and all $t \in [0,T]$.

Theorem 4.2. *Suppose* $0 < T \leq T_*$. *Then*

(i) $S_t^{\sigma,\nu} : \mathbb{R}^d \to \mathbb{R}^d$ *is a bijection for* $\sigma \in AC_2(0,T;\mathcal{P}(\mathbb{T}^d))$, $\nu \in \mathcal{P}(\mathbb{T}^d)$ *and* $t \in [0,T]$.

(ii) Denote by $R_t^{\sigma,\nu}$ *the inverse of* $S_t^{\sigma,\nu}$. *We have*

$$\sup_{\sigma,\nu,T} \Big\{ ||\nabla_{(t,q)} R^{\sigma,\nu}||_\infty \,|\, \sigma \in AC_2(0,T;\mathcal{P}(\mathbb{T}^d)), \nu \in \mathcal{P}(\mathbb{T}^d), T \in (0,T_*] \Big\} < \infty.$$

Proof. (i) In light of Remark 6, it suffices to show that if $r_0 > 0$ and $y \in \bar{B}_{r_0}$, where $\bar{B}_{r_0}$ is the closed ball of radius r_0, then for all r large enough, the equation $y = S_t^{\sigma,\nu} q$ admits at most one solution in $\bar{B}_{2r}$. By (4.12), $\nabla_{(t,q)} S^{\sigma,\nu}$ is of class $W^{1,\infty}$ and so, $S^{\sigma,\nu}$ is of class C^1. Inequality (4.15), combined with the fact that $\deg\,(S_t^{\sigma,\nu}, \bar{B}_{2r}, y) = 1$ (cf. e.g. [37]), implies the existence of a unique $q \in \bar{B}_{2r}$ such that $y = S_t^{\sigma,\nu} q$. This concludes the proof of (i).

(ii) Since $S_t^{\sigma,\nu} : \mathbb{R}^d \to \mathbb{R}^d$ is an invertible function of class C^1 with a positive determinant, its inverse $R_t^{\sigma,\nu} : \mathbb{R}^d \to \mathbb{R}^d$ is of class C^1. We have

$$\nabla R_t^{\sigma,\nu} = \frac{(\mathrm{cof} \nabla S_t^{\sigma,\nu})^T}{\det \nabla S_t^{\sigma,\nu}} \big(R_t^{\sigma,\nu}\big), \quad \dot{R}_t^{\sigma,\nu} = -\nabla R_t^{\sigma,\nu} \dot{S}_t^{\sigma,\nu} \big(R_t^{\sigma,\nu}\big).$$

Hence, exploiting (4.10), (4.12), (4.14) and (4.15) one concludes the proof of (ii). □

We define

$$\mathbf{v}_t^{\sigma,\nu} y = \dot{S}_t^{\sigma,\nu} \Big(R_t^{\sigma,\nu} y\Big) \qquad t \in [0,T],\ y \in \mathbb{R}^d.$$

Using (4.5), (4.12) and Theorem 4.2 (ii) we conclude that

$$\sup_{\sigma,\nu,T} \Big\{ ||\nabla_{(t,q)} \mathbf{v}^{\sigma,\nu}||_\infty \,|\, \sigma \in AC_2(0,T;\mathcal{P}(\mathbb{T}^d)), \nu \in \mathcal{P}(\mathbb{T}^d), T \in (0,T_*] \Big\} < \infty. \tag{4.16}$$

Lemma 4.3. *Suppose* $\{\nu^n\}_n \subset \mathcal{P}(\mathbb{T}^d)$ *converges narrowly to* ν *and* $0 < T \le T_*$. *Suppose* $\{\sigma\} \cup \{\sigma^n\}_n \subset AC_2(0,T;\mathcal{P}(\mathbb{T}^d))$, $\{\sigma^n_t\}_n$ *converges narrowly to* σ_t *for every* $t \in [0,T]$ *and*

$$\sup_n \int_0^T |(\sigma^n)'|^2(t)dt < \infty.$$

Then $\{S^{\sigma^n,\nu^n}\}_n$ *converges uniformly to* $S^{\sigma,\nu}$ *on* $[0,T]\times\mathbb{T}^d$ *and* $\{\mathbf{v}^{\sigma^n,\nu^n}\}_n$ *converges uniformly to* $\mathbf{v}^{\sigma,\nu}$ *on* $[0,T]\times\mathbb{R}^d$.

Proof. Assumption (U1) ensures that the ranges of the $\nabla_\mu \mathcal{U}_0(\nu^n)$ are contained in a ball whose radius is independent of n. Next, (U2) and Corollary 1 ensure that $\{S^{\sigma^n,\nu^n}\}_n$ converges uniformly to $S^{\sigma,\nu}$ on $[0,T]\times\mathbb{T}^d$ and $\{\dot S^{\sigma^n,\nu^n}\}_n$ converges uniformly to $\dot S^{\sigma,\nu}$ on $[0,T]\times\mathbb{T}^d$. Thus, $\{S^{\sigma^n,\nu^n}\}_n$ converges uniformly to $S^{\sigma,\nu}$ on $[0,T]\times\mathbb{R}^d$ and $\{\dot S^{\sigma^n,\nu^n}\}_n$ converges uniformly to $\dot S^{\sigma,\nu}$ on $[0,T]\times\mathbb{R}^d$. We conclude that $\{R^{\sigma^n,\nu^n}\}_n$ converges uniformly to $R^{\sigma,\nu}$ on $[0,T]\times\mathbb{T}^d$ and so, it converges on $[0,T]\times\mathbb{R}^d$. These facts show that $\{\mathbf{v}^{\sigma^n,\nu^n}\}_n$ converges uniformly to $\mathbf{v}^{\sigma,\nu}$ on $[0,T]\times\mathbb{R}^d$. □

5. **Many particle Hamiltonian systems.**

As in Section 4 we assume throughout this section that

(U1) and (U2) hold, together with (4.1).

We define on $\mathcal{P}(\mathbb{T}^d)$

$$\mathcal{V}(\mu) = \int_{\mathbb{T}^d} V d\mu, \quad \mathcal{W}(\mu) = \frac{1}{2}\int_{\mathbb{T}^d} W * \mu d\mu.$$

We define the Lagrangian L and the Hamiltonian H

$$L(\mu,\mathbf{v}) = \frac{1}{2}||\mathbf{v}||^2_\mu - \mathcal{V}(\mu) - \mathcal{W}(\mu), \quad H(\mu,\mathbf{v}) = \frac{1}{2}||\mathbf{v}||^2_\mu + \mathcal{V}(\mu) + \mathcal{W}(\mu),$$

for $\mu \in \mathcal{P}(\mathbb{T}^d)$, $\mathbf{v} \in L^2(\mu)$ and we define the value function

$$\mathcal{U}(t,\nu) = \inf_{(\sigma,\mathbf{v})}\Big\{\int_0^t L(\sigma_s,\mathbf{v}_s)ds + \mathcal{U}_0(\sigma_0) \mid \sigma_t = \nu\Big\}. \tag{5.1}$$

We also define the costs

$$C_0^t(\mu,\nu) = \inf_{(\sigma,\mathbf{v})}\Big\{\int_0^t L(\sigma_s,\mathbf{v}_s)ds \mid \sigma_0 = \mu,\ \sigma_t = \nu\Big\}. \tag{5.2}$$

In Equations (5.1–5.2) the infimum is taken over the set of pairs $(\sigma,\mathbf{v})$ such that $\sigma \in AC_2(0,t;\mathcal{P}_2(\mathbb{T}^d))$ and $\mathbf{v}$ is a velocity for σ.

If $x^1,\cdots,x^n \in \mathbb{R}^d$ we set

$$U_0^n(x^1,\cdots,x^n) = \mathcal{U}_0\big(\mu^{\mathbf{x}}\big), \quad \text{where} \quad \mu^{\mathbf{x}} := \frac{1}{n}\sum_{i=1}^n \delta_{x^i}.$$

We further assume that for all integers $n \ge 1$

(U3) $U_0^n \in C^3\big((\mathbb{T}^d)^n\big)$ and for all $x^1,\cdots,x^n \in \mathbb{R}^d$

$$\frac{1}{n}\nabla_\mu \mathcal{U}_0(\mu^{\mathbf{x}})(x^i) = \nabla_{x^i} U_0^n(x^1,\cdots,x^n). \tag{5.3}$$

5.1. **Uniform estimates for finite dimensional systems.** In this subsection we review results of the theory of finite dimensional dynamical systems which can be found in [10] or [31] and then provide uniform estimate on dynamical systems consisting of finitely many indistinguishable particles.

For $t \in (0, T]$ we define

$$U^n(t, \mathbf{x}) = \frac{1}{n} \min_{\mathbf{r}} \Big\{ \sum_{i=1}^n \int_0^t \big(\frac{1}{2}|\dot{r}^i|^2 - V(r^i) - \frac{1}{2n} \sum_{j=1}^n W(r^i - r^j) \big) ds + U_0^n(\mathbf{r}(0)) \Big\},$$

where the minimum is performed over the set of $\mathbf{r} \in W^{1,2}\big(0, t; (\mathbb{T}^d)^n\big)$ such that $\mathbf{r}(t) = \mathbf{x}$. Observe that $U^n(t, \mathbf{x})$ is invariant under the permutation of the x^i's and so, we can define

$$\mathcal{U}^n(t, \mu^{\mathbf{x}}) := U^n(t, \mathbf{x}).$$

There exists $\mathbf{r}^n \in W^{1,2}\big(0, t; (\mathbb{T}^d)^n\big)$ which achieves the minimum in $U^n(t, \mathbf{x})$. We have

$$\ddot{r}_s^{n,i} = -\nabla V\big(r_s^{n,i}\big) - \frac{1}{n} \sum_{i=1}^n \nabla W\big(r_s^{n,i} - r_s^{n,j}\big) \quad (i = 1, \cdots, n). \tag{5.4}$$

We set

$$\sigma_s^n = \frac{1}{n} \sum_{i=1}^n \delta_{r_s^{n,i}}, \quad \mathbf{v}_s^n = \frac{1}{n} \sum_{i=1}^n \dot{r}_s^{n,i} \delta_{r_s^{n,i}}.$$

In general, we have $1/n\dot{\mathbf{r}}_0^n$ belongs to the super differential of U_0^n at $\mathbf{r}_0^n$. Since U_0^n is assumed to be differentiable we have, thanks to Equation (5.3)

$$\dot{r}_0^{n,i} = \nabla_\mu \mathcal{U}_0\big(\mu^{\mathbf{r}_0^n}\big)\big(r_0^{n,i}\big). \tag{5.5}$$

Hence,

$$(r_t^{n,i}, \dot{r}_t^{n,i}) = \Phi_t^{\sigma^n} \Big(r_0^{n,i}, \nabla_\mu \mathcal{U}_0\big(\mu^{\mathbf{r}_0^n}\big)\big(r_0^{n,i}\big) \Big) \tag{5.6}$$

and so,

$$\sigma_t^n = \Big(S_t^{\sigma^n, \sigma_0^n} \Big)_{\#} \sigma_0^n \quad \text{and} \quad \mathbf{v}_t^n = \mathbf{v}_t^{\sigma^n, \sigma_0^n}. \tag{5.7}$$

Equations (4.5), (4.6), (5.6) and (5.7) yield

$$\frac{1}{n} |\dot{\mathbf{r}}_t^n|^2 = ||\mathbf{v}_t^{\sigma^n, \sigma_0^n}||_{\sigma^n}^2 \leq s_\infty. \tag{5.8}$$

By the fact that $V, W \in C^3(\mathbb{T}^d)$ there exists a constant $C > 0$ such that

$$-C \leq \nabla^2 V, \nabla^2 W \leq C.$$

If $x^1, \cdots, x^n, y^1, \cdots, y^n \in \mathbb{R}^d$ permuting the order of the y^i's if necessary, we may assume that

$$\frac{1}{n} \sum_{i=1}^n |x^i - y^i|^2 = W_2^2(\mu^{\mathbf{x}}, \mu^{\mathbf{y}}).$$

Set

$$\gamma^n := \frac{1}{n} \sum_{i=1}^n \delta_{(x^i, y^i)}.$$

We have $\gamma^n \in \Gamma_o(\mu^{\mathbf{x}}, \mu^{\mathbf{y}})$. Furthermore,

$$\mathcal{V}(\mu^{\mathbf{y}}) - \mathcal{V}(\mu^{\mathbf{x}}) = \frac{1}{n} \sum_{i=1}^n V(y^i) - \frac{1}{n} \sum_{i=1}^n V(x^i)$$

$$\geq \frac{1}{n}\sum_{i=1}^{n}\langle \nabla V(x^i), y^i - x^i\rangle - \frac{C}{2n}\sum_{i=1}^{n}|y^i - x^i|^2$$

which means

$$\mathcal{V}(\mu^{\mathbf{y}}) \geq \mathcal{V}(\mu^{\mathbf{x}}) + \int_{\mathbb{T}^d\times\mathbb{T}^d}\langle \nabla_\mu \mathcal{V}(\mu^{\mathbf{x}})(q), r-q\rangle \gamma^n(dq,dr) - \frac{C}{2}W_2^2(\mu^{\mathbf{x}},\mu^{\mathbf{y}}). \quad (5.9)$$

Similarly,

$$\mathcal{V}(\mu^{\mathbf{y}}) \leq \mathcal{V}(\mu^{\mathbf{x}}) + \int_{\mathbb{T}^d\times\mathbb{T}^d}\langle \nabla_\mu \mathcal{V}(\mu^{\mathbf{x}})(q), r-q\rangle \gamma^n(dq,dr) + \frac{C}{2}W_2^2(\mu^{\mathbf{x}},\mu^{\mathbf{y}}). \quad (5.10)$$

For $\mathcal{W}$, we lose the coefficient $1/2$ in front of C to obtain

$$\left|\mathcal{W}(\mu^{\mathbf{y}}) - \mathcal{W}(\mu^{\mathbf{x}}) - \int_{\mathbb{T}^d\times\mathbb{T}^d}\langle \nabla_\mu \mathcal{W}(\mu^{\mathbf{x}})(q), r-q\rangle \gamma^n(dq,dr)\right| \leq CW_2^2(\mu^{\mathbf{x}},\mu^{\mathbf{y}}). \quad (5.11)$$

Since $\mathcal{P}(\mathbb{T}^d)$ is a bounded set, Theorem 6.1 yields that U is a κ_T–Lipschitz function on $[0,T]\times\mathcal{P}(\mathbb{T}^d)$, where κ_T depends only on T and the Lipschitz constant of $\mathcal{U}_0$. The bounds in Equations (5.8–5.11) are what is needed to obtain the following standard theorem with uniform estimates in n.

Theorem 5.1. *For $t \in [0,T]$*

(i)

$$\mathcal{U}^n(t,\sigma_t^n) = \mathcal{U}_0^n\big(\sigma_0^n\big) + \int_0^t \Big(\frac{1}{2}|(\sigma^n)'|^2(s) - \mathcal{V}(\sigma_s^n) - \mathcal{W}(\sigma_s^n)\Big)ds$$

(ii) The Lipschitz constant of $\mathcal{U}^n$ on $[0,T]\times\mathcal{P}(\mathbb{T}^d)$ is less than or equal to κ_T.

(iii) If $\mu^{\mathbf{y}} \in \mathcal{P}^n(\mathbb{T}^d)$ and $t\in(0,T)$ then there exists $\gamma_t^n \in \Gamma_o(\sigma_t^n,\mu^{\mathbf{y}})$ such that

$$\left|\mathcal{U}^n(t,\mu^{\mathbf{y}}) - \mathcal{U}^n(t,\sigma_t^n) - \int_{\mathbb{T}^d\times\mathbb{T}^d}\langle \mathbf{v}_t^n(q), r-q\rangle\gamma_t^n(dq,dr)\right| \leq \frac{4(C+1)}{t}W_2^2(\sigma_t^n,\mu^{\mathbf{y}}).$$

Proof. (i) The optimality of $\mathbf{r}^n$ in $U^n(T,\cdot)$ yields (i).

(ii) The standard theory of Hamiltonian systems ensures that (ii) holds with uniform estimates resulting from Equation (5.8) and the fact that the diameter of $\mathcal{P}(\mathbb{T}^d)$ is finite. However, a proof of (ii) in a more general setting has been provided in Subsection 6.3.

(iii) Under conditions (5.9–5.11) the theory of Hamiltonian systems yields (iii). □

By the fact that $\mathcal{P}^n(\mathbb{T}^d)\subset\mathcal{P}(\mathbb{T}^d)$ we have that

$$\mathcal{U}^n \geq \mathcal{U} \quad \text{on} \quad [0,T]\times\mathcal{P}^n(\mathbb{T}^d). \quad (5.12)$$

Set

$$\breve{\mathcal{U}}^n(s,\nu) = \inf_{t,\mu}\Big\{\mathcal{U}^n(t,\mu) + \kappa_T\big(|s-t| + W_2(\mu,\nu)\big) \mid t\in[0,T], \mu\in\mathcal{P}^n(\mathbb{T}^d)\Big\}.$$

Note that $\breve{\mathcal{U}}^n$ is a Lipschitz extension of $\mathcal{U}^n$ over $[0,T]\times\mathcal{P}(\mathbb{T}^d)$, with a Lipschitz constant less than or equal to κ_T.

5.2. **Optimal paths and their properties.** Fix $\mu \in \mathcal{P}(\mathbb{T}^d)$. The goal of this subsection is to construct a special path $\sigma \in AC_2(0,T;\mathcal{P}(\mathbb{T}^d))$ such that

$$\mathcal{U}(T,\mu) - \mathcal{U}_0(\sigma_0) = \int_0^T L(\sigma_t, \mathbf{v}_t)dt$$

and along which $\mathcal{U}$ is differentiable.

