ON A CLASS OF FIRST ORDER HAMILTON-JACOBI EQUATIONS IN METRIC SPACES

LUIGI AMBROSIO, AND JIN FENG

ABSTRACT. We establish well-posedness of a class of first order Hamilton-Jacobi equation in geodesic metric spaces. The result is then applied to solve a Hamilton-Jacobi equation in the Wasserstein space of probability measures, which arises from the variational formulation of a compressible Euler equation.

1. INTRODUCTION

Let (X, d) be a complete metric space and a geodesic space. We are interested in a class of minimization problem which includes in particular the following:

$$\int_0^T \Big(\frac{1}{2}|x'|^2 - V(x)\Big) dr$$

where $V : \mathsf{X} \mapsto \mathbb{R}$ is uniformly continuous and bounded from above, $x = x(t) \in AC([0,T]; \mathsf{X})$ is an absolutely continuous path in X, and

$$|x'|(t) := \lim_{s \to t} \frac{\mathsf{d}(x(s), x(t))}{|s-t|}$$

denotes its metric derivative. See Chapter 1, Ambrosio, Gigli and Savaré [2] for definitions and properties of absolutely continuous curves in metric spaces. Given $U_0 : X \to \mathbb{R}$, we define

$$U(t,x) := \sup \Big\{ U_0(z(t)) - \int_0^t \Big(\frac{1}{2}|z'|^2(r) - V(z(r))\Big) dr : z(0) = x,\ z(\cdot) \in AC([0,t]; \mathsf{X}) \Big\}.$$

Then U solves a Hamilton-Jacobi equation, formally written as

$$\partial_t U(t,x) = \frac{1}{2}|D_x U(t,x)|^2 + V(x), \tag{1.1}$$

where the *slope* (also called *local Lipschitz constant*) for a function $f : \mathsf{X} \mapsto \mathbb{R}$ is defined as

$$|Df|(x) := \limsup_{y \to x} \frac{|f(y) - f(x)|}{\mathsf{d}(y,x)}. \tag{1.2}$$

We are interested in a well-posedness theory for (1.1) and related equations. To fix the ideas and to separate difficulties of different nature, we do not pursue generality and only focus on the case of V with uniformly continuity in balls of finite radius, with possible growth to $-\infty$ at certain rate with respect to the metric distance, and uniformly bounded from above. The case $V = 0$ is of special interest, as the corresponding U then defines a Hopf-Lax semigroup which has applications to transportation inequalities in abstract metric space settings. Pointwise solution of (1.1) has been considered in Chapters 7 and 22 of Villani [31], in Section 3 of Ambrosio, Gigli and Savaré [3], Gozlan, Roberto and Samson [26]. The pointwise solution

Date: December 20, 2013.

is sufficient for the purpose of the applications considered in these references. However, as is known in the case even when $\mathsf{X} \subset \mathbb{R}^d$, it is not good enough to ensure uniqueness for (1.1). Here we will generalize the notion of viscosity solution to metric space setting and develop a well-posedness theory.

The theory of viscosity solutions for (first order) Hamilton-Jacobi equation in infinite dimensions was initiated by Crandall and Lions [10]. One of the structural assumption is that X is a Hilbert space, or slightly more generally, certain Banach space with smooth norm. However, recent applications start to witness a situation where X is space of probability measures. This includes examples in statistical mechanics [4, 18], optimal control and probability [15, 16], classical game theory [8, 9], fluid mechanics [19, 20, 14, 17, 21, 22], and mean-field games (we refer to [7] for a compilation of references). In sections 2 and 3 of this article, we extend the first order viscosity theory to general metric spaces setting by exploring maximum principles of the Hamiltonian operator. Other alternatives exist. For instance, [24, 29] emphasize a formulation on path (hence the Lagrangian) by considering a sub-class of the Hamiltonians considered here. During the preparation of this article, we also received a preprint from the authors of [22] where their last section considers a related Cauchy problem using ideas of the same kind as in the first three sections of this article. Definition of viscosity solution is given for general Hamiltonians but well-posedness is treated for the case of $H(x,p) = H(p)$ and $H(x,p) = H(p) + f(x)$ only. Using Perron's method, solution is constructed implicitly. There is no convexity on p assumption on H. In this article, under a rather general structural assumption on H in (1.5), we treat Hamiltonians with much more general x-dependence. Growth estimate of solution is also provided. Our assumption implies that $p \mapsto H(x,p)$ is convex. However, the existence part of our well-posedness result is explicitly constructed using dynamical programming and value functions. Moreover, the proof of our uniqueness result does not critically rely on such convexity assumption. Finally, in Section 4, we give well-posedness for the resolvent equation relative to a special Hamiltonian in space of probability measures. Such issue, in the form of Cauchy problem, had been considered by [19, 22, 27] but no well-posedness was given. In particular, the relation between the metric definition of viscosity solution and a possible definition in Wasserstein space used in the above references was left open. See the concluding comments at the end of section 7 in [22]. We settle this issue in Section 4. We mention that, although we only treat the resolvent problem in detail, in principle the Cauchy problem should follow similarly.

The spirit of this article is closer to parts I, II and III of [10], but different than the rest of that series. By this, we mean that our Hamiltonian will only depend on f through $|Df|$. There is no notion of $\langle Df(x), v(x)\rangle_x$ for some velocity field v, which requires a notion of duality. This feature unfortunately excludes the examples in [5, 18, 14, 17]. However, at least in the case where X is space of probability measures, many such problems can be treated effectively by another method different than the one discussed here [15, 16, 18]. That method is closer in spirit to parts IV, V, VI, VII of [10]. Further generalization of the viscosity method to more general models in metric space settings is worthwhile.

The rest of this introduction summarizes well-posedness results for resolvent type and Cauchy problem of Hamilton-Jacobi equations in metric spaces. Detailed developments are given in Sections 2 and 3. In Section 4, we apply these results to an example concerning the variational formulation of a compressible Euler equation and derive well-posedness for an associated Hamilton-Jacobi equation in the space of probability measures.

Acknowledgement. Both authors acknowledge the support of the ERC ADG GeMeThNES. The second author also thanks the hospitaliy of the Classe di Scienze, Scuola Normale Superiore di Pisa, Italy, and of the Mathematics Department of Ecole Normale Supérieure de Lyon, France, during his sabbatical leave. He also acknowledges support by the LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, France.

The authors had useful discussions with Andrzej Swiech and Wilfrid Gangbo after the article was written. In particular, W. Gangbo discovered a gap in Section 4 in an earlier version of the paper. These inputs are incorporated in the current version.

1.1. **Basic setup.** We assume that (X, d) is a complete metric space and a geodesic space in the following sense: for every x, $y \in \mathsf{X}$, there exists a continuous curve $z : [0,1] \to \mathsf{X}$ such that

$$z(0) = x, \quad z(1) = y, \quad \mathsf{d}(z(s), z(t)) = (t-s)\mathsf{d}(x,y), \quad \forall 0 \le s \le t \le 1. \tag{1.3}$$

In fact, from the above identity, we conclude that $z(\cdot)$ has to be a constant-speed curve, namely

$$|z'|(r) = \mathsf{d}(x,y), \text{ for all } r \in (0,1). \tag{1.4}$$

See also Remark 1.8 for more general assumptions on X.

We set $\bar{\mathbb{R}} := [-\infty, +\infty]$, $\mathbb{R}_+ = [0, \infty)$ and use the notation $B_r(x)$ and $\overline{B}_r(x)$ respectively for open and closed balls. Let $BUC(\mathsf{X}; \mathbb{R})$ denote the space of bounded uniformly continuous functions on X, $LSC(\mathsf{X}; \bar{\mathbb{R}})$ (respectively $USC(\mathsf{X}; \bar{\mathbb{R}})$) denote the space of lower (respectively, upper) semi-continuous functions on X, $M^u(\mathsf{X}; \bar{\mathbb{R}})$ denote the space of measurable functions from $\mathsf{X} \mapsto \bar{\mathbb{R}}$ which are bounded from above.

If $g : \mathsf{X} \mapsto \bar{\mathbb{R}}$ and $\zeta : \mathsf{X} \mapsto [0, \infty]$, g is called with growth at most ζ if for some constant $C \in \mathbb{R}_+$ it holds

$$|g(x)| \le C(1 + \zeta(x)), \qquad \text{for all } x \in \mathsf{X}.$$

Let $L := L(x,q) : \mathsf{X} \times \mathbb{R}_+ \mapsto \bar{\mathbb{R}}$, we define $H : \mathsf{X} \times \mathbb{R}_+ \mapsto \bar{\mathbb{R}}$ by

$$H(x,p) := \sup_{q \ge 0} \Big(pq - L(x,q) \Big), \quad \forall p \ge 0. \tag{1.5}$$

For notational convenience, we also introduce an extension of H allowing $p < 0$:

$$\bar{H}(x,p) := \sup_{q \ge 0} \Big(pq - L(x,q) \Big), \quad \forall p \in \mathbb{R}.$$

We fix a basepoint $\bar{x}$ in X and assume:

Condition 1.1.

(1) *L is lower semicontinuous from $\mathsf{X} \times [0,\infty)$ into $\mathbb{R} \cup \{+\infty\}$ and $\inf L > -\infty$.*
(2) *$\ell(q) := \inf_{x \in \mathsf{X}} L(x,q)$ is super-linear, namely*

$$\lim_{q \to +\infty} \frac{\ell(q)}{q} = +\infty.$$

(3) *Either $L(\cdot, q) \equiv +\infty$ or it is real-valued and continuous.*

(4) *For each $R > 0$, there is a local modulus $\omega_R(r) : \mathbb{R}_+ \mapsto \mathbb{R} \cup \{+\infty\}$ (i.e. ω_R is continuous at $r = 0$ and $\omega_R(0) = 0$) and some constant $C_R > 0$ such that: for every x, $y \in B_R(\bar{x})$ and for every $q \geq 0$ with $L(x, q) < \infty$, there exists $q' \geq 0$ satisfying*

$$\frac{|q - q'|}{1 - \inf L + \ell(q)} \leq C_R \mathsf{d}(x, y),$$

and

$$\frac{L(y, q') - L(x, q)}{1 - \inf L + \ell(q)} \leq \omega_R(\mathsf{d}(x, y)).$$

(5) *There exists $q_0 > 0$ such that $L(\cdot, q_0)$ is finite on bounded sets: we denote by $\zeta : \mathbb{R}_+ \mapsto \mathbb{R}_+$ and a non-negative, non-decreasing continuous function $\zeta : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that*

$$L(x, q_0) \leq \zeta\big(\mathsf{d}(x, \bar{x})\big), \quad \forall x \in \mathsf{X}. \tag{1.6}$$

Moreover, there exists a nondecreasing non-negative function $\beta \in C^1(\mathbb{R}_+)$ which grows to infinity faster than ζ (i.e. $\lim_{r \to +\infty} \beta(r)/(1 + \zeta(r)) = +\infty$) such that

$$\sup_{x \in \mathsf{X},\, x \neq \bar{x}} H\Big(x, |D_x \beta \circ \mathsf{d}(\cdot, \bar{x})|\Big) = \sup_{x \in \mathsf{X},\, x \neq \bar{x}} H\Big(x, \beta'(\mathsf{d}(x, \bar{x}))\Big) < \infty.$$

Remark 1.2. (1) In most cases, Condition 1.1(4) can be verified by taking $q' = q$. By doing so, the first inequality is trivially satisfied. In the model case when $L(x, q) = \mathfrak{l}(q) - V(x)$, it is sufficient to have V bounded from above, $\mathfrak{l}$ bounded from below and V be uniformly continuous in bounded regions. Note that the $\mathfrak{l}$ does not need to be continuous.
(2) The motivation for the growth estimate type Condition 1.1(5) is more involved to explain. We offer the following example to illustrate its usefulness and limitation.

Example 1.3. Consider the case

$$L(x, q) := \mathfrak{l}(q) - V(x), \quad q \geq 0, \quad H(x, p) := \mathfrak{h}(p) + V(x), \quad p \geq 0.$$

where $\mathfrak{h}(p) := \sup_{q \geq 0}[qp - \mathfrak{l}(q)]$ is non-decreasing in $p \geq 0$. Assume that $\mathfrak{l} : \mathbb{R}_+ \mapsto \mathbb{R} \cup \{+\infty\}$ is super-linear and lower semi-continuous, and that V is bounded from above and uniformly continuous in balls of finite radius. Then Conditions 1.1(1)-(4) are all satisfied. In addition, we assume that $\mathfrak{l}$ is not trivial: there exists $q_0 > 0$ such that $\mathfrak{l}(q_0) < \infty$. Then, Condition 1.1(5) is satisfied in the following important situations:

(a) V is bounded: $\sup_{\mathsf{X}} |V| < \infty$. In this case, one can just take

$$\zeta(r) := \sup_{x \in \mathsf{X}} L(x, q_0), \quad \beta(r) = r.$$

(b) There exist C_0, $C_1 > 0$, $\theta \in (0, 1)$ such that

$$V(x) \geq -C_0 - C_1 \mathsf{d}^\theta(x, \bar{x}).$$

In this case, we can take $\zeta(r) = \mathfrak{l}(q_0) + C_0 + C_1 r^\theta$ and $\beta(r) = r$. Then it follows that

$$H\Big(x, \beta'(\mathsf{d}(x, \bar{x}))\Big) = V(x) + \mathfrak{h}(1) \leq \sup_{x \in \mathsf{X}} V(x) + \mathfrak{h}(1) < \infty.$$

(c) There exist C_0, C_1, C_2, $C_3 > 0$, $0 \leq \theta_2 \leq \theta_1$ such that

$$-C_0 - C_1 \mathsf{d}^{1+\theta_1}(x, \bar{x}) \leq V(x) \leq C_2 - C_3 \mathsf{d}^{1+\theta_2}(x, \bar{x}), \tag{1.7}$$

and that

$$\lim_{r\to+\infty}\frac{\mathfrak{h}(r^{\theta_3})}{r^{1+\theta_2}}=0,\quad \exists \text{ some } \theta_3>\theta_1. \tag{1.8}$$

In this case, we can take $\zeta(r)=\mathfrak{l}(q_0)+C_0+C_1r^{1+\theta_1}$, $\beta(r)=r^{1+\theta_3}$. Then

$$\begin{aligned} H\Big(x,\beta'(\mathsf{d}(x,\bar{x}))\Big) &= V(x)+\mathfrak{h}(\mathsf{d}^{\theta_3}(x,\bar{x})) \\ &\le C_2-C_3\mathsf{d}^{1+\theta_2}(x,\bar{x})+\mathfrak{h}(\mathsf{d}^{\theta_3}(x,\bar{x})) \\ &\le \sup_{r\ge 0}(C_2-C_3r^{1+\theta_2}+\mathfrak{h}(r^{\theta_3}))<\infty. \end{aligned}$$

More specifically, focusing on the example of

$$\mathfrak{l}(q)=\frac{|q|^2}{2},\quad \mathfrak{h}(p)=\frac{|p|^2}{2},$$

then (1.8) is implied by

$$\theta_1<\theta_3<\frac{1+\theta_2}{2}. \tag{1.9}$$

Assume that (1.7) holds with

$$0\le\theta_2\le\theta_1<1,\quad 2\theta_1<1+\theta_2.$$

Then one can always find such a θ_3, implying (1.7) and (1.8), consequently Condition 1.1.5.

It is important to note that the above conditions allow the case of

$$H(x,p)=|p|+V(x),\quad p\ge 0,$$

which arises if we take $L(x,q)=\mathfrak{l}(q)-V(x)$ with

$$\mathfrak{l}(1):=0,\quad \mathfrak{l}(q)=+\infty,\forall q\neq 1,$$

or

$$\mathfrak{l}(q):=0,\forall q\in[0,1],\quad \mathfrak{l}(q)=+\infty,\forall q\in(1,\infty).$$

Note that the function $H(x,p)$ in (1.5) is only defined for $p\ge 0$.

There are two types of equations that we are interested in.

1.2. **Resolvent equation.** For $\alpha>0$ and $h\in BUC(\mathsf{X})$, we define the value function

$$f(x) := \sup\Big\{\int_0^\infty e^{-r/\alpha}\Big(\frac{h(z)}{\alpha}-L(z,|z'|)\Big)dr : z(\cdot)\in AC_{\rm loc}(\mathbb{R}_+;\mathsf{X}),\ z(0)=x\Big\}. \tag{1.10}$$

We expect that f formally solves the equation

$$f(x)-\alpha H\Big(x,|Df(x)|\Big)=h(x). \tag{1.11}$$

To make this precise, we proceed in several steps. First, we define two collections of test functions D_0, $D_1\subset C(\mathsf{X})$. By D_0, we mean the class of functions of the following form

$$\varphi(\cdot) := \lambda^{-1}\Big(\frac{\mathsf{d}^2(\cdot,y)}{2\delta}+\kappa\beta\circ\mathsf{d}(\cdot,\bar{x})+\epsilon\mathsf{d}(\cdot,x_0)\Big) \tag{1.12}$$

where $\lambda>1$, y, $x_0\in\mathsf{X}$, and $\delta>0$, κ, $\epsilon\ge 0$ are free parameters, while the function β is the one in Condition 1.1(5). We further denote $D_0^{\lambda,\kappa,\epsilon}$ the subset of functions in D_0 built with parameters λ, κ and ϵ, we define an operator $H_0^{\lambda,\kappa,\epsilon}$ which formally estimates $H(x, |D\varphi(x)|)$ from above

$$H_0^{\lambda,\kappa,\epsilon}\varphi(x) := H\Big(x, \frac{1}{\lambda}\Big(\Big|D\frac{\mathsf{d}^2(\cdot, y)}{2\delta}(x)\Big| + \kappa\beta' \circ \mathsf{d}(x, \bar{x}) + \epsilon\Big)\Big) \qquad \varphi \in D_0^{\lambda,\kappa,\epsilon}.$$

By D_1, we mean the collection of test functions of the form

$$\psi(\cdot) := -\frac{\mathsf{d}^2(x, \cdot)}{2\delta} - \kappa\beta \circ \mathsf{d}(\cdot, \bar{x}) - \epsilon\mathsf{d}(\cdot, y_0), \tag{1.13}$$

where β is still the same function, x, $y_0 \in \mathsf{X}$, and $\delta > 0$, κ, $\epsilon \geq 0$. Denoting by $D_1^{\kappa,\epsilon}$ the subset of functions in D_1 built with parameters κ and ϵ, we define another operator $H_1^{\kappa,\epsilon}$ which formally estimates $H(y, |D\psi(y)|)$ from below

$$H_1^{\kappa,\epsilon}\psi(y) := \bar{H}\Big(y, \Big|D\frac{\mathsf{d}^2(x, \cdot)}{2\delta}(y)\Big| - \kappa\beta' \circ \mathsf{d}(y, \bar{x}) - \epsilon\Big) \qquad \psi \in D_1^{\kappa,\epsilon}.$$

Notice the scaling parameter $\lambda > 1$ presents only in D_0, and that the two collections of test functions are of opposite signs.

Definition 1.4 (Resolvent equation)**.** A function f is called a viscosity sub-solution to (1.11), if for each $\varphi \in D_0^{\lambda,\kappa,\epsilon}$, and for each $x_0 \in \mathsf{X}$ such that $(f - \varphi)(x_0) = \sup_{x\in\mathsf{X}}(f - \varphi)(x)$, we have

$$\alpha^{-1}(f - h)(x_0) \leq (H_0^{\lambda,\kappa,\epsilon}\varphi)^*(x_0),$$

where g^* denotes upper semicontinuous envelope of g. Analogously, a function f is called a viscosity super-solution to (1.11), if for each $\psi \in D_1^{\kappa,\epsilon}$, and for each $y_0 \in \mathsf{X}$ such that $(\psi - f)(y_0) = \sup_{y\in\mathsf{X}}(\psi - f)(y)$, we have

$$\alpha^{-1}(f - h)(y_0) \geq (H_1^{\kappa,\epsilon}\psi)_*(y_0),$$

where g_* denotes the lower semicontinuous envelope of g.

The specific form of the functions in D_0 and D_1 is motivated by Ekeland's perturbed optimization principle, which does not need compactness or local compactness of the space (Proposition 5.1 in the Appendix). As a matter of fact, if the ambient space is compact one can work fixing the parameters $\kappa = \epsilon = 0$, and if bounded closed sets are compact, one can work with $\epsilon = 0$. Suppose that sub- and super-solutions are at most growth of $\zeta \circ \mathsf{d}(\cdot, \bar{x})$, where β has a slower growth than ζ (as in Condition 1.1(5)), then such principle guarantees that, with every given δ, κ, ϵ, we can always choose x_0 such that it becomes the global strict maximum of $f - \varphi$. The case of ψ is similar.

Our first result is:

Theorem 1.5. *The value function f defined in* (1.10) *is the only continuous viscosity solution of* (1.11) *with growth at most $\zeta \circ \mathsf{d}(\cdot, \bar{x})$.*

This proof is a combination of Lemma 2.3 and Lemma 2.15.

1.3. **The Cauchy problem in finite time** $[0, T]$**.** Let $h \in BUC(\mathsf{X})$. We define

$$U(t, x) := \sup\Big\{h(x(t)) - \int_0^t L(x, |x'|)dr : x(\cdot) \in AC([0, t]; X),\ x(0) = x\Big\}. \tag{1.14}$$

Formally, it solves the Cauchy problem written as

$$\partial_t U(t, x) = H\Big(x, |D_x U|(t, x)\Big), \text{ for } t \in (0, T], \quad U(0, x) = h(x). \tag{1.15}$$

We make this precise, developing a well-posedness theory by extending viscosity solution techniques.

