THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITE BOURGOGNE FRANCHE-COMTE
PREPAREE A l'Université de Bourgogne

Ecole doctorale n° ED553 Carnot-Pasteur

Doctorat de Mathématiques

Par Monsieur DING Hao

Newton flows, Stochastic parallel translations, Q-Wiener processes and Dean-Kawasaki equations on the Wasserstein space

Thèse présentée et soutenue à Academy of Mathematics and systems Sciences, CAS, Beijing,

le 28 juin 2022 à 10h00

Composition du Jury :

Jaramillo, José-Luis	Professeur, Université de Bourgogne,	France	Président
Wang, Fengyu	Professeur, Université de Tianjin,	Chine	Rapporteur
Juillet, Nicolas	Professeur, Université de Haute Alsace,	France	Rapporteur
Dong, Zhao	Professeur, Academy of Mathematics and systems Sciences, CAS,	Chine,	Examinateur
Fang, Shizan	Professeur, Université de Bourgogne,	France	Directeur de thèse
Li, Xiang-dong	Professeur, Academy of Mathematics and systems Sciences, CAS,	Chine	Codirecteur de thèse

Titre: Flots de Newton, transports parallèles stochastiques, Q-processus de Wiener, et équation de Dean- Kawasaki sur l'espace de Wasserstein

Résumé:

Nous allons introduire des flots de Newton sur l'espace de Wasserstein. L'existence et l'unicité de l'équation de Newton avec le problème de Cauchy sera établie. Nous allons éclaircir également les liens entre l'équation d'écoulement de Newton relâchée et l'équation de Keller-Segel.

Nous allons étendre lad éfinition de la connexion de Levi-Civita de Lott à l'espace de Wasserstein des mesures de probabilité ayant densité et divergence de tel sorte que les transports parallèles puissent être définis comme en géométrie différentielle. Nous allons démontrer l'existence des transports parallèles au sens fort de Lott pour le cas du tore.

Nous allons établir un formalisme intrinsèque pour le calcul stochastique d'Itô sur l'espace de Wasserstein à travers les trois fonctionnelles typiques. Nous allons construire la forme faible et la forme forte de l'équation différentielle partielle stochastique définissant le transport parallèle, dont l'existence et l'unicité est démontrée dans le cas du tore. Des processus de diffusion non-dégénérée sont construits en utilisant les fonctions propres du laplacian.

Nous allons construire une nouvelle approche du système d'interaction de particules aux solutions du problème de martingale pour l'équation de Dean-Kawasaki sur le tore sous une condition plus faible portant sur l'intensité de corrélation spatiale.

Title: Newton flows, stochastic parallel translations, Q-Wiener processes and Dean-Kawasaki equations on the Wasserstein space

Abstract:

We introduce Newton flows on the Wasserstein space and prove the well-posedness of Cauchy problem of the Newton flow equation. We show the connections between the relaxed Newton flow equation and the Keller-Segel equation.

We extend the definition of Lott's Levi-Civita connection to the Wasserstein space of probability measures having density and divergence so that parallel translations for can be introduced as done in differential geometry. In the case of torus, we prove the well-posedness of Lott's equation for parallel translations.

We establish an intrinsic formalism for Itô stochastic calculus on the Wasserstein space throughout three kinds of functionals. We construct the weak and strong form of stochastic partial differential equations for stochastic parallel translations, the well-posedness is also proved in the case of torus. As a kind of non-degenerated diffusion process on Wasserstein spaces, Q-Wiener process is constructed using the eigenfunctions of the Laplacian.

We construct a new interactive particle model approximation to the solution to the regularized martingale problem of the diffusive Dean-Kawasaki equation on the one-dimensional torus under a weaker condition on the spatial correlation intensity of the noise than the classical one.

Contents

Acknowledgments

This thesis was completed under the careful guidance of professor Xiangdong Li and professor Shizan Fang. First of all, I would like to thank professor Li for his careful cultivation from the beginning of my doctoral career. He led me into fields of optimal transport, random matrices, stochastic differential geometry and so on. He encouraged me to do original work. His profound academic achievements and the deep insights deeply influenced me. When I was in the third year of my Ph.D. career, he recommended me to visit France for a joint training program. When COVID-19 was spreading, He helped me a lot and support me unconditionally when I returned to China. In the last year of Ph.D., my knees injuried, He helped me solving problems in medical reimbursement and living housing. Professor Li is just like one of my relatives in Beijing. I sincerely thank him for his care for me over the past six years! I will continue to work hard in the future.

I would like to thank professor Shizan Fang. During my joint training program in France, professor Fang gave me the greatest freedom both in study and life. He always demonstrated how to deal with the details of problems. He also teached me to start from the root of the problem and lay a solid foundation. His rigorous attitude towards scientific research influenced me and benefited me for the rest of my life. He also led me into the local life in Dijon and solved the practical difficulties for me. I remember that when I arrived at France for the first time, He patiently helped me dealing with various procedures. During the spreading of COVID-19, I felt frustrated while he talked to me and encouraged me to persist in research. I sincerely thank professor Fang for his caring for me!

I would like to thank professor Fengyu Wang and professor Nicolas Juillet. They gave me valuable suggestions on revision of my thesis. I would like to thank José-Luis Jaramillo. He is the president of the committee of my doctor thesis defence and also second year mid-term report. I would like to thank professor Zhao Dong. He is one of the guarantors of my joint training program, and also one of the reviewers of my mid-term defence. He provided valuable advices to my research work. Xia Liu, Yonghua Yin and other teachers from the Academy of Mathematics and Systems Science helped me a lot when I was recovering from my knee injury. My colleagues helped me in scientific research. During my knee's injury, my classmates Houqi Su, Chengjie Wang and Kaiyuan Cui found medical treatments for me and accompanied with me. Fan Jiang, Yiming Bao, Chaojie Qian and Zihan Jiang gave me valuable advices when I was confused about my future career choice.

I would like to thank the Academy of Mathematics and Systems Science, Chinese Academy of Sciences for providing me with a good living and working environment. I would like to thank the China Scholarship Council for giving me the opportunity to go to France for joint training and providing financial support. I would like to thank the institute of Mathematics of Burgundy for the hospitality during my stay in France.

I would like to thank my girlfriend Yue Xu, who always encouraged me when I was in France. After returning to China, every time I experienced difficulties in life, she was always by my side. The future is too far, the past has gone, I wish she would be with me at every particular moment.

