HAMILTONIAN SYSTEMS AND THE CALCULUS OF DIFFERENTIAL FORMS ON THE WASSERSTEIN SPACE

A Thesis
Presented to
The Academic Faculty

by

Hwa Kil Kim

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology
August 2009

HAMILTONIAN SYSTEMS AND THE CALCULUS OF DIFFERENTIAL FORMS ON THE WASSERSTEIN SPACE

Approved by:

Professor Wilfrid Gangbo, Advisor
School of Mathematics
Georgia Institute of Technology

Professor Michael Loss
School of Mathematics
Georgia Institute of Technology

Professor Ronghua Pan
School of Mathematics
Georgia Institute of Technology

Professor Andrzej Swiech
School of Mathematics
Georgia Institute of Technology

Professor Allen Tannenbaum
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: 8 April 2009

To my wife and my daughter.

ACKNOWLEDGEMENTS

First, I want to thank my advisor Dr. Wilfrid Gangbo. He brought me into the field of Optimal Transportation which is flourishing and has many applications to other areas in mathematics. During the years of my PhD, he was always patient and supporting. I learned a lot from him not only as a mathematician but also as a human being. Without him, it wouldn't be possible for me to finish PhD degree successfully.

I also want to thank to all the committee members, Dr Michael Loss, Ronghua Pan, Andrzej Swiech and Allen Tannenbaum for their time on reviewing my thesis and mathematics which I learned from them. Especially, I can't forget about many fruitful suggestions on my thesis from Dr. Swiech. What stimulated my interest in the area of PDE was also his PDE class which I took at the first year of PhD degree.

I have spent not small portion of my life in Atlanta. There are many people who made the time as a valuable memory which I can never forget. First, I want to thank to friends from the school of mathematics at GaTech. We came from different countries and have different culture and life style but we have same warm heart. Because of you, I was able to start foreign life without any difficulty. Second, brothers and sisters from Yedam church and SFC, I know it was God who led me to you. God used you guys to make me know him and to make me hear his voice. We met every week and tried to know God more. It was precious time and I will miss you and the time we had together.

Finally I would like to thank to my family. I thank to my parents for their support and sacrifice from the beginning of my life until now. They have given me all they have and still they are trying to give me anything left. My brothers and sisters, through my life, provided me with a fence which is called Family. Inside of this fence, I was in peace no matter how strong thunder storms were outside of it. My wife, Hye Jeong and my daughter, Yeju, you are gifts from God. You may not know how much I love you. You also may not know how much you make me happy.

Above all, I thank my Lord Jesus Christ for his love and mercy. To him, I am just a humble servant.

TABLE OF CONTENTS

DEDICATION iii

ACKNOWLEDGEMENTS iv

SUMMARY viii

I PRELIMINARY 1

1.1 Notations 1

1.2 Wasserstein space 2

II HAMILTONIAN SYSTEMS 10

2.1 Moreau-Yosida approximation 11

2.2 Variational problem in connection with the Moreau-Yosida approximation 16

2.2.1 A special class of Hamiltonian 16

2.2.2 Properties of minimizers 19

2.3 Hamiltonian flow 22

2.3.1 Stability of Hamiltonian flows 31

2.3.2 Examples 32

2.4 Uniqueness of Hamiltonian flows 33

III CALCULUS OF DIFFERENTIAL FORMS ON THE WASSERSTEIN SPACE 41

3.1 Tangent and Cotangent bundles 41

3.1.1 Tangent space 41

3.1.2 Differential forms on $\mathcal{M}$ 44

3.2 Calculus of pseudo differential 1-forms 48

3.2.1 Green's formula for smooth surfaces and 1-forms 49

3.2.2 Regularity and differentiability of pseudo 1-forms 51

3.2.3 Further continuity and differentiability properties of regular forms 56

3.2.4 Mollification of absolutely continuous paths in $\mathcal{M}$ 62

3.2.5 Integration of regular pseudo 1-forms 66

3.2.6 Green's formula for annuli, the first cohomology of regular pseudo 1-forms 68

3.2.7 Example: Restriction of 1-forms to the space of discrete measures 75

REFERENCES 77

SUMMARY

This thesis consists of two parts. In the first part, we study stability properties of Hamiltonian systems on the Wasserstein space. Let H be a Hamiltonian satisfying conditions imposed in [2]. We regularize H via Moreau-Yosida approximation to get H_τ and denote by μ_τ a solution of system with the new Hamiltonian H_τ. Suppose H_τ converges to H as τ tends to zero. We show μ_τ converges to μ and μ is a solution of a Hamiltonian system which is corresponding to the Hamiltonian H. At the end of first part, we give a sufficient condition for the uniqueness of Hamiltonian systems.

In the second part, we develop a general theory of differential forms on the Wasserstein space. Our main result is to prove an analogue of Green's theorem for 1-forms and show that every closed 1-form on the Wasserstein space is exact. If the Wasserstein space were a manifold in the classical sense, this result wouldn't be worthy of mention. Hence, the first cohomology group, in the sense of de Rham, vanishes.

CHAPTER I

PRELIMINARY

In this preliminary chapter, I will fix notations and introduce terminology and some already known facts which I will use later(cfr.[3])

1.1 Notations

- $\mathcal{P}(\mathbb{R}^D) = \{\mu | \mu$ is a *Borel probability measure* on $\mathbb{R}^D\}$

- Let $\mathcal{M}$ be the subspace of $\mathcal{P}(\mathbb{R}^D)$ with bounded second moment, *i.e.*

$$\mathcal{M} := \big\{\mu \in \mathcal{P}(\mathbb{R}^D) : \mu \geq 0, \int_{\mathbb{R}^D} d\mu = 1, \int_{\mathbb{R}^D} |x|^2 \, d\mu < \infty\big\}.$$

- Let $\mu \in \mathcal{P}(\mathbb{R}^D)$ and let $f : \mathbb{R}^D \to \mathbb{R}^k$ be a *Borel map*. Then $\nu := f_{\#}\mu$ is a *Borel measure* on $\mathbb{R}^k$ characterized by $\nu[B] = \mu[f^{-1}(B)]$ for all *Borel sets* $B \subset \mathbb{R}^k$. In this case, we say f pushes μ forward to ν.

- $C_c^\infty(\mathbb{R}^D)$ is the collection of all infinitely differentiable functions with compact support.

- We denote $C_b(\mathbb{R}^D)$ the collection of all continuous and bounded functions.

- $C_o(\mathbb{R}^D) := \{f | f : \mathbb{R}^D \to \mathbb{R}\}$ where f is continuous and $f(x) \to 0$ as $x \to \infty$.

- $Id : \mathbb{R}^D \to \mathbb{R}^D$ is the identity map i.e $Id(x) = x$ for all $x \in \mathbb{R}^D$.

- $\mathrm{Diff}_c(\mathbb{R}^D)$ denotes the set of diffeomorphisms of $\mathbb{R}^D$ with compact support, *i.e.* those which coincide with the identity map Id outside of a compact subset of $\mathbb{R}^D$.

- Let $\mathcal{X}_c$ denote the space of compactly supported smooth vector fields on $\mathbb{R}^D$.

- Let $\mu \in \mathcal{P}(\mathbb{R}^D)$ and let $f : \mathbb{R}^D \to \mathbb{R}^k$ be $f \in L^2(\mu)$. We denote the L^2 norm of f by $||f||_\mu$ i.e

$$||f||_\mu^2 := ||f||_{L^2(\mu)}^2 = \int_{\mathbb{R}^D} |f(x)|^2 d\mu(x)$$

- Let $\mu \in \mathcal{P}(\mathbb{R}^D)$, we define $\overline{\nabla C_c^\infty}^\mu$ as the closure of $\{\nabla\phi : \phi \in C_c^\infty(\mathbb{R}^D)\}$ in the $L^2(\mu)$ topology.
- Let $\mu \in \mathcal{P}(\mathbb{R}^D)$, we denote the support of μ by $supp(\mu)$.
- Let $r > 0$ and $x \in \mathbb{R}^D$ then $B_r(x)$ denotes the ball in $\mathbb{R}^D$ of center x and radius r.
- Let $f : \mathbb{R}^D \to \mathbb{R}^k$ be a Lipschitz function then we denote the Lipschitz constant by $\mathrm{Lip}(f)$.

1.2 Wasserstein space

Recall that $\mathcal{M}$ is the subspace of $\mathcal{P}(\mathbb{R}^D)$ with bounded second moment.

Definition 1.2.1. Let μ, $\nu \in \mathcal{M}$. Consider

$$W_2(\mu,\nu) := \big(\inf_{\gamma\in\Gamma(\mu,\nu)} \int_{\mathbb{R}^D\times\mathbb{R}^D} |x-y|^2 d\gamma(x,y)\big)^{1/2}. \tag{1}$$

Here, $\Gamma(\mu,\nu)$ denotes the set of Borel measures γ on $\mathbb{R}^D \times \mathbb{R}^D$ which have μ and ν as *marginals*, *i.e.* satisfying $\pi^1_\#(\gamma) = \mu$ and $\pi^2_\#(\gamma) = \nu$ where π^1 and π^2 denote the standard projections $\pi^1, \pi^2 : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}^D$ defined by $\pi^1(x_1, x_2) = x_1$ and $\pi^2(x_1, x_2) = x_2$.

Equation (1) defines a distance on $\mathcal{M}$ which is called Wasserstein metric. It is known that the infimum in the right hand side of equation (1) is always achieved. We will denote by $\Gamma_o(\mu,\nu)$ the set of γ which minimize this expression. We also denote

$$\Gamma_o(\gamma,\mu) = \{\mathbf{u} \in \mathcal{P}(\mathbb{R}^{3D}) : \pi^{1,2}_\#\mathbf{u} = \gamma, \quad \pi^{1,3}_\#\mathbf{u} \in \Gamma_o(\pi^1_\#\gamma,\mu)\} \tag{2}$$

where $\pi^{1,2}$, $\pi^{1,3} : \mathbb{R}^D \times \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}^D \times \mathbb{R}^D$ are defined by $\pi^{1,2}(x_1, x_2, x_3) = (x_1, x_2)$ and $\pi^{1,3}(x_1, x_2, x_3) = (x_1, x_3)$.

Remark 1.2.2. In the equation (2), $\Gamma_o(\gamma,\mu)$ has nothing to do with any cost functions. It is a conventional notation.

Recall that μ is *absolutely continuous* with respect to Lebesgue measure $\mathcal{L}^D$, written $\mu << \mathcal{L}^D$, if it is of the form $\mu = \rho(x)\,\mathcal{L}^D$ for some function $\rho \in L^1(\mathbb{R}^D)$. In this case for any $\nu \in \mathcal{M}$ there exists a unique map $T : \mathbb{R}^D \to \mathbb{R}^D$ such that $T_\#\mu = \nu$ and

$$W_2^2(\mu,\nu) = \int_{\mathbb{R}^D} |x - T(x)|^2 d\mu(x), \tag{3}$$

cfr. *e.g.* [3] [5] or [15]. One refers to T as the *optimal map* that pushes μ forward to ν. It can be shown that $(\mathcal{M}, W_2)$ is a separable complete metric space, cfr. *e.g.* [3] Proposition 7.1.5. The following theorem is an important result from Monge-Kantorovich theory

Theorem 1.2.3. *[3] or [29] Let μ, $\nu \in \mathcal{M}$, we have*

$$W_2^2(\mu,\nu) = \sup_{u,v\in C(\mathbb{R}^D)} \Big\{ \int_{\mathbb{R}^D} u d\mu + \int_{\mathbb{R}^D} v d\nu : \ u(x) + v(y) \leq |x-y|^2 \ \forall x, y \in \mathbb{R}^D \Big\}. \tag{4}$$

Definition 1.2.4. [3] Let $\mu, \nu \in \mathcal{M}$ and let $\gamma \in \Gamma_o(\mu,\nu)$. The *barycentric projection* $\bar{\gamma}_\mu^\nu : \mathbb{R}^D \to \mathbb{R}^D$ of γ with respect to the first marginal μ is characterized by

$$\int \psi(x) \bar{\gamma}_\mu^\nu(x) d\mu(x) = \int \psi(x) y d\gamma(x,y) \quad \forall \psi \in C_b \tag{5}$$

where C_b is a collection of all continuous and bounded functions. Similarly, the *barycentric projection* $\bar{\gamma}_\nu^\mu : \mathbb{R}^D \to \mathbb{R}^D$ of γ with respect to the second marginal ν is defined by

$$\int \psi(y) \bar{\gamma}_\nu^\mu(y) d\nu(y) = \int \psi(y) x d\gamma(x,y) \quad \forall \psi \in C_b \tag{6}$$

Theorem 1.2.5. *[3] If $\gamma \in \Gamma_o(\mu_0,\mu_1)$ then the curve $t \to \mu_t := ((1-t)\pi_1 + t\pi_2)_\#\gamma$ is a constant speed geodesic connecting μ_0 to μ_1. Conversely, any constant speed geodesic $\mu_t : [0,1] \to \mathcal{M}$ connecting μ_0 to μ_1 has this representation for a suitable $\gamma \in \Gamma_o(\mu_0,\mu_1)$.*

Definition 1.2.6. [24] Let $H : \mathcal{M} \to (-\infty, \infty)$ and $\lambda \in \mathbb{R}$. We say that H is λ-*convex* if for every $\mu_0, \mu_1 \in \mathcal{M}$ and every optimal transport plan $\gamma \in \Gamma_o(\mu_0,\mu_1)$ we have

$$H(\mu_t) \leq (1-t)H(\mu_0) + tH(\mu_1) - \frac{\lambda}{2} t(1-t) W_2^2(\mu_0,\mu_1) \quad \forall t \in [0,1] \tag{7}$$

Here $\mu_t = ((1-t)\pi_1 + t\pi_2)_\#\gamma$.

Remark 1.2.7. In [3], λ- *convexity* is defined in a weaker way.

Definition 1.2.8. Let $F : \mathcal{M} \to \mathbb{R}$ be a function on $\mathcal{M}$. We say that $\xi \in L^2(\mu)$ belongs to the *subdifferential* $\partial_- F(\mu)$ if

$$F(\nu) \geq F(\mu) + \sup_{\gamma \in \Gamma_o(\mu,\nu)} \iint_{\mathbb{R}^D \times \mathbb{R}^D} \langle \xi(x), y - x \rangle \, d\gamma(x,y) + o(W_2(\mu,\nu)),$$

as $\nu \to \mu$. If $-\xi \in \partial_-(-F)(\mu)$ we say that ξ belongs to the *superdifferential* $\partial^+ F(\mu)$.

If $\xi \in \partial_- F(\mu) \cap \partial^+ F(\mu)$ then, for any $\gamma \in \Gamma_o(\mu,\nu)$,

$$F(\nu) = F(\mu) + \iint_{\mathbb{R}^D \times \mathbb{R}^D} \langle \xi(x), y - x \rangle \, d\gamma(x,y) + o(W_2(\mu,\nu)). \tag{8}$$

If such ξ exists we say that F is *differentiable* at μ and we define the *gradient vector* $\nabla_\mu F := \pi_\mu(\xi)$ where $\pi_\mu(\xi)$ is the projection defined by the equation (116). Using barycentric projections one can show that, for $\gamma \in \Gamma_o(\mu,\nu)$,

$$\iint_{\mathbb{R}^D \times \mathbb{R}^D} \langle \xi(x), y - x \rangle \, d\gamma(x,y) = \iint_{\mathbb{R}^D \times \mathbb{R}^D} \langle \pi_\mu(\xi)(x), y - x \rangle \, d\gamma(x,y) \tag{9}$$

Thus $\pi_\mu(\xi) \in \partial_- F(\mu) \cap \partial^+ F(\mu) \cap T_\mu \mathcal{M}$ and it satisfies the analogue of equation (8). It can be shown that the gradient vector is unique, *i.e.* that $\partial_- F(\mu) \cap \partial^+ F(\mu) \cap T_\mu \mathcal{M} = \{\pi_\mu(\xi)\}$.

Remark 1.2.9. (i) Let $\gamma \in \Gamma_o(\mu,\nu)$ then

$$\begin{aligned} \iint_{\mathbb{R}^D \times \mathbb{R}^D} \langle \xi(x), y - x \rangle \, d\gamma(x,y) &= \int_{\mathbb{R}^D} \langle \xi(x), \bar{\gamma}_\mu^\nu(x) - x \rangle \, d\mu(x) \\ &= \int_{\mathbb{R}^D} \langle \pi_\mu(\xi)(x), \bar{\gamma}_\mu^\nu(x) - x \rangle \, d\mu(x) \\ &= \iint_{\mathbb{R}^D \times \mathbb{R}^D} \langle \pi_\mu(\xi)(x), y - x \rangle \, d\gamma(x,y) \end{aligned} \tag{10}$$

We used the fact $\bar{\gamma}_\mu^\nu - Id \in T_\mu \mathcal{M}$(cfr. [3]) in the second equality in equation (10).

(ii) Let $\phi \in C_c^\infty(\mathbb{R}^D)$ then $x \to |x|^2/2 + t\phi(x)$ is a convex function for small $t \in (0,1)$.

Define $v_t(x) := x + t\nabla\phi(x)$ and let $\nu_t = v_{t\#}\mu$ then v_t is an optimal map that pushes forward μ to ν_t i.e $(Id \times v_t)_{\#}\mu \in \Gamma_o(\mu, \nu_t)$ for all small t.

Suppose $\xi, \eta \in \partial_{-}F(\mu) \cap \partial^{+}F(\mu)$ then by equation (8), we have

$$F(\nu_t) = F(\mu) + t\int_{\mathbb{R}^D} \langle \xi(x), \nabla\phi(x)\rangle \, d\mu(x) + o(W_2(\mu, \nu_t)). \tag{11}$$

for all small $t \in (0, 1)$. Similarly, we have

$$F(\nu_t) = F(\mu) + t\int_{\mathbb{R}^D} \langle \eta(x), \nabla\phi(x)\rangle \, d\mu(x) + o(W_2(\mu, \nu_t)). \tag{12}$$

for all small $t \in (0, 1)$. We combine equations (11) and (12) to get

$$0 = \int_{\mathbb{R}^D} \langle \eta(x) - \xi(x), \nabla\phi(x)\rangle \, d\mu(x) = \int_{\mathbb{R}^D} \langle \pi_\mu(\eta)(x) - \pi_\mu(\xi)(x), \nabla\phi(x)\rangle \, d\mu(x). \tag{13}$$

Since ϕ is arbitrary, equation (13) gives $\pi_\mu(\eta) = \pi_\mu(\xi)$. This shows the uniqueness of the gradient vector.

Lemma 1.2.10. *Let $H : \mathcal{M} \to (-\infty, \infty)$ be lower semicontinuous and $\lambda-$convex for some $\lambda \in \mathbb{R}$ and let $\mu \in \mathcal{M}$. Then the following statements are equivalent:*

(i) $\xi \in \partial_{-}H(\mu)$

(ii) For all $\nu \in \mathcal{M}$

$$H(\nu) \geq H(\mu) + \inf_{\gamma \in \Gamma_o(\mu,\nu)} \int_{\mathbb{R}^D} < \xi(x), y - x > d\gamma(x, y) + o(W_2(\mu, \nu))$$

(iii) For all $\nu \in \mathcal{M}$

$$H(\nu) \geq H(\mu) + \sup_{\gamma \in \Gamma_o(\mu,\nu)} \int_{\mathbb{R}^D} < \xi(x), y - x > d\gamma(x, y) + \frac{\lambda}{2}W_2^2(\mu, \nu)$$

Proof. Proposition 4.2 of [2] ☐

Lemma 1.2.11. *[3] Let $\gamma^{12}, \gamma^{13} \in \mathcal{P}(\mathbb{R}^{2D})$ such that $\pi_{\#}\gamma^{12} = \pi_{\#}\gamma^{13} = \mu^1$. Then there exists $\mathbf{u} \in \mathcal{P}(\mathbb{R}^{3D})$ such that $\pi^{1,2}_{\#}\mathbf{u} = \gamma^{12}$, and $\pi^{1,3}_{\#}\mathbf{u} = \gamma^{13}$. Moreover, if $\mu^2 = \pi^2_{\#}\gamma^{12}, \mu^3 = \pi^2_{\#}\gamma^{13} \in \mathcal{M}$ and $\gamma^{12} \in \Gamma_0(\mu^1, \mu^2), \gamma^{13} \in \Gamma_0(\mu^1, \mu^3)$ then $\mathbf{u} \in \Gamma_0(\gamma^{12}, \mu^3)$.*

Proof. Let $\gamma^{12} = \int_{\mathbb{R}^D} \gamma^{12}_{x_1} d\mu^1, \gamma^{13} = \int_{\mathbb{R}^D} \gamma^{13}_{x_1} d\mu^1$ be the disintegration of γ^{12}, γ^{13} then the measure $\mathbf{u}$ whose disintegration w.r.t x_1 is

$$\int_{\mathbb{R}^D} \gamma^{12}_{x_1} \times \gamma^{13}_{x_1} d\mu^1(x_1)$$

has the required property. Let us denote this $\mathbf{u}$ by $\gamma^{12} \times \gamma^{13}$. $\square$

Definition 1.2.12. [3] Given $\mu \in \mathcal{M}$, let $T_\mu\mathcal{M}$ denote the closure of ∇C_c^∞ in $L^2(\mu)$. We call it the *tangent space* of $\mathcal{M}$ at μ. The *tangent bundle* $T\mathcal{M}$ is defined as the disjoint union of all $T_\mu\mathcal{M}$.

Following [3] we now provide an analytic justification for the above definition of tangent spaces for $\mathcal{M}$.

Suppose we are given a curve $\sigma : (a,b) \to \mathcal{M}$ and a Borel vector field $X : (a,b) \times \mathbb{R}^D \to \mathbb{R}^D$ such that $X_t \in L^2(\sigma_t)$. Here, we have written σ_t in place of $\sigma(t)$ and X_t in place of $X(t)$. We will write

$$\frac{\partial \sigma}{\partial t} + div_\sigma(X) = 0 \tag{14}$$

if the following condition holds: for all $\phi \in C_c^\infty((a,b) \times \mathbb{R}^D)$,

$$\int_a^b \int_{\mathbb{R}^D} \Big(\frac{\partial \phi}{\partial t} + \nabla\phi \cdot X_t \Big) \, d\sigma_t \, dt = 0, \tag{15}$$

i.e. if equation (14) holds in the sense of distributions. Given σ_t, notice that if equation (14) holds for X then it holds for $X+W$, for any Borel map $W : (a,b)\times\mathbb{R}^D \to \mathbb{R}^D$ such that $W_t \in \mathrm{Ker}(div_{\sigma_t})$ i.e $\int \nabla\phi \cdot W_t d\sigma_t = 0$ for all $\psi \in C_c^\infty(\mathbb{R}^D)$ and $t \in (a,b)$.

The following definition and remark can be found in [3] Chapter 1.

Definition 1.2.13. Let $(\mathbb{S}, \mathrm{dist})$ be a metric space. A curve $t \in (a,b) \mapsto \sigma_t \in \mathbb{S}$ is *2–absolutely continuous* if there exists $\beta \in L^2(a,b)$ such that

$$\mathrm{dist}(\sigma_t, \sigma_s) \leq \int_s^t \beta(\tau) d\tau \tag{16}$$

for all $a < s < t < b$. We then write $\sigma \in AC_2(a,b;\mathbb{S})$. For such curves the limit $|\sigma'|(t) := \lim_{s\to t} \text{dist}(\sigma_t,\sigma_s)/|t-s|$ exists for $\mathcal{L}^1$–almost every $t \in (a,b)$. We call this limit the *metric derivative* of σ at t. It satisfies $|\sigma'| \leq \beta\ \mathcal{L}^1$–almost everywhere.

Remark 1.2.14. (i) If $\sigma \in AC_2(a,b;\mathbb{S})$ then $|\sigma'| \in L^2(a,b)$ and $\text{dist}(\sigma_s,\sigma_t) \leq \int_s^t |\sigma'|(\tau)d\tau$ for $a < s < t < b$. We can apply Hölder's inequality to conclude that $\text{dist}^2(\sigma_s,\sigma_t) \leq c|t-s|$, where $c = \int_a^b |\sigma'|^2(\tau)d\tau$.
(ii) It follows from (i) that $\{\sigma_t|\ t \in [a,b]\}$ is a compact set, so it is bounded. For instance, $\text{dist}(\sigma_s,\sigma_a) \leq \sqrt{c|s-a|}$.

We now recall [3] Theorem 8.3.1. It shows that the definition of tangent space given above is flexible enough to include the velocities of any "good" curve in $\mathcal{M}$.

Proposition 1.2.15. *If $\sigma \in AC_2(a,b;\mathcal{M})$ then there exists a Borel map $v : (a,b) \times \mathbb{R}^D \to \mathbb{R}^D$ such that $\frac{\partial\sigma}{\partial t} + div_\sigma(v) = 0$ and $v_t \in L^2(\sigma_t)$ for $\mathcal{L}^1$–almost every $t \in (a,b)$. We call v a velocity for σ. If w is another velocity for σ then the projections $\pi_{\sigma_t}(v_t)$, $\pi_{\sigma_t}(w_t)$ coincide for $\mathcal{L}^1$–almost every $t \in (a,b)$ where π_{σ_t} is defined by the equation (116). One can choose v such that $v_t \in \overline{\nabla C_c^\infty}^{\sigma_t}$ and $||v_t||_{\sigma_t} = |\sigma'|(t)$ for $\mathcal{L}^1$–almost every $t \in (a,b)$. In that case, for $\mathcal{L}^1$–almost every $t \in (a,b)$, v_t is uniquely determined. We denote this velocity $\dot\sigma$ and refer to it as the velocity of minimal norm, since if w_t is any other velocity associated to σ then $||\dot\sigma_t||_{\sigma_t} \leq ||w_t||_{\sigma_t}$ for $\mathcal{L}^1$–almost every $t \in (a,b)$ and so $\text{dist}(\sigma_t,\sigma_s) \leq \int_s^t ||\dot\sigma_\tau||_{\sigma_\tau} d\tau \leq \int_s^t ||w_\tau||_{\sigma_\tau} d\tau$ for all $a < s < t < b$.*

We also recall [3] Proposition 8.4.6. which gives another taste of tangent space.

Proposition 1.2.16. *If $\sigma \in AC_2(a,b;\mathcal{M})$ and let $v_t = \dot\sigma_t \in T_{\sigma_t}\mathcal{M}$ be the velocity of minimal norm. Then, for $\mathcal{L}^1-$ a.e. $t \in (a,b)$ the following property holds: for any choice of $\gamma_h \in \Gamma_o(\mu_t,\mu_{t+h})$ we have*

$$\lim_{h\to 0}\big(\pi^1, \frac{1}{h}(\pi^2-\pi^1)\big)_{\#}\gamma_h = (Id\times v_t)_{\#}\mu_t \qquad in \quad \mathcal{P}_2(\mathbb{R}^D\times\mathbb{R}^D) \tag{17}$$

The following remark can be found in [3] Lemma 1.1.4 in a more general context.

Remark 1.2.17 (Lipschitz reparametrization). Let $\sigma \in AC_2(a, b; \mathcal{M})$ and v be a velocity associated to σ. Fix $\alpha > 0$ and define $S(t) = \int_a^t \big(\alpha + ||v_\tau||_{\sigma_\tau}\big) d\tau$. Then $S : [a, b] \to [0, L]$ is absolutely continuous and increasing, with $L = S(b)$. The inverse of S is a function whose Lipschitz constant is less than or equal to $1/\alpha$. Define

$$\bar{\sigma}_s := \sigma_{S^{-1}(s)}, \qquad \bar{v}_s := \dot{S}^{-1}(s) v_{S^{-1}(s)}.$$

One can check that $\bar{\sigma} \in AC_2(0, L; \mathcal{M})$ and that $\bar{v}$ is a velocity associated to $\bar{\sigma}$. Fix $t \in (a, b)$ and set $s := S(t)$. Then $v_t = \dot{S}(t)\bar{v}_{S(t)}$ and $||\bar{v}_s||_{\sigma_s} = \frac{||v_t||_{\sigma_t}}{\alpha + ||v_t||_{\sigma_t}} < 1$.

Now we introduce some preliminary lemmas which will be used in the construction of Hamiltonian flows.

Lemma 1.2.18. *Suppose $\{f_n\}_{n=1}^\infty$ weakly converges to f in $L^2(0, T)$ then we have*

$$f(t) \leq \limsup_{n\to\infty} f_n(t) \quad \text{for} \quad \text{a.e} \quad t \in (0, t)$$

Proof. By slightly modifying theorem 2.13 in [19], we have a sequence $\{F_n\}_{n=1}^\infty$ of convex combinations of f_n converging to f a.e $t \in (0, t)$ i.e

$$F_n := \sum_{i=n}^{m_n} \lambda_i^n f_i \quad \text{where} \quad \sum_{i=1}^{n} \lambda_i^n = 1, \quad \lambda_i^n \geq 0$$

converges to f a.e $t \in (0, t)$. So we have

$$\limsup_{n\to\infty} f_n = \lim_{n\to\infty} \sup_{i \geq n} f_i \geq \lim_{n\to\infty} F_n = f$$

□

Lemma 1.2.19. *Let $f : \mathbb{R}^D \to \mathbb{R}^D$ be a Borel map, $\mu \in \mathcal{P}(\mathbb{R}^D)$, and let $v \in L^2(\mu; \mathbb{R}^D)$. Then, setting $\nu = f_\# \mu$, we have $f_\#(v\mu) = w\nu$ for some $w \in L^2(\nu)$ with*

$$||w||_{L^2(\nu)} \leq ||v||_{L^2(\mu)}$$

Proof. Lemma 7.1 in [2] □

Lemma 1.2.20. *Let* $T > 0$, $C \geq 0$, $\mu_t^n : [0,T] \to \mathcal{M}$ *and* $v_t^n : \mathbb{R}^D \to \mathbb{R}^k$. *Let* $v_t^n \in L^2(\mu_t^n; R^k)$ *be satisfying*

(a) $\mu_t^n \to \mu_t$ *narrowly as* $n \to \infty$ *for all* $t \in [0,T]$

(b) $||v_t^n||_{L^2(\mu_t^n;R^k)} \leq C$ *for a.e* $t \in [0,T]$

(c) The R^k *valued space-time measures* $v_t^n \mu_t^n dt$ *are weakly** *converging in* $(0,T) \times \mathbb{R}^D$ *to* σ.