We extend the classes of test functions in D_0, D_1 by adding a time variable. The set $\hat{D}_0^{\lambda,\kappa,\epsilon}$ consists of test functions

$$\varphi := \varphi(t,x) = \phi(t) + \varphi(x) \tag{1.16}$$

where $\phi : [0,T] \mapsto \mathbb{R}$ is Lipschitz and $\varphi \in D_0^{\lambda,\kappa,\epsilon}$. We denote $\hat{D}_0 := \cup_{\lambda>1,\kappa,\epsilon>0} \hat{D}_0^{\lambda,\kappa,\epsilon}$. Similarly, the class $\hat{D}_1^{\kappa,\epsilon}$ consists of

$$\psi := \psi(s,y) = \phi(t) + \psi(y) \tag{1.17}$$

where ϕ is locally Lipschitz and $\psi \in D_1^{\kappa,\epsilon}$. The class $\hat{D}_1$ is defined analogously. For locally Lipschitz function $\phi : [0,T] \mapsto \mathbb{R}$, we denote

$$\partial_t^+ \phi(t) := \limsup_{s\to t} \frac{\phi(s)-\phi(t)}{s-t}, \quad \partial_t^- \phi(t) := \liminf_{s\to t} \frac{\phi(s)-\phi(t)}{s-t}.$$

Definition 1.6 (Cauchy Problem)**.** Let $U : [0,T] \times \mathsf{X} \mapsto \bar{\mathbb{R}}$.

U is called a viscosity sub-solution to (1.15), if

(a) (Initial condition) For all $C \in \mathbb{R}_+$ it holds

$$\limsup_{t\to 0^+} \sup_{\{x\in \mathsf{X}:\ \zeta\circ \mathsf{d}(x,\bar{x})\le C\}} \Big(U(t,x) - h(x) \Big) \le 0.$$

(b) For each $\varphi \in \hat{D}_0^{\lambda,\kappa,\epsilon}$ and for each $(t_0,x_0) \in (0,T] \times \mathsf{X}$ such that

$$(U-\varphi)(t_0,x_0) = \sup_{[0,T]\times\mathsf{X}} (U-\varphi),$$

we have

$$\Big((-\partial_t^- + H_0^{\lambda,\kappa,\epsilon})\varphi \Big)^*(t_0,x_0) \ge 0.$$

U is called a super-solution to (1.15), if

(a) (Initial condition) For all $C \in \mathbb{R}_+$ it holds

$$\liminf_{s\to 0^+} \inf_{\{y\in \mathsf{X}:\ \zeta\circ \mathsf{d}(y,\bar{x})\le C\}} \Big(U(s,y) - h(y) \Big) \ge 0.$$

(b) For each $\psi \in \hat{D}_1^{\kappa,\epsilon}$ and for each $(s_0,y_0) \in (0,T] \times \mathsf{X}$ such that

$$(\psi-U)(s_0,y_0) = \sup_{[0,T]\times\mathsf{X}} (\psi-U),$$

we have

$$\Big((-\partial_s^+ + H_1^{\kappa,\epsilon})\psi \Big)_*(s_0,y_0) \le 0.$$

If U is both a sub- and super- solution, then it is called a solution.

Theorem 1.7 (Cauchy problem)**.** *Suppose that $h \in BUC(\mathsf{X})$. Then the function U defined in* (1.14) *is continuous, bounded from above, and has growth at most $\zeta \circ \mathsf{d}(\cdot,\bar{x})$ from below. It is also the unique continuous viscosity solution to* (1.15) *satisfying the above properties.*

The proof is a combination of Lemma 3.5 and Lemma 3.8.

Remark 1.8 (Length spaces)**.** It is interesting to remark that the geodesic assumption can be relaxed to a length space assumption, namely for all x, $y \in \mathsf{X}$ the infimum of the length of continuous curves joining x to y is equal to $\mathsf{d}(x,y)$. For proper spaces (i.e. when closed balls are compact) the two notions are equivalent, but in general geodesic is strictly stronger than length. In all proofs suffices to replace geodesics with curves having almost minimal length to obtain, with minor modifications, the same results. The key property $|D^-\mathsf{d}^2(\cdot,y)|(x) = \mathsf{d}(x,y)$ (see Lemma 2.1) we use of the slope of the distance function is still true for length spaces with minor modifications in the proof, see for instance [28, Lemma 2.8].

2. The resolvent problem

2.1. **Slope estimates.** We introduce two more notions of slope

$$|D^+f|(x) := \limsup_{z\to x}\frac{(f(z)-f(x))^+}{\mathsf{d}(z,x)}, \quad |D^-f|(x) := \limsup_{z\to x}\frac{(f(z)-f(x))^-}{\mathsf{d}(z,x)}.$$

We use convention $0/0 = 0$. It follows that $|Df|(x) = \max\{|D^+f|(x), |D^-f|(x)\}$. The following is an elementary but important property of geodesic spaces (see also Remark 1.8), which we will use critically in the proof of comparison principle for viscosity solutions. We include the proof for the reader's convenience.

Lemma 2.1.

$$\left|D_x^+\frac{\mathsf{d}^2}{2}(x,y)\right| \le \mathsf{d}(x,y), \quad \left|D_x^-\frac{\mathsf{d}^2}{2}(x,y)\right| = \mathsf{d}(x,y).$$

Hence

$$\left|D_x\frac{\mathsf{d}^2}{2}(x,y)\right| = \mathsf{d}(x,y).$$

Proof. Fix x, $y \in \mathsf{X}$. By triangle inequality,

$$\left|\mathsf{d}^2(z,y)-\mathsf{d}^2(x,y)\right| = (\mathsf{d}(z,y)+\mathsf{d}(x,y))\left|\mathsf{d}(z,y)-\mathsf{d}(x,y)\right| \le (\mathsf{d}(z,y)+\mathsf{d}(x,y))\mathsf{d}(z,x).$$

Hence $|D_x^{\pm}\mathsf{d}^2(x,y)| \le 2\mathsf{d}(x,y)$.

Next, we prove that $|D_x^-\mathsf{d}^2(x,y)| \ge 2\mathsf{d}(x,y)$. We only need to prove the case when $\mathsf{d}(x,y) > 0$. Take $w(t)$ to be a constant speed geodesic with

$$w(0) = x, \quad w(1) = y, \quad \mathsf{d}(w(s),x) = s\mathsf{d}(x,y), \quad \mathsf{d}(w(s),y) = (1-s)\mathsf{d}(x,y).$$

Then

$$\frac{\mathsf{d}^2(y,x)-\mathsf{d}^2(y,w(s))}{\mathsf{d}(w(s),x)} = \frac{\mathsf{d}^2(x,y)(1-(1-s)^2)}{s\mathsf{d}(x,y)} = 2\mathsf{d}(x,y) - s\mathsf{d}(x,y)$$

and taking the limit as $s \downarrow 0$ provides the inequality. □

2.2. **Comparison principles.** Let $\lambda > 1$, δ, κ, $\epsilon \in (0,1)$ and y, $x_0 \in \mathsf{X}$. We consider test functions $\varphi(x)$ of the form (1.12). Then

$$|D\varphi(x)| \;\le\; \frac{1}{\lambda}\left(\frac{\mathsf{d}(x,y)}{\delta} + \kappa\beta'\circ\mathsf{d}(x,\bar{x}) + \epsilon\right).$$

Therefore, since $p \mapsto H(x,p)$ is non-decreasing in $p \geq 0$, we get

$$\begin{aligned}
\lambda\Big(H_0^{\lambda,\kappa,\epsilon}\varphi\Big)^*(x) &\leq \limsup_{\hat{x}\to x} \lambda H\Big(\hat{x}, \frac{1}{\lambda}\Big(\frac{\mathsf{d}(\hat{x},y)}{\delta} + \kappa\beta' \circ \mathsf{d}(\hat{x},\bar{x}) + \epsilon\Big)\Big) \\
&= \limsup_{\hat{x}\to x} \sup_{q\geq 0} \Big(\big(\frac{\mathsf{d}(\hat{x},y)}{\delta} + \kappa\beta' \circ \mathsf{d}(\hat{x},\bar{x}) + \epsilon\big)q - \lambda L(\hat{x},q)\Big) \\
&\leq \sup_{q\geq 0} \Big(\big(\frac{\mathsf{d}(x,y)}{\delta} + \kappa\beta' \circ \mathsf{d}(x,\bar{x}) + \epsilon\big)q - \lambda L(x,q)\Big).
\end{aligned}$$

In deriving the last step, we used the fact that $\ell(q)$ is super-linear in q and that $(x,q) \mapsto L(x,q)$ is lower semicontinuous.

Similarly, for $\psi(y)$ of the form (1.13) with the parameters x, y_0 and ϵ, δ, κ, we may use $|D\mathsf{d}^2(\cdot,y)| = 2\mathsf{d}(\cdot,y)$ to get

$$(H_1^{\kappa,\epsilon}\psi)_*(y) \geq \sup_{q\geq 0}\Big[\Big(\frac{\mathsf{d}(x,y)}{\delta} - \kappa\beta' \circ \mathsf{d}(y,\bar{x}) - \epsilon\Big)q - L(y,q)\Big].$$

Lemma 2.2. *Assume that positive parameters κ, λ, R are fixed with $1+4\kappa < \lambda$, and consider functions $\varphi_\delta \in D_0$, $\psi_\delta \in D_1^{\kappa,\delta}$ built with these parameters and $y = y_\delta$ for φ_δ, $x = x_\delta$ for ψ_δ. Then, if x_δ, $y_\delta \in B_R(\bar{x})$ and $\lim_\delta \mathsf{d}^2(x_\delta,y_\delta)/\delta = 0$, one has*

$$\begin{aligned}
&\limsup_{\delta\to 0^+}\Big(\lambda(H_0^{\lambda,\kappa,\delta}\varphi_\delta)^*(x_\delta) - (H_1^{\kappa,\delta}\psi_\delta)_*(y_\delta)\Big) \\
\leq\ & (\lambda - 1 - 2\kappa)\max\{0, -\inf L\} + 2\kappa \sup_{x\in \mathsf{X}} H\Big(x, \beta' \circ \mathsf{d}(x,\bar{x})\Big).
\end{aligned}$$

Proof. Using the above estimates, suffices to estimate from above, uniformly in q as $\delta \downarrow 0$, the difference

$$\Big(\big(\frac{\mathsf{d}(x_\delta,y_\delta)}{\delta} + \kappa\beta' \circ \mathsf{d}(x_\delta,\bar{x}) + \delta\big)q - \lambda L(x_\delta,q)\Big) - \Big(\big(\frac{\mathsf{d}(x_\delta,y_\delta)}{\delta} - \kappa\beta' \circ \mathsf{d}(y_\delta,\bar{x}) - \delta\big)\hat{q} - L(y_\delta,\hat{q})\Big),$$

choosing conveniently $\hat{q}$ in terms of q. Writing it in the more convenient form

$$\frac{\mathsf{d}(x_\delta,y_\delta)}{\delta}(q-\hat{q}) + \kappa\beta' \circ \mathsf{d}(x_\delta,\bar{x})q + \kappa\beta' \circ \mathsf{d}(y_\delta,\bar{x})\hat{q} + \delta(q+\hat{q}) - \lambda L(x_\delta,q) + L(y_\delta,\hat{q})$$

and using Condition 1.1(4) for the choice of $\hat{q}$, we can further estimate

$$\begin{aligned}
&C_R\frac{\mathsf{d}^2(x_\delta,y_\delta)}{\delta}\Big(1 - \inf L + L(x_\delta,q)\Big) + \kappa\beta' \circ \mathsf{d}(x_\delta,\bar{x})q + \kappa\beta' \circ \mathsf{d}(y_\delta,\bar{x})q \\
&\quad + \kappa\beta' \circ \mathsf{d}(y_\delta,\bar{x})C_R\mathsf{d}(x_\delta,y_\delta)\Big(1 - \inf L + \ell(q)\Big) + 2\delta q + \delta C_R \mathsf{d}(x_\delta,y_\delta)\Big(1 - \inf L + \ell(q)\Big) \\
&\quad + \omega_R(\mathsf{d}(x_\delta,y_\delta))\Big(1 - \inf L + \ell(q)\Big) - (\lambda-1)L(x_\delta,q).
\end{aligned}$$

Now we see that in this expression the leading term is $L(x_\delta,q)$, which appears with the negative factor $(1-\lambda)$, while in all other terms q appears either linearly or with an infinitesimal factor times $\ell(q)$, which is smaller than $L(x_\delta,q)$. Therefore, this proves that we can restrict ourselves to a set of uniformly bounded q's. Hence, taking limits, by our assumptions on x_δ and y_δ, we have only to take care of the terms

$$\kappa\beta' \circ \mathsf{d}(x_\delta,\bar{x})q + \kappa\beta' \circ \mathsf{d}(y_\delta,\bar{x})q - (\lambda-1)L(x_\delta,q). \tag{2.1}$$

Now, adding and subtracting $2\kappa L(x_\delta,q)$, we reach the conclusion. $\square$

We are now ready to prove a comparison principle, under the assumptions of Condition 1.1.

Lemma 2.3 (Comparison principle)**.** *Let $\overline{f} \in USC(\mathsf{X};\mathbb{R})$ be an upper semicontinuous sub-solution to* (1.11) *with h replaced by $h_0 \in C_b(\mathsf{X})$, with growth from above at most $\zeta \circ \mathsf{d}(\cdot, \bar{x})$. Let $\underline{f} \in LSC(\mathsf{X};\mathbb{R})$ be a lower semicontinuous super-solution to* (1.11) *with h replaced by $h_1 \in C_b(\mathsf{X})$, with growth from below at most $\zeta \circ \mathsf{d}(\cdot, \bar{x})$.*
If either h_0 or h_1 are uniformly continuous, it holds

$$\sup(\overline{f} - \underline{f}) \leq \sup(h_0 - h_1).$$

Proof. Let $\lambda > 1$, $\kappa > 0$ be fixed, with $\lambda > 1 + 4\kappa$; we will take limits with respect to these parameters only at the end of the proof, so we ignore this dependence, emphasizing instead the dependence on δ. Let

$$\Psi_\delta(x, y) := \lambda \overline{f}(x) - \underline{f}(y) - \frac{1}{2\delta}\mathsf{d}^2(x, y) - \kappa\beta \circ \mathsf{d}(x, \bar{x}) - \kappa\beta \circ \mathsf{d}(y, \bar{x}), \quad \forall x\,, y \in \mathsf{X},$$

where β is the one in Condition 1.1(5). By the growth conditions on the sub- and super-solutions, $\sup_{\mathsf{X}\times\mathsf{X}} \Psi_\delta < \infty$. In particular, we can find $\bar{x}_\delta$ and $\bar{y}_\delta \in \mathsf{X}$ such that

$$\Psi_\delta(\bar{x}_\delta, \bar{y}_\delta) \geq \sup_{\mathsf{X}\times\mathsf{X}} \Psi_\delta - \delta. \tag{2.2}$$

By the above mentioned Ekeland's perturbed optimization principle (applied in $\mathsf{X} \times \mathsf{X}$ with the distance $\mathsf{d}(x_1, y_1) + \mathsf{d}(x_2, y_2)$ between pairs (x_1, y_1), (x_2, y_2)), we can then further find a point $(x_\delta, y_\delta) \in \mathsf{X} \times \mathsf{X}$ with $\mathsf{d}(\bar{x}_\delta, x_\delta) + \mathsf{d}(\bar{y}_\delta, y_\delta) \leq 1$ such that (x_δ, y_δ) is the strict global maximum M_δ of

$$(x, y) \mapsto \Psi_\delta(x, y) - \delta\mathsf{d}(x, x_\delta) - \delta\mathsf{d}(y, y_\delta).$$

In particular,

$$\Psi_\delta(\bar{x}_\delta, \bar{y}_\delta) \leq \Psi_\delta(x_\delta, y_\delta) + \delta\mathsf{d}(\bar{x}_\delta, x_\delta) + \delta\mathsf{d}(\bar{y}_\delta, y_\delta) \leq \Psi_\delta(x_\delta, y_\delta) + \delta. \tag{2.3}$$

Consequently, since $M_\delta = \Psi_\delta(x_\delta, y_\delta)$, we get

$$M_\delta \leq \sup_{\mathsf{X}\times\mathsf{X}} \Psi_\delta \leq M_\delta + 2\delta. \tag{2.4}$$

The growth assumptions on $\overline{f}$ and $\underline{f}$ and β (with respect to ζ) imply

$$R = R_{\kappa,\lambda} := \sup_{\delta\in(0,1)} \mathsf{d}(x_\delta, \bar{x}) + \mathsf{d}(y_\delta, \bar{x}) < \infty. \tag{2.5}$$

Notice that $\sup \Psi_\delta \leq \sup \Psi_{\delta'}$ for $0 < \delta < \delta'$; consequently (2.4) gives $M_\delta \leq M_{\delta'} + 2\delta'$ for $0 < \delta < \delta'$, so that M_δ has a limit as $\delta \downarrow 0$. Evaluating Ψ_δ on the diagonal and using (2.4) give that the limit is finite. On the other hand, using (x_δ, y_δ) as an admissible point in the maximization of $\Psi_{2\delta}$ and (2.4) gives

$$M_{2\delta} - M_\delta \geq \big(\frac{1}{2\delta} - \frac{1}{4\delta}\big)\mathsf{d}^2(x_\delta, y_\delta) - 4\delta,$$

so that

$$\lim_{\delta\downarrow 0} \frac{1}{\delta}\mathsf{d}^2(x_\delta, y_\delta) = 0. \tag{2.6}$$

Set φ as in (1.12) with y replaced by y_δ, x_1 by x_δ. Then

$$(\overline{f} - \varphi)(x_\delta) = \sup_{\mathsf{X}}(\overline{f} - \varphi),$$

so that, by viscosity sub-solution property, we have

$$\lambda \frac{\overline{f} - h_0}{\alpha}(x_\delta) \leq \lambda (H_0^{\lambda,\kappa,\delta}\varphi)^*(x_\delta).$$

Similarly, set ψ as in (1.13) with x replaced by x_δ, y_1 by y_δ. Then

$$(\psi - \underline{f})(y_\delta) = \sup_{\mathsf{X}}(\psi - \underline{f}),$$

so that, by viscosity super-solution property, we have

$$\frac{\underline{f} - h_1}{\alpha}(y_\delta) \geq (H_1^{\kappa,\delta}\psi)_*(y_\delta).$$

Using first (2.2), then (2.3) and the above mentioned sub- and super- solution bounds, choosing $x = y$ we get

$$\begin{aligned}
\lambda \overline{f}(x) - \underline{f}(x) - 2\kappa\beta \circ \mathsf{d}(x,\bar{x}) - 2\delta &\leq \Psi_\delta(x,x) - 2\delta \leq \Psi_\delta(\bar{x}_\delta, \bar{y}_\delta) - \delta \leq \Psi_\delta(x_\delta, y_\delta) \\
&\leq \lambda \overline{f}(x_\delta) - \underline{f}(y_\delta) \\
&\leq \lambda h_0(x_\delta) - h_1(y_\delta) + \alpha\Big(\lambda (H_0^{\lambda,\kappa,\delta}\varphi)^*(x_\delta) - (H_1^{\kappa,\delta}\psi)_*(y_\delta)\Big).
\end{aligned}$$

Without loss of generality, we assume that h_1 is uniformly continuous with a modulus ω_{h_1} (otherwise the argument is similar), so that the last term above does not exceed

$$(\lambda - 1)\sup_{\mathsf{X}} |h_0| + \sup_{\mathsf{X}}(h_0 - h_1) + \omega_{h_1}(\mathsf{d}(x_\delta, y_\delta))) + \alpha\Big(\lambda (H_0^{\lambda,\kappa,\delta}\varphi)^*(x_\delta) - (H_1^{\kappa,\delta}\psi)_*(y_\delta)\Big).$$

Noting (2.5) and (2.6), we can first let $\delta \downarrow 0$ and then use the estimate in Lemma 2.2, letting $\kappa \to 0$ and finally $\lambda \downarrow 1$. $\square$

2.3. **Semigroup property and some estimates.** First, we give some growth estimate on the $f := f_\alpha$ in (1.10).

For $h \in M^u(\mathsf{X}; \bar{\mathbb{R}})$, the space of measurable functions which are bounded from above, we define

$$T_t h(x) := \sup \Big\{ h(z(t)) - \int_0^t L(z(r), |z'|(r))dr : z \in AC([0,t]; \mathsf{X}),\ z(0) = x \Big\} \tag{2.7}$$

for every $t \geq 0$. From the boundedness from below of L, we know that $T_t h$ is well defined and that

$$T_t : M^u(\mathsf{X}; \bar{\mathbb{R}}) \mapsto M^u(\mathsf{X}; \bar{\mathbb{R}}).$$

Lemma 2.4 (Semigroup property)**.** *For $s, t \geq 0$ it holds*

$$T_s T_t f = T_{s+t} f, \quad f \in M^u(\mathsf{X}; \bar{\mathbb{R}}).$$

In addition,

$$\begin{aligned}
T_t(h + c) &= (T_t h) + c, \quad \forall h \in M^u(\mathsf{X}; \bar{\mathbb{R}}), c \in \mathbb{R}, \\
T_t h &\leq T_t g, \quad \textit{whenever } h \leq g, \quad h,\, g \in M^u(\mathsf{X}; \bar{\mathbb{R}}).
\end{aligned}$$

Finally, T is a a contraction in $B(\mathsf{X})$:

$$\sup_{\mathsf{X}} |T_t h_1 - T_t h_2| \leq \sup_{\mathsf{X}} |h_1 - h_2|, \quad h_1,\, h_2 \in B(\mathsf{X}).$$

Proof. Since the concatenation of two absolutely continuous paths gives another absolutely continuous path, the standard proof for the semigroup property in the case of $\mathsf{X} = \mathbb{R}^d$ transfers verbatim. Note that existence of optimal paths is not needed, using ϵ-optimal ones is sufficient. For more details, one can adapt the proof of Lemma 5.7 and the first two parts of Proposition 5.8 in [17] to the current setting. $\square$

Lemma 2.5 (Dynamic programming)**.** *The function* f_α *defined in* (1.10) *satisfies*

$$f_\alpha(x) \;=\; \sup\Big\{ \int_0^t e^{-s/\alpha}\Big(\frac{h(z(s))}{\alpha} - L(z(s),|z'|(s))\Big)ds + e^{-t/\alpha} f_\alpha(z(t)) : \\ z(\cdot) \in AC([0,t];\mathsf{X}),\ z(0)=x\Big\}, \quad \forall x \in \mathsf{X}.$$

Proof. It follows from similar arguments as in the proof of semigroup property. $\square$

Next, we give some local Lipschitz estimate on f_α.