Finally, I would like to thank my parents for their selfless dedication and deep love for me. I can not achieve anything without them.

Chapter 1

Introduction

1.1 Research background

1.1.1 Optimal transport

Optimal transport problem is firstly proposed by French mathematician Monge form practical engineering problems. In general, assume that X, Y are two Polish spaces(complete separable metric space), $T : X \to Y$ is a Borel map and $\mu \in \mathbb{P}(X)$ is a probability measure, then we say the probability measure $T_{\#}\mu \in \mathbb{P}(Y)$ is a pushforward measure of μ by T, if

$$T_{\#}\mu(E) = \mu(T^{-1}(E)), \quad \forall\, E \subset Y\, Borel.$$

The pushforward satisfies, for all Borel function $f \in L^1(T_{\#}\mu)$,

$$\int f dT_{\#}\mu = \int f \circ T d\mu.$$

$T_{\#}\mu$ is also called the image measure of μ under T, or T transports μ to $T_{\#}\mu$.
Let $c :\ X \times Y \to \mathbb{R} \cup +\{\infty\}$ and $\mu \in \mathbb{P}(X)$, $\nu \in \mathbb{P}(Y)$, then the Monge optimal transport problem is to find the optimal transport map T such that

$$\text{minimize } \ I[T] = \int_X c(x, T(x))d\mu(x) \tag{1.1.1}$$

among all the measurable map satisfying $T_{\#}\mu = \nu$. Monge optimal transport problem is ill-posed because

1. there may not exist T satisfying $T_{\#}\mu = \nu$, for example, if μ is a $Dirac$ measure while ν is not.

2. $T_{\#}\mu = \nu$ is not weakly closed in general weak topology, i.e. if $T^n_{\#}\mu = \nu$ and T^n weakly converges to T, it

is not necessary that $T_{\#}\mu = \nu$.

In 1940s, Kantorovich proposed a relaxed version of optimal transport problem in the optimal allocation of national resources. Let $\mathcal{C}(\mu, \nu) = \{\gamma \in \mathbb{P}(X \times Y) \mid \pi^X_{\#}\gamma = \mu, \pi^Y_{\#}\gamma = \nu\}$, where π^X, π^Y are projection maps form $X \times Y$ to X and Y respectively. The Kantorovich optimal transport plan problem is to find $\gamma \in \mathcal{C}(\mu, \nu)$ such that it

$$\text{minimize } \; I[\gamma] = \int_{X \times Y} c(x, y) d\gamma(x, y). \tag{1.1.2}$$

Usually, we call the minimizer as the optimal transport plan. When the cost function c is lower semi-continuous and bounded from below, there always exists a optimal transport plan. From Monge-Kantorovich optimal transport problem, Kantorovich introduced the 2-Wasserstein distance in the probability measure space: for $\mu, \nu \in \mathbb{P}_2(X)$ and $c(x, y) = d^2(x, y)$, define 2-Wasserstein distance W_2 as

$$W_2^2(\mu, \nu) := \inf_{\gamma \in \Gamma(\mu, \nu)} \int_{X \times X} d^2(x, y)\gamma(dx, dy).$$

Since we always consider 2-Wasserstein distance in this thesis, we will call W_2 as Wasserstein distance without additional requirements. We also call $(\mathbb{P}(X), W_2)$ as Wasserstein space.

1.1.2 Geometry and differential equations on the Wasserstein space

Denote $\mathbb{P}_{2,ac}(\mathbb{R}^d)$ as the set containing all of the absolutely continuous probability measures with respect to Lebesgue measure on $\mathbb{R}^d$ and finite second moments. When it is constrained in $\mathbb{P}_{2,ac}(\mathbb{R}^d)$ with the cost function $c = d^2$, Brenier [Bre91] used convex functions to describe optimal transport maps of Monge-Kantorovich optimal transport problems. This result built a bridge between the fields of optimal transport and Monge-Ampère equation, fluid dynamics, metric measure geometry, probability etc. We introduced some parts of related works .

- Partial differential equations: A class of diffusive equations can be seen as gradient flows on $\mathbb{P}(M)$. This viewpoint brought new development to contraction of diffusion semigroup, log-sobolev inequality and other related fields(see [Vil09]) .

- Infinite dimensional differential geometry: Let $\mathbb{P}_2^\infty(\mathbb{R}^d)$ be the set of probability measures which have strictly positive smooth densities. Otto defined a Riemannian metric on $\mathbb{P}_2^\infty(\mathbb{R}^d)$, which makes $(\mathbb{P}_2^\infty(\mathbb{R}^d), W_2)$ a infinite dimensional Riemannian manifold. Also, Otto got the geodesic equation and calculated the lower bound of section curvature, so that he formally showed that $\mathbb{P}_2(\mathbb{R}^d)$ has nonnegative section curvature. Based on these works, J. Lott [Lot06] derived the Riemann curvature of $\mathbb{P}_2^\infty(M)$, where the base space M is a complete simple connected Riemannian manifold without boundary.

- Metric measure geometry: Sturm, Lott, Villani etc. proved that nonnegativeness of Ricci curvature of the manifold M is equivalent to the convexity of Boltzmann entropy along Wasserstein geodesics(see [Stu06, LV09]). This means one can use the geodesic convexity of Boltzmann entropy to give the lower bound of Ricci curvature of M, even when M is not a smooth Riemannian manifold.