We choose a $\mu^n = 1/n \sum_{i=1}^n \delta_{x^{n,i}}$ such that $\{\mu^n\}_n$ converges to μ in the W_2–metric (cf. Lemma 8.2). Let $\{\sigma^n\}_n$ be the optimal paths obtained in Subsection 5.1. The metric W_2 being bounded on $\mathcal{P}(\mathbb{T}^d)$, thanks to Proposition 2 there exists an increasing sequence of integers $\{n_k\}_k$ (depending on μ) and paths $\sigma^\mu \in AC_2(0,T;\mathcal{P}(\mathbb{T}^d))$ such that for all $t \in [a,b]$, $\{\sigma_t^{n_k}\}_k$ converges narrowly to σ_t^μ. Futhermore, (3.3) holds.

Note that $|\breve{\mathcal{U}}^n(0,\cdot)| \leq ||\mathcal{U}_0||_\infty$. Since for each $\breve{\mathcal{U}}^n$ is κ_T–Lipschitz, we obtain that $\{\breve{\mathcal{U}}^n\}_n$ is equicontinuous and bounded in $[0,T]\times\mathcal{P}(\mathbb{T}^d)$. The latter set being compact (cf. [43]), we use the Ascoli–Arzela Theorem to obtain that $\{\breve{\mathcal{U}}^n\}_n$ is pre–compact for the uniform convergence. Any of its points of accumulation will be κ_T–Lipschitz.

Theorem 5.2. *The following hold:*

(i) the sequence $\{\breve{\mathcal{U}}^n\}_n$ converges uniformly to $\mathcal{U}$ on $[0,T]\times\mathcal{P}(\mathbb{T}^d)$ as $n\to\infty$.
(ii) There exists $\sigma \in AC_2(0,T;\mathcal{P}(\mathbb{T}^d))$ such that

$$\left(S_t^{\sigma,\sigma_0}\right)_{\#}\sigma_0 = \sigma_t$$

and

$$\mathcal{U}(T,\mu) = \mathcal{U}_0(\sigma_0) + \int_0^T L\big(\sigma_t, \mathbf{v}_t^{\sigma,\sigma_0}\big)dt. \tag{5.13}$$

(iii) $\mathbf{v}_t^{\sigma,\sigma_0}$ is the velocity of minimal norm for σ.
(iv) If $\nu \in \mathcal{P}(\mathbb{T}^d)$ and $t\in(0,T)$ then there exists $\gamma_t \in \Gamma_o(\sigma_t,\nu)$ such that

$$\Big|\mathcal{U}(t,\nu) - \mathcal{U}(t,\sigma_t) - \int_{\mathbb{T}^d\times\mathbb{T}^d}\langle \mathbf{v}_t^{\sigma,\sigma_0}(q), r-q\rangle \gamma_t(dq,dr)\Big| \leq \frac{4(C+1)}{t}W_2^2(\sigma_t,\nu).$$

Proof. (i) Let $\breve{\mathcal{U}}$ be a point of accumulation of $\{\breve{\mathcal{U}}^n\}_n$ for the uniform convergence, so that a subsequence of $\{\breve{\mathcal{U}}^n\}_n$ converges to $\breve{\mathcal{U}}$. To alleviate the notation we assume that the whole sequence $\{\breve{\mathcal{U}}^n\}_n$ converges to $\breve{\mathcal{U}}$ and will show that $\breve{\mathcal{U}} = \mathcal{U}$.

Fix $\nu \in \mathcal{P}(\mathbb{T}^d)$ and $t\in[0,T]$. Then choose $\nu^n \in \mathcal{P}^n(\mathbb{T}^d)$ such that $\{\nu^n\}_n$ converges to ν in the W_2–metric. We use Equation (5.12) to conclude that up to an appropriate subsequence

$$\breve{\mathcal{U}}(t,\nu) = \lim_{n\to\infty}\mathcal{U}^n(t,\nu^n) \geq \lim_{n\to\infty}\mathcal{U}(t,\nu^n) = \mathcal{U}(t,\nu). \tag{5.14}$$

Above, we have used the fact that $\mathcal{U}$ is Lipschitz as stated right before Theorem 5.1.

Let δ be an arbitrary positive number and let $\sigma \in AC_2(0,t;\mathcal{P}(\mathbb{T}^d))$ be such that $\sigma_t = \nu$ and

$$\mathcal{U}(t,\nu) \geq -\delta + \mathcal{U}_0(\sigma_0) + \int_0^t \Big(\frac{1}{2}|\sigma'|^2(s) - \int_{\mathbb{T}^d}(V + \frac{1}{2}W*\sigma_s)d\sigma_s\Big)ds. \tag{5.15}$$

By Lemma 8.2 there exist $\bar\sigma^n \in AC_2(0,T;\mathcal{P}^n(\mathbb{T}^d))$ such that

$$\lim_{n\to\infty}\sup_{s\in[0,t]} W_2(\sigma_s,\bar\sigma^n_s)=0 \quad\text{and}\quad \lim_{n\to\infty}\Big|\int_0^T |(\bar\sigma^n)'|^2(s)ds-\int_0^T|\sigma'|^2(s)ds\Big|=0. \tag{5.16}$$

Furthermore, we can find $\bar x^{n,i}\in AC_2(0,T,\mathbb{T}^d)$ $(i=1,\cdots,n)$ such that

$$\bar\sigma^n_s=\frac1n\sum_{i=1}^n \delta_{\bar x^{n,i}(s)}, \quad\text{and}\quad |(\bar\sigma^n)'|^2(s)=\frac1n\sum_{i=1}^n|\dot{\bar x}^{n,i}|^2(s).$$

Thus,

$$\mathcal{U}^n(t,\bar\sigma^n_t)-\mathcal{U}_0(\bar\sigma^n_0)\le\int_0^t\Big(\frac12|(\bar\sigma^n)'|^2(s)-\int_{\mathbb{T}^d}(V+\frac12 W*\bar\sigma^n_s)d\bar\sigma^n_s\Big)ds. \tag{5.17}$$

We first combine (5.15) and the second identity in (5.16) and then combine the first identity in (5.16) and (5.17) to obtain

$$\begin{aligned}\mathcal{U}(t,\nu) &\ge -\delta+\mathcal{U}_0(\sigma_0)+\lim_{n\to\infty}\int_0^t\Big(\frac12|(\bar\sigma^n)'|^2(s)-\int_{\mathbb{T}^d}(V+\frac12W*\bar\sigma^n_s)d\bar\sigma^n_s\Big)ds\\ &\ge -\delta+U_0(\sigma_0)+\limsup_{n\to\infty}\mathcal{U}^n(t,\sigma^n_t)-\mathcal{U}_0(\sigma^n_0)\\ &= -\delta+\breve{\mathcal{U}}(t,\sigma_t).\end{aligned} \tag{5.18}$$

Since δ is an arbitrary positive number, (5.14) and (5.18) establish (i).

(ii) We use (5.8), the fact that W_2 is uniformly bounded on $\mathcal{P}(\mathbb{T}^d)$ in Proposition 2 to obtain $\sigma\in AC_2(0,T;\mathcal{P}(\mathbb{T}^d))$ and an increasing sequence of integers $\{n_k\}_k$ such that for all $t\in[0,T]$, $\{\sigma^{n_k}_t\}_k$ converges narrowly to σ_t. To alleviate the notation, we assume that the whole sequence converges. By assumption (U2), $\{\nabla_\mu\mathcal{U}_0(\sigma^n_0)\}_n$ converges uniformly $\nabla_\mu\mathcal{U}_0(\sigma_0)$ on $\mathbb{T}^d$. By Lemma 4.3, $\{S^{\sigma^n,\nu^n}\}_n$ converges uniformly to S^{σ,σ_0} on $[0,T]\times\mathbb{T}^d$ and $\{\mathbf{v}^{\sigma^n,\nu^n}\}_n$ converges uniformly to $\mathbf{v}^{\sigma,\sigma_0}$ on $[0,T]\times\mathbb{R}^d$. By (5.7), $\left(S^{\sigma,\sigma_0}_t\right)_\#\sigma_0=\sigma_t$. We use Theorem 5.1 (i) to conclude the proof of (ii).

(iii) The fact that $\mathbf{v}^{\sigma^n,\nu^n}$ is a velocity for σ^n implies that $\mathbf{v}^{\sigma,\sigma_0}$ is a velocity for σ. The optimality condition in Equation (5.13) imposes that $\mathbf{v}^{\sigma,\sigma_0}_t$ is the velocity of minimal norm for σ.

(iv) Let $\{\nu^n\}_n\subset\mathcal{P}(\mathbb{T}^d)$ be a sequence converging narrowly to ν. For $t\in(0,T)$, Theorem 5.1 (iii) provides us with $\gamma^n_t\in\Gamma_o(\sigma^n_t,\nu^n)$ such that

$$\Big|\mathcal{U}^n(t,\nu^n)-\mathcal{U}^n(t,\sigma^n_t)-\int_{\mathbb{T}^d\times\mathbb{T}^d}\langle\mathbf{v}^n_t(q),r-q\rangle\gamma^n_t(dq,dr)\Big|\le\frac{4(C+1)}{t}W_2^2(\sigma^n_t,\mu^{\mathbf{y}}). \tag{5.19}$$

By Theorem 2.2, there exists a subsequence $\{\gamma^{n_k}_t\}_k$ (depending on t) that converges narrowly to some $\gamma_t\in\Gamma_o(\sigma_t,\nu)$. We use the fact that $\{\mathcal{U}^n\}_n$ converges uniformly, that $\{\mathbf{v}^{\sigma^n,\nu^n}\}_n$ converges uniformly to $\mathbf{v}^{\sigma,\sigma_0}$, and (5.19) to conclude the proof of (iv). □

Remark 7. In fact Theorem 5.2 proves the following (we write $\mathbf{v}$ instead of $\mathbf{v}^{\sigma,\sigma_0}$):

(i) For each $t\in(0,T)$, $\mathcal{U}_t$ is differentiable at σ_t,

$$\nabla_\mu\mathcal{U}_t(\sigma_t)=\mathbf{v}_t$$

and by (4.16), $\mathbf{v}$ is Lipschitz.

(ii) Since σ^μ satisfies the optimality condition (5.13), it then satisfies the PDEs (cf. [40])

$$\partial_t(\sigma \mathbf{v}) + \nabla \cdot \big(\sigma \mathbf{v} \otimes \mathbf{v}\big) = -\sigma_t \nabla\big(V + W * \sigma\big),$$

with the initial condition

$$\mathbf{v}_0 = \nabla_\mu \mathcal{U}_0(\sigma_0).$$

6. **Value functions and Hamilton Jacobi equations in the sense of differentials.** In the previous sections we have used that $\mathcal{P}(\mathbb{T}^d)$ is compact for the Wasserstein metric, a property which fails for $\mathcal{P}_2(\mathbb{R}^d)$. The results obtained in this section do not require such a compactness property and so, in the sequel $M = \mathbb{R}^d$ or $M = \mathbb{T}^d$. We also consider the potential functions which are more general than the ones considered in the previous sections. We only assume that $\mathcal{W} : \mathcal{P}_2(M) \to \mathbb{R}$ be a Borel function that is bounded below on bounded sets of $\big(\mathcal{P}_2(M), W_2\big)$. The main results of this section are Proposition 5 and Theorems 6.1, 6.4.

If $\mu \in \mathcal{P}_2(M)$ and $\mathbf{v}, \zeta \in L^2(\mu)$, we define

$$L(\mu, \mathbf{v}) = \frac{1}{2}||\mathbf{v}||^2_\mu - \mathcal{W}(\mu) \quad \text{and} \quad H(\mu, \zeta) = \frac{1}{2}||\zeta||^2_\mu + \mathcal{W}(\mu).$$

For $t \in (0, T]$ we define

$$\mathcal{U}(t, \nu) = \inf_{(\sigma, \mathbf{v})} \Big\{ \int_0^t L(\sigma_s, \mathbf{v}_s) ds + \mathcal{U}_0(\sigma_0) \mid \sigma_t = \nu \Big\}, \tag{6.1}$$

where the infimum is taken over the set of pairs $(\sigma, \mathbf{v})$ such that $\sigma \in AC_2(0, t; \mathcal{P}_2(M))$ and $\mathbf{v}$ is a velocity for σ.

6.1. **Conditions (I) and Lipschitz value function $\mathcal{U}(t, \cdot)$.** Assume that $\mathcal{U}_0, \mathcal{W} : \mathcal{P}_2(M) \to \mathbb{R}$ have a modulus of continuity $\omega \in C([0, \infty))$. In other words, ω is monotone nondecreasing, $\omega(0) = 0 \leq \omega(y)$ for all $y \geq 0$ and

$$|\mathcal{U}_0(\mu_1) - \mathcal{U}_0(\mu_0)|,\ |\mathcal{W}(\mu_1) - \mathcal{W}(\mu_0)| \leq \omega\big(W_2(\mu_0, \mu_1)\big)$$

for all $\mu_0, \mu_1 \in \mathcal{P}_2(M)$.

Proposition 3. *Assume $\mathcal{U}$ has only finite values for $t \in (0, T]$ and $\mu \in \mathcal{P}_2(M)$. Under the assumption that $\mathcal{W}$ and $\mathcal{U}_0$ have ω as a modulus of continuity, $\mathcal{U}(t, \cdot)$ has $(t+1)\omega$ as a modulus of continuity. In particular, if $\mathcal{U}_0$ and $\mathcal{W}$ are l–Lipschitz then $\mathcal{U}(t, \cdot)$ is $(t+1)l$–Lipschitz.*

Proof. Let ϵ be an arbitrary positive number and let $\nu_0, \nu_1 \in \mathcal{P}_2(M)$. Interchanging ν_0 with ν_1 if necessary, we assume without loss of generality that $\mathcal{U}(t, \nu_1) \geq \mathcal{U}(t, \nu_0)$. Let $\sigma \in AC_2(0, t; \mathcal{P}_2(M))$ and let $\mathbf{v}$ be a velocity for σ such that $\sigma_0 = \mu_1$, $\sigma_t = \nu_0$ and

$$\mathcal{U}(t, \nu_0) \geq -\epsilon + \int_0^t L(\sigma_s, \mathbf{v}_s) ds + \mathcal{U}_0(\sigma_0). \tag{6.2}$$

By Lemma 8.3 there exist $\sigma^* \in AC_2(0, t; \mathcal{P}_2(M))$ and a velocity $\mathbf{v}^*$ for σ^* such that $\sigma^*_t = \nu_1$,

$$\int_0^t ||\mathbf{v}^*_s||^2_{\sigma^*_s} ds \leq \int_0^t ||\mathbf{v}_s||^2_{\sigma_s} ds$$

and for all $s \in [0, t]$

$$W_2(\sigma_s, \sigma^*_s) \leq W_2(\sigma_t, \sigma^*_t). \tag{6.3}$$

We have

$$\mathcal{U}(t,\nu_1) \leq \int_0^t \Big(\frac{1}{2}||\mathbf{v}_s^*||^2_{\sigma_s^*} - \mathcal{W}(\sigma_s^*)\Big)ds + \mathcal{U}_0(\sigma_0^*) \leq \int_0^t \Big(\frac{1}{2}||\mathbf{v}_s||^2_{\sigma_s} - \mathcal{W}(\sigma_s^*)\Big)ds + \mathcal{U}_0(\sigma_0^*).$$

This, together with (6.3), implies

$$\mathcal{U}(t,\nu_1) \leq \int_0^t \Big(\frac{1}{2}||\mathbf{v}_s||^2_{\sigma_s} - \mathcal{W}(\sigma_s) + \omega\big(W_2(\sigma_s,\sigma_s^*)\big)\Big)ds + \mathcal{U}_0(\sigma_0) + \omega\big(W_2(\sigma_0,\sigma_0^*)\big). \tag{6.4}$$

We combine (6.2–6.4) to obtain

$$|\mathcal{U}(t,\nu_1) - \mathcal{U}(t,\nu_0)| \leq \epsilon + (t+1)\omega\big(W_2(\sigma_t,\sigma_t^*)\big) = \epsilon + (t+1)\omega\big(W_2(\nu_0,\nu_1)\big).$$

Since ϵ is an arbitrary positive number, this concludes the proof of the proposition. □

6.2. **Continuity of $(T,\mu,\nu) \to C_0^T(\mu,\nu)$ under conditions (II).** We suppose $\mathcal{W}$ is a Borel function, bounded from below on balls. We suppose that

$$\limsup_{n\to\infty} \mathcal{W}(\mu^n) \leq \mathcal{W}(\mu) \tag{6.5}$$

for all bounded sequences $\{\mu^n\}_n \subset \mathcal{P}_2(M)$ that converge narrowly to μ. We further assume there exist constants $C_0 > 0$ and $\beta \in [1,2)$ such that

$$\mathcal{W}(\mu) \leq C_0\Big(1 + \int_M |x|^\beta \mu(dx)\Big) \tag{6.6}$$

for all $\mu \in \mathcal{P}_2(M)$.

For $\epsilon_0 > 0$, let D_{ϵ_0} be a positive number depending only on ϵ_0 and β such that $|x|^\beta \leq \epsilon_0|x|^2 + D_{\epsilon_0}$. Throughout this subsection we assume that

$$2C_0\epsilon_0 T^2 < 1/4. \tag{6.7}$$

Set

$$\bar\lambda(\mu) = C_0\Big(1 + D_{\epsilon_0} + 2\epsilon_0\int_M |x|^2\mu(dx)\Big), \quad \lambda^*(T,\mu) = 4\Big(\bar\lambda(\mu) + L(\mu,\vec{0}) + \frac{1}{T}\mathcal{U}_0(\mu)\Big).$$

Since $\mathcal{W}$ is bounded from below on bounded sets, there exists a monotone nondecreasing function $\mathcal{W}^o \in C([0,\infty))$ such that for each $R > 0$,

$$\mathcal{W}^o(R) \geq \sup_\mu\{-\mathcal{W}(\mu) \mid \int_M |x|^2\mu(dx) \leq R^2\}. \tag{6.8}$$

Examples of $\mathcal{W}$ include

$$\mathcal{W}(\mu) = \int_M \varphi(x)\mu(dx) + \int_{M\times M}\phi(x-y)\mu(dx)\mu(dy),$$

where $\varphi, \phi \in C^1(M)$ are semiconcave and satisfy

$$|\varphi(x)| \leq \frac{C_0}{2}(1+|x|^\beta) \quad \text{and} \quad |\phi(x)| \leq \frac{C_0}{4}(1+|x|^\beta)$$

for all $x \in M$.