Lemma 2.6 (Local Lipschitz estimate)**.** *If* $\sup_{\mathsf{X}} |h| < \infty$*, then* f_α *is continuous. More precisely, we have the local Lipschitz estimate*

$$|f_\alpha(x) - f_\alpha(y)| \le (e^{\mathsf{d}(x,y)/(q_0\alpha)} - 1)\Big(2\sup_{\mathsf{X}}|h| - \alpha \inf L + \alpha\zeta(\mathsf{d}(\bar{x},x) + \mathsf{d}(x,y) + \mathsf{d}(y,\bar{x}))\Big),$$

where $q_0 > 0$ *and* ζ *are given in* (1.6).

Proof. Let $x \neq y \in \mathsf{X}$. First, by Condition 1.1(5), there exists q_0 such that $\ell(q_0) < \infty$. Take $z = z(r)$ to be a constant speed geodesic between x and y in time interval $(0,\delta)$ with $q_0\delta := \mathsf{d}(x,y)$:

$$z : [0,\delta] \mapsto \mathsf{X}, \quad z(0) = x, \quad z(\delta) = y, \quad |z'| \equiv q_0 \text{ in } (0,\delta).$$

By dynamic programming, we have

$$f_\alpha(x) \ge \int_0^\delta \frac{e^{-r/\alpha}}{\alpha} h(z(r))dr - \int_0^\delta e^{-r/\alpha} L(z,|z'|)dr + e^{-\delta/\alpha} f_\alpha(z(\delta)).$$

Then

$$f_\alpha(y) - f_\alpha(x) \;\le\; (e^{\delta/\alpha}-1)f_\alpha(x) + (e^{\delta/\alpha}-1)\sup_{z\in\mathsf{X}}|h(z)| \\ +\alpha(e^{\delta/\alpha}-1)\sup_{r\in[0,\delta]} L(z(r),q_0).$$

By (1.6), for $r \in [0,\delta]$,

$$L(z(r),q_0) \le \zeta \circ \mathsf{d}(z(r),\bar{x}) \le \zeta \circ \Big(\mathsf{d}(\bar{x},x) + \mathsf{d}(x,z(r)) + \mathsf{d}(z(r),y) + \mathsf{d}(y,\bar{x})\Big).$$

$\square$

Lemma 2.7. *For any* $x \in \mathsf{X}$ *we have*

$$\inf_{\mathsf{X}} h - \alpha\zeta \circ (\mathsf{d}(x,\bar{x})) \le f_\alpha(x) \le \sup_{\mathsf{X}} h - \alpha \inf L,$$

where ζ *is the function in* (1.6).

Proof. For every absolutely continuous path z, $-L(z,|z'|) \le -\inf L$, giving the upper bound estimate of f.

For the lower bound, by Condition 1.1(5), let q_0 be the positive number appearing in (1.6) holds. For $x \in \mathsf{X} \setminus \{\bar{x}\}$, we construct a path $z(\cdot) \in AC(\mathbb{R}_+;\mathsf{X})$ as follows: let $\delta := \mathsf{d}(x,\bar{x})/q_0 > 0$ and define $z(r)$ to be the constant speed geodesic with $z(0) = x$, $z(\delta) = \bar{x}$, $|z'|(r) \equiv q_0$. Then on $r \in [\delta, 2\delta]$, we let $z(r) := z(2\delta - r)$ be the reversal of the path in $[0,\delta]$; and we continue this process to define a locally absolutely continuous path in $\mathbb{R}_+$. It follows in particular, except for $r = k\delta$ for $k = 0,1,2,\ldots$,

$$L(z(r),|z'|(r)) = L(z(r),q_0) \le \zeta \circ \mathsf{d}(z(r),\bar{x}) \le \zeta \circ \mathsf{d}(x,\bar{x}).$$

Therefore

$$\begin{aligned} f(x) &\ge \inf_{\mathsf{X}} h - \int_0^\infty e^{-r/\alpha} L(z(r),q)dr \ge \inf_{\mathsf{X}} h - \int_0^\infty e^{-r/\alpha} \zeta \circ \mathsf{d}(z(r),\bar{x})dr \\ &\ge \inf_{\mathsf{X}} h - \alpha\zeta \circ (\mathsf{d}(x,\bar{x})). \end{aligned}$$

If $x = \bar{x}$ we use the continuity of f to conclude. □

Following the same proof of Lemma 2.7, we also have the following estimates for T_t.

Lemma 2.8. *For each $h \in M^u(\mathsf{X},\bar{\mathbb{R}})$, $t \ge 0$ and $x \in \mathsf{X}$ it holds*

$$\inf_{B_r(x)} h - t\zeta \circ \mathsf{d}(x,\bar{x}) \le T_t h(x) \le \sup_{\mathsf{X}} h - t \inf L,$$

where $r := \mathsf{d}(x,\bar{x})$.

The function φ of the form (1.12), defining D_0, are never bounded from above, unless X has finite diameter. For each $x \in \mathsf{X}$, we introduce the following localization: for $M > 0$, let $\eta_M \in C^1(\mathbb{R}_+)$ be such that $0 \le \eta_M'(r) \le 1$, and

$$\eta_M(r) = r, \ \ r \le M, \qquad \eta_M(r) = 2M, \ \ r \ge 3M. \tag{2.8}$$

Using η_M we define

$$\varphi_M(x) = \eta_M \circ \varphi(x). \tag{2.9}$$

Then, since $\varphi_M \le 2M$, φ_M is bounded from above. Moreover, $|D\varphi(x)| = |D\varphi_M(x)|$ whenever $\varphi(x) < M$.

Lemma 2.9 (Upper estimate for the generator)**.** *For $\varphi \in D_0^{\lambda,\kappa,\epsilon}$ of the form (1.12) and $x \in \mathsf{X}$, it holds*

$$\limsup_{t \to 0^+} t^{-1}\Big(T_t\varphi_M(x) - \varphi_M(x)\Big) \le (H_0^{\lambda,\kappa,\epsilon}\varphi)^*(x) \qquad \forall M > \varphi(x).$$

Proof. Since φ_M is locally Lipschitz it is well known and easy to check (see for instance [2, Theorem 1.2.5]) that $|D\varphi_M|$ is a strong upper gradient of φ_M, namely $|\varphi_M(a) - \varphi_M(b)| \le \int_a^b |D\varphi_M|(z(r))|z'|(r)dr$ for all $z \in AC([a,b];\mathsf{X})$. Consequently, for each $t > 0$ and $x \in \mathsf{X}$, we can find $z = z_t \in AC([0,t];\mathsf{X})$ with $z(0) = x$ such that

$$\begin{aligned} T_t\varphi_M(x) - \varphi_M(x) &\leq t^2 + \varphi_M(z(t)) - \varphi_M(z(0)) - \int_0^t L(z(r), |z'|(r))dr \\ &\leq t^2 + \int_0^t \Big(|D\varphi_M(z(r))||z'|(r) - L(z(r), |z'|(r))\Big)dr \\ &\leq t^2 + \int_0^t \Big(H(\cdot, |D\varphi_M|(\cdot))\Big)^*(z(r))dr, \end{aligned}$$

where H is defined in (1.5). From the first inequality in the above,

$$\int_0^t \ell(|z'|(r))dr \leq \int_0^t L(z(r), |z'|(r))dr \leq t^2 + \varphi_M(z(t)) - T_t\varphi_M(x) \leq C_M,$$

where $C_M > 0$ is independent of $t \in [0,1]$. Note that (e.g. Theorem 1.1.2 [2]),

$$\mathsf{d}(z(s), x) = \mathsf{d}(z(s), z(0)) \leq \int_0^s |z'|(r)dr \leq \int_0^t |z'|(r)dr, \quad \forall s \in [0,t].$$

By super-linearity growth assumption on ℓ we get

$$\lim_{t\to 0^+} \sup_{0\leq s\leq t} \mathsf{d}(z_t(s), x) = 0.$$

Hence

$$\limsup_{t\to 0^+} t^{-1}\Big(T_t\varphi_M(x) - \varphi_M(x)\Big) \leq \Big(H(\cdot, |D\varphi_M|(\cdot))\Big)^*(x).$$

Since $|D\varphi_M|(x) = \eta'_M \circ \varphi(x)|D\varphi|(x)$, for $M > \varphi(x)$ we obtain $|D\varphi_M|(x) = |D\varphi|(x)$ and the conclusion follows. ◻

Lemma 2.10 (Lower estimate for the generator)**.** *For $\psi \in D_1^{\kappa,\epsilon}$ and $y \in \mathsf{X}$, we have*

$$\liminf_{t\to 0^+} t^{-1}\Big(T_t\psi(y) - \psi(y)\Big) \geq H_1^{\kappa,\epsilon}\psi(y).$$

Proof. Fix $y \in \mathsf{X}$ and $q \geq 0$. For each $t > 0$, we consider those $x \in \mathsf{X}$ on sphere of radius qt to center y:

$$qt = \mathsf{d}(x,y).$$

Because X contains more than one point and it is a geodesic space, when t is small enough, one can always find such $x \in \mathsf{X}$. We select a constant-speed geodesic z satisfying

$$z : [0,t] \mapsto \mathsf{X}, \quad z(0) = y, \quad z(t) = x, \quad |z'| \equiv q \text{ in } (0,t).$$

Recall that

$$\psi(\cdot) := -\frac{\mathsf{d}^2(\cdot, x)}{2\delta} - \kappa\beta \circ \mathsf{d}(\cdot, \bar{x}) - \epsilon\mathsf{d}(\cdot, y_1).$$

By the definition of T_t and by optimizing x over the sphere,

$$\begin{aligned} T_t\psi(y) - \psi(y) &\geq \sup_z \Big(\frac{\psi(z(t)) - \psi(y)}{\mathsf{d}(x,y)}\mathsf{d}(x,y) - \int_0^t L(z(r), q)dr\Big) \\ &= \sup_{x\in\mathsf{X}:\mathsf{d}(x,y)=tq} \Big(\frac{\psi(x) - \psi(y)}{\mathsf{d}(x,y)}\Big)\int_0^t q dr - \int_0^t L(z(r), q)dr. \end{aligned}$$

Using rough estimate $|D\mathsf{d}(\cdot,z)|\le 1$, we have that

$$\limsup_{s\to 0^+}\sup_{x\in\mathsf{X}:\mathsf{d}(x,y)=s}\Big(\frac{\psi(x)-\psi(y)}{\mathsf{d}(x,y)}\Big)\ge\Big|D\frac{\mathsf{d}^2(x,\cdot)}{2\delta}(y)\Big|-\kappa\beta'\circ\mathsf{d}(y,\bar{x})-\epsilon. \tag{2.10}$$

Since

$$\sup_{0\le r\le t}\mathsf{d}(z(r),y)\le\mathsf{d}(z(t),y)\le\mathsf{d}(x,y)\to 0,\qquad\text{as } t\to 0,$$

if $\ell(q)<\infty$ we can use condition 1.1(3) to get

$$\liminf_{t\to 0^+}t^{-1}\Big(T_t\psi(y)-\psi(y)\Big)\ge q\Big(\Big|D\frac{\mathsf{d}^2(x,\cdot)}{2\delta}(y)\Big|-\kappa\beta'\circ\mathsf{d}(y,\bar{x})-\epsilon\Big)-L(y,q).$$

Optimizing over $q\ge 0$ yields

$$\liminf_{t\to 0^+}t^{-1}\Big(T_t\psi(y)-\psi(y)\Big)\quad\ge\quad\bar{H}\Big(y,\Big|D\frac{\mathsf{d}^2(x,\cdot)}{2\delta}(y)\Big|-\kappa\beta'\circ\mathsf{d}(y,\bar{x})-\epsilon\Big).$$

□

2.4. **Existence of viscosity solutions.** Let f_α be defined according to (1.10). We note that f_α has growth estimates as in Lemma 2.7.

Lemma 2.11. *For all $x\in\mathsf{X}$ it holds*

$$\limsup_{t\to 0^+}t^{-1}\Big(T_tf_\alpha(x)-f_\alpha(x)\Big)\le\alpha^{-1}(f_\alpha(x)-h(x)).$$

Proof. By definition of T_tf_α, for each $t\in(0,1)$ there exists an absolutely continuous path $z=z_t$ in $[0,t]$ with $z(0)=x$ such that

$$T_tf_\alpha(x)\le t^2+f_\alpha(z(t))-\int_0^t L(z(r),|z'|(r))dr.$$

This implies, together with definition of ℓ,

$$\sup_{t\in[0,1]}\int_0^t\ell(|z'|(r))dr\le\sup_{t\in[0,1]}\int_0^t L(z(r),|z'|(r))dr\le\sup_{\mathsf{X}}f_\alpha-\inf_{0\le t\le 1}T_tf_\alpha(x)+1<\infty. \tag{2.11}$$

In addition, combine the above upper estimate of T_tf_α with dynamical programming principle (Lemma 2.5), we have

$$\begin{aligned}
&\frac{T_tf_\alpha(x)-f_\alpha(x)}{t}\\
\le\;& t-\frac{1}{t}\int_0^t\frac{e^{-r/\alpha}}{\alpha}h(z(r))dr+\frac{1}{t}\int_0^t(e^{-r/\alpha}-1)L\Big(z(r),|z'|(r)\Big)dr+\frac{1-e^{-t/\alpha}}{t}f(z(t))\\
\le\;& t-\frac{1}{t}\int_0^t\frac{e^{-r/\alpha}}{\alpha}h(z(r))dr+\frac{1-e^{-t/\alpha}}{t}f_\alpha(z(t))+\inf L\Big(\alpha\frac{1-e^{-t/\alpha}}{t}-1\Big).
\end{aligned}$$

Note that (e.g. [2, Theorem 1.1.2])

$$\mathsf{d}(z(s),x)=\mathsf{d}(z(s),z(0))\le\int_0^s|z'|(r)dr\le\int_0^t|z'|(r)dr,\quad\forall s\in[0,t].$$

In view of (2.11) and super-linearity assumption on ℓ, we have

$$\lim_{t\to 0^+}\sup_{0\le s\le t}\mathsf{d}(z_t(s),x)=0,$$

implying (by the continuity of h and f_α)

$$\limsup_{t\to 0^+} \frac{T_t f_\alpha(x) - f_\alpha(x)}{t} \le -\frac{1}{\alpha} h(x) + \frac{1}{\alpha} f_\alpha(x).$$

$\square$

Lemma 2.12. *If f_α is defined according to* (1.10)*, then*

$$\liminf_{t\to 0^+} t^{-1}\Big(T_t f_\alpha(x) - f_\alpha(x)\Big) \ge \alpha^{-1}(f_\alpha(x) - h(x)).$$

Proof. By Lemma 2.5, for each $t \in (0,1)$, there exists an absolutely continuous path $z = z_t$ with $z(0) = x$ such that

$$f_\alpha(x) \le t^2 + \int_0^t \frac{e^{-r/\alpha}}{\alpha} h(z(r)) - \int_0^t e^{-r/\alpha} L(z(r), |z'|(r)) dr + e^{-t/\alpha} f_\alpha(z(t)). \tag{2.12}$$

Combined with the definition of $T_t f_\alpha$, this gives

$$e^{-t/\alpha} T_t f_\alpha(x) \ge f_\alpha(x) - t^2 + \int_0^t (e^{-r/\alpha} - e^{-t/\alpha}) L(z(r), |z'|(r)) dr - \int_0^t \frac{e^{-r/\alpha}}{\alpha} h(z(r)) dr.$$

Then

$$\begin{aligned} \frac{1}{t}\Big(T_t f_\alpha(x) - f_\alpha(x)\Big) \quad &\ge \quad \frac{e^{t/\alpha} - 1}{t} f_\alpha(x) - t e^{t/\alpha} - (\inf L)(1 + \alpha \frac{1 - e^{t/\alpha}}{t}) \\ &\quad - \frac{1}{t} \int_0^t \frac{e^{(t-r)/\alpha}}{\alpha} h(z(r)) dr. \end{aligned}$$

Similar to the proof in Lemma 2.11, from (2.12), we have that

$$\sup_{t\in[0,1]} \int_0^t \ell(|z'|(r)) dr \le \sup_{t\in[0,1]} \int_0^t L(z(r), |z'|(r)) dr < \infty,$$

which implies

$$\lim_{t\to 0^+} \sup_{0\le s\le t} \mathsf{d}(z_t(s), x) = 0.$$

By continuity of h,

$$\liminf_{t\to 0^+} \frac{1}{t}\Big(T_t f_\alpha(x) - f_\alpha(x)\Big) \ge \alpha^{-1} f_\alpha(x) - \alpha^{-1} h(x).$$

$\square$

Lemma 2.13. *The function f_α in* (1.10) *is a viscosity super-solution to* (1.11)*.*

Proof. Let $\psi \in D_1^{\kappa,\epsilon}$, hence continuous and bounded from above. Let $x \in \mathsf{X}$ be such that $(\psi - f_\alpha)(x) = \sup_{\mathsf{X}}(\psi - f_\alpha)$, then by Lemma 2.4 we get

$$T_t\psi(x) - \psi(x) = T_t(\psi(\cdot) - \psi(x))(x) \le T_t(f_\alpha(\cdot) - f_\alpha(x))(x) = T_t f_\alpha(x) - f_\alpha(x), \quad t \ge 0.$$

Combined with results in Lemmas 2.10 and 2.11, we get $H_1^{\kappa,\epsilon}\psi(x) \le (f(x) - h(x))/\alpha$. $\square$

Lemma 2.14. *The function f_α in* (1.10) *is a viscosity sub-solution to* (1.11)*.*

Proof. Let φ be of form (1.12) which is locally Lipschitz continuous, bounded from below. Let $x \in \mathsf{X}$ be such that $(f_\alpha - \varphi)(x) = \sup_{\mathsf{X}}(f_\alpha - \varphi)$ and choose $M \geq \sup_{\mathsf{X}} f_\alpha - f_\alpha(x) + \varphi(x)$. Then, we claim that $\sup(f_\alpha - \varphi_M) \leq \sup(f_\alpha - \varphi)$. Indeed, by our choice of M we have

$$f_\alpha(y) - M \leq f(x) - \varphi(x) \qquad \forall y \in \mathsf{X}$$

and $f_\alpha(y) - M \geq f_\alpha(y) - \varphi_M(y)$ wherever $\varphi_M(y) \neq \varphi(y)$. Since $\varphi_M(x) = \varphi(x)$, it follows that

$$(f_\alpha - \varphi)(x) = (f_\alpha - \varphi_M)(x) = \sup_{\mathsf{X}}(f_\alpha - \varphi_M).$$

Now, similar to the proof of Lemma 2.13,

$$T_t f_\alpha(\cdot) - f_\alpha(x) \leq T_t \varphi_M(\cdot) - \varphi_M(x).$$

Combined with Lemma 2.12 and Lemma 2.9, we get

$$\alpha^{-1}(f_\alpha - h)(x) \leq \liminf_{t \to 0^+} \frac{1}{t}\Big(T_t \varphi_M(x) - \varphi_M(x)\Big) \leq (H_0 \varphi)^*(x).$$

□

In summary, we proved the

Theorem 2.15 (Existence)**.** *The function f_α defined by* (1.10) *is continuous, bounded from above and has at most $\zeta \circ \mathsf{d}(\cdot, \bar{x})$ growth to $-\infty$. It is the unique viscosity solution to* (1.11) *in the class of functions satisfying these bounds.*

3. The Cauchy problem

3.1. **Further estimates on the time dependent value function.**

Lemma 3.1. *For every $h \in BUC(\mathsf{X})$ and $t \geq 0$, we have*

$$\limsup_{s \to 0^+} \sup_{\overline{B}_R(\bar{x})} \Big(T_{t+s}h(x) - T_t h(x)\Big) \leq 0, \quad \forall R \in \mathbb{R}_+.$$

Proof. From (2.7) and the semigroup identity $T_{s+t} = T_s T_t$ in Lemma 2.4, there exists $\hat{z} := \hat{z}_{s,t,x} \in AC([0, s+t]; \mathsf{X})$ such that $\hat{z}(0) = x$ and

$$\begin{aligned} T_{t+s}h(x) - T_t h(x) &\leq s + h(\hat{z}(t+s)) - h(\hat{z}(t)) - \int_t^{t+s} L(\hat{z}, |\hat{z}'|) dr \\ &\leq s + h(\hat{z}(t+s)) - h(\hat{z}(t)) - s \inf L. \end{aligned}$$

In addition, the first inequality above implies that, for $s \in [0, 1]$,

$$\begin{aligned} \int_t^{t+s} \ell(|\hat{z}'|) dr &\leq \int_t^{t+s} L(\hat{z}, |\hat{z}'|) dr \\ &\leq 1 + h(\hat{z}(t+s)) - h(\hat{z}(t)) - T_{t+s}h(x) + T_t h(x) \\ &\leq 1 + 4 \sup_{\mathsf{X}} |h| - t \inf L + (t+s)\zeta \circ \mathsf{d}(x, \bar{x}), \end{aligned}$$

where we used Lemma 2.8 in obtaining the last line. Using the super-linear growth of ℓ,

$$\lim_{s \to 0^+} \sup_{r \in [0,s],\, x \in \overline{B}_R(\bar{x})} \mathsf{d}(\hat{z}_{s,t,x}(t+r), \hat{z}_{s,t,x}(t)) = 0, \quad \forall R \in \mathbb{R}_+,$$

implying (by the uniform continuity of h) the conclusion. □

Lemma 3.2. *Let $h \in BUC(\mathsf{X})$ and $t \geq 0$, then $T_t h$ is locally uniformly continuous (i.e. uniformly in balls of finite radius).*