1.1.3 Stochastic analysis on Wasserstein spaces

In 2013, Prof. Xiangdong Li constructed a Langevin deformation connecting geodesic flows and gradient flows, and collaborated with Songzi Li to prove the W-entropy formula about the Langevin deformation([LL16]) .In 2017, Prof. Xiangdong Li proposed a research plan for constructing Brown motion and Langevin diffusion process on Wasserstein spaces in his application for the funding from the National Natural Science Foundation of China. In 2018, Prof. Xiangdong Li suggested me studying the construction of Brownian motion on Wasserstein spaces. In this subsection, starting from Brownian motion on Wasserstein spaces, we introduce some developments on related studies on the stochastic analysis and stochastic differential equations on Wasserstein spaces.
von Renesse and Sturm [vRS09] constructed an entropic measure $\mathcal{P}^\beta$ on $\mathbb{P}(\mathbb{T})$, and proved that the Wasserstein Dirichlet form

$$\mathbf{E}(u,v)=\int_{\mathbb{P}(\mathbb{T})}<\bar{D}u(\mu),\bar{D}v(\mu)>_{L^2(\mu)}\,d\mathcal{P}^\beta(\mu)$$

is closable, so that they can construct a reversible markov process with respect to $\mathcal{P}^\beta$ on $\mathbb{P}(\mathbb{T})$: $(\mu_t)_{t\in[0,T]}$. It satisfies Itô type formula and Varadhan type formula. In detail, for a smooth function u on $\mathbb{P}(\mathbb{T})$,

$$u(\mu_t)-u(\mu_0)-\frac{1}{2}\int_0^t \mathbb{L}u(\mu_s)ds$$

is a martingale, where $\mathbb{L}$ is a second order differential operator. And its quadratic variation is square of Wasserstein gradient of u. This property is similar to the Itô formula for Brownian motion in Euclidean space. $(\mu_t)_{t\in[0,T]}$ also satisfies, for any Borel subset A of $\mathbb{P}(\mathbb{T})$,

$$\lim_{\epsilon\to 0^+}\epsilon\ln\mathcal{P}(\mu_{t+\epsilon}\in A\big|\mu_t)=-\frac{1}{2}W_2^2(\mu_t,A).$$

This also shows that $(\mu_t)_{t\in[0,T]}$ is a "Brownian motion" under W_2 metric.
von Renesse and Sturm called (μ_t) as Wasserstein diffusion. Their original construction is quite abstract. To know more about its dynamic properties, we can study it by describing it by stochastic partial differential equations or particle model approximation. We introduce these two aspects of works for the Wasserstein diffusion and other related stochastic process on the Wasserstein space.

- von Renesse, Sturm etc. [AvR10, Stu14] gave particle model approximation to the Wasserstein diffusion by finite dimensional approximation. However, since the entropic measure only supports on singular measure without discrete parts, $(\mu_t)_{t\in[0,T]}$ do not have a absolutely continuous part and a discrete part almost surely. This property shows the bad analytic property of μ_t. In order to improve this point, Konaroskyi [Kon17, Kon11, KvR15, KvR17] constructed a new particle model to approximate a class of diffusion process on the Wasserstein space, which still shares the main feature of Wasserstein diffusion but has better analytic properties. However, the process satisfying Konaroskyi's model is not necessarily unique. Marx [Mar18] rectified the original model and constructed a unique diffusion process which satisfies that model.

- Konarovskyi and Von Renesse [KvR17, KvR15, vRLK19] proved that all the diffusion process on the Wasserstein space which shares the features of Brownian motion are all satisfied by a regularized form of Dean-Kawasaki type stochastic partial differential equation:

$$\partial_t \mu = \alpha \Delta \mu + \Xi(\mu) + \mathrm{div}(\sqrt{\mu}\dot{W}),$$

where Ξ is some nonlinear operator, $\dot{W}$ is a white noise both in space and time. In particular, [vRLK19] proved that if one wants to get a non-trivial solution to the Dean-Kawasaki equation, the regularization term Ξ is necessary. Dean-Kawasaki is a stochastic Fokker-Planck equation, the related problems about McKean-Vlasov equations also attract much attentions. [Wan21, BLPR17] studied a class of mean-field stochastic differential equations and the corresponding partial differential equations on the measure space. Stochastic differential equations on the Wasserstein space are also related to mean-field game theory. In short, mean-field game theory investigates the Nash equilibrium of the mean-field limit of interactive particle systems, whose interaction is determined by the distribution of the particles. To study such problems, Larry and Lions [LL06a, LL06b, LL07] developed differential calculus on the Wasserstein space.

1.2 Main contents

Inspired by the works mentioned above, This paper mainly studies some topics on the geometry and stochastic analysis on the Wasserstein space.

In Chapter 2, we mainly introduce some preliminaries. Firstly, we review the basic topological facts about the Wasserstein space. Secondly, we introduce Brenier's and McCann's works on the optimal transport map. Then, starting from Benamou-Brenier formula, we describe geodesics on the Wasserstein space from viewpoint of displacement interpolation and Riemannian geometry. As a remark, we explain the relation between geodesic equations and zero-pressure Euler equation. Finally, we introduce a gradient flow equation on $\mathbb{P}(M)$ and implicit Euler approximation.

In Chapter 3, we mainly introduce Newton flows on Wasserstein spaces. We firstly give a brief review on the Newton flow equation on $\mathbb{R}^d$, and use implicit Euler approximation method to prove the existence of solutions. Using a similar method, we prove the existence of solutions to the Newton flow equation on $\mathbb{P}(\mathbb{T}^d)$ under certain conditions (Theorem 3.2.6) and give the conditions for uniqueness. In particular, when the base space is $\mathbb{R}$, we give conditions for the uniqueness of the limiting point of Newton flows, i.e. there exists a unique minimizer of

the potential functional.. It is known that gradient flows on Wasserstein spaces are equivalent to Fokker-Planck equations. As a comparison, we introduce the corresponding partial differential equations of Newton flows of some calsses of calssical functionals in section 3.3. We also reveal the connection between the Newton flow equation on $\mathbb{P}(\mathbb{T})$ and the Keller-Segel equation.
The main contributions of this chapter:

- Under certain conditions, we prove the existence and uniqueness of the solution to the Newton flow equation on $\mathbb{P}(\mathbb{T}^d)$. The conditions applies to the common functional $F(\mu) = \int V d\mu + \int W * \mu d\mu$.
- When the base space is $\mathbb{R}$, we give conditions for the uniqueness of the limiting point of Newton flows, i.e. there exists a unique minimizer of the potential functional.
- on $\mathbb{P}(\mathbb{T})$, we reveal the connection between the Newton flow equation and the Keller-Segel equation.