Remark 8. Let $\sigma \in AC_2(0,T;\mathcal{P}_2(M))$ be such that $\mathbf{v}$ is one of its velocities. We have

$$\mathcal{W}(\sigma_t) \leq \bar\lambda(\mu) + \frac{1}{4T}\int_0^T ||\mathbf{v}_\tau||^2_{\sigma_\tau} d\tau \tag{6.9}$$

and

$$\int_0^T L(\sigma_t, v_t)dt \geq -T\bar{\lambda}(\mu) + \frac{1}{4}\int_0^T ||\mathbf{v}_t||^2_{\sigma_t} dt. \tag{6.10}$$

Let $\sigma \in AC_2(0,T;\mathcal{P}_2(M))$ has $\mathbf{v}$ as a velocity. First,

$$\int_M |z|^\beta \sigma_t(dz) \leq D_{\epsilon_0} + \epsilon_0 \int_M |z|^2 \sigma_t(dz). \tag{6.11}$$

We use Remark 1 and Hölder's inequality to obtain

$$W_2^2(\sigma_t, \delta_{\vec{0}}) \leq 2W_2^2(\sigma_s, \delta_{\vec{0}}) + 2T\int_0^T ||\mathbf{v}_\tau||^2_{\sigma_\tau} d\tau. \tag{6.12}$$

We have by (6.12)

$$\int_M |x|^\beta \sigma_t(dx) \leq D_{\epsilon_0} + 2\epsilon \int_M |x|^2 \sigma_s(dx) + 2T\epsilon_0 \int_0^T ||\mathbf{v}_\tau||^2_{\sigma_\tau} d\tau. \tag{6.13}$$

Setting $s = T$ in (6.13) and using (6.6) we conclude that if $2C_0\epsilon_0 T^2 < 1/4$, then (6.9) holds. A direct integration over $[0,T]$ yields (6.10).

Proposition 4 (Existence of optimal paths and velocity estimate)**.** *Suppose $\mathcal{W}$ satisfies (6.5), $\mathcal{U}_0$ is bounded below by a constant u_- and lower semicontinuous for the narrow convergence topology. Then Equation (6.1) admits a minimizer $(\sigma, \mathbf{v})$ such that $\mathbf{v}$ is the velocity of minimal norm for σ and $H(\sigma_t, \mathbf{v}_t)$ is time independent. We have*

$$W_2^2(\mu, \sigma_t) \leq T^2\lambda^*(T,\mu) - 4Tu_-, \tag{6.14}$$

$$W_2^2(\sigma_t, \delta_{\vec{0}}) \leq 2T^2\lambda^*(T,\mu) - 8Tu_- + 2W_2^2(\mu, \delta_{\vec{0}}). \tag{6.15}$$

Furthermore,

$$\begin{aligned} ||\mathbf{v}_t||^2_{\sigma_t} \quad &\leq \quad \lambda^*(T,\mu) - \frac{6u_-}{T} + 2\mathcal{W}^o\Big(\sqrt{2T^2\lambda^*(T,\mu) - 8Tu_- + 2W_2^2(\mu, \delta_{\vec{0}})}\Big) \\ &+ \quad 2\bar{\lambda}(\mu) + \frac{1}{2}\lambda^*(T,\mu). \end{aligned} \tag{6.16}$$

Proof. The proof of Lemma 5.3 [42] can be adapted to obtain existence of a minimizer $(\sigma, \mathbf{v})$. Observe that $\mathbf{v}$ must be the velocity of minimal norm and so, by Proposition 3.11 [41], we may assume without loss of generality that $H(\sigma_t, \mathbf{v}_t)$ is time independent.

Existence of a minimizer $(\sigma, \mathbf{v})$ in (6.1) was proved in [41]. Setting

$$\sigma_t^* = \mu, \quad \mathbf{v}_t^* = \vec{0}$$

for all $t \in [0,T]$ we have

$$\int_0^T L(\sigma_t, \mathbf{v}_t)dt + \mathcal{U}_0(\sigma_0) = \mathcal{U}(t,\mu) \leq \int_0^T L(\sigma_t^*, \mathbf{v}_t)dt + \mathcal{U}_0(\sigma_0^*) = TL(\mu, \vec{0}) + \mathcal{U}_0(\mu).$$

We exploit (6.10) to conclude that

$$\int_0^T ||\mathbf{v}_t||^2_{\sigma_t} dt \leq T\lambda^*(T,\mu) - 4u_-. \tag{6.17}$$

This, together with Remark 1 implies that (6.14) holds. We combine (6.12) (with $\sigma_s = \mu$) with (6.14) to obtain (6.15). Hence, by (6.8) and (6.15)

$$-\mathcal{W}(\sigma_t) \leq \mathcal{W}^o\Big(\sqrt{2T^2\lambda^*(T,\mu) - 8Tu_- + 2W_2^2(\mu, \delta_{\vec{0}})}\Big). \tag{6.18}$$

We use the first inequality in Remark 8 and (6.17) to conclude that

$$\mathcal{W}(\sigma_t) \leq \bar{\lambda}(\mu) + \frac{1}{4}\lambda^*(T,\mu) - \frac{u_-}{T}. \tag{6.19}$$

By (6.17), the set of $t_0 \in [0,T]$ such that

$$||\mathbf{v}_{t_0}||^2_{\sigma_{t_0}} \leq \lambda^*(T,\mu) - \frac{4u_-}{T}$$

is a set of positive measure. Choose such a t_0 and use the fact that $H(\sigma_t, \mathbf{v}_t)$ is independent of t to conclude

$$||\mathbf{v}_t||^2_{\sigma_t} = ||\mathbf{v}_{t_0}||^2_{\sigma_{t_0}} + 2\big(\mathcal{W}(\sigma_{t_0}) - \mathcal{W}(\sigma_t)\big) \leq \lambda^*(T,\mu) - \frac{4u_-}{T} + 2\big(\mathcal{W}(\sigma_{t_0}) - \mathcal{W}(\sigma_t)\big).$$

This together with (6.18) and (6.19) yields (6.16). □

Remark 9 (The discrete case)**.** Suppose $\mathcal{W}$ satisfies (6.5), $\mathcal{U}_0$ is bounded from below by a constant u_- and is lower semicontinuous for the narrow convergence topology. For an integer $n \geq 1$, $\mu \in \mathcal{P}^n(M)$ we define

$$\mathcal{U}^n(t,\mu) = \min_{(\sigma,\mathbf{v})}\Big\{\int_0^t L(\sigma_\tau, \mathbf{v}_\tau)d\tau + \mathcal{U}_0(\mu) \mid \sigma_t = \mu\Big\}, \tag{6.20}$$

where the minimum is performed over the set of $(\sigma, \mathbf{v})$ such that $\sigma \in AC_2(0,t;\mathcal{P}^n(M))$ and $\mathbf{v}$ is a velocity for σ. Existence of a minimizer $(\sigma,\mathbf{v})$ in the finite dimensional problem (6.20) is obtained by standard methods of the calculus of variations. As above, $H(\sigma_t,\mathbf{v}_t)$ is time independent and (6.14, 6.15, 6.16) continue to hold.

Assume $\sigma : [0,1] \to \mathcal{P}_2(\mathbb{R}^d)$ is a geodesic of constant speed connecting μ to ν. Then, the velocity $\mathbf{v}$ of minimal norm for σ is such that $||\mathbf{v}_t||_{\sigma_t} = W_2(\mu,\nu)$. Given $\epsilon > 0$ we consider the path $\sigma^\epsilon : [0,1] \to \mathcal{P}_2(\mathbb{R}^d)$ obtained by the reparametrization $\sigma^\epsilon_\tau = \sigma_{\tau\epsilon^{-1}}$. Its velocity of minimal norm $\mathbf{v}^\epsilon$ satisfies $\mathbf{v}^\epsilon_\tau = \epsilon^{-1}\mathbf{v}_{\tau\epsilon^{-1}}$ and so,

$$\int_0^\epsilon ||\mathbf{v}^\epsilon_\tau||^2_{\sigma^\epsilon_\tau} = \frac{W_2^2(\mu,\nu)}{\epsilon}. \tag{6.21}$$

We have

$$W_2(\sigma^\epsilon_t, \sigma^\epsilon_0) = W_2(\sigma_{t\epsilon^{-1}}, \sigma_0) = t\epsilon^{-1}W_2(\mu,\nu)$$

and

$$W_2(\sigma_t, \delta_{\vec{0}}) \leq W_2(\sigma_t,\sigma_0) + W_2(\sigma_0,\delta_{\vec{0}}) = tW_2(\sigma_1,\sigma_0) + W_2(\sigma_0,\delta_{\vec{0}}).$$

Hence,

$$W_2(\sigma_t,\delta_{\vec{0}}) \leq 2(W_2(\sigma_1,\delta_{\vec{0}}) + W_2(\delta_{\vec{0}},\sigma_0)). \tag{6.22}$$

By (6.22)

$$-\int_0^\epsilon \mathcal{W}(\sigma^\epsilon_\tau)dt \leq \epsilon\mathcal{W}^o\Big(2\big(W_2(\mu,\delta_{\vec{0}}) + W_2(\nu,\delta_{\vec{0}})\big)\Big)$$

and so, by (6.21)

$$C^\epsilon_0(\mu,\nu) \leq \frac{W_2^2(\mu,\nu)}{2\epsilon} + \epsilon\mathcal{W}^o\Big(2\big(W_2(\mu,\delta_{\vec{0}}) + W_2(\nu,\delta_{\vec{0}})\big)\Big). \tag{6.23}$$

By (6.10), C^T_0 never achieves the value $-\infty$ on $\mathcal{P}_2(M)\times\mathcal{P}_2(M)$.

Assume $\sigma \in AC_2(0,T;\mathcal{P}_2(M))$, $\mathbf{v}$ is a velocity for σ, $\sigma_0 = \mu$, $\sigma_T = \nu$ and

$$\int_0^T L(\sigma_t,\mathbf{v}_t)dt \leq C^T_0(\mu,\nu) + T.$$

By (6.10)

$$\int_0^T ||\mathbf{v}_t||^2_{\sigma_t} dt \le 4T\lambda(\nu) + C_0^T(\mu,\nu) + T. \tag{6.24}$$

Hence, the set of $t \in [0,T]$ such that

$$||\mathbf{v}_t||^2_{\sigma_t} \le 4\bar{\lambda}(\nu) + \frac{C_0^T(\mu,\nu)}{T} + 1 \tag{6.25}$$

is of positive measure. We use (6.12) and (6.24), then replace ϵ by T in (6.23) to obtain

$$W_2^2(\sigma_t,\nu) \le T^2 + 4T^2\lambda(\nu) + \frac{W_2^2(\mu,\nu)}{2} + T^2\mathcal{W}^o\Big(2\big(W_2(\mu,\delta_{\vec{0}}) + W_2(\nu,\delta_{\vec{0}})\big)\Big). \tag{6.26}$$

Remark 10. Since $\mathcal{W}$ satisfies (6.5), as in Proposition 4, (6.1) admits a minimizer $(\sigma,\mathbf{v})$. By (6.26), the range of $\mathcal{W}(\sigma)$ is contained in an interval centered at the origin and whose length $l(\mu,\nu)$ is a monotone nondecreaasing function of $W_2(\mu,\delta_{\vec{0}})+W_2(\nu,\delta_{\vec{0}})$. By Proposition 3.11 [41], we may assume without loss of generality that $H(\sigma_t,\mathbf{v}_t)$ is independent of t. Choose t_0 such that (6.25) holds. We have

$$||\mathbf{v}_t||^2_{\sigma_t} = ||\mathbf{v}_{t_0}||^2_{\sigma_{t_0}} + 2\big(\mathcal{W}(\sigma_{t_0}) - \mathcal{W}(\sigma_t)\big).$$

This, together with (6.24-6.25), yields existence of a function $R \in C\big([0,\infty)^2\big)$, monotone, nondecreasing in each of their variables, such that

$$\sup_{t\in[0,T]} ||\mathbf{v}_t||^2_{\sigma_t} \le \frac{C_0^T(\mu,\nu)}{T} + R\Big(T, W_2(\mu,\delta_{\vec{0}}) + W_2(\nu,\delta_{\vec{0}})\Big). \tag{6.27}$$

Proposition 5. *The function $F : (T,\mu,\nu) \to C_0^T(\mu,\nu)$ is continuous on the metric space $\mathcal{S} = (0,\infty)\times\mathcal{P}_2(M)\times\mathcal{P}_2(M)$. Suppose $\mathcal{U}_0 : \mathcal{P}_2(M) \to \mathbf{R}$ is continuous, bounded from below and set*

$$\mathcal{U}(T,\mu) = \inf_{\nu\in\mathcal{P}_2(M)} \{C_T(\nu,\mu) + \mathcal{U}_0(\nu)\}.$$

Then $\mathcal{U}$ is continuous on $[0,\infty)\times\mathcal{P}_2(M)$.

Proof. We are to show that F is sequentially lower and upper semicontinuous at each point $(T,\mu,\nu)\in\mathcal{S}$. Suppose $\{T^n\}_n \subset (0,\infty)$ converges to $T\in(0,\infty)$, $\{\mu^n\}_n$ converges to μ in $\mathcal{P}_2(M)$ and $\{\nu^n\}_n$ converges to ν in $\mathcal{P}_2(M)$.

1. Let $\delta > 0$ and let $\sigma \in AC_2(0,T;\mathcal{P}_2(M))$ and let $\mathbf{v}$ be its velocity of minimal norm such that

$$C_0^T(\mu,\nu) > \int_0^T L(\sigma_t,\mathbf{v}_t)dt - \delta, \quad \sigma_0 = \mu, \quad \sigma_T = \nu. \tag{6.28}$$

Fix $\epsilon > 0$ small enough and assume without loss of generality that $|T - T^n| < \epsilon$. Then,

$$C_0^{T^n}(\mu^n,\nu^n) \le C_0^\epsilon(\mu^n,\sigma_\epsilon) + C_\epsilon^{T-\epsilon}(\sigma_\epsilon,\sigma_{T-\epsilon}) + C_{T-\epsilon}^{T^n}(\sigma_{T-\epsilon},\nu^n). \tag{6.29}$$

By (6.23)

$$\limsup_{n\to\infty} C_0^\epsilon(\mu^n,\sigma_\epsilon) \le \frac{W_2^2(\mu,\sigma_\epsilon)}{2\epsilon} + \epsilon\mathcal{W}^o\Big(2\big(W_2(\mu,\delta_{\vec{0}}) + W_2(\sigma_\epsilon,\delta_{\vec{0}})\big)\Big). \tag{6.30}$$

Similarly,

$$\limsup_{n\to\infty} C_{T-\epsilon}^{T^n}(\sigma_{T-\epsilon},\nu^n) \le \frac{W_2^2(\sigma_{T-\epsilon},\nu)}{2\epsilon} + \epsilon\mathcal{W}^o\Big(2\big(W_2(\nu,\delta_{\vec{0}}) + W_2(\sigma_{T-\epsilon},\delta_{\vec{0}})\big)\Big). \tag{6.31}$$

By Remark 1

$$\frac{W_2^2(\mu,\sigma_\epsilon)}{\epsilon} \le \int_0^\epsilon ||\mathbf{v}_\tau||^2_{\sigma_\tau} d\tau, \quad \frac{W_2^2(\sigma_{T-\epsilon},\nu)}{\epsilon} \le \int_{T-\epsilon}^T ||\mathbf{v}_\tau||^2_{\sigma_\tau} d\tau.$$

This, together with (6.29–6.31), implies

$$\begin{aligned}\limsup_{n\to\infty} C_0^{T^n}(\mu^n,\nu^n) &\le \liminf_{\epsilon\to 0^+} C_\epsilon^{T-\epsilon}(\sigma_\epsilon,\sigma_{T-\epsilon}) \\ &\le \lim_{\epsilon\to 0^+}\int_\epsilon^{T-\epsilon} L(\sigma_t,\mathbf{v}_t)dt \\ &= \int_0^T L(\sigma_t,\mathbf{v}_t)dt.\end{aligned}$$

Hence by (6.28)

$$\limsup_{n\to\infty} C_0^{T^n}(\mu^n,\nu^n) \le C_0^T(\mu,\nu)+\delta.$$

Since $\delta>0$ is arbitrary, we conclude that F is upper semicontinuous.