Proof. Since $T_0 h = h$, we only need to prove the case $t > 0$. For each $x \in \mathsf{X}$, there exists $z := z_{t,x,\epsilon} \in AC([0,t];\mathsf{X})$, $z(0) = x$ such that

$$T_t h(x) \leq \epsilon + h(z(t)) - \int_0^t L(z, |z'|)dr.$$

Then for any $s \geq 0$ and every $\hat{z} \in AC((0, t+s);\mathsf{X})$ with $\hat{z}(0) = y$, we have

$$\begin{aligned} T_t h(x) - T_t h(y) &= T_t h(x) - T_{s+t} h(y) + T_{s+t} h(y) - T_t h(y) \\ &\leq \epsilon + h(z(t)) - h(\hat{z}(s+t)) - \int_0^t L(z, |z'|)dr \\ &\quad + \int_0^s L(\hat{z}, |\hat{z}'|)dr + \int_s^{s+t} L(\hat{z}, |\hat{z}'|)dr \\ &\quad + T_{s+t} h(y) - T_t h(y). \end{aligned}$$

Next, we choose a special $s := \mathsf{d}(x,y)/q_0$ where $q_0 > 0$ satisfies $\ell(q_0) < \infty$ (whose existence is ensured by Condition 1.1(5)) and a special $\hat{z}$ such that in time interval $(s, s+t)$:

$$\hat{z}(r+s) = z(r), \quad \forall r \in [0,t],$$

and in time interval $(0,s)$: $\hat{z} \in AC([0,s];\mathsf{X})$ is a constant speed geodesic connecting x to y:

$$\hat{z}(0) = y, \quad \hat{z}(s) = x, \quad |\hat{z}'| \equiv \frac{\mathsf{d}(x,y)}{s}.$$

Then

$$\begin{aligned} T_t h(x) - T_t h(y) &\leq \epsilon + \int_0^{\mathsf{d}(x,y)/q_0} L(\hat{z}(r), q_0)dr + T_{s+t} h(y) - T_t h(y) \\ &\leq \epsilon + \zeta(\mathsf{d}(x,y) + \mathsf{d}(x,\bar{x}))\frac{\mathsf{d}(x,y)}{q_0} + \Big(T_{s+t} h(y) - T_t h(y)\Big). \end{aligned}$$

Using the result in Lemma 3.1, we conclude. □

Lemma 3.3. *For every $h \in BUC(\mathsf{X})$ and $t \geq 0$,*

$$\liminf_{s \to 0^+} \inf_{\overline{B}_R(\bar{x})} \Big(T_{s+t} h(x) - T_t h(x)\Big) \geq 0, \quad \forall R \in \mathbb{R}_+.$$

Proof. Let $g = T_t h$, then g is bounded from above with possible growth to $-\infty$ at most at the rate of $\zeta \circ \mathsf{d}(\cdot, \bar{x})$ (Lemma 2.8), moreover, g is locally uniformly continuous (Lemma 3.2). Therefore, using $T_{s+t} = T_s T_t$, we only need to prove

$$\liminf_{s \to 0^+} \inf_{\overline{B}_R(\bar{x})} \Big(T(s)g(x) - g(x)\Big) \geq 0, \quad \forall R \in \mathbb{R}_+. \tag{3.1}$$

Let $q_0 > 0$ be such that $\ell(q_0) < \infty$, whose existence is ensured by Condition 1.1(5). Suppose $z \in \mathsf{X} \setminus \{\bar{x}\}$. Then by triangle inequality

$$\mathsf{d}(x,\bar{x}) \vee \mathsf{d}(x,z) \geq \frac{1}{2}\mathsf{d}(z,\bar{x}) > 0, \quad \forall x \in \mathsf{X}.$$

Take

$$\epsilon_0 := \frac{1}{2q_0}\mathsf{d}(z,\bar{x}) > 0.$$

Then whenever $s < \epsilon_0$, for every $x \in \mathsf{X}$,

$$\mathsf{d}(x,\bar{x}) \vee \mathsf{d}(x,z) \geq q_0\epsilon_0 > q_0 s,$$

implying existence of $y \neq x$ such that $\mathsf{d}(x,y) = q_0 s$. Let $w : [0,s] \mapsto \mathsf{X}$ be a constant-speed geodesic with $w(0) = x$, $w(s) = y$, $|w'| \equiv q_0$ in $(0,s)$. Then for each x fixed with $\mathsf{d}(x,\bar{x}) \leq R$,

$$\begin{aligned} T_s g(x) - g(x) &\geq g(w(s)) - g(w(0)) - \int_0^s L(w(r), |w'|(r))dr \\ &\geq g(y) - g(x) - s\zeta(\mathsf{d}(x,\bar{x}) + \mathsf{d}(x,y)). \end{aligned}$$

Noticing that $s \to 0^+$ implies that $\mathsf{d}(x,y) \to 0^+$, by local uniform continuity of g, the conclusion follows. □

Lemma 3.4. *Let $h \in BUC(\mathsf{X})$ and let U be defined according to* (1.14)*. Then $U \in C([0,T] \times \mathsf{X})$.*

Proof. Since

$$\begin{aligned} |U(s+t,x) - U(t,y)| &= |T_{s+t}h(x) - T_t h(y)| \\ &\leq |T_{s+t}h(x) - T_t h(x)| + |T_t h(x) - T_t h(y)|, \end{aligned}$$

the conclusion follows from Lemma 3.1, Lemma 3.2 and Lemma 3.3. □

3.2. **Comparison principle.**

Lemma 3.5. *Let $U \in USC([0,T] \times \mathsf{X}; \mathbb{R})$ be an upper semicontinuous sub-solution to* (1.15) *with initial data $U(0,x) = U_0(x)$. Moreover, suppose that it has growth from above at most $\zeta \circ \mathsf{d}(\cdot,\bar{x})$. Let $V \in LSC([0,T] \times \mathsf{X}; \mathbb{R})$ be a lower semicontinuous super-solution to* (1.15) *with initial data and $V(0,y) = V_0(y)$. Suppose that V has growth from below at most $\zeta \circ \mathsf{d}(\cdot,\bar{x})$. Then*

$$\sup_{\mathsf{X}} (U - V)(t,\cdot) \leq \sup_{\mathsf{X}} (U_0 - V_0), \quad \forall t \in [0,T].$$

Proof. The idea of the proof is identical to the resolvent equation case in Lemma 2.3. Therefore, we only highlights details which are different.

Let $\lambda > 1$, $c > 0$, $\kappa > 0$ be fixed and define

$$\Psi_\delta(t,x;s,y) = \lambda\Big(U(t,x) - ct\Big) - V(s,y) - \frac{\mathsf{d}^2(x,y)}{2\delta} - \frac{1}{2\delta}|s-t|^2 - \kappa\beta \circ \mathsf{d}(x,\bar{x}) - \kappa\beta \circ \mathsf{d}(y,\bar{x}).$$

With all the above parameters fixed, by the Ekeland's principle, we can find $(t_\delta, x_\delta; s_\delta, y_\delta)$ which is the global strict maximum of

$$(t,x;s,y) \mapsto \Psi_\delta(t,x;s,y) - \delta\mathsf{d}(x,x_\delta) - \delta\mathsf{d}(y,y_\delta) - \delta|t - t_\delta| - \delta|s - s_\delta|.$$

Moreover, similar relations to (2.2), (2.3) and (2.4) hold:

$$\sup_{([0,T]\times\mathsf{X})^2} \Psi_\delta \leq \Psi_\delta(t_\delta, x_\delta, s_\delta, y_\delta) + 2\delta. \tag{3.2}$$

By the growth condition on U and V and the relation between β and ζ as formulated in Condition 1.1.5,

$$R := R_{\lambda,\kappa,c} := \sup_{\delta\in(0,1)} \mathsf{d}(x_\delta,\bar{x}) + \mathsf{d}(y_\delta,\bar{x}) < \infty.$$

Moreover,

$$\lim_{\delta\to 0^+}\Big(\frac{1}{\delta}\mathsf{d}^2(x_\delta,y_\delta)+\frac{1}{\delta}|t_\delta-s_\delta|^2\Big)=0.$$

Without loss of generality, selecting a subsequence if necessary,

$$\lim_{\delta\to 0^+}s_\delta=\lim_{\delta\to 0^+}t_\delta=:r_*.$$

First, suppose that $r_*>0$. Then when δ small enough, we can always assume $t_\delta>0$, and $s_\delta>0$. Set

$$\varphi(t,x):=ct+\lambda^{-1}\Big(\frac{\mathsf{d}^2(x,y_\delta)}{2\delta}+\frac{1}{2\delta}|t-s_\delta|^2+\kappa\beta\circ\mathsf{d}(x,\bar{x})+\delta\mathsf{d}(x,x_\delta)+\delta|t-t_\delta|\Big)\in\hat{D}_0.$$

Then

$$(U-\varphi)(t_\delta,x_\delta)=\sup_{[0,T]\times\mathsf{X}}\Big(U-\varphi\Big).$$

By viscosity sub-solution property, we have

$$c+\frac{t_\delta-s_\delta}{\delta}-\delta\le\lambda(H_0^{\lambda,\kappa,\delta}\varphi)^*(x_\epsilon).$$

Similarly, set

$$\psi(s,y):=-\frac{\mathsf{d}^2(x_\delta,y)}{2\delta}-\frac{1}{2\delta}|s-t_\delta|^2-\kappa\beta\circ\mathsf{d}(y,\bar{x})-\delta\mathsf{d}(y,y_\delta)-\delta|s-s_\delta|\in\hat{D}_1.$$

Then

$$(\psi-V)(s_\delta,y_\delta)=\sup_{[0,T]\times\mathsf{X}}\Big(\psi-V\Big),$$

and by viscosity super-solution property,

$$\frac{t_\delta-s_\delta}{\delta}+\delta\ge(H_1^{\kappa,\delta}\psi)_*(y_\delta).$$

Consequently

$$c-2\delta\le\lambda(H_0\varphi)^*(x_\delta)-(H_1^{\kappa,\delta}\psi)_*(y_\delta).$$

Using the estimate in Lemma 2.2 and proceed as in the proof of Lemma 2.3, we have

$$0<c\le\liminf_{\lambda\to 1^+}\liminf_{\kappa\to 0^+}\liminf_{\delta\to 0^+}\Big(\lambda(H_0\varphi)^*(x_\delta)-(H_1\psi)_*(y_\delta)\Big)\le 0.$$

The above contradiction leads us to conclude that the earlier assumption $r_*>0$ cannot be incorrect, hence $r_*=0$.

Now suppose $r_*=0$. For each $(t,x)\in[0,T]\times\mathsf{X}$, by (3.2) and the definition of Ψ,

$$\begin{aligned}\lambda\Big(U(t,x)-ct\Big)-V(t,x)-2\kappa\beta\circ\mathsf{d}(x,\bar{x})-2\delta&=\Psi(t,x;t,x)-2\delta\\ &\le\Psi(t_\delta,x_\delta;s_\delta,y_\delta)\le\lambda\Big(U(t_\delta,x_\delta)-ct_\delta\Big)-V(s_\delta,y_\delta).\end{aligned}$$

Taking limits first as $\delta\downarrow 0$, then as $\lambda\downarrow 1$ and eventually as $c\downarrow 0$ on both sides, we get

$$U(t,x)-V(t,x)\le\sup_{\mathsf{X}}\Big(U_0-V_0\Big).$$

$\square$

3.3. **Existence.** Let $h \in BUC(\mathsf{X})$. By Lemma 3.4, the function U in (1.14) is continuous, by Lemma 2.8, it is bounded from above with possible growth to $-\infty$ with rate at most $\zeta \circ \mathsf{d}(\cdot, \bar{x})$.

Lemma 3.6. *The function U is a super-solution to* (1.15)*.*

Proof. First of all, Lemma 3.3 (in the case of $t = 0$) verifies the initial condition part of viscosity super-solution. To verify the other part of super-solution property, we let $\psi \in \hat{D}_1^{\kappa,\epsilon}$ and $(s_0, y_0) \in (0, T] \times \mathsf{X}$ be such that

$$U(s,y) - U(s_0,y_0) \geq \psi(s,y) - \psi(s_0,y_0), \quad \forall (s,y) \in [0,T] \times \mathsf{X}.$$

By the semigroup, the order preserving and the translation invariance properties of T (Lemma 2.4), for $0 < r < s_0$,

$$\begin{aligned} 0 = U(s_0,y_0) - U(s_0,y_0) &= T_r\Big(U(s_0 - r, \cdot) - U(s_0,y_0)\Big)(y_0) \\ &\geq T_r\Big(\psi(s_0 - r, \cdot) - \psi(s_0,y_0)\Big)(y_0). \end{aligned}$$

In order to conclude, we only need to prove

$$\liminf_{r \to 0^+} \frac{1}{r} T_r\Big(\psi(s_0 - r, \cdot) - \psi(s_0,y_0)\Big)(y_0) \geq (-\partial_s^+ + H_1^{\kappa,\epsilon})\psi(s_0,y_0).$$

Note that by the form of $\psi = \psi(s,y)$ in (1.17),

$$\psi(s_0 - r, y) - \psi(s_0, y) = \phi(s_0 - r) - \phi(s_0) \geq -r(\partial_s^+\phi)(s_0) + o(r).$$

Combining the above with Lemma 2.10, we conclude. $\square$

Lemma 3.7. *The function U is a sub-solution to* (1.15)*.*

Proof. The proof is almost symmetric with respect to that of Lemma 3.6, except a truncation argument, which we highlight next.

Lemma 3.1 implies that the initial condition of viscosity sub-solution is satisfied. Let $\varphi \in \hat{D}_0$. We assume that there exists $(t_0, x_0) \in (0, T] \times \mathsf{X}$ such that

$$\varphi(s,y) - \varphi(s_0,y_0) \geq U(s,y) - U(s_0,y_0), \quad \forall (s,y) \in [0,T] \times \mathsf{X}. \tag{3.3}$$

The function φ is not bounded from above, hence we cannot simply apply the semigroup T on it. However, for each $M > 0$, we can always find a $\eta = \eta_M \in C^2(\mathbb{R})$ with $0 \leq \eta' \leq 1$ satisfying (2.8) and define a localized version φ_M of φ according to (2.9). By choosing M large enough, we will have the conclusion of Lemma 2.9, in addition, (3.3) holds with φ replaced by φ_M. Then, repeating the same procedure as in the proof of Lemma 3.6 gives the sub-solution property. $\square$

Combine the above results, we have the following

Lemma 3.8. *Let $h \in BUC(\mathsf{X})$. Then the function U in* (1.14) *is a continuous viscosity solution to* (1.15) *with initial data $U(0,x) = h(x)$. U is bounded from above and grows to $-\infty$ at most at the rate of $\zeta \circ \mathsf{d}(\cdot, \bar{x})$.*

4. Application to a compressible Euler equation

Let $\mathsf{X} := \mathcal{P}_2(\mathbb{R}^d)$ be the space of probability measures on $\mathbb{R}^d$ with finite second moment, and let $\mathsf{d}(\rho, \gamma)$ denote the 2-Wasserstein so that, (X, d) is a complete and geodesic metric space (chapter 7 of [2]). Unlike other parts of this article, we use ρ, γ to denote typical elements in X. We choose $\bar{\rho} \in \mathsf{X}$ as an arbitrary but fixed probability measure with smooth and compactly supported Lebesgue density. It plays the role of the base point $\bar{x}$.

Feng and Nguyen studied in [17] an action, defined on space of paths $C([0,T];\mathsf{X})$ (with $T \in \mathbb{R}_+$), whose minimizers are weak solution to the following system of d-dimensional compressible Euler equation (Theorem 1.10 in [17]):

$$
(4.1) \quad \begin{cases} \partial_t \rho + \operatorname{div}(\rho u) &= 0 \\ \partial_t(\rho u) + \operatorname{div}(\rho u \otimes u) + \nabla P(\rho) &= -\rho \nabla(\phi + \Phi * \rho) - 2\nu^2 \rho \nabla\left(\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}} - \frac{1}{4}\psi\right) \\ P(\rho) &= \rho F'(\rho) - F(\rho). \end{cases}
$$

In the above, $\rho = \rho(t,x) : \mathbb{R}_+ \times \mathbb{R}^d \mapsto \mathbb{R}$ and $u = u(t,x) : \mathbb{R}_+ \times \mathbb{R}^d \mapsto \mathbb{R}^d$ are the unknowns, $\nu > 0$ is a given constant and the functions ψ, ϕ, $\Phi \in C^1(\mathbb{R})$, $F \in C^2(\mathbb{R}_+)$ are prescribed. Precise requirements on ψ, ϕ, Φ and F can be found in Condition 1.5 of [17]. The term $\operatorname{div}(\rho u \otimes u)$ is understood as a vector whose i-th component is $\operatorname{div}(\rho u u_i)$. Associated with the above equation are Hamilton-Jacobi partial differential equations in X. Well-posedness for both resolvent formulation as well as Cauchy problem are proved in Theorems 1.13 and 1.14 of [17]. Next, we apply results of this article to study the action, its minimizer and the associated Hamilton-Jacobi equations in the limiting case $\nu = 0$. To simplify and streamline the main ideas, we only consider the simpler case of $F = 0$. We will work under sufficient regularity assumptions on ϕ, Φ. Note that if Φ is allowed to be singular, then (4.1) is related to the Euler-Poisson problem considered in [20].

This section is organized as follow. We first introduce a Riemannian structure to the Wasserstein space X by following mostly the formalism of Otto [30], and borrowing some technical results from Ambrosio, Gigli and Savaré [2] to relate metric and differential point of view. A precise connection is given in Lemma 4.1. Then we show that every action minimizer is a weak solution to a compressible pressure-less Euler equation (4.7) in distributional solution sense. Finally, we consider well-posednes for the associated Hamilton-Jacobi partial differential equations. From a metric space point of view, since (X, d) is a geodesic space, it is no surprising that our earlier results apply. What we want to show is that there is another geometric based formulation of the equation. Considerable efforts were given to versions of such formulation in earlier literature (e.g. [19, 22]), with absence of a uniqueness result. Making use of the metric level result, we demonstrate that the choice of tangent (and co-tangent) space structure in these earlier literature is inadequate for adapting the metric proof of comparison principle. We will explore the geometric tangent cone concept defined in chapter 12 of [2] to redefine another Hamiltonian. We show that this new one is compatible with our earlier metric formulation. Well-posedness for the PDE, in the metric as well as as in the geometric formulations, then follow.

4.1. Lagrangian and existence of action minimizer.

There is more than one way of introducing tangent/co-tangent spaces of $X := \mathcal{P}_2(\mathbb{R}^d)$. Chapter 8 of [2] examines a set of equivalent ones. We will use one of them (the set T_ρ in (4.4)) next to study the problem of action minimization corresponding to the case $\nu = 0$.

Let $\|p\|_{1,\rho}^2 := \int_{\mathbb{R}^d} |\nabla p|^2 d\rho$ for $p \in C_c^\infty(\mathbb{R}^d)$ and

$$\begin{aligned} H_{1,\rho}(\mathbb{R}^d) &:= \text{ abstract completion of } C_c^\infty(\mathbb{R}^d), \text{ as a pre-Hilbert space, under } \|\cdot\|_{1,\rho}, \\ H_{-1,\rho}(\mathbb{R}^d) &:= \{m \in \mathcal{D}'(\mathbb{R}^d) : \|m\|_{-1,\rho} < \infty\}, \end{aligned}$$

where

$$\frac{1}{2}\|m\|_{-1,\rho}^2 := \sup_{\varphi \in C_c^\infty(\mathbb{R}^d)} \{\langle \varphi, m\rangle - \frac{1}{2}\int_{\mathbb{R}^d} |\nabla \varphi|^2 d\rho\}. \tag{4.2}$$

We also denote

$$L^2_{\nabla,\rho}(\mathbb{R}^d) := L^2_\rho \text{ closure of } \{\nabla p : p \in C_c^\infty(\mathbb{R}^d)\} \ ,$$

Note that

$$|\langle p, q\rangle_{1,\rho}| \le \|q\|_{1,\rho}\|p\|_{1,\rho} = \|q\|_{1,\rho}\|\nabla p\|_{L^2_\rho(\mathbb{R}^d)}, \quad \forall p \in C_c^\infty(\mathbb{R}^d),$$

hence the linear operator ∇ can be extended from $q \in C_c^\infty(\mathbb{R}^d)$ to all $q \in H_{1,\rho}(\mathbb{R}^d)$. We denote such extension $\hat{\nabla}$. By Lemma D.34 in Appendix D of [16], for each $m_i \in H_{-1,\rho}(\mathbb{R}^d)$, we can identify a unique $p_i \in H_{1,\rho}(\mathbb{R}^d)$ such that $m_i = -\text{div}(\rho\hat{\nabla}p_i) \in H_{-1,\rho}(\mathbb{R}^d)$ (equivalently $\hat{\nabla}p_i \in L^2_{\nabla,\rho}(\mathbb{R}^d)$) and

$$\langle m_1, m_2\rangle_{-1,\rho} = \langle p_1, m_2\rangle_{H_{1,\rho}(\mathbb{R}^d)\times H_{-1,\rho}(\mathbb{R}^d)} = \langle p_2, m_1\rangle_{H_{1,\rho}(\mathbb{R}^d)\times H_{-1,\rho}(\mathbb{R}^d)} = \langle \hat{\nabla}p_1, \hat{\nabla}p_2\rangle_{L^2_\rho \times L^2_\rho}.$$

We refer to chapter 8 of [2] or appendix D.5 of [16] for further properties and relations of these spaces and we just quote here the elementary inequality

$$\|\text{div}(v\rho)\|_{-1,\rho}^2 \le \int_{\mathbb{R}^d} |v|^2 d\rho \qquad \forall v \in L^2_\rho(\mathbb{R}^d). \tag{4.3}$$

Viewing $\rho(t,dx)dt$ as a measure on $(0,\infty)\times\mathbb{R}^d$, its distributional time derivative $\partial_t\rho(t,dx)$ exists. For each $t \in \mathbb{R}_+$, we define $\dot{\rho}(t)$ as the unique element in $\mathcal{D}'(\mathbb{R}^d)$ satisfying

$$\langle \varphi, \dot{\rho}\rangle := \frac{d}{dt}\langle \rho(t), \varphi\rangle, \quad \forall \varphi \in C_c^\infty(\mathbb{R}^d),$$

whenever the right hand side in the above exists. For $\rho(\cdot) \in AC([0,T], \mathsf{X})$, by Theorem 8.3.1 of [2], $\partial_t\rho = -\text{div}_x(\rho v)$ in the sense of distributions for some $v := v(t,x)$ such that $\int_{[0,T]\times\mathbb{R}^d} |v(t,x)|^2\rho(t,dx)dt < \infty$. In particular, there exists a Lebesgue measure set zero $\mathcal{N} \subset \mathbb{R}_+$ such that

$$\frac{d}{dt}\langle \rho(t), \varphi\rangle = \langle \varphi, -\text{div}_x(\rho v)\rangle, \quad t \in [0,T]\setminus\mathcal{N}, \quad \forall \varphi \in C_c^\infty(\mathbb{R}^d).$$

Hence we conclude the existence of $\dot{\rho}(t) \in H_{-1,\rho(t)}(\mathbb{R}^d)$ almost everywhere in $t \in \mathbb{R}_+$. In fact, if $|\rho'|$ denotes the metric derivative (with respect to Wasserstein distance) as before, Theorem 8.3.1 of [2] provides the following more precise result connecting the Riemannian and metric points of view.