Then there exists $v_t \in L^2(\mu_t; R^k)$ *with* $||v_t||_{L^2(\mu_t;R^k)} \leq C$ *for a.e* t *such that* $\sigma = v_t \mu_t dt$

Proof. Lemma 7.2 in [2] □

CHAPTER II

HAMILTONIAN SYSTEMS

In contrast to the theory of gradient flows on the Wasserstein space [3][9][17][25], the theory of Hamiltonian systems presents many more additional difficulties which have not yet been well understood. The first systematic study addressing evolution problems on $\mathcal{M}$ of the Hamiltonian type was made by Ambrosio and Gangbo [2]. In [2], they studied Hamiltonian systems of the form

$$\begin{cases} \frac{d}{dt}\mu_t + \nabla \cdot (Jv_t\mu_t) = 0, \qquad t \in (0,T) \\ v_t \in \partial.H(\mu_t) \cap T_{\mu_t}\mathcal{M}, \end{cases} \tag{18}$$

where the given function $H : \mathcal{M} \to \mathbb{R}$ is referred to as a Hamiltonian. Here $J : \mathbb{R}^D \to \mathbb{R}^D$ is a matrix satisfying $Jv \perp v$ for all $v \in \mathbb{R}^D$. When $D = 2d$ then we can simply set J to be the $(2d) \times (2d)$ canonical symplectic matrix. The theory in [2] covers a large class of systems which have recently generated a lot of interest, including the Vlasov-Poisson in one space dimension [6][30], the Vlasov-Monge-Ampere [7][12] and the semigeostophic systems [4][10][11][12].

Our goal is to study the stability properties of the system (18) for Hamiltonians H satisfying properties imposed in [2]. More precisely, we replace H by a Hamiltonian H_τ and denote by μ_τ the solution of the system with the new Hamiltonian. We suppose that H_τ converges to H and μ_τ converges to μ as τ tends to zero. The topology used for these convergences are to be made specific soon. Our stability property is that the limiting solution μ solves (18). There are not so many natural way to regularize a function H on an infinite dimensional manifold to obtain a new function H_τ. The Moreau-Yosida regularization is one way of regularizing in that context. In this study, we keep our focus on the Moreau-Yosida regularization.

The main result of this chapter is Theorem 2.3.4. In Theorem 2.3.3, we show there exists a solution μ^τ of equation (18) where H is substituted by H_τ and subdifferential is replaced by superdifferential. Here we don't need the condition (H2') on H. Theorem 2.3.4 says that $\{\mu^\tau\}_{\tau>0}$ converges to the solution μ of equation (18) by imposing the condition (H2'). This result establishes the stability property linking the solution of H_τ to those of H.

We view our study as a preliminary step for gaining enough insights which could be exploited to extend the above results to Hamiltonian systems where H fails to satisfy properties imposed in [2].

2.1 Moreau-Yosida approximation

Definition 2.1.1. Let $H : \mathcal{M} \to (-\infty, \infty]$ be a lower semicontinuous functional. For $\tau > 0$, the Moreau-Yosida approximation H_τ of H is defined as

$$H_\tau(\mu) := \inf_{\nu\in\mathcal{M}} \mathbf{H}(\tau,\mu;\nu) = \inf_{\nu\in\mathcal{M}} \{\frac{1}{2\tau}W_2^2(\mu,\nu) + H(\nu)\} \tag{19}$$

where

$$\mathbf{H}(\tau,\mu;\nu) = \frac{1}{2\tau}W_2^2(\mu,\nu) + H(\nu) \tag{20}$$

We also set

$$J_\tau[\mu] := \{\mu_\tau : \mathbf{H}(\tau,\mu;\mu_\tau) \le \mathbf{H}(\tau,\mu;\nu) \quad \forall \nu \in \mathcal{M}\} \tag{21}$$

Lemma 2.1.2. *If $H_\tau(\mu_o) = \frac{1}{2\tau}W_2^2(\mu_o,\nu_o) + H(\nu_o)$ i.e $\nu_o \in J_\tau[\mu_o]$ then*

$$H_\tau(\mu) - H_\tau(\mu_o) \le \int_{\mathbb{R}^d\times\mathbb{R}^d} < \frac{Id - \bar{\gamma}_{\mu_o}^{\nu_o}}{\tau}(x), y - x > d\gamma_{\mu_o}^{\mu}(x,y) + \frac{1}{2\tau}W_2^2(\mu_o,\mu) \quad \forall \quad \mu \tag{22}$$

where $\gamma_{\mu_o}^{\mu} \in \Gamma_o(\mu_o,\mu)$. In particular, $\frac{Id-\bar{\gamma}_{\mu_o}^{\nu_o}}{\tau}$ belongs to the super-differential of H_τ at μ_o. Here $\bar{\gamma}_{\mu_o}^{\nu_o}$ is the barycentric projection of $\gamma_{\mu_o}^{\nu_o} \in \Gamma_o(\mu_o,\nu_o)$ with respect to μ_o which is defined by equation (5).

Proof.

$$H_\tau(\mu_o) = \frac{1}{2\tau}W_2^2(\mu_o,\nu_o) + H(\nu_o), \qquad H_\tau(\mu) \le \frac{1}{2\tau}W_2^2(\mu,\nu_o) + H(\nu_o) \quad \forall \quad \mu$$

So we have

$$\begin{aligned} H_\tau(\mu) - H_\tau(\mu_o) &\leq \frac{1}{2\tau}W_2^2(\mu,\nu_o) - \frac{1}{2\tau}W_2^2(\mu_o,\nu_o) \\ &\leq \int < \frac{x-\bar{\gamma}_{\mu_o}^{\nu_o}(x)}{\tau}, y-x > d\gamma_{\mu_o}^{\mu}(x,y) + \frac{1}{2\tau}W_2^2(\mu_o,\mu) \end{aligned}$$

The second inequality follows from (-1)-convexity of $\mu \to -\frac{1}{2}W_2^2(\mu,\nu_o)$ and Proposition 4.2 and 4.3 of [2] □

Lemma 2.1.3. *Suppose* $\psi : \mathcal{M} \to (-\infty,\infty]$, $\lambda \in \mathbb{R}$ *and for every* $\bar{\mu} \in \mathcal{M}$, *there exists* $\bar{\xi} \in T_{\bar{\mu}}\mathcal{M}$ *such that*

$$\psi(\mu) \geq \psi(\bar{\mu}) + \int < \bar{\xi}(x), y-x > d\gamma_{\bar{\mu}}^{\mu}(x,y) + \frac{\lambda}{2}W_2^2(\mu,\bar{\mu}). \tag{23}$$

Then ψ *is* λ- *convex.*

Proof. Let $\mu_o, \mu_1 \in \mathcal{M}$ and $t \to \mu_t$ be a geodesic between μ_o and μ_1. We want to show

$$\psi(\mu_t) \leq (1-t)\psi(\mu_o) + t\psi(\mu_1) - \frac{\lambda}{2}t(1-t)W_2^2(\mu_o,\mu_1)$$

Let $\gamma_{\mu_o}^{\mu_1} \in \Gamma_0(\mu_0,\mu_1)$, we define $\gamma_{\mu_o}^{\mu_t} = (\pi_1,(1-t)\pi_1+t\pi_2)_{\#}\gamma_{\mu_o}^{\mu_1}$ and $\gamma_{\mu_t}^{\mu_1} = ((1-t)\pi_1+t\pi_2,\pi_2)_{\#}\gamma_{\mu_o}^{\mu_1}$ then

$$\gamma_{\mu_o}^{\mu_t} \in \Gamma_o(\mu_o,\mu_t), \quad \gamma_{\mu_t}^{\mu_1} \in \Gamma_o(\mu_t.\mu_1),$$

So

$$(1-t)\psi(\mu_o) \geq (1-t)\psi(\mu_t) + (1-t)\int < \xi_t(z), x-z > d\gamma_{\mu_t}^{\mu_o}(z,x) + \frac{\lambda}{2}(1-t)W_2^2(\mu_o,\mu_t) \tag{24}$$

and

$$t\psi(\mu_1) \geq t\psi(\mu_t) + t\int < \xi_t(z), x-z > d\gamma_{\mu_t}^{\mu_1}(z,x) + \frac{\lambda}{2}tW_2^2(\mu_1,\mu_t) \tag{25}$$

We combine equations (24) and (25) to conclude

$$\begin{aligned}(1-t)\psi(\mu_o)+t\psi(\mu_1) &\geq \psi(\mu_t)-(1-t)t\int <\xi_t((1-t)x+ty),y-x> d\gamma_{\mu_o}^{\mu_1}(x,y)\\ &+ (1-t)t^2\frac{\lambda}{2}W_2^2(\mu_o,\mu_1)+t\frac{\lambda}{2}(1-t)^2W_2^2(\mu_o,\mu_1)\\ &+ t(1-t)\int <\xi_t((1-t)x+ty),y-x> d\gamma_{\mu_o}^{\mu_1}(x,y)\\ &= \psi(\mu_t)+\frac{\lambda}{2}t(1-t)W_2^2(\mu_o,\mu_1)\end{aligned}$$

□

Corollary 2.1.4. *Let $H:\mathcal{M}\to(-\infty,\infty]$ be a lower semicontinuous functional and H_τ be the Moreau-Yosida approximation of H. If $J_\tau[\mu]\neq\emptyset$ for all $\mu\in\mathcal{M}$ then H_τ is $\frac{1}{\tau}$ - concave.*

Proof. Equations (22) and (23) give that $-H_\tau$ is $-\frac{1}{\tau}$ convex which means that H_τ is $\frac{1}{\tau}$ - concave. □

Remark 2.1.5. Suppose H is lower semicontinuous. If H is bounded below or λ-convex then $J_\tau[\mu]\neq\emptyset$ for all $\mu\in\mathcal{M}$ [3].

Lemma 2.1.6. *Let $H:\mathcal{M}\to(-\infty,+\infty)$ be a λ- convex continuous functional and H_τ be the Moreau-Yosida approximation of H. Let $\nu_\tau\in J[\mu_\tau]$ i.e*

$$H_\tau(\mu_\tau)=\frac{1}{2\tau}W_2^2(\mu_\tau,\nu_\tau)+H(\nu_\tau)$$

Suppose there exists a function S_τ such that $(Id\times S_\tau)_{\#}\mu_\tau\in\Gamma_o(\mu_\tau,\nu_\tau)$ for each τ, and μ_τ and ν_τ have uniformly bounded support(independent on τ) i.e supp(μ_τ), supp$(\nu_\tau)\subset B_r(0)$ for some $r>0$. Define

$$v_\tau:=\frac{Id-\bar{\gamma}_{\mu_\tau}^{\nu_\tau}}{\tau}=\frac{Id-S_\tau}{\tau}\in\partial^+H_\tau(\mu_\tau)\cap T_{\mu_\tau}\mathcal{M}$$

Suppose $\mu_\tau\to\tilde{\mu}$ in $(\mathcal{M},W_2)$ and $||v_\tau||_{L^2(\mu_\tau)}\leq C$ for all τ then $\nu_\tau\to\tilde{\mu}$ in $(\mathcal{M},W_2)$. Furthermore, if

$$v\tilde{\mu}\in\bigcap_{M=1}^{\infty}\bar{co}(\{v_{\tau_n}\mu_{\tau_n}:n\geq M\})$$

where $\bar{co}$ denotes the closed convex hull with respect to weak$*$- topology, then

$$v\tilde{\mu} \in \bigcap_{M=1}^{\infty} \bar{co}(\{\xi_{\tau_n}\nu_{\tau_n} : n \geq M\})$$

where $\xi_\tau := \frac{\bar{\gamma}_{\nu_\tau}^{\mu_\tau} - Id}{\tau} \in \partial_- H(\nu_\tau) \cap T_{\nu_\tau}\mathcal{M}$.

Proof. By Lemma 2.1.2, we have for all μ

$$\begin{aligned} H_\tau(\mu) - H_\tau(\mu_\tau) &\leq \int < \frac{x - \bar{\gamma}_{\mu_\tau}^{\nu_\tau}(x)}{\tau}, y - x > d\gamma_\tau^1(x,y) + \frac{1}{2\tau} W_2^2(\mu_\tau, \mu) \\ &= \int < v_\tau(x), y - x > d\gamma_\tau^1(x,y) + \frac{1}{2\tau} W_2^2(\mu_\tau, \mu) \end{aligned} \tag{26}$$

where $\gamma_\tau^1 \in \Gamma_o(\mu_\tau, \mu)$ and

$$\begin{aligned} H(\nu) - H(\nu_\tau) &\geq -\frac{1}{2\tau} W_2^2(\mu_\tau, \nu) + \frac{1}{2\tau} W_2^2(\mu_\tau, \nu_\tau) \\ &\geq \int < \frac{x_2 - x_1}{\tau}, x_3 - x_1 > d\hat{\mu} - \frac{1}{2\tau} W_2^2(\nu_\tau, \nu) \\ &= \int < \frac{\bar{\gamma}_{\nu_\tau}^{\mu_\tau}(x) - x}{\tau}, y - x > d\gamma_\tau^2(x,y) - \frac{1}{2\tau} W_2^2(\nu_\tau, \nu) \\ &= \int < \xi_\tau(x), y - x > d\gamma_\tau^2(x,y) - \frac{1}{2\tau} W_2^2(\nu_\tau, \nu) \end{aligned} \tag{27}$$

where $\hat{\gamma}_\tau \in \Gamma_o(\nu_\tau, \mu_\tau), \gamma_\tau^2 \in \Gamma_o(\nu_\tau, \nu)$ and $\hat{\mu} = \hat{\gamma}_\tau \times \gamma_\tau^2 \in \Gamma_o(\hat{\gamma}_\tau, \nu)$ constructed as in lemma 1.2.11. Define $\xi_\tau(x) := \frac{\bar{\gamma}_{\nu_\tau}^{\mu_\tau}(x) - x}{\tau}$ then the equation (27) says $\xi_\tau \in \partial_- H(\nu_\tau)$ and trivially $\xi_\tau \in T_{\nu_\tau}M$. Since H is λ - convex, Proposition 4.2 of [2] gives

$$H(\nu) - H(\nu_\tau) \geq \int < \xi_\tau(x), y - x > d\gamma_\tau^2(x,y) + \frac{\lambda}{2} W_2^2(\nu_\tau, \nu) \tag{28}$$

Notice that $\{H(\nu_\tau)\}_{\tau>0}$ is bounded below. Otherwise $H_\tau(\tilde{\mu}) = \inf_\nu\{\frac{1}{2\tau} W_2^2(\tilde{\mu}, \nu) + H(\nu)\} = -\infty$ for all $\tau > 0$ and this is a contradiction to the fact that $\{H_\tau(\tilde{\mu})\}_{\tau>0}$ converges to $H(\tilde{\mu}) > -\infty$ as $\tau \to 0$. Since $\nu_\tau \in J[\mu_\tau]$,

$$\begin{aligned} W_2^2(\mu_\tau, \nu_\tau) &= 2\tau(H_\tau(\mu_\tau) - H(\nu_\tau)) \\ &\leq 2\tau(H(\mu_\tau) - H(\nu_\tau)) \end{aligned}$$

this together with $H(\nu_\tau)$ is bounded below and $\{\mu_\tau\}_{\tau>0}$ converges to $\tilde{\mu}$ in W_2 gives that $\{\nu_\tau\}_{\tau>0}$ also converges to $\tilde{\mu}$ in W_2 as $\tau \to 0$. Next, let $\gamma_\tau = (Id \times S_\tau)_\# \mu_\tau \in \Gamma_o(\mu_\tau, \nu_\tau)$

then

$$\begin{aligned}\int < \psi(y), \xi_\tau(y) > d\nu_\tau(y) &= \int < \psi(y), \frac{x-y}{\tau} > d\gamma_\tau(x,y) \qquad (29)\\ &= \int < \psi(x) + \nabla\psi(\xi_{x,y}) \cdot (y-x), \frac{x-y}{\tau} > d\gamma_\tau(x,y)\end{aligned}$$

where $\xi_{x,y}$ is in the line segment between x and y.

Since $\psi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$

$$\begin{aligned}\Big|\int < \nabla\psi(\xi_{x,y}) \cdot (y-x), \frac{x-y}{\tau} > d\gamma_\tau(x,y)\Big| &\leq C \int \Big|< (y-x), \frac{x-y}{\tau} >\Big| d\gamma_\tau(x,y)\\ &= C \int \tau |v_\tau(x)|^2 d\mu_\tau(x)\\ &= O(\tau) \qquad (30)\end{aligned}$$

We combine equations (29) and (30) to get

$$\begin{aligned}\int < \psi(y), \xi_\tau(y) > d\nu_\tau(y) &= \int < \psi(x), \frac{x-y}{\tau} > d\gamma_\tau(x,y) + O(\tau) \qquad (31)\\ &= \int < \psi(x), v_\tau(x) > d\mu_\tau(x) + O(\tau)\end{aligned}$$

Similarly,

$$\sum_{i=n}^{m(n)} \lambda_i \int < \psi(y), \xi_{\tau_i}(y) > d\nu_{\tau_i}(y) = \sum_{i=n}^{m(n)} \lambda_i \int < \psi(x), v_{\tau_i}(x) > d\mu_{\tau_i} + O(\tau)$$

Notice that $O(\tau)$ is independent of convex combination of $v_{\tau_i}\mu_{\tau_i}$ because of uniform bound on $||v_{\tau_i}||_{L^2(\mu_{\tau_i})}$. And

$$\begin{aligned}\int < \psi(x), v(x) > d\tilde{\mu}(x) &= \lim_{n\to\infty} \sum_{i=n}^{m(n)} \lambda_i \int < \psi(x), v_{\tau_i}(x) > d\mu_{\tau_i} + O(\tau)\\ &= \lim_{n\to\infty} \sum_{i=n}^{m(n)} \lambda_i \int < \psi(y), \xi_{\tau_i}(y) > d\nu_{\tau_i}(y) \qquad (32)\end{aligned}$$

This concludes proof. □

2.2 *Variational problem in connection with the Moreau-Yosida approximation*

We first introduce some notation which will be frequently used through this section. Recall $\mathcal{M}$ is the space of probability measures on $\mathbb{R}^D$ with finite second moment. We define

$$\mathcal{M}_r := \{\mu \in \mathcal{M} : supp(\mu) \subset B_r(0)\} \tag{33}$$

2.2.1 A special class of Hamiltonian

For a given $\lambda \in \mathbb{R}$, we define

$$H(\mu) = -\sup_{(v,B)\in\mathcal{C}} \Big\{ \int v d\mu + B \Big\} \tag{34}$$

where $\mathcal{C} \subset \{(v, B); \frac{\lambda|x|^2}{2} - v(x) \text{ is convex}, B \in (-\infty, \infty)\}$. Without loss of generality, we may assume $v(0) = 0$. Let $\tau \in (0, 1)$, we define Moreau-Yosida approximation of H

$$H_\tau(\mu) = \inf_{\nu\in\mathcal{M}} \Big\{ \frac{1}{2\tau} W_2^2(\mu, \nu) + H(\nu) \Big\} \tag{35}$$

From the definition of H_τ, we have

$$\begin{aligned} H_\tau(\mu) &= \inf_{\nu\in\mathcal{M}} \Big\{ \frac{1}{2\tau} W_2^2(\mu, \nu) - \sup_{(v,B)\in\mathcal{C}} \{ \int v d\nu + B \} \Big\} \\ &= \inf_{(v,B)\in\mathcal{C}} \Big\{ \inf_{\nu\in\mathcal{M}} \{ \frac{1}{2\tau} W_2^2(\mu, \nu) + \int -v d\nu \} - B \Big\} \end{aligned} \tag{36}$$

The problem in the equation (36) suggests that we need to well study problem of the form

$$\inf_{\nu\in\mathcal{M}} \Big\{ \frac{1}{2\tau} W_2^2(\mu, \nu) + \int -v d\nu \Big\} \tag{37}$$

Lemma 2.2.1. *For given $\mu \in \mathcal{M}$ and $(v, B) \in \mathcal{C}$, problem (37) has a minimizer μ_v. Further more there exists a convex function ϕ_v such that $\mu_v = T_{v\#}\mu$ where $T_v = \nabla\phi_v$ and T_v is Lipschitz continuous with $Lip(T_v) \le \frac{1}{\tau(1-\lambda\tau)}$ which is independent of v. So* *we have*

$$\inf_{\nu\in\mathcal{M}}\{\frac{1}{2\tau}W_2^2(\mu,\nu)+\int -v d\nu\} = \frac{1}{2\tau}W_2^2(\mu,T_{v\#}\mu)+\int -v dT_{v\#}\mu \tag{38}$$
$$= \int (-v)_\tau d\mu$$

where $(-v)_\tau$ *is defined by equation (41). This implies*

$$H_\tau(\mu) = \inf_{(v,B)\in\mathcal{C}}\{\int(-v)_\tau d\mu - B\} \tag{39}$$

Proof. For given $(v,B)\in\mathcal{C}$ and $\mu\in\mathcal{M}$, we have

$$\inf_{\nu\in\mathcal{M}}\frac{1}{2\tau}W_2^2(\mu,\nu)-\int v d\nu = \inf_{\nu\in\mathcal{M}}\inf_{\gamma\in\Gamma_o(\mu,\nu)}\int\frac{|y-x|^2}{2\tau}-v(y)d\gamma(x,y) \tag{40}$$

Set

$$(-v)_\tau(x) := \inf_{y\in\mathbb{R}^D}\{\frac{|y-x|^2}{2\tau}-v(y)\} \tag{41}$$

Trivially we have

$$\inf_{\nu\in\mathcal{M}}\frac{1}{2\tau}W_2^2(\mu,\nu)-\int v d\nu \geq \inf_{\nu\in\mathcal{M}}\inf_{\gamma\in\Gamma_o(\mu,\nu)}\int(-v)_\tau(x)d\gamma(x,y) \tag{42}$$
$$= \inf_{\nu\in\mathcal{M}}\inf_{\gamma\in\Gamma_o(\mu,\nu)}\int(-v)_\tau(x)d\mu(x)$$
$$= \int(-v)_\tau d\mu$$

Let us study (41) more carefully. We fix $x\in\mathbb{R}^D$ and define $\phi_x:\mathbb{R}^D\to(-\infty,\infty)$ as

$$\phi_x(y) := \frac{|y-x|^2}{2\tau}-v(y) = \frac{|y-x|^2}{2\tau}-\frac{\lambda}{2}|y|^2+\frac{\lambda}{2}|y|^2-v(y) \tag{43}$$

Since v is $\lambda-$ convex, ϕ_x is strictly convex if $1/\tau-\lambda>0$ and $\phi_x(y)\to\infty$ as $y\to\infty$. So equation (41) has a unique minimizer y_x ie,

$$(-v)_\tau(x) = \frac{|x-y_x|^2}{2\tau}-v(y_x) \tag{44}$$

Next

$$(-v)_\tau(x) = \inf_y\{\frac{|y|^2}{2\tau}+\frac{|x|^2}{2\tau}-\frac{x\cdot y}{\tau}-v(y)\} \tag{45}$$
$$= \frac{|x|^2}{2\tau}-\sup_y\{\frac{x\cdot y}{\tau}+v(y)-\frac{|y|^2}{2\tau}\}$$

So

$$\begin{aligned}\frac{|x|^2}{2\tau} - (-v)_\tau(x) &= \frac{1}{\tau}\sup_y\{x \cdot y + \tau v(y) - \frac{|y|^2}{2}\} \\ &= \frac{1}{\tau}\psi^*(x)\end{aligned} \tag{46}$$

where $\psi(y) := \frac{|y|^2}{2} - \tau v(y)$ and ψ^* is the Legendre transform of ψ. Since

$$\psi(y) = \frac{|y|^2}{2} - \tau v(y) = (1-\lambda\tau)\frac{|y|^2}{2} + \tau(\frac{\lambda}{2}|y|^2 - v(y)) \tag{47}$$

if $1-\lambda\tau > 0$ then ψ is strictly convex function and $\nabla^2\psi(y) \geq I(1-\tau\lambda)$. This implies $\psi^* \in C^1$ and $Lip(\nabla\psi^*) \leq 1/(1-\lambda\tau)$.

From equation (46), we have

$$(-v)_\tau \in C^1, \qquad Lip(\nabla(-v)_\tau) \leq \frac{1}{\tau(1-\lambda\tau)} + \frac{1}{\tau} \tag{48}$$

We combine equations (44), (45) and (48) to get

$$\begin{aligned}\nabla(-v)_\tau(x) = \frac{x}{2} - \frac{y_x}{2} \Rightarrow y_x &= x - \tau\nabla(-v)_\tau(x) \\ &= \nabla(\frac{|x|^2}{2} - \tau(-v)_\tau(x)) = \frac{1}{\tau}\nabla\psi^*(x)\end{aligned} \tag{49}$$

Let us define $T_v(x) := x - \tau\nabla(-v)_\tau(x) = 1/\tau\nabla\psi^*(x)$ then T_v is the gradient of a convex function. This says T_v is the optimal transport map between μ and $T_{v\#}\mu$ i.e $(Id \times T_v)_{\#}\mu \in \Gamma_0(\mu, T_{v\#}\mu)$.

Furthermore, from equations (44) and (49) we have

$$(-v)_\tau(x) = \frac{|x - T_v(x)|^2}{2\tau} - v(T_v(x)) \tag{50}$$

Define $\phi_v := \frac{1}{\tau}\psi^*$ where ψ^* is defined as in (46). Then the rest of the statement in Lemma 2.2.1 was already proven and we only need to show that the equation (38) holds. Since $(Id \times T_v)_{\#}\mu \in \Gamma_0(\mu, T_{v\#}\mu)$,

$$\begin{aligned}\inf_{\nu\in\mathcal{M}} \frac{1}{2\tau}W_2^2(\mu,\nu) - \int v d\nu &\leq \inf_{\nu=T_v\#\mu}\int \frac{|T_v(x)-x|^2}{2\tau} - v(T_v(x))d\mu(x) \\ &= \int(-v)_\tau(x)d\mu(x)\end{aligned} \tag{51}$$

We combine (42) and (51) to conclude proof. □

Remark 2.2.2. For later use, we estimate the minimizer of $-v_\tau(0)$ in the equation (41).

$$(-v)_\tau(0) = \inf_{y\in\mathbb{R}^D}\{\frac{|y|^2}{2\tau} - v(y)\} = \inf_{y\in\mathbb{R}^D}\phi_0(y) \tag{52}$$

Let y_0 be the minimizer then , recall that $v(0) = 0$

$$\begin{aligned} 0 = \phi_0(0) \geq \phi_0(y_0) &= (\frac{1}{2\tau} - \frac{\lambda}{2})|y_0|^2 + \frac{\lambda}{2}|y_0|^2 - v(y_0) \\ &\geq (\frac{1}{2\tau} - \frac{\lambda}{2})|y_0|^2 + < \xi, y_0 > \end{aligned} \tag{53}$$

where $\xi \in \partial_-\psi(0)$ and $\psi(y) = \frac{\lambda}{2}|y|^2 - v(y)$. This means

$$|T_v(0)| = |y_0| \leq \frac{2\tau}{1-\tau\lambda}|\xi| \tag{54}$$

Next, $\phi_0(y_0) = (-v)_\tau(0)$ together equation (53) gives an estimate

$$0 \geq (-v)_\tau(0) \geq (\frac{1}{2\tau} - \frac{\lambda}{2})|y_0|^2 - |\xi||y_0| \geq -(\frac{|\xi|\tau}{1-\lambda\tau})^2 \tag{55}$$

So we have

$$0 \leq \frac{\psi^*(0)}{\tau} = -(-v)_\tau(0) \leq (\frac{|\xi|\tau}{1-\lambda\tau})^2 \tag{56}$$

2.2.2 Properties of minimizers

Now we impose an extra condition on $\mathcal{C}$ in equation (34) namely *" locally uniform L_1 bound"* condition ;

For every compact $K \subset \mathbb{R}^d$, there exists a constant $M(K)$ only depending on K such that $\int_K |v|dx \leq M(K)$ for all $(v, B) \in \mathcal{C}$.

Lemma 2.2.3. *Let $H : \mathcal{M} \to (-\infty, \infty)$ be defined by equation (34). If, for every compact $K \subset \mathbb{R}^d$, we have a constant $M(K)$ only depending on K such that $\int_K |v|dx \leq M(K)$ for all $(v, B) \in \mathcal{C}$ then $|\nabla v|_{L^\infty(B_r(0))} \leq C_1(r)$ for all $(v, B) \in \mathcal{C}$ and $H : \mathcal{M}_r \to (-\infty, \infty)$ is continuous w.r.t W_2 distance.*

Proof. If $f \in C$ is convex then we have (Theorem 6.3.1 of [13])

$$ess \sup_{B_r(0)} |\nabla f| \leq \frac{C_o}{r} \frac{1}{|B_r(0)|} \int_{B_{2r}(0)} |f| dx$$

Since $\frac{\lambda |x|^2}{2} - v$ is convex, we have

$$ess \sup_{B_r(0)} |\nabla v| \leq |\lambda| r + \frac{C_0}{r} \frac{1}{|B_r(0)|} \int_{B_{2r}(0)} |v - \frac{|x|^2}{2}| dx := C_1(r) \tag{57}$$

So $|\nabla v|_{L^\infty(B_r(0))} \leq C_1(r)$ for all $v \in \mathcal{C}$.

Suppose

$$H(\mu_1) = \int v_1 d\mu_1 + B_1$$

then

$$H(\mu_2) - H(\mu_1) \leq \int v_1 (d\mu_2 - d\mu_1) \leq C_1(r) W_1(\mu_1, \mu_2) \leq C_1(r) W_2(\mu_1, \mu_2)$$

$\square$

Lemma 2.2.4. *Let $\mu \in \mathcal{M}$ and $(v, B) \in \mathcal{C}$ then the problem 37 has a minimizer μ_v. Further more there exists a convex function ϕ_v such that $\mu_v = T_{v\#}\mu$ where $T_v = \nabla \phi_v$ and T_v is Lipschitz continuous with Lip(T_v) $\leq \frac{1}{\tau(1-\lambda\tau)}$ which is independent on v. Suppose $\mathcal{C}$ satisfies " locally uniform L_1 bound" condition and μ has a bounded support i.e supp(μ) $\subset B_R(0)$ then there exist constants C_2 and C_3 which are independent on v such that for all $x \in B_R(0)$ i.e $|x| \leq R$*

$$|T_v(x)| \leq C_2, \qquad |\phi_v(x)| \leq C_3 \tag{58}$$

and supp(μ_v) $\subset B_{C_2}(0)$.

Proof. With Lemma 2.2.1, we only need to show equation (58). From equations (54) and (57), we have

$$|T_v(0)| \leq \frac{2\tau}{1 - \lambda\tau} C_1(r)$$

where r is any fixed number between 0 and R. So for $|x| \leq R$,

$$|T_v(x)| \leq |T_v(0)| + \text{Lip}(T_v)|x| \leq \frac{2\tau}{1 - \lambda\tau} C_1(r) + \frac{1}{\tau(1 - \lambda\tau)} R := C_2 \tag{59}$$

Similarly we have

$$|\phi_v(x)| \leq |\phi_v(0)| + |\nabla\phi_v(y)||x| \tag{60}$$

where y is on the line segment between0 and x. Equations (56), (59), (60) together $T_v = \nabla\phi_v$ give

$$|\phi_v(x)| \leq (\frac{\tau C_1(r)}{1-\lambda\tau})^2 + C_2 R := C_3 \tag{61}$$

□

Theorem 2.2.5. *Let $H : \mathcal{M} \to (-\infty, \infty)$ be defined by equation (34) and let $\mu \in \mathcal{M}$ have a bounded support. Then*

$$\inf_{\nu\in\mathcal{M}}\{\frac{1}{2\tau}W_2^2(\mu,\nu) + H(\nu)\} \tag{62}$$

(62) has a minimizer ν_o and there exists a Lipschitz function $T = \nabla\psi$ where ψ is a convex function such that $\nu_o = T_{\#}\mu$, $\mathrm{Lip}(T) \leq \frac{1}{\tau(1-\lambda\tau)}$ and $|T(x)| \leq C_2$ if $x \in B_R(0)$.