In Chapter 4, we mainly introduce the Riemannian geometry and parallel translation on $\mathbb{P}_2(M)$. We revisit the intrinsic differential geometry of the Wasserstein space $(\mathbb{P}_2(M), W_2)$. In detail, we fristly introduce the tangent space of $\mathbb{P}_2(M)$ from Ambrosio's theorem on the representation of absolutely continuous curves on $\mathbb{P}_2(M)$. Next, we prove the existence (Theorem 4.2.4) and uniqueness (Theorem 4.6.3) of solutions to ordinary differential equations on $\mathbb{P}_2(M)$. In section 4.3, we rewrite Lie bracket, Levi-Civita connection, proposed by J. Lott in [Lot06], in an intrinsic geometric way. We also extend the domain of Levi-Civita to more general vector fields in tangent spaces of the measure included in $\mathbb{P}_{\text{div}}(M)$ (Theorem 4.3.6). In section 4.4, we prove that when $\sigma \in \mathbb{P}_{2,ac}(M)$, the square of Wasserstein distance $W_2^2(\sigma, \mu)$ is derivable along any constant vector field at any μ . At last, in section 4.5, based on the pointwise derivability of W_2^2, we obtain the extension of vector fields along good curves on $\mathbb{P}_2(M)$ (Theorem 4.5.1), and introduce the calssical results on parallel translation. We also prove the existence and uniqueness of the smooth solution to the parallel translation equation on $\mathbb{P}_2(\mathbb{T})$ (Theorem 4.5.7).
The main contributions of this chapter:

- We extend the domain of Levi-Civita connection on $\mathbb{P}_2^\infty(M)$, so that one can introduce Levi-Civita connection for more general vector fields on $\mathbb{P}_{\text{div}}(M)$.
- We extend vector fields on $\mathbb{P}_2(M)$, so that one can introduce parallel translations as in differential geometry.
- We prove the existence and uniqueness of the smooth solution to the parallel translation equation on $\mathbb{P}_2(\mathbb{T})$, and improve the regularity results on the solution proposed by Ambrosio.

In Chapter 5, we mainly introduce stochastic parallel translations and $Q-$Wiener process on the Wasserstein space. First of all, we do Itô stochastic calculus for three kinds of functional on the Wasserstein space: potential functional, interaction functional and Entropy functional, along the image measure process induced by some stochastic differential equation. We also prove the existence and uniqueness of the solution to the stochastic gradient flow equation when the noise is finite dimensional (Theorem 5.2.8). Next, we construct stochastic parallel

translation, along the image measure process induced by some stochastic differential equation with enough regularity, as a L^2 limit of Euler approximation (Proposition 5.3.3). To get more information about the dynamics of stochastic parallel translation, we prove that stochastic parallel translation is a weak solution, both in sense of probability and analysis, of a stratanovich form of stochastic partial differential equation (Theorem 5.3.4). Then, in the spirit of Wong-Zakai approximation, we find the strong form of stochastic partial differential equation satisfied by stochastic parallel translation (Theorem 5.3.5) and prove the conservation of norm (Theorem 5.3.8). In section 5.4, we pick a base on M so that we can construct a $Q-$Wiener process on $\mathbb{P}_2(M)$ (Theorem 5.4.5). Finally, as an example, we prove the well-posedness of stochastic parallel translation on $\mathbb{P}_2(\mathbb{T})$ (Theorem 5.5.1).
The main contributions of this chapter:

- We prove the existence of stochastic parallel translation along the image measure process induced by a stochastic differential equation. And we construct the weak and strong form of stochastic partial differential equations satisfied by stochastic parallel translation. Also, we can prove the regular solution to the strong form equation preserves norm.
- We construct a $Q-$Wiener process on the Wasserstein space.
- We prove well-posedness of strong form of stochastic partial differential equations satisfied by stochastic parallel translation on $\mathbb{P}_2(\mathbb{T})$.

In Chapter 6, we mainly study the diffusive Dean-Kawasaki equation on one dimensional Torus with colored noise. Using the idea of $Q-$Wiener process and interaction particle system, we give a new particle approximation model to the regularized martingale problem $(RMP)^{\alpha,\beta}_{\mathbb{1}_{\mathbb{T}}dx}$ of the diffusive Dean-Kawasaki equation on one dimensional Torus driven by a white noise, whose spatial correlated intensity is larger than 1 (Theorem 6.3.1). Under such conditions, we prove the existence of solutions to the regularized martingale problem $(RMP)^{\alpha,\beta}_{\mathbb{1}_{\mathbb{T}}dx}$ (Theorem 6.4.1). We also prove that the solution $\{\mu_t,\, t \in [0,T]\}$ is non-atomic for all $t \in [0,T]$ almost surely (Lemma 6.4.2).
The main contributions of this chapter:

- We proposed a new particle approximation model to solutions to the regularized martingale problem of the diffusive Dean-Kawasaki equation on one dimensional Torus.
- We prove the existence of nontrivial solutions to the regularized martingale problem of the diffusive Dean-Kawasaki equation on one dimensional Torus under a weaker condition on noise than other classical conditions.

Chapter 2

Preliminaries

In this chapter, we will introduce some preliminaries about optimal transport theory. We will firstly introduce the basic topological facts about the Wasserstein space, then we will introduce Brenier's optimal transport map theorem and Benamou-Brenier formula. As a remark, We will explain the connection between fluid mechanics and optimal transport theory. Benamou-Brenier formula can be seen as a representation of the geodesic on the Wasserstein space. To illustrate this point of view, we introduce displacement interpolation and infinite dimensional Riemannian metric. At last, we will apply implicit Euler approximation method to approximate a gradient flow equation on the Wasserstein space.

2.1 Optimal transport and geodesics on the Wasserstein space

Theorem 2.1.1. *X is a metric space, then*

- *$(\mathbb{P}_2(X), W_2)$ is a metric space;*
- *convergence in W_2 is equivalent to weak convergence plus convergence of second moments;*
- *if X is a Polish space, then $(\mathbb{P}_2(X), W_2)$ is also a Polish space.*

Proof. see [Vil03] . □

This theorem shows the topology properties of $(\mathbb{P}_2(X), W_2)$. When the base space is a connected compact manifold, W_2 metrizes weak convergence. In this paper, we always consider the optimal transport problem when the cost function is the square of distance.
Now, we come back to Monge-Kantorovich transportation problem and denote $\mathcal{C}_o(\mu, \nu)$ as the set containing all of the optimal transport plans $\gamma \in \mathbb{P}(X \times Y)$. It is natural to ask when there is a unique minimizer and when the minimizer of Kantorovich transportation problem can be a minimizer of Monge transportation problem? The following theorem gives the answer:

Theorem 2.1.2. *(Brenier) Let $\mu, \nu \in \mathbb{P}_2(\mathbb{R}^d)$, then,*

1. *If μ is absolutely continuous, Then there exists a unique optimal transport plan*

$$\gamma = (Id \times \nabla\varphi)_{\#}\mu,$$

where $\nabla\varphi$ is the unique(uniquely determined $d\mu-$almost everywhere) gradient of a convex function φ which satisfies $\nabla\varphi_{\#}\mu = \nu$.

