

Viability and Invariance in Wasserstein Space $\mathcal{P}_2(\mathbb{R}^d)$

Hélène Frankowska

CNRS and Sorbonne Université

Dynamical Systems and Control Theory program review

September 20-22, 2021

Thanks to AFOSR (and Frederick) for the support

Outline of the talk

1 Wasserstein space

- Solutions to Controlled Continuity Equation
- Some Geometric Concepts in Wasserstein Space

2 Viability and Invariance

3 Hamilton-Jacobi-Bellman equation

- Hadamard Sub and Superdifferentials
- Optimal Control Problem
- Viscosity Solutions of HJB Equation

4 Maximum Principle and Sensitivity Relations

- Badreddine Z. & Frankowska H. (under revision) *Solutions to Hamilton-Jacobi Equation on a Wasserstein Space*
- Badreddine Z. & Frankowska H. (submitted) *Viability and Invariance of Systems on Metric Spaces*

Setting

Objective : to develop control theory in **metric spaces**. Today I will illustrate some of achievements on a Wasserstein space.

Below U denotes a fixed **compact metric space** and any Lebesgue measurable $u : [0, +\infty) \rightarrow U$ is called a control.

\mathcal{U} denotes the set of all controls.

$\mathcal{P}(\mathbb{R}^d)$ - the set of Borel probability measures on \mathbb{R}^d

The **second momentum** of $\mu \in \mathcal{P}(\mathbb{R}^d)$:

$$\mathcal{M}_2(\mu) := \int_{\mathbb{R}^d} |x|^2 d\mu(x)$$

$\mathcal{P}_2(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ - measures with finite second momentum

$\mathcal{P}_c(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ - measures with compact support

$\mathcal{P}(K) \subset \mathcal{P}(\mathbb{R}^d)$ - measures with the support contained in $K \subset \mathbb{R}^d$

$L^2(\mu; \mathbb{R}^d)$ - the space of $\psi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ with $\int_{\mathbb{R}^d} |\psi(x)|^2 d\mu(x) < \infty$

Transport Plans

For $\mu \in \mathcal{P}(\mathbb{R}^d)$, its **pushforward** $T_{\#}\mu$ through a Borel map $T : \mathbb{R}^d \rightarrow \mathbb{R}^m$ is: $T_{\#}\mu(B) := \mu(T^{-1}(B))$ for any Borel $B \subset \mathbb{R}^m$

Let $\pi^1, \pi^2 : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ be the projection operations onto the first and second variable. A measure $\gamma \in \mathcal{P}(\mathbb{R}^{2d})$ is a **transport plan** between $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ if $\pi_{\#}^1 \gamma = \mu$ and $\pi_{\#}^2 \gamma = \nu$. The set of all transport plans is denoted by $\Gamma(\mu, \nu)$.

The **Wasserstein distance** $W_2(\mu, \nu)$ between $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ is

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

The set of all **optimal transport plans** is denoted by $\Gamma_0(\mu, \nu)$.

Wasserstein spaces are used in models of **collective dynamics**.

Solutions to Continuity Equation

An absolutely continuous curve $\mu(\cdot) \in AC([0, 1], \mathcal{P}_2(\mathbb{R}^d))$ solves a **continuity equation** driven by a velocity field $(t, x) \mapsto F(t, x) \in \mathbb{R}^d$ with the initial condition $\mu_0 \in \mathcal{P}_c(\mathbb{R}^d)$ if

$$\partial_t \mu(t) + \operatorname{div}(F(t)\mu(t)) = 0 \quad \mu(0) = \mu_0,$$

in the sense of **distributions**, i.e. for any $\phi \in C_c^\infty((0, 1) \times \mathbb{R}^d; \mathbb{R})$

$$\int_0^1 \int_{\mathbb{R}^d} (\partial_t \phi(t, x) + \langle \nabla_x \phi(t, x), F(t, x) \rangle) d\mu(t)(x) dt = 0$$

$Lip(\mathbb{R}^d, \mathbb{R}^d)$ - space of bounded Lipschitz functions from \mathbb{R}^d into \mathbb{R}^d with the topology of local uniform convergence.