Proof. Let

$$\begin{aligned} l &:= \inf_{\nu\in\mathcal{M}}\{\frac{1}{2\tau}W_2^2(\mu,\nu) + H(\nu)\} \\ &= \inf_{(v,B)\in\mathcal{C}}\inf_{\nu\in\mathcal{M}}\{\frac{1}{2\tau}W_2^2(\mu,\nu) - \int v d\nu - B\} \end{aligned}$$

Define $L(\nu, v, B) := \frac{1}{2\tau}W_2^2(\mu,\nu) - \int v d\nu - B$, then there exists a minimizing sequence (ν_k, v_k, B_k) such that

$$l = \liminf_{k\to 0} L(\nu_k, v_k, B_k) \tag{63}$$

Let $\bar{\nu}_k$ be a minimizer of equation (37) w.r.t v_k i.e $\bar{\nu}_k$ satisfies the following,

$$\begin{aligned} \frac{1}{2\tau}W_2^2(\mu,\bar{\nu}_k) - \int v_k d\bar{\nu}_k &= \inf_{\nu\in\mathcal{M}}\{\frac{1}{2\tau}W_2^2(\mu,\nu) - \int v_k d\nu)\} \\ &\leq \frac{1}{2\tau}W_2^2(\mu,\nu_k) - \int v_k d\nu_k \end{aligned} \tag{64}$$

By Lemma 2.2.4, there exists a convex function ϕ_k such that $\bar{\nu}_k = T_{k\#}\mu$ where $T_k = \nabla\phi_k$ and T_k is Lipschitz continuous with $\mathrm{Lip}(T_k) \leq \frac{1}{\tau(1-\tau\lambda)}$. And Lemma 2.2.4

also says $|T_k| \leq C_2$(independent on k) on the bounded set $B_R(0)$. Equations (63) and (64) give

$$l \leq L(\bar{\nu}_k, v_k, B_k) \leq L(\nu_k, v_k, B_k)$$

and

$$l \leq \frac{1}{2\tau} W_2^2(\mu, \bar{\nu}_k) + H(\bar{\nu}_k) \leq L(\bar{\nu}_k, v_k, B_k) \leq L(\nu_k, v_k, B_k)$$

Since supp $(\bar{\nu}_k)$ are uniformly bounded, up to a subsequence $\{\bar{\nu}_k\}_{k=1}^{\infty}$ converges to $\nu_o \in \mathcal{M}$. Since T_k and ϕ_k are equi-Lipschitz and equi-bounded, up to a subsequence $\{T_k\}_{k=1}^{\infty}$ (respectively $\{\phi_k\}_{k=1}^{\infty}$) converges uniformly to T(respectively ϕ) and $T = \nabla\phi$. Uniform convergence preserves Lipschitz constant and from $T_{k\#}\mu = \bar{\nu}_k$, we have $T_{\#}\mu = \nu_0$. □

2.3 *Hamiltonian flow*

Definition 2.3.1. Let $H : \mathcal{M} \rightarrow (-\infty, \infty]$ be a proper, lower semicontinuous function. We say that an absolutely continuous curve $\mu_t : [0, T] \rightarrow D(H)$ is a solution of the *Hamiltonian system relative to* H, starting from $\bar{\mu} \in \mathcal{M}$, if there exist $v_t \in L^2(\mu_t)$ with $||v_t||_{L^2(\mu_t)} \in L^1(0, T)$, such that

$$\begin{cases} \frac{d}{dt}\mu_t + \nabla \cdot (Jv_t\mu_t) = 0, \quad \mu_0 = \bar{\mu} \qquad t \in (0, T) \\ v_t \in \partial_{.} H(\mu_t) \cap T_{\mu_t}\mathcal{M} \quad a.e \quad t \in (0, T) \end{cases} \tag{65}$$

Here $J : \mathbb{R}^D \rightarrow \mathbb{R}^D$ is a linear map such that $Jv \perp v$ for all $v \in \mathbb{R}^D$. In other words, we need to find $t \longrightarrow \mu_t, \quad t \longrightarrow v_t$ satisfying (65) in the sense of distribution: For any $\eta \in C_c^{\infty}(0, T)$ and $\zeta \in C_c^{\infty}(R^d)$, we have

$$\int_0^T \int_{R^d} \eta'(t)\zeta(x) + \eta(t) < \nabla\zeta(x) : Jv_t(x) > d\mu_t(x)dt = 0$$

Now we are ready to construct a Hamiltonian flow in $\mathcal{M}$ where Hamiltonian H is given by (34). The aim of Lemma 2.3.2 is to show that the Moreau-Yosida approximation H_τ satisfies assumption which was imposed in [2] to ensure existence of solutions in equation (18) where subdifferential is replaced by superdifferential.

Let $\mu, \nu \in \mathcal{M}$ and $\gamma \in \Gamma_o(\mu, \nu)$. Recall from the definition 1.2.4 that $\bar{\gamma}_\mu^\nu$ is the barycentric projection of γ with respect to μ.

Lemma 2.3.2. *Let H be of the form in the equation (34). Suppose $supp(\mu_n) \subset B_R(0)$ and recall that by Theorem 2.2.5, for each n there exists ν_n such that ν_n is a minimizer of (35) w.r.t μ_n i.e $\nu_n \in J_\tau[\mu_n]$. Let $\mu_n \to \mu$ in W_2. Then $\mathcal{K}_o \subset \tilde{\mathcal{K}}_o$ where*

$$\mathcal{K}_o := \bigcap_{m=1}^{\infty} \overline{co}(\{\frac{Id - \bar{\gamma}_{\mu_n}^{\nu_n}}{\tau}\mu_n : n \geq m\})$$

and

$$\tilde{\mathcal{K}}_o := \{w\mu : w = \frac{Id - \bar{\gamma}_\mu^\nu}{\tau} \in \partial^\cdot H_\tau(\mu), \quad \nu \in J_\tau[\mu], \quad \text{Lip}(w) \leq \frac{1}{\tau^2}\}.$$

Further more $\bar{\gamma}_\mu^\nu = \nabla\psi$ for some convex function ψ such that $(Id \times \nabla\psi)_\# \mu \in \Gamma_o(\mu, \nu)$. Here $\overline{co}$ denotes the closed convex hull, with respect to weak - topology.*

Proof. By Theorem 2.2.5, ν_n also has uniformly bounded support. So

$$supp(\mu_n), supp(\nu_n) \subset B_R(0) \quad \text{i.e} \quad \mu_n, \nu_n \in \mathcal{M}_R$$

where $\mathcal{M}_R$ is given as in the equation (33). Let $\mathbf{H}$ and H_τ be as in equations (19) and (20)

$$\mathbf{H}(\tau, \mu; \nu) = \frac{1}{2\tau} W_2^2(\mu, \nu) + H(\nu), \quad H_\tau(\mu) = \inf_\nu \mathbf{H}(\tau, \mu, \nu)$$

Since

$$H_\tau(\mu_n) = \mathbf{H}(\tau, \mu_n, \nu_n)$$

using Lemma 2.1.2, we have

$$H_\tau(\mu) \leq \qquad H_\tau(\mu_n) + \int < \frac{x - \bar{\gamma}_{\mu_n}^{\nu_n}(x)}{\tau}, y - x > d\gamma(x, y) + \frac{1}{2\tau} W_2^2(\mu_n, \mu)$$

where $\gamma \in \Gamma_o(\mu_n, \mu)$. So

$$\frac{Id - \bar{\gamma}_{\mu_n}^{\nu_n}}{\tau} \in \partial^+ H_\tau(\mu_n).$$

Let $\tilde{\mathbf{u}} \in \mathcal{K}_o$ then,

$$\tilde{\mathbf{u}} = \lim_{n\to\infty} \sum_{i=n}^{l(n)} \lambda_i^n \frac{Id - \bar{\gamma}_{\mu_i}^{\nu_i}}{\tau} \mu_i = \lim_{n\to\infty} \sum_{i=n}^{l(n)} \lambda_i^n w_i \mu_i$$

where $\sum_{i=n}^{l(n)} \lambda_i^n = 1$ with $\lambda_i^n \geq 0$, $-\frac{I}{\tau^2} \leq \nabla w_i \leq \frac{I}{\tau^2}$ and $w_i = \frac{Id - \bar{\gamma}_{\mu_i}^{\nu_i}}{\tau}$

From assumption

$$\lim_{i\to\infty} \mu_i = \mu \quad in \quad W_2$$

We define

$$\hat{\mu}_n = \sum_{i=n}^{l(n)} \lambda_i^n \mu_i, \quad \hat{\nu}_n = \sum_{i=n}^{l(n)} \lambda_i^n \nu_i, \quad \gamma_i \in \Gamma_o(\mu_i, \nu_i), \quad \hat{\gamma}_n = \sum_{i=n}^{l(n)} \lambda_i^n \gamma_i$$

Set

$$\begin{aligned} E(\mu) = \{(Id \times \nabla\psi)_{\#\mu} \quad \in \quad & \Gamma_o(\mu, \nu) : \psi - convex, \nu \in \mathcal{M}_R, \mathbf{H}(\tau, \mu, \nu) = H_\tau(\mu), \\ & \frac{Id - \nabla\psi}{\tau} \in \partial^+ H_\tau(\mu) \quad and \quad -\frac{I}{\tau^2} \leq \nabla(\frac{Id - \nabla\psi}{\tau}) \leq \frac{I}{\tau^2}\} \end{aligned}$$

Claim 1. For any i, there exists $\eta_i \in E(\mu)$ such that $\delta(\eta_i, \gamma_i) \to 0$ as $i \to \infty$, where δ is the metric of Remark 5.1.1 of [3] whose topology coincides with the narrow convergence.

(proof) Suppose there exist $\epsilon_o > 0$ and a subsequence $\{\gamma_{i_j}\}_{j=1}^{\infty}$ such that $\delta(\eta, \gamma_{i_j}) \geq \epsilon_o$ for all $\eta \in E(\mu)$. Since μ_{i_j} and ν_{i_j} have uniformly bounded supports, $\gamma_{i_j} \in \Gamma_o(\mu_{i_j}, \nu_{i_j})$ also have uniformly bounded supports. This implies, up to a subsequence we have

$$\nu_{i_j} \to \tilde{\nu}, \qquad \gamma_{i_j} \to \gamma \qquad in \quad W_2 \tag{66}$$

Notice that the cyclical monotonicity gives $\gamma \in \Gamma_o(\mu, \tilde{\nu})$. Recall $\mathbf{H}(\tau, \mu_i, \nu_i) = H_\tau(\mu_i)$ and $(\bar{\gamma}_{\mu_i}^{\nu_i})_{\#}\mu_i = \nu_i$ which means

$$\frac{1}{2\tau} W_2^2(\mu_{i_j}, \nu_{i_j}) + H(\nu_{i_j}) \quad \leq \quad \frac{1}{2\tau} W_2^2(\mu_{i_j}, \nu) + H(\nu) \quad \forall \nu \in \mathcal{M} \tag{67}$$

We combine equations (66) and (67) to get

$$\frac{1}{2\tau}W_2^2(\mu,\tilde{\nu}) + H(\tilde{\nu}) \quad \leq \quad \frac{1}{2\tau}W_2^2(\mu,\nu) + H(\nu) \quad \forall \nu \in \mathcal{M} \tag{68}$$

So we have $H_\tau(\mu) = \mathbf{H}(\tau,\mu,\tilde{\nu})$. This together with $\gamma \in \Gamma_o(\mu,\tilde{\nu})$ gives $\gamma \in E(\mu)$ which contradicts the fact that $\delta(\eta,\gamma_{i_j}) \geq \epsilon_o$ for all $\eta \in E(\mu)$.

Claim 2. $\tilde{\mathbf{u}} = w\mu$ where $w \in \partial^+ H_\tau(\mu)$ and $\mathrm{Lip}(w) \leq \frac{1}{\tau^2}$

(proof) Set $\hat{\eta}_n = \sum_{i=n}^{l(n)} \lambda_i^n \eta_i$, then

$$\delta(\hat{\gamma}_n,\hat{\eta}_n) \leq \sum_{i=n}^{l(n)} \lambda_i^n \delta(\eta_i,\gamma_i) \leq \sup_{i\geq n} \delta(\eta_i,\gamma_i) \to 0 \quad as \quad n \to \infty$$

By definition,

$$\eta_i \in \Gamma_o(\mu,\tilde{\nu}_i) \quad \mathbf{H}(\tau,\mu,\tilde{\nu}_i) = H_\tau(\mu) \quad \eta_i = (Id \times \nabla\psi_i)_{\#\mu}$$

Set $\bar{\psi}_n = \sum_{i=n}^{l(n)} \lambda_i^n \psi_i$ then $\bar{\psi}_n$ is convex and $-\frac{I}{\tau^2} \leq \nabla(\frac{Id-\nabla\bar{\psi}_n}{\tau}) \leq \frac{I}{\tau^2}$. We extract a subsequence $\{\bar{\psi}_{n_k}\}$ such that

$$\frac{Id-\nabla\bar{\psi}_{n_k}}{\tau} \xrightarrow{unif} \frac{Id-\nabla\bar{\psi}}{\tau}, \qquad \frac{Id-\nabla\psi_{n_k}}{\tau} \xrightarrow{unif} \frac{Id-\nabla\bar{\psi}}{\tau} \tag{69}$$

on the bounded set$B_R(0)$. Since $\frac{Id-\nabla\psi_i}{\tau} \in \partial^+ H_\tau(\mu)$ for all i and $\partial^+ H_\tau(\mu)$ is convex, we have $\frac{Id-\nabla\bar{\psi}_n}{\tau} \in \partial^+ H_\tau$ for all n. The uniform convergence in equation (69) implies $\frac{Id-\nabla\bar{\psi}}{\tau} \in \partial^+ H_\tau$. By abusing notation, we rename $\{\bar{\psi}_{n_k}\}$ by $\{\bar{\psi}_n\}$

So, for any continuous function F with compact support, we have

$$\begin{aligned}
\int_{\mathbb{R}^D} F \cdot d\tilde{\mathbf{u}} &= \lim_{n\to\infty} \sum_{i=n}^{l(n)} \lambda_i^n \int_{\mathbb{R}^D} F \cdot w_i d\mu_i \\
&= \lim_{n\to\infty} \sum_{i=n}^{l(n)} \lambda_i^n \int_{\mathbb{R}^D} F(x) \cdot \frac{x-y}{\tau} d\gamma_i(x,y) \\
&= \lim_{n\to\infty} \int_{\mathbb{R}^D\times\mathbb{R}^D} F(x) \cdot \frac{x-y}{\tau} d\hat{\gamma}_n(x,y)
\end{aligned}$$

Since we know $\delta(\hat{\gamma}_n, \hat{\eta}_n) \to 0$ and supports of μ_n and ν_n are uniformly bounded,

$$\begin{aligned}
\lim_{n\to\infty} \int_{\mathbb{R}^D\times\mathbb{R}^D} F(x)\cdot \frac{x-y}{\tau} d\hat{\gamma}_n(x,y) &= \lim_{n\to\infty} \int_{\mathbb{R}^D\times\mathbb{R}^D} F(x)\cdot \frac{x-y}{\tau} d\hat{\eta}_n(x,y) \\
&= \lim_{n\to\infty} \sum_{i=n}^{l(n)} \lambda_i^n \int_{\mathbb{R}^D} F(x)\cdot \frac{x-y}{\tau} d\eta_i(x,y)
\end{aligned}$$

Since $\eta_i \in E(\mu)$, we combine this with (69) to conclude

$$\begin{aligned}
\lim_{n\to\infty} \sum_{i=n}^{l(n)} \lambda_i^n \int_{\mathbb{R}^D} F(x)\cdot \frac{x-y}{\tau} d\eta_i(x,y) &= \lim_{n\to\infty} \sum_{i=n}^{l(n)} \lambda_i^n \int_{\mathbb{R}^D} F(x)\cdot \frac{x-\nabla\psi_i(x)}{\tau} d\mu(x) \\
&= \lim_{n\to\infty} \int_{\mathbb{R}^D} F(x)\cdot \frac{x-\nabla\bar{\psi}_n(x)}{\tau} d\mu(x) \\
&= \int_{\mathbb{R}^D} F(x)\cdot \frac{x-\nabla\bar{\psi}(x)}{\tau} d\mu(x)
\end{aligned}$$

□

Recall that Hamiltonian $H : \mathcal{M} \to (-\infty, \infty]$ is given by

$$H(\mu) = -\sup_{(v,B)\in\mathcal{C}} \{\int v d\mu + B\} \tag{70}$$

where $\mathcal{C} \subset \{(v, B); \frac{\lambda|x|^2}{2} - v(x) \text{ is convex}, B \in (-\infty, \infty)\}$ and λ is a given real number. Also recall that we imposed *" locally uniform L_1 bound"* condition on $\mathcal{C}$:

(H1') *For every compact $K \subset \mathbb{R}^D$, there exists a constant $M(K)$ only depending on K such that $\int_K |v|dx \le M(K)$ for all $(v, B) \in \mathcal{C}$.*

By combining Lemma 2.3.2, Theorem 2.3.3 follows directly from a result in [2] but for the completeness we give a full proof.

Theorem 2.3.3. *Let $H : \mathcal{M} \to (-\infty, \infty]$ be defined by equation (70) and let H_τ be Moreau-Yosida approximation of H. Suppose $\mathcal{C}$ satisfies the (H1')" locally uniform L_1 bound" condition. If $\bar{\mu} \in \mathcal{M}$ has a bounded support, then we have a Hamiltonian flow satisfying*

$$\begin{cases} \frac{d}{dt}\mu_t^\tau + \nabla\cdot(Jv_t^\tau \mu_t^\tau) = 0, & t \in (0, T) \\ v_t^\tau \in \partial^+ H_\tau(\mu_t^\tau) \cap T_{\mu_t^\tau} M & \mathrm{Lip}(v_t^\tau) \le \frac{1}{\tau^2}, \quad \mu_0^\tau = \bar{\mu} \end{cases} \tag{71}$$

Proof. **Step 1 : Construction of a discrete solution**

We build discrete solutions $\mu^N_{t,\tau}$ satisfying :

$$\begin{cases} \frac{d}{dt}\mu^N_{t,\tau} + \nabla\cdot(Jv^N_{t,\tau}\mu^N_{t,\tau}) = 0, \qquad t\in(0,T) \\ v^N_{t,\tau} \in \partial^+ H_\tau(\mu^N_{t,\tau})\cap T_{\mu^N_{t,\tau}}\mathcal{M}, \quad t = nh, \quad n = 0,1,\cdots,N \end{cases}$$

with the following conditions :

(a) The Lipschitz constant of $t\to\mu^N_{t,\tau}$ is less than for some fixed C_o

(b) $W_2(\mu^N_{t,\tau},\bar\mu)\le C_oT$

(c) The delayed Hamiltonian equation

$$\frac{d}{dt}\mu^N_{t,\tau} + \nabla\cdot(w^N_{t,\tau}\mu^N_{t,\tau}) = 0 \tag{72}$$

holds in the sense of distributions in $(0,T)\times\mathbb{R}^D$.

We build first the solution in $[0,h]$ by setting $w^N_{0,\tau} = J\frac{Id-\bar\gamma^{\hat\mu}_{\bar\mu}}{\tau}$, where $\hat\mu\in J_\tau[\bar\mu]$ *i.e*

$$\begin{aligned} H_\tau(\bar\mu) &= \inf_{\nu\in P_2(\mathbb{R}^D)}\{\frac{1}{2\tau}W_2^2(\bar\mu,\nu)+H(\nu)\} \\ &= \frac{1}{2\tau}W_2^2(\bar\mu,\hat\mu)+H(\hat\mu) \end{aligned}\tag{73}$$

By Theorem 2.2.5, we have $\hat\mu = \nabla\psi_\#\bar\mu$ for some convex function ψ, that is $\bar\gamma^{\hat\mu}_{\bar\mu} = \nabla\psi$. We then set

$$\mu^N_{t,\tau} = (Id+tw^N_{0,\tau})_\#\bar\mu, \qquad w^N_{t,\tau} = \frac{(Id+tw^N_{0,\tau})_\#(w^N_{0,\tau}\bar\mu)}{\mu^N_{t,\tau}}, \qquad t\in[0,h]$$

We claim that $w^N_{t,\tau}$ is an admissible velocity field for $\mu^N_{t,\tau}$. Indeed, for any $\phi\in C_c^\infty(\mathbb{R}^D)$, we have

$$\begin{aligned} \frac{d}{dt}\int\phi d\mu^N_{t,\tau} &= \frac{d}{dt}\int\phi(Id+tw^N_{o,\tau})d\bar\mu \\ &= \int <\nabla\phi(x+tw^N_{o,\tau}), w^N_{o,\tau}> d\bar\mu = \int <\nabla\phi, w^N_{t,\tau}> d\mu^N_{t,\tau} \end{aligned}$$

Since ϕ is arbitrary, equation (72) is fulfilled in $[0,h]$. Notice also that Lemma 1.2.19 gives

$$\int |w_{t,\tau}^N|^2 d\mu_{t,\tau}^N \leq \int |w_{0,\tau}^N|^2 d\mu_{0,\tau}^N = \int |w_{0,\tau}^N|^2 d\bar{\mu}$$

and Jensen's inequality gives

$$\int |w_{0,\tau}^N|^2 d\bar{\mu} = \int |\frac{Id - \bar{\gamma}_{\bar{\mu}}^{\hat{\bar{\mu}}}}{\tau}|^2 d\bar{\mu} \leq \frac{1}{\tau^2} W_2^2(\bar{\mu},\hat{\mu}) \tag{74}$$

Let's estimate $\frac{1}{\tau^2} W_2^2(\bar{\mu},\hat{\mu})$. From equation (73), we have for all $\nu \in P_2$

$$\frac{1}{2\tau} W_2^2(\bar{\mu},\hat{\mu}) + H(\hat{\mu}) \quad \leq \quad \frac{1}{2\tau} W_2^2(\bar{\mu},\nu) + H(\nu)$$

If we choose $\nu = \bar{\mu}$ then

$$\frac{1}{2\tau} W_2^2(\bar{\mu},\hat{\mu}) + H(\hat{\mu}) \quad \leq \quad H(\bar{\mu})$$

By lemma 2.2.3

$$\frac{1}{2\tau} W_2^2(\bar{\mu},\hat{\mu}) \quad \leq \quad H(\bar{\mu}) - H(\hat{\mu}) \leq C W_2(\bar{\mu},\hat{\mu})$$

and so, $W_2(\bar{\mu},\hat{\mu}) \leq 2\tau C$. We combine this with (74) to conclude

$$\int |w_{t,\tau}^N|^2 d\mu_{t,\tau}^N \leq \int |w_{0,\tau}^N|^2 d\mu_{0,\tau}^N \leq (2C)^2$$

Hence by the Proposition 1.2.15, the Lipschitz constant of $t \to \mu_{t,\tau}^N$ in $[0,h]$ is bounded by $2C$. We can repeat this process, setting $w_{h,\tau}^N = J \frac{Id - \bar{\gamma}_{\mu_{h,\tau}}^{\hat{\mu}_{h,\tau}}}{\tau}$, where

$$\begin{aligned} H_\tau(\mu_{h,\tau}) &= \inf_{\nu \in \mathcal{M}} \{ \frac{1}{2\tau} W_2^2(\mu_{h,\tau},\nu) + H(\nu) \} \\ &= \frac{1}{2\tau} W_2^2(\mu_{h,\tau},\hat{\mu_{h,\tau}}) + H(\hat{\mu_{h,\tau}}) \end{aligned}$$

and introduce the following extension on $(h, 2h]$

$$\mu_{t,\tau}^N = (Id + (t-h) w_{h,\tau}^N)_\# (\mu_{h,\tau}^N) \qquad w_{t,\tau}^N = \frac{(Id + (t-h) w_{h,\tau}^N)_\# (w_{h,\tau}^N \mu_{h,\tau}^N)}{\mu_{t,\tau}^N}$$

As before, we have

$$\int |w^N_{t,\tau}|^2 d\mu^N_{t,\tau} \leq \int |w^N_{h,\tau}|^2 d\mu^N_{h,\tau} \leq (2C)^2$$

and for any $\phi \in C^\infty_c(\mathbb{R}^D)$, we have

$$\begin{aligned} \frac{d}{dt}\int \phi d\mu^N_{t,\tau} = \frac{d}{dt}\int \phi(Id + (t-h)w^N_{h,\tau})d\mu^N_{h,\tau} &= \int < \nabla\phi(x+(t-h)w^N_{h,\tau}), w^N_{h,\tau} > d\mu^N_{h,\tau} \\ &= \int < \nabla\phi, w^N_{t,\tau} > d\mu^N_{t,\tau} \end{aligned}$$

For $t \in [h, 2h]$, the Lipschitz constant of $t \to \mu^N_{t,\tau}$ is bounded by $2C$ and the continuity equation (72) holds in $(h, 2h]$. By iteration, we have obtained $t \to \mu^N_{t.\tau}$ satisfying ;

(i). For any $\phi \in C^\infty_c(\mathbb{R}^D)$,

$$\frac{d}{dt}\int \phi d\mu^N_{t,\tau} = \int < \nabla\phi, w^N_{t,\tau} > d\mu^N_{t,\tau} \quad for \quad t \neq 0, h, 2h, \cdots, Nh \tag{75}$$

Since equation (75) holds point wise sense in (0,T) except finite points,(75) holds in the distribution sense in (0,T). which gives equation (72) is true in $\mathbb{R}^D \times (0, T)$

(ii). $\mu^N_{t.\tau}$ have uniformly bounded supports and $t \to \mu^N_{t.\tau}$ has a lipshchitz constant bounded by $2C$

So we have that $t \to \mu^N_{t,\tau}$ is a solution of equation (72) with Lipschitz bound $C_0 = 2C$

Step 2. $N \to \infty$

By (ii), $t \to \mu^N_{t,\tau}$ are equi-bounded in $\mathcal{M}$ and equi-Lipschitz continuous. Furthermore we have a uniform Lipshcitz bound C_0 which is independent of τ and N. Since $\mu^N_{t.\tau}$ have uniformly bounded supports, we may assume(up to a subsequence) that $\mu^N_{t,\tau} \to \mu^\tau_t$ in W_2 as $N \to \infty$. This gives that the C_0- Lipschitz bound independent of τ in (a) and the distance bound in (b) are preserved in the limit.

It remains to show that μ^τ_t solves the equation (71). To this aim, taking into account Lemma 1.2.20 and possibly extracting a subsequence we can assume that there exist $w^\tau_t \in L^2(\mu^\tau_t)$, with $||w^\tau_t||_{L^2(\mu^\tau_t)} \leq C_0$ for a.e t, such that the space-time measures $\{w^N_{t,\tau}\mu^N_{t,\tau}dt\}^\infty_{N=1}$ weak* converge to $w^\tau_t\mu^\tau_t dt$. We have to show that

$w_t^\tau = Jv_t^\tau$ for some $v_t^\tau \in \partial^+ H_\tau(\mu_t^\tau) \cap T_{\mu_t^\tau}\mathcal{M}$. To this aim, notice that for all $\psi \in C_c^\infty(0,T), \phi \in C_c^\infty(\mathbb{R}^D;\mathbb{R}^D)$ we have

$$\lim_{N\to\infty} \int_0^T \psi(t) < \phi, w_{t,\tau}^N \mu_{t,\tau}^N > dt = \int_0^T \psi(t) < \phi, w_t^\tau \mu_t^\tau > dt$$

For ϕ fixed, this means that the maps $t \to < \phi, w_{t,\tau}^N \mu_{t,\tau}^N >$ weakly converge in $L^2(0,T)$ to $< \phi, w_t^\tau \mu_t^\tau >$. Lemma 1.2.18 gives

$$< \phi, w_t^\tau \mu_t^\tau > \leq \limsup_{N\to\infty} < \phi, w_{t,\tau}^N \mu_{t,\tau}^N > \tag{76}$$

for a.e $t \in (0,T)$. Let $\{\phi_n\}_{n=1}^\infty$ be a countable dense subset of $C_c^\infty(\mathbb{R}^D;\mathbb{R}^D)$ in the sup norm. Since $C_c^\infty(\mathbb{R}^D;\mathbb{R}^D)$ is dense subset of $C_c(\mathbb{R}^D;\mathbb{R}^D)$ in the sup norm, $\{\phi_n\}$ is dense in $C_c(\mathbb{R}^D;\mathbb{R}^D)$ and $C_0(\mathbb{R}^D;\mathbb{R}^D)$(the closure, in the sup norm of $C_c(\mathbb{R}^D;\mathbb{R}^D)$). Let $A_n \subset (0,T)$ be such that equation (76) holds with ϕ_n for $t \in (0,T)\backslash A_n$ and A_n is Lebesgue negligible. Define $\mathcal{N} = \cup A_n$ then $\mathcal{N}$ is Lebesgue negligible and if $t \in (0,T)\backslash\mathcal{N}$ then equation (76) holds for all ϕ_n. Since $\{\phi_n\}_{n=1}^\infty$ is dense subset of $C_0(\mathbb{R}^D;\mathbb{R}^D)$, equation (76) holds for all $\phi \in C_0(\mathbb{R}^D;\mathbb{R}^D)$ and $t \in (0,T)\backslash\mathcal{N}$.