2. *Under the assumption of 1, $\nabla\varphi$ is the unique ($d\mu-$a.s.) solution to the Monge transportation problem:*

$$\int_{\mathbb{R}^d} |x - \nabla\varphi(x)|^2 \, d\mu(x) = \inf_{\{T: T_{\#}\mu = \nu\}} \int_{\mathbb{R}^d} |x - T(x)|^2 \, d\mu(x).$$

.

3. *If ν is also absolutely continuous, then, for $d\mu-$ almost all x and $d\nu-$ almost all y ,*

$$\nabla\varphi^* \circ \nabla\varphi = x, \quad \nabla\varphi \circ \nabla\varphi^*(y) = y,$$

where $\nabla\varphi^*$ is the ($d\nu-$ almost everywhere) unique gradient of a convex function which push ν forward to μ .

Brenier considered the optimal transport problem when $c = d^2$ on $\mathbb{P}_2(\mathbb{R}^d)$, and gave a sufficient condition for the uniqueness of the optimal transport plan: the initial probability measure is absolutely continuous with respect to Lebesgue measure. In this case, the optimal transport plan of Kantorovich transportation problem is also the optimal transport map of Monge transportation problem, which can be represented by a gradient of some convex function. McCann gave the optimal transport map theorem when the base space is a complete connected Riemannian manifold, so that one can see more clearly the geometric feature of optimal transport maps. Here, we briefly introduce a part of his results:

Theorem 2.1.3. *(McCann) Let M be a complete connected smooth Riemannian manifold, dx is a standard Riemannian measure. The cost function $c(x, y) = d^2(x, y)$, where d is the Riemannian diatance. Given $\mu, \nu \in \mathbb{P}_2(M)$, and suppose that μ is absolutely continuous with respect to dx, then there is a unique optimal transport plan γ from μ to ν such that*

$$\gamma = (Id \times T)_{\#}\mu,$$

where T is uniquely ($d\mu-$almost surely) determined. And there is a $\frac{d^2}{2}-$concave function φ such that

$$T(x) = \exp_x(-\nabla\varphi).$$

Proof. See [McC01]. □

These results describe the static optimal transport problem, while the theorem below deals with the optimal transport problem from the viewpoint of dynamics, which can be seen as a representation of geodesics on the Wasserstein space.

Theorem 2.1.4. *(Benamou-Brenier formula) For $(\mu, v) := (\mu_t, v_t)_{t\in[0,1]}$, define the energy functional $A[\mu, v] := \int_0^1 |v_t|^2 \mu_t dt$, then*

$$\inf_{(\mu,v)\in V(\mu_0,\mu_1)} A[\mu, v] = W_2^2(\mu_0, \mu_1), \tag{2.1.1}$$

where $V(\mu_0, \mu_1)$ is a set contains all the pairs $(\mu, v) := (\mu_t, v_t)_{t\in[0,1]}$ which satisfies the following conditions:

1. $\mu \in C([0, 1], \mathbb{P}_{2,ac}(\mathbb{R}^d))$ *, where $\mathbb{P}_{2,ac}(\mathbb{R}^d)$ is equipped with weak* topology.*
2. $v \in L^2(d\mu_t dt)$.
3. $\bigcup_{t\in[0,1]} supp(\mu_t)$ *is bounded.*
4. *The following mass transportation equation*

   $$\partial_t \mu_t + \mathsf{div}(\mu_t v_t) = 0$$

   holds in sense of distribution.
5. $\mu(t = 0, \cdot) = \mu_0(\cdot), \ \ \mu(t = 1, \cdot) = \mu_1(\cdot).$

Proof. See [BB00]. □

Remark 2.1.5. *The theorem above showed a connection between fluid dynamics and optimal transport. We think of μ_0 and μ_1 as the density of particles in a given region in $\mathbb{R}^d$ at time $t = 0$ and $t = 1$. If we assume that for any $t \in [0, 1]$, there exists a vector field v_t , which is smooth in time t and uniformly Lipschitz in space, describing how particles move around, i.e. we can describe the time evolution of the particles position by*

$$\frac{dX_t}{dt} = v_t(X_t). \tag{2.1.2}$$

According to the ordinary differential equation theory, given $x_0 \in \mathbb{R}^d$, (2.1.2)*has a unique solution $X_{x_0}(t)$ for $t \in [0, 1]$. Also, the map $(t, x_0) \mapsto X_{x_0}(t)$ is a one-to-one Lipschitz map. Then $(T_t)_{0\le t\le 1} = (x \mapsto X_x(t))$ is a diffeomorphic flow on $\mathbb{R}^d$. By the method of characteristics, $\mu_t = (T_t)_{\#}\mu_0$ is the unique weak solution of the following mass transportation equation:*

$$\partial_t \mu_t + \mathsf{div}(\mu_t v_t) = 0. \tag{2.1.3}$$

The fluid's kinetic energy at time t is $E(t) = \int_{\mathbb{R}^d} \mu_t |v_t|^2 dx$. The total energy for all the particles moving with speed v_t from $t = 0$ to $t = 1$ is $A[\mu, v] = \int_0^1 E(t)dt$. In fact, (2.1.3) *is the Eulerian representation of the fluid dynamic,* (2.1.2) *is the Lagranian representation. These two representations are equivalent when T_t is diffeomorphic.*
Each pair (μ, v) in $V(\mu_0, \mu_1)$ represents a continuous curve from μ_0 to μ_1 in $\mathbb{P}_2(\mathbb{R}^d)$. It also represent a dynamic process of a fluid field transporting μ_0 to μ_1 according to the velocity field v_t. The formula (2.1.1) *reveals that the geodesic in probability measure space under W_2 distance corresponds to the fluid dynamic process with the lowest total kinetic energy, in which case the W_2 distance is the lowest kinetic energy.*