Lemma 4.1. *For $\rho(\cdot) \in AC([0,T];\mathsf{X})$, we have $\|\dot{\rho}(r)\|_{-1,\rho(r)} = |\rho'|(r)$ for a.e. $r \in (0,T)$. Furthermore, there exist $p(r) \in H_{1,\rho(r)}(\mathbb{R}^d)$ with $v(r) := \hat{\nabla}p(r) \in L^2_{\nabla,\rho(r)}(\mathbb{R}^d)$ satisfying*

$$-\text{div}(v(r)\rho(r)) = \dot{\rho}(r), \quad \|v(r)\|_{L^2_{\rho(r)}(\mathbb{R}^d)} = |\rho'|(r) \textit{ for a.e. } r \in (0,T).$$

Let

$$\mathsf{T}_\rho := \Big\{(\rho, p) : p \in H_{1,\rho}(\mathbb{R}^d)\Big\}, \qquad \mathsf{T} := \sqcup_{\rho \in \mathsf{X}} \mathsf{T}_\rho. \tag{4.4}$$

We define the Lagrangian

$$L(\rho, p) := \frac{1}{2}\|p\|_{1,\rho}^2 - V(\rho), \qquad (\rho, p) \in \mathsf{T}, \tag{4.5}$$

where

$$V(\rho) := \int_{\mathbb{R}^d} \phi(x)\rho(dx) + \frac{1}{2}\int\int_{\mathbb{R}^d\times\mathbb{R}^d} \Phi(x-y)\rho(dx)\rho(dy).$$

We assume the following throughout this section.

Condition 4.2.

(1) $\Phi \in C^1(\mathbb{R}^d)$ *is bounded, with globally bounded gradient;*

(2) $-\phi \in C(\mathbb{R}^d)$ *has compact finite sub-levels with sub-linear growth in the sense that there exist* $\theta \in (0,1)$ *and* $C \in \mathbb{R}_+$ *such that*

$$-\phi(x) \le C(1 + |x|^\theta) \quad \forall x \in \mathbb{R}^d;$$

(3) $\phi \in C^1(\mathbb{R}^d)$*, with globally bounded gradient;*

The above Lagrangian induces an action function defined on every path $\rho(\cdot) \in AC([0,T];\mathsf{X})$: let $p(r)$ be chosen according to Lemma 4.1,

$$\begin{aligned} A[\rho(\cdot)] &:= \int_0^T L(\rho(r), p(r))dr \\ &= \int_0^T \Big(\frac{1}{2}|\rho'|^2 + V(\rho)\Big)dr = \int_0^T \Big(\frac{1}{2}\|\dot\rho\|_{-1,\rho}^2 + V(\rho)\Big)dr. \end{aligned} \tag{4.6}$$

Lemma 4.3 (Existence of an action minimizer)**.** *For every* $\rho_0, \gamma_0 \in \mathsf{X}$*, there exists a path* $\rho(\cdot) \in AC([0,T];\mathsf{X})$ *with* $\rho(0) = \rho_0$ *and* $\rho(T) = \gamma_0$ *such that*

$$A[\rho(\cdot)] = \inf\left\{A[\sigma(\cdot)] : \sigma(0) = \rho_0, \sigma(T) = \gamma_0, \sigma(\cdot) \in AC([0,T];\mathsf{X})\right\}.$$

Proof. Let $\sigma_\epsilon(\cdot) \in AC([0,T];\mathsf{X})$ be ϵ-optimizers of the action satisfying $\sigma_\epsilon(0) = \rho_0$ and $\sigma_\epsilon(T) = \gamma_0$. By Lemma 4.1, there exists $u_\epsilon(t) := u_\epsilon(t,\cdot) \in L^2_{\nabla,\sigma_\epsilon(t)}$ such that $\|u_\epsilon(t)\|_{L^2(\sigma_\epsilon(t))} = |\sigma_\epsilon'(t)|$ for a.e. t. Let

$$\mathbf{m}_\epsilon(t; dx, d\xi) := \delta_{u_\epsilon(t,x)}(d\xi)\sigma_\epsilon(t, dx).$$

Then

$$\partial_t \mathbf{m}_\epsilon + \mathrm{div}_x(\xi \mathbf{m}_\epsilon) = 0,$$

and

$$A[\sigma_\epsilon(\cdot)] = \int_{r\in[0,T]}\int_{(x,\xi)\in\mathbb{R}^d\times\mathbb{R}^d}\int_{y\in\mathbb{R}^d} \Big(\frac{1}{2}|\xi|^2 - (\phi(x) + \Phi(x-y))\Big)\sigma_\epsilon(r; dy)\mathbf{m}_\epsilon(r; dx, d\xi)dr.$$

From $\sup_\epsilon A[\sigma_\epsilon(\cdot)] < \infty$ and the assumption on compact finite sub-levels of $-\phi$, we obtain that $\{\mathbf{m}_\epsilon(dr, dx, d\xi) := \mathbf{m}_\epsilon(r, dx, d\xi) \times dr : \epsilon > 0\}$ is tight in the weak convergence of probability measure topology. Let $\mathbf{m}_0(dr, dx, d\xi)$ be a limit point. Since the time marginal is always the Lebesgue measure along the sequence, let $\sigma(r, dx) := \mathbf{m}_0(r, dx, \mathbb{R}^d)$, we have the disintegration

$$\mathbf{m}_0(dr, dx, d\xi) = \mathbf{m}_0(d\xi; r, x)\sigma(r, dx)dr$$

allowing us to define a measurable function

$$u(r, x) := \int_{\mathbb{R}^d} \xi \mathbf{m}_0(d\xi; r, x).$$

Then

$$\partial_t \sigma + \mathrm{div}_x(\sigma u) = 0.$$

Using (4.3), Jensen's inequality and Fatou's lemma, we get

$$\begin{aligned} A[\sigma(\cdot)] &\leq \int_0^T \Big(\int_{\mathbb{R}^d} \frac{1}{2}|u(r,x)|^2 \sigma(r, dx) - V(\sigma(r)) \Big) dr \\ &\leq \int_{[0,T]} \int_{\mathbb{R}^d \times \mathbb{R}^d} \Big(\frac{1}{2}|\xi|^2 - (\phi(x) + \Phi * \sigma(x)) \Big) \mathbf{m}_0(r; dx, d\xi) dr \\ &\leq \liminf_{\epsilon \to 0+} \int_{[0,T]} \int_{\mathbb{R}^d \times \mathbb{R}^d} \Big(\frac{1}{2}|\xi|^2 - (\phi(x) + \Phi * \sigma_\epsilon(x)) \Big) \mathbf{m}_\epsilon(r; dx, d\xi) dr \\ &= \liminf_{\epsilon \to 0+} A[\sigma_\epsilon(\cdot)] = \inf\{A[\sigma(\cdot)] : \sigma(0) = \rho_0, \, \sigma(T) = \gamma_0, \, \sigma(\cdot) \in AC([0,T]; \mathsf{X})\}. \end{aligned}$$

□

The following result relates the variational problem we consider to a compressible Euler equation. Note that this is different than the Cauchy problem of Euler equation where constructing a mono kinetic solution is much harder (see for instance Gangbo and Westdickenberg [23]).

Theorem 4.4. *A minimizer of the action $A[\cdot]$, with initial value $\rho(0)$ and terminal value $\rho(T)$ exists. Moreover, any such minimizer $\rho(\cdot)$ is a weak (i.e. distributional) solution in $(0,T) \times \mathbb{R}^d$ to the following pressure-less compressible Euler equation*

$$\left\{ \begin{array}{rcl} \partial_t \rho + \mathrm{div}(\rho u) &=& 0, \\ \partial_t(\rho u) + \mathrm{div}(\rho u \otimes u) + \rho \nabla(\phi + \Phi * \rho) &=& 0. \end{array} \right. \tag{4.7}$$

Proof. Combined with the existence of minimizer in Lemma 4.3, the arguments in Section 3.2 in [17] can be adapted to give the proof, with some simplifications. Indeed, in the present case $\nu = 0$, $F = 0$ and $\nabla\Phi$, $\nabla\phi$ are bounded continuous, so that the *a priori* estimates in Lemmas 3.5 and Lemma 3.6 in [17] are not needed anymore. □

4.2. **An inadequate choice of Hamiltonian in the case $\nu = 0$.** In view of the tangent space structure, we choose cotangent space

$$\mathsf{T}^*_\rho := \Big\{ (\rho, n) : n \in H_{-1,\rho}(\mathbb{R}^d) \Big\}, \quad \mathsf{T}^* := \sqcup_{\rho \in \mathsf{X}} \mathsf{T}^*_\rho.$$

C_c^∞ is dense in the tangent space T_ρ and the cotangent space $\mathsf{T}^*_\rho \subset \mathcal{D}'(\mathbb{R}^d)$. For every $(\rho, n) \in \mathsf{T}^*_\rho$ and $p \in C_c^\infty(\mathbb{R}^d)$,

$$\langle p, n \rangle_{H_{1,\rho}(\mathbb{R}^d) \times H_{-1,\rho}(\mathbb{R}^d)} = \langle p, n \rangle_{\mathcal{D}(\mathbb{R}^d) \times \mathcal{D}'(\mathbb{R}^d)}.$$

These motivate the following definition of gradient.

Definition 4.5. Let $f : \mathsf{X} \mapsto \bar{\mathbb{R}}$ and $\rho_0 \in \mathsf{X}$. The gradient $n := \mathrm{grad} f(\rho_0) \in \mathcal{D}'(\mathbb{R}^d)$ is defined as the unique distribution n such that

$$\langle p, n \rangle_{\mathcal{D}(\mathbb{R}^d) \times \mathcal{D}'(\mathbb{R}^d)} = \lim_{t \to 0^+} \frac{1}{t} (f(\sigma(t)) - f(\sigma(0)))$$

holds for every $p \in C_c^\infty(\mathbb{R}^d)$ and $\sigma(\cdot) := \sigma^p(\cdot) \in AC(\mathbb{R}_+; \mathsf{X})$ satisfying

$$\partial_t \sigma + \mathrm{div}(\sigma \nabla p) = 0, \quad \sigma(0) = \rho_0. \tag{4.8}$$

We define

$$H(\rho, n) := \frac{1}{2} \|n\|_{-1,\rho}^2 + V(\rho), \quad (\rho, n) \in \mathsf{X} \times \mathcal{D}'(\mathbb{R}^d).$$

Note that (4.2) variationally defines $\|n\|_{-1,\rho}$ for all $n \in \mathcal{D}'(\mathbb{R}^d)$. We now define the operator

$$Hf(\rho) := H(\rho, \mathrm{grad} f(\rho)) = \frac{1}{2} \|\mathrm{grad} f(\rho)\|_{-1,\rho}^2 + V(\rho). \tag{4.9}$$

This is formally the $\nu = 0$ limit of the case considered in [17]. We claim that this is not the correct analogue of the metric version of Hamiltonians studied earlier. In particular, it is an open problem to establish comparison principle for viscosity solution of

$$f - \alpha H f = h.$$

A few computations will clarify.

We consider the function $\rho \mapsto \frac{1}{2} \mathsf{d}^2(\rho, \gamma)$. First, we introduce some notation from mass transportation theory. Let π^i be a projection from $(x_1, x_2, \ldots) \in \mathbb{R}^d \times \mathbb{R}^d \times \ldots$ onto the i-th coordinate $\pi^i(x_1, x_2, \ldots) := x_i$. Similarly, we define $\pi^{i,j}(x_1, x_2, \ldots) := (x_i, x_j)$ for $i \neq j$. Let ρ, $\gamma \in \mathsf{X}$. We denote

$$\Gamma(\rho, \gamma) \; := \; \Big\{ \boldsymbol{\mu} \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d) : \pi^1_\# \boldsymbol{\mu} = \rho, \pi^2_\# \boldsymbol{\mu} = \gamma \Big\}$$

and

$$\Gamma_o(\rho, \gamma) \; := \; \Big\{ \boldsymbol{\mu} \in \Gamma(\rho, \gamma) : \mathsf{d}^2(\rho, \gamma) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \boldsymbol{\mu}(dx, dy) \Big\}.$$

That is, Γ_0 is the collection of probability measures solving the Kantorovich problem (Chapter 6 of [2]).

Let $\rho_0 \in \mathsf{X}$ and $p \in C_c^\infty(\mathbb{R}^d)$. We define $\sigma := \sigma(t)$ as in (4.8) for $t \in \mathbb{R}$. Let $\boldsymbol{\mu}(t) \in \Gamma_o(\sigma(t), \gamma)$ and $\boldsymbol{\mu}(t; dx, dy) := \boldsymbol{\mu}(t; dy|x) \sigma(t; dx)$ be a disintegration of $\boldsymbol{\mu}$ with a Borel selection of $\boldsymbol{\mu}(t; dy|x)$, so that the function u below is Borel:

$$u(t, x) := \int_{\mathbb{R}^d} (x - y) \boldsymbol{\mu}(t; dy|x).$$

Lemma 4.6. *Let the p, σ, u be as above. Then $\frac{d}{dt} \mathsf{d}^2(\sigma(t), \gamma)$ exists for $t \in \mathbb{R} \setminus \mathcal{N}$, where $\mathcal{N}$ is some Lebesgue measure zero set, and*

$$\frac{d}{dt} \frac{1}{2} \mathsf{d}^2(\sigma(t), \gamma) = \int_{\mathbb{R}^d} u(t, x) \nabla p(x) \sigma(t, dx) = \langle p, -\mathrm{div}(\sigma u) \rangle, \quad t \in \mathbb{R} \setminus \mathcal{N}. \tag{4.10}$$

Moreover, if $\Gamma_o(\rho_0, \gamma) = \{\boldsymbol{\mu}_0\}$ consists of only one element, then $u_0 := u(0, x)$ gives

$$\mathrm{grad}_{\rho_0} \frac{1}{2} \mathsf{d}^2 = -\mathrm{div}(\rho_0 u_0) \in H_{-1,\rho_0}(\mathbb{R}^d). \tag{4.11}$$

Proof. The identity (4.10) follows directly from Theorem 8.4.7 in [2]. Since $\lim_{t\to 0} \mathsf{d}(\sigma(t), \rho_0) = 0$, we have d-relative compactness of $\{\boldsymbol{\mu}_{t_n}\}$ for any sequence $t_n \to 0$, whose limit point has to belong to $\Gamma_o(\rho_0, \gamma)$. Suppose that $\Gamma_o(\rho_0, \gamma)$ consists of a singleton, it follows that $\lim_{t\to 0^+} \mathsf{d}(\boldsymbol{\mu}_t, \boldsymbol{\mu}_0) = 0$. Combined with (4.10), we arrive that $\mathsf{d}^2(\sigma(t), \gamma)$ is right differentiable at $t = 0$ and (4.10) holds at $t = 0$. $\square$

We now consider the optimal transportation problem $\boldsymbol{\mu}_t \in \Gamma_o(\sigma(t), \gamma)$ time by time. In view of the definition of $u(t)$ and $\boldsymbol{\mu}_t$, by Jensen's inequality, we have

$$
\begin{aligned}
\|\mathrm{div}(\sigma(t)u(t))\|^2_{-1,\sigma(t)} &\le \int_{\mathbb{R}^d} |u(t,x)|^2 \sigma(t,dx) \\
&\le \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \boldsymbol{\mu}(t;dx,dy) = \mathsf{d}^2(\sigma(t),\gamma).
\end{aligned}
\tag{4.12}
$$

for every $t \in \mathbb{R} \setminus \mathcal{N}$. We note that, unless $\boldsymbol{\mu}(t)$ is given by an optimal transportation map $T_t := T_t(x)$ (i.e. $\boldsymbol{\mu}(t; dy|x) = \delta_{T_t(x)}(dy)$ for x a.e.-$\sigma(t)$), the last inequality in (4.12) is a *strict* one. Recall that in proving uniqueness of viscosity solution (metric formulation) through the comparison principle, we made critical use of the identities in Lemma 2.1. The appearance of a strict inequality in (4.12) suggests that the above notion of gradient of functions in the space of probability measures is not compatible with metric definition, at least when $\Gamma_o(\rho_0, \gamma)$ is a singleton and ρ_0 is not absolutely continuous. Moreover, the previous strategy of proving comparison principle cannot be replicated, unless ρ is absolutely continuous w.r.t. Lebesgue measure.

In [19] (Section 3) and [22] (Sections 4 and 6), although a slightly different notion of viscosity solution is used, the choice of tangent space is still T, hence the notion of sub-super-gradient which is defined on T^* bring the same problem as mentioned above.

Note that for the viscosity solutions considered in [17], gradients of test functions are only evaluated at ρ's with Lebesgue density. The existence and the uniqueness of the optimal map T are then guaranteed by Brenier's theorem (e.g. Theorem 6.2.4 in [2] or Theorem D.25 in Appendix D of [16]). This can be done because that, when $\nu > 0$, for any path $\rho := \rho(t)$ with finite action as defined in [17], *a priori* estimates give the following trajectory regularity: $\rho(t; dx) = \rho(t,x)dx$ has Lebesgue density for all $t > 0$. The proof of this property uses entropy function and related interpolation inequalities. This trajectorial level regularity is then translated to the Hamiltonian formulation, through the definition of viscosity solutions in [17], by an appropriate choice of test functions (different than the ones here) with entropy as part of the perturbation. In the case of $\nu = 0$, however, such feature is lost (e.g. [20]).

Next, we refine the above approach by augmenting the tangent space. We will also use a different notion of viscosity solution (Definition 4.14) based on sub- and super-differentials and directly defined on candidate solutions at every point, instead of defined indirectly using test functions at certain maximum/minimum points (Definition 1.4). This allows us to link our results next directly with those in literature [19, 22].

4.3. **Geometric tangent cone on X and sub-, super-differentials of functions.** A close inspection on the short proof of Lemma 2.1 reveals the following. The tangent space in previous paragraph does not contain sufficiently many tangent directions to distinguish among certain paths which are the geodesics used in the metric slope calculations. When we define differentiation of the $\rho \mapsto \mathsf{d}^2(\rho, \gamma)$ along these paths, and ρ is singular, the difference shows up. Next, we will use the geometric tangent cones concept, introduced in sections 12.3

and 12.4 of [2], to remedy this problem. Then we reformulate a new Hamiltonian. This allows us to establish a link with the earlier metric slope formulation of the Hamilton-Jacobi PDE, consequently deriving well-posedness of the new formulation of the equation as a by-product.

Let $\rho \in \mathsf{X}$ and

$$G(\rho) := \Big\{ \mathbf{m} := \mathbf{m}(dx; d\xi) \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d) : \pi^1_\# \mathbf{m} = \rho, (\pi^1, \pi^1 + \epsilon \pi^2)_\# \mathbf{m} \in \Gamma_o(\rho, \gamma), \\ \text{for some } \gamma \in \mathsf{X}, \epsilon > 0 \Big\}.$$

For each $\mathbf{m}_i \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ with $\pi^1_\# \mathbf{m}_i = \rho$, $i = 1, 2$, we define a distance

$$\mathsf{D}_\rho(\mathbf{m}_1, \mathbf{m}_2) := \inf \Big\{ \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} |\xi - \eta|^2 \mathbf{M}(dx; d\xi, d\eta) : \mathbf{M} \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d) \\ \pi^{1,2}_\# \mathbf{M} = \mathbf{m}_1, \pi^{1,3}_\# \mathbf{M} := \mathbf{m}_2 \Big\},$$

and

$$\langle \mathbf{m}_1, \mathbf{m}_2 \rangle_\rho := \max \Big\{ \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} (\xi \cdot \eta) \mathbf{M}(dx; d\xi, d\eta) : \mathbf{M} \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d), \\ \pi^{1,2}_\# \mathbf{M} = \mathbf{m}_1, \pi^{1,3}_\# \mathbf{M} := \mathbf{m}_2 \Big\}.$$

When $\mathbf{m}_1 = \mathbf{m}_2 = \mathbf{m} = \mathbf{m}(dx; d\xi)$, the above maximum is trivially attained at

$$\mathbf{M}(dx; d\xi, d\eta) := \delta_\xi(d\eta) \mathbf{m}(dx, d\xi).$$

Hence

$$\|\mathbf{m}\|^2_\rho := \langle \mathbf{m}, \mathbf{m} \rangle_\rho = \int_{\mathbb{R}^d} |\xi|^2 \mathbf{m}(dx; d\xi).$$

We now define

$$\mathsf{Tan}_\rho := \overline{G(\rho)}^{D_\rho(\cdot,\cdot)}, \quad \mathsf{Tan} := \sqcup_{\rho \in \mathsf{X}} \mathsf{Tan}_\rho. \tag{4.13}$$

Recall that T_ρ is identified using $H_{1,\rho}$. By Theorem 12.4.4 of [2], $\mathsf{T}_\rho \subset \mathsf{Tan}_\rho$, via the embedding $q \mapsto (Id \times \hat{\nabla} q)_\# \rho$. The embedding of T_ρ to Tan_ρ is isometric and one-to-one when ρ has Lebesgue density. However, in general, the inclusion can be strict.