Now, fix $t \in (0,T)\backslash\mathcal{N}$ where equation (76) holds for all $\phi \in C_0(\mathbb{R}^D;\mathbb{R}^D)$ and apply Hahn-Banach theorem to obtain that

$$w_t^\tau \mu_t^\tau \in \cap_{M=1}^\infty \bar{co}\{w_{t,\tau}^N \mu_{t,\tau}^N : N \geq M\}$$

where $\bar{co}$ is the closed convex hull with respect to the weak* topology on $C_c(\mathbb{R}^D;\mathbb{R}^D)$. Indeed, fix M and assume by contradiction that $w_t^\tau \mu_t^\tau$ does not belong to $\bar{co}\{w_{t,\tau}^N \mu_{t,\tau}^N : N \geq M\}$. Then we can strongly separate $w_t^\tau \mu_t^\tau$ and $\bar{co}\{w_{t,\tau}^N \mu_{t,\tau}^N : N \geq M\}$ by a continuous linear functional induced by some function $\phi \in C_c(\mathbb{R}^D;\mathbb{R}^D)$ to obtain a contradiction with equation (76). As

$$\begin{aligned} w_{t,\tau}^N \mu_{t,\tau}^N &= \big(Id + (t-[Nt]/N) w_{[Nt]/N,\tau}^N\big)_\# (w_{[Nt]/N,\tau}^N \mu_{[Nt]/N,\tau}^N) \\ &= \big(Id + (t-[Nt]/N) J \frac{Id - \bar{\gamma}_{\mu_{[Nt]/N,\tau}^N}^{\hat{\mu}_{[Nt]/N,\tau}^N}}{\tau}\big)_\# \big(J \frac{Id - \bar{\gamma}_{\mu_{[Nt]/N,\tau}^N}^{\hat{\mu}_{[Nt]/N,\tau}^N}}{\tau} \mu_{[Nt]N,\tau}^N\big) \end{aligned}$$

We obtain also that

$$w_t^\tau \mu_t^\tau \in \cap_{M=1}^\infty \bar{co}\{J \frac{Id - \bar{\gamma}_{\mu^N_{[Nt]/N,\tau}}^{\hat{\mu}^N_{[Nt]/N,\tau}}}{\tau} \mu^N_{[Nt]N,\tau} : N \geq M\}$$

Hence Lemma 2.3.2 gives that $w_t^\tau \mu_t^\tau = J v_t^\tau \mu_t^\tau$ for some $v_t^\tau = \frac{Id - \bar{\gamma}_{\mu_t^\tau}^{\nu_t^\tau}}{\tau} \in \partial^+ H(\mu_t^\tau) \cap T_{\mu_t^\tau}\mathcal{M}$ where $\nu_t^\tau \in J[\mu_t^\tau]$ and $\mathrm{Lip}(v_t^\tau) \leq \frac{1}{\tau^2}$.

So we have obtained a Hamiltonian system satisfying the equation (71). □

2.3.1 Stability of Hamiltonian flows

We impose a condition on H :

(H2') If $W_2(\mu_n, \mu) \to 0$, then

$$\bigcap_{m=1}^\infty \bar{co}\{v_n \mu_n : v_n \in \partial_- H(\mu_n) \cap T_{\mu_n}\mathcal{M}, n \geq m\} \subset \{w\mu : w \in \partial_- H(\mu) \cap T_\mu \mathcal{M}\} \quad (77)$$

Theorem 2.3.4. *Let $H : \mathcal{M} \to (-\infty, \infty]$ be defined by equation (70) and $\bar{\mu} \in \mathcal{M}$ has a bounded support. For each $\tau \in (0,1)$, let μ^τ be the solution of equation (71) in Theorem 2.3.3. Suppose H satisfies condition $(H2')$. Let τ converge to 0 then $\{\mu^\tau\}_{\tau>0}$ converges to a Hamiltonian flow μ which satisfies*

$$\begin{cases} \frac{d}{dt}\mu_t + \nabla \cdot (J v_t \mu_t) = 0, & t \in (0,T) \\ v_t \in \partial_- H(\mu_t) \cap T_{\mu_t}\mathcal{M} \quad a.e \quad t \in (0,T), & \mu_0 = \bar{\mu} \end{cases} \quad (78)$$

Proof. By step 2 in the proof of Theorem 2.3.3, $t \to \mu_t^\tau$ are equi-bounded in $\mathcal{M}$ and equi-Lipschitz continuous. Hence we may assume that $\mu_t^{\tau_n} \to \mu_t$ in W_2 as $\tau_n \to 0$ for some sequence τ_n. This gives that the C_0- Lipschitz bound in (a) and the distance bound in (b) are preserved in the limit. By the same reasoning as in step 2 in the proof of Theorem 2.3.3, $\{J v_t^{\tau_n} \mu_t^{\tau_n} dt\}_{n=1}^\infty$ weak* converges to $J v_t \mu_t dt$. So μ_t satisfies

$$\begin{cases} \frac{d}{dt}\mu_t + \nabla \cdot (J v_t \mu_t) = 0, & t \in (0,T) \\ \mu_0 = \bar{\mu} \end{cases}$$

It remains to show $v_t \in \partial_- H(\mu_t) \cap T_{\mu_t} M$. As in step 2 in the proof of Theorem 2.3.3, for a.e $t \in (0,T)$

$$v_t \mu_t \in \bigcap_{M=1}^{\infty} \bar{co}\{v_t^{\tau_n} \mu_t^{\tau_n} : n \geq M\}$$

From Lemma 2.1.6, we have $v_t \mu_t \in \bigcap_{M=1}^{\infty} \bar{co}\{\xi_t^{\tau_n} \nu_t^{\tau_n} : n \geq M\}$, where $\nu_t^{\tau} \in J_{\tau}[\mu_t^{\tau}]$ and $\xi_t^{\tau_n} = \frac{\bar{\gamma}_{\nu_t^{\tau_n}}^{\mu_t^{\tau_n}} - Id}{\tau_n} \in \partial_- H(\nu_t^{\tau_n}) \cap T_{\nu_t^{\tau_n}} \mathcal{M}$. From (H2') condition, we get $v_t \in \partial_- H(\mu_t) \cap T_{\mu_t} \mathcal{M}$. So we have

$$\begin{cases} \frac{d}{dt}\mu_t + \nabla \cdot (J v_t \mu_t) = 0, & t \in (0,T) \\ v_t \in \partial_- H_{\tau}(\mu_t) \cap T_{\mu_t}\mathcal{M} \quad a.e \quad t \in (0,T) & \mu_0 = \bar{\mu} \end{cases}$$

□

2.3.2 Examples

(i) Let H be a potential energy functional defined as follows

$$H(\mu) = \int v(x) d\mu(x) \tag{79}$$

where $v : \mathbb{R}^D \to (-\infty, \infty]$ be a C^1, $\lambda-$ convex function. Then it trivially satisfies (H1')*" locally uniform L_1 bound"* condition and (H2') also holds since $\partial_- H(\mu) = \nabla v$ for all μ.

(ii) Let $H : \mathcal{M} \to (\infty, \infty)$ be a Hamiltonian defined as follows

$$H(\mu) = -\frac{1}{2} W_2^2(\mu, m_0) \tag{80}$$

where $m_0 \in \mathcal{M}$ has a bounded support. Then by the duality formula 4, we have

$$\frac{1}{2} W_2^2(m_0, \mu) = \sup_{u \oplus v \leq c} \{ \int u dm_0 + \int v d\mu \} = \sup_{(u,v) \in \mathcal{V}} \{ \int v d\mu + \int u dm_0 \} \tag{81}$$

where $\mathcal{V}$ be the set of pairs (u, v), such that $x \to |x|^2/2 - u(x), |x|^2/2 - v(x)$ are Legendre transform of each other and so, are convex. Let

$$\mathcal{C} := \big\{ (v, B) | \frac{|x|^2}{2} - v(x) = \phi^*, \quad v(0) = 0, \quad B = \int \frac{|y|^2}{2} - \phi(y) dm_0(y) \big\} \tag{82}$$

where $\phi \equiv \infty$ outside the convex hull of the support of m_0. As a consequence, if $(c, B) \in \mathcal{C}$ then $(Id - \nabla v) \subset B_R(0)$ where $B_R(0)$ is any ball containing the support of m_0. Thus

$$|v(x)| \leq |v(0)| + \text{Lip}(v)|x| \leq (1 + R)|x| \tag{83}$$

So equations (81) and (82) gives

$$H(\mu) = -\sup_{(v,B)\in\mathcal{C}} \{\int v d\mu + B : B = \int u dm_0\} \tag{84}$$

and if μ has a bounded support then equation (83) says $\mathcal{C}$ satisfies the *" locally uniform L_1 bound"* condition in Lemma 2.2.3. (H2') is satisfied from the fact $\partial_- H(\mu) \cap T_\mu\mathcal{M} = \{\bar{\gamma} - Id : \gamma \in \Gamma_0(\mu, m_0)\}$ together the Lemma 7.6 in [2].

Remark 2.3.5. As another interesting Hamiltonian, suppose $H : \mathcal{M} \to (\infty, \infty)$ is given by an interaction energy as follows

$$H(\mu) = \iint_{\mathbb{R}^D \times \mathbb{R}^D} w(x - y) d\mu(x) d\mu(y) \tag{85}$$

where $w : \mathbb{R}^D \to \mathbb{R}$ is given.

We expect we can do analysis similar to that in subsections 2.2.1 and 2.2.2 to characterize the minimizers of Moreau-Yosida approximation equation (19) by showing

$$H_\tau(\mu) = \iint_{\mathbb{R}^D \times \mathbb{R}^D} w_\tau(x - y) d\mu(x) d\mu(y)$$

In this case, $\nabla w_\tau * \mu_\tau \in \partial^+ H_\tau(\mu_\tau) \cap T_{\mu_\tau}\mathcal{M}$ and we let τ converge to 0 and might get the stability result $\nabla w * \mu \in \partial_- H(\mu) \cap T_\mu\mathcal{M}$ which gives the Hamiltonian flow satisfying equation (78).

2.4 Uniqueness of Hamiltonian flows

Lemma 2.4.1. *For $i = 1, 2$, let $\mu^i \in AC_2(0, 1 : \mathcal{M})$ and $v^i \in L^2(\mu^i)$ be velocity for μ^i i.e*

$$\partial_t \mu^i + \nabla \cdot (\mu^i v^i) = 0$$

Set $a(t) = \frac{1}{2}W_2^2(\mu_t^1, \mu_t^2)$ *then* $a \in W^{1,2}(0,1)$ *and*

$$\dot{a} \le \int < v_t^1(x) - v_t^2(y), x - y > d\gamma_t(x,y) \quad \gamma_t \in \Gamma_o(\mu_t^1, \mu_t^2) \tag{86}$$

Proof. Let w^i be the velocity of minimal norm of μ^i. We know there exists a Lebesgue measure zero set $\mathcal{N} \subset [0,1]$ such that for $t \in [0,1] \setminus \mathcal{N}$, we have

$$\lim_{h\to 0}(\pi^1 \times \frac{\pi^2 - \pi^1}{h})_{\#}\gamma_h^i = (Id \times w_t^i)_{\#}\mu_t^i \quad \forall \quad \gamma_h^i \in \Gamma_o(\mu_t^i, \mu_{t+h}^i) \tag{87}$$

and

$$\lim_{h\to 0}\frac{W_2(\tilde{\mu}_{t+h}^1, \mu_{t+h}^i)}{h} = 0 \tag{88}$$

where $\tilde{\mu}_{t+h}^i = (Id + hw_t^i)_{\#}\mu_t^i$. Define $b(t) := W_2(\mu_t^1, \mu_t^2)$ then

$$\begin{aligned} b(t+h) - b(t) &= W_2(\mu_{t+h}^1, \mu_{t+h}^2) - W_2(\mu_t^1, \mu_t^2) \\ &= W_2(\mu_{t+h}^1, \mu_{t+h}^2) - W_2(\mu_{t+h}^1, \mu_t^2) + W_2(\mu_{t+h}^1, \mu_t^2) - W_2(\mu_t^1, \mu_t^2) \\ &\le W_2(\mu_{t+h}^2, \mu_t^2) + W_2(\mu_{t+h}^1, \mu_t^1) \\ &\le \int_t^{t+h} F(\tau)d\tau \end{aligned}$$

where $F(\tau) = ||v_\tau^1||_{\mu_\tau^1} + ||v_\tau^2||_{\mu_\tau^2}$. Notice that $b \in L^\infty(0,1)$ and $F \in L^2(0,1)$. This implies $b \in H^1(0,1)$ and $\dot{a}(t) = b(t)\dot{b}(t) \in L^2 \Rightarrow a \in H^1(0,1)$

$$\begin{aligned} \lim_{h\to 0}\frac{b^2(t+h) - b^2(t)}{h} &= \lim_{h\to 0}\frac{W_2^2(\mu_{t+h}^1, \mu_{t+h}^2) - W_2^2(\mu_t^1, \mu_t^2)}{h} \\ &= \lim_{h\to 0}\frac{W_2^2(\mu_{t+h}^1, \mu_{t+h}^2) - W_2^2(\tilde{\mu}_{t+h}^1, \mu_{t+h}^2)}{h} \\ &+ \lim_{h\to 0}\frac{W_2^2(\tilde{\mu}_{t+h}^1, \mu_{t+h}^2) - W_2^2(\tilde{\mu}_{t+h}^1, \tilde{\mu}_{t+h}^2)}{h} \\ &+ \lim_{h\to 0}\frac{W_2^2(\tilde{\mu}_{t+h}^1, \tilde{\mu}_{t+h}^2) - W_2^2(\mu_t^1, \mu_t^2)}{h} \end{aligned} \tag{89}$$

First,

$$\begin{aligned} &\lim_{h\to 0}\left|\frac{W_2^2(\mu_{t+h}^1, \mu_{t+h}^2) - W_2^2(\tilde{\mu}_{t+h}^1, \mu_{t+h}^2)}{h}\right| \\ &\le \lim_{h\to 0}\frac{W_2(\mu_{t+h}^1, \tilde{\mu}_{t+h}^1)}{h}(W_2(\mu_{t+h}^1, \mu_{t+h}^2) + W_2(\tilde{\mu}_{t+h}^1, \mu_{t+h}^2)) = 0 \end{aligned} \tag{90}$$

Similarly, we have

$$\lim_{h\to 0}\frac{W_2^2(\tilde{\mu}^1_{t+h},\mu^2_{t+h})-W_2^2(\tilde{\mu}^1_{t+h},\tilde{\mu}^2_{t+h})}{h}=0 \tag{91}$$

and

$$\begin{aligned}&\lim_{h\to 0}\frac{W_2^2(\tilde{\mu}^1_{t+h},\tilde{\mu}^2_{t+h})-W_2^2(\mu^1_t,\mu^2_t)}{h} \\ &\le \lim_{h\to 0}\frac{1}{h}\int |(x+hw_t^1(x))-(y+hw_t^2(y))|^2-|x-y|^2 d\gamma_t(x,y) \\ &= 2\int < w_t^1(x)-w_t^2(y), x-y > d\gamma_t(x,y)\end{aligned} \tag{92}$$

Notice that

$$\begin{aligned}\int < w_t^1(x), x-y > d\gamma_t(x,y) &= \int < w_t^1(x), x-\bar{\gamma}_t(x) > d\mu_t^1(x) \\ &= \int < v_t^1(x), x-\bar{\gamma}_t(x) > d\mu_t^1(x) \\ &= \int < v_t^1(x), x-y > d\gamma_t(x,y)\end{aligned} \tag{93}$$

Similarly we have

$$\int < w_t^2(y), x-y > d\gamma_t(x,y) = \int < v_t^2(y), x-y > d\gamma_t(x,y) \tag{94}$$

Combining (92)(93) and (94), we have

$$\lim_{h\to 0}\frac{W_2^2(\tilde{\mu}^1_{t+h},\tilde{\mu}^2_{t+h})-W_2^2(\mu^1_t,\mu^2_t)}{h} \le 2\int < v_t^1(x)-v_t^2(y), x-y > d\gamma_t(x,y) \tag{95}$$

We combine (89),(90),(91) and (95) to get (86). This concludes proof. □

Corollary 2.4.2. *For $i=1,2$, let $\mu^i \in AC_2(0,1:\mathcal{M})$ and $v^i \in L^2(\mu^i)$ be velocity for μ^i i.e*

$$\partial_t\mu^i + \nabla\cdot(\mu^i v^i) = 0$$

Suppose

$$\int |v_t^1(x)-v_t^2(y)| d\gamma_t(x,y) \le CW_2^2(\mu_t^1,\mu_t^2)\big(1+\log\frac{1}{W_2^2(\mu_t^1,\mu_t^2)}\big) \tag{96}$$

for fixed nonnegative constant C and $\mu_0^1=\mu_0^2$ then $\mu_t^1=\mu_t^2$ for all $t\in[0,1]$.

Proof. Let $a(t) = \frac{1}{2}W_2^2(\mu_t^1, \mu_t^2)$ then (86) gives

$$\dot{a}(t) \le \sqrt{C} W_2^2(\mu_t^1, \mu_t^2) \sqrt{1 + \log \frac{1}{W_2^2(\mu_t^1, \mu_t^2)}}$$
$$\le \tilde{C} a(t) \big(1 + \log \frac{1}{W_2^2(\mu_t^1, \mu_t^2)}\big) \tag{97}$$

Since $a(0) = 0$, this together with (97) concludes proof. $\square$

In general, it is hard to expect conditions like (96) on arbitrary velocity. Furthermore with minimal norm condition on velocity, we have an example of Hamiltonian system which has multiple solutions (cfr. Example 2.4.4). So we restrict Hamiltonian system on some subset $\mathcal{Q}$ of $\mathcal{M}$.

Corollary 2.4.3. *Let $\mathcal{Q} \subset \mathcal{M}$ and suppose $H : \mathcal{M} \to (-\infty, \infty]$ is subdifferentiable on $\mathcal{Q}$.*

(i) Suppose H satisfies, for all $\mu, \nu \in \mathcal{Q}$

$$\int |v_\nu - v_\mu| d\gamma(x, y) \le C W_2^2(\mu, \nu) \big(1 + \log \frac{1}{W_2^2(\mu, \nu)}\big)$$

where v_μ and is the minimal subdifferential of H at μ in the sense of $v_\mu \in \partial_- H(\mu)$ and $||v_\mu|| \le ||\xi||$ for any $\xi \in \partial_- H(\mu)$. Then there exists at most one solution $\mu \in \mathcal{Q}$ of

$$\frac{d}{dt}\mu_t + \nabla \cdot (J v_t \mu_t) = 0, \qquad \mu_0 \in \mathcal{Q} \tag{98}$$

where v_t is the minimal subdifferential of H at μ_t.

(ii) Suppose H is differentiable and its exterior derivative i.e 1-form is regular in the sense of (126) on $\mathcal{Q}$ then there exists at most one solution $\mu \in \mathcal{Q}$ of

$$\frac{d}{dt}\mu_t + \nabla \cdot (J \nabla H(\mu_t)) = 0, \qquad \mu_0 \in \mathcal{Q} \tag{99}$$

Proof. (i) follows directly from Corollary 2.4.2. To prove (ii), it is enough to show ∇H satisfies (96). Actually, regularity assumption (126) implies

$$\int |\nabla H(\mu) - \nabla(\nu)|^2 d\gamma \le C W_2^2(\mu, \nu) \qquad \forall \quad \mu, \nu \in \mathcal{Q}$$

where $\gamma \in \Gamma_o(\mu, \nu)$. This concludes proof. □

Example 2.4.4. Non-uniqueness

Let $\nu_0 = (-1/2, 1/2)_{|\mathcal{L}^1}$, define $M : [0,1] \times \mathbb{R} \to \mathbb{R}$ by

$$M(t,x) = \begin{cases} x - \cos(t - \frac{1}{2})x & t \in [0, \frac{1}{2}] \\ 0 & t \in [\frac{1}{2}, \frac{3}{4}] \\ x - \cos(t - \frac{3}{4})x & t \in [\frac{3}{4}, 1] \end{cases}$$

We denote $M_t(\cdot) := M(t, \cdot)$ and define $\mu_t = (M_t \times \dot{M}_t)_\# \nu_0 \in \mathcal{P}_2(\mathbb{R}^2)$ then μ_t solves

$$\frac{d}{dt}\mu_t + \nabla \cdot (J v_t \mu_t) = 0, \qquad \mu_0 = (M_0 \times \dot{M}_0)_\# \nu_0 \tag{100}$$

where v_t is the minimal subdifferential of H at μ_t and H is defined by

$$H(\mu) = -\frac{1}{2} W_2^2(\mu, \nu) \tag{101}$$

where $\nu = \nu_0 \times \delta_0$.

Define $\tilde{M} : [0,1] \times \mathbb{R} \to \mathbb{R}$ by

$$\tilde{M}(t,x) = \begin{cases} x - \cos(t - \frac{1}{2})x & t \in [0, \frac{1}{2}] \\ 0 & t \in [\frac{1}{2}, 1] \end{cases}$$

Similarly, we denote $\tilde{M}_t(\cdot) := \tilde{M}(t, \cdot)$ and define $\tilde{\mu}_t = (\tilde{M}_t \times \dot{\tilde{M}}_t)_\# \nu_0 \in \mathcal{P}_2(\mathbb{R}^2)$ then $\tilde{\mu}_t$ also solves (100).

Proof. (1) Let us show μ_t satisfies

$$\frac{d}{dt}\mu_t + \partial_y(\xi \mu_t) = \partial_\xi((-\bar{\gamma}_\rho + id)\mu_t) \tag{102}$$

where $\rho(t,y) = \int_{\mathbb{R}} d\mu_t(y, \xi)$, $\gamma_\rho \in \Gamma_o(\rho, \nu_0)$.

To see the equation (102) is true, it is enough to show for a.e $t \in (0,1)$

$$\frac{d}{dt} \int_{\mathbb{R}^2} \phi(y,\xi) d\mu_t(y,\xi) = \int_{\mathbb{R}^2} \xi \partial_y \phi + (\bar{\gamma}_\rho(y) - y)\partial_\xi \phi d\mu_t(y,\xi) \tag{103}$$

for any $\phi \in C_c^\infty(\mathbb{R}^2)$. We may choose $\phi(y,\xi) = g(y)h(\xi)$ with $g, h \in C_c^\infty(\mathbb{R})$.

(a) For $t \in (0, \frac{1}{2})$, we have

$$\mu_t = \big([1-\cos(t-\frac{1}{2})]x, \sin(t-\frac{1}{2})x\big)_{\#}\nu_0,$$
$$\rho = \frac{1}{1-\cos(t-\frac{1}{2})}\big(-\frac{1-\cos(t-\frac{1}{2})}{2}, \frac{1-\cos(t-\frac{1}{2})}{2}\big)|_{\mathcal{L}^1}$$
$$\bar{\gamma}_\rho(y) = \int_{\mathbb{R}^2} x d\gamma(y,x) = \int_{\mathbb{R}^2} x\delta_{y/(1-\cos(t-\frac{1}{2}))} \times d\rho(y) = \frac{y}{1-\cos(t-\frac{1}{2})}$$

By direct computation, we have

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^2} g(y)h(\xi)d\mu_t(y,\xi) &= \frac{d}{dt}\int_{\mathbb{R}} g\big([1-\cos(t-\frac{1}{2})]x\big)h\big(\sin(t-\frac{1}{2})x\big)d\nu_0(x) \\ &= \int_{\mathbb{R}} g'\big([1-\cos(t-\frac{1}{2})]x\big)\sin(t-\frac{1}{2})xh\big(\sin(t-\frac{1}{2})x\big)d\nu_0(x) \\ &+ \int_{\mathbb{R}} g\big([1-\cos(t-\frac{1}{2})]x\big)h'\big(\sin(t-\frac{1}{2})x\big)\cos(t-\frac{1}{2})xd\nu_0(x)\end{aligned} \tag{104}$$

From the definition of μ_t, we have

$$\int_{\mathbb{R}^2} \xi g'(y)h(\xi)d\mu_t(y,\xi) = \int_{\mathbb{R}} \sin(t-\frac{1}{2})xg'\big([1-\cos(t-\frac{1}{2})]x\big)h\big(\sin(t-\frac{1}{2})x\big)d\nu_0(x) \tag{105}$$

and

$$\int_{\mathbb{R}^2} (\bar{\gamma}_\rho(y)-y)g(y)h'(\xi)d\mu_t(y,\xi) = \int_{\mathbb{R}} \cos(t-\frac{1}{2})xg\big([1-\cos(t-\frac{1}{2})]x\big)h'\big(\sin(t-\frac{1}{2})x\big)d\nu_0(x) \tag{106}$$

We combine (104),(105) and (106) to show (103)holds for all $t \in (0, \frac{1}{2})$.

(b) For $t \in (\frac{1}{2}, \frac{3}{4})$, we have

$$\mu_t = \delta_0 \times \delta_0, \quad \rho = \delta_0, \quad \bar{\gamma}_\rho(0) = \int_{R^2} xd\gamma_\rho(0,x) = \int_{-1/2}^{1/2} xdx = 0$$

So

$$\frac{d}{dt}\int_{\mathbb{R}^2} g(y)h(\xi)d\mu_t(y,\xi) = \frac{d}{dt}g(0)h(0) = 0 \tag{107}$$

and

$$\int_{\mathbb{R}^2} \xi g'(y)h(\xi)d\mu_t(y,\xi) = \int_{\mathbb{R}} 0g'(0)h(0)d\nu_0(x) = 0 \tag{108}$$

$$\int_{\mathbb{R}^2} (\bar{\gamma}_\rho(y) - y) g(y) h'(\xi) d\mu_t(y,\xi) = \int_{\mathbb{R}} 0 g(0) h'(0) d\nu_0(x) = 0 \tag{109}$$

From (107),(108) and (109) we have (103) is true for all $t \in (\frac{1}{2}, \frac{3}{4})$.

(c) Similarly, for $t \in (\frac{3}{4}, 1)$, we can show (103).

(2) Define $H(\mu) = -\frac{1}{2}W_2^2(\mu, \nu)$ where $\nu = \nu_0 \times \delta_0$ then

$$\begin{aligned} H(\mu) &= -\frac{1}{2}W_2^2(\mu^1, \nu_0) - \frac{1}{2}\int \xi^2 d\mu^2(\xi) \qquad \mu^1 = \pi^1_{\#}\mu, \quad \mu^2 = \pi^2_{\#}\mu \\ &=: H_1(\mu^1) + H_2(\mu^2) \end{aligned}$$

(a) For $t \in (0, \frac{1}{2})$, since $\mu_t = ([1 - \cos(t - \frac{1}{2})]x)_{\#}\nu_0$ is absolutely continuous with respect to the Lebesgue measure, we do the same calculus as in (112) with $\tilde{\mu} = (id, T^{\mu^1}_{\mu^1_t}, T^{\nu_0}_{\mu^1_t})_{\#}\mu^1_t$ then we have

$$\begin{aligned} H_1(\mu^1) - H_1(\mu^1_t) &\geq \int_{\mathbb{R}^3} < y - x; z - x > -\frac{(y-x)^2}{2} d\tilde{\mu}(x,y,z) \\ &= \int_{\mathbb{R}^3} < y - x; T^{\nu_0}_{\mu^t_1}(x) - x > d\gamma(x,y) - \frac{1}{2}W_2^2(\mu^t_1, \mu^1) \end{aligned} \tag{110}$$

This says $T^{\nu_0}_{\mu^1_t} - id \in \partial_- H_1(\mu^1_t)$ and it is easy to see $T^{\nu_0}_{\mu^1_t} = \bar{\gamma}_\rho$. For $\gamma \in \Gamma_o(\mu^2_t, \mu^2)$

$$\begin{aligned} H_2(\mu^2) - H_2(\mu^2_t) &= -\frac{1}{2}\int \eta^2 - \xi^2 d\gamma(\xi,\eta) \\ &= \int < -\xi; \eta - \xi > d\gamma(\xi,\eta) - \frac{1}{2}W_2^2(\mu^2, \mu^2_t) \end{aligned} \tag{111}$$

We combine (110) and (111) to get $(\bar{\gamma}_\rho(y) - y, -\xi) \in \partial_- H(\mu_t)$.

(b) For $t \in (\frac{1}{2}, \frac{3}{4})$, we define $\tilde{\mu} := \mu^1_t \times \mu^1 \times \nu_0$ then, since $\mu^1_t = \delta_0$, we have

$$\begin{aligned} H_1(\mu^1) - H_1(\mu^1_t) &= -\frac{1}{2}W_2^2(\mu^1, \nu_0) + \frac{1}{2}W_2^2(\mu^1_t, \nu_0) \\ &\geq \int_{\mathbb{R}^3} -\frac{(y-z)^2}{2} + \frac{(x-z)^2}{2} d\tilde{\mu}(x,y,z) \\ &= \int_{\mathbb{R}^3} < y - x; z - x > -\frac{(y-x)^2}{2} d\tilde{\mu}(x,y,z) \\ &= \int_{\mathbb{R}^2} < y; z > d\mu_1(y) \times \nu_0(z) - \frac{1}{2}W_2^2(\mu^1_t, \mu^1) \\ &= -\frac{1}{2}W_2^2(\mu^1_t, \mu^1) \end{aligned} \tag{112}$$

and

$$H_2(\mu^2) - H_2(\mu_t^2) = -\frac{1}{2}\int y^2 d\mu_2(y) = -\frac{1}{2}W_2^2(\mu^2, \mu_t^2) \tag{113}$$

From (112) we have $0 \in \partial_- H_1(\mu_t^1)$ and (113) gives $0 \in \partial_- H_2(\mu_t^2)$. This implies $(0,0) \in \partial_- H(\mu_t)$ which coincides with $(\bar{\gamma}_\rho(y) - y, -\xi)$ at $\mu_t = \delta_0 \times \delta_0$.

(c) For $t \in (\frac{1}{2}, 1)$, similarly we have $(\bar{\gamma}_\rho(y) - y, -\xi) \in \partial_- H(\mu_t)$.

It is well known that $\bar{\gamma}_\rho - id \in T_\rho$ which means there exist a sequence $\phi_n \in C_c^\infty$ such that $\nabla\phi_n \to \bar{\gamma}_\rho - id$ in $L^2(\rho)$ so we define $\psi_n(y,\xi) = \phi_n(y) + \xi^2/2$ then $\nabla\psi_n \to (\bar{\gamma}_\rho(y) - y, -\xi)$ in $L^2(\mu_t)$

From (1) and (2), we know (102) is equivalent to

$$\begin{cases} \frac{d}{dt}\mu_t + \nabla \cdot (Jv_t\mu_t) = 0, & t \in (0,1) \\ v_t \in \partial_- H(\mu_t) \cap T_{\mu_t}\mathcal{M} & \mu_0 = (M_0 \times \dot{M}_0)_\# \nu_0 \end{cases}$$

where $H(\mu) = -\frac{1}{2}W_2^2(\mu, \nu)$ with $\nu = \nu_0 \times \delta_0$

(3) Now, let's check the minimal norm property.

(a) For $t \in (\frac{1}{2}, \frac{3}{4})$, we have $v_t = (\bar{\gamma}_\rho(y) - y, -\xi) = (0,0)$ and trivially $(0,0)$ has minimal $L^2(\mu_t)$ norm among all subdifferentials of H.