2. For each n let $\sigma^n \in AC_2(0,T^n;\mathcal{P}_2(M))$ and let $\mathbf{v}^n$ be its velocity of minimal norm such that

$$C_0^{T^n}(\mu^n,\nu^n) > \int_0^{T^n} L(\sigma_t^n,\mathbf{v}_t^n)dt - \frac{1}{n}, \quad \sigma_0^n=\mu^n, \quad \sigma_T^n=\nu^n. \tag{6.32}$$

Since $\{(T^n,\mu^n,\nu^n)\}_n$ is bounded in $\mathcal{S}$, (6.23) implies that $\{C_0^{T^n}(\mu^n,\nu^n)\}_n$ is bounded above in $\mathbb{R}$. Thus by (6.24) and (6.26), for each $\delta>0$ small enough, the following suprema are not only independent of δ but they are finite:

$$\sup_n \int_0^{T-\delta} ||\mathbf{v}_t^n||^2_{\sigma_t^n}dt, \quad \sup_{n,t}\{W_2(\sigma_t^n,\delta_{\vec{0}}) \mid t\in[0,T-\delta], n\in\mathbb{N}\}<\infty. \tag{6.33}$$

We refer to Propositions 3 and 4 in [40] to infer the existence of $\sigma \in AC_2(0,T;\mathcal{P}_2(M))$ such that, up to a subsequence which is independent of t, $\{\sigma_t^n\}_n$ converges narrowly to σ_t for each $t\in[0,T)$ and

$$\liminf_{n\to\infty}\int_0^{T_n} ||\mathbf{v}_t^n||^2_{\sigma_t^n}dt \ge \liminf_{n\to\infty}\int_0^{T-\delta} ||\mathbf{v}_t^n||^2_{\sigma_t^n}dt \ge \int_0^{T-\delta}||\mathbf{v}_t||^2_{\sigma_t}dt. \tag{6.34}$$

Here, $\mathbf{v}$ is the velocity of minimal norm for σ. Letting δ tend to 0 in (6.34) we have

$$\liminf_{n\to\infty}\int_0^{T_n} ||\mathbf{v}_t^n||^2_{\sigma_t^n}dt \ge \liminf_{n\to\infty}\int_0^{T-\delta} ||\mathbf{v}_t^n||^2_{\sigma_t^n}dt \ge \int_0^{T-\delta}||\mathbf{v}_t||^2_{\sigma_t}dt. \tag{6.35}$$

It is apparent that we can define univoquely σ_T and obtain

$$\sigma\in AC_2(0,T;\mathcal{P}_2(M)), \quad \sigma_0=\mu \quad \text{and} \quad \sigma_T=\nu.$$

By (6.33), $\{\sigma_t^n\}_{n,t}$ is a bounded subset of $\mathcal{P}_2(M)$. Thus, by (6.6), $\{-\mathcal{W}(\sigma_t^n)\}_{n,t}$ is bounded from below in $\mathbb{R}$ by a certain number b. We then apply Fatou's Lemma and use (6.5) to conclude that

$$\liminf_{n\to\infty}\int_0^{T_n}(-\mathcal{W}(\sigma_t^n)-b)dt \ge \liminf_{n\to\infty}\int_0^{T-\delta}(-\mathcal{W}(\sigma_t^n)-b)dt \ge \int_0^{T-\delta}(-\mathcal{W}(\sigma_t)-b)dt.$$

Letting δ tend to 0, we obtain

$$\liminf_{n\to\infty}\int_0^{T_n} -\mathcal{W}(\sigma_t^n) \ge \int_0^T -\mathcal{W}(\sigma_t)dt. \tag{6.36}$$

Thus, combining (6.32) (6.35) and (6.36) we infer that

$$\liminf_{n\to\infty} C_0^{T^n}(\mu^n,\nu^n) \geq \int_0^T L(\sigma_t,\mathbf{v}_t)dt \geq C_0^T(\mu,\nu).$$

Consequently, F is also lower semicontinuous and so, it is continuous.

3. Suppose that $T=0$. Then

$$\mathcal{U}(T^n,\mu^n) \leq C_0^{T^n}(\mu^n,\mu^n) + \mathcal{U}_0(\mu^n) \leq -T^n\mathcal{W}(\mu^n) + \mathcal{U}_0(\mu^n).$$

Since $\mathcal{W}$ is bounded from below on bounded sets, we have that $\{\mathcal{U}(T^n,\mu^n)\}_n$ is bounded above in $\mathbb{R}$ by a constant which we denote by λ. We first conclude that

$$\limsup_{n\to\infty} \mathcal{U}(T^n,\mu^n) \leq \limsup_{n\to\infty}\{-T^n\mathcal{W}(\mu^n) + \mathcal{U}_0(\mu^n)\} \leq \mathcal{U}_0(\mu).$$

Hence, $\mathcal{U}$ is upper semicontinuous at $(0,\mu)$.

Let $\{\eta^n\}_n \subset \mathcal{P}_2(M)$ be such that

$$\lambda \geq \mathcal{U}(T^n,\mu^n) \geq -\frac{1}{n} + \int_0^{T^n} L(\sigma_t^n,\mathbf{v}_t^n)dt + \mathcal{U}_0(\eta^n) \geq -\frac{1}{n} + C_0^{T^n}(\eta^n,\mu^n) + \mathcal{U}_0(\eta^n),$$

where

$$\sigma^n \in AC_2(0,T^n;\mathcal{P}_2(M)), \quad \sigma_0^n = \eta^n, \quad \text{and} \quad \sigma_{T^n}^n = \mu^n.$$

By (6.33) and the fact that $\mathcal{U}_0$ is bounded from below, we have that $\{\eta^n\}_n$ is a bounded sequence. As above

$$\sup_n \int_0^{T^n} ||\mathbf{v}_t^n||^2_{\sigma_t^n} dt, \quad \sup_n W_2(\sigma_t^n,\delta_{\vec{0}}) < \infty. \tag{6.37}$$

By Remark 1

$$W_2^2(\eta^n,\mu^n) \leq T^n \int_0^{T^n} ||\mathbf{v}_t^n||^2_{\sigma_t^n} dt.$$

We conclude that $\{\eta^n\}_n$ converges to μ and so, by (6.6), $\{\mathcal{W}(\sigma_t^n)\}_n$ is bounded from below. Hence,

$$\begin{aligned}
\liminf_{n\to\infty} \mathcal{U}(T^n,\mu^n) &\geq \liminf_{n\to\infty} C_0^{T^n}(\eta^n,\mu^n) + \mathcal{U}_0(\eta^n) \\
&\geq \liminf_{n\to\infty} -\int_0^{T^n} \mathcal{W}(\sigma_t^n)dt + \mathcal{U}_0(\eta^n) \\
&\geq \mathcal{U}_0(\mu).
\end{aligned}$$

Hence, $\mathcal{U}$ is also lower semicontinuous at $(0,\mu)$ and so, it is continuous there.

4. Arguments similar to those used in steps 1–3 yield that $\mathcal{U}$ is continuous at (T,μ) if $T>0$. □

6.3. **Lipschitz properties of $\mathcal{U}$ in all variables under conditions (I) and (II).** Throughout this subsection we assume that

$$\mathcal{U}_0, \mathcal{W} : \mathcal{P}_2(M) \to \mathbb{R}$$

are κ–Lipschitz, U_0 is lower semicontinuous for the narrow convergence, $\mathcal{W}$ satisfies (6.5) and (6.6). We assume that

$$T>0, \quad \epsilon_0>0, \quad 8\kappa\epsilon_0<1, \quad 8C_0\epsilon_0T^2<1$$

and D_{ϵ_0} is such that

$$|x|^\beta \leq \epsilon_0|x|^2 + D_{\epsilon_0}$$

for all $x \in M$. For each $r > 0$, we define $\mathcal{S}_r$ to be the Cartesian product of $[0, T]$ and the closed ball of center $\delta_{\vec{0}}$ and radius r in $\mathcal{P}_2(M)$. The purpose of this section is to show that the value function U in (6.1) is Lipschitz on $\mathcal{S}_r$.

We will use the fact that $\mathcal{U}$ satisfies the following property (cf. Lemma 2.4 [41]):

$$\mathcal{U}(s, \mu) = \min_{\sigma} \Big\{ \int_t^s L(\sigma_\tau, \mathbf{v}_\tau) d\tau + \mathcal{U}(t, \sigma_t) \Big\} \quad 0 \le t < s \le T, \tag{6.38}$$

where, the infimum is performed over the set of $(\sigma, \mathbf{v})$ such that $\sigma \in AC_2(t, s; \mathcal{P}_2(M))$, $\mathbf{v}$ is a velocity for σ and $\sigma_s = \mu$.

If $\mu \in \mathcal{P}_2(M)$,

$$\mathcal{U}(t, \mu) \le C_0^t(\mu, \mu) + \mathcal{U}_0(\mu) \le -t\mathcal{W}(\mu) + \mathcal{U}_0(\mu). \tag{6.39}$$

Let $\sigma \in AC_2(0, t; \mathcal{P}_2(M))$, let $\mathbf{v}$ be a velocity for σ and assume that $\sigma_t = \mu$. If

$$\int_0^t L(\sigma_\tau, \mathbf{v}_\tau) d\tau + \mathcal{U}_0(\sigma_0) \le \mathcal{U}(t, \mu) + t$$

then by (6.39)

$$\int_0^t L(\sigma_\tau, \mathbf{v}_\tau) d\tau \le -t\mathcal{W}(\mu) + \mathcal{U}_0(\mu) + t - \mathcal{U}_0(\sigma_0) \le (1 - \mathcal{W}(\mu))t + \kappa W_2(\sigma_0, \mu).$$

We use Remark 1 to conclude that

$$\begin{aligned}\int_0^t L(\sigma_\tau, \mathbf{v}_\tau) d\tau &\le (1 - \mathcal{W}(\mu))t + \kappa \int_0^t ||\mathbf{v}_\tau||_{\sigma_\tau} d\tau \\ &\le (1 - \mathcal{W}(\mu))t + \kappa\epsilon_0 \int_0^t ||\mathbf{v}_\tau||^2_{\sigma_\tau} d\tau + \frac{\kappa t}{\epsilon_0}.\end{aligned}$$

By (6.10)

$$\int_0^t ||\mathbf{v}_\tau||^2_{\sigma_\tau} d\tau \le 4t\bar{\lambda}(\sigma_t) + 4(1 - \mathcal{W}(\mu))t + 4\kappa\epsilon_0 \int_0^t ||\mathbf{v}_\tau||^2_{\sigma_\tau} d\tau + 4\frac{\kappa t}{\epsilon_0}.$$

Thus,

$$\int_0^t ||\mathbf{v}_\tau||^2_{\sigma_\tau} d\tau \le 8t\bar{\lambda}(\mu) + 8(1 - \mathcal{W}(\mu))t + 8\frac{\kappa t}{\epsilon_0}. \tag{6.40}$$

By Remark 1 and (6.40)

$$W_2^2(\sigma_t, \mu) \le 8t^2\bar{\lambda}(\mu) + 8(1 - \mathcal{W}(\mu))t^2 + 8\frac{\kappa}{\epsilon_0}. \tag{6.41}$$

Theorem 6.1. *The restriction of $\mathcal{U}$ to $\mathcal{S}_r$ is a Lipschitz continuous function.*

Proof. Recall that by Proposition 3, for each $t \in [0, T]$, $\mathcal{U}(t, \cdot)$ is $((T+1)\kappa)$–Lipschitz. It remains to show that for each $\mu \in \mathcal{P}_2(M)$ such that $W_2(\mu, \delta_{\vec{0}}) \le r$, $\mathcal{U}(\cdot, \mu)$ is Lipschitz, with a Lipschitz constant independent of μ and depending only on r.

As done in Subsection 6.2, we use (6.40–6.41) and the fact that $\mathcal{U}_0$ is lower semicontinuous for the narrow convergence topology to obtain the following: if $\mu \in \mathcal{P}_2(M))$ and $t \in [0, T]$, there exists $\sigma^{\mu,t} \in AC_2(0, t; \mathcal{P}_2(M))$ and a velocity $\mathbf{v}^{\mu,t}$ for $\sigma^{\mu,t}$ such that

$$\mathcal{U}(t, \mu) = \int_0^t L(\sigma_\tau^{\mu,t}, \mathbf{v}_\tau^{\mu,t}) d\tau + \mathcal{U}_0(\sigma_0^{\mu,t})$$

and

$$\sup_{\tau,t,\mu} \{ ||\mathbf{v}_\tau^{\mu,t}||_{\sigma_\tau^t} \mid 0 \le \tau \le t \le T,\ W_2(\mu, \delta_{\vec{0}}) \le r \} < \infty.$$

Hence, by Remark 1

$$\sup_{\tau,t,\mu}\{W_2(\sigma^t_\tau,\mu) \mid 0\le \tau\le t\le T,\ W_2(\mu,\delta_{\vec 0})\le r\}<\infty.$$

Thus,

$$s_1=\sup_{\tau,t,\mu}\Big\{||\mathbf{v}^{\mu,t}_\tau||_{\sigma^{\mu,t}_\tau}+|L(\sigma^t_\tau,\mathbf{v}^{\mu,t}_\tau)| \;\Big|\; 0\le\tau\le t\le T,\ W_2(\mu,\delta_{\vec 0})\le r\Big\}<\infty.$$

Let $s\in[0,t)$. By equation (6.38)

$$\mathcal{U}(t,\mu)=\int_s^t L(\sigma^{\mu,t}_\tau,\mathbf{v}^{\mu,t}_\tau)d\tau+\mathcal{U}(s,\sigma^{\mu,t}_s)$$

and so,

$$|\mathcal{U}(t,\mu)-\mathcal{U}(s,\mu)|\le\Big|\int_s^t L(\sigma^{\mu,t}_\tau,\mathbf{v}^{\mu,t}_\tau)d\tau\Big|+|\mathcal{U}(s,\sigma^{\mu,t}_s)-\mathcal{U}(s,\mu)|. \tag{6.42}$$

We use the fact that $\mathcal{U}(s,\cdot)$ is $(1+s)\kappa$–Lipschitz and Remark 1 to obtain

$$|\mathcal{U}(s,\sigma^{\mu,t}_s)-\mathcal{U}(s,\mu)|\le W_2(\sigma^{\mu,t}_s,\mu)\le\int_s^t||\mathbf{v}^{\mu,t}_\tau||_{\sigma^{\mu,t}_\tau}d\tau\le s_1|t-s|. \tag{6.43}$$

We combine (6.42) and (6.43) to conclude that

$$|\mathcal{U}(t,\mu)-\mathcal{U}(s,\mu)|\le s_1|t-s|+(1+s)\kappa s_1|t-s|\le s_1|t-s|+(1+T)\kappa s_1|t-s|.$$

□

6.4. **Hamilton Jacobi equations.** Let $V,W\in C^1(\mathbb{R}^d)$ be such that there exist $\beta\in[1,2)$ and $C_0>0$ such that

$$4|W(z)|,\ 2|V(z)|\le C_0(|z|^{2-\epsilon}+1) \tag{6.44}$$

and assume

$$\mathcal{W}(\mu)=\int_{\mathbb{R}^d}\Big(V(x)+W*\mu\Big)\mu(dx).$$

In this subsection we consider viscosity solutions of the equation

$$\partial_t\mathcal{U}+H(\mu,\nabla_\mu\mathcal{U})=0,\qquad \mathcal{U}(0,\cdot)=\mathcal{U}_0. \tag{6.45}$$

Definition 6.2. Let $\mathcal{U}:\mathcal{P}_2(M)\to\mathbb{R}\cup\{\pm\infty\}$, let $\mu\in\mathcal{P}_2(M)$ and let $\xi\in T_\mu\mathcal{P}_2(M)$ (cf. Section 2).

(i) We say that ξ is in the subdifferential of $\mathcal{U}$ at μ and we write $\xi\in\partial_\cdot\mathcal{U}(\mu)$ if

$$\mathcal{U}(\nu)-\mathcal{U}(\mu)\ge\sup_{\gamma\in\Gamma_o(\mu,\nu)}\int_{M\times M}\xi(q)\cdot(r-q)\gamma(dq,dr)+o\big(W_2(\mu,\nu)\big)\quad\forall\nu\in\mathcal{P}_2(M). \tag{6.46}$$

(ii) We say that ξ is in the superdifferential of $\mathcal{U}$ at μ and we write $\xi\in\partial^\cdot\mathcal{U}(\mu)$ if $-\xi\in\partial_\cdot(-\mathcal{U})(\mu)$.

(iii) When $\partial_\cdot\mathcal{U}(\mu)$ and $\partial^\cdot\mathcal{U}(\mu)$ are both nonempty then they are equal and reduce to a single element (cf. e.g. [41]) which we denote by $\nabla_\mu\mathcal{U}(\mu)$, and refer to as the Wasserstein gradient of $\mathcal{U}$.

Definition 6.3. Let $T>0$ and let $\mathcal{U}:[0,T)\times\mathcal{P}_2(M)\to\mathbb{R}$.

(i) We say that $\mathcal{U}$ is a viscosity subsolution for (6.45) if $\mathcal{U}$ is upper semicontinuous on $[0,T)\times\mathcal{P}_2(M)$, if for all $(t,\mu)\in(0,T)\times\mathcal{P}_2(M)$ and all $\theta,\zeta)\in\partial^\cdot\mathcal{U}(t,\mu)$

$$\mathcal{U}(\cdot,0)\le\mathcal{U}_0,\ \text{and}\ \theta+H(\mu,\zeta)\le 0. \tag{6.47}$$

(ii) We say that $\mathcal{U}$ is a viscosity supersolution for (6.45) if $\mathcal{U}$ is lower semicontinuous on $[0,T)\times\mathcal{P}_2(M)$, if for all $(t,\mu)\in(0,T)\times\mathcal{P}_2(M)$ and all $(\theta,\zeta)\in\partial\mathcal{U}(t,\mu)$

$$\mathcal{U}(\cdot,0)\geq\mathcal{U}_0,\ \text{ and } \theta+H(\zeta,\mu)\geq 0. \tag{6.48}$$

(iii) We say that $\mathcal{U}$ is a viscosity solution for (6.45) if $\mathcal{U}$ is both a viscosity subsolution and a viscosity supersolution.

Denote by $\mathcal{L}^d$ the Lebesgue measure on $(0,1)^d$. Given $f\in L^2\big((0,1)^d\big)$ we set

$$\bar{U}_0(f)=\mathcal{U}_0(f_{\#}\mathcal{L}^d)$$

Theorem 6.4. *Suppose $\mathcal{U}_0:\mathcal{P}_2(M)\to\mathbb{R}$ is bounded below and lower semicontinuous for the narrow convergence. Let $\mathcal{U}$ be the value function in Equation (6.1). Then:*

(i) The infimum in (6.1) is a minimum.
(ii) $\mathcal{U}$ is a viscosity subsolution of Equation (6.45).
(iii) Suppose $d=1$, $\bar{U}_0$ is Frechet differentiable and λ–convex for some $\lambda\in\mathbb{R}$ and $T\lambda^-<1$. We assume that the gradient of $\bar{U}_0$ is a continuous map of the Hilbert space $L^2\big((0,1)^d\big)$ into itself. Then $\mathcal{U}$ is a viscosity solution of Equation (6.45).