We note that, as definitions, the notations $\|\mathbf{m}\|_\rho$ and $\langle \mathbf{m}_1, \mathbf{m}_2 \rangle_\rho$ remain valid for general $\mathbf{m}, \mathbf{m}_1, \mathbf{m}_2 \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ with common first marginal measure ρ, they do not have to be in $G(\rho)$. We will use the notations in such more general context later.

Definition 4.7. [Fréchet super- and sub-differentials] Let $f : \mathsf{X} \mapsto \bar{\mathbb{R}}$ and that $\rho \in \mathsf{X}$ be a point such that $|f(\rho)| < \infty$. We denote super-, sub-differentials and differential of f at ρ respectively by $\partial^+_\rho f := \partial^+ f(\rho)$, $\partial^-_\rho f := \partial^- f(\rho)$, and $\partial_\rho f := \partial f(\rho)$. These are subsets of $\{\mathbf{n} \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d) : \pi^1_\# \mathbf{n} = \rho\}$ satisfying the following.

We say that $\mathbf{n} \in \partial^+_\rho f$ if there exists a modulus of continuity $\omega_\mathbf{n}$ such that, for every $\rho_1 \in \mathsf{X}$ and every $\mathbf{M} \in \mathcal{P}_2((\mathbb{R}^d)^3)$ such that $(\pi^1, \pi^1 + \pi^2)_\# \mathbf{M} \in \Gamma_o(\rho, \rho_1)$ and $\pi^{1,3}_\# \mathbf{M} = \mathbf{n}$, we have

$$f(\rho_1) - f(\rho) \leq \int_{(\mathbb{R}^d)^3} (\xi \cdot \eta) \mathbf{M}(dx; d\xi, d\eta) + \mathsf{d}(\rho, \rho_1) \omega_\mathbf{n}(\mathsf{d}(\rho, \rho_1)). \tag{4.14}$$

Analogously, we say that $\mathbf{n} \in \partial_\rho^- f$ if there exists a modulus of continuity $\omega_{\mathbf{n}}$ such that, for every $\rho_1 \in \mathsf{X}$ and every $\mathbf{M} \in \mathcal{P}_2((\mathbb{R}^d)^3)$ with $(\pi^1, \pi^1 + \pi^2)_\# \mathbf{M} \in \Gamma_o(\rho, \rho_1)$ and $\pi^{1,3}_\# \mathbf{M} \in \Gamma_o(\rho, \rho_1)$, we have

$$f(\rho_1) - f(\rho) \geq \int_{(\mathbb{R}^d)^3} (\xi \cdot \eta) \mathbf{M}(dx; d\xi, d\eta) + \mathsf{d}(\rho, \rho_1) \omega_{\mathbf{n}}(\mathsf{d}(\rho, \rho_1)). \tag{4.15}$$

Finally, we define

$$\partial_\rho f := \partial_\rho^+ f \cap \partial_\rho^- f.$$

The definitions above are closely related to the ones in Definition 10.3.1 in [2] and in Chapter 5 of Gigli [25]. Let $\boldsymbol{\mu} \in \Gamma_o(\rho, \gamma)$,

$$\mathbf{n}_1 := (\pi^1, (\pi^1 - \pi^2))_\# \boldsymbol{\mu}, \qquad \mathbf{n}_2 := (\pi^2, (\pi^1 - \pi^2))_\# \boldsymbol{\mu}.$$

Then by Lemma 2.1,

$$\|\mathbf{n}_1\|_\rho = \mathsf{d}(\rho, \gamma) = \|\mathbf{n}_2\|_\gamma = |D_\rho \frac{1}{2} \mathsf{d}^2(\rho, \gamma)| = |D_\gamma(-\frac{1}{2} \mathsf{d}^2(\rho, \gamma))|, \tag{4.16}$$

where $|D_\rho \frac{1}{2}\mathsf{d}^2|$ and $|D_\gamma \frac{1}{2}\mathsf{d}^2|$ mean metric slopes, respectively w.r.t. ρ and γ. Next, we show that $\mathbf{n}_1 \in \partial_\rho^+ \frac{1}{2}\mathsf{d}^2(\rho, \gamma)$ and $\mathbf{n}_2 \in \partial_\gamma^-(-\frac{1}{2}\mathsf{d}^2(\rho, \gamma))$. Such result is the key to establishing a comparison principle for our geometric-based formulation of the Hamilton-Jacobi equation.

Lemma 4.8. *The following inclusions hold:*

$$\begin{aligned} \partial_\rho^+ \frac{1}{2} \mathsf{d}^2(\cdot, \gamma) \quad &\supset \quad \Big\{ (\pi^1, (\pi^1 - \pi^2))_\# \boldsymbol{\mu} : \boldsymbol{\mu} \in \Gamma_o(\rho, \gamma) \Big\}, \\ \partial_\gamma^- \Big(- \frac{1}{2} \mathsf{d}^2(\rho, \cdot) \Big) \quad &\supset \quad \Big\{ (\pi^2, (\pi^1 - \pi^2))_\# \boldsymbol{\mu} : \boldsymbol{\mu} \in \Gamma_o(\rho, \gamma) \Big\}. \end{aligned}$$

Proof. The conclusions all follow from Theorem 10.2.2 of [2]. □

Definition 4.9. For any $t \in \mathbb{R}$ and $\mathbf{n} \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$, we define scalar multiplication

$$t \cdot \mathbf{n} := (\pi^1, t\pi^2)_\# \mathbf{n}.$$

For $\mathbf{n}_i \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ we define a multi-valued addition by

$$\begin{aligned} \mathbf{n}_1 \oplus \mathbf{n}_2 \quad := \quad \{ &\mathbf{n} \in \mathcal{P}_2((\mathbb{R}^d)^2) : \mathbf{n} = (\pi^1, \pi^2 + \pi^3)_\# \mathbf{N}, \\ &\mathbf{N} \in \mathcal{P}_2((\mathbb{R}^d)^3), \pi^{1,2}_\# \mathbf{N} = \mathbf{n}_1, \pi^{1,3}_\# \mathbf{N} = \mathbf{n}_2 \}. \end{aligned} \tag{4.17}$$

We also define, for φ_i, $i = 1, 2$,

$$\partial_\rho^+ \varphi_1 \oplus \partial_\rho^+ \varphi_2 := \cup_{\mathbf{n}_i \in \partial_\rho^+ \varphi_i} \mathbf{n}_1 \oplus \mathbf{n}_2$$

and we define $\partial_\rho^- \varphi_1 \oplus \partial_\rho^- \varphi_2$ similarly.

It directly follows then

(1) $\|t \cdot \mathbf{n}\|_\rho = |t| \|\mathbf{n}\|_\rho$,
(2) if $\mathbf{n} \in \partial_\rho^\pm f$ and $t > 0$, then $t \cdot \mathbf{n} \in \partial_\rho^\pm (tf)$,
(3) if $\mathbf{n} \in \mathbf{n}_1 \oplus \mathbf{n}_2$, then

$$\|\mathbf{n}\|_\rho \leq \|\mathbf{n}_1\|_\rho + \|\mathbf{n}_2\|_\rho, \quad \|\mathbf{n}_1\|_\rho \leq \|\mathbf{n}\|_\rho + \|\mathbf{n}_2\|_\rho. \tag{4.18}$$

Lemma 4.10. *Suppose that* $\rho := \pi^1_\# \mathbf{n}_1 = \pi^1_\# \mathbf{n}_2 \in \mathcal{P}_2(\mathbb{R}^d)$. *Then* $\mathbf{n}_1 \oplus \mathbf{n}_2 \neq \emptyset$.

Proof. Using conditional probability measures, we can write $\mathbf{n}_i(dx; d\eta_i) = \mathbf{n}_i(d\eta_i|x)\rho(dx)$, $i = 1, 2$. If we define

$$\mathbf{N}(dx; d\eta_1, d\eta_2) := \mathbf{n}_1(d\eta_1|x)\mathbf{n}_2(d\eta_2|x)\rho(dx),$$

it is immediate to check that $\mathbf{n} := (\pi^1, \pi^2 + \pi^3)_\# \mathbf{N} \in \mathbf{n}_1 \oplus \mathbf{n}_2$. □

Lemma 4.11. *Let* $\mathbf{n}_i \in \mathsf{Tan}_\rho$, $i = 1, 2$. *Then*

(1) $\mathbf{n}_1 \oplus \mathbf{n}_2 \subset \mathsf{Tan}_\rho$;
(2) $t \cdot \mathbf{n}_i \in \mathsf{Tan}_\rho$, *for all* $t \in \mathbb{R}$ *(in particular,* $(-1) \cdot \mathbf{n}_i \in \mathsf{Tan}_\rho$*).*

Proof. These are Propositions 4.25 and 4.29 of Gigli [25]. □

The fact that $(-1) \cdot \mathbf{n} \in \mathsf{Tan}_\rho$ whenever $\mathbf{n} \in \mathsf{Tan}_\rho$ leads to a nontrivial consequence that will help us simplifying later arguments considerably.

Lemma 4.12. *Let* $\mathbf{n} \in \partial^i_\rho f \cap \mathsf{Tan}_\rho$ *where* $i \in \{+, -\}$. *Then* $\|\mathbf{n}\|_\rho \le |Df|(\rho)$.

Proof. We prove the case of $\mathbf{n} \in \partial^+_\rho f \cap \mathsf{Tan}_\rho$. Then other one follows by noting that $\mathbf{n} \in \partial^-_\rho f$ if and only if $(-1) \cdot \mathbf{n} \in \partial^+_\rho(-f)$ and that $|Df|(\rho) = |D(-f)|(\rho)$.

First, by Lemma 4.11, $\mathbf{m} := (-1) \cdot \mathbf{n} \in \mathsf{Tan}_\rho$. By the density of $G(\rho)$ in Tan_ρ, there exist $\mathbf{m}_k \in G(\rho)$ and $\mathbf{M}_k \in \mathcal{P}_2((\mathbb{R}^d)^3)$ such that, setting

$$\mathbf{n}_k := (-1) \cdot \mathbf{m}_k,$$

there holds $\pi^{1,2}_\# \mathbf{M}_k = (-1) \cdot \mathbf{n}_k$, $\pi^{1,3}_\# \mathbf{M}_k = \mathbf{n}$ and

$$\lim_{k\to\infty} \mathsf{D}_\rho(\mathbf{n}, \mathbf{n}_k) = \lim_{k\to\infty} \int_{(\mathbb{R}^d)^3} |(-\xi) - \eta|^2 \mathbf{M}_k(dx; d\xi, d\eta) = 0.$$

Now, if $\epsilon_k > 0$ satisfy the property that $(\pi^1, \pi^1 + \epsilon_k \pi^2)_\# \mathbf{m}_k$ is an optimal plan, we can assume with no loss of generality that $\epsilon_k \to 0$ (because if this property holds for ϵ_k, it holds for all $\epsilon \in (0, \epsilon_k)$).

We now define

$$\tilde{\mathbf{M}}_k := (\pi^1, \epsilon_k \pi^2, \pi^3)_\# \mathbf{M}_k, \qquad \rho_k := (\pi^1 + \pi^2)_\# (\epsilon_k \cdot \mathbf{m}_k).$$

Then

$$(\pi^1, \pi^1 + \pi^2)_\# \tilde{\mathbf{M}}_k \in \Gamma_o(\rho, \rho_k), \qquad \pi^{1,3}_\# \tilde{\mathbf{M}}_k = \mathbf{n}.$$

By the definition of super-gradient,

$$\begin{aligned}
f(\rho_k) - f(\rho) &\le \int_{(\mathbb{R}^d)^3} (\xi \cdot \eta) \tilde{\mathbf{M}}_k(dx; d\xi, d\eta) + \mathsf{d}(\rho, \rho_k) \omega_{\mathbf{n}}(\mathsf{d}(\rho, \rho_k)) \\
&\le -\epsilon_k \|\mathbf{n}_k\|^2_\rho + \epsilon_k \mathsf{D}_\rho(\mathbf{n}, \mathbf{n}_k) \|\mathbf{n}_k\|_\rho + \mathsf{d}(\rho, \rho_k) \omega_{\mathbf{n}}(\mathsf{d}(\rho, \rho_k)) \\
&= \Big(- \|\mathbf{n}_k\|_\rho + \mathsf{D}_\rho(\mathbf{n}, \mathbf{n}_k) \Big) \|\epsilon_k \cdot \mathbf{m}_k\|_\rho + \mathsf{d}(\rho, \rho_k) \omega_{\mathbf{n}}(\mathsf{d}(\rho, \rho_k)) \\
&= \Big(- \|\mathbf{n}_k\|_\rho + \mathsf{D}_\rho(\mathbf{n}, \mathbf{n}_k) \Big) \mathsf{d}(\rho, \rho_k) + \mathsf{d}(\rho, \rho_k) \omega_{\mathbf{n}}(\mathsf{d}(\rho, \rho_k)).
\end{aligned}$$

Therefore

$$\|\mathbf{n}\|_\rho \le \limsup_{k\to\infty} \frac{(f(\rho) - f(\rho_k))^+}{\mathsf{d}(\rho_k, \rho)} \le \limsup_{\rho_1 \to \rho,\, \rho_1 \neq \rho} \frac{|f(\rho_1) - f(\rho)|}{\mathsf{d}(\rho_1, \rho)} = |Df|(\rho).$$

□

Lemma 4.13. *Let $\varphi := \varphi_1 + \varphi_2$, then $\partial^i_\rho \varphi_1 \oplus \partial^i_\rho \varphi_2 \subset \partial^i_\rho \varphi$, $i \in \{+,-\}$.*

Proof. We prove the case of super-gradients only, as the other case follows by symmetry.

Let $\mathbf{n} = \mathbf{n}_1 \oplus \mathbf{n}_2 \in \partial^+_\rho \varphi_1 \oplus \partial^+_\rho \varphi_2$ where $\mathbf{n}_i \in \partial^+_\rho \varphi_i$ for $i = 1, 2$. That is, there exists $\mathbf{N} \in \mathcal{P}_2((\mathbb{R}^d)^3)$ such that $\pi^{1,2}_\# \mathbf{N} = \mathbf{n}_1$, $\pi^{1,3}_\# \mathbf{N} = \mathbf{n}_2$ and

$$\mathbf{n}(dx; d\eta) = \int_{\eta_2} \int_{\eta_1} \delta_{\eta_1 + \eta_2}(d\eta) \mathbf{N}(dx; d\eta_1, d\eta_2).$$

That is,

$$\int_x \int_{\eta_2} \int_{\eta_1} \varphi(x, \eta_1 + \eta_2) \mathbf{N}(dx; d\eta_1, d\eta_2) = \int \int \varphi(x, \eta) \mathbf{n}(dx; d\eta). \tag{4.19}$$

For every $\rho_1 \in \mathsf{X}$ and every $\mathbf{M} \in \mathcal{P}_2((\mathbb{R}^d)^3)$ such that $(\pi^1, \pi^1 + \pi^2)_\# \mathbf{M} \in \Gamma_o(\rho, \rho_1)$ and $\pi^{1,3}_\# \mathbf{M} = \mathbf{n}$, we conclude the lemma by showing that

$$\varphi(\rho_1) - \varphi(\rho) \leq \int_{(\mathbb{R}^d)^3} (\xi \cdot \eta) \mathbf{M}(dx; d\xi, d\eta) + \mathsf{d}(\rho, \rho_1)(\omega_{\mathbf{n_1}} + \omega_{\mathbf{n_2}})(\mathsf{d}(\rho, \rho_1)), \tag{4.20}$$

where $\omega_{\mathbf{n_i}}$, $i = 1, 2$ are the modulus appearing in the definition of $\mathbf{n}_i \in \partial^+_\rho \varphi_i$.

To verify (4.20), we first construct a $\mathbf{P} \in \mathcal{P}_2((\mathbb{R}^d)^4)$ with certain desired properties. First, using $\pi^{1,3}_\# \mathbf{M} = \mathbf{n}$, we decompose

$$\mathbf{M}(dx; d\xi, d\eta) = \mathbf{M}_{x,\eta}(d\xi) \mathbf{n}(dx; d\eta)$$

and define

$$\mathbf{P}(dx; d\xi, d\eta_1, d\eta_2) := \mathbf{M}_{x, \eta_1 + \eta_2}(d\xi) \mathbf{N}(dx; d\eta_1, d\eta_2).$$

Then,

$$(\pi^1, \pi^2, \pi^3 + \pi^4)_\# \mathbf{P} = \mathbf{M}. \tag{4.21}$$

Indeed,

$$\begin{aligned}
& \int \int \int \int \phi(x, \xi, \eta_1 + \eta_2) \mathbf{P}(dx; d\xi, d\eta_1, d\eta_2) \\
= \ & \int_x \int_{\eta_2} \int_{\eta_1} \Big(\int_\xi \phi(x; \xi, \eta_1 + \eta_2) \mathbf{M}_{x, \eta_1 + \eta_2}(d\xi) \Big) \mathbf{N}(dx; d\eta_1, d\eta_2) \\
= \ & \int_x \Big[\int_\eta \Big(\int_\xi \phi(x; \xi, \eta) \mathbf{M}_{x,\eta}(d\xi) \Big) \mathbf{n}(dx; d\eta) \\
= \ & \int_x \int_\eta \int_\xi \phi(x; \xi, \eta) \mathbf{M}(dx; d\xi; d\eta).
\end{aligned}$$

In the above, from the first to the second equalities, we used (4.19) by taking

$$\varphi(x, \eta) = \int_\xi \phi(x; \xi, \eta) \mathbf{M}_{x,\eta}(d\xi). \tag{4.22}$$

Next, we define

$$\mathbf{M}_1(dx; d\xi, d\eta_1) := \pi^{1,2,3}_\# \mathbf{P}(dx; d\xi, d\eta_1) = \int_{\eta_2} \mathbf{M}_{x, \eta_1 + \eta_2}(d\xi) \mathbf{N}(dx; d\eta_1, d\eta_2),$$

$$\mathbf{M}_2(dx;d\xi,d\eta_2) := \pi^{1,2,4}_{\#}\mathbf{P}(dx;d\xi,d\eta_2) = \int_{\eta_1} \mathbf{M}_{x,\eta_1+\eta_2}(d\xi)\mathbf{N}(dx;d\eta_1,d\eta_2).$$

Then it is immediate to check that

$$\pi^{1,3}_{\#}\mathbf{M}_1 = \pi^{1,3}_{\#}\mathbf{P} = \pi^{1,2}_{\#}\mathbf{N} = \mathbf{n}_1, \qquad \pi^{1,3}_{\#}\mathbf{M}_2 = \pi^{1,4}_{\#}\mathbf{P} = \pi^{1,3}_{\#}\mathbf{N} = \mathbf{n}_2,$$

while (4.21) gives

$$(\pi^1,\pi^1+\pi^2)_{\#}\mathbf{M}_i = (\pi^1,\pi^1+\pi^2)_{\#}\mathbf{P} = (\pi^1,\pi^1+\pi^2)_{\#}\mathbf{M} \in \Gamma_o(\rho,\rho_1) \qquad i = 1,\, 2.$$

With the above properties, writing $\omega_{\mathbf{n}} := \omega_{\mathbf{n_1}} + \omega_{\mathbf{n_2}}$, we have

$$\begin{aligned} & \varphi_1(\rho_1) - \varphi_1(\rho) + \varphi_2(\rho_1) - \varphi_2(\rho) \\ \le\ & \int_{(\mathbb{R}^d)^3} (\xi\cdot\eta_1)\mathbf{M}_1(dx;d\xi,d\eta_1) + \int_{(\mathbb{R}^d)^3} (\xi\cdot\eta)\mathbf{M}_2(dx;d\xi,d\eta_2) + \mathsf{d}(\rho,\rho_1)\omega_{\mathbf{n}}(\mathsf{d}(\rho,\rho_1)) \\ =\ & \int_{(\mathbb{R}^d)^4} (\xi\cdot(\eta_1+\eta_2))\mathbf{P}(dx;d\xi,d\eta_1,d\eta_2) + \mathsf{d}(\rho,\rho_1)\omega_{\mathbf{n}}(\mathsf{d}(\rho,\rho_1)) \\ =\ & \int_{(\mathbb{R}^d)^3} (\xi\cdot\eta)\mathbf{M}(dx;d\xi,d\eta) + \mathsf{d}(\rho,\rho_1)\omega_{\mathbf{n}}(\mathsf{d}(\rho,\rho_1)). \end{aligned}$$

This establishes the validity of (4.20). □

4.4. **Augmented Lagrangian and Hamiltonian, Viscosity solution of a Hamilton-Jacobi equation in $\mathcal{P}_2(\mathbb{R}^d)$ and uniqueness.** Now, we define an *augmented* Lagrangian

$$L(\mathbf{m}) := \frac{1}{2}\|\mathbf{m}\|^2_\rho - V(\pi^1_{\#}\mathbf{m}), \qquad \forall \mathbf{m} \in \mathsf{Tan}.$$

We also define, for $\mathbf{n} \in \mathcal{P}_2(\mathbb{R}^d\times\mathbb{R}^d)$,

$$\mathbf{H}(\mathbf{n}) := \frac{1}{2}\|\mathbf{n}\|^2_\rho + V(\pi^1_{\#}\mathbf{n}).$$

We study well-posedness of a Hamilton-Jacobi partial differential equation formally written as

$$f(\rho) - \alpha\mathbf{H}(\rho,\partial_\rho f) = h(\rho), \tag{4.23}$$

where $\alpha > 0$ and $h \in BUC(\mathsf{X})$. The Cauchy problem formulated using this augmented form can be treated similarly, we do not pursue details here to avoid repetition.