(b) For $t \in (0, \frac{1}{2})$, we have $v_t = (\bar{\gamma}_\rho(y) - y, -\xi)$. Since μ_t^1 is absolutely continuous with respect to the Lebesgue measure, $\bar{\gamma}_\rho - id = T_{\mu_t^1}^{\nu_0} - id$ is the only element of $\partial H_1(\mu_1^t) \cap T_{\mu_t^1}$ (see for instance Proposition 4.3 in [2]) that means $\bar{\gamma}_\rho - id$ has minimal $L^2(\mu_t^1)$ norm among all subdifferentials of H_1 at μ_t^1. And Remark 10.4.3 in [3] says that $-id$ has the minimal $L^2(\mu_t^2)$ norm among all subdifferentials of H_2 at μ_t^2. So $v_t = (\bar{\gamma}_\rho(y) - y, -\xi)$ has the the minimal $L^2(\mu_t)$ norm among all subdifferentials of H at μ_t. Finally, (1),(2) and (3) says μ_t solves

$$\frac{d}{dt}\mu_t + \nabla \cdot (Jv_t\mu_t) = 0, \qquad \mu_0 = (M_0 \times \dot{M}_0)_\# \nu_0 \tag{114}$$

where v_t is the minimal subdifferential of H at μ_t. Similarly, we can show that $\tilde{\mu}_t$ also solves (114). □

CHAPTER III

CALCULUS OF DIFFERENTIAL FORMS ON THE WASSERSTEIN SPACE

As we saw in chapter II, many interesting classes of PDE's can be viewed as Hamiltonian flows on $\mathcal{M}$. Developing this idea ,however, requires a rigorous symplectic formalism for $\mathcal{M}$. In a joint work with Gangbo and Pacini [14], we provide the basis for a framework to define and to study Hamiltonian systems on $\mathcal{M}$ and we achieve two main goals. First, we developed a general theory of differential forms on $\mathcal{M}$. Next, we show that there exists a natural symplectic and Hamiltonian formalism for $\mathcal{M}$ which is compatible with this calculus of curves and forms. This chapter is part of [14]. One of our result is an analogue of Green's theorem for 1-forms(cfr. Theorems 3.2.3 and 3.2.33). A corollary of that is that every closed 1-form on $\mathcal{M}$ is exact(cfr. Corollary 3.2.35). Hence, the first cohomology group, in the sense of de Rham, vanishes. We point out the recent paper by Gangbo-Tudorascu [16] showing that if we replace $\mathbb{R}^D$ by π^D although $\mathcal{P}_2(\pi^D)$ remains a convex set its first de Rham cohomology is not trivial. Hence the fact that the first de Rham cohomology group of $\mathcal{M}$ is trivial is not a mere consequence of the convexity of $\mathcal{M}$.

3.1 Tangent and Cotangent bundles

3.1.1 Tangent space

Recall that $\mathcal{X}_c$ denote the space of compactly supported smooth vector fields on $\mathbb{R}^D$. And also recall, for given $\mu \in \mathcal{M}$, $T_\mu\mathcal{M}$ denote the closure of ∇C_c^∞ in $L^2(\mu)$ and we call it the *tangent space* of $\mathcal{M}$ at μ.

Definition 3.1.1. Given $\mu \in \mathcal{M}$ we define the *divergence operator*

$$div_\mu : \mathcal{X}_c \to (C_c^\infty)^*, \quad \langle div_\mu(X), f\rangle := -\int_{\mathbb{R}^D} df(X)\, d\mu.$$

Notice that the divergence operator is linear and that $\langle div_\mu(X), f\rangle \leq ||\nabla f||_\mu ||X||_\mu$. This proves that the operator div_μ extends to $L^2(\mu)$ by continuity; we will continue to use the same notation for the extended operator, so that $\mathrm{Ker}(div_\mu)$ is now a closed subspace of $L^2(\mu)$.

It follows from [3] Lemma 8.4.2 that, given any $\mu \in \mathcal{M}$, there is an orthogonal decomposition

$$L^2(\mu) = \overline{\nabla\, C_c^\infty}^\mu \oplus \mathrm{Ker}(div_\mu). \tag{115}$$

We will denote by

$$\pi_\mu : L^2(\mu) \to \overline{\nabla\, C_c^\infty}^\mu \tag{116}$$

the corresponding projection. Notice that each tangent space has a natural Hilbert space structure.

Remark 3.1.2. Decomposition 115 shows that $T_\mu\mathcal{M}$ can also be identified with the quotient space $L^2(\mu)/\mathrm{Ker}(div_\mu)$: the map π_μ provides a Hilbert space isomorphism between these two spaces.

Lemma 3.1.3. *If $\phi : \mathbb{R}^D \to \mathbb{R}^D$ is a Lipschitz map with Lipschitz constant $Lip\,\phi$ then $\phi_\# : \mathcal{M} \to \mathcal{M}$ is also a Lipschitz map with the same Lipschitz constant.*

Proof. Let $\mu, \nu \in \mathcal{M}$. Note that if $u(x) + v(y) \leq |x-y|^2$ for all $x, y \in \mathbb{R}^D$ then

$$u \circ \phi(a) + v \circ \phi(b) \leq |\phi(a) - \phi(b)|^2 \leq (Lip\,\phi)^2 |a-b|^2.$$

This, together with equation (4) yields

$$\int_{\mathbb{R}^D} u d\phi_\#\mu + \int_{\mathbb{R}^D} v d\phi_\#\nu = \int_{\mathbb{R}^D} u \circ \phi d\mu + \int_{\mathbb{R}^D} v \circ \phi d\nu \leq (Lip\,\phi)^2 W_2^2(\mu,\nu). \tag{117}$$

We maximize the expression at the left handside of equation (117) over the set of pairs (u, v) such that $u(x) + v(y) \leq |x-y|^2$ for all $x, y \in \mathbb{R}^D$. Then we use again equation (4) to conclude the proof.

□

The next results concern the lifted action of $\mathrm{Diff}_c(\mathbb{R}^D)$ on $T\mathcal{M}$.

Lemma 3.1.4. *For any $\mu \in \mathcal{M}$ and $\phi \in Diff_c(\mathbb{R}^D)$, the map $\phi_* : \mathcal{X}_c(\mathbb{R}^D) \to \mathcal{X}_c(\mathbb{R}^D)$ has a unique continuous extension $\phi_* : L^2(\mu) \to L^2(\phi_\# \mu)$. Furthermore $\phi_*\big(Ker(div_\mu)\big) \leq Ker(div_{\varphi_\# \mu})$. Thus ϕ_* induces a continuous map $\phi_* : T_\mu\mathcal{M} \to T_{\phi_\#\mu}\mathcal{M}$.*

Proof. Let $\mu \in \mathcal{M}$, $\phi \in \mathrm{Diff}_c(\mathbb{R}^D)$, $f \in C_c^\infty(\mathbb{R}^D)$ and let $X \in \mathrm{Ker}(div_\mu)$. If C_ϕ is the L^∞-norm of $\nabla\phi$ we have $||\phi_* X||_{\phi_\#\mu} \leq C_\phi ||X||_\mu$. Hence ϕ_* admits a unique continuous linear extension. Furthermore

$$\begin{aligned}
\int_{\mathbb{R}^D} \langle \nabla f, \varphi_* X\rangle d\varphi_\#\mu = \int_{\mathbb{R}^D} \langle \nabla f \circ \varphi, \varphi_* X \circ \varphi\rangle d\mu &= \int_{\mathbb{R}^D} \langle \nabla f \circ \varphi, \nabla\varphi X\rangle d\mu \\
&= \int_{\mathbb{R}^D} \langle (\nabla\varphi)^T \nabla f \circ \varphi, X\rangle d\mu \\
&= \int_{\mathbb{R}^D} \langle \nabla[f \circ \varphi], X\rangle d\mu = 0.
\end{aligned}$$

□

Lemma 3.1.5. *Let $\sigma \in AC_2(a, b; \mathcal{M})$ and let v be a velocity for σ. Let $\varphi \in Diff_c(\mathbb{R}^D)$. Then $t \to \varphi_\#(\sigma_t) \in AC_2(a, b; \mathcal{M})$ and $\varphi_* v$ is a velocity for $\varphi_\#\sigma$.*

Proof. If $a < s < t < b$, by Lemma 3.1.3, $W_2(\varphi_\#\sigma_t, \varphi_\#\sigma_s) \leq (Lip\,\varphi)\, W_2(\sigma_t, \sigma_s)$. Because $\sigma \in AC_2(a, b; \mathcal{M})$ one concludes that $d\varphi_\#(\sigma) \in AC_2(a, b; \mathcal{M})$. If $f \in C_c^\infty((a, b)\times \mathbb{R}^D)$ we have

$$\begin{aligned}
\int_a^b \int_{\mathbb{R}^D} \Big(\frac{\partial f_t}{\partial t} + df_t(\phi_* v_t)\Big) d(\varphi_\#\sigma_t) dt &= \int_a^b \int_{\mathbb{R}^D} \Big(\frac{\partial f_t}{\partial t} \circ \varphi + (df_t(\phi_* v_t)) \circ \varphi\Big) d\sigma_t dt \\
&= \int_a^b \int_{\mathbb{R}^D} \Big(\frac{\partial (f \circ \varphi)_t}{\partial t} + d(f \circ \varphi)_t(v_t)\Big) d\sigma_t dt \\
&= 0.
\end{aligned}$$

To obtain the last equality we have used that $(t, x) \to f(t, \varphi(x))$ is in $C_c^\infty((a, b) \times \mathbb{R}^D)$.

□

3.1.2 Differential forms on $\mathcal{M}$

Recall from Definition 1.2.12 that the tangent bundle $T\mathcal{M}$ of $\mathcal{M}$ is defined as the union of all spaces $T_\mu\mathcal{M}$, for $\mu \in \mathcal{M}$. We now define the *pseudo tangent bundle* $\mathcal{T}\mathcal{M}$ to be the union of all spaces $L^2(\mu)$. Analogously, the union of the dual spaces $T_\mu^*\mathcal{M}$ defines the *cotangent bundle* $T^*\mathcal{M}$; we define the *pseudo cotangent bundle* $\mathcal{T}^*\mathcal{M}$ to be the union of the dual spaces $L^2(\mu)^*$.

It is clear from the definitions that we can think of $T\mathcal{M}$ as a subbundle of $\mathcal{T}\mathcal{M}$. Decomposition 115 allows us also to define an injection $T^*\mathcal{M} \to \mathcal{T}^*\mathcal{M}$ by extending any covector $T_\mu\mathcal{M} \to \mathbb{R}$ to be zero on the complement of $T_\mu\mathcal{M}$ in $L^2(\mu)$. In this sense we can also think of $T^*\mathcal{M}$ as a subbundle of $\mathcal{T}^*\mathcal{M}$. The projections π_μ combine to define a surjection $\pi : \mathcal{T}\mathcal{M} \to T\mathcal{M}$. Likewise, restriction yields a surjection $\mathcal{T}^*\mathcal{M} \to T^*\mathcal{M}$.

Definition 3.1.6. A *1-form* on $\mathcal{M}$ is a section of the cotangent bundle $T^*\mathcal{M}$, *i.e.* a collection of maps $\mu \mapsto F_\mu \in T_\mu^*\mathcal{M}$. A *pseudo 1-form* is a section of the pseudo cotangent bundle $\mathcal{T}^*\mathcal{M}$.

Analogously, a *2-form* on $\mathcal{M}$ is a collection of alternating multilinear maps

$$\mu \mapsto \Lambda_\mu : T_\mu\mathcal{M} \times T_\mu\mathcal{M} \to \mathbb{R},$$

continuous for each μ in the sense that $|\Lambda_\mu(X_1, X_2)| \leq c_\mu \|X_1\|_\mu \cdot \|X_2\|_\mu$, for some $c_\mu \in \mathbb{R}$. A *pseudo 2-form* is a collection of continuous alternating multilinear maps

$$\mu \mapsto \bar{\Lambda}_\mu : L^2(\mu) \times L^2(\mu) \to \mathbb{R}.$$

For $k = 1, 2$ we let $\Lambda^k\mathcal{M}$ (respectively, $\bar{\Lambda}^k\mathcal{M}$) denote the space of k-forms (respectively, pseudo k-forms). We define a *0-form* to be a function $\mathcal{M} \to \mathbb{R}$.

For $k = 1, 2$, the continuity condition implies that any k-form is uniquely defined by its values on any dense subset of $T_\mu\mathcal{M}$ or $T_\mu\mathcal{M} \times T_\mu\mathcal{M}$. For instance we can take the dense subset defined by smooth gradient vector fields. The analogue is true for pseudo k-forms. Once again, using Decomposition 115 yields an injection $\Lambda^k\mathcal{M} \to \bar{\Lambda}^k\mathcal{M}$ and, by restriction, a surjection $\bar{\Lambda}^k\mathcal{M} \to \Lambda^k\mathcal{M}$. In this sense every pseudo k-form defines a natural k-form.

Since $T_\mu\mathcal{M}$ is a Hilbert space, by the Riesz representation theorem every 1-form Λ_μ on $T_\mu\mathcal{M}$ can be written $\Lambda_\mu(Y) = \int_{\mathbb{R}^D}\langle A_\mu, Y\rangle d\mu$ for a unique $A_\mu \in T_\mu\mathcal{M}$ and all $Y \in T_\mu\mathcal{M}$. The analogous fact is true also for pseudo 1-forms.

Example 3.1.7. Any $f \in C_c^\infty$ defines a function on $\mathcal{M}$, *i.e.* a 0-form, as follows:

$$F(\mu) := \int_{RD} f d\mu.$$

We will refer to these as the *linear* functions on $\mathcal{M}$, in that the natural extension to the space $(C_c^\infty)^*$ defines a function which is linear with respect to μ.

Any $\bar{A} \in \mathcal{X}_c$ defines a pseudo 1-form on $\mathcal{M}$ as follows:

$$\bar{\Lambda}_\mu(X) := \int_{\mathbb{R}^D}\langle \bar{A}, X\rangle d\mu.$$

We will refer to these as the *linear* pseudo 1-forms. Notice that if $\bar{A} = \nabla f$ for some $f \in C_c^\infty$ then $\bar{\Lambda}$ is actually a 1-form.

Any bounded field $B = B(x)$ on $\mathbb{R}^D$ of $D \times D$ matrices defines a *linear* pseudo 2-form via

$$\bar{B}(X,Y) := \int_{\mathbb{R}^D} B(X,Y) d\mu.$$

Recall that $\text{Diff}_c(\mathbb{R}^D)$ denote the set of diffeomorphisms of $\mathbb{R}^D$ with compact support, *i.e.* those which coincide with the identity map Id outside of a compact subset of $\mathbb{R}^D$. The action of $\text{Diff}_c(\mathbb{R}^D)$ on $\mathcal{M}$ can be lifted to forms and pseudo forms as follows.

Definition 3.1.8. For $k = 1, 2$, let $\bar{\Lambda}$ be a pseudo k-form on $\mathcal{M}$. Then any $\phi \in \text{Diff}_c(\mathbb{R}^D)$ defines a *pull-back* k-multilinear map $\phi^*\bar{\Lambda}$ on $\mathcal{M}$ as follows:

$$(\phi^*\bar{\Lambda})_\mu(X_1, \ldots, X_k) := \bar{\Lambda}_{\phi_{\#}\mu}(\phi_* X_1, \ldots, \phi_* X_k).$$

It is simple to check that $\phi^*\bar{\Lambda}$ is indeed continuous in the sense of Definition 3.1.6 and is thus a pseudo k-form.

It follows from Lemma 3.1.4 that the push-forward operation preserves Decomposition 115. This implies that the pull-back preserves the space of k-forms, *i.e.* the pull-back of a k-form is a k-form.

Definition 3.1.9. Let $F : \mathcal{M} \to \mathbb{R}$ be a function on $\mathcal{M}$. Suppose F is differentiable and let $\nabla_\mu F$ be the gradient vector at $\mu \in \mathcal{M}$ as defined in the definition 1.2.8.

If the gradient vector exists for every $\mu \in \mathcal{M}$, we can define the *differential* or *exterior derivative* of F to be the 1-form dF determined, for any $\mu \in \mathcal{M}$ and $Y \in T_\mu\mathcal{M}$, by $dF(\mu)(Y) := \int_{\mathbb{R}^D} \langle \nabla_\mu F, Y \rangle \, d\mu$. To simplify the notation we will sometimes write $Y(F)$ rather then $dF(Y)$.

Remark 3.1.10. Assume $F : \mathcal{M} \to \mathbb{R}$ is differentiable. Given $X \in \nabla C_c^\infty(\mathbb{R}^D)$, let ϕ_t denote the flow of X i.e $\dot{\phi}_t = X(\phi_t)$, $\phi_0 = Id$. Fix $\mu \in \mathcal{M}$.
(i) Set $\nu_t := (Id + tX)_{\#}\mu$. Then

$$F(\nu_t) = F(\mu) + t \int_{\mathbb{R}^D} \langle \nabla_\mu F, X \rangle d\mu + o(t).$$

(ii) Set $\mu_t := \phi_{t\#}\mu$. If $||\nabla_\mu F(\mu)||_\mu$ is bounded on compact subsets of $\mathcal{M}$ then

$$F(\mu_t) = F(\mu) + t \int_{\mathbb{R}^D} \langle \nabla_\mu F, X \rangle d\mu + o(t).$$

Proof. The proof of (i) is a direct consequence of equation (8) and of the fact that, if $r > 0$ is small enough, $\big(Id \times (Id + tX)\big)_{\#}\mu \in \Gamma_o(\mu, \nu_t)$ for $t \in [-r, r]$ which was shown in Remark 1.2.9.

To prove (ii), set

$$A(s,t) := (1-s)(Id+tX) + s\phi_t \qquad s \in [0,1]$$

Notice that $||\phi_t - Id - tX||_\mu \leq t^2||(\nabla X)X||_\infty$ and that $(s,t) \to m(s,t) := A(s,t)_\#\mu$ defines a continuous map of the compact set $[0,1]\times[-r,r]$ into $\mathcal{M}$. Hence the range of m is compact so $||\nabla_\mu F(\mu)||_\mu$ is bounded there by a constant C. It can be shown that F is Lipschitz on the range of m. Once we have this, let $\bar{\gamma}_t := \big((Id+tX)\times\phi_t\big)_\#\mu$. We have $\bar{\gamma}_t \in \Gamma(\nu_t,\mu_t)$ so $W_2(\mu_t,\nu_t) \leq ||\phi_t - Id - tX||_\mu = 0(t^2)$. We conclude that

$$|F(\nu_t) - F(\mu_t)| \leq CW_2(\mu_t,\nu_t) = 0(t^2).$$

This, together with (i), yields (ii).

It remains to prove that F is Lipschitz on the range of m.

Claim 1. If there is a compact set $\mathcal{K} \subset \mathcal{M}$ which contains the range of m and all the constant geodesic between any two elements of it, then F is Lipschitz on the range of m.

(proof) Let μ_0 and μ_1 be in the range of m and suppose that $t \to \mu_t$ is a constant speed geodesic between μ_0 and μ_1. Then

$$\begin{aligned} F(\mu_1) - F(\mu_0) = \int_0^1 \frac{d}{dt}F(\mu_t)dt &= \int_0^1 <\nabla F(\mu_t), \dot{\mu}_t>_{\mu_t} dt \\ &= \int_0^1 dt \int <\nabla F(\mu_t), v_t> d\mu_t \\ &\leq \int_0^1 dt||\nabla F(\mu_t)||||v_t|| \end{aligned}$$

Since $\mu_t \in \mathcal{K}$ and $\mathcal{K}$ is compact, $||\nabla F(\mu_t)||$ is bounded by a constant $C_\mathcal{K}$. So we have

$$F(\mu_1) - F(\mu_0) \leq C_\mathcal{K}\int_0^1 ||v_t||dt = C_\mathcal{K}W_2(\mu_0,\mu_1)$$

This concludes Claim 1.

Claim 2. There exists a compact set $\mathcal{K} \subset \mathcal{M}$ satisfying the assumption in the claim 1.

(proof) Let us denote the range of m by $\mathcal{R}$. Notice that $\mathcal{R} \subset \mathcal{M}$ is tight and $\mathcal{R}$ has uniformly integrable 2-moments i.e

$$\lim_{n\to\infty} \int_{\mathbb{R}^D \setminus B(0,n)} |x|^2 d\mu(x) = 0 \quad uniformly \quad w.r.t. \quad \mu \in \mathcal{R}$$

Define $\mathcal{C}$ to be the union of $\mathcal{R}$ and all the constant geodesic between any two elements of $\mathcal{R}$. It is easy to see that $\mathcal{C}$ is tight and $\mathcal{R}$ has uniformly integrable 2-moments. This implies $\mathcal{C}$ is relatively compact. Finally we define $\mathcal{K}$ to be the union of $\mathcal{C}$ and all the limit points of it. Then $\mathcal{K}$ is a set satisfying the assumption in the claim 1.

We combine claim 1 and claim 2 to get F is Lipschitz on the range of m.

□

Example 3.1.11. Fix $f \in C_c^\infty$ and let $F : \mathcal{M} \to \mathbb{R}$ be the corresponding linear function, as in Example 3.1.7. Then F is differentiable with gradient $\nabla_\mu F \equiv \nabla f$. Thus dF is a linear 1-form on $\mathcal{M}$. Viceversa, every linear 1-form Λ is *exact*. In other words, if $\Lambda_\mu(X) = \int_{\mathbb{R}^D} \langle A, X \rangle d\mu$ for some $A = \nabla f$ then $\Lambda = dF$ for $F(\mu) := \int_{\mathbb{R}^D} f \, d\mu$.

Definition 3.1.12. Let $\bar{\Lambda}$ be a pseudo 1-form on $\mathcal{M}$. We say that $\bar{\Lambda}$ is *differentiable* with *exterior derivative* $d\bar{\Lambda}$ if (i) for all $X \in \nabla C_c^\infty$, the function $\bar{\Lambda}(X)$ is differentiable and (ii)for all $X, Y \in \nabla C_c^\infty$

$$d\bar{\Lambda}(X,Y) := X\bar{\Lambda}(Y) - Y\bar{\Lambda}(X) - \bar{\Lambda}([X,Y]) \tag{118}$$

yields a well-defined pseudo 2-form $d\bar{\Lambda}$ on $\mathcal{M}$ (see Definition 1.2.8 for notation).

Example 3.1.13. Assume $\bar{\Lambda}$ is a linear pseudo 1-form, *i.e.* $\bar{\Lambda}(\cdot) = \int_{\mathbb{R}^D} \langle \bar{A}, \cdot \rangle d\mu$ for some $\bar{A} \in \mathcal{X}_c$. Then $\bar{\Lambda}$ is differentiable and $d\bar{\Lambda}(X,Y) = \int_{\mathbb{R}^D} \langle (\nabla \bar{A} - \nabla \bar{A}^T)X, Y \rangle d\mu$. In particular $d\bar{\Lambda}$ is a linear pseudo 2-form. Furthermore if $\bar{\Lambda}$ is a linear 1-form, *i.e.* $\bar{A} = \nabla f$ for some $f \in C_c^\infty$, then $d\bar{\Lambda} = 0$.

3.2 Calculus of pseudo differential 1-forms

Given a 1–form α on a finite-dimensional manifold, Green's formula compares the integral of $d\alpha$ along a surface to the integral of α along the boundary curves. In

Section 3.2.1 we show that an analogous result for $\mathcal{M}$ is rather simple when strong regularity assumptions are imposed on the surface in $\mathcal{M}$. However, from the point of view of applications it is important to establish Green's formula under weaker assumptions. This is the main goal of this section which we achieve this mainly working with pseudo 1-forms.

3.2.1 Green's formula for smooth surfaces and 1-forms

Let $S : [0,1] \times [0,T] \to \mathcal{M}$ denote a map such that, for each $s \in [0,1]$, $S(s,\cdot) \in AC_2(0,T;\mathcal{M})$ and, for each $t \in [0,T]$, $S(\cdot,t) \in AC_2((0,1);\mathcal{M})$. Let $v(s,\cdot,\cdot)$ denote the velocity of minimal norm for $S(s,\cdot)$ and $w(\cdot,t,\cdot)$ denote the velocity of minimal norm for $S(\cdot,t)$. We assume that $v, w \in C^2([0,1] \times [0,T] \times \mathbb{R}^D, \mathbb{R}^D)$ and that their derivatives up to third order are bounded. We further assume that v and w are gradient vector fields so that $\partial_s v$ and $\partial_t w$ are also gradients.

Let Λ be a differentiable pseudo 1–form on $\mathcal{M}$ such that $\Lambda_\mu(u) = 0$ whenever $u \in L^2(\mu)$ and $div_\mu u = 0$. Because of this, we may view Λ as a 1–form on $\mathcal{M}$. Assume that

$$\sup_{\mu \in \mathcal{K}} ||\Lambda_\mu|| < \infty \tag{119}$$

for all compact subsets $\mathcal{K} \subset \mathcal{M}$, where $||\Lambda_\mu|| := \sup_v \{\Lambda_\mu(v) \ : v \in T_\mu\mathcal{M}, ||v||_\mu \leq 1\}$. We also assume that for all compact subsets $\mathcal{K} \subset \mathcal{M}$ there exists a constant $C_\mathcal{K}$ such that

$$|\Lambda_\nu(u) - \Lambda_\mu(u)| \leq C_\mathcal{K} W_2(\mu,\nu)(||u||_\infty + ||\nabla u||_\infty) \tag{120}$$

for $\mu, \nu \in \mathcal{K}$ and $u \in C_b(\mathbb{R}^D, \mathbb{R}^D)$ such that ∇u is bounded.

Using Remark 1.2.14, Proposition 1.2.15 and the bound on v, w and on their derivatives, we find that S is $1/2$–Hölder continuous. Hence its range is compact so $||\Lambda_{S(s,t)}||$ is bounded. We then use equations (119), (120) and Taylor expansions for w^s_{t+h} and v^{s+h}_t to obtain that

$$\partial_t \Big(\Lambda_{S(s,t)}(w^s_t) \Big)_{|s=\bar{s},t=\bar{t}} = v^{\bar{s}}_{\bar{t}}(\Lambda_{S(s,t)}(w^{\bar{s}}_{\bar{t}})) + \Lambda_{S(\bar{s},\bar{t})}(\partial_t w^{\bar{s}}_t)_{|t=\bar{t}}, \tag{121}$$

where we use the notation of Definition 1.2.8. Similarly,

$$\partial_s\Big(\Lambda_{S(s,t)}(v_t^s)\Big)_{|s=\bar{s},t=\bar{t}} = w_{\bar{t}}^{\bar{s}}(\Lambda_{S(s,t)}(v_{\bar{t}}^{\bar{s}})) + \Lambda_{S(\bar{s},\bar{t})}(\partial_s v_{\bar{t}}^{s})_{|s=\bar{s}}. \tag{122}$$

Now suppose that $S(s,t) = \rho(s,t,\cdot)\mathcal{L}^D$ for some $\rho \in C^1([0,1]\times[0,T]\times\mathbb{R}^D)$ which is bounded with bounded derivatives. Then the following lemma holds.

Lemma 3.2.1. *For $(s,t) \in (r,1)\times(0,T)$ we have $(\partial_t w_t^s - \partial_s v_t^s) - [w_t^s, v_t^s] \in Ker(div_{S(s,t)})$.*

Proof. We have, in the sense of distributions,

$$\partial_t \rho_t^s + \nabla\cdot(\rho_t^s v_t^s) = 0, \quad \partial_s \rho_t^s + \nabla\cdot(\rho_t^s w_t^s) = 0 \tag{123}$$

and so

$$\nabla\cdot\partial_s(\rho_t^s v_t^s) = -\partial_s\partial_t\rho_t^s = \nabla\cdot(\partial_t\rho_t^s w_t^s).$$

We use that ρ, v and w are smooth to conclude that

$$\nabla\cdot\Big(v_t^s\partial_s\rho_t^s + \rho_t^s\partial_s v_t^s\Big) = \nabla\cdot\Big(w_t^s\partial_t\rho_t^s + \rho_t^s\partial_t w_t^s\Big).$$

This implies that if $\varphi \in C_c^\infty(\mathbb{R}^D)$ then

$$\int_{\mathbb{R}^D}\langle\nabla\varphi, v_t^s\partial_s\rho_t^s + \rho_t^s\partial_s v_t^s\rangle = \int_{\mathbb{R}^D}\langle\nabla\varphi, w_t^s\partial_t\rho_t^s + \rho_t^s\partial_t w_t^s\rangle. \tag{124}$$

We use again that ρ, v and w are smooth to obtain that equation (123) holds pointwise. We use this fact on equation (124) to obtain

$$\int_{\mathbb{R}^D}\langle\nabla\varphi, -v_t^s\nabla\cdot(\rho_t^s w_t^s) + \rho_t^s\partial_s v_t^s\rangle = \int_{\mathbb{R}^D}\langle\nabla\varphi, -w_t^s\nabla\cdot(\rho_t^s v_t^s) + \rho_t^s\partial_t w_t^s\rangle.$$

Rearranging, this leads to

$$\int_{\mathbb{R}^D}\langle\nabla\varphi, \partial_s v_t^s - \partial_t w_t^s\rangle\rho_t^s d\mathcal{L}^D = \int_{\mathbb{R}^D}\big\langle\nabla\varphi, v_t^s\big\rangle\nabla\cdot(\rho_t^s w_t^s) - \big\langle\nabla\varphi, w_t^s\big\rangle\nabla\cdot(\rho_t^s v_t^s).$$

Integrating by parts and substituting $\rho_t^s\mathcal{L}^D$ with $S(s,t)$ we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^D}\langle\nabla\varphi, \partial_s v_t^s - \partial_t w_t^s\rangle dS(s,t) \\
&\qquad = \int_{\mathbb{R}^D}\Big(\big\langle\nabla^2\varphi w_t^s + (\nabla w_t^s)^T\nabla\varphi, v_t^s\big\rangle - \big\langle\nabla^2\varphi v_t^s + (\nabla v_t^s)^T\nabla\varphi, w_t^s\big\rangle\Big)dS \\
&\qquad = \int_{\mathbb{R}^D}\Big\langle\nabla\varphi, [v_t^s, w_t^s]\Big\rangle dS(s,t).
\end{aligned}$$

Since $\varphi \in C_c^\infty(\mathbb{R}^D)$ is arbitrary, the proof is finished. □

We combine equations (121) and (122) and use Lemma 3.2.1 to conclude the following.

Proposition 3.2.2. *For each* $t \in (0,T)$ *and* $s \in (a,b)$ *we have*

$$\partial_t\Big(\Lambda_{S(s,t)}(w_t^s)\Big) - \partial_s\Big(\Lambda_{S(s,t)}(v_t^s)\Big) = d\Lambda_{S(s,t)}(v_t^s, w_t^s). \tag{125}$$

Next, we define $||d\Lambda_\mu||$ to be the smallest nonnegative number λ such that $|d\Lambda_\mu(X,Y)| \leq \lambda||X||_\mu||Y||_\mu$ for $X, Y \in \nabla C_c^\infty(\mathbb{R}^D)$.