Proof. (i) It suffices to verify that the assumptions of Proposition 4 are satisfied. Only (6.5) remains to be checked. However, in fact a statement stronger which we need in the proof of (ii), can be made. Indeed, By (6.44) and by the fact that $\beta<2$, $\mathcal{W}$ is bounded from below on bounded subsets of $\mathcal{P}_2(M)$ and

$$\lim_{n\to\infty}\mathcal{W}(\mu^n)=\mathcal{W}(\mu)$$

whenever $\{\mu^n\}_n\subset\mathcal{P}_2(M)$ is a bounded sequence that converges narrowly to μ. In particular, $\mathcal{W}$ is continuous.

(ii) Inequality (6.44) yields (6.6). Since $\beta<2$ we obtain the existence of $e_0,e_1>0$ such that $8e_0T^2<\pi^2$ and

$$\mathcal{W}(\nu)\leq e_0\int_M|x|^2\nu(dx)+e_1$$

for all $\nu\in\mathcal{P}_2(M)$. We apply Theorem 3.9-(i) of [41] to conclude the proof of (ii).

(iii) Corollary 5.3 of [41] yields (iii). □

Remark 11. We learned from R. Hynd and H-K. Kim that when $d\geq 1$ and $W\equiv 0$, the value function in Theorem 6.4 is a viscosity solution of Equation (6.45) [53].

7. Metric viscosity solutions.

In this section we want to show that with little effort one can define a notion of a metric viscosity solution, based on local slopes, for a class of Hamilton–Jacobi equations that only depend on the "length" of the gradient variable. We present one possible definition but the readers should be free to experiment with it by possibly choosing different sets of test functions or by interpreting some terms differently. This section was motivated by [8, 48]. We do not know if the results here are completely new. N. Gigli mentioned to the second author a year ago that he had a notion of a viscosity solution for which he was able to show uniqueness. The second author was also told that L. Ambrosio and J. Feng are working on a notion of viscosity solution for similar equations and obtained existence and uniqueness results [4].

7.1. **Definition and comparison.** Let $(\mathbb{S}, d)$ be a complete metric space which is a geodesic space. By this we mean that for every $x, y \in \mathbb{S}$ there exists a geodesic of constant speed $x_t, 0 \le t \le 1$, connecting x and y, i.e. a curve such that

$$x_0 = y, x_1 = x,\ d(x_s, x_t) = |s-t| d(x,y),\ 0 \le t \le s \le 1.$$

Let $T > 0$. We consider an equation

$$\begin{cases} \partial_t u + H(t, x, |\nabla u|) = 0, & \text{in } (0,T) \times \mathbb{S}, \\ u(0,x) = g(x) & \text{on } \mathbb{S}, \end{cases} \tag{7.1}$$

where $H : [0,T] \times \mathbb{S} \times [0, +\infty) \to \mathbb{R}$ is continuous, and $|\nabla u|$ is the local slope of u. Let $x_0 \in \mathbb{S}$ be a fixed point.

Following [7, 8, 48, 64], for $v : (0,T) \times \mathbb{S} \to \mathbb{R}$ we define the upper and lower local slopes of v

$$|\nabla^+ v(t,x)| := \limsup_{y \to x} \frac{[v(t,y) - v(t,x)]_+}{d(y,x)}, |\nabla^- v(t,x)| := \limsup_{y \to x} \frac{[v(t,y) - v(t,x)]_-}{d(y,x)}, \tag{7.2}$$

and its local slope

$$|\nabla v(t,x)| := \limsup_{y \to x} \frac{|v(t,y) - v(t,x)|}{d(y,x)}.$$

It is easy to see that $|\nabla^- v| = |\nabla^+(-v)|$. We also define

$$|\nabla v(t,x)|^* = \limsup_{(s,y) \to (t,x)} |\nabla v(s,y)|.$$

Equation (7.1) must be interpreted in a proper viscosity sense. We first define a class of test functions.

Definition 7.1. A function $\psi : (0,T) \times \mathbb{S} \to \mathbb{R}$ is a subsolution test function ($\psi \in \underline{\mathcal{C}}$) if $\psi(t,x) = \psi_1(t,x) + \psi_2(t,x)$, where ψ_1, ψ_2 are Lipschitz on every bounded and closed subset of $(0,T) \times \mathbb{S}$, $|\nabla \psi_1(t,x)| = |\nabla^- \psi_1(t,x)|$ is continuous, and $\partial_t \psi_1, \partial_t \psi_2$ are continuous. A function $\psi : (0,T) \times \mathbb{S} \to \mathbb{R}$ is a supersolution test function ($\psi \in \overline{\mathcal{C}}$) if $-\psi \in \underline{\mathcal{C}}$.

Lemma 7.2. *Let $\psi_1(t,x) = k(t) + k_1(t)\varphi(d^2(x,y))$, where $y \in \mathbb{S}$, $\varphi \in C^1([0,+\infty))$, $\varphi' \ge 0$, $k, k_1 \in C^1((0,T)), k_1 \ge 0$. Then*

$$|\nabla^- \psi_1(t,x)| = |\nabla \psi_1(t,x)| = 2k_1(t)\varphi'(d^2(x,y))d(x,y).$$

In particular $|\nabla \psi_1(t,x)|$ is continuous and thus the function can be used as the ψ_1 part of a test function.

Proof. We have

$$\psi_1(t,z) - \psi_1(t,x) = k_1(t)\varphi'(d^2(x,y))(d^2(z,y) - d^2(x,y)) + o(d^2(z,y) - d^2(x,y)).$$

Therefore by triangle inequality

$$\begin{aligned} |\nabla \psi_1(t,x)| &\le \limsup_{z \to x} k_1(t)\varphi'(d^2(x,y)) \frac{2d(z,x)d(x,y) + d^2(z,x)}{d(z,x)} \\ &= 2k_1(t)\varphi'(d^2(x,y))d(x,y). \end{aligned}$$

Let $x_s, 0 \le t \le 1$ be a geodesic of constant speed connecting x and y, i.e. a curve such that $x_0 = y, x_1 = x, d(x_s, x_\tau) = |s - \tau| d(x,y)$. Then $d(x_s, y) =$

$sd(x,y), d(x_s,x) = (1-s)d(x,y)$. Then

$$\begin{aligned} |\nabla^-\psi_1(t,x)| &\geq \limsup_{s\to 1} k_1(t)\varphi'(d^2(x,y))\frac{d^2(x,y)-d^2(x_s,y)}{d(x_s,x)} \\ &= \lim_{s\to 1} k_1(t)\varphi'(d^2(x,y))d(x,y)\frac{1-s^2}{1-s} \\ &= 2k_1(t)\varphi'(d^2(x,y))d(x,y). \end{aligned}$$

This proves the claim since $|\nabla^-\psi_1(t,x)| \leq |\nabla\psi_1(t,x)|$. □

Remark 12. Our choice of test functions is rather arbitrary. All of the results would still be true if we restricted the class of test functions so that we had enough test functions to prove comparison principle. In particular we could take the ψ_1 part of test functions to be composed of the functions from Lemma 7.2.

We define for $r \geq 0$

$$H_r(t,x,s) := \inf_{|\tau - s|\leq r} H(t,x,\tau), \quad H^r(t,x,s) := \sup_{|\tau - s|\leq r} H(t,x,\tau).$$

Definition 7.3. An upper semicontinuous function $u : [0,T)\times\mathbb{S} \to \mathbb{R}$ is a metric viscosity subsolution of (7.1) if $u(0,x) \leq g(x)$ on $\mathbb{S}$, and whenever $u-\psi$ has a local maximum at (t,x) for some $\psi \in \underline{\mathcal{C}}$, then

$$\partial_t\psi(t,x) + H_{|\nabla\psi_2(t,x)|^*}(t,x,|\nabla\psi_1(t,x)|) \leq 0. \tag{7.3}$$

A lower semicontinuous function $u : [0,T)\times\mathbb{S} \to \mathbb{R}$ is a metric viscosity supersolution of (7.1) if $u(0,x) \geq g(x)$ on X, and whenever $u-\psi$ has a local minimum at (t,x) for some $\psi \in \overline{\mathcal{C}}$, then

$$\partial_t\psi(t,x) + H^{|\nabla\psi_2(t,x)|^*}(t,x,|\nabla\psi_1(t,x)|) \geq 0. \tag{7.4}$$

A continuous function $u : [0,T)\times\mathbb{S} \to \mathbb{R}$ is a metric viscosity solution of (7.1) if it is both a metric viscosity subsolution and a metric viscosity supersolution of (7.1).

Remark 13. We stated the definition of viscosity solution for equations defined in the whole space, however we can define metric viscosity subsolutions/supersolutions in any open subset Q of $(0,T)\times\mathbb{S}$ by requiring that (7.3)/(7.3) be satisfied whenever a local maximum/minimum is in Q. Initial condition is disregarded in such cases. The definition can also be applied in an obvious way to stationary equations $H(x,u,|\nabla u|) = 0$.

We recall a variational principle of Borwein-Preiss (see [18], Theorem 2.6 and Remark 2.7 about the result in a metric space) formulated in a form suitable for us. It can be obtained following the proof of Theorem 2.6 of [18] using the metric

$$\overline{d}((t,s,x,y),(t',s',x',y')) = (|t-t'|^2 + |s-s'|^2 + d^2(x,x') + d^2(y,y'))^{\frac{1}{2}}.$$

We remark that it would be enough for our purposes to use a version of Ekeland's variational principle but the perturbation function from Lemma 7.4 is more regular. Lemma 7.4 was also used in [36].

Lemma 7.4. *Let $\Phi : [0,T]\times[0,T]\times\mathbb{S}\times\mathbb{S} \to [-\infty,+\infty)$ be upper semicontinuous and bounded from above. Let for $n \geq 1$, $(\hat{t}_n,\hat{s}_n,\hat{x}_n,\hat{y}_n)$ be such that*

$$\Phi(\hat{t}_n,\hat{s}_n,\hat{x}_n,\hat{y}_n) > \sup\Phi - \frac{1}{n}.$$

Then there exist sequences x_k^n, y_k^n *such that* $d(x_k^n, \hat{x}_n) \le 1, d(y_k^n, \hat{y}_n) \le 1, k \ge 1$, *points* $\bar{t}_n, \bar{s}_n \in [0,T], \bar{x}_n, \bar{y}_n \in \mathbb{S}$, *such that* $(x_k^n, y_k^n) \to (\bar{x}_n, \bar{y}_n)$, *sequences of non-negative numbers* β_k^n *such that* $\sum_{k=1}^{+\infty} \beta_k^n = 1$, *and quadratic polynomials* $p_1^n, p_2^n \ge 0$ *with* $|(p_1^n)'(\bar{t}_n)| \le 4/n, |(p_2^n)'(\bar{s}_n)| \le 4/n$, *such that*

$$\Phi(\bar{t}_n, \bar{s}_n, \bar{x}_n, \bar{y}_n) > \sup \Phi - \frac{1}{n}$$

and

$$\Phi(\bar{t}_n, \bar{s}_n, \bar{x}_n, \bar{y}_n) - \frac{1}{n}\sum_{k=1}^{\infty} \beta_k^n (d^2(\bar{x}_n, x_k^n) + d^2(\bar{y}_n, y_k^n)) - p_1^n(\bar{t}_n) - p_2^n(\bar{s}_n)$$
$$\ge \Phi(t,s,x,y) - \frac{1}{n}\sum_{k=1}^{\infty} \beta_k^n (d^2(x, x_k^n) + d^2(y, y_k^n)) - p_1^n(t) - p_2^n(s).$$

for all $(t,s,x,y) \in [0,T] \times [0,T] \times \mathbb{S} \times \mathbb{S}$.

From now on we will restrict our attention to equations

$$\begin{cases} \partial_t u + H(|\nabla u|) + f(x) = 0, & \text{in } (0,T) \times \mathbb{S}, \\ u(0,x) = g(x) & \text{on } \mathbb{S}. \end{cases} \tag{7.5}$$

We assume that $H : [0, +\infty) \to \mathbb{R}$ is continuous and

$$f : \mathbb{S} \to \mathbb{R}, \quad g : \mathbb{S} \to \mathbb{R}$$

are uniformly continuous, i.e. there exists a modulus ω such that

$$|f(x) - f(y)| + |g(x) - g(y)| \le \omega(d(x,y)) \quad \text{for } x, y \in \mathbb{S}. \tag{7.6}$$

We could assume that f also depends on t but we do not do so for simplicity.

We will only present the proof of comparison for equation (7.5) since it is the most relevant for the class of Hamilton-Jacobi equations studied in this paper and since we do not want to make any assumptions about the growth and continuity of H. Once the basic techniques are in place the proof is not much different from typical viscosity proofs in finite dimensions [23] or in Hilbert spaces and can be modified to general equations (7.1) under typical assumptions on H and growth conditions for sub- and supersolutions. The proof would be much easier if $\mathbb{S}$ was compact (or locally compact) since we could avoid the use of Lemma 7.4.

Theorem 7.5. *Let (7.6) hold and H be continuous. Let u be a metric viscosity subsolution of (7.5) and v be a metric viscosity supersolution of (7.5) satisfying*

$$|u(t,x)| + |v(t,x)| \le K(1 + d(x_0, x)) \tag{7.7}$$

for some $K \ge 0$, *and*

$$\lim_{t \to 0} ([u(t,x) - g(x)]_+ + [v(t,x) - g(x)]_-) = 0 \quad \textit{uniformly on bounded sets of } \mathbb{S}. \tag{7.8}$$

Then $u \le v$.

Proof. We first notice that the functions $u_1(t,x) = e^{-t}u(t,x), v_1(t,x) = e^{-t}v(t,x)$ are respectively a viscosity subsolution and a viscosity supersolution of the equation

$$\begin{cases} \partial_t u + u + e^{-t} H(e^t |\nabla u|) + e^{-t} f(x) = 0 \\ u(0,x) = g(x). \end{cases} \tag{7.9}$$

Let $L > 0$ be such that $\omega(s) \le 1 + Ls$. For $0 < \mu < 1$ we define

$$u_\mu(t,x) = u_1(t,x) - \frac{\mu}{T-t}, \quad v_\mu(s,y) = v_1(s,y) + \frac{\mu}{T-s}.$$

Step 1. We will first show that for every μ

$$\lim_{R\to+\infty}\lim_{r\to 0}\sup_{t,s,x,y}\{u_\mu(t,x)-v_\mu(s,y)-2Ld(x,y):|t-s|<r,$$
$$d(x_0,x)+d(x_0,y)<R\}<+\infty. \tag{7.10}$$

Let $\gamma_R\in C^1([0,+\infty)),\gamma_R\geq 0,\gamma_R'\geq 0,R\geq 1$, be a family of functions such that

$$\liminf_{r\to\infty}\frac{\gamma_R(r)}{r}\geq 3K \quad \text{for every } R\geq 1, \tag{7.11}$$

$$|\gamma_R'(r)|\leq C \quad \text{for all } R\geq 1, r\in[0,+\infty), \tag{7.12}$$

$$\gamma_R(r)=0 \quad \text{for } r\in[0,R], R\geq 1. \tag{7.13}$$

For $R\geq 1,\beta>0,\mu>0$ we define the function

$$\begin{aligned}\Phi_{R,\beta}(t,s,x,y) &= u_\mu(t,x)-v_\mu(s,y)-2L(1+d^2(x,y))^{\frac{1}{2}}\\ &- \gamma_R(d(x_0,x))-\gamma_R(d(x_0,y))-\frac{(t-s)^2}{2\beta}.\end{aligned}$$

The function Φ is upper semicontinuous on $[0,T]\times[0,T]\times\mathbb{S}\times\mathbb{S}$ and, by (7.7), (7.11), is bounded from above. If (7.10) is not satisfied, then (7.13) implies that for every n there exist $R_n,(t_n^i,s_n^i,x_n^i,y_n^i)$ such that $d(x_0,x_n^i)+d(x_0,y_n^i)<R_n,|t_n^i-s_n^i|\to 0$ as $i\to+\infty$, and $u_\mu(t_n^i,x_n^i)-v_\mu(s_n^i,y_n^i)-2Ld(x_n^i,y_n^i)\geq n$. Thus for every $\beta>0,n$, $\limsup_{i\to+\infty}\Phi_{R_n,\beta}(t_n^i,s_n^i,x_n^i,y_n^i)\geq n-2L$, and thus

$$\lim_{R\to+\infty}\limsup_{\beta\to 0}\sup\Phi_{R,\beta}\geq\lim_{n\to+\infty}\limsup_{i\to+\infty}\Phi_{R_n,\beta}(t_n^i,s_n^i,x_n^i,y_n^i)=\infty. \tag{7.14}$$

Therefore, Lemma 7.4 applied with $n=1$ implies that for large R, there exist β_k,x_k,y_k,p_1,p_2 satisfying conditions of Lemma 7.4 such that

$$\Phi_{R,\beta}(t,s,x,y)-\sum_{k=1}^{\infty}\beta_k(d^2(x,x_k)+d^2(y,y_k))-p_1(t)-p_2(s)$$

has a maximum at a point $(\bar t,\bar s,\bar x,\bar y)$ such that

$$\Phi_{R,\beta}(\bar t,\bar s,\bar x,\bar y)>\sup\Phi_{R,\beta}-1, \tag{7.15}$$

and hence

$$u_1(\bar t,\bar x)-v_1(\bar s,\bar y)\geq u_\mu(\bar t,\bar x)-v_\mu(\bar s,\bar y)\geq 2Ld(\bar x,\bar y). \tag{7.16}$$

It follows from (7.14) and (7.15) that

$$\limsup_{R\to+\infty}\limsup_{\beta\to 0}\Phi_{R,\beta}(\bar t,\bar s,\bar x,\bar y)=+\infty. \tag{7.17}$$

We also notice that since $\tilde\Phi_{R,\beta}$ is bounded by a constant depending on R,

$$\frac{(\bar t-\bar s)^2}{2\beta}\leq C_R$$

for some constant C_R.