Solutions to Non-Local Continuity Equation

Let $f : \mathcal{P}_2(\mathbb{R}^d) \times U \rightarrow Lip(\mathbb{R}^d, \mathbb{R}^d)$ be continuous. Consider a **non-local controlled continuity equation** with $u(\cdot) \in \mathcal{U}$

$$\partial_t \mu(t) + \operatorname{div}(f(\mu(t), u(t))\mu(t)) = 0 \quad \mu(0) = \mu_0$$

We assume that $\exists \rho \geq 0, A \geq 0, k \geq 0$ s.t. $\forall \mu \in \mathcal{P}_2(\mathbb{R}^d), u \in U$:
 $Lip f(\mu, u) \leq A, \|f(\mu, u)\|_\infty \leq \rho, f(\cdot, u)$ is k -Lipschitz.

Theorem (Benoît Bonnet, HF; 2021 JDE)

$\forall r > 0, \mu_0 \in \mathcal{P}(B(0, r)), u(\cdot) \in \mathcal{U}$ there exists a unique $\mu(\cdot)$ solving the continuity equation with $F(t) = f(\mu(t), u(t))$.
 Furthermore $\operatorname{supp}(\mu(t)) \subset B(0, r + \rho t)$ for every $t \geq 0$.

Barycentric Projection and Proximal Normals

Let $\alpha \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$, $\mu = \pi_{\#}^1 \alpha$. Denote by $\bar{\alpha} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ the **barycentric projection** of α with respect to μ .

Let $K \subset \mathcal{P}_2(\mathbb{R}^d)$, $\mu \in K$ and $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ satisfy

$$W_2(\nu, \mu) = \text{dist}(\nu, K)$$

For any optimal transport plan $\alpha \in \Gamma_0(\mu, \nu)$, consider its **barycentric projection** $\bar{\alpha} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ with respect to μ . Any $p \in L^2(\mu; \mathbb{R}^d)$ satisfying

$$p(x) := \bar{\alpha}(x) - x, \text{ for } \mu - \text{a.e. } x \in \mathbb{R}^d,$$

is called a **proximal normal** to K at μ .

Tangents and Normals in the Wasserstein space

Example

Let $Q \subset \mathbb{R}^d$ be closed, $y_0 \notin Q$ and let $x_0 \in Q$ be such that $|y_0 - x_0| = \text{dist}(y_0, Q)$. Consider $K = \{\delta_q : q \in Q\} \subset \mathcal{P}_2(\mathbb{R}^d)$. Then $p(\cdot) = y_0 - x_0$ δ_{x_0} - a.e. and we recovered the classical notion of proximal normal to Q at x_0 .

Denote by $N_K^{pr}(\mu)$ the set of all proximal normals to K at μ .

Define the **contingent cone** to K at $\mu \in K$:

$$\mathring{T}_K(\mu) := \left\{ F \in \text{Lip}(\mathbb{R}^d, \mathbb{R}^d) \mid \liminf_{h \rightarrow 0+} \frac{1}{h} \text{dist}((Id + hF)_\# \mu, K) = 0 \right\}$$

and the **normal cone** to K at μ by

$$[\mathring{T}_K(\mu)]^- := \{p \in L^2(\mu; \mathbb{R}^d) \mid \int_{\mathbb{R}^d} \langle p(x), F(x) \rangle d\mu(x) \leq 0, \forall F \in \mathring{T}_K(\mu)\}$$

Viability and Invariance

Let $K \subset \mathcal{P}_2(\mathbb{R}^d)$ be **proper**.

From now on we also assume that $f(\mu, U)$ is convex for any μ .
Consider the non-local controlled continuity equation

$$(E) \quad \partial_t \mu(t) + \operatorname{div}(f(\mu(t), u(t))\mu(t)) = 0$$

K is **viable under (E)** if for every $\mu_0 \in K$ having a **compact support**, **there exists** $u(\cdot) \in \mathcal{U}$ such that the solution μ of (E) with $\mu(0) = \mu_0$ satisfies $\mu(t) \in K$ for all $t \geq 0$.