Theorem 3.2.3 (Green's formula for smooth surfaces)**.** *Let* S *be the surface in* $\mathcal{M}$ *defined above and let its boundary* ∂S *be the union of the negatively oriented curves* $S(r,\cdot)$, $S(\cdot,T)$ *and the positively oriented curves* $S(1,\cdot)$, $S(\cdot,0)$. *Suppose that* $\mu \to ||d\Lambda_\mu||$ *is also bounded on compact subsets of* $\mathcal{M}$. *Then*

$$\int_S d\Lambda = \int_{\partial S} \Lambda.$$

Proof. Recall that v_t^s, w_t^s and their derivatives are bounded. This, together with equations (119) and (120), implies that the functions $(s,t) \to \Lambda_{S(s,t)}(v_t^s)$ and $(s,t) \to \Lambda_{S(s,t)}(w_t^s)$ are continuous. Hence, by Proposition 3.2.2, $(s,t) \to d\Lambda_{S(s,t)}(v_t^s, w_t^s)$ is Borel measurable as it is a limit of quotients of continuous functions. The fact that $\mu \to ||d\Lambda_\mu||$ is bounded on compact subsets of $\mathcal{M}$ gives that $(s,t) \to d\Lambda_{S(s,t)}(v_t^s, w_t^s)$ is bounded. The rest of the proof of this theorem is identical to that of Theorem 3.2.33 when we use Proposition 3.2.2 in place of Corollary 3.2.31. □

3.2.2 Regularity and differentiability of pseudo 1-forms

Definition 3.2.4. Let $\mu \to \bar{\Lambda}_\mu = \int_{\mathbb{R}^D} \langle \bar{A}_\mu, \cdot \rangle d\mu$ be a pseudo 1-form on $\mathcal{M}$. We will say that $\bar{\Lambda}$ is *regular* if for each $\mu \in \mathcal{M}$ there exists a Borel $D \times D$ matrix valued function $B_\mu \in L^\infty(\mathbb{R}^D \times \mathbb{R}^D, \mu)$ and a function $O_\mu \in C(\mathbb{R})$ with $O_\mu(0) = 0$ such that

$$\sup_\gamma \Big\{ \int_{\mathbb{R}^D \times \mathbb{R}^D} |\bar{A}_\nu(y) - \bar{A}_\mu(x) - B_\mu(x)(y-x)|^2 d\gamma(x,y),\ \gamma \in \Gamma_o(\mu,\nu) \Big\}$$
$$\leq W_2^2(\mu,\nu) \min\{O_\mu(W_2(\mu,\nu)), c(\bar{\Lambda})\}^2. \tag{126}$$

where $\Gamma_o(\mu,\nu)$ is the set of γ minimizers in equation (1) and $c(\bar{\Lambda}) > 0$ is a constant independent of μ. We also assume that $||B_\mu||_{L^\infty(\mu)}$ is uniformly bounded. Taking $c(\bar{\Lambda})$ large enough, there is no loss of generality in assuming that

$$\sup_{\mu\in\mathcal{M}} ||B_\mu||_{L^\infty(\mu)} \leq c(\bar{\Lambda}). \tag{127}$$

Remark 3.2.5. Assumption 126 could be substantially weakened for our purposes. We only make such a strong assumption to avoid introducing more notation and making longer computations.

Example 3.2.6. Every linear pseudo 1-form is regular. In other words, given $\bar{A} \in \mathcal{X}_c$, if we define $\bar{\Lambda}_\mu(Y) := \int_{\mathbb{R}^D} \langle \bar{A}, Y\rangle d\mu$ then $\bar{\Lambda}$ is regular. Indeed, setting $B_\mu := \nabla \bar{A}$ we use Taylor expansion and the fact that the second derivatives of A are bounded to obtain equation (126).

Remark 3.2.7. Let $\bar{\Lambda}$ be as in Example 3.2.6. Then the restriction of $\bar{\Lambda}$ to $T\mathcal{M}$ gives a 1-form Λ defined by

$$\Lambda_\mu(Y) := \int_{\mathbb{R}^D} \langle \pi_\mu(\bar{A}), Y\rangle d\mu \qquad \forall Y \in T_\mu\mathcal{M}.$$

It is not clear what smoothness properties the projections $\mu \to \pi_\mu$ might have with respect to $\mu \in \mathcal{M}$. This is one reason why in this context it seems more practical to work with $\bar{A}$ rather than with its projections.

From now till the end of Section 3.2 we assume $\bar{\Lambda}$ is a regular pseudo 1-form on $\mathcal{M}$ and we use the notation $\bar{A}_\mu$, B_μ as in Definition 3.2.4.

Remark 3.2.8. If $\mu,\nu \in \mathcal{M}$, $X \in L^2(\mu)$, $Y \in L^2(\nu)$ and $\gamma \in \Gamma_o(\mu,\nu)$ then

$$\begin{aligned}\bar{\Lambda}_\nu(Y) - \bar{\Lambda}_\mu(X) - \int_{\mathbb{R}^D\times\mathbb{R}^D} \Big(\langle \bar{A}_\mu(x), Y(y) - X(x)\rangle + \langle B_\mu(x)(y-x), Y(y)\rangle\Big) d\gamma(x,y) \\ = \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_\nu(y) - \bar{A}_\mu(x) - B_\mu(x)(y-x), Y(y)\rangle d\gamma(x,y).\end{aligned} \tag{128}$$

By equation (126) and Hölder's inequality

$$\Big|\int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_\nu(y) - \bar{A}_\mu(x) - B_\mu(x)(y-x), Y(y)\rangle\Big| \leq W_2(\mu,\nu)c(\bar{\Lambda})\, ||Y||_\nu. \tag{129}$$

Similarly, equation (127) and Hölder's inequality yield

$$\Big|\int_{\mathbb{R}^D\times\mathbb{R}^D}\langle B_\mu(x)(y-x), Y(y)\rangle\Big| \le W_2(\mu,\nu)c(\bar\Lambda)\,||Y||_\nu. \tag{130}$$

We use equations (129) and (130) to obtain

$$\Big|\bar\Lambda_\nu(Y)-\bar\Lambda_\mu(X)-\int_{\mathbb{R}^D\times\mathbb{R}^D}\langle \bar A_\mu(x), Y(y)-X(x)\rangle d\gamma(x,y)\Big| \le 2c(\bar\Lambda)W_2(\mu,\nu)\,||Y||_\nu. \tag{131}$$

Remark 3.2.9. Let $Y \in C^1_c(\mathbb{R}^D)$ and define $F(\mu) := \bar\Lambda_\mu(Y)$. Then

$$|F(\nu)-F(\mu)| \le W_2(\nu,\mu)\Big(||\bar A_\nu||_\nu ||\nabla Y||_\infty + 2c(\bar\Lambda)||Y||_\infty\Big)$$

Proof. By Hölder's inequality

$$\Big|\int_{\mathbb{R}^D\times\mathbb{R}^D}\langle \bar A_\mu(x), Y(y)-Y(x)\rangle d\gamma(x,y)\Big| \le ||\bar A_\mu||_\mu ||\nabla Y||_\infty W_2(\nu,\mu).$$

We apply Remark 3.2.8 with $Y = X$ and we exchange the role of μ and ν to conclude the proof. □

Lemma 3.2.10. *The function*

$$\mathcal{M}\to\mathbb{R},\quad \mu\mapsto ||\bar A_\mu||_\mu$$

is continuous on $\mathcal{M}$ *and bounded on bounded subsets of* $\mathcal{M}$*. Suppose* $S : [r,1]\times[a,b]\to\mathcal{M}$ *is continuous. Then*

$$\sup_{(s,t)\in[r,1]\times[a,b]} ||\bar A_{S(s,t)}||_{S(s,t)} < \infty.$$

Proof. Fix $\mu_0 \in \mathcal{M}$. For each $\mu \in \mathcal{M}$ we choose $\gamma_\mu \in \Gamma_o(\mu_0,\mu)$. We have

$$\big|\, ||\bar A_\mu||_\mu - ||\bar A_{\mu_0}||_{\mu_0}\big| = \big|\, ||\bar A_\mu(y)||_{\gamma_\mu} - ||\bar A_{\mu_0}(x)||_{\gamma_\mu}\big| \le ||\bar A_\mu(y)-\bar A_{\mu_0}(x)||_{\gamma_\mu}. \tag{132}$$

where

$$||\bar A_\mu||^2_\mu = \int_{\mathbb{R}^D}|\bar A_\mu(y)|^2 d\mu(y) = \iint_{\mathbb{R}^D\times\mathbb{R}^D}|\bar A_\mu(y)|^2 d\gamma_\mu(x,y) =: ||\bar A_\mu(y)||^2_{\gamma_\mu}$$

Similarly

$$||\bar{A}_{\mu_0}||^2_{\mu_0} = \int_{\mathbb{R}^D} |\bar{A}_{\mu_0}(x)|^2 d\mu_0(x) = \iint_{\mathbb{R}^D\times\mathbb{R}^D} |\bar{A}_{\mu_0}(x)|^2 d\gamma_{\mu_0}(x,y) =: ||\bar{A}_{\mu_0}(y)||^2_{\gamma_{\mu_0}}$$

Equation (132), together with equations (126) and (127), yields

$$\Big| ||\bar{A}_\mu||_\mu - ||\bar{A}_{\mu_0}||_{\mu_0} \Big| \le ||B_{\mu_0}(x)(y-x)||_{\gamma_\mu} + c(\bar{\Lambda})W_2(\mu_0,\mu) \le 2c(\bar{\Lambda})W_2(\mu_0,\mu).$$

To obtain the last inequality we have used Hölder's inequality. This proves the first claim.

Notice that $(s,t) \to ||\bar{A}_{S(s,t)}||_{S(s,t)}$ is the composition of two continuous functions and is defined on the compact set $[r,1]\times[a,b]$. Hence it achieves its maximum. □

Lemma 3.2.11. *Let $Y \in C^2_c(\mathbb{R}^D,\mathbb{R}^D)$ and define $F(\mu) := \bar{\Lambda}_\mu(Y)$. Then F is differentiable with gradient $\nabla_\mu F = \pi_\mu(\nabla Y^T(x)\bar{A}_\mu(x) + B^T_\mu(x)Y(x))$.*

Furthermore, assume $X \in \nabla C^2_c(\mathbb{R}^D)$ and let $\varphi_t(x) = x + tX(x) + t\overline{O}_t(x)$, where $\overline{O}_t$ is any continuous function on $\mathbb{R}^D$ such that $||\overline{O}_t||_\infty$ tends to 0 as t tends to 0. Set $\mu_t := \varphi(t,\cdot)_\#\mu$. Then

$$F(\mu_t) = F(\mu)+t\int_{\mathbb{R}^D}\Big[\langle\bar{A}_\mu(x),\nabla Y(x)X(x)\rangle+\langle B_\mu(x)X(x),Y(x)\rangle\Big]d\mu(x)+o(t). \quad (133)$$

Proof. Choose $\mu,\nu\in\mathcal{M}$ and $\gamma\in\Gamma_0(\mu,\nu)$. As in Remark 3.2.8,

$$\begin{aligned}\bar{\Lambda}_\nu(Y)-\bar{\Lambda}_\mu(Y) &- \int_{\mathbb{R}^D\times\mathbb{R}^D}\Big(\langle\bar{A}_\mu(x),Y(y)-Y(x)\rangle+\langle B_\mu(x)(y-x),Y(y)\rangle\Big)d\gamma(x,y)\\ &= \int_{\mathbb{R}^D\times\mathbb{R}^D}\langle\bar{A}_\nu(y)-\bar{A}_\mu(x)-B_\mu(x)(y-x),Y(y)\rangle d\gamma(x,y).\end{aligned}$$

By equation (126) and Hölder's inequality,

$$\Big|\int_{\mathbb{R}^D\times\mathbb{R}^D}\langle\bar{A}_\nu(y)-\bar{A}_\mu(x)-B_\mu(x)(y-x),Y(y)\rangle\Big| \le o(W_2(\mu,\nu))\,||Y||_\nu.$$

Since $Y\in C^2_c(\mathbb{R}^D,\mathbb{R}^D)$ we can write $Y(y)=Y(x)+\nabla Y(x)(y-x)+R(x,y)(y-x,y-x)$, for some continuous quadratic form $R = R(x,y)$ with compact support. Then

$$\int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_\mu(x), Y(y) - Y(x)\rangle d\gamma(x,y) = \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_\mu(x), \nabla Y(x)(y-x)\rangle d\gamma(x,y) + \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle A_\mu(x), R(x,y)(y-x, y-x)\rangle d\gamma(x,y). \tag{134}$$

We now want to show that the term in equation (134) is of the form $o(W_2(\mu,\nu))$ as ν tends to μ. For any $\epsilon > 0$, choose a smooth compactly supported vector field $Z = Z(x)$ such that $\|\bar{A}_\mu - Z\|_\mu < \epsilon$. Then, using Hölder's inequality,

$$\begin{aligned} |\int_{\mathbb{R}^D\times\mathbb{R}^D} \langle A_\mu(x), R\cdot(y-x)^2\rangle d\gamma(x,y)| \leq& \int_{\mathbb{R}^D\times\mathbb{R}^D} |\langle Z(x), R\cdot(y-x)^2\rangle| d\gamma(x,y) \\ &+ \int_{\mathbb{R}^D\times\mathbb{R}^D} |\langle (y-x)^T R^T (A_\mu(x) - Z(x)), y-x\rangle| d\gamma(x,y) \\ &\leq \|(y-x)^T R^T\|_\infty \epsilon W_2(\mu,\nu) + \|R^T Z\|_\infty W_2^2(\mu,\nu). \end{aligned}$$

Since ϵ and $\|Z\|_\infty$ are independent of ν, this gives the required estimate. Likewise,

$$\begin{aligned} \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle B_\mu(x)(y-x), Y(y)\rangle d\gamma(x,y) &= \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle B_\mu(x)(y-x), Y(y) - Y(x)\rangle d\gamma(x,y) \\ &\quad + \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle B_\mu(x)(y-x), Y(x)\rangle d\gamma(x,y) \\ &= \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle B_\mu(x)(y-x), Y(x)\rangle d\gamma(x,y) \\ &\quad + o(W_2(\mu,\nu)). \end{aligned}$$

Combining these results shows that

$$\bar{\Lambda}_\nu(Y) = \bar{\Lambda}_\mu(Y) + \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \nabla Y^T(x)\bar{A}_\mu(x) + B_\mu^T(x)Y(x), y-x\rangle d\gamma(x,y) + o(W_2(\mu,\nu)).$$

As in Definition 1.2.8, this proves that F is differentiable and that

$$\nabla_\mu F = \pi_\mu(\nabla Y^T(x)\bar{A}_\mu(x) + B_\mu^T(x)Y(x)).$$

Now assume that ϕ_t is the flow of X. Notice that the curve $t \to \mu_t$ belongs to $AC_2(-r, r; \mathcal{M})$ for $r > 0$. We could choose for instance $r = 1$. Hence the curve is continuous on $[-1,1]$. By Lemma 3.2.10, the composed function $t \to ||\bar{A}_{\mu_t}||_{\mu_t}$ is also continuous. Hence its range is compact in $\mathbb{R}$, so there exists $\bar{C} > 0$ such that $||\bar{A}_{\mu_t}||_{\mu_t} \leq \bar{C}$ for all $t \in [-1,1]$. We may now use Remark 3.1.10 to conclude.

The general case of ϕ_t as in the statement of Lemma 3.2.11 can be studied using analogous methods. □

Lemma 3.2.12. *Any regular pseudo 1-form is differentiable in the sense of Definition 3.1.12. Furthermore,* $\forall X, Y \in T_\mu\mathcal{M}$*,*

$$d\bar{\Lambda}_\mu(X,Y) = \int_{\mathbb{R}^D} \langle (B_\mu - B_\mu^T)X, Y\rangle d\mu. \tag{135}$$

Proof. We need to check the validity of Definition 3.1.12. Choose $X, Y \in C_c^2(\mathbb{R}^D)$. By Lemma 3.2.11, $\bar{\Lambda}(X)$ and $\bar{\Lambda}(Y)$ are differentiable functions on M. Using the expression given in Lemma 3.2.11 for their gradients, it is simple to check that

$$X\bar{\Lambda}(Y) - Y\bar{\Lambda}(X) - \bar{\Lambda}([X,Y]) = \int_{\mathbb{R}^D} \langle (B_\mu - B_\mu^T)X, Y\rangle d\mu. \tag{136}$$

Since the right hand side of equation (136) is continuous, multilinear and alternating, $d\bar{\Lambda}(X,Y)$ is a well-defined pseudo 2-form on $\mathcal{M}$. □

3.2.3 Further continuity and differentiability properties of regular forms

We collect here various other regularity properties of regular pseudo 1-forms.

Corollary 3.2.13. *Choose* $\sigma \in AC_2(a,b;\mathcal{M})$*. For* $r > 0$ *and* $s \in [r,1]$*, define*

$$D_s : \mathbb{R}^D \to \mathbb{R}^D, \quad D_s(x) := sx.$$

Set $\sigma_t^s = D_{s\#}\sigma_t$*. Then there exists a constant* $C_\sigma(r)$ *depending only on* σ *and* r *such that* $||\bar{A}_{\sigma_t^s}||_{\sigma_t^s} \leq C_\sigma(r)$ *for all* $(s,t) \in [r,1]\times[a,b]$*.*

Proof. By Remark 1.2.14 (i), $\sigma : [a,b] \to \mathcal{M}$ is 1/2–Hölder continuous: there exists a constant $c > 0$ such that $W_2^2(\sigma_{t_2}, \sigma_{t_1}) \leq c|t_2 - t_1|$. Together with lemma 3.1.3 and the fact that $Lip(D_s) = s \leq 1$, this gives that $t \to \sigma_t^s$ is uniformly 1/2–Hölder continuous:

$$W_2^2(\sigma_{t_2}^s, \sigma_{t_1}^s) \leq W_2^2(\sigma_{t_2}, \sigma_{t_1}) \leq c|t_2 - t_1|.$$

Remark 1.2.14 (ii) ensures that $\{\sigma_t |\ t \in [a,b]\}$ is bounded and so there exists $\bar{c} > 0$ such that $W_2(\sigma_t, \delta_0) \leq \bar{c}$ for all $t \in [a,b]$.

One can readily check that $\gamma := \big(D_{s_1} \times D_{s_2}\big)_{\#}\sigma_t \in \Gamma(\sigma_t^{s_1}, \sigma_t^{s_2})$, so

$$
\begin{aligned}
W_2^2(\sigma_t^{s_1}, \sigma_t^{s_2}) \leq \int_{\mathbb{R}^D \times \mathbb{R}^D} |x-y|^2 d\gamma &= \int_{\mathbb{R}^D} |D_{s_1}x - D_{s_2}x|^2 d\sigma_t(x) \\
&= |s_2 - s_1|^2 \int_{\mathbb{R}^D} |x|^2 d\sigma_t(x) \leq \bar{c}|s_2 - s_1|^2.
\end{aligned}
$$

Thus $s \to \sigma_t^s$ is 1–Lipschitz. Consequently $(t,s) \to \sigma_t^s$ is 1/2–Hölder continuous. This, together with Lemma 3.2.10, yields the proof.

∎

Lemma 3.2.14. *Assume $\{\mu_\epsilon\}_{\epsilon \in E} \subset \mathcal{M}$ and $v_\epsilon \in L^2(\mu_\epsilon)$ are such that $C := \sup_{\epsilon \in E} ||v_\epsilon||_{L^2(\mu_\epsilon)}$ is finite. Assume $\{\mu_\epsilon\}_{\epsilon \in E}$ converges to μ in $\mathcal{M}$ as ϵ tends to 0 and that there exists $v \in L^2(\mu)$ such that $\{v_\epsilon \mu_\epsilon\}_{\epsilon \in E}$ converges weak-$*$ to $v\mu$, as $\epsilon \to 0$. If $\gamma_\epsilon \in \Gamma_o(\mu, \mu_\epsilon)$ then $\lim_{\epsilon \to 0} a_\epsilon = 0$, where $a_\epsilon = \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_\mu(x), v_\epsilon(y) - v(x) \rangle d\gamma_\epsilon(x,y)$.*

Proof. It is easy to obtain that $||v||_{L^2(\mu)} \leq C$. Let $\gamma_\epsilon \in \Gamma_o(\mu, \mu_\epsilon)$ and $\xi \in \mathcal{X}_c$. Then there exists a bounded function $C_\xi \in C(\mathbb{R}^D \times \mathbb{R}^D)$ and a real number M such that

$$
\xi(x) - \xi(y) = \nabla\xi(y)(x-y) + |x-y|^2 C_\xi(x,y), \quad |C_\xi(x,y)| \leq M, \tag{137}
$$

for $x, y \in \mathbb{R}^D$. We use the first equality in equation (137) to obtain that

$$
\begin{aligned}
\langle \bar{A}_\mu(x), v_\epsilon(y) - v(x) \rangle = \langle \bar{A}_\mu(x) - \xi(x), v_\epsilon(y) - v(x) \rangle + \langle \xi(y), v_\epsilon(y) \rangle - \langle \xi(x), v(x) \rangle \\
+ \langle \nabla\xi(y)(x-y) + |x-y|^2 C_\xi(x,y), v_\epsilon(y) \rangle.
\end{aligned} \tag{138}
$$

Hence,

$$
\begin{aligned}
|a_\epsilon| \leq ||\bar{A}_\mu(x) - \xi(x)||_{L^2(\gamma_\epsilon)} ||v_\epsilon(y) - v(x)||_{L^2(\gamma_\epsilon)} + b_\epsilon \\
+ \Big| \int_{\mathbb{R}^D \times \mathbb{R}^D} \big(\nabla\xi(y)(x-y) + |x-y|^2 C_\xi(x,y)\big) d\gamma_\epsilon(x,y) \Big|.
\end{aligned} \tag{139}
$$

Above, we have set $b_\epsilon := |\int_{\mathbb{R}^D\times\mathbb{R}^D}\big(\langle\xi(y), v_\epsilon(y)\rangle - \langle\xi(x), v(x)\rangle\big)d\gamma_\epsilon(x,y)|$. By the second inequality in equation (137) and by equation (139)

$$|a_\epsilon| \le 2C||\bar{A}_\mu - \xi||_{L^2(\mu)} + b_\epsilon + ||\nabla\xi||_\infty W_2(\mu,\mu_\epsilon) + MW_2^2(\mu,\mu_\epsilon). \tag{140}$$

By assumption $\{W_2(\mu,\mu_\epsilon)\}_{\epsilon\in E}$ tends to 0 and $\{b_\epsilon\}_{\epsilon\in E}$ tends to 0 as ϵ tends to 0. These facts, together with equation (140), yield $\limsup_{\epsilon\to 0}|a_\epsilon| \le 2C||\bar{A}_\mu - \xi||_{L^2(\mu)}$ for arbitrary $\xi \in \mathcal{X}_c$. We use that $\mathcal{X}_c$ is dense in $L^2(\mu)$ to conclude that $\lim_{\epsilon\to} a_\epsilon = 0$.

□

Corollary 3.2.15. *Assume* $\{\mu_\epsilon\}_{\epsilon\in E}\subset\mathcal{M}$, μ, $v_\epsilon\in L^2(\mu_\epsilon)$ *and* v *satisfy the assumptions of Lemma 3.2.14. Then* $\lim_{\epsilon\to 0}\bar{\Lambda}_{\mu_\epsilon}(v_\epsilon) = \bar{\Lambda}_\mu(v)$.

Proof. Let $\gamma_\epsilon\in\Gamma_o(\mu,\mu_\epsilon)$. Observe that

$$\begin{aligned}\langle\bar{A}_{\mu_\epsilon}(y), v_\epsilon(y)\rangle - \langle\bar{A}_\mu(x), v(x)\rangle &= \langle\bar{A}_\mu(x), v_\epsilon(y) - v(x)\rangle + \langle B_\mu(x)(y-x), v_\epsilon(y)\rangle \\ &\quad + \Big\langle \bar{A}_{\mu_\epsilon}(y) - \bar{A}_\mu(x) - B_\mu(x)(y-x), v_\epsilon(y)\Big\rangle.\end{aligned} \tag{141}$$

We now integrate equation (141) over $\mathbb{R}^D\times\mathbb{R}^D$ and use equations (126)–(127) and the fact that $\gamma_\epsilon\in\Gamma_o(\mu,\mu_\epsilon)$. We obtain

$$\begin{aligned}|\bar{\Lambda}_{\mu_\epsilon}(v_\epsilon) - \bar{\Lambda}_\mu(v)| &\le |a_\epsilon| + ||B_\mu||_{L^\infty(\mu)}W_2(\mu,\mu_\epsilon)||v_\epsilon||_{\mu_\epsilon} + o(W_2(\mu,\mu_\epsilon))||v_\epsilon||_{\mu_\epsilon} \\ &\le |a_\epsilon| + C||B_\mu||_{L^\infty(\mu)}W_2(\mu,\mu_\epsilon) + C\,o(W_2(\mu,\mu_\epsilon)).\end{aligned} \tag{142}$$

Letting ϵ tend to 0 in equation (142) we conclude the proof of the corollary. □

Lemma 3.2.16 (continuity of $\bar{\Lambda}_{\sigma_t}(X_t)$)**.** *Suppose* $\sigma\in AC_2(a,b;\mathcal{M})$. *If* $X\in C((a,b)\times\mathbb{R}^D,\mathbb{R}^D)$ *such that* $\sup_t ||X_t||_{\sigma_t}$ *is finite then* $t\to\bar{\Lambda}_{\sigma_t}(X_t) =: \lambda(t)$ *is continuous on* (a,b).

Proof. Fix $t \in (a,b)$ so that t belongs to the interior of a compact set $K^* \subset (a,b)$. Let $\varphi \in C_c(\mathbb{R}^D, \mathbb{R}^D)$ and denote by K a compact set containing its support. Observe that X is uniformly continuous on $K^* \times K$ so

$$\limsup_{h\to 0} |\int_{\mathbb{R}^D} \langle \varphi(x), X_{t+h}(x) - X_t(x)\rangle d\sigma_{t+h}(x)| \leq \limsup_{h\to 0} ||\varphi||_\infty \sup_{x\in K} |X_{t+h}(x) - X_t(x)| = 0. \tag{143}$$

Since $\langle X_t, \varphi\rangle \in C_c$ and σ is continuous at t by Remark 1.2.14, we also see that

$$\lim_{h\to 0} \int_{\mathbb{R}^D} \langle \varphi(x), X_t(x)\rangle d\sigma_{t+h}(x) = \int_{\mathbb{R}^D} \langle \varphi(x), X_t(x)\rangle d\sigma_t(x). \tag{144}$$

Since $\varphi \in C_c(\mathbb{R}^D, \mathbb{R}^D)$ is arbitrary, equations (143) and (144) give that $\{X_{t+h}\sigma_{t+h}\}_{h>0}$ converges weak-$*$ to $\sigma_t X_t$ as h tends to zero. Corollary 3.2.15 yields that λ is continuous at t. □

Lemma 3.2.17 (Lipschitz property of $\bar\Lambda_{\sigma_t}(X_t)$)**.** *Suppose that $\sigma \in AC_2(a,b;\mathcal{M})$ and v is a velocity for σ. Let $X \in C^1([a,b]\times\mathbb{R}^D, \mathbb{R}^D)$ and $\tilde C > 0$ be such that*

$$\sup_{t\in[a,b]} ||\bar A_{\sigma_t}||_{\sigma_t},\ ||v_t||_{\sigma_t},\ ||X_t||_{\sigma_t},\ ||\partial_t X_t||_\infty,\ ||\nabla X_t||_\infty \leq \tilde C. \tag{145}$$

Then $t \to \bar\Lambda_{\sigma_t}(X_t) =: \lambda(t)$ is L–Lipschitz for a constant L which is an increasing function of $\tilde C$.

Proof. By equation (145)

$$|X(t+h,y) - X(t,x)| = \Big|\int_0^1 \big(h\partial_t X + \nabla X\cdot(y-x)\big)(t+lh, x+l(y-x))dl\Big| \leq \tilde C(|h| + |y-x|). \tag{146}$$

Let $\gamma_h \in \Gamma_o(\sigma_t, \sigma_{t+h})$. We exploit equation (131) where we substitute Y by X_{t+h} and use equations (145) and (146) to obtain

$$\begin{aligned} |\lambda(t+h) - \lambda(t)| &\leq |\int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar A_{\sigma_t}(x), X_{t+h}(y) - X_t(x)\rangle d\gamma_h(x,y)| \\ &\qquad + 2c(\bar\Lambda) W_2(\sigma_t, \sigma_{t+h})\, ||X_{t+h}||_{\sigma_{t+h}} \\ &\leq \tilde C^2(|h| + W_2(\sigma_t,\sigma_{t+h})) + 2c(\bar\Lambda)W_2(\sigma_t,\sigma_{t+h})\,\tilde C \\ &\leq 2|h|\tilde C^2\big(1 + \tilde C + 2c(\bar\Lambda)\big). \end{aligned} \tag{147}$$

The last inequality in equation (147) is a consequence of equation (145) and Remark 1.2.14, which yield $W_2(\sigma_t, \sigma_{t+h}) \leq \tilde{C}|h|$. Thus λ is L–Lipschitz with $L := \tilde{C}^2\big(1 + \tilde{C} + 2c(\bar{\Lambda})\big)$. □

One can identify points where λ is differentiable by making additional assumptions on X. We next show that the set of differentiability of λ contains $(a, b) \setminus \mathcal{N}$. Here, $\mathcal{N}$ is the set of $t \in (a, b)$ for which there exists $\gamma_h \in \Gamma_o(\sigma_t, \sigma_{t+h})$ such that $\big(\pi^1 \times (\pi^2 - \pi^1)/h\big)_{\#}\gamma_h$ fails to converge to $(Id \times \bar{v}_t)_{\#}\sigma_t$ in $\mathcal{P}_2(\mathbb{R}^D \times \mathbb{R}^D)$ as h tends to 0. The derivative of λ at t will be written in terms of the projection $\bar{v}_t$ of v_t onto the tangent space $T_{\sigma_t}\mathcal{M}$, *i.e.* $\bar{v}_t := \pi_{\sigma_t}(v_t)$.