Therefore, when (7.15) holds, it is easy to see from (7.6) and (7.8), that $0<\bar t,\bar s<T$ for sufficiently small β. Therefore using the definition of viscosity solution we have

$$u_1(\bar t,\bar x)+\frac{\bar t-\bar s}{\beta}+\frac{\mu}{(T-\bar t)^2}+p_1'(\bar t)+e^{-\bar t}H_{e^{\bar t}|\nabla\psi_2(\bar x)|^*}(e^{\bar t}|\nabla\psi_1(\bar x)|)+e^{-\bar t}f(\bar x)\leq 0,$$

$$v_1(\bar s,\bar y)+\frac{\bar t-\bar s}{\beta}-\frac{\mu}{(T-\bar s)^2}-p_2'(\bar s)++e^{-\bar s}H^{e^{\bar s}|\nabla\tilde\psi_2(\bar y)|^*}(e^{\bar s}|\nabla\tilde\psi_1(\bar y)|)+e^{-\bar s}f(\bar y)\geq 0,$$

where

$$\psi_1(x) = 2L(1+d^2(x,\bar{y}))^{\frac{1}{2}}, \quad \psi_2(x) = \gamma_R(d(x_0,x)) + \sum_{k=1}^{\infty} \beta_k d^2(x,x_k),$$

$$\tilde{\psi}_1(y) = -2L(1+d^2(\bar{x},y))^{\frac{1}{2}}, \quad \tilde{\psi}_2(y) = -\gamma_R(d(x_0,y)) - \sum_{k=1}^{\infty} \beta_k d^2(y,y_k).$$

The function ψ_1 is globally Lipschitz and since from Lemma 7.4 we have $d(\bar{x},x_k) \le 2$ for all k, it is easy to see that $|\nabla\psi_1(\bar{x})| + |\nabla\psi_2(\bar{x})|^* \le C$ for some C independent of R,μ,β. Similarly we have $|\nabla\tilde{\psi}_1(\bar{y})| + |\nabla\tilde{\psi}_2(\bar{y})|^* \le C$.

Using the continuity of H we thus obtain

$$u_1(\bar{t},\bar{x}) - v_1(\bar{s},\bar{y}) + e^{-\bar{t}} f(\bar{x}) - e^{-\bar{s}} f(\bar{y}) \le C_1,$$

where C_1 is independent of R,μ,β. It thus follows from (7.6), and (7.16) that

$$\begin{aligned} u_1(\bar{t},\bar{x}) - v_1(\bar{s},\bar{y}) &\le 1 + C_1 + (e^{-\bar{s}} - e^{-\bar{t}}) f(\bar{y}) + L d(\bar{x},\bar{y}) \\ &\le 1 + C_1 + (e^{-\bar{s}} - e^{-\bar{t}}) f(\bar{y}) + \frac{1}{2}(u_1(\bar{t},\bar{x}) - v_1(\bar{s},\bar{y})), \end{aligned}$$

and hence

$$u_1(\bar{t},\bar{x}) - v_1(\bar{s},\bar{y}) \le 2(1+C_1) + 2(e^{-\bar{s}} - e^{-\bar{t}}) f(\bar{y}).$$

Therefore,

$$\Phi(\bar{t},\bar{s},\bar{x},\bar{y}) \le u_1(\bar{t},\bar{x}) - v_1(\bar{s},\bar{y}) \le 2(1+C_1) + 2(e^{-\bar{s}} - e^{-\bar{t}}) f(\bar{y}),$$

which, noticing that for fixed μ, R, the distances $d(x_0,\bar{y})$ remain bounded, implies

$$\limsup_{R\to\infty} \limsup_{\beta\to 0} \Phi(\bar{t},\bar{s},\bar{x},\bar{y}) \le 2(1+C_1),$$

which contradicts (7.17).

Step 2. Suppose that $u_1(\tilde{t},\tilde{x}) - v_1(\tilde{t},\tilde{x}) > 2\nu$ for some $\nu > 0$ and $\tilde{t},\tilde{x}$. Then the function

$$\Psi(t,s,x,y) := u_\mu(t,x) - v_\mu(s,y) - \frac{d^2(x,y)}{2\epsilon} - \delta(d^2(x_0,x) + d^2(x_0,y)) - \frac{(t-s)^2}{2\beta}$$

is upper semicontinuous on $[0,T]\times[0,T]\times\mathbb{S}\times\mathbb{S}$ and bounded from above. Define

$$m_{\mu,\epsilon,\delta,\beta} := \sup_{t,s,x,y} \Psi(t,s,x,y).$$

We have $m_{\mu,\epsilon,\delta,\beta} > 3\nu/2$ for small $\mu,\epsilon,\delta,\beta > 0$. Thus, for small $\mu,\epsilon,\delta,\beta > 0$ and large n there exist $\beta_k^n, x_k^n, y_k^n, p_1^n, p_2^n$ as in Lemma 7.4 such that

$$\Psi(t,s,x,y) - \frac{1}{n}\sum_{k=1}^{\infty} \beta_k^n (d^2(x,x_k^n) + d^2(y,y_k^n)) - p_1^n(t) - p_2^n(s)$$

has a maximum at a point $(\bar{t},\bar{s},\bar{x},\bar{y})$ such that

$$\Psi(\bar{t},\bar{s},\bar{x},\bar{y}) \ge m_{\mu,\epsilon,\delta,\beta} - \frac{1}{n} \ge \nu. \tag{7.18}$$

Defining

$$m_{\mu,\epsilon,\delta} := \lim_{r\to 0}\{\tilde{\Psi}(t,s,x,y) : |t-s| < r\}, \text{where } \tilde{\Psi}(t,s,x,y) = \Psi(t,s,x,y) + \frac{(t-s)^2}{2\beta}.$$

we claim that

$$m_{\mu,\epsilon,\delta} = \lim_{\beta\to 0} m_{\mu,\epsilon,\delta,\beta}. \tag{7.19}$$

To see this let (t_r, s_r, x_r, y_r) be such that $|t_r - s_r| < r$ and

$$m_{\mu,\epsilon,\delta} = \lim_{r\to 0} \tilde{\Psi}(t_r, s_r, x_r, y_r).$$

Then for every $\beta > 0$

$$\lim_{r\to 0} \tilde{\Psi}(t_r, s_r, x_r, y_r) = \lim_{r\to 0} \Psi(t_r, s_r, x_r, y_r) \le m_{\mu,\epsilon,\delta,\beta},$$

which implies

$$m_{\mu,\epsilon,\delta} \le \lim_{\beta\to 0} m_{\mu,\epsilon,\delta,\beta}.$$

Now let $(t_\beta, s_\beta, x_\beta, y_\beta)$ be such that

$$m_{\mu,\epsilon,\delta,\beta} < \Psi(t_\beta, s_\beta, x_\beta, y_\beta) + \beta \le \tilde{\Psi}(t_\beta, s_\beta, x_\beta, y_\beta) + \beta.$$

Since $\tilde{\Psi}$ is bounded by a constant depending on R, there is a constant $\tilde{C}_R$ such that

$$\frac{(t_\beta - s_\beta)^2}{2\beta} \le \tilde{C}_R.$$

This implies

$$m_{\mu,\epsilon,\delta} < \sup\{\tilde{\Psi}(t, s, x, y) : |t - s| \le (2\tilde{C}_R\beta)^{\frac{1}{2}}\} + \beta.$$

Letting $\beta \to 0$ above it thus follows that

$$m_{\mu,\epsilon,\delta} \ge \lim_{\beta\to 0} m_{\mu,\epsilon,\delta,\beta}$$

which completes the proof of the claim.

Now

$$m_{\mu,\epsilon,\delta,\beta} \le \Psi(\bar{t}, \bar{s}, \bar{x}, \bar{y}) + \frac{1}{n}$$

and thus

$$m_{\mu,\epsilon,\delta,\beta} + \frac{(\bar{t} - \bar{s})^2}{4\beta} \le \Psi(\bar{t}, \bar{s}, \bar{x}, \bar{y}) + \frac{1}{n} + \frac{(\bar{t} - \bar{s})^2}{4\beta} \le m_{\mu,\epsilon,\delta,2\beta} + \frac{1}{n}.$$

This implies

$$\lim_{\beta\to 0} \limsup_{n\to+\infty} \frac{(\bar{t} - \bar{s})^2}{\beta} = 0 \quad \text{for fixed } \mu, \epsilon, \delta. \tag{7.20}$$

By (7.7) we also have

$$d(x_0, \bar{x}) + d(x_0, \bar{y}) \le R_\delta \quad \text{for fixed } \mu, \epsilon, \tag{7.21}$$

for some $R_\delta > 0$. Therefore, by (7.6), and (7.20), for sufficiently small $\mu, \epsilon, \delta, \beta$, we must have $0 < \bar{t}, \bar{s} < T$. Now, by (7.18),

$$\frac{d^2(\bar{x}, \bar{y})}{2\epsilon} + \delta(d^2(x_0, \bar{x}) + d^2(x_0, \bar{y})) + \frac{(\bar{t} - \bar{s})^2}{2\beta} \le u_\mu(\bar{t}, \bar{x}) - v_\mu(\bar{s}, \bar{y})$$

and thus, taking $\limsup_{\beta\to 0} \limsup_{n\to+\infty}$ above and using (7.10), (7.20) and (7.21), we obtain for every μ, ϵ, δ

$$\begin{aligned}
&\limsup_{\beta\to 0} \limsup_{n\to\infty} \frac{d^2(\bar{x}, \bar{y})}{2\epsilon} + \delta(d^2(x_0, \bar{x}) + d^2(x_0, \bar{y})) \\
\le\ &\limsup_{\beta\to 0} \limsup_{n\to\infty} (u_\mu(\bar{t}, \bar{x}) - v_\mu(\bar{s}, \bar{y})) \\
\le\ &\limsup_{\beta\to 0} \limsup_{n\to\infty} 2Ld(\bar{x}, \bar{y}) + C_2 \\
\le\ &\limsup_{\beta\to 0} \limsup_{n\to\infty} \frac{d^2(\bar{x}, \bar{y})}{4\epsilon} + C_3,
\end{aligned}$$

where C_2, C_3 may depend on μ. This in particular implies that

$$\limsup_{\beta\to 0}\limsup_{n\to+\infty}\frac{d(\bar{x},\bar{y})}{\epsilon}\le 2\left(\frac{C_3}{\epsilon}\right)^{\frac{1}{2}}, \tag{7.22}$$

$$\delta(d(x_0,\bar{x})+d(x_0,\bar{y}))\le C_\mu\sqrt{\delta} \tag{7.23}$$

for some constant C_μ. Using the definition of viscosity solution and Lemma 7.2 we obtain

$$u_1(\bar{t},\bar{x})+\frac{\bar{t}-\bar{s}}{\beta}+\frac{\mu}{(T-\bar{t})^2}+(p_1^n)'(\bar{t})+e^{-\bar{t}}H_{e^{\bar{t}}|\nabla\psi_2(\bar{x})|^*}\left(e^{\bar{t}}\frac{d(\bar{x},\bar{y})}{\epsilon}\right)+e^{-\bar{t}}f(\bar{x})\le 0,$$

$$v_1(\bar{s},\bar{y})+\frac{\bar{t}-\bar{s}}{\beta}-\frac{\mu}{(T-\bar{s})^2}-(p_2^n)'(\bar{s})+e^{-\bar{s}}H^{e^{\bar{s}}|\nabla\tilde{\psi}_2(\bar{y})|^*}\left(e^{\bar{s}}\frac{d(\bar{x},\bar{y})}{\epsilon}\right)+e^{-\bar{s}}f(\bar{y})\ge 0,$$

where

$$\psi_2(x)=\delta d^2(x_0,x)+\frac{1}{n}\sum_{k=1}^{\infty}\beta_k^n d^2(x,x_k^n),$$

$$\tilde{\psi}_1(y)=-\delta d^2(x_0,y)-\frac{1}{n}\sum_{k=1}^{\infty}\beta_k^n d^2(y,y_k^n).$$

In particular, (7.21) and (7.23) give

$$\limsup_{\delta\to 0}\limsup_{\beta\to 0}\limsup_{n\to+\infty}(|\nabla\psi_2(\bar{x})|^*+|\nabla\tilde{\psi}_2(\bar{y})|^*)=0.$$

We now subtract the above inequalities, use the continuity of H, and (7.20), (7.21), (7.23), (7.22) to get

$$e^{-\bar{t}}H\left(e^{\bar{t}}\frac{d(\bar{x},\bar{y})}{\epsilon}\right)-e^{-\bar{t}}H\left(e^{\bar{t}}\frac{d(\bar{x},\bar{y})}{\epsilon}\right)+e^{-\bar{t}}f(\bar{x})-e^{-\bar{t}}f(\bar{y})\le-\frac{2\mu}{T^2}+\sigma(\delta,\beta,n),$$

where $\limsup_{\delta\to 0}\limsup_{\beta\to 0}\limsup_{n\to+\infty}\sigma(\delta,\beta,n)=0$ for fixed μ,ϵ. It remains to take

$$\limsup_{\epsilon\to 0}\limsup_{\delta\to 0}\limsup_{\beta\to 0}\limsup_{n\to+\infty}$$

in the above inequality and use (7.6), (7.22) to obtain a contradiction. □

Corollary 2. *Let u be a metric viscosity subsolution of*

$$\begin{cases} \partial_t u+H(|\nabla u|)+f_1(x)=0\\ u(0,x)=g_1(x),\end{cases} \tag{7.24}$$

and v be a metric viscosity supersolution of

$$\begin{cases} \partial_t v+H(|\nabla v|)+f_2(x)=0\\ v(0,x)=g_2(x),\end{cases} \tag{7.25}$$

where f_1, g_1, f_2, g_2 satisfy (7.6), H is continuous, and u, v satisfy (7.7) and (7.8) with g_1 and g_2 respectively. Then

$$u-v\le\sup_x\{g_1(x)-g_2(x)\}+t\sup_x\{f_2(x)-f_1(x)\}. \tag{7.26}$$

Proof. The result follows from Theorem 7.5 upon noticing that the function

$$v_1(t,x)=v(t,x)+\sup_x\{g_1(x)-g_2(x)\}+t\sup_x\{f_2(x)-f_1(x)\}$$

is a viscosity supersolution of (7.24). □

It is easy to see that the notion of metric viscosity solution has good limiting properties. In particular it is stable with respect to uniform limits. Moreover, if the metric space $\mathbb{S}$ is locally compact, the method of half-relaxed limits of Barles-Perthame (see [23]) also works for it.

7.2. **Existence of solutions.** We first show that a version of Perron's method can be applied to produce a viscosity solution of (7.5) without any additional restrictions on H. Let us first recall that the upper semicontinuous envelope of a function f is denoted by f^* and is the least upper semicontinuous function which is greater than or equal to f. Similarly, the lower semicontinuous envelope of a function f is denoted by f_* and is the largest lower semicontinuous function which is less than or equal to f. We say that a function f has a strict maximum at (t,x) over a set $A \subset [0,T] \times \mathbb{S}$ if $f(s,y) \le f(t,x)$ for all $(s,y) \in A$ and whenever (t_n,x_n) is a sequence in A such that $f(t_n,x_n) \to f(t,x)$ then $(t_n,x_n) \to (t,x)$. Strict minimum is defined similarly.

Theorem 7.6. *Let (7.6) hold and H be continuous. Let $\underline{u}$ be a metric viscosity subsolution of (7.5) and $\overline{v}$ be a metric viscosity supersolution of (7.5) satisfying (7.7),*

$$\lim_{t\to 0} \left([\overline{v}^*(t,x) - g(x)]_+ + [\underline{u}_*(t,x) - g(x)]_-\right) = 0 \quad \textit{uniformly on bounded sets of } \mathbb{S}, \tag{7.27}$$

and $\underline{u} \le \overline{v}$. Denote

$$\mathcal{S} := \{w : \underline{u} \le w \le \overline{v},\ w \textit{ is a metric viscosity subsolution of } (7.5)\}.$$

Then

$$v := \sup_{w\in\mathcal{S}} w$$

is a metric viscosity solution of (7.5).

Proof. **Step 1.** Suppose that $v^* - \psi$ has a maximum at a point (t,x) over some set $A = \{(s,y) : |t-s|^2 + d^2(y,x) \le \eta$ for some $\eta > 0$ and $\psi = \psi_1 + \psi_2 \in \underline{\mathcal{C}}$. Replacing $\psi(s,y)$ by $\psi(s,y) + (s-t)^2 + d^2(y,x)$ we can assume that the maximum is strict. By the definition of v^* there exist $w_n \in \mathcal{S}$ and $(\tilde{t}_n, \tilde{x}_n) \to (t,x)$ such that $w_n(\tilde{t}_n,\tilde{x}_n) \to v^*(t,x)$, and thus

$$\sup_A (w_n - \psi) \to v^*(t,x) - \psi(t,x).$$

Applying Lemma 7.4 on A, there exist points $(t_n,x_n) \in A$, and perturbation functions $\varphi_n(s,y) = \frac{1}{n}\sum_{k=1}^\infty \beta_k^n d^2(y,x_k^n) + p_1^n(t)$ from Lemma 7.4 such that

$$|\partial_t \varphi_n(t_n,x_n)| \le 1/n,\ |\nabla\varphi_n(t_n,x_n)|^* \le 1/n$$

and such that $w_n - \psi - \varphi_n$ has a maximum over A at (t_n,x_n), and

$$\sup_A (w_n - \psi) - \frac{1}{n} < w_n(t_n,x_n) - \psi(t_n,x_n) \le v^*(t_n,x_n) - \psi(t_n,x_n) \le v^*(t,x) - \psi(t,x).$$

Letting $n \to +\infty$ above we thus obtain

$$\lim_{n\to+\infty} (v^*(t_n,x_n) - \psi(t_n,x_n)) = v^*(t,x) - \psi(t,x).$$

Since the maximum at (t,x) was strict this implies $(t_n,x_n) \to (t,x)$.