K is **invariant under (E)** if for every $\mu_0 \in K$ having a compact support and **for every** $u(\cdot) \in \mathcal{U}$ the solution of (E) with $\mu(0) = \mu_0$ satisfies $\mu(t) \in K$ for all $t \geq 0$.

Characterisations of Viability and Invariance

Proposition (necessary conditions for viability and invariance)

If K is viable under (E) , then $f(\mu, U) \cap \overset{\circ}{T}_K(\mu) \neq \emptyset$ for any $\mu \in K$ having a compact support.

If K is invariant under (E) , then $f(\mu, U) \subset \overset{\circ}{T}_K(\mu)$ for any $\mu \in K$ having a compact support.

Theorem (sufficient conditions for viability and invariance)

If $f(\mu, U) \cap \overset{\circ}{T}_K(\mu) \neq \emptyset$ for each $\mu \in K$, then K is viable under (E) .

If $f(\mu, U) \subset \overset{\circ}{T}_K(\mu)$ for each $\mu \in K$, then K is invariant under (E) .

Viability and Invariance are used to investigate uniqueness of solutions to **Hamilton-Jacobi-Bellman equation** and also to study **stability** of control systems.

Equivalent Conditions for Viability

The following statements are **equivalent**:

- for any $\mu \in K$, $f(\mu, U) \cap \overset{\circ}{T}_K(\mu) \neq \emptyset$;
- for any $\mu \in K$, and for every $p \in [\overset{\circ}{T}_K(\mu)]^-$ we have

$$\min_{u \in U} \int_{\mathbb{R}^d} \langle p(x), f(\mu, u)(x) \rangle d\mu(x) \leq 0$$
;
- for any $\mu \in K$, and for **every** $p \in N_K^{pr}(\mu)$ we have

$$\min_{u \in U} \int_{\mathbb{R}^d} \langle p(x), f(\mu, u)(x) \rangle d\mu(x) \leq 0$$
;
- for any $\mu \in K$ and $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ satisfying
 $W_2(\nu, \mu) = \text{dist}(\nu, K)$, **there exists** a proximal normal p
 associated to ν, μ such that

$$\min_{u \in U} \int_{\mathbb{R}^d} \langle p(x), f(\mu, u)(x) \rangle d\mu(x) \leq 0.$$

A **similar result** holds also for the relation $f(\mu, U) \subset \overset{\circ}{T}_K(\mu)$
 with min replaced by max

Hadamard Subdifferentials

Consider $w : [0, 1] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and $(t, \mu) \in \text{dom}(w)$.

We say that $(p_t, p_\mu) \in \mathbb{R} \times L^2(\mu; \mathbb{R}^d)$ belongs to the

Hadamard subdifferential $\partial^- w(t, \mu)$ of w at (t, μ) if

$\forall F \in Lip(\mathbb{R}^d, \mathbb{R}^d)$ and any $\kappa \in \mathbb{R}$,

$$p_t \kappa + \int_{\mathbb{R}^d} \langle p_\mu(x), F(x) \rangle d\mu(x) \leq D_\uparrow w(t, \mu)(\kappa, F),$$

Hadamard superdifferential $\partial^+ w(t, \mu) := -\partial^-(-w)(t, \mu)$,

where for $\kappa \in \mathbb{R}$, $F \in Lip(\mathbb{R}^d, \mathbb{R}^d)$ with $(\kappa, F) \in \overset{\circ}{T}_{\text{dom}(w)}(t, \mu)$

$$D_\uparrow w(t, \mu)(\kappa, F) := \liminf_{\substack{h \rightarrow 0+, \kappa' \rightarrow \kappa \\ W_2(\mu', (Id + hF)_\# \mu) = o(h) \\ (t + h\kappa', \mu') \in \text{dom}(w)}} \frac{w(t + h\kappa', \mu') - w(t, \mu)}{h}$$

If $(\kappa, F) \notin \overset{\circ}{T}_{\text{dom}(w)}(t, \mu)$, then set $D_\uparrow w(t, \mu)(\kappa, F) = +\infty$.