The formula (2.1.1) can be seen as geodesic equation on the Wasserstein space from two points of view. Firstly, it is a random version of action minimizing curves. In detail, this viewpoint starts from the time dependent optimal transport problem and uses displacement interpolation to describe geodesics on the Wasserstein space. Secondly, from the viewpoint of Riemannian geometry, if we equip $\mathbb{P}$ with suitable topology, tangent bundles and Riemannian metric, (2.1.1) can be realized as a energy variation formula for C^1-curves. We will introduce these two viewpoints.
We firstly introduce the time dependent Monge optimal transport problem on $\mathbb{P}_{ac}$:

$$\inf_T \left\{ \int_X \int_0^1 \left| \frac{dT_t(x)}{dt} \right|^2 dt\mu(dx) \Big| T_0 = Id,\ (T_1)_\# \mu = \nu \right\}. \tag{2.1.4}$$

Due to convexity of $\frac{d^2(x,\cdot)}{2}$, It is easy to see that for any $x \in \mathbb{R}^d$, the trajectory $\{T_t(x),\ t \in [0,1]\}$ with lowest cost is always a straight line([Vil03]) . Combined with the optimal transport map $T(x) = \nabla\varphi(x)$ given by theorem 2.1.2, we get the expression of the trajectories $T_t(x)$ with lowest cost:

Theorem 2.1.6. *(McCann [McC97]) Let $\mu, \nu \in \mathbb{P}_{ac}(\mathbb{R}^d)$, $\nabla\varphi$ is the unique($d\mu-$a.s.) gradient of convex function φ satisfying $(\nabla\varphi)_\# \mu = \nu$. Then the solution to the time dependent Monge optimal transport problem 2.1.4 is*

$$T_t(x) = t\nabla\varphi(x) + (1-t)x, \quad 0 \le t \le 1. \tag{2.1.5}$$

Proof. See [McC97] . □

$\mu_t = (T_t)_\# \mu$ is called the displacement interpolation from μ to ν . It shows the dynamic process of optimal transport. In general, we can still define displacement interpolation for $\mu \in \mathbb{P}(\mathbb{R}^d)$.

Definition 2.1.7. *Let $\mu_0, \mu_1 \in \mathbb{P}_2(\mathbb{R}^d)$, $\gamma \in \mathcal{C}(\mu_0, \mu_1)$ is a transport plan. We say that a curve $[\gamma](t) : [0,1] \to \mathbb{P}_2(\mathbb{R}^d)$ on $\mathbb{P}_2(\mathbb{R}^d)$ is a displacement interpolation from μ to ν induced by γ, if*

$$[\gamma](t) := \left((1-t)\pi^1 + t\pi^2\right)_{\#}\gamma.$$

where π^1, π^2 are projection maps to the first variable and the second variable respectively.

It can be proved that the displacement interpolation between μ and ν is equivalent to the geodesic between μ and ν (see [Gig11]).

Going back to the case for $\mathbb{P}_{ac}(\mathbb{R}^d)$, we can derive the geodesic equation by displacement interpolation. Note that although $\mathbb{P}_{ac}(\mathbb{R}^d)$ is not general, it has a obvious geometry feature and a clear correspondence with geometry structure and differential calculus on the Euclidean space or finite dimensional manifold. In detail, suppose that $t = 0$ the initial velocity field $v_0 = \nabla\varphi - Id$ at time 0, then $v_t = (\nabla\varphi - Id) \circ T_t^{-1}$ due to the displacement interpolation. Combined with (2.1.2) , the Lagranian representation of the geodesic from μ to ν is

$$\begin{cases} \dfrac{d}{dt}T_t = v_t(T_t) \\ \dfrac{d^2}{dt^2}T_t = 0. \end{cases} \tag{2.1.6}$$

Using (**??**), we have

$$0 = \frac{d^2}{dt^2}T_t = \frac{\partial v}{\partial t}(T_t) + v(t, T_t) \cdot \nabla v(t, T_t).$$

Then, since (2.1.3), we give the Eulerian representation of the geodesic:

$$\begin{cases} \partial_t \mu_t + \operatorname{div}(\mu_t v_t) = 0 \\ \dfrac{\partial v}{\partial t} + v \cdot \nabla v = 0, \end{cases} \tag{2.1.7}$$

The initial condition is totally determined by μ and ν:

$$\mu_0 = \mu; \quad v(0, x) = \nabla\varphi(x) - x. \tag{2.1.8}$$

Remark 2.1.8. *There is a long history for researches on displacement interpolation, which is firstly proposed by McCann(see [Vil09]). Here, we only consider the simplest case for the kinetic energy $E(t)$, while for a general Lagranian action , displacement interpolation can also be introduced. We refer to [Vil09] for more details.*

Remark 2.1.9. (2.1.7)*, in which the first equation is mass conservation and the second one is movement conservation, is a compressible Euler equation for zero pressure. Generally, the well-posedness of* (2.1.7) *is not obvious. Even when the initial value is smooth enough, the solution may explode in finite time because of the intersection of characteristics, or in other word, mass concentration. However, in the discussion above, since φ is convex, characteristics will never intersect with each other during $t \in [0, 1)$.*

In dimension one, (2.1.7) *is also an inviscid Burger's equation .*

Remark 2.1.10. *All the theorems above are valid when the base space is a complete connected compact Riemannian manifold with certain conditions on curvature.*

2.2 Riemannian structure on the Wasserstein space

Next, we introduce another point of view: Riemannian geometry. This viewpoint is also one of the starting points of our works. In the early 21st century, Otto firstly proposed a Riemannian metric on $\mathbb{P}_2^\infty(\mathbb{R}^d)$. In this section, we introduce the tangent space and Riemannian metric on $\mathbb{P}^\infty(M)$, where M is a compact Riemannian manifold.