We now have

$$\partial_t\psi(t_n,x_n) + \partial_t\varphi_n(t_n,x_n) + H_{|\nabla\tilde{\psi}_n(t_n,x_n)|^*}(|\nabla\psi_1(t_n,x_n)|) + f(x_n) \le 0, \tag{7.28}$$

where $\tilde{\psi}_n = \psi_2 + \varphi_n$. It follows from the definition that

$$|\nabla \tilde{\psi}_n(t_n, x_n)|^* \leq |\nabla \psi_2(t_n, x_n)|^* + |\nabla \varphi_n(t_n, x_n)|^*.$$

Therefore

$$\limsup_{n \to +\infty} |\nabla \tilde{\psi}_n(t_n, x_n)|^* \leq \limsup_{n \to +\infty} |\nabla \psi_2(t_n, x_n)|^* + \frac{1}{n} \leq |\nabla \psi_2(t, x)|^*, \tag{7.29}$$

where we used the upper semicontinuity of $|\nabla \psi_2|^*$. It is not difficult to see that since H is continuous, the function $H_r(s)$ is continuous in r, s (and hence uniformly continuous on bounded sets) and is non-increasing in r. It thus remains to let $n \to +\infty$ in (7.28) and use (7.29) to get

$$\partial_t \psi(t, x) + H_{|\nabla \psi_2(t,x)|^*}(|\nabla \psi_1(t, x)|) + f(x) \leq 0.$$

It now follows from Theorem 7.5 that $v^* \leq \overline{v}$ and hence $v = v^* \in \mathcal{S}$.

We remark that it is obvious from the definition of metric viscosity subsolution that the maximum of two metric viscosity subsolutions in any open subset of $(0, T) \times \mathbb{S}$ is a metric viscosity subsolution, a fact which we will use in Step 2.

Step 2. If v_* is not a viscosity supersolution then there exist (t, x) and $\psi = \psi_1 + \psi_2 \in \overline{\mathcal{C}}$ such that $v_* - \psi$ has a local minimum at (t, x) and

$$\partial_t \psi(t, x) + H^{|\nabla \psi_2(t,x)|^*}(|\nabla \psi_1(t, x)|) + f(x) < -2\epsilon. \tag{7.30}$$

for some $\epsilon > 0$. If $v_*(t, x) = \overline{v}(t, x)$ then, since $v_* \leq \overline{v}$, this would mean that $\overline{v} - \psi$ has a local minimum at (t, x). But $\overline{v}$ is a viscosity supersolution and hence (7.30) could not be true. Therefore we must have $v_*(t, x) < \overline{v}(t, x)$. Moreover

$$\psi_t(s, y) + H^{|\nabla \psi_2(s,y)|^*}(|\nabla \psi_1((s, y)|) + f(y) < -\epsilon \quad \text{if } |t - s|^2 + d^2(x, y) < r^2 \tag{7.31}$$

for some $t > r > 0$. Without loss of generality we can assume that $v_*(t, x) - \psi(t, x) = 0$ and the minimum is strict. Therefore, by possibly making r smaller, there exists $0 < \eta$ such that $v_*(s, y) > \psi(s, y) + \eta$ for $r^2/2 \leq |t - s|^2 + d^2(x, y) < r^2$ and $\psi + \eta < \overline{v}$ if $|t - s|^2 + d^2(x, y) < r^2$. Define a function

$$w(s, y) = \begin{cases} \max(\psi + \eta, v) & \text{if } |t - s|^2 + d^2(x, y) < r^2, \\ v \quad \text{otherwise.} \end{cases}$$

We claim that w is a viscosity subsolution of (7.5). To prove this it is enough to show that the function ψ (and hence $\psi + \eta$) is a viscosity subsolution of (7.5) in $\{(s, y) : |t - s|^2 + d^2(x, y) < r^2\}$. Let then $\psi - \tilde{\psi}$ have a local maximum at (s, y) for some $\tilde{\psi} = \tilde{\psi}_1 + \tilde{\psi}_2 \in \underline{\mathcal{C}}$. Then obviously $\partial_t \tilde{\psi}(s, y) = \partial_t \psi(s, y)$ and

$$|\nabla^+(\psi_1 - \tilde{\psi}_1)(s, y)| \leq |\nabla^+(\tilde{\psi}_2 - \psi_2)(s, y)| \leq |\nabla(\tilde{\psi}_2 - \psi_2)(s, y)|. \tag{7.32}$$

Therefore

$$\begin{aligned} & |\nabla^+(\tilde{\psi}_2 - \psi_2)(s, y)| \\ \geq\ & \limsup_{z \to y} \frac{[(\psi_1 - \tilde{\psi}_1)(s, z)) - (\psi_1 - \tilde{\psi}_1)(s, y))]_+}{d(z, y)} \\ \geq\ & \limsup_{z \to y} \frac{[\psi_1(s, z)) - \psi_1(s, y)]_+}{d(z, y)} - \limsup_{z \to y} \frac{|\tilde{\psi}_1(s, z) - \tilde{\psi}_1(s, y)|}{d(z, y)} \\ =\ & |\nabla^+ \psi_1(s, y)| - |\nabla \tilde{\psi}_1(s, y)| = |\nabla \psi_1(s, y)| - |\nabla \tilde{\psi}_1(s, y)|. \end{aligned}$$

Likewise we obtain

$$\begin{aligned}|\nabla^+(\tilde\psi_2-\psi_2)(s,y)| &\geq |\nabla^+(-\tilde\psi_1)(s,y)|-|\nabla\psi_1(s,y)|\\ &= |\nabla^-\tilde\psi_1(s,y)|-|\nabla\psi_1(s,y)|\\ &= |\nabla\tilde\psi_1(s,y)|-|\nabla\psi_1(s,y)|.\end{aligned}$$

It thus follows from the above two inequalities and (7.32) that

$$||\nabla\tilde\psi_1(s,y)|-|\nabla\psi_1(s,y)||\leq|\nabla(\tilde\psi_2-\psi_2)(s,y)|\leq|\nabla\psi_2(s,y)|+|\nabla\tilde\psi_2(s,y)|$$

which, together with (7.31), implies

$$\partial_t\tilde\psi(s,y)+H_{|\nabla\tilde\psi_2(s,y)|^*}(|\nabla\tilde\psi_1((s,y)|)+f(y)<-\epsilon.$$

Therefore w is a viscosity subsolution of (7.5) and hence $w\in\mathcal{S}$ (since $w\leq\overline{v}$). However, it is clear from the definition of w that $w(\tau,z)>v(\tau,z)$ for some (τ,z) close to (t,x). This is a contradiction so v_* must be a viscosity supersolution of (7.5). Since by Theorem 7.5 we must have $v\leq v_*$ it finally follows that $v=v^*=v_*$ is a viscosity solution of (7.5). □

We remark that under the assumptions $\underline{u}\leq\overline{v}$ and (7.6), condition (7.27) is equivalent to

$$\lim_{t\to0}\left(|\overline{v}(t,x)-g(x)|+|\underline{u}(t,x)-g(x)|\right)=0$$

uniformly on bounded sets of $\mathbb{S}$.

Corollary 3. *Let g be Lipschitz continuous and f satisfy (7.6) and be bounded, and H be continuous. Then there exists a viscosity solution of (7.5).*

Proof. We notice that for sufficiently big C, the functions

$$\underline{u}(t,x)=-Ct+g(x),\quad \overline{u}(t,x)=Ct+g(x)$$

are respectively a viscosity subsolution and a viscosity supersolution of (7.5) satisfying (7.7) and (7.27). To see this for the subsolution case, suppose that $\underline{u}-\psi$ has a local maximum at a point (t,x) for some $\psi=\psi_1+\psi_2\in\underline{\mathcal{C}}$. Then

$$\psi_i(t,y)-\psi_1(t,x)\geq(\underline{u}(t,y)-\psi_2(t,y))-(\underline{u}(t,x)-\psi_2(t,x)).$$

Therefore

$$\begin{aligned}|\nabla\psi_1(t,x)| &= |\nabla^-\psi_1(t,x)|\\ &\leq |\nabla^-(\underline{u}-\psi_2)(t,x)|\\ &\leq |\nabla\underline{u}(t,x)|+|\nabla\psi_2(t,x)|\leq C_1+|\nabla\psi_2(t,x)|^*,\end{aligned}$$

where C_1 is the Lipschitz constant of g. Therefore $|\nabla\psi_1(t,x)|-|\nabla\psi_2(t,x)|^*\leq C_1$ and hence

$$H_{|\nabla\psi_2(t,x)|^*}(|\nabla\psi_1(t,x)|)\leq\sup_{0\leq s\leq C_1}H(s)=:C_2$$

which implies that $\underline{u}$ is a viscosity subsolution if $C\geq C_2+\sup f$.

The result thus follows from Theorem 7.6. □

Let us now consider a simpler case of equation

$$\begin{cases}\partial_t u+H(|\nabla u|)=0\\ u(0,x)=g(x),\end{cases}\tag{7.33}$$

where the Hamiltonian H is convex and g is bounded and uniformly continuous on bounded subsets of $\mathbb{S}$. More precisely, suppose that

$$H(s) = \sup_{r \geq 0} \{sr - \alpha(r)\}, \quad \text{for } s \geq 0,$$

where $\alpha : [0, +\infty) \to [0, +\infty)$ is an increasing convex function such that $\alpha(0) = 0$ and $\alpha(r)/r \to +\infty$ as $r \to +\infty$. In particular $H(0) = 0$ and H is increasing. The solution of (7.33) should be given by the Hopf-Lax formula

$$u(t, x) = \inf_{y \in \mathbb{S}} \left\{ g(y) + t\alpha\left(\frac{d(y, x)}{t}\right)\right\}. \tag{7.34}$$

Indeed it was proved in [8, 48] (see also [64]) that u satisfies

$$\frac{d}{dt_+} u(t, x) + H(|\nabla u(t, x)|) = 0 \quad \text{for every } t > 0, x \in \mathbb{S}.$$

We will prove that u is a metric viscosity solution of (7.33). First we observe that, since the space is geodesic, it is easy to see that u satisfies the semigroup property

$$u(t + h, x) = \inf_{y \in \mathbb{S}} \left\{ u(t, y) + h\alpha\left(\frac{d(y, x)}{h}\right)\right\} \quad 0 \leq t < t + h \leq T. \tag{7.35}$$

Theorem 7.7. *Under the above assumptions on H and g, the function u given by (7.34) is a metric viscosity solution of (7.33) on $[0, +\infty) \times \mathbb{S}$.*

Proof. It is standard to see that u is continuous on $[0, +\infty) \times \mathbb{S}$.
Step 1. Suppose that $u - \psi$ has a local maximum at a point (t, x) for some $\psi = \psi_1 + \psi_2 \in \underline{\mathcal{C}}$. Set $r > 0$. By the definition of test functions, there must exist points x_n such that $d(x, x_n) \to 0$ and

$$|\nabla \psi_1(t, x)| = |\nabla^- \psi_1(t, x)| = \lim_{n \to +\infty} \frac{\psi_1(t, x) - \psi_1(t, x_n)}{d(x, x_n)}.$$

Denote $\epsilon_n = d(x, x_n)/r$ and let $s = t - \epsilon_n$. Then by (7.35) we have

$$\psi(t, x) - \psi(t - \epsilon_n, x_n) \leq u(t, x) - u(t - \epsilon_n, x_n) \leq \epsilon_n \alpha\left(\frac{d(x_n, x)}{\epsilon_n}\right) = \epsilon_n \alpha(r). \tag{7.36}$$

Now

$$\begin{aligned}
\frac{\psi(t, x) - \psi(t - \epsilon_n, x_n)}{\epsilon_n} &= \frac{\psi(t, x) - \psi(t, x_n)}{\epsilon_n} + \frac{\psi(t, x_n) - \psi(t - \epsilon_n, x_n)}{\epsilon_n} \\
&\geq (|\nabla \psi_1(t, x)| - |\nabla \psi_2(t, x)| + \sigma_1(n))r \\
&+ \frac{1}{\epsilon_n} \int_{t - \epsilon_n}^{t} \partial_t \psi(s, x_n) ds \\
&= (|\nabla \psi_1(t, x)| - |\nabla \psi_2(t, x)|)r + \partial_t \psi(t, x) \\
&+ \sigma_1(n).
\end{aligned} \tag{7.37}$$

where $\lim_{n \to +\infty} \sigma_1(n) = 0$. Combining (7.36) and (7.37) and letting $n \to +\infty$ we thus obtain for every $r > 0$

$$\partial_t \psi(t, x) + (|\nabla \psi_1(t, x)| - |\nabla \psi_2(t, x)|)r - \alpha(r) \leq 0$$

This obviously implies that

$$\partial_t \psi(t, x) + H_{|\nabla \psi_2(t, x)|}(|\nabla \psi_1(t, x)|) \leq 0.$$

Step 2. Suppose that $u - \psi$ has a local minimum at a point (t, x) for some $\psi = \psi_1 + \psi_2 \in \overline{\mathcal{C}}$. By (7.35), for every $\epsilon > 0$ there exists $x_\epsilon, d(x, x_\epsilon) \to 0$ as $\epsilon \to 0$, such that

$$\psi(t,x) - \psi(t-\epsilon, x_\epsilon) \geq u(t,x) - u(t-\epsilon, x_\epsilon) \geq \epsilon\alpha\left(\frac{d(x_\epsilon, x)}{\epsilon}\right) - \epsilon^2. \tag{7.38}$$

We have

$$\begin{aligned} \frac{\psi(t,x) - \psi(t-\epsilon, x_\epsilon)}{\epsilon} &= \frac{\psi(t,x) - \psi(t, x_\epsilon)}{\epsilon} + \frac{\psi(t, x_\epsilon) - \psi(t-\epsilon, x_\epsilon)}{\epsilon} \\ &\leq (|\nabla\psi_1(t,x)| + |\nabla\psi_2(t,x)| + \sigma_2(\epsilon))\frac{d(x_\epsilon, x)}{\epsilon} \\ &+ \frac{1}{\epsilon}\int_{t-\epsilon}^{t} \partial_t\psi(s, x_\epsilon)ds \\ &= (|\nabla\psi_1(t,x)| + |\nabla\psi_2(t,x)| + \sigma_2(\epsilon))\frac{d(x_\epsilon, x)}{\epsilon} \\ &+ \partial_t\psi(t,x) + \sigma_2(\epsilon), \end{aligned} \tag{7.39}$$

where $\lim_{\epsilon\to 0}\sigma_2(\epsilon) = 0$. Combining (7.38) and (7.39) it thus follows

$$\begin{aligned} -\epsilon - \sigma_2(\epsilon) &\leq \partial_t\psi(t,x) + (|\nabla\psi_1(t,x)| + |\nabla\psi_2(t,x)| + \sigma_2(\epsilon))\frac{d(x_\epsilon, x)}{\epsilon} \\ &- \alpha\left(\frac{d(x_\epsilon, x)}{\epsilon}\right) \\ &\leq \partial_t\psi(t,x) + H(|\nabla\psi_1(t,x)| + |\nabla\psi_2(t,x)| + \sigma_2(\epsilon)) \\ &= \partial_t\psi(t,x) + H^{|\nabla\psi_2(t,x)|}(|\nabla\psi_1(t,x)| + \sigma_2(\epsilon)). \end{aligned}$$

It remains to let $\epsilon \to 0$ above to conclude the proof. □

We expect that value functions for more general problems, like these studied in Section 6, are metric viscosity solutions of the associated Hamilton–Jacobi equations in our, or perhaps slightly different sense. The relationship between the notion of metric viscosity solution and the notion from Section 6 is also yet to be investigated.

8. Appendix.

8.1. Gronwall type inequality.

Lemma 8.1. *Let ω be a nonnegative Borel function defined on $[0, a]$ such that $\omega(y) > \omega(0) = 0$ for $y \in (0, a)$. Assume*

$$\int_0^a \frac{dy}{\omega(y)} = \infty.$$

Suppose $Q : [0, T] \to [0, a]$ is a Lipschitz function such that $Q(0) = 0$ and $\dot{Q} \leq \omega(Q)$ almost everywhere. Then $Q \equiv 0$ on $(0, T)$.

Proof. Suppose on the contrary that the open set $O = \{t \in (0, T) \mid Q(t) > 0\}$ is not empty. Let (α, β) be a connected component of O, where $0 \leq \alpha < \beta \leq T$. If $Q(\alpha) > 0$, then $\alpha \neq 0$ and so, there exists $\epsilon > 0$ such that $(a - \epsilon, \beta) \subset O$, which contradicts the maximality property of (α, β). Hence, $Q(\alpha) = 0$. Since almost everywhere on (α, β) we have $\dot{Q} \leq \omega(Q)$ and $\omega(Q) > 0$ we conclude that if $\alpha < t_0 < t_1 < \beta$ then

$$(t_1 - t_0) \geq \int_{t_0}^{t_1} \frac{\dot{Q}}{\omega(Q)}dt = \int_{Q(t_0)}^{Q(t_1)} \frac{dy}{\omega(y)}.$$

Thus,

$$t_1 - \alpha \geq \int_0^{Q(t_1)} \frac{dy}{\omega(y)} = \infty,$$

which leads a contradiction. □

8.2. **Shift of a curve in $\mathcal{P}_2(M)$.** Let $\sigma \in AC_2(0,T;\mathcal{P}_2(M))$ and let $\mathbf{v}$ be a velocity for σ. The following lemma can be derived from the Appendix in [42].

Lemma 8.2. *There exists an increasing sequence of integers $\{n_k\}_k$ and paths $\sigma^k \in AC_2(0,T;\mathcal{P}^{n_k}(M))$ such that $\mathbf{v}^k$ is a velocity for σ^k such that*

$$W_2(\sigma_t,\sigma_t^k) \leq \frac{1}{k} \quad \textit{and} \quad \Big|\int_0^T ||\mathbf{v}_t^k||^2_{\sigma_t^k} dt - \int_0^T ||\mathbf{v}_t||^2_{\sigma_t} dt\Big| \leq \frac{1}{k}. \tag{8.1}$$

Furthermore, we can find $x^{i,k} \in AC_2(0,T,M)$ ($i = 1,\cdots,n_k$) such that

$$\sigma_t^k = \frac{1}{n_k}\sum_{i=1}^{n_k} \delta_{x^{i,k}(t)}.$$

For almost every $t \in (0,T)$

$$||\mathbf{v}_t^k||^2_{\sigma_t^k} = \frac{1}{n_k}\sum_{i=1}^{n_k} |\dot{x}^{i,k}(t)|^2.$$

We prove the following lemma.