Lemma 3.2.18 (Differentiability property of $\bar{\Lambda}_{\sigma_t}(X_t)$)**.** *Suppose that σ, v and X are as in Lemma 3.2.17. We further suppose that $X \in C^2([a, b] \times \mathbb{R}^D, \mathbb{R}^D)$ and*

$$||\partial^2_{tt} X_t||_\infty,\ ||\nabla \partial_t X_t||_\infty,\ ||\nabla^2 X_t||_\infty \leq \tilde{C}. \tag{148}$$

If $t \in (a, b) \setminus \mathcal{N}$ then

$$\lambda'(t) = \int_{\mathbb{R}^D} \Big\langle \bar{A}_{\sigma_t}(x), \partial_t X_t(x) + \nabla X_t(x) \cdot \bar{v}_t(x) \Big\rangle d\sigma_t(x) + \int_{\mathbb{R}^D} \big\langle B_{\sigma_t}(x) \cdot \bar{v}_t(x), X_t(x) \big\rangle d\sigma_t(x). \tag{149}$$

Proof. We shall show that equation (155) holds by establishing a series of inequalities. First, by equations (145) and (148)

$$|X(t + h, y) - X(t, x) - h\partial_t X(t, x) - \nabla X(t, x) \cdot (y - x)| \leq \tilde{C}(|h|^2 + |y - x|^2). \tag{150}$$

We exploit equation (150) to obtain

$$\Big|\int_{\mathbb{R}^D \times \mathbb{R}^D} \Big(\langle \bar{A}_{\sigma_t}(x), X_{t+h}(y) - X_t(x)\rangle - h\big\langle \bar{A}_{\sigma_t}(x), \partial_t X_t(x) + \nabla X_t(x) \cdot \frac{y - x}{h}\rangle\Big) d\gamma_h(x, y)\Big|$$
$$\leq \tilde{C}^2\Big(|h|^2 + W_2^2(\sigma_t, \sigma_{t+h})\Big).$$

This, together with the fact that $t \in (a,b) \setminus \mathcal{N}$ yields

$$\lim_{h\to 0} \int_{\mathbb{R}^D \times \mathbb{R}^D} \Big\langle \bar{A}_{\sigma_t}(x), \frac{X_{t+h}(y) - X_t(x)}{h} \Big\rangle d\gamma_h(x,y)$$
$$= \int_{\mathbb{R}^D} \Big\langle \bar{A}_{\sigma_t}(x), \partial_t X_t(x) + \nabla X_t(x) \cdot \bar{v}_t(x) \Big\rangle d\sigma_t(x). \quad (151)$$

By equations (145) and (150)

$$|X(t+h,y) - X(t,x)| \leq \tilde{C}(|h| + |y-x| + |h|^2 + |y-x|^2)$$

So Hölder's inequality yields

$$\Big| \int_{\mathbb{R}^D \times \mathbb{R}^D} \Big\langle B_{\sigma_t}(x) \cdot (y-x), X_{t+h}(y) - X_t(x) \Big\rangle d\gamma_h(x,y) \Big|$$
$$\leq ||B_{\sigma_t}||_{\sigma_t} \tilde{C} W_2(\sigma_t, \sigma_{t+h}) \cdot \Big(|h| + |h|^2 + W_2(\sigma_t, \sigma_{t+h}) + W_2^2(\sigma_t, \sigma_{t+h}) \Big)$$
$$\leq c(\bar{\Lambda}) \tilde{C}^2 |h| \Big(|h| + |h|^2 + \tilde{C}|h| + \tilde{C}^2 |h|^2 \Big). \quad (152)$$

To obtain equation (152) we have used equation (127) to bound $||B_{\sigma_t}||_{\sigma_t}$. As before, we have also used Remark 1.2.14 to control $W_2(\sigma_t, \sigma_{t+h})$ with $\tilde{C}|h|$. By equation (152) and the fact that $t \in (a,b) \setminus \mathcal{N}$

$$\lim_{h\to 0} \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_{\sigma_t}(x) \cdot \frac{y-x}{h}, X_{t+h}(y) \rangle d\gamma_h(x,y)$$
$$= \lim_{h\to 0} \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle B_{\sigma_t}(x) \cdot \frac{y-x}{h}, X_t(x) \rangle d\gamma_h(x,y)$$
$$= \int_{\mathbb{R}^D} \langle B_{\sigma_t}(x) \cdot \bar{v}_t(x), X_t(x) \rangle d\sigma_t(x). \quad (153)$$

If we substitute ν by σ_{t+h}, μ by σ_t, Y by X_{t+h} and X by X_t in equation (129) and as before control $W_2(\sigma_t, \sigma_{t+h})$ with $\tilde{C}|h|$, we obtain

$$\lim_{h\to 0} \frac{1}{h} \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle \bar{A}_{\sigma_{t+h}}(y) - \bar{A}_{\sigma_t}(x) - B_{\sigma_t}(x)(y-x), \bar{A}_{\sigma_{t+h}}(y) \rangle \Big) d\gamma_h(x,y) = 0. \quad (154)$$

We make the same substitution in equation (128) and use equation (154) to obtain

$$\lambda'(t) = \lim_{h\to 0} \int_{\mathbb{R}^D \times \mathbb{R}^D} \Big(\langle \bar{A}_{\sigma_t}(x), \frac{X_{t+h}(y) - X_t(x)}{h} \rangle + \langle B_{\sigma_t}(x) \cdot \frac{y-x}{h}, X_{t+h}(y) \rangle \Big) d\gamma_h(x,y). \quad (155)$$

Thanks to equations (155), (151) and (153) we obtain equation (149). $\square$

3.2.4 Mollification of absolutely continuous paths in $\mathcal{M}$

Throughout this section we suppose that $\eta_D^\epsilon \in C^\infty(\mathbb{R}^D)$ is a mollifier : $\eta_D^\epsilon(x) = 1/\epsilon^D \eta(x/\epsilon)$, for some bounded symmetric function $\eta \in C^\infty(\mathbb{R}^D)$ whose derivatives of all orders are bounded. We also impose that $\eta > 0$, $\int_{\mathbb{R}^D} |x|^2 \eta(x)dx < \infty$ and $\int_{\mathbb{R}^D} \eta = 1$. We fix $\mu \in \mathcal{M}$ and define $f^\epsilon(x) := \int_{\mathbb{R}^D} \eta_D^\epsilon(x-y)d\mu(y)$. Observe that $f^\epsilon \in C^\infty(\mathbb{R}^D)$ is bounded, all its derivatives are bounded and $\int_{\mathbb{R}^D} f^\epsilon = 1$.

We suppose that $\eta_1^\epsilon \in C^\infty(\mathbb{R})$ is a standard mollifier: $\eta_1^\epsilon(t) = 1/\epsilon \eta_1(t/\epsilon)$, for some bounded symmetric function $\eta_1 \in C^\infty(\mathbb{R})$ which is positive on $(-1,1)$ and vanishes outside $(-1,1)$. We also impose that $\int_{\mathbb{R}} \eta_1 = 1$ and assume that $|\epsilon| < 1$.

Suppose $\sigma \in AC_2(a,b;\mathcal{M})$ and $v : (a,b) \times \mathbb{R}^D \to \mathbb{R}^D$ is a velocity associated to σ so that $t \to ||v_t||_{\sigma_t} \in L^\infty(a,b)$. Suppose that for each $t \in (a,b)$ there exists $\rho_t > 0$ such that $\sigma_t = \rho_t \mathcal{L}^D$.

We can extend σ and v in time on an interval larger than $[a,b]$. For instance, set $\tilde{\sigma}_t = \sigma_a$ for $t \in (a-1,a)$ and set $\tilde{\sigma}_t = \sigma_b$ for $t \in (b,b+1)$. Observe that $\tilde{\sigma} \in AC_2(a-1,b+1;\mathcal{M})$ and we have a velocity $\tilde{v}$ associated to $\tilde{\sigma}$ such that $\tilde{v}_t = v_t$ for $t \in [a,b]$. We can choose $\tilde{v}$ such that $||\tilde{v}_t||^2_{\tilde{\sigma}_t} = 0$ for t outside (a,b). In particular, $\int_{a-1}^{b-1} ||\tilde{v}_t||^2_{\tilde{\sigma}_t} dt = \int_a^b ||v_t||^2_{\sigma_t} dt$. If σ is time periodic i.e $\sigma_a = \sigma_b$ and $v_a = v_b$ then we set $\tilde{\sigma}_t = \sigma_{t+b-a}$ (respectively $\tilde{v}_t = v_{t+b-a}$) for $t \in (a-1,a)$ and $\tilde{\sigma}_t = \sigma_{t-(b-a)}$ ($\tilde{v}_t = v_{t-(b-a)}$) for $t \in (b,b+1)$. In the sequel we won't distinguish between σ, $\tilde{\sigma}$ on the one hand and v, $\tilde{v}$ on the other hand. This extension becomes useful when we try to define ρ_t^ϵ as it appears in equation (156). The new density functions are meaningful if we substitute σ by $\tilde{\sigma}$ and impose that $\epsilon \in (0,1)$.

For $\epsilon \in (0,1)$, set

$$\rho_t^\epsilon(x) := \int_{\mathbb{R}} \eta_1^\epsilon(t-\tau)\rho_\tau(x)d\tau,\ \sigma_t^\epsilon := \rho_t^\epsilon \mathcal{L}^D,\ \rho_t^\epsilon(x)v_t^\epsilon(x) := \int_{\mathbb{R}} \eta_1^\epsilon(t-\tau)\rho_\tau(x)v_\tau(x)d\tau. \tag{156}$$

Note that $\rho_t^\epsilon(x) > 0$ for all $t \in (a,b)$ and $x \in \mathbb{R}^D$ and ρ_t^ϵ is a probability density.

Also, $v^\epsilon : (a,b) \times \mathbb{R}^D \to \mathbb{R}^D$ is a velocity associated to σ^ϵ.

In the sequel we set

$$C^2 := \int_{\mathbb{R}^D} |x|^2 \eta(x) dx, \quad C_1 = \int_{\mathbb{R}} \eta_1(\tau) \tau d\tau, \quad C_v := \sup_{\tau \in (a-1, b+1)} ||v_\tau||_{\sigma_\tau}.$$

Lemma 3.2.19. *We assume that for each $t \in (a,b)$ there exists $\rho_t > 0$ such that $\sigma_t = \rho_t \mathcal{L}^D$. Then $\sigma^\epsilon \in AC_2(a,b;\mathcal{M})$. For $a < s < t < b$,*

$$(i)\ W_2(\mu, f^\epsilon \mathcal{L}^D) \le \epsilon C, \quad (ii)\ ||v_t^\epsilon||_{\sigma_t^\epsilon} \le C_v \quad \text{and} \quad (iii)\ W_2(\sigma_t^\epsilon, \sigma_t) \le \epsilon C_1 C_v.$$

Proof. We denote by $\mathcal{U}$ the set of pairs (u,v) such that $u, v \in C(\mathbb{R}^D)$ are bounded and $u(x) + v(y) \le |x-y|^2$ for all $x, y \in \mathbb{R}^D$. Fix $(u,v) \in \mathcal{U}$. By Fubini's theorem one gets the well-known identity

$$\int_{\mathbb{R}^D} u(x) f^\epsilon(x) dx = \int_{\mathbb{R}^D} d\mu(y) \int_{\mathbb{R}^D} u(x) \eta_\epsilon(x-y) dx. \tag{157}$$

Since $v(y) = \int_{\mathbb{R}^D} v(y) \eta_\epsilon(x-y) dx$, equation (157) yields that

$$\begin{aligned}
\int_{\mathbb{R}^D} u(x) f^\epsilon(x) dx + \int_{\mathbb{R}^D} v(y) d\mu(y) &= \int_{\mathbb{R}^D} d\mu(y) \int_{\mathbb{R}^D} \eta_\epsilon(x-y) \big(u(x) + v(y)\big) dx \\
&\le \int_{\mathbb{R}^D} d\mu(y) \int_{\mathbb{R}^D} \eta_\epsilon(x-y) |x-y|^2 dx \qquad (158) \\
&= \int_{\mathbb{R}^D} d\mu(y) \int_{\mathbb{R}^D} \frac{1}{\epsilon^D} \eta(\frac{z}{\epsilon}) |z|^2 dz = C^2 \epsilon^2.
\end{aligned}$$

To obtain equation (158) we have used that $(u,v) \in \mathcal{U}$. We have proven that $\int_{\mathbb{R}^D} u(x) f^\epsilon(x) dx + \int_{\mathbb{R}^D} v(y) d\mu(y) \le C^2 \epsilon^2$ for arbitrary $(u,v) \in \mathcal{U}$. Thanks to the dual formulation of the Wasserstein distance equation (4), we conclude the proof of (i).

Note that for each $t \in (a,b)$ and $x \in \mathbb{R}^D$, $\eta_1^\epsilon(t-\tau)\rho_\tau(x)/\rho_t^\epsilon(x)$ is a probability density on $\mathbb{R}$. Hence, by Jensen's inequality

$$|v_t^\epsilon(x)|^2 = \left| 1/\rho_t^\epsilon(x) \int_{\mathbb{R}} \eta_1^\epsilon(t-\tau) \rho_\tau(x) v_\tau(x) d\tau \right|^2 \le 1/\rho_t^\epsilon(x) \int_{\mathbb{R}} \eta_1^\epsilon(t-\tau) \rho_\tau(x) |v_\tau(x)|^2 d\tau.$$

We multiply both sides of the previous inequality by $\rho_t^\epsilon(x)$. We integrate the subsequent inequality over $\mathbb{R}^D$ and use Fubini's theorem to conclude the proof of (ii).

We use (ii) and Remark 1.2.14 (i) to obtain that $\sigma^\epsilon \in AC_2(a,b;\mathcal{M})$.

We have

$$\int_{\mathbb{R}^D} u(x)d\sigma_t^\epsilon(x) = \int_{\mathbb{R}^D} u(x)dx \int_{\mathbb{R}} \eta_1^\epsilon(\tau)\rho_{t-\tau}(x)d\tau = \int_{\mathbb{R}} \eta_1^\epsilon(\tau)d\tau \int_{\mathbb{R}^D} u(x)d\sigma_{t-\tau}(x).$$

Hence, using that $v(y) = \int_{\mathbb{R}} \eta_1^\epsilon(\tau)v(y)d\tau$, we obtain

$$\int_{\mathbb{R}^D} u(x)d\sigma_t^\epsilon(x) + \int_{\mathbb{R}^D} v(y)d\sigma_t(y) = \int_{\mathbb{R}} \eta_1^\epsilon(\tau)d\tau\Big(\int_{\mathbb{R}^D} ud\sigma_{t-\tau} + \int_{\mathbb{R}^D} vd\sigma_\tau\Big) \tag{159}$$

$$\leq \int_{\mathbb{R}} \eta_1^\epsilon(\tau)W_2^2(\sigma_{t-\tau},\sigma_t)d\tau \tag{160}$$

$$\leq \int_{\mathbb{R}} \eta_1^\epsilon(\tau)\tau^2 C_v^2 d\tau = \epsilon^2 C_1 C_v^2. \tag{161}$$

To obtain equation (160) we have used the dual formulation of the Wasserstein distance equation (4) and the fact that $(u,v) \in \mathcal{U}$. We have used Remark 1.2.14 to obtain equation (161). Since $\int_{\mathbb{R}^D} ud\sigma_t^\epsilon + \int_{\mathbb{R}^D} vd\sigma_t \leq \epsilon CC_v$ for arbitrary $(u,v) \in \mathcal{U}$, we conclude that (iii) holds. □

Remark 3.2.20. Assume that for each $t \in (a,b)$ there exists $\rho_t > 0$ such that $\sigma_t = \rho_t\mathcal{L}^D$. Let $\phi \in C_c(\mathbb{R}^D)$. Setting $I_\phi(t) := \int_{\mathbb{R}^D}\langle\phi, v_t\rangle\rho_t d\mathcal{L}^D$, we have

$$|\int_{\mathbb{R}^D}\langle\phi, v_t^\epsilon\rangle\rho_t^\epsilon d\mathcal{L}^D| = |\eta_1^\epsilon * I_\phi(t)| \leq ||\phi||_\infty\, C_v. \tag{162}$$

Corollary 3.2.21. *Suppose that for each $t \in (a,b)$ there exists $\rho_t > 0$ such that $\sigma_t = \rho_t\mathcal{L}^D$. Then, for each $t \in [a,b]$, $\{\sigma_t^\epsilon\}_{\epsilon>0}$ converges to σ_t in $\mathcal{M}$ as ϵ tends to zero. For $\mathcal{L}^1$–almost every $t \in [a,b]$, $\{\sigma_t^\epsilon v_t^\epsilon\}_{\epsilon>0}$ converges weak-$*$ to $\sigma_t v_t$ as ϵ tends to zero.*

Proof. By Lemma 3.2.19 (iii), $\{\sigma_t^\epsilon\}_{\epsilon>0}$ converges to σ_t in $\mathcal{M}$ as ϵ tends to zero.

Let $\mathcal{C}$ be a countable family in $C_c(\mathbb{R}^D)$. For each $\phi \in C_c(\mathbb{R}^D)$, the set of Lebesgue points of I_ϕ is a set of full measure in $[a,b]$. For these points $\eta_1^\epsilon * I_\phi(t)$ tends to $I_\phi(t)$ as ϵ tends to zero. Thus there is a set S of full measure in $[a,b]$ such that for all $\phi \in \mathcal{C}$ and all $t \in S$, $\eta_1^\epsilon * I_\phi(t)$ tends to $I_\phi(t)$ as ϵ tends to zero. Fix $\varphi \in C_c(\mathbb{R}^D)$ and choose

$\delta > 0$ arbitrary. Let $\phi \in \mathcal{C}$ be such that $||\varphi - \phi||_\infty \leq \delta$. Note that

$$|\eta_1^\epsilon * I_\varphi(t) - I_\varphi(t)| \leq |\eta_1^\epsilon * I_\phi(t) - I_\phi(t)| + |\eta_1^\epsilon * I_{\phi-\varphi}(t)| + |I_{\phi-\varphi}(t)|.$$

We use inequality 162 to conclude that

$$|\eta_1^\epsilon * I_\varphi(t) - I_\varphi(t)| \leq |\eta_1^\epsilon * I_\phi(t) - I_\phi(t)| + 2\delta C_v.$$

If $t \in S$, the previous inequality gives that $\limsup_{\epsilon\to 0} |\eta_1^\epsilon * I_\varphi(t) - I_\varphi(t)| \leq 2\delta C_v$. Since $\delta > 0$ is arbitrary we conclude that $\lim_{\epsilon\to 0} |\eta_1^\epsilon * I_\varphi(t) - I_\varphi(t)| = 0$. ☐

Corollary 3.2.22. *Suppose $\sigma \in AC_2(a, b; \mathcal{M})$ for all $a < b$, v is a velocity associated to σ and $\infty > C := \sup_{t\in[a,b]} ||v_t||_{\sigma_t}$. Define*

$$f_t^r(x) := \int_{\mathbb{R}^D} \eta_D^r(x-y)d\sigma_t(y), \quad \sigma_t^r := f_t^r \mathcal{L}^D, \quad f_t^r(x)v_t^r(x) := \int_{\mathbb{R}^D} \eta_D^r(x-y)v_t(y)d\sigma_t(y).$$

As in equation (156), we define for $0 < \epsilon < 1$,

$$\rho_t^{\epsilon,r}(x) := \int_{\mathbb{R}} \eta_1^\epsilon(t-\tau)f_\tau^r(x)d\tau, \quad \sigma_t^{\epsilon,r} := \rho^{\epsilon,r}\mathcal{L}^D, \quad \rho_t^{\epsilon,r}(x)v_t^{\epsilon,r}(x) := \int_{\mathbb{R}} \eta_1^\epsilon(t-\tau)f_\tau^r(x)v_\tau^r(x)d\tau.$$

Then,

(i) v^r is a velocity associated to σ^r and, for each $t \in (a, b)$, $\{\sigma_t^r\}_r$ converges to σ_t in $\mathcal{M}$ as r tends to zero. For $\mathcal{L}^1$–almost every $t \in (a, b)$, $||v_t^r||_{\sigma_t^r} \leq C$ and $\{v_t^r\sigma_t^r\}_{r>0}$ converges weak-∗ to $\bar{v}_t\bar{\sigma}_t$ as r tends to zero.

(ii) $v^{\epsilon,r}$ is a velocity associated to $\sigma^{\epsilon,r}$ and, for each $t \in (a, b)$, $\{\sigma_t^{\epsilon,r}\}_\epsilon$ converges to σ_t^r in $\mathcal{M}$ as ϵ tends to zero. For every $t \in (a, b)$, $||v_t^{\epsilon,r}||_{\sigma_t^{\epsilon,r}} \leq C$ while for $\mathcal{L}^1$–almost every $t \in (a, b)$, $\{v_t^{\epsilon,r}\sigma_t^{\epsilon,r}\}_{r>0}$ converges weak-∗ to $v_t^r\sigma_t^r$ as ϵ tends to zero.

(iii) The function $t \to \bar{\Lambda}_{\sigma_t^{\epsilon,r}}(v_t^{\epsilon,r})$ is continuous while $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ is measurable on $(0, T)$.

(iv) Suppose in addition that σ and v are time–periodic: $\sigma_a = \sigma_b, \qquad v_a = v_b$. Then $\sigma_a^{\epsilon,r} = \sigma_b^{\epsilon,r}$ and $v_a^{\epsilon,r} = v_b^{\epsilon,r}$.

Proof. It is well known that $||v_t^r||_{\sigma_t^r} \le ||v_t||_{\sigma_t} \le C$ (see [3] Lemma 8.1.10) so, by Remark 1.2.14 (i), $\sigma \in AC_2(a,b;\mathcal{M})$. One can readily check that v^r is a velocity associated to σ^r. Lemma 3.2.19 shows that, for each $t \in (a,b)$, $\{\sigma_t^r\}_r$ converges to σ_t in $\mathcal{M}$ as r tends to zero. Let $\varphi \in C_c(\mathbb{R}^D, \mathbb{R}^D)$. Set $\varphi^r := \eta_D^r * \varphi$. Since $\{\varphi^r\}_{r>0}$ converges uniformly to φ,

$$\lim_{r\to 0} \int_{\mathbb{R}^D} \langle \varphi, v_t^r \rangle d\sigma_t^r = \int_{\mathbb{R}^D} \langle v_t, \varphi \rangle d\sigma_t.$$

Thus $\{v_t^r \sigma_t^r\}_{r>0}$ converges weak-$*$ to $v_t\sigma_t$ as r tends to zero. This proves (i).

We next fix $r > 0$. For a moment we won't display the dependence in r. For instance we write v^ϵ instead of $v_t^{\epsilon,r}$ as in equation (156). Note that $\rho^\epsilon \in C^1([a,b]\times\mathbb{R}^D)$, $\rho^\epsilon > 0$ and ρ_t^ϵ is a probability density. Also $v_t^\epsilon \in C^1([a,b]\times\mathbb{R}^D, \mathbb{R}^D)$ and v^ϵ is a velocity associated to σ^ϵ. Fix $t \in [\bar{a}, \bar{b}] \subset (a,b)$. Lemma 3.2.19 gives that $||v_t^\epsilon||_{\sigma_t^\epsilon} \le C$ for all $\epsilon > 0$ small enough. By Corollary 3.2.21 $\{v_t^\epsilon \sigma_t^\epsilon\}_{\epsilon>0}$ converges weak-$*$ to $v_t\sigma_t$ as ϵ tends to zero. This proves (ii).

By Lemma 3.2.16, $t \to \bar{\Lambda}_{\sigma_t^\epsilon}(v_t^\epsilon)$ is continuous in (a,b). Hence by (ii) $t \to \bar{\Lambda}_{\sigma_t^r}(v_t^r)$ is measurable as a pointwise limit of measurable functions. We then use (i) to conclude that $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ is measurable as a pointwise limit of measurable functions. This proves (iii). The proof of (iv) is straightforward. □

3.2.5 Integration of regular pseudo 1-forms

We can now study the properties of regular pseudo 1-forms with respect to integration.

Corollary 3.2.23. *Let $\sigma \in AC_2(a,b;\mathcal{M})$ and let v be a velocity associated to σ. Suppose $t \to ||v_t||_{\sigma_t}$ is square integrable on (a,b). Then $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ is measurable and square integrable on (a,b).*

Proof. Let $\bar{\sigma}$ be the reparametrization of σ as introduced in Remark 1.2.17 and let $\bar{v}$ be the associated velocity. By Corollary 3.2.22 (iii), because $\sup_{s\in[0,L]} ||\bar{v}_s||_{\bar{\sigma}_s} \le 1$, we have that $s \to \bar{\Lambda}_{\bar{\sigma}_s}(\bar{v}_s)$ is measurable. But $\bar{\Lambda}_{\sigma_t}(v_t) = \dot{S}(t)\bar{\Lambda}_{\bar{\sigma}_{S(t)}}(\bar{v}_{S(t)})$. Thus $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ is measurable.

By Corollary 3.2.10 there exists a constant C_σ independent of t such that $||\bar{A}_{\sigma_t}||_{\sigma_t} \leq C_\sigma$ for all $t \in [a, b]$. Thus

$$|\bar{\Lambda}_{\sigma_t}(v_t)| = \Big| \int_{\mathbb{R}^D} \langle \bar{A}_{\sigma_t}, v_t \rangle d\sigma_t \Big| \leq ||\bar{A}_{\sigma_t}||_{\sigma_t} ||v_t||_{\sigma_t} \leq C_\sigma ||v_t||_{\sigma_t}.$$

Since $t \to ||v_t||_{\sigma_t}$ is square integrable, the previous inequality yields the proof. □

Corollary 3.2.24. *Suppose* $\{\sigma^r\}_{0\leq r\leq c} \subset AC_2(a, b; \mathcal{M})$, v^r *is a velocity associated to* σ^r *and* $\infty > C := \sup_{(t,r)\in E} ||v^r_t||_{\sigma^r_t}$ *where* $E := [a, b] \times [0, c]$. *Suppose that, for* $\mathcal{L}^1$*–almost every* $t \in (a, b)$, $\{v^r_t \sigma^r_t\}_{r>0}$ *converges weak-∗ to* $v_t\sigma_t$ *and* $\{\sigma^r_t\}_{r>0}$ *converges in* $\mathcal{M}$ *to* σ_t *as* r *tends to zero. If* $(t, r) \to \sigma^r_t$ *is continuous at every* $(t, 0) \in [a, b] \times \{0\}$ *then* $\lim_{r\to 0} \int_a^b \bar{\Lambda}_{\sigma^r}(v^r)dt = \int_a^b \bar{\Lambda}_\sigma(v)dt$. *Here we have set* $\sigma_t := \sigma^0_t$.

Proof. By Lemma 3.2.10 we may assume without loss of generality that $||\bar{A}_{\sigma^r_t}||_{\sigma^r_t}$ is bounded on E by a constant $\bar{C}$ independent of $(t, r) \in E$.

We obtain

$$\sup_{(t,r)\in E} |\bar{\Lambda}_{\sigma^r_t}(v^r_t)| \leq \sup_{(t,r)\in E} ||\bar{A}_{\sigma^r_t}||_{\sigma^r_t} ||v^r_t||_{\sigma^r_t} \leq \bar{C}C. \tag{163}$$

Corollary 3.2.15 ensures that $\lim_{r\to 0} \bar{\Lambda}_{\sigma^r_t}(v^r_t) = \bar{\Lambda}_{\sigma_t}(v_t)$ for $\mathcal{L}^1$–almost every $t \in [a, b]$. This, together with equation (163) shows that, as r tends to 0, the sequence of functions $t \to \bar{\Lambda}_{\sigma^r_t}(v^r_t)$ converges to the function $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ in $L^1(a, b)$. This proves the corollary. □

Definition 3.2.25. Let $\sigma \in AC_2(a, b; \mathcal{M})$ and let v be a velocity associated to σ. Suppose $t \to ||v_t||_{\sigma_t}$ is square integrable on (a, b). By Corollary 3.2.23, $t \to \bar{\Lambda}_{\sigma_t}(v_t)$ is also square integrable on (a, b). It is thus meaningful to calculate the integral $\int_a^b \bar{\Lambda}_{\sigma_t}(v_t)dt$.

We will call $\int_a^b \bar{\Lambda}_{\sigma_t}(v_t)dt$ the *integral of* $\bar{\Lambda}$ *along* (σ, v). When v is the velocity of minimal norm we will write this simply as $\int_a^b \bar{\Lambda}$ and call it the *integral of* $\bar{\Lambda}$ *along* σ.

Remark 3.2.26. Suppose that $r : [c,d] \to [a,b]$ is invertible and Lipschitz. Define $\bar\sigma_s = \sigma_{r(s)}$. Then $\bar\sigma \in AC_2(c,d;\mathcal{M})$ and $\bar v_s(x) = \dot r(s) v_{r(s)}(x)$ is a velocity for $\bar\sigma$. Furthermore, $\int_c^d \bar\Lambda_{\bar\sigma_t}(\bar v_t)dt = \int_a^b \bar\Lambda_{\sigma_t}(v_t)dt$.

Proof. Let $\beta \in L^2(a,b)$ be as in Definition 1.2.13. Then

$$W_2(\sigma_{r(s+h)}, \sigma_{r(s)}) \le \Big|\int_{r(s)}^{r(s+h)} \beta(t)dt\Big| = \Big|\int_s^{s+h} \bar\beta(\tau)d\tau\Big| \quad \text{where} \quad \bar\beta(s) := |\dot r(s)|\beta(r(s)).$$

Because $\bar\beta \in L^2(c,d)$ we conclude that $\bar\sigma \in AC_2(c,d;\mathcal{M})$. Direct computations give that, for $\mathcal{L}^1$ a.e. $s \in (c,d)$,

$$\lim_{h\to 0} W_2(\sigma_{r(s+h)}, \sigma_{r(s)})/|h| = |\dot r(s)|\,|\sigma'|(r(s)).$$

Thus $|\bar\sigma'|(s) = |\dot r(s)|\,|\sigma'|(r(s))$. Let $\phi \in C_c^\infty(\mathbb{R}^D)$ and let v be a velocity for σ (see Proposition 1.2.15). The chain rule shows that, in the sense of distributions,

$$\frac{d}{ds}\int_{\mathbb{R}^D} \phi d\sigma_{r(s)} = \dot r(s)\langle \nabla\phi, v_{r(s)}\rangle_{\sigma_{r(s)}} = \langle \nabla\phi, \bar v_s\rangle_{\bar\sigma_s},$$

where $\bar v_s(x) = \dot r(s) v_{r(s)}(x)$. Thus $\bar v$ is a velocity for $\bar\sigma$. Using the linearity of $\bar\Lambda$ we have

$$\int_c^d \bar\Lambda_{\bar\sigma_s}(\bar v_s)ds = \int_c^d \dot r(s)\bar\Lambda_{\sigma_{r(s)}}(v_{r(s)})ds = \int_a^b \bar\Lambda_{\sigma_t}(v_t)dt.$$

This concludes the proof. □

3.2.6 Green's formula for annuli, the first cohomology of regular pseudo 1-forms

Let $\sigma \in AC_2(a,b;\mathcal{M})$ and let v be its velocity of minimal norm (see Proposition 1.2.15). The following proposition is extracted from [3] Theorem 8.3.1 and Proposition 8.4.5.

Proposition 3.2.27. *Let $\mathcal{N}_1$ be the set of t such that v_t fails to be in $T_{\sigma_t}\mathcal{M}$. Let $\mathcal{N}_2$ be the set of $t \in [a,b]$ such that $\Big(\pi^1 \times (\pi^2 - \pi^1)/h\Big)_{\#}\eta_h$ fails to converge to $(Id \times v_t)_{\sigma_t}$ in the Wasserstein space $\mathcal{M}(\mathbb{R}^D \times \mathbb{R}^D)$, for some $\eta_h \in \Gamma_o(\sigma_t, \sigma_{t+h})$. Let $\mathcal{N}$ be the union of $\mathcal{N}_1$ and $\mathcal{N}_2$. Then $\mathcal{L}^1(\mathcal{N}) = 0$.*

As in Section 3.2.3, for $r \in (0,1)$ and $s \in [r,1]$ we define

$$D_s z := sz, \quad \sigma(s,t) = \sigma_t^s := D_{s\#}\sigma_t \tag{164}$$

and

$$w(s,t,\cdot) = w_t^s(z) := \frac{z}{s} = D_s^{-1}z, \quad v(s,t,\cdot) = v_t^s := D_{s*}v_t \tag{165}$$

According to Lemma 3.1.4, for each $s \in [r,1]$, $\sigma(s,\cdot) \in AC_2(a,b;\mathcal{M})$ admits $v(s,\cdot)$ as a velocity. For each t and $\phi \in C_c^\infty(\mathbb{R}^D)$, in the sense of distributions,

$$\frac{d}{ds}\int_{\mathbb{R}^D} \phi d\sigma_t^s = \frac{d}{ds}\int_{\mathbb{R}^D} \phi(sx)d\sigma_t(x) = \int_{\mathbb{R}^D} \langle \nabla\phi(sx), x\rangle d\sigma_t(x) = \int_{\mathbb{R}^D} \langle \nabla\phi, w_t^s\rangle d\sigma_t^s.$$

Thus $w(\cdot,t)$ is a velocity for $\sigma(\cdot,t)$.