Relation of Normals and Subdifferentials

Let $\mathcal{E}p(w)$ denote the epigraph of w .

Proposition

For any $0 \neq q \in \mathbb{R}$ the following statements are equivalent:

- $(p_t, p_\mu, q) \in [\overset{\circ}{T}_{\mathcal{E}p(w)}(t, \mu, w(t, \mu))]^-;$
- $(\frac{p_t}{|q|}, \frac{p_\mu}{|q|}) \in \partial^- w(t, \mu).$

The main difficulty in application of this result to Hamilton-Jacobi-Bellman equations comes from **horizontal normals**, i.e. when

$$[\overset{\circ}{T}_{\mathcal{E}p(w)}(t, \mu, w(t, \mu))]^- \subset \mathbb{R} \times L^2(\mu; \mathbb{R}^d) \times \{0\}$$

Optimal Control Problem

Given the **cost function** $g : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$ and the **initial condition** $\mu_0 \in \mathcal{P}_c(\mathbb{R}^d)$, consider the Mayer problem

$$(P) \quad \text{minimize } g(\mu(1))$$

over all the controls $u(\cdot) \in \mathcal{U}$ and corresponding solutions of (E) defined on $[0, 1]$ with $\mu(0) = \mu_0$.

The associated **value function** $\tilde{V} : [0, 1] \times \mathcal{P}_c(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\pm\infty\}$:

$$\tilde{V}(t, \nu_0) = \inf \left\{ g(\mu(1)) \mid \mu(\cdot) : [t, 1] \rightarrow \mathcal{P}_2(\mathbb{R}^d) \text{ is a solution of (E)} \right. \\ \left. \text{on } [t, 1] \text{ for some } u(\cdot) \in \mathcal{U}, \mu(t) = \nu_0 \right\}.$$

Hamilton-Jacobi-Bellman equation

Define the **Hamiltonian** $\mathcal{H} : \mathcal{P}_2(\mathbb{R}^d) \times L^2(\mu; \mathbb{R}^d) \rightarrow \mathbb{R}$ by

$$\mathcal{H}(\mu, p) = \max_{u \in U} \int_{\mathbb{R}^d} \langle p(x), f(\mu, u)(x) \rangle d\mu(x)$$

Hamilton-Jacobi-Bellman equation associated to (P)

$$[HJB] \quad -\partial_t w(t, \mu) + \mathcal{H}(\mu, -\nabla_\mu w(t, \mu)) = 0$$

For any $r > 0$, consider the compact metric space

$$\Delta_r := \{(t, \nu) \in [0, 1] \times \mathcal{P}_2(\mathbb{R}^d) : \text{supp}(\nu) \subset B(0, r + \rho t)\}.$$

Consider the **restricted value function** $V : \Delta_r \rightarrow \mathbb{R} \cup \{\pm\infty\}$ defined by $V(t, \nu) = \tilde{V}(t, \nu)$ for $(t, \nu) \in \Delta_r$. We do not use the notation V_r because for all $r < R$ it holds $V_R = V_r$ on Δ_r .

Viscosity Solutions

A continuous map $w : \Delta_r \rightarrow \mathbb{R}$ is called a **viscosity solution** to $[HJB]$ on Δ_r if $\forall (t, \mu) \in \Delta_r$ with $t < 1$

$$-p_t + \mathcal{H}(\mu, -p_\mu) \geq 0 \quad \forall (p_t, p_\mu) \in \partial^- w(t, \mu)$$

$$-p_t + \mathcal{H}(\mu, -p_\mu) \leq 0 \quad \forall (p_t, p_\mu) \in \partial^+ w(t, \mu)$$

$w : [0, 1] \times \mathcal{P}_c(\mathbb{R}^d) \rightarrow \mathbb{R}$ is called a viscosity solution to $[HJB]$ on $[0, 1] \times \mathcal{P}_c(\mathbb{R}^d)$ if it is a viscosity solution on every Δ_r , $r > 0$.