We define, for every $f : \mathsf{X} \mapsto \mathbb{R}$,

$$\mathbf{H}_0 f(\rho) \ := \ \inf\Big\{\mathbf{H}(\mathbf{n}) :\ \mathbf{n} \in \partial^+_\rho f \cap \mathsf{Tan}_\rho\Big\}, \tag{4.24}$$

$$\mathbf{H}_1 f(\gamma) \ := \ \sup\Big\{\mathbf{H}(\mathbf{n}) :\ \mathbf{n} \in \partial^-_\gamma f \cap \mathsf{Tan}_\gamma\Big\}. \tag{4.25}$$

We recall the conventions that $\inf\emptyset = +\infty$ and $\sup\emptyset = -\infty$. By density of $G(\rho)$ in Tan_ρ, the sup on Tan_ρ (respectively Tan_γ) in the above defining equalities can be replaced by sup on $G(\rho)$ (respectively by $G(\gamma)$).

The restriction of $\mathbf{n} \in \mathsf{Tan}$ in (4.24) and (4.25) deserves an explanation. We observe that the roles of $\pi^{1,2}_{\#}\mathbf{M}$ and $\pi^{1,3}_{\#}\mathbf{M}$ in Definition 4.7 are not symmetric. In particular, if $\mathbf{n} = \pi^{1,3}_{\#}\mathbf{M} \notin \mathsf{Tan}$, then

$$\mathbf{M}(dx;d\xi,d\eta) := \delta_\eta(d\xi)\mathbf{n}(dx;d\eta)$$

won't belong to the admissible class appearing in definition of super- sub- differentials. A practical significance of the above $\mathbf{M}$ is the property that

$$\int_{(\mathbb{R}^d)^3} (\xi \cdot \eta)\mathbf{M}(dx; d\xi, d\eta) = \|\mathbf{n}\|_\rho^2, \quad \text{where } \rho = \pi^1_\# \mathbf{M}.$$

We will use this in the existence of super-solution in Lemma 4.23 near the very end. In principle, one could have developed a theory without the restriction $\mathbf{n} \in \mathsf{Tan}$. However, this will complicate the definition of $\mathbf{H}_0$ and $\mathbf{H}_1$ requiring the use of measures $\mathbf{M}$ with 3 variables $(x; \xi, \eta)$, instead of the $\mathbf{n}$ with 2-variables (x, ξ).

Definition 4.14 (Viscosity solution using sub, super-differentials)**.** f is a viscosity sub-solution to (4.23) if $f \in USC(\mathsf{X}; \mathbb{R})$ and

$$\alpha^{-1}(f - h)(\rho) \le (\mathbf{H}_0 f)^*(\rho), \quad \rho \in \mathsf{X};$$

f is a viscosity super-solution to (4.23) if $f \in LSC(\mathsf{X}; \mathbb{R})$ and

$$\alpha^{-1}(f - h)(\rho) \ge (\mathbf{H}_1 f)_*(\rho), \quad \rho \in \mathsf{X}.$$

If $f \in C(\mathsf{X}; \mathbb{R})$ is both a viscosity sub- as well as super-solution, it is a solution.

We will refer the above equation and its definition of solution as *geometric-based* formulation. In contrast, we refer to the earlier equation (1.11) in conjunction with Definition 1.4 as *metric-based* formulation. We reveal next that uniqueness of the metric based formulation, which follows from Lemma 2.3, implies uniqueness of the geometric-based formulation.

Let φ and ψ be the special test functions in (1.12) and (1.13), and $H_0^{\lambda,\kappa,\epsilon}$ and $H_1^{\kappa,\epsilon}$ be the metric-based formulation of Hamiltonians defined in Section 1.2 using the notion of local slope. Next, we prove an estimate of $\mathbf{H}_0\varphi$ from above by the $H_0^{\lambda,\kappa,\epsilon}\varphi$, and prove another estimate of $\mathbf{H}_1\psi$ from below by the $H_1^{\kappa,\epsilon}\psi$. To be notationally consistent throughout this section, we always use the probability-measure-space notations, even when dealing with the metric formulation. Specifically, we recall that a base point $\bar\rho \in \mathsf{X}$ is chosen. It has smooth compactly supported Lebesgue density, as defined in the beginning of this section. We also recall that, for $\delta, \kappa, \epsilon > 0$,

$$\varphi(\rho) := \lambda^{-1}\Big(\frac{\mathsf{d}^2(\rho,\gamma)}{2\delta} + \kappa\beta \circ \mathsf{d}(\rho,\bar\rho) + \epsilon \mathsf{d}(\rho,\rho_1)\Big), \quad \lambda > 1, \gamma, \rho_1 \in \mathsf{X}, \tag{4.26}$$

and that

$$H_0^{\lambda,\kappa,\epsilon}\varphi(\rho) \;:=\; H\Big(\rho, \frac{1}{\lambda}(\frac{\mathsf{d}(\rho,\gamma)}{\delta} + \kappa\beta' \circ \mathsf{d}(\rho,\bar\rho) + \epsilon)\Big),$$

where

$$H(\rho, p) := \frac{1}{2}p^2 + V(\rho), \quad (\rho, p) \in \mathsf{X} \times \mathbb{R}_+ \mapsto \bar{\mathbb{R}}.$$

Similarly,

$$\psi(\gamma) := -\frac{\mathsf{d}^2(\rho,\gamma)}{2\delta} - \kappa\beta \circ \mathsf{d}(\gamma,\bar\rho) - \epsilon \mathsf{d}(\gamma,\gamma_1), \quad \rho, \gamma_1 \in \mathsf{X}, \tag{4.27}$$

and

$$H_1^{\kappa,\epsilon}\psi(\gamma) \;:=\; H\Big(\gamma, (\frac{\mathsf{d}(\rho,\gamma)}{\delta} - \kappa\beta' \circ \mathsf{d}(\gamma,\bar\rho) - \epsilon) \vee 0\Big).$$

We note that $\partial^+_\rho \mathsf{d}(\cdot, \rho_1)$ is empty at $\rho = \rho_1$. Therefore, $\partial^+_\rho\varphi$ and $\partial^-_\gamma\psi$ may be empty when evaluated at some important points, making the use of $\mathbf{H}_0\varphi$ and $\mathbf{H}_1\psi$ non-informative at these points. Next, we consider further perturbations which are smooth in the sense of super- sub-differentials.

Let

$$\mathsf{d}_\alpha(\rho,\gamma) := \sqrt{\mathsf{d}^2(\rho,\gamma)+\alpha}, \quad \rho,\gamma \in \mathsf{X}, \alpha > 0.$$

Let α, $\theta > 0$ be small, and $\beta_k \in [0,1]$ with $\sum_k \beta_k = 1$ and $\rho_k \in \mathsf{X}$ where $k = 1, 2, \ldots$. We build a bump function

$$\Delta(\rho) := \sum_{k=1}^{\infty} \beta_k \mathsf{d}^2(\rho,\rho_k). \tag{4.28}$$

and define

$$\varphi_{\alpha,\theta}(\rho) := \lambda^{-1}\Big(\frac{\mathsf{d}^2(\rho,\gamma)}{2\delta} + \kappa\beta \circ \mathsf{d}_\alpha(\rho,\bar\rho) + \epsilon \mathsf{d}_\alpha(\rho,\rho_1)\Big) + \theta\Delta(\rho), \qquad \forall \rho \in \mathsf{X}. \tag{4.29}$$

By Lemma 4.8 and Lemma 4.11(2), each summand is superdifferentiable and it admits tangent elements in the superdifferential. By Lemma 4.13 and Lemma 4.11(1), we obtain

$$\partial_\rho^+ \varphi_{\alpha,\theta} \cap \mathsf{Tan}_\rho \neq \emptyset \qquad \forall \rho \in \mathsf{X}.$$

Although there are countable sum of distances in Δ, the conclusions of the above lemmas still holds when applied to Δ due to the summability of β_ks and the special form of sub-super-differentials identified in Lemma 4.8.

Lemma 4.15. *Suppose that* $\lim_{\alpha,\theta\to 0} \mathsf{d}(\rho_{\alpha,\theta},\rho) = 0$ *and that* $\sup_k \mathsf{d}(\rho,\rho_k) \le M < \infty$ *with* M *independent of* θ, α. *Then*

$$\limsup_{\alpha\to 0^+,\theta\to 0^+} \mathbf{H}_0 \varphi_{\alpha,\theta}(\rho_{\alpha,\theta}) \le H_0^{\lambda,\kappa,\epsilon}\varphi(\rho).$$

Proof. For notational simplicity, we re-write the $\varphi_{\alpha,\theta}$ in (4.29) into four terms corresponding to each of the terms in that expression

$$\varphi_{\alpha,\theta} := \lambda^{-1}(\varphi_0 + \kappa\varphi_1 + \epsilon\varphi_2) + \theta\Delta.$$

For all $\mathbf{n} \in \partial_\rho^+ \varphi_{\alpha,\theta} \cap \mathsf{Tan}_\rho$, invoking Lemma 4.12,

$$\|\mathbf{n}\|_\rho \le \frac{1}{\lambda}\Big(\frac{\mathsf{d}(\rho,\gamma)}{\delta} + \kappa\beta' \circ \mathsf{d}_\alpha(\rho,\bar\rho) + \epsilon\Big) + 2\theta \sup_{k=1,2,\ldots} \mathsf{d}(\rho,\rho_k). \tag{4.30}$$

Since $\mathbb{R}_+ \ni p \mapsto H(\rho,p)$ is nondecreasing, we have

$$\begin{aligned} \mathbf{H}_0\varphi_{\alpha,\theta}(\rho_{\alpha,\theta}) &:= \inf\Big\{\mathbf{H}(\mathbf{n}) : \mathbf{n} \in \partial^+_{\rho_{\alpha,\theta}}\varphi \cap \mathbf{Tan}_{\rho_{\alpha,\theta}}\Big\} \\ &\le H\Big(\rho_{\alpha,\theta}, \frac{1}{\lambda}(\frac{\mathsf{d}(\rho_{\alpha,\theta},\gamma)}{\delta} + \kappa\beta' \circ \mathsf{d}_\alpha(\rho_{\alpha,\theta},\bar\rho) + \epsilon) + 2\theta \sup_k \mathsf{d}(\rho_{\alpha,\theta},\rho_k)\Big). \end{aligned}$$

We conclude by taking the limit. $\square$

Similarly, we construct a smoothly perturbed version of the function ψ in (4.27):

$$\psi_{\alpha,\theta}(\gamma) := -\frac{\mathsf{d}^2(\rho,\gamma)}{2\delta} - \kappa\beta \circ \mathsf{d}_\alpha(\gamma,\bar\gamma) - \epsilon\mathsf{d}_\alpha(\gamma,\gamma_1) - \theta\Delta(\gamma). \tag{4.31}$$

Lemma 4.16. *Suppose that* $\lim_{\alpha,\theta\to 0} \mathsf{d}(\gamma_{\alpha,\theta},\gamma)=0$ *and that* $\sup_k \mathsf{d}(\gamma,\rho_k)\le M<\infty$ *with* M *independent of* θ, α. *Then*

$$\liminf_{\theta\to 0^+,\alpha\to 0^+} \mathbf{H}_1\psi_{\alpha,\theta}(\gamma_{\alpha,\theta}) \ge H_1^{\kappa,\epsilon}\psi(\gamma).$$

Proof. We re-write the test function ψ in (4.31) into four terms $\psi_{\alpha,\theta}:=\psi_0+\kappa\psi_1+\epsilon\psi_2-\theta\Delta$, where in particular, $\psi_0(\gamma)=-\frac{1}{2\delta}\mathsf{d}^2(\rho,\gamma)$.

By Lemmas 4.8 and 4.11(2), for any $\boldsymbol{\mu}\in\Gamma_o(\rho,\gamma)$ the plan

$$\mathbf{n}_0 := \Big(\pi^2, \frac{1}{\delta}(\pi^1-\pi^2)\Big)_{\#}\boldsymbol{\mu}$$

belongs to $\partial_\gamma^-\psi_0\cap\mathsf{Tan}_\gamma$. Also, Lemma 4.13 provides

$$\mathbf{n}_1\in\Big(\partial_\gamma^-(\kappa\psi_1+\epsilon\psi_2-\theta\Delta)\Big)\cap\mathsf{Tan}_\gamma.$$

Take any $\mathbf{n}\in\mathbf{n}_0\oplus\mathbf{n}_1$ (whose existence is guaranteed by Lemma 4.10). By Lemmas 4.13 and 4.11(1), $\mathbf{n}\in\partial_\gamma^-\psi_{\alpha,\theta}\cap\mathsf{Tan}_\gamma$. Additionally, by Lemma 4.12,

$$\|\mathbf{n}_1\|_\gamma\le|D(\kappa\psi_1+\epsilon\psi_2-\theta\Delta)|(\gamma)\le\epsilon':=(\kappa\beta'\circ\mathsf{d}_\alpha(\gamma,\bar\rho)+\epsilon+2\theta M).$$

Hence in view of (4.18), we have that $\|\mathbf{n}_0\|_\gamma\le\|\mathbf{n}\|_\gamma+\epsilon'$, that is,

$$\|\mathbf{n}\|_\gamma\ge(\|\mathbf{n}_0\|_\gamma-\epsilon')\vee 0=\Big(\frac{\mathsf{d}(\rho,\gamma)}{\delta}-\epsilon'\Big)\vee 0.$$

By monotonicity of $\mathbb{R}_+\ni p\mapsto H(\rho,p)$,

$$\begin{aligned}\mathbf{H}_1\psi_{\alpha,\theta}(\gamma) &:= \sup\{\mathbf{H}(\hat{\mathbf{n}}):\hat{\mathbf{n}}\in\partial_\gamma^-\psi_{\alpha,\theta}\cap\mathsf{Tan}_\gamma\}\\ &\ge \mathbf{H}(\mathbf{n})\\ &\ge H\Big(\gamma,\Big(\frac{\mathsf{d}(\rho,\gamma)}{\delta}-\kappa\beta'\circ\mathsf{d}_\alpha(\gamma,\bar\rho)-\epsilon-\theta M\Big)\vee 0\Big).\end{aligned}$$

Replace the γ above by $\gamma_{\alpha,\theta}$ and taking the limit, the conclusion follows. □

Next, we link $\mathbf{H}_i f$, $i=1$, 2 with $\mathbf{H}_0\varphi_{\alpha,\theta}$ and $\mathbf{H}_1\psi_{\alpha,\theta}$, at those special points ρ, γ appearing in the maximum principle.

Lemma 4.17. *Let* $f:\mathsf{X}\mapsto\mathbb{R}$, $\alpha\ge 0,\theta\ge 0$, *and* $\rho_0,\gamma_0\in\mathsf{X}$ *satisfy*

$$(f-\varphi_{\alpha,\theta})(\rho)=\sup_{\mathsf{X}}(f-\varphi_{\alpha,\theta}),\qquad(\psi_{\alpha,\theta}-f)(\gamma)=\sup_{\mathsf{X}}(\psi_{\alpha,\theta}-f).$$

Then $\partial_\rho^+\varphi_{\alpha,\theta}\subset\partial_\rho^+ f$ *and* $\partial_\gamma^-\psi_{\alpha,\theta}\subset\partial_\gamma^- f$. *Consequently, at such* ρ *and* γ *we respectively have*

$$\mathbf{H}_0 f(\rho)\le\mathbf{H}_0\varphi_{\alpha,\theta}(\rho),\qquad\mathbf{H}_1 f(\gamma)\ge\mathbf{H}_1\psi_{\alpha,\theta}(\gamma).$$

Proof. By assumption, $\varphi_{\alpha,\theta}(\rho_1)-\varphi_{\alpha,\theta}(\rho)\ge f(\rho_1)-f(\rho)$ for all $\rho_1\in\mathsf{X}$. Hence $\partial_\rho^+\varphi_{\alpha,\theta}\subset\partial_\rho^+ f$ by the defining inequality of super-differential in Definition 4.7.

Similarly, $f(\gamma_1)-f(\gamma)\ge\psi_{\alpha,\theta}(\gamma_1)-\psi_{\alpha,\theta}(\gamma)$ for all $\gamma_1\in\mathsf{X}$. Hence $\partial_\gamma^-\psi_{\alpha,\theta}\subset\partial_\gamma^- f$ follows from definition of sub-differential. □

Combining the above lemmas, we have the following relation between the metric-based and geometric-based formulations of viscosity solutions.

Lemma 4.18. *Every viscosity sub- (resp. super-) solution to* (4.23) *in the sense of Definition 4.14, is a viscosity sub- (resp. super-) solution to* (1.11) *in the sense of Definition 1.4.*

As a consequence, the comparison principle for (1.11) *implies the comparison principle for* (4.23).

Proof. We only prove the sub-solution case, the other case is similar.

Let $(f-\varphi)(\rho_0) = \sup_{\mathsf{X}}(f-\varphi)$, where φ is defined as in (4.26). We write

$$\varphi_\alpha := \lambda^{-1}\Big(\frac{\mathsf{d}^2(\rho,\gamma)}{2\delta} + \kappa\beta \circ \mathsf{d}_\alpha(\rho,\bar{\rho}) + \epsilon \mathsf{d}_\alpha(\rho,\rho_1)\Big).$$

Noting $\mathsf{d}_\alpha \geq \mathsf{d}$, we have

$$(f-\varphi_\alpha)(\rho_0) \geq \sup_{\mathsf{X}}(f-\varphi_\alpha) - \theta.$$

where

$$\theta := \theta(\alpha) := \frac{\kappa}{\lambda}\Big(\beta\circ \mathsf{d}_\alpha(\rho_0,\bar{\rho}) - \beta\circ\mathsf{d}(\rho_0,\bar{\rho})\Big) + \frac{\epsilon}{\lambda}\Big(\mathsf{d}_\alpha(\rho_0,\rho_1) - \mathsf{d}(\rho_0,\rho_1)\Big) \in \mathbb{R}_+.$$

By applying the Borwein-Preiss variational principle in Proposition 5.2 with $\epsilon = \theta$ and $F = f - \varphi_\alpha$, we can now find $(\beta_k,\rho_k) \in [0,1]\times\mathsf{X}$, $k = 1,2,\ldots$ (this sequence may depend on all the earlier parameters α,θ and ρ_0) and $\rho_{\alpha,\theta}\in\mathsf{X}$ such that

$$(f-\varphi_{\alpha,\theta})(\rho_{\alpha,\theta}) = \sup_{\mathsf{X}}(f-\varphi_{\alpha,\theta}),$$

where $\varphi_{\alpha,\theta} := \varphi_\alpha - \sqrt{\theta}\Delta$ and Δ is defined as in (4.28):

$$\Delta(\rho) := \sum_{k=1}^\infty \beta_k \mathsf{d}^2(\rho,\rho_k).$$

Moreover,

$$\sup_k \mathsf{d}(\rho_k,\rho_0) \leq \theta^{1/4}, \qquad \mathsf{d}(\rho_0,\rho_{\alpha,\theta}) \leq \theta^{1/4}.$$

Now, combining Lemmas 4.15 and 4.17 the conclusion follows. □

4.5. **Existence of viscosity solution for Hamilton-Jacobi PDEs in $\mathcal{P}_2(\mathbb{R}^d)$.** In view of Lemma 4.1, the value function f defined in (1.10) becomes

$$f(\rho) = \sup\Big\{\int_0^\infty e^{-r/\alpha}\Big(\frac{h(\sigma(r))}{\alpha} - L(\sigma,\dot{\sigma})\Big)dr : \sigma(\cdot)\in AC_{loc}(\mathbb{R}_+;\mathsf{X}),\ \sigma(0)=\rho\Big\}. \tag{4.32}$$

Lemma 4.19. *Condition 4.2 implies Condition 1.1.5. Moreover, the value function f is continuous, bounded from above and has at most sub-linear growth rate to $-\infty$.*

Proof. For any $0<\theta<2$,

$$\int_{\mathbb{R}^d}|x|^\theta d\rho \leq \Big(\int_{\mathbb{R}^d}|x|^2 d\rho\Big)^{\theta/2} \leq C(1+\mathsf{d}(\rho,\bar{\rho}))^\theta.$$

With the assumptions on ϕ,Φ in Condition 4.2 and the choice of $\theta\in(0,1)$, it follows that there exist C_0, $C_1\in\mathbb{R}_+$ such that

$$-C_0(1+\mathsf{d}(\rho,\bar{\rho}))^\theta \leq V(\rho) \leq C_1.$$

Hence we may choose $\zeta(r) = C(1+r^\theta)$ for some $C\in\mathbb{R}_+$ and $\beta(r) = r$ to verify Condition 1.1.5.

The continuity on f follows from Lemma 2.6, while the growth estimates follows from Lemma 2.7. $\square$

We already know, through Theorem 2.15, that f is a viscosity solution to the metric formulation of Hamilton-Jacobi equation (1.11). Next, we show that it is also a viscosity solution to (4.23) as given in the sense, a priori stronger, of Definition 4.14. First, we observe the following small time behavior of semigroup T_t. Denote

$$C_{\phi,\Phi} := \|\nabla \phi\|_{L^\infty} + \|\nabla \Phi\|_{L^\infty}.$$

Under Condition 4.2, $\sup_\rho |D_\rho V| \le C_{\phi,\Phi}$. Then we have the following.