We assume that

$$||\sigma'||_\infty := \sup_{t\in[a,b]} ||v_t||_{\sigma_t} < \infty.$$

By Remark 1.2.14,

$$c_\sigma^0 := \sup_{t\in[a,b]} W_2(\sigma_t, \delta_0) < \infty.$$

By the fact that $D_{s\#}\sigma_t = \sigma_t^s$ we have

$$W_2^2(\sigma_t^s, \delta_0) = s^2 W_2^2(\sigma_t, \delta_0) \le s^2 c_\sigma^0 \le \bar{C}_\sigma. \tag{166}$$

Here, we are free to choose $\bar{C}_\sigma$ to be any constant greater than c_σ^0.

Remark 3.2.28. Note that $(1+h/s)Id$ pushes σ_t^s forward to σ_t^{s+h} and is the gradient of a convex function. Thus

$$\gamma^h := \Big(Id \times (1+h/s)Id\Big)_\# \sigma_t^s \in \Gamma_o(\sigma_t^s, \sigma_t^{s+h}).$$

For γ^h–almost every $(x,y) \in \mathbb{R}^D \times \mathbb{R}^D$ we have $y = (1+h/s)x$, so

$$v_t^{s+h}(y) = (s+h)v_t(\frac{y}{s+h}) = (1+\frac{h}{s})v_t^s(\frac{sy}{s+h}) = (1+\frac{h}{s})v_t^s(x). \tag{167}$$

Using the definition of σ_t^s and v_t^s we obtain the identities

$$||Id||_{\sigma_t^s} = s||Id||_{\sigma_t} \le s\bar{C}_\sigma, \quad ||v_t^s||_{\sigma_t^s} = s||v_t||_{\sigma_t} \le s||\sigma'||_\infty. \tag{168}$$

We use the first identity in equation (168) and the fact that $(1+h/s)Id$ pushes σ_t^s forward to σ_t^{s+h} to obtain

$$W_2^2(\sigma_t^s, \sigma_t^{s+h}) = \frac{h^2}{s^2}||Id||^2_{\sigma_t^s} = h^2||Id||^2_{\sigma_t} = h^2 W_2^2(\sigma_t, \delta_0) \leq h^2 \bar{C}_\sigma^2. \tag{169}$$

Set

$$V(s,t) := \bar{\Lambda}_{\sigma_t^s}(v_t^s), \qquad W(s,t) := \bar{\Lambda}_{\sigma_t^s}(w_t^s)$$

Lemma 3.2.29. *For each $t \in (a,b) \setminus \mathcal{N}$, the function $V(t,\cdot)$ is differentiable on $(r,1)$ and its derivative is bounded by a constant $L_1(r)$ depending only on σ and r. Furthermore*

$$\partial_s V(s,t) = \int_{\mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(x), \frac{v_t^s(x)}{s} \rangle d\sigma_t^s(x) + \int_{\mathbb{R}^D} \langle B_{\sigma_t^s}(x) w_t^s(x), v_t^s(x) \rangle d\sigma_t^s(x).$$

Proof. Let $C_\sigma(r)$ be as in Corollary 3.2.13 and let $\bar{C}_\sigma$ be as in equation (166). We use equations (131), (167) and then Hölder's inequality to obtain

$$|V(s+h,t) - V(s,t)| \leq \frac{h}{s}||\bar{A}_{\sigma_t^s}||_{\sigma_t^s}\, ||v_t^s||_{\sigma_t^s} + 2c(\bar{\Lambda}) W_2(\sigma_t^s, \sigma_t^{s+h}) ||v_t^{s+h}||_{\sigma_t^{s+h}}. \tag{170}$$

We combine equations (168), (169) and (170) to conclude that

$$|V(s+h,t) - V(s,t)| \leq h C_\sigma(r)||\sigma'||_\infty + 2hc(\bar{\Lambda})\bar{C}_\sigma(s+h)||\sigma'||_\infty. \tag{171}$$

This proves that $V(\cdot,t)$ is Lipschitz on $(r,1)$ and that its derivative is bounded by a constant $L_1(r)$. As in Remark 3.2.8,

$$\begin{aligned} \lim_{h\to 0} & \frac{V(s+h,t) - V(s,t)}{h} \\ &= \lim_{h\to 0} \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_{\sigma_t^s}(x), \frac{v_t^{s+h}(y) - v_t^s(x)}{h} \rangle d\gamma^h(x,y) \\ &+ \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle B_{\sigma_t^s}(x) \frac{y-x}{h}, v_t^{s+h}(y) \rangle d\gamma^h(x,y) \\ &+ \frac{1}{h} \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_{\sigma_t^{s+h}}(y) - \bar{A}_{\sigma_t^s}(x) - B_{\sigma_t^s}(x)(y-x), v_t^{s+h}(y) \rangle \gamma^h(x,y). \end{aligned} \tag{172}$$

By equation (129), the last inequality in equation (168) and equation (169) we have

$$\lim_{h\to 0} \frac{1}{h} \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_{\sigma_t^{s+h}}(y) - \bar{A}_{\sigma_t^s}(x) - B_{\sigma_t^s}(x)(y-x), v_t^{s+h}(y) \rangle \gamma^h(x,y) = 0. \tag{173}$$

We use equations (167), (172), (173) and the fact that, for $\gamma_t^{s,h}$–almost every $(x,y)\in\mathbb{R}^D\times\mathbb{R}^D$, $y=(1+h/s)x$ to conclude that

$$\begin{aligned}
&\lim_{h\to 0}\frac{V(s+h,t)-V(s,t)}{h}\\
&=\int_{\mathbb{R}^D}\langle\bar{A}_{\sigma_t^s}(x),\frac{v_t^s(x)}{s}\rangle d\sigma_t^s(x)+\lim_{h\to 0}\int_{\mathbb{R}^D}\langle B_{\sigma_t^s}(x)\frac{x}{s},(1+\frac{h}{s})v_t^s(x)\rangle d\sigma_t^s(x)\\
&=\int_{\mathbb{R}^D}\langle\bar{A}_{\sigma_t^s}(x),\frac{v_t^s(x)}{s}\rangle d\sigma_t^s(x)+\int_{\mathbb{R}^D}\langle B_{\sigma_t^s}(x)w_t^s(x),v_t^s(x)\rangle d\sigma_t^s(x).
\end{aligned}\tag{174}$$

This proves the lemma. □

Lemma 3.2.30. *For each $s\in[r,1]$ and $t\in(a,b)\setminus\mathcal{N}$, the function $W(s,\cdot)$ is differentiable at t and its derivative is bounded by a constant $L_2(r)$ depending only on σ and r. Furthermore*

$$\partial_t W(s,t)=\int_{\mathbb{R}^D}\langle\bar{A}_{\sigma_t^s}(x),\frac{v_t^s(x)}{s}\rangle d\sigma_t^s(x)+\int_{\mathbb{R}^D}\langle w_t^s(x),B_{\sigma_t^s}(x)v_t^s(x)\rangle d\sigma_t^s(x).$$

Proof. We would like to apply Lemmas 3.2.17 and 3.2.18 with X_t substituted by w_t^s and σ_t substituted by σ_t^s. It suffices to show that if $t\in(a,b)\setminus\mathcal{N}$ and $\gamma_h^s\in\Gamma_o(\sigma_t^s,\sigma_{t+h}^s)$ then $\Big(\pi^1\times(\pi^2-\pi^1)/h\Big)_{\#}\gamma_h^s$ converges to $(Id\times v_t^s)_{\sigma_t^s}$ in $\mathcal{M}(\mathbb{R}^D\times\mathbb{R}^D)$ as h tends to 0. Set

$$\gamma_h:=\big(D_s^{-1}\times D_s^{-1}\big)_{\#}\gamma_h^s.$$

where D_s is given in equation (164). Since

$$\pi^1\circ\big(D_s^{-1}\times D_s^{-1}\big)=D_s^{-1}\circ\pi^1\quad\text{and}\quad\pi^2\circ\big(D_s^{-1}\times D_s^{-1}\big)=D_s^{-1}\circ\pi^2,$$

we conclude that $\gamma_h\in\Gamma(\sigma_t,\sigma_{t+h})$. By the fact that the support of γ_h^s is cyclically monotone, we have that the support of γ_h is also cyclically monotone. Hence $\gamma_h\in\Gamma_o(\sigma_t,\sigma_{t+h})$. We have

$$\big(\pi^1\times\frac{\pi^2-\pi^1}{h}\big)_{\#}\gamma_h^s=\big(D_s\times D_s\big)_{\#}\Big((\pi^1\times\frac{\pi^2-\pi^1}{h})_{\#}\gamma_h\Big)\to\big(D_s\times D_s\big)\circ(Id\times v_t)_{\#}\sigma_t=(Id\times v_t^s)_{\#}\sigma_t^s.$$

□

Corollary 3.2.31. *For each $s \in (r,1)$ and $t \in (a,b) \setminus \mathcal{N}$ we have*

$$\partial_t\Big(\bar\Lambda_{\sigma_t^s}(w_t^s)\Big) - \partial_s\Big(\bar\Lambda_{\sigma_t^s}(v_t^s)\Big) = d\bar\Lambda_{\sigma_t^s}(v_t^s, w_t^s).$$

Proof. This corollary is a direct consequence of Lemmas 3.2.12, 3.2.29 and 3.2.30. □

Remark 3.2.32. Notice that, unlike in Proposition 3.2.2, in Lemma 3.2.29 and Corollary 3.2.31 we don't assume that $v \in C^1((r,1]\times(a,b)\times\mathbb{R}^D, \mathbb{R}^D)$. Although possibly neither ∇v_t^s nor $\partial_s v_t^s$ exist, equation (167) ensures that $||v_t^{s+h}\circ\pi^2 - v_t^s\circ\pi^1||_{\gamma^h} \le h||\sigma'||_\infty$. That inequality was crucial in the proof of Lemma 3.2.29.

Theorem 3.2.33 (Green's formula on the annulus)**.** *Consider in $\mathcal{M}$ the surface $S(s,t) = D_{s\#}\sigma$ for $(s,t) \in [r,1]\times[0,T]$ and its boundary ∂S which is the union of the negatively oriented curves $S(r,\cdot)$, $S(\cdot,T)$ and the positively oriented curves $S(1,\cdot)$, $S(\cdot,0)$. Then*

$$\int_S d\bar\Lambda = \int_{\partial S}\bar\Lambda.$$

Proof. We use Corollary 3.2.31 to obtain

$$\begin{aligned}
\int_S d\bar\Lambda &= \int_0^T dt\int_r^1 d\bar\Lambda_{S(s,t)}(v_t^s, w_t^s)ds\\
&= \int_0^T dt\int_r^1 \Big[\partial_t\Big(\bar\Lambda_{S(s,t)}(w_t^s)\Big) - \partial_s\Big(\bar\Lambda_{S(s,t)}(v_t^s)\Big)\Big]ds\\
&= \int_r^1\Big(\bar\Lambda_{S(s,T)}(w_T^s) - \bar\Lambda_{S(s,0)}(w_0^s)\Big)ds - \int_0^T\Big(\bar\Lambda_{S(1,t)}(v_t^1) - \bar\Lambda_{S(r,t)}(v_t^r)\Big)dt\\
&= \int_{\partial S}\bar\Lambda. \qquad (175)
\end{aligned}$$

□

Corollary 3.2.34. *If we further assume that $\bar\Lambda$ is a closed pseudo 1–form and that $\sigma_0 = \sigma_T$, then $\int_0^T \bar\Lambda_{\sigma_t}(v_t)dt = 0$.*

Proof. For $s \in [r,1]$ define

$$l(s) = \int_0^T \bar\Lambda_{S(s,t)}(v_t^s)dt, \qquad \bar l(t) = \int_r^1 \bar\Lambda_{S(s,t)}(w_t^s)ds.$$

Since $w_T^s = w_0^s$ and $\sigma_T^s = D_{s\#}\sigma_T = D_{s\#}\sigma_0 = \sigma_0^s$, we have $\bar{l}(T) = \bar{l}(0)$. This, together with equation (175) and the fact that $d\bar{\Lambda} = 0$, yields $\int_0^T \bar{\Lambda}_{\sigma_t}(v_t)dt = l(1) = l(r)$. But

$$|l(r)| \leq \int_0^T |\bar{\Lambda}_{S(s,t)}(v_t^r)|dt \leq \int_0^T ||\bar{A}_{S(s,t)}||_{S(s,t)}||v_t^r||_{S(s,t)}dt \leq r||\sigma'||_\infty \int_0^T ||\bar{A}_{S(s,t)}||_{S(s,t)}dt, \tag{176}$$

where we have used the last inequality in equation (168). The first inequality in equation (166) shows that, for r small enough, $\{S(s,t)\}_{t\in\times[0,T])}$ is contained in a small ball centered at δ_0. But Lemma 3.2.10 gives that $\mu \to ||\bar{A}_\mu||_\mu$ is continuous at δ_0. Thus there exist constants c and r_0 such that $||\bar{A}_{S(s,t)}||_{S(s,t)} \leq c$ for all $t \in [0,T]$ and all $r < r_0$. We can now exploit equation (176) to obtain

$$|l(1)| = \liminf_{r\to 0} |l(r)| \leq \liminf_{r\to 0} rTc||\sigma'||_\infty = 0.$$

□

Corollary 3.2.35. *Let $\bar{\Lambda}$ be a regular pseudo 1–form on $\mathcal{M}$. Let Λ denote the corresponding 1–form on $\mathcal{M}$, defined by restriction. Assume $\bar{\Lambda}$ is closed, i.e. $d\bar{\Lambda} = 0$. Then Λ is exact, i.e. there exists a differentiable function F on $\mathcal{M}$ such that $dF = \Lambda$.*

Proof. Fix $\mu \in \mathcal{M}$. Let σ be any curve in $AC_2(a,b;\mathcal{M})$ such that $\sigma_a = \delta_0$ and $\sigma_b = \mu$. Assume that v is its velocity of minimal norm and that $\sup_{(a,b)} ||v_t||_{\sigma_t} < \infty$. By Corollary 3.2.34, $\int_\sigma \bar{\Lambda}$ depends only on μ, *i.e.* it is independent of the path σ. Also, Remark 3.2.26 ensures that $\int_\sigma \bar{\Lambda}$ is independent of a, b. It is thus meaningful to define

$$F(\mu) := \int_\sigma \bar{\Lambda}.$$

We now want to show that F is differentiable. Fix $\mu, \nu \in \mathcal{M}$ and $\gamma \in \Gamma_o(\mu,\nu)$. Define $\sigma_t := ((1-t)\pi^1 + t\pi^2)_\#\gamma$. Then $\sigma : [0,1] \to \mathcal{M}$ is a constant speed geodesic between μ and ν. Let v_t denote its velocity of minimal norm. Clearly,

$$F(\nu) - F(\mu) = \int_0^1 \bar{\Lambda}_{\sigma_t}(v_t)dt. \tag{177}$$

Let $\bar{\gamma} : \mathbb{R}^D \to \mathbb{R}^D$ denote the *barycentric projection* of γ. Set $v := \bar{\gamma} - Id$. Then $\gamma_t := (\pi^1, (1-t)\pi^1 + t\pi^2)_{\#}\gamma \in \Gamma_o(\sigma_0, \sigma_t)$ and

$$\begin{aligned}\bar{\Lambda}_{\sigma_t}(v_t) - \bar{\Lambda}_{\sigma_0}(v) &= \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_{\sigma_0}(x), v_t(y) - v(x)\rangle + \langle B_{\sigma(0)}(x)(y-x), v_t(y)\rangle d\gamma_t(x,y) \\ &\quad + \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_{\sigma_t}(y) - \bar{A}_{\sigma_0}(x) - B_{\sigma_0}(x)(y-x), v_t(y)\rangle d\gamma_t(x,y).\end{aligned}$$

By equation (126) and Hölder's inequality,

$$\Big|\int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_{\sigma_t}(y) - \bar{A}_{\sigma_0}(x) - B_{\sigma_0}(x)(y-x), v_t(y)\rangle d\gamma_t(x,y)\Big| \leq o(W_2(\sigma_0,\sigma_t))\, ||v_t||_{\sigma_t}.$$

It is well known (cfr. [3] Lemma 7.2.1) that if $0 < t \leq 1$ then there exists a unique optimal transport map T_t^1 between σ_t and σ_1, *i.e.* $\Gamma_o(\sigma_t,\sigma_1) = \{(Id \times T_t^1)_{\#}\sigma_t\}$. One can check that $v_t(y) = \frac{T_t^1(y) - y}{1-t}$ and $||v_t||_{\sigma_t} = W_2(\sigma_t,\sigma_1)/(1-t) = W_2(\sigma_0,\sigma_1)$. Thus

$$\begin{aligned}&\int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_{\sigma_0}(x), v_t(y) - v(x)\rangle d\gamma_t(x,y) \\ &= \int \langle \bar{A}_{\sigma_0}(x), \frac{T_t^1(y) - y}{1-t} - (\bar{\gamma}(x) - x)\rangle d\gamma_t(x,y) \\ &= \int \langle \bar{A}_{\sigma_0}(x), \frac{z - ((1-t)x + tz)}{1-t} - (z-x)\rangle d\gamma(x,z) \\ &= 0.\end{aligned}$$

Similarly,

$$\begin{aligned}\int_{\mathbb{R}^D\times\mathbb{R}^D} \langle B_{\sigma_0}(x)(y-x), v_t(y)\rangle d\gamma_t(x,y) &= t\int_{\mathbb{R}^D\times\mathbb{R}^D} \langle B_{\sigma_0}(x)(z-x), z-x\rangle d\gamma(x,y) \\ &= o(W_2(\sigma_0,\sigma_1)) = o(W_2(\mu,\nu)).\end{aligned}$$

Combining these equations shows that

$$\bar{\Lambda}_{\sigma_t}(v_t) - \bar{\Lambda}_{\sigma_0}(v) = o(W_2(\mu,\nu)). \tag{178}$$

Notice that (178) is independent of t. Combining (177) and (178) we find

$$\begin{aligned}F(\nu) &= F(\mu) + \bar{\Lambda}_{\sigma_0}(v) + \int_0^1 \bar{\Lambda}_{\sigma_t}(v_t) - \bar{\Lambda}_{\sigma_0}(v)dt \\ &= F(\mu) + \bar{\Lambda}_{\sigma_0}(v) + o(W_2(\mu,\nu)) \\ &= F(\mu) + \int_{\mathbb{R}^D\times\mathbb{R}^D} \langle \bar{A}_{\sigma_0}(x), y-x\rangle d\gamma(x,y) + o(W_2(\mu,\nu)).\end{aligned}$$

As in Definition 1.2.8, this proves that F is differentiable and that $\nabla_\mu F = \pi_\mu(\bar{A}_\mu)$. Thus $dF = \Lambda$. $\square$

3.2.7 Example: Restriction of 1-forms to the space of discrete measures

Fix an integer $n \geq 1$. Given $x_1, \cdots, x_n \in \mathbb{R}^D$, set $\mathbf{x} := (x_1, \cdots, x_n)$ and $\mu_\mathbf{x} := 1/n \sum_{i=1}^n \delta_{x_i}$. Let M denote the set of such measures and TM denote its tangent bundle. It is easy to see $T_{\mu_\mathbf{x}} M = L^2(\mu_\mathbf{x})$. Choose a regular pseudo 1–form $\bar{\Lambda}$ on $\mathcal{M}$. By restriction we obtain a 1–form α on M, defined by $\alpha_\mathbf{x} := \bar{\Lambda}_{\mu_\mathbf{x}}$.

Let $A : \mathbb{R}^{nD} \to \mathbb{R}^{nD}$ be defined by

$$A(\mathbf{x}) = (A_1(\mathbf{x}), \cdots, A_n(\mathbf{x})) := \Big(\bar{A}_{\mu_\mathbf{x}}(x_1), \cdots, \bar{A}_{\mu_\mathbf{x}}(x_n)\Big).$$

Let $X = (X_1, \cdots, X_n) \in \mathbb{R}^{nD}$ be such that $X_i = X_j$ whenever $x_i = x_j$. We identify X with $\tilde{X} \in L^2(\mu_\mathbf{x})$ defined by $\tilde{X}(x_i) = X_i$. Note that $\alpha_\mathbf{x}(X) = \frac{1}{n}\langle A(\mathbf{x}), X\rangle$. Now define a $nD \times nD$ matrix $B(\mathbf{x})$ by setting

$$B_{k+i,k+j} := \Big(B_{\mu_\mathbf{x}}(x_{k+1})\Big)_{ij}, \qquad \text{for } k = 0, \cdots n-1, \quad i, j = 1, \cdots, D, \tag{179}$$

$$B_{l,m} := 0 \quad \text{if} \quad (l,m) \notin \{(k+i, k+j) : k = 0, \cdots n-1, \quad i, j = 1, \cdots, D\}. \tag{180}$$

Proposition 3.2.36. *The map $A : \mathbb{R}^{nD} \to \mathbb{R}^{nD}$ is differentiable and $\nabla A(\mathbf{x}) = B(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^{nD}$.*

Proof. Let $\mathbf{x} = (x_1, \cdots, x_n) \in \mathbb{R}^{nD}$. Set $r := \min_{x_i \neq x_j} |x_i - x_j|$. If $\mathbf{y} = (y_1, \cdots, y_n) \in \mathbb{R}^{nD}$ and $|\mathbf{y} - \mathbf{x}| < r/2$ then $\Gamma_o(\mu_\mathbf{x}, \mu_\mathbf{y})$ has a single element $\gamma_\mathbf{y} = 1/n \sum_{i=1}^n \delta_{(x_i, y_i)}$ and $nW_2^2(\mu_\mathbf{x}, \mu_\mathbf{y}) = |\mathbf{y} - \mathbf{x}|^2$. By equation (126),

$$|A(\mathbf{y}) - A(\mathbf{x}) - B(\mathbf{x})(\mathbf{y} - \mathbf{x})|^2 = n\, o(\frac{|\mathbf{y} - \mathbf{x}|^2}{n}). \tag{181}$$

This concludes the proof. $\square$

Lemma 3.2.37. *Suppose* $\mathbf{x} = (x_1, \cdots, x_n) \in \mathbb{R}^{nD}$ *and* $X = (X_1, \cdots, X_n)$*,* $Y = (Y_1, \cdots, Y_n) \in \mathbb{R}^{nD}$ *are such that* $X_i = X_j$*,* $Y_i = Y_j$ *whenever* $x_i = x_j$*. Recall that by abuse of notation, define* X *with a vector field we still denote* $X \in L^2(\mu_{\mathbf{x}})$*. Similarly, we identify* Y *with a vector field we still denote* $Y \in L^2(\mu_{\mathbf{y}})$*.*

$$d\bar{\Lambda}_{\mu_{\mathbf{x}}}(X, Y) = d\alpha_{\mathbf{x}}(X, Y).$$

Proof. We use Lemma 135 and equations (179)– (180) to obtain

$$d\bar{\Lambda}_{\mu_{\mathbf{x}}}(X, Y) = \sum_{k=1}^{n} \Big\langle (B_{\mu_{\mathbf{x}}}(x_k) - B_{\mu_{\mathbf{x}}}(x_k)^T) X_k, Y_k \Big\rangle = d\alpha_{\mathbf{x}}(X, Y).$$

□

Corollary 3.2.38. *Suppose that* $\mathbf{r} = (r_1, \cdots, r_n) \in C^2([0, T], \mathbb{R}^{nD})$ *and set* $\sigma_t := 1/n \sum_{i=1}^n \delta_{r_i(t)}$*. If* $\bar{\Lambda}$ *is closed and* $\sigma_0 = \sigma_T$ *then* $\int_\sigma \alpha = 0$*.*

Proof. This is a direct consequence of Corollary 3.2.34.

□

Remark 3.2.39. One can check by direct computation that the familiar identity $\partial_t(\alpha_{\mathbf{x}}(\partial_s \mathbf{x})) - \partial_s(\alpha_{\mathbf{x}}(\partial_t \mathbf{x})) = d\alpha_{\mathbf{x}}(\partial_t \mathbf{x}, \partial_s \mathbf{x})$ holds. Together with Lemma 3.2.37 this shows that

$$\partial_t\Big(\bar{\Lambda}_{\sigma_t^s}(w_t^s)\Big) - \partial_s\Big(\bar{\Lambda}_{\sigma_t^s}(v_t^s)\Big) = d\bar{\Lambda}_{\sigma_t^s}(v_t^s, w_t^s),$$

which we used to prove Theorem 3.2.34.

Remark 3.2.40. Notice that the assumption $\sigma_0 = \sigma_T$ is weaker than $\mathbf{r}(0) = \mathbf{r}(T)$.

REFERENCES

[1] Agueh, M., "Asymptotic behavior for doubly degenerate parabolic equations," *C. R. Math. Acad. Sci. Paris*, vol. 337, no. 5, pp. 331–336, 2003.

[2] Ambrosio, L. and Gangbo, W., "Hamiltonian ODEs in the Wasserstein space of probability measures," *Comm. Pure Appl. Math.*, vol. 61, no. 1, pp. 18–53, 2008.

[3] Ambrosio, L., Gigli, N., and Savaré, G., *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich, Basel: Birkhäuser Verlag, second ed., 2008.

[4] Benamou, J.-D. and Brenier, Y., "Weak existence for the semigeostrophic equations formulated as a coupled Monge-Ampère/transport problem," *SIAM J. Appl. Math.*, vol. 58, no. 5, pp. 1450–1461 (electronic), 1998.

[5] Brenier, Y., "Polar factorization and monotone rearrangement of vector-valued functions," *Comm. Pire Appl. Math*, vol. 44, no. 4, pp. 475–417, 1991.

[6] Brenier, Y., "Convergence of the Vlasov-Poisson system to the incompressible Euler equations," *Comm. Partial Differential Equations*, vol. 25, no. 3-4, pp. 737–754, 2000.

[7] Brenier, Y. and Loeper, G., "A geometric approximation to the Euler equations: the Vlasov-Monge-Ampère system," *Geom. Funct. Anal.*, vol. 14, no. 6, pp. 1182–1218, 2004.

[8] Carlen, E. A. and Gangbo, W., "Constrained steepest descent in the 2-Wasserstein metric," *Ann. of Math. (2)*, vol. 157, no. 3, pp. 807–846, 2003.

[9] Carrillo, J. A., McCann, R. J., and Villani, C., "Contractions in the 2-Wasserstein length space and thermalization of granular media," *Arch. Ration. Mech. Anal.*, vol. 179, no. 2, pp. 217–263, 2006.

[10] Cullen, M. and Feldman, M., "Lagrangian solutions of semigeostrophic equations in physical space," *SIAM J. Math. Anal.*, vol. 37, no. 5, pp. 1371–1395 (electronic), 2006.

[11] Cullen, M. and Gangbo, W., "A variational approach for the 2-dimensional semi-geostrophic shallow water equations," *Arch. Ration. Mech. Anal.*, vol. 156, no. 3, pp. 241–273, 2001.

[12] Cullen, M., Gangbo, W., and Pisante, G., "The semigeostrophic equations discretized in reference and dual variables," *Arch. Ration. Mech. Anal.*, vol. 185, no. 2, pp. 341–363, 2007.

[13] Evans, L. C. and Gariepy, R. F., *Measure theory and fine properties of functions*. Studies in Advanced Mathematics, Boca Raton, FL: CRC Press, 1992.

[14] Gangbo, w., Kim, H. K., and Pacini, T., "Differential forms on wasserstein space and infinite dimensional hamiltonian systems," *Preprint*.

[15] Gangbo, W. and McCann, R. J., "The geometry of optimal transportation," *Acta Math.*, vol. 177, no. 2, pp. 113–161, 1996.

[16] Gangbo, W. and Tudorascu, A., *A weak KAM theorem for the nonlinear Vlasov equation*. in preparation.

[17] Jordan, R., Kinderlehrer, D., and Otto, F., "The variational formulation of the Fokker-Planck equation," *SIAM J. Math. Anal.*, vol. 29, no. 1, pp. 1–17 (electronic), 1998.

[18] Khesin, B. and Lee, P., "Poisson geometry and first integrals of geostrophic equations," *Phys. D*, vol. 237, no. 14-17, pp. 2072–2077, 2008.

[19] Lieb, E. H. and Loss, M., *Analysis*, vol. 14 of *Graduate Studies in Mathematics*. Providence, RI: American Mathematical Society, 2001.

[20] Loeper, G., "Uniqueness of the solution to the Vlasov-Poisson system with bounded density," *J. Math. Pures Appl. (9)*, vol. 86, no. 1, pp. 68–79, 2006.

[21] Lott, J., "Some geometric calculations on Wasserstein space," *Comm. Math. Phys.*, vol. 277, no. 2, pp. 423–437, 2008.

[22] Lott, J. and Villani, C., "Ricci curvature for metric-measure space via optimal transport," *Ann. of Math.*

[23] Marsden, J. E. and Weinstein, A., "The Hamiltonian structure of the Maxwell-Vlasov equations," *Phys. D*, vol. 4, no. 3, pp. 394–406, 1981/82.

[24] McCann, R. J., "A convexity principle for interacting gases," *Adv. Math.*, vol. 128, no. 1, pp. 153–179, 1997.

[25] Otto, F., "The geometry of dissipative evolution equations: the porous medium equation," *Comm. Partial Differential Equations*, vol. 26, no. 1-2, pp. 101–174, 2001.

[26] Pauli, W., *General principles of quantum mechanics*. Berlin: Springer-Verlag, 1980. Translated from the German by P. Achuthan and K. Venkatesan, With an introduction by Charles P. Enz.

[27] Sturm, K.-T., "On the geometry of metric measure spaces. I," *Acta Math.*, vol. 196, no. 1, pp. 65–131, 2006.

[28] Sturm, K.-T., "On the geometry of metric measure spaces. II," *Acta Math.*, vol. 196, no. 1, pp. 133–177, 2006.

[29] Villani, C., *Topics in optimal transportation*, vol. 58 of *Graduate Studies in Mathematics*. Providence, RI: American Mathematical Society, 2003.

[30] Zheng, Y. X. and Majda, A., "Existence of global weak solutions to one-component Vlasov-Poisson and Fokker-Planck-Poisson systems in one space dimension with measures as initial data," *Comm. Pure Appl. Math.*, vol. 47, no. 10, pp. 1365–1401, 1994.