Theorem

Let $r > 0$. If the restriction of g to $\mathcal{P}_2(B(0, r + \rho))$ is continuous, then V is a viscosity solution to $[HJB]$ on Δ_r and $V(1, \cdot) = g(\cdot)$ on $\mathcal{P}(B(0, r + \rho))$.

Uniqueness of Viscosity Solution

Theorem

Let $r > 0$. If $w : \Delta_r \rightarrow \mathbb{R}$ is a viscosity solution to [HJB] on Δ_r with $w(1, \cdot) = g(\cdot)$ on $\mathcal{P}(B(0, r + \rho))$, then $w = V$ on Δ_r .

Proofs are based on

1. Rewriting the definition of a solution w of [HJB] via **viability and invariance like** properties of the epigraph and hypograph of w for the map $\mu \mapsto \{1\} \times f(\mu, U) \times \{0\}$.
2. Applying the viability and invariance theorems to

$$\begin{cases} t' = 1, & t(t_0) = t_0, \\ \partial_t \mu(t) + \operatorname{div}(f(\mu(t), u(t))\mu(t)) = 0, & \mu(t_0) = \mu_0, \quad u(\cdot) \in \mathcal{U}, \\ z' = 0, & z(t_0) = w(t_0, \mu_0). \end{cases}$$

to deduce that w is **nondecreasing along trajectories** of controlled continuity equation and is **constant along at least one** such trajectory.

Maximum Principle

Assume $g(\cdot)$, $f(\cdot, u, \cdot)$ are **locally continuously differentiable**.

Define the **Hamiltonian** and the symplectic matrix of \mathbb{R}^{2d}

$$H(\zeta, u) = \int_{\mathbb{R}^{2d}} \langle r, f(\pi_{\#}^1 \zeta, u, x) \rangle d\zeta(x, r), \quad J_{2d} = \begin{pmatrix} 0 & \text{Id} \\ -\text{Id} & 0 \end{pmatrix}$$

for any $\zeta \in \mathcal{P}_c(\mathbb{R}^{2d})$, $u \in U$.

Theorem (Benoît Bonnet, HF; 2021 AMO + submitted)

Let (μ^*, u^*) be **optimal** for (P) . Then the unique solution ν^* of

$$\begin{cases} \partial_t \nu^*(t) + \text{div}(J_{2d} H_{\nu}(\nu^*(t), u^*(t)) \nu(t)) = 0, & \pi_{\#}^1 \nu^*(t) = \mu^*(t) \\ \nu^*(1) = (\text{Id}, -\nabla g(\mu^*(1))_{\#} \mu^*(1)) \end{cases}$$

satisfies for a.e. $t \in [0, 1]$ the **maximality condition**

$$H(\nu^*(t), u^*(t)) = \max_{u \in U} H(\nu^*(t), u)$$

Sensitivity Relation. BB & HF; 2021 CDC IEEE

Theorem (necessary optimality condition)

Let $(\mu^(\cdot), u^*(\cdot))$ be optimal and $\nu^*(\cdot)$ be the corresponding state-costate curve. Then,*

$$(H(\nu^*(t), u^*(t)), -\bar{\nu}^*(t)) \in \partial^+ \tilde{V}(t, \mu^*(t))$$

for almost every $t \in [0, 1]$, where $\bar{\nu}^(t) \in L^2(\mathbb{R}^d; \mu^*(t))$ denotes the barycentric projection of $\nu^*(t)$ onto $\mu^*(t)$.*

Theorem (sufficient optimality condition)

Let $(\mu^(\cdot), u^*(\cdot))$ be a trajectory-control pair and $\nu^*(\cdot)$ be a state-costate curve satisfying the PMP and the sensitivity relation. Then $(\mu^*(\cdot), u^*(\cdot))$ is optimal for (P) .*

*Thank you for
your attention
!!!*