Definition 2.2.1. *Given $\mu \in \mathbb{P}^\infty(M)$ with $d\mu = \rho dx$, define the tangent space $\mathbf{T}_\mu$ at μ as*

$$\mathbf{T}_\mu \mathbb{P}^\infty(M) := \{\nabla\psi, \;\; \psi \in C^\infty(M)\}$$

For any $\nabla\psi_1, \nabla\psi_2 \in \mathbf{T}_\mu \mathbb{P}^\infty(M)$, the Riemannian metric is defined as

$$\langle \nabla\psi_1, \nabla\psi_2 \rangle_\mu = \int_M \langle \nabla\psi_1, \nabla\psi_2 \rangle \rho dx$$

Theorem 2.2.2 (Geodesics)**.** *If $c : [0,1] \to \mathbb{P}^\infty(M)$ is a smooth immersed curve, and suppose that $c(t) = \rho(t)dx$. ρ satisfies*

$$\partial_t \rho = -\nabla.(\rho\nabla\phi),$$

where $\nabla\phi(t) \neq 0$ and $\int_M \phi\rho dx = 0$. Then, the length of c, denoted as $L(c)$, under Wasserstein distance satisfies:

$$L(c) = \int_0^1 \Big(\int_M |\nabla\phi(t)|^2 \rho(t) dx \Big)^{\frac{1}{2}} dt.$$

Remark 2.2.3. *$\mathbb{P}^\infty(M)$ can become a infinite dimensional smooth Riemannian manifold if equipped with a topology induced by smooth curves(see [Ott01], [KM97]). The definition of tangent space and Riemannian metric can be naturally extended to $\mathbb{P}_2(M)$, which we will introduce in Chapter 4. However, $\mathbb{P}_2(M)$ can not be a differentiable Riemannian manifold. This can be seen by a simple observation: At discrete probability measure, the exponential map can not give a one-to-one local map from its tangent space to its neighbourhood.*

There is an open problem: Can one find a subspace of $\mathbb{P}(M)$, larger than $\mathbb{P}^\infty(M)$, so that it can become a infinite dimensional Riemannian manifold? Or can the formal Riemannian structure and Riemannian calculus be extended to a larger space? In Chapter 4, we will try to find the answer to the second question from the point of analysis.

2.3 Gradient flow equation on the Wasserstein space

A huge class of partial differential equations can be seen as gradient flows on the Wasserstein space. This is firstly proposed by Otto in [Ott01]. In this section, we briefly introduce the gradient flow equation of the following functional

$$E(\rho) = \begin{cases} \int \rho \log \rho dx + \int V \rho dx, & \rho \in \mathbb{P}_{ac}(\mathbb{R}^d) \\ +\infty, & \text{otherwise.} \end{cases}$$

Its gradient under Wasserstein metric is $gradF(\rho) = \nabla \log \rho + \nabla V$, which we will explained later in Chapter 3 . Suppose that V is smooth and $\lambda-$convex for $\lambda > 0$. We will use implicit Euler approximation method to derive gradient flow equation.

At first, given time step $\tau > 0$ and initial measure $\rho_\tau^0 = \rho^0$. We construct discrete solution $\{\rho_\tau^n\}$. Given ρ_τ^n, define $\rho_\tau^{n+1} = \text{argmin}\, E(\rho) + \frac{W_2^2(\rho_\tau^n,\rho)}{2\tau}$. Since E is strictly convex([Vil03]), ρ_τ^{n+1} is unique. Because

$$\int \rho_\tau^{n+1} \log \rho_\tau^{n+1} dx + \int V \rho_\tau^{n+1} dx + \frac{W_2^2(\rho_\tau^n, \rho_\tau^{n+1})}{2\tau} \leq \int \rho_\tau^n \log \rho_\tau^n dx + \int V \rho_\tau^n dx, \tag{2.3.1}$$

This means

$$\sup_{n \leq 0} E(\rho_\tau^n) \leq E(\rho^0).$$

Thus, we get the uniform boundedness of $E(\rho_\tau)$, so that ρ_τ is weakly compact in L^1. At the same time, by summing together the inequalities 3.2.4, we have the following energy estimate:

$$\sum_{n \geq 0} W_2^2(\rho_\tau^n, \rho_\tau^{n+1}) \leq 2\tau(E(\rho^0) - \inf E),$$

Also, from this last estimate, we can get equi-continuity by Cauchy-Schwarz inequality. Then by Ascoli's theorem, there exists a subsequence $\{\rho_{\tau_k}\}_{k\geq 0}$ uniformly converging to some ρ under $C([0,T], \mathbb{P}_{ac}(\mathbb{R}^d) - \omega - L^1)$.

Next, we want to prove ρ satisfies

$$\partial_t \rho = \Delta \rho + \nabla.(\rho \nabla V) \tag{2.3.2}$$

in distribution. Let $\xi \in C_c^\infty(\mathbb{R}_+^d)$, we operate a small perturbation around ρ_τ^{n+1}:

$$\rho_\epsilon = (Id + \epsilon\xi)_\# \rho_\tau^{n+1}.$$

When ϵ is small enough, $Id + \epsilon\xi$ is a C^1 difeomorphism. We have

$$E(\rho_\epsilon) = \int \rho_\tau^{n+1} \log \frac{\rho_\tau^{n+1}}{\det(I_d + \epsilon \nabla \xi)} dx + \int \rho_\tau^{n+1}(x) V(x + \epsilon\xi(x)) dx.$$

Thus, On the other hand, since ρ_τ^n, ρ_τ^{n+1} are absolutely continuous, there exists an optimal map $\nabla\varphi$ such that

$\nabla\varphi_{\#}\rho_\tau^n = \rho_\tau^{n+1}$. Then

$$\rho_\epsilon = [(Id + \epsilon\xi)\circ\nabla\varphi]_{\#}\rho_\tau^n,$$

so

$$W_2(\rho_\tau^n, \rho_\epsilon) \leq \int \rho_\tau^n(x)|x - \nabla\varphi(x) - \epsilon\xi\circ\nabla\varphi(x)|^2 dx.$$

Therefore, we obtain

$$\begin{aligned}
&E(\rho_\epsilon) - E(\rho_\tau^{n+1}) + \frac{W_2^2(\rho_\tau^n, \rho_\epsilon)}{2\tau} - \frac{W_2^2(\rho_\tau^n, \rho_\tau^{n+1})}{2\tau} \\
&\leq \int \rho_\tau^n(x)\frac{1}{2\tau}\left(|x - \nabla\phi(x) - \epsilon\xi\circ\nabla\phi(x)|^2 - |x - \nabla\phi(x)|^2\right)dx \\
&+ \int \rho_\tau^{n+1}[V(x+\epsilon\xi) - V(x)]dx - \int \rho_\tau^{n+1}(x)\log\det(Id + \epsilon\nabla\xi(x))dx.
\end{aligned}$$

Since ρ_τ^{n+1} is the minimizer of $E(\rho) + \frac{W_2^2(\rho_\tau^n,\rho)}{2\tau}$, the left hand side of the above inequality must be larger than 0. Let $\epsilon \to 0^+$, we get the Euler-Lagrange equation:

$$\frac{1}{\tau}\int \rho_\tau^n(x)\langle\nabla\phi(x) - x, \xi(\nabla\phi(x))\rangle dx = \int \rho_\tau^{n+1}[\langle -\nabla\log(\rho_\tau^{n+1}) - \nabla V, \xi\rangle]dx. \tag{2.3.3}$$

According to the energy estimate, we can prove, without every details which can be seen in [Vil03],

$$\int \rho(t)\xi - \int \rho(s)\xi = \int_s^t \int \rho(r)(\Delta\xi - \nabla V\cdot\nabla\xi)dr.$$

(2.3.2) is proved .