Lemma 8.3. *Given $\nu \in \mathcal{P}_2(M)$ there exist $\sigma^* \in AC_2(0,T;\mathcal{P}_2(M))$ and a velocity $\mathbf{v}^*$ for σ^* such that $\sigma_T^* = \nu$,*

$$\int_0^T ||\mathbf{v}_t^*||^2_{\sigma_t^*} dt \leq \int_0^T ||\mathbf{v}_t||^2_{\sigma_t} dt, \tag{8.2}$$

and for all $t \in [0,T]$

$$W_2(\sigma_t,\sigma_t^*) \leq W_2(\sigma_T,\sigma_T^*). \tag{8.3}$$

Proof. Let $(\sigma^k,\mathbf{v}^k)$ be as in Lemma 8.2 and let $\{y^{i,k}\}_{i=1}^{n_k} \subset M$ be such that

$$\lim_{k\to\infty} W_2(\nu^k,\nu) = 0, \quad \text{where} \quad \nu^k = \frac{1}{n_k}\sum_{i=1}^{n_k} \delta_{y^{i,k}(T)}. \tag{8.4}$$

Reordering $\{y^{i,k}\}_{i=1}^{n_k}$ if necessary, we may assume without loss of generality that

$$W_2^2(\nu^k,\sigma_T^k) = \frac{1}{n_k}\sum_{i=1}^{n_k} |y^{i,k}(T) - x_T^{i,k}|^2. \tag{8.5}$$

Set

$$\sigma_t^{*,k} = \frac{1}{n_k}\sum_{i=1}^{n_k} \delta_{y^{i,k}(t)}, \quad \text{where} \quad y^{i,k}(t) = x^{i,k}(t) - x^{i,k}(T) + y^{i,k}(T).$$

We have $\sigma^{*,k} \in AC_2(0,T;\mathcal{P}^{n_k}(M))$ and it has a unique velocity $\mathbf{v}^{*,k}$ such that for almost every $t \in (0,T)$ (cf. Section 7.3 [42])

$$\mathbf{v}_t^{*,k}(y^{i,k}(t)) = \dot{y}^{i,k}(t).$$

For these t

$$||\mathbf{v}_t^{*,k}||^2_{\sigma_t^{*,k}} = \frac{1}{n_k}\sum_{i=1}^{n_k}|\mathbf{v}_t^{*,k}(y^{i,k}(t))|^2 = \frac{1}{n_k}\sum_{i=1}^{n_k}|\dot{y}^{i,k}(t)|^2 = \frac{1}{n_k}\sum_{i=1}^{n_k}|\dot{x}^{i,k}(t)|^2 = ||\mathbf{v}_t^k||^2_{\sigma_t^k}. \tag{8.6}$$

Observe that

$$\mathcal{W}(\sigma_t^{*,k}) = \frac{1}{n^2}\sum_{i,j=1}^{n} W\Big((x^{i,k}(t)-x^{j,k}(t))+(y^{i,k}(T)-y^{j,k}(T))-(x^{i,k}(T)-x^{j,k}(T))\Big). \tag{8.7}$$

Thanks to (8.5) we conclude that for all $t\in[0,T]$,

$$W_2^2\big(\sigma_t^k,\sigma_t^{*,k}\big) \le \frac{1}{n_k}\sum_{i=1}^{n_k}|y^{i,k}(t)-x^{i,k}(t)|^2 = \frac{1}{n_k}\sum_{i=1}^{n_k}|y^{i,k}(T)-x^{i,k}(T)|^2 = W_2^2(\nu^k,\sigma_T^k). \tag{8.8}$$

By the triangle inequality

$$W_2\big(\sigma_t^{*,k},\sigma_0\big) \le W_2\big(\sigma_t^{*,k},\sigma_t^k\big) + W_2\big(\sigma_t^k,\sigma_t\big) + W_2\big(\sigma_t,\sigma_0\big).$$

We use (8.8), the first inequality in (8.1) and Remark 1 to conclude that

$$W_2\big(\sigma_t^{*,k},\sigma_0\big) \le W_2(\nu^k,\sigma_T^k) + \frac{1}{k} + \int_0^t ||\mathbf{v}_s||_{\sigma_s}ds \le m+1+\int_0^T ||\mathbf{v}_s||_{\sigma_s}ds, \tag{8.9}$$

where

$$m = \sup_k W_2(\nu^k,\sigma_T^k) \le \sup_k W_2(\nu^k,\nu) + W_2(\nu,\sigma_T) + W_2(\sigma_T,\sigma_T^k) < \infty.$$

By the second inequality in (8.1) and (8.6),

$$\sup_{k\in\mathbb{N}}\int_0^T ||\mathbf{v}_s^{*,k}||^2_{\sigma_s^{*,k}}ds \le 1 + \int_0^T ||\mathbf{v}_s||^2_{\sigma_s}ds. \tag{8.10}$$

Passing to a subsequence if necessary and applying the refined version of the Ascoli–Arzela Theorem in [7] (cf. also Proposition 3.20 [43]) may assume without loss of generality that there exists $\sigma^*\in AC_2(0,T;\mathcal{P}_2(M))$ such that $\{\sigma_t^{*,k}\}_k$ converges narrowly to σ_t^* for each $t\in[0,T]$. Since $W_2(\sigma_T^{*,k},\nu^k)=0$, (8.4) implies that $W_2(\sigma_T^*,\nu)=0$. We let k tend to ∞ in (8.8) to obtain (8.3).

By (8.1) and (8.6),

$$\int_0^T ||\mathbf{v}_s||^2_{\sigma_s}ds = \liminf_{k\to\infty}\int_0^T ||\mathbf{v}_s^k||^2_{\sigma_s^k}ds = \liminf_{k\to\infty}\int_0^T ||\mathbf{v}_s^{*,k}||^2_{\sigma_s^{*,k}}ds. \tag{8.11}$$

Since $||\mathbf{v}_s^{*,k}||_{\sigma_s^{*,k}} \ge |(\sigma^{*,k})'|(s)$ almost everywhere (cf. Proposition 1), we first use (8.11) and then Proposition 3 [40] to conclude that

$$\int_0^T ||\mathbf{v}_s||^2_{\sigma_s}ds \ge \liminf_{k\to\infty}\int_0^T |(\sigma^{*,k})'|^2ds \ge \int_0^T |(\sigma^*)'|^2ds.$$

If $\mathbf{v}^*$ is the velocity of minimal norm for σ^* we observe that we have established (8.2). ☐

Acknowledgments. The research of W. Gangbo was supported by NSF grants DMS–0901070 and DMS–1160939. The research of A. Święch was supported by NSF grant DMS-0856485. Both authors would like to thank R. Hynd, H-K. Kim and C. Roberto for useful discussions.

REFERENCES

[1] Y. Achdou and I. Capuzzo Dolcetta, *Mean field games: Numerical methods*, SIAM J. Numer. Anal., **48** (2010), 1136–1162.

[2] Y. Achdou, F. Camilli and I. Capuzzo Dolcetta, *Mean field games: Numerical methods for the planning problem*, SIAM J. Control Optim., **50** (2012), 77–109.

[3] G. Alberti and L. Ambrosio, *A geometrical approach to monotone functions in* $\mathbf{R}^n$, Math. Z., **230** (1999), 259–316.

[4] L. Ambrosio and J. Feng, *On a class of first order Hamilton-Jacobi equations in metric spaces*, preprint.

[5] L. Ambrosio, N. Fusco and D. Pallara, "Functions of Bounded Variation and Free Discontinuity Problems," Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2000.

[6] L. Ambrosio and W. Gangbo, *Hamiltonian ODEs in the Wasserstein space of probability measures*, Comm. Pure Appl. Math., **61** (2007), 18–53.

[7] L. Ambrosio, N. Gigli and G. Savaré, "Gradient Flows in Metric Spaces and in the Space of Probability Measures," Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005.

[8] L. Ambrosio, N. Gigli and G. Savaré, *Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces,* Rev. Mat. Iberoam., **29** (2013), 969–996.

[9] V. Barbu and G. Da Prato, "Hamilton-Jacobi Equations in Hilbert Spaces," Research Notes in Mathematics, **86**, Pitman (Advanced Publishing Program), Boston, MA, 1983.

[10] P. Bernard and B. Buffoni, *Optimal mass transportation and Mather theory*, Journal of the European Mathematical Society, **9** (2007), 85–121.

[11] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio and C. Landim, *Macroscopic fluctuation theory for stationary non-equilibrium states*, J. Statist. Phys., **107** (2002), 635–675.

[12] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio and C. Landim, *Large deviations for the boundary driven symmetric simple exclusion process*, Math. Phys. Anal. Geom., **6** (2003), 231–267.

[13] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio and C. Landim, *Minimum dissipation principle in stationary non-equilibrium states*, J. Statist. Phys., **116** (2004), 831–841.

[14] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio and C. Landim, *Stochastic interacting particle systems out of equilibrium*, J. Stat. Mech. Theory Exp., **2007** (2007), P07014, 35 pp.

[15] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio and C. Landim, *Action functional and quasi-potential for the Burgers equation in a bounded interval*, Comm. Pure Appl. Math., **64** (2011), 649–696.

[16] L. Bertini, D. Gabrielli and J. L. Lebowitz, *Large deviations for a stochastic model of heat flow*, J. Stat. Phys., **121** (2005), 843–885.

[17] A. J. Bertozzi and A. Majda, "Vorticity and Incompressible Flow," Cambridge Texts in Applied Mathematics, **27**, Cambridge University Press, Cambridge, 2002.

[18] J. M. Borwein and D. Preiss, *A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions*, Trans. Amer. Math. Soc., **303** (1987), 517–527.

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

[20] Y. Brenier, W. Gangbo, G. Savaré and M. Westdickenberg, *Sticky particle dynamics with interactions*, Journal de Math. Pures et Appliquées (9), **99** (2013), 577–617.

[21] P. Cardialaguet, J.-M. Lasry, P.-L. Lions and A. Porretta, *Long times average of mean field games*, Networks and Heterogeneous Media, **7** (2012), 279–301.

[22] P. Cardialaguet and M. Quincampoix, *Deterministic differential games under probability knowledge of initial condition*, Int. Game Theory Rev., **10** (2008), 1–16.

[23] M. G. Crandall, H. Ishii and P.-L. Lions, *User's Guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.), **27** (1992), 1–67.

[24] M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations in Banach spaces,* in "Trends in the Theory and Practice of Nonlinear Analysis" (Arlington, Tex., 1984), North-Holland Math. Stud., **110**, North-Holland, Amsterdam, (1985), 115–119.

[25] M. G. Crandall and P.-L. Lions, *Hamilton-Jacobi equations in infinite dimensions. I. Uniqueness of viscosity solutions*, J. Funct. Anal., **62** (1985), 379–396.

[26] M. G. Crandall and P.-L. Lions, *Hamilton-Jacobi equations in infinite dimensions. II. Existence of viscosity solutions*, J. Funct. Anal., **63** (1986), 368–405.

[27] M. G. Crandall and P.-L. Lions, *Hamilton-Jacobi equations in infinite dimensions. III*, J. Funct. Anal., **68** (1986), 214–247.

[28] M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations in infinite dimensions. IV. Hamiltonian with unbounded linear terms*, J. Funct. Anal., **90** (1990), 237–283.

[29] M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations in infinite dimensions. V. Unbounded linear terms and B-continuous functions*, J. Funct. Anal., **97** (1991), 417–465.

[30] M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations in infinite dimensions. VII. The HJB equation is not always satisfied*, J. Funct. Anal., **125** (1994), 111–148.

[31] A. Fathi, "Weak KAM Theory in Lagrangian Dynamics," Cambridge Studies in Advanced Mathematics, Cambridge University Press, New York, 2012.

[32] J. Feng, *A Hamilton-Jacobi PDE in the space of measures and its associated compressible Euler equations*, C. R. Math. Acad. Sci. Paris, **349** (2011), 973–976.

[33] J. Feng and M. Katsoulakis, *A comparison principle for Hamilton-Jacobi equations related to controlled gradient flows in infinite dimensions*, Arch. Ration. Mech. Anal., **192** (2009), 275–310.

[34] J. Feng and T. Kurtz, "Large Deviations for Stochastic Processes," Mathematical Surveys and Monographs, **131**, American Mathematical Society, Providence, RI, 2006.

[35] J. Feng and T. Nguyen, *Hamilton-Jacobi equations in space of measures associated with a system of conservation laws*, J. Math. Pures Appl. (9), **97** (2012), 318–390.

[36] J. Feng and A. Święch, *Optimal control for a mixed flow of Hamiltonian and gradient type in space of probability measures*, With an appendix by Atanas Stefanov, Trans. Amer. Math. Soc., **365** (2013), 3987–4039.

[37] I. Fonseca and W. Gangbo, "Degree Theory in Analysis and Its Applications," Oxford Lecture Series in Mathematics and its Applications, **2**, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1995.

[38] W. Gangbo, H. K. Kim and T. Pacini, *Differential forms on Wasserstein space and infinite-dimensional Hamiltonian systems*, Memoirs of the AMS, **211** (2011).

[39] W. Gangbo and R. McCann, *The geometry of optimal transportation*, Acta Math., **177** (1996), 113–161.

[40] W. Gangbo, T. Nguyen and A. Tudorascu, *Euler-Poisson systems as action-minimizing paths in the Wasserstein space*, Arch. Ration. Mech. Anal., **192** (2009), 419–452.

[41] W. Gangbo, T. Nguyen and A. Tudorascu, *Hamilton-Jacobi equations in the Wasserstein space*, Methods and Applications of Analysis, **15** (2008), 155–183.

[42] W. Gangbo and A. Tudorascu, *Homogenization for a class of integral functionals in spaces of probability measures*, Advances in Mathematics, **230** (2012), 1124–1173.

[43] W. Gangbo and A. Tudorascu, *Weak KAM theory on the Wasserstein with multi-dimensional underlying space*, to appear in Comm. Pure Applied Math.

[44] Y. Giga, N. Hamamuki and A. Nakayasu, *Eikonal equations in metric spaces*, preprint.

[45] D. A. Gomes, J. Mohr and R. R. Souza, *Discrete time, finite state space mean field games*, J. Math. Pures Appl. (9), **93** (2010), 308–328.

[46] D. A. Gomes and L. Nurbekyan, *Weak kam theory on the d-infinite dimensional torus*, preprint.

[47] D. A. Gomes and L. Nurbekyan, *On the minimizers of calculus of variations problems in Hilbert spaces*, preprint.

[48] N. Gozlan, C. Roberto and P. M. Samson, *Hamilton Jacobi equations on metric spaces and transport-entropy inequalities*, preprint, `arXiv:1203.2783`.

[49] O. Guéant, J.-M. Lasry and P.-L. Lions, *Mean field games and applications*, in "Paris-Princeton Lectures on Mathematical Finance 2010," Lecture Notes in Math., Springer, Berlin, (2003), 205–266.

[50] O. Guéant, J.-M. Lasry and P.-L. Lions, *Application of mean field games to growth theory*, preprint.

[51] O. Guéant, *A reference case for mean field games models*, J. Math. Pures Appl. (9), **92** (2009), 276–294.

[52] O. Guéant, "Mean Field Games and Applications to Economics," Ph.D Thesis, Université Paris Dauphine, 2010.

[53] R. Hynd and H.-K. Kim, work in progress.

[54] H. Ishii, *Viscosity solutions for a class of Hamilton-Jacobi equations in Hilbert spaces*, J. Funct. Anal., **105** (1992), 301–341.

[55] B. Khesin and P. Lee, *Poisson geometry and first integrals of geostrophic equations*, Phys. D, **237** (2008), 2072–2077.

[56] J.-M. Lasry and P.-L. Lions, *Large investor trading impacts on volatility*, Ann. Inst. H. Poincaré Anal. Non Linéaire, **24** (2007), 311–323.

[57] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. I. Le cas stationnaire,* (French) [Mean field games. I. The stationary case], C. R. Math. Acad. Sci. Paris, **343** (2006), 619–625.

[58] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. II. Horizon fini et controle optimal*, (French) [Mean field games. II. Finite horizon and optimal control], C. R. Math. Acad. Sci. Paris, **343** (2006), 679–684.

[59] J.-M. Lasry and P.-L. Lions, *Mean field games*, Japan. J. Math., **2** (2007), 229–260.

[60] P.-L. Lions, *Cours au Collège de France*. Available from: `http://www.college-de-france.fr`.

[61] G. Loeper, "Applications de l'Équation de Monge-Ampère à la Modélisation des Fluides et des Plasmas," Thesis dissertation, Université de Nice Sophia Antipolis.

[62] G. Loeper, *A fully nonlinear version of the incompressible Euler equations: The semi-geostrophic system*, SIAM Journal on Mathematical Analysis, **38** (2006), 795–823.

[63] J. Lott, *Some geometric calculations on Wasserstein space*, Comm. Math. Phys., **277** (2008), 423–437.

[64] J. Lott and C. Villani, *Hamilton-Jacobi semigroup on length spaces and applications*, J. Math. Pures Appl. (9), **88** (2007), 219–229.

[65] D. Tataru, *Viscosity solutions of Hamilton-Jacobi equations with unbounded nonlinear terms*, J. Math. Pures Appl., **163** (1992), 345–392.

[66] D. Tataru, *Viscosity solutions for Hamilton-Jacobi equations with unbounded nonlinear term: A simplified approach*, J. Differential Equations, **111** (1994), 123–146.

[67] C. Villani, "Topics in Optimal Transportation," Graduate Studies in Mathematics, **58**, American Mathematical Society, Providence, RI, 2003.

[68] V. Judovič, *Non-stationary flows of an ideal incompressible fluid*, Žh. Vyčisl. Mat. i Mat. Fiz., **3** (1963), 1032–1066.

Received January 2013; revised April 2013.

E-mail address: `gangbo@math.gatech.edu`
E-mail address: `swiech@math.gatech.edu`