Lemma 4.20. *Let $f : \mathsf{X} \mapsto \mathbb{R}$ be an arbitrary Borel function which is bounded from above. Then for each $t > 0$ and $\rho_0 \in \mathsf{X}$ with $f(\rho_0) > -\infty$, we can find $\rho_1 := \rho_{1,t} \in \mathsf{X}$ satisfying*

$$\mathsf{d}^2(\rho_0, \rho_1) \le C_f t, \quad t \in [0,1] \tag{4.33}$$

where C_f is a constant only depending on f, such that

$$T_t f(\rho_0) - f(\rho_0) \;\le\; t^2 + f(\rho_1) - f(\rho_0) - tL\Big(\rho_0, \frac{\mathsf{d}(\rho_0,\rho_1)}{t}\Big) + C_{\phi,\Phi}\sqrt{C_f}\, t\sqrt{t}. \tag{4.34}$$

Proof. For every absolutely continuous curve $\rho(\cdot)$,

$$\int_0^t (V(\rho(s)) - V(\rho(0)))ds \le C_{\phi,\Phi} \int_0^t \int_0^s |\rho'|(r) dr ds \le t C_{\phi,\Phi} \int_0^t |\rho'(r)| dr.$$

Recall $L(\rho, q) = q^2/2 - V(\rho)$, we denote

$$L_t(\rho, q) := L(\rho, q) - tC_{\phi,\Phi} q.$$

Then

$$\begin{aligned}
& T_t f(\rho_0) - f(\rho_0) \\
= \;& \sup \Big\{ f(\rho(t)) - f(\rho(0)) - \int_0^t L(\rho_0, |\rho'|(s)) ds \\
& \qquad + \int_0^t (V(\rho(s)) - V(\rho_0)) ds : \rho(\cdot) \in AC([0,t]; \mathsf{X}),\ \rho(0) = \rho_0 \Big\} \\
\le \;& \sup \Big\{ f(\rho(t)) - f(\rho_0) - \int_0^t L_t(\rho_0, |\rho'|(s)) ds : \rho(\cdot) \in AC([0,t]; \mathsf{X}),\ \rho(0) = \rho_0 \Big\} \\
= \;& \sup_{\rho_1 \in \mathsf{X}} \Big\{ f(\rho_1) - f(\rho_0) - tL_t\Big(\rho_0, \frac{\mathsf{d}(\rho_0,\rho_1)}{t}\Big) \Big\},
\end{aligned}$$

where the last line follows from convexity of L_t in $q \in \mathbb{R}_+$.

Since f is bounded from above and $L(\rho_0, q) = q^2/2 - V(\rho_0)$, we can always restrict the ρ_1s in the last maximization problem to satisfy (4.33), uniformly in t in bounded time interval. Hence the conclusion (4.34) follows. $\square$

Lemma 4.21. *The value function f in* (4.32) *is a continuous viscosity sub-solution to* (4.23) *in the sense of Definition 4.14.*

Proof. In view of Lemma 2.12, we only need to show that for the value function f and every $\rho_0 \in \mathsf{X}$, we have

$$\liminf_{t \to 0^+} \frac{1}{t} \Big(T_t f(\rho_0) - f(\rho_0) \Big) \le (\mathbf{H}_0 f)(\rho_0). \tag{4.35}$$

Without loss of generality, we assume $\partial^+_{\rho_0} f$ is non-empty, otherwise by convention, inf over the empty set is $+\infty$, hence $\mathbf{H}_0 f(\rho_0) = +\infty$ and the above inequality holds trivially. We apply the estimate in Lemma 4.20. For each fixed $t > 0$ and the $\rho_1 := \rho_{1,t}$ appearing in that lemma, let $\mathbf{n} \in \partial^+_{\rho_0} f$ and let $\mathbf{M} \in \mathcal{P}_2((\mathbb{R}^d)^3)$ such that $(\pi^1, \pi^1 + \pi^2)_\# \mathbf{M} \in \Gamma_o(\rho, \rho_1)$ and $\pi^{1,3}_\# \mathbf{M} = \mathbf{n}$. From (4.34) and the defining relation of super-differential, we have

$$\begin{aligned}
T_t f(\rho_0) - f(\rho_0) &\leq t^2 + \|\mathbf{n}\|_{\rho_0} \mathsf{d}(\rho_0, \rho_{1,t}) + \mathsf{d}(\rho_0, \rho_{1,t}) \omega_{\mathbf{n}}(\mathsf{d}(\rho_0, \rho_{1,t})) - tL\Big(\rho_0, \frac{\mathsf{d}(\rho_0, \rho_{1,t})}{t}\Big) + Ct\sqrt{t} \\
&= t^2 + t\Big[\Big(\|\mathbf{n}\|_{\rho_0} + \omega_{\mathbf{n}}(\mathsf{d}(\rho_0, \rho_{1,t}))\Big) \frac{\mathsf{d}(\rho_0, \rho_{1,t})}{t} - L\Big(\rho_0, \frac{\mathsf{d}(\rho_0, \rho_{1,t})}{t}\Big)\Big] + Ct\sqrt{t} \\
&\leq t^2 + Ct\sqrt{t} + t\frac{1}{2}\Big(\|\mathbf{n}\|_{\rho_0} + \omega_{\mathbf{n}}(\mathsf{d}(\rho_0, \rho_{1,t}))\Big)^2 + tV(\rho_0).
\end{aligned}$$

Consequently (4.35) follows. □

Next, we also establish a scaling property for the defining inequality in super-differentials. Recall that, for $\gamma_0 \in \mathsf{X}$, $\mathbf{m} \in G(\gamma_0)$ means that $\epsilon > 0$ such that $(\pi^1, \pi^1 + \epsilon\pi^2)_\# \mathbf{m}$ is optimal between its first marginal, namely γ_0, and its second marginal, namely $(\pi^1 + \epsilon\pi^2)_\# \mathbf{m}$.

Lemma 4.22. *Let $f : \mathsf{X} \mapsto \mathbb{R}$, $\gamma_0 \in \mathsf{X}$, $\mathbf{m} \in G(\gamma_0)$ and $\mathbf{n}_0 \in \partial^-_{\gamma_0} f$. Given $\epsilon > 0$ such that $(\pi^1, \pi^1 + \epsilon\pi^2)_\# \mathbf{m}$ is an optimal plan, we define*

$$\gamma(t) := (\pi^1 + t\pi^2)_\# \mathbf{m}, \qquad t \in [0, \epsilon]. \tag{4.36}$$

Then $\gamma(t)$ is a constant speed geodesic between $\gamma(0) = \gamma_0$ and $\gamma(\epsilon)$ and there exists a modulus of continuity $\omega_{\mathbf{n}_0}$ satisfying this property: for all $\mathbf{M}_0 \in \mathcal{P}((\mathbb{R}^d)^3)$ such that $\pi^{1,2}_\# \mathbf{M}_0 = \mathbf{m}$ and $\pi^{1,3}_\# \mathbf{M}_0 = \mathbf{n}_0$, there holds

$$f(\gamma(t)) - f(\gamma_0) \geq t \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} (\xi \cdot \eta) \mathbf{M}_0(dy; d\xi, d\eta) + \mathsf{d}(\gamma_0, \gamma(t)) \omega_{\mathbf{n_0}}(\mathsf{d}(\gamma_0, \gamma(t))), \quad \forall t \in [0, \epsilon].$$

Proof. The fact that $\gamma(t)$ is a constant speed geodesic is well-known, see e.g. Lemma 7.2.1 of [2]. For $t \in [0, \epsilon]$ we define $\mathbf{M}_t := (\pi^1, t\pi^2, \pi^3)_\# \mathbf{M}_0$. Then $\pi^{1,3}_\# \mathbf{M}_t = \mathbf{n}_0$ and

$$(\pi^1, \pi^1 + \pi^2)_\# \mathbf{M}_t = (\pi^1, \pi^1 + t\pi^2)_\# \mathbf{M}_0 \in \Gamma_o(\gamma_0, \gamma(t)).$$

Consequently, by definition of sub-differential of f,

$$\begin{aligned}
f(\gamma(t)) - f(\gamma_0) &\geq \int_{(\mathbb{R}^d)^3} (\xi \cdot \eta) \mathbf{M}_t(dy; d\xi, d\eta) + \mathsf{d}(\gamma_0, \gamma(t)) \omega_{\mathbf{n_0}}(\mathsf{d}(\gamma_0, \gamma(t))) \\
&= t \int_{(\mathbb{R}^d)^3} (\xi \cdot \eta) \mathbf{M}_0(dy; d\xi, d\eta) + \mathsf{d}(\gamma_0, \gamma(t)) \omega_{\mathbf{n_0}}(\mathsf{d}(\gamma_0, \gamma(t))).
\end{aligned}$$

□

Lemma 4.23. *The value function f in (4.32) is a continuous viscosity super-solution to (4.23) in the sense of Definition 4.14.*

Proof. In view of Lemma 2.11, we only need to verify that for the value function f and every $\gamma_0 \in \mathsf{X}$, we have

$$\limsup_{t \to 0^+} \frac{1}{t}\Big(T_t f(\gamma_0) - f(\gamma_0)\Big) \geq (\mathbf{H}_1 f)(\gamma_0).$$

Let $\mathbf{n}_0 \in \partial^-_{\gamma_0} f$ (if no such $\mathbf{n}_0$ exists by convention the sup over the empty set is $-\infty$, hence $\mathbf{H}_1 f(\gamma_0) = -\infty$ and the above inequality holds trivially). Let $\mathbf{m} \in G(\gamma_0)$, let $\epsilon > 0$ be such that $(\pi^1, \pi^1 + \epsilon\pi^2)_{\#}\gamma_0$ is an optimal plan, and define $\gamma(t)$ as in (4.36). Since $\gamma(t)$ is a constant speed geodesic,

$$|\gamma'|(t) = \epsilon^{-1}\mathsf{d}(\gamma_0, \gamma(\epsilon)) = \|\mathbf{m}\|_{\gamma_0}, \qquad 0 < t < \epsilon.$$

Taking this into account, by the definition of T_t in (2.7), we also have

$$T_t f(\gamma_0) - f(\gamma_0) \;\; \geq \;\; f(\gamma(t)) - f(\gamma_0) - \int_0^t \Big(\frac{1}{2}\|\mathbf{m}\|^2_{\gamma_0} - V(\gamma(s))\Big) ds.$$

Now, let $\mathbf{M}_0 \in \mathcal{P}((\mathbb{R}^d)^3)$ be such that $\pi^{1,2}_{\#}\mathbf{M}_0 = \mathbf{m}$ and $\pi^{1,3}_{\#}\mathbf{M}_0 = \mathbf{n}_0$. From Lemma 4.22, for all $t \in [0, \epsilon]$ we conclude that

$$f(\gamma(t)) - f(\gamma_0) \geq t\Big(\int_{(\mathbb{R}^d)^3} (\xi \cdot \eta)\mathbf{M}_0(dy; d\xi, d\eta) + \epsilon^{-1}\mathsf{d}(\gamma_0, \gamma(\epsilon))\omega_{\mathbf{n_0}}(\mathsf{d}(\gamma_0, \gamma(t)))\Big).$$

Consequently, the lower semicontinuity of V give

$$\limsup_{t\to 0^+} \frac{1}{t}\Big(T_t f(\gamma_0) - f(\gamma_0)\Big) \;\; \geq \;\; \int_{(\mathbb{R}^d)^3} (\xi \cdot \eta - \frac{1}{2}|\xi|^2)\mathbf{M}_0(dy; d\xi, d\eta) + V(\gamma_0).$$

Now we make the extra assumption that $\mathbf{n}_0 \in \partial^-_{\gamma_0} f \cap \mathsf{Tan}_{\gamma_0}$. If the slightly stronger condition $\mathbf{n}_0 \in \partial^-_{\gamma_0} f \cap G(\gamma_0)$ holds, then taking

$$\mathbf{M}_0(dx; d\xi, d\eta) := \delta_\eta(d\xi)\mathbf{n}_0(dx; d\eta),$$

(which is admissible since $\pi^{1,2}_{\#}\mathbf{M}_0 = \mathbf{n_0} \in G(\gamma_0)$) and noting

$$\int_{(\mathbb{R}^d)^3} (\xi \cdot \eta - \frac{1}{2}|\xi|^2)\mathbf{M}_0(dy; d\xi, d\eta) = \int_{(\mathbb{R}^d)^3} \frac{1}{2}|\eta|^2 \mathbf{M}_0(dx; d\xi, d\eta) = \frac{1}{2}\|\mathbf{n}_0\|^2_{\gamma_0},$$

the conclusion follows. In the general situation $\mathbf{n}_0 \in \mathsf{Tan}_{\gamma_0}$, if $\mathbf{m}_n \in G(\gamma_0)$ converge to $\mathbf{n}_0$ we construct $\mathbf{M}^n_0$ satisfying $\pi^{1,2}_{\#}\mathbf{M}^n_0 = \mathbf{m}_n$, $\pi^{1,3}_{\#} = \mathbf{n}_0$ and use the fact that

$$\lim_{n\to\infty} \int_{(\mathbb{R}^d)^3} (\xi \cdot \eta - \frac{1}{2}|\xi|^2)\mathbf{M}^n_0(dy; d\xi, d\eta) = \frac{1}{2}\|\mathbf{n}_0\|^2_{\gamma_0}$$

to conclude. □

In summary, we have

Lemma 4.24. *The value function f is a viscosity solution to both* (4.23) *and* (1.11).

4.6. **Well-posedness of Hamilton-Jacobi PDE in $\mathcal{P}_2(\mathbb{R}^d)$.** In view of the comparison result in Lemma 4.18, we conclude with the following.

Theorem 4.25. *The value function f is the unique continuous viscosity solution with at most sub-linear growth, both to the metric-based formulation* (1.11) *as well as to the geometric-based formulation* (4.23).

In principle, we can also treat the Hamilton-Jacobi PDE associated with Vlasov-Monge-Ampére equation as described in the last section of [1], using the same methodology. We do not pursue details here further.

5. Appendix - perturbed optimization principles

Let $(\mathsf{Y}, \mathsf{d}_Y)$ be a complete metric space and let $F : \mathsf{Y} \mapsto \mathbb{R} \cup \{-\infty\}$ be an upper semi-continuous function, $\not\equiv -\infty$, uniformly bounded from above.

Proposition 5.1 (Ekeland). *Let $\epsilon > 0$ and $y_0 \in \mathsf{Y}$ be such that*

$$F(y_0) \geq \sup_{y \in \mathsf{Y}} F(y) - \epsilon.$$

Then there exist $y_\epsilon \in \mathsf{Y}$ such that

$$F(y_\epsilon) \geq F(y_0), \quad \mathsf{d}_Y(y_\epsilon, y_0) \leq 1, \quad F(y) < F(y_\epsilon) + \epsilon \mathsf{d}_Y(y, y_\epsilon), \ \forall y \neq y_\epsilon.$$

In particular, y_ϵ is the global strict (hence unique) maximum of $y \mapsto F(y) - \epsilon \mathsf{d}_Y(y, y_\epsilon)$.

Proof. The Proposition is an adaptation of Theorem 1 in Ekeland [13]. □

Proposition 5.2 (Borwein-Preiss). *Let $\epsilon > 0$ and $y_0 \in \mathsf{Y}$ be such that*

$$F(y_0) \geq \sup_{y \in \mathsf{Y}} F(y) - \epsilon.$$

Then there exist $y_k \in \mathsf{Y}$, $y_\epsilon \in \mathsf{Y}$ and non-negative numbers β_k with $\sum_{k=1}^\infty \beta_k = 1$ such that

$$\lim_{k \to \infty} \mathsf{d}_Y(y_k, y_\epsilon) = 0, \qquad \sup_{k=1,2,\ldots} \mathsf{d}_Y(y_k, y_\epsilon) \leq \epsilon^{1/4}, \qquad \mathsf{d}_Y(y_\epsilon, y_0) \leq \epsilon^{1/4},$$

$$F(y_\epsilon) \geq \sup_{y \in \mathsf{Y}} F(y) - \epsilon, \quad F(y_\epsilon) - \sqrt{\epsilon}\Delta(y_\epsilon) \geq F(y) - \sqrt{\epsilon}\Delta(y) \ \ \forall y \in \mathsf{Y},$$

where

$$\Delta(y) := \Delta_{\{\beta_k, y_k\}}(y) := \sum_{k=1}^\infty \beta_k \mathsf{d}_Y^2(y, y_k).$$

Proof. The conclusion is an adaptation of Theorem 2.6, estimates (2.8), (2.13) and (2.14) in its proof, and Remark 2.7, all of Borwein-Preiss [6], in the special case where $p = 2$, $\lambda = \epsilon^{1/4}$. □

In the above result, we have

$$|\Delta(y_\epsilon)| \leq \sum_k \beta_k \mathsf{d}_Y^2(y_\epsilon, y_k) \leq \sqrt{\epsilon}.$$

In particular, we also have

$$|D\Delta|(y_\epsilon) \leq 2 \sum_k \beta_k \mathsf{d}_Y(y_\epsilon, y_k) \leq 2\epsilon^{1/4}.$$

References

1. L. Ambrosio, and W. Gangbo. Hamiltonian ODEs in the Wasserstein space of Probability Measures *Commun. Pure. Appl. Math.* **LXI** (2008), 18-53
2. L. Ambrosio, N. Gigli and G. Savaré. *Gradient flows.* Second edition, Lectures in Mathematics ETH Zürich, Birkhäuser, Verlag, Basel, 2008.
3. L. Ambrosio, N. Gigli and G. Savaré. Heat flow and calculus on metric measure spaces with Ricci curvature bounded below - the compact case. *Bollettino U.M.I.* (9), **V** (2012) 575-629.

4. 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), no. 3-4, 635-675. Large deviations for the boundary driven symmetric simple exclusion process. *Math. Phys. Anal. Geom.***6** (2003), no. 3, 231-267. Minimum dissipation principle in stationary non-equilibrium states, *J. Statist. Phys.* **116** (2004), no. 1-4, 831-841. Stochastic interacting particle systems out of equilibrium, *J. Stat. Mech. Theory Exp.* **7** (2007), P07014, 35. Action functional and quasi-potential for the Burgers equation in a bounded interval. *Comm. Pure Appl. Math.***64** (2011), no. 5, 649-696.
5. L. Bertini, D. Gabrielli, and J. L. Lebowitz. Large deviations for a stochastic model of heat flow. *J. Stat. Phys.* **121** (2005), no. 56, 843 -885.
6. 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), no. 2, 517-527.
7. P. Cardialaguet. *Notes on Mean-Field Games (from P.L.Lions' lectures at College de France).* Jan (2012), http://www.ceremade.dauphine.fr/∼cardalia
8. P. Cardialaguet and M. Quincampoix. Deterministic differential games under probability knowledge of initial condition. *Int. Game Theory Rev.* **10** (2008), no. 1, 116.
9. P. Cardaliaguet. Information issues in differential game theory. *ESAIM: Proceedings* **35.** March (2012) 1-13
10. M. G. Crandall, and P.L. Lions. Hamilton-Jacobi Equations in Infinite Dimensions, PART I *J. Funct. Anal.* **62** (1985) 3. 379-396. PART II *J. Funct. Anal.* **65** (1986) 3. 368-405. PART III *J. Funct. Anal.* **68** (1986) 2. 214-247. PART IV *J. Funct. Anal.* **90** (1990) 2. 237-283. PART V *J. Funct. Anal.* **97** (1991) 2. 417-465. PART VII *J. Funct. Anal.* **125** (1994) 1. 111-148.
11. M. G. Crandall, and P.L. Lions. Hamilton-Jacobi Equations in Infinite Dimensions, PART VI: Nonlinear A and Tataru's Method Refined. *Evolution equations, control theory, and biomathematics (Han sur Lesse, 1991), 51–89, Lecture Notes in Pure and Appl. Math.* **155** Dekker, New York, 1994.
12. 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.* **27**, 1–67, 1992.
13. I. Ekeland. Nonconvex minimization problems. *Bull. Amer. Math. Soc. (N. S.)***1**, 443-474. (1979)
14. 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), no. 17-18, 973-976.
15. J. Feng, and M. Katsoulakis. A comparison principle for Hamilton-Jacobi equations related to controlled gradient flows in infinite dimensions. *Arch. Rational Mech. Anal.* **192** 275-310. (2009)
16. J. Feng, and T. G. Kurtz. *Large deviation for stochastic processes* Mathematical Survey and Monographs, American Mathematical Society, Vol. **131**, Rhode Island, 2006.
17. J. Feng, T. Nguyen. Hamilton-Jacobi equations in space of measures associated with a system of conservation laws. *J. Math. Pures Appl.* (9) **97** (2012), no. 4, 318-390
18. J. Feng, A. Swiech with Appendix B by A. Stefanov. Optimal control for a mixed flow of Hamiltonian and gradient type in space of probability measures. *Trans. A.M.S.* **365** (2013) 3987-4039.
19. W. Gangbo, T. Nguyen, and A. Tudorascu. Hamilton-Jacobi equations in the Wasserstein space. *Methods and Applications of Analysis* **15** (2008), no. 2, 155-184.
20. W. Gangbo, T. Nguyen, and A. Tudorascu. Euler-Poisson systems as action-minimizing paths in the Wasserstein space *Arch. Ration. Mech. Anal.* **192** (2009), no. 3, 419-452.
21. W. Gangbo and A. Tudorascu. Homogenization for a class of integral functionals in spaces of probability measures. *Advances in Mathematics* **230**, no. 3 (2012), 1124-1174.
22. W. Gangbo, and A. Swiech. Optimal transport and large number of particles. *Discrete and Continuous Dynamical Systems* **34**, no. 4 (2014), 1397-1441.
23. W. Gangbo, and M. Westdickenberg. Optimal transport for the system of isentropic Euler equations *Commun. P.D.E.* **34** (2009), 1041-1073.
24. Y. Giga, N. Hamamuki, and A. Nakayasu. Eikonal equations in metric spaces, *Tran. A.M.S.* Accepted.
25. N. Gigli. On the geometry of the space of probability measures endowed with the quadratic optimal transportation distance *Ph.D. Thesis, Scuola Normale Superiore, Pisa, Italy* 2008. http://cvgmt.sns.it/paper/491/
26. N. Gozlan, C. Roberto, and P. Samson, Hamilton-Jacobi equations on metric spaces and transport entropy inequalities. Preprint 2011, to appear on Revista Matematica Iberoamericana.

27. R. Hynd, H.K. Kim, Work in progress.
28. A.Mondino, A new notion of angle between three points in a metric space. *Preprint, http://cvgmt.sns.it*, 2013.
29. A. Nakayasu. On metric viscosity solutions for Hamilton-Jacobi equations of evolution type *In preparation*
30. F. Otto. The geometry of dissipative evolution equations: The porous medium equation. *Commun. P.D.E.s* **26**, Issue 1-2, 101-174. (2001)
31. C. Villani. *Optimal transport.* Springer, Berlin. (2009)

Classe di Scienze, Scuola Normale Superiore, Pisa, Italy, Department of Mathematics, University of Kansas, Lawrence, Kansas 66045, U.S.A.