Although we have not strictly prove (2.3.2) is exactly the gradient flow of E on the Wasserstein space, it still use the viewpoint of gradient flow to approximate the solution to the diffusive equation (2.3.2). In Chapter 3 , we will use a similar method to approximate the Newton flow equation.

Chapter 3

Newton Flow on the Wasserstein Space

Recently, gradient flows on the Wasserstein space attract much attention and get fruitful results. In 1998, using implicit Euler approximations to gradient flows on the Wasserstein space, [JKO98] gave a time-discreted interation method for a class of Fokker-Planck equations. [Ott01] introduced Riemannian geometry on the Wasserstein space and proved that porous medium equations are gradient flows of Renyi's entropy on the Wasserstein space. Then, applying the ideas of gradient flows, Otto proved the contraction of diffusion semigroups under W_2 distance. Otto and Villani [Ott01] proved Talagrand inequalities and HWI inequalities for Fokker-Planck equations.

On the other hand, in Calculus, Newton method is an important algorithm to find solutions of $f(x) = 0$ for differentiable functions. It also plays an important role in proving implicit function theorem. Former soviet mathematician Kantorovich introduced generalized Newton method on Banach space, which can be used to solve a huge class of integral and differential equations. In May 2011, Fields Medal Winner Villani mentioned Newton method's application in nonequilibrium statistical mechanics in a public report.

In 2019, inspired by Villani's report, professor Xiangdong Li suggested me studying Newton flow on the Wasserstein space and related topics. In detail, we consider

- How to reasonably define Newton flow equations on Wasserstein spaces;
- the connections between Newton flows and differential equations;
- existence of solutions to Newton flow equations;
- uniqueness of solutions to Newton flow equations;
- convergence of Newton methods;

- applications of Newton flows and Newton methods.

In July 2019, under guidance of professor Xiangdong Li, we derived Newton flow equations on the Wasserstein space and got the conditions for uniqueness of solutions to Newton flow equations. In August 2019, professor Xiangdong Li mentioned our works on the joint meeting of Chinese Academy of Mathematics and System Sciences and Huawei company. After that, we further improved the results under the guidance of professor Xiangdong Li and studied the existence of Newton flows. Next, we briefly introduce the main contents of this chapter. Firstly, in section 3.1, we give a short review on Newton flow equations on $\mathbb{R}^d$, and use the implicit Euler approximation to prove the existence of solutions. In section 3.2, we prove the well-posedness of Newton flow equations(theorem 3.2.6, theorem 3.2.12). Especially, when the base space is $\mathbb{R}$, we give the sufficient conditions for the uniqueness of limiting points of Newton flows of potential functionals (theorem 3.2.13), i.e. uniqueness of minimizer of potential functional on the Wasserstein space. It is known that gradient flows are equivalent to Fokker-Planck equations. As a comparison, we give the partial differential equations corresponding to Newton flows of several classes of classical functionals on the Wasserstein space. In the last section, we reveal the connection between relaxed Newton flow equations and Keller-Segel equations on $\mathbb{P}(\mathbb{T}^1)$.
In general, consider the operator P on a Banach space and suppose that x^* is a zero point of P, i.e.

$$P(x^*) = 0.$$

Starting from a given point x_0, assuming that $[P'(x_0)]^{-1}$ exists, define

$$x_1 = x_0 - [P'(x_0)]^{-1}(P(x_0)),$$

If we define in this way recursively , we can construct $\{x_n\}$ satisfying

$$x_{n+1} = x_n - [P'(x_n)]^{-1}(P(x_n)). \tag{3.0.1}$$

$\{x_n\}$ is a approximation solution to $P(x) = 0$. The sequence generating method introduced above is called Newton's method (see [KA82]). Its continued equation is called Newton flow equation. The convergence problem of Newton methods has been studied in [KA82] and other related works. It is interesting that the Newton's method usually has a faster convergence speed than another algorithm: gradient descent method.
We firstly introduce Newton flow equation on Euclidean space and the corresponding implicit Euler approximation.

3.1 Review of Newton flow equations on $\mathbb{R}^d$

We first study the easiest case to see how to use implicit Euler method to approximate the Newton flow equation. At the same time, we compare with the process to approximate the gradient flow equation.

We assume that $F : \mathbb{R}^d \to \mathbb{R}$ is second order differentiable and the operator $\nabla^2 F$ is bounded uniformly by $0 < \lambda_1 < \nabla^2 F(x) < \lambda_2$. Given time step $\tau > 0$ and x_τ^n , let

$$F_{x_\tau^n,\tau} := F(x) + \frac{1}{2\tau}\langle \nabla^2 F(x_\tau^n) \cdot (x - x_\tau^n), x - x_\tau^n \rangle,$$

We assume it has a unique minimizer and let x_τ^{n+1} be the unique solution of the minimization problem:

$$\min F_{x_\tau^n,\tau} \tag{3.1.1}$$

Then the corresponding Euler-Lagrange equation is

$$\nabla F(x_\tau^{n+1}) + \frac{1}{\tau}\nabla^2 F(x_\tau^n) \cdot (x_\tau^{n+1} - x_\tau^n) = 0$$

Let the partition of $[0, T]$ be $\{0, \tau, 2\tau, ...n\tau, ...\}$, we construct $x_\tau(t)$ by connecting x_τ^n and x_τ^{n+1} by straight line. Also, define $V_\tau(t) = \frac{x_\tau^{n+1} - x_\tau^n}{\tau}$, when $t \in [n\tau, (n + 1)\tau)$. Our goal is to prove that there exists one solution to the Newton flow equation

$$\nabla^2 F(x(t)) \cdot \dot{x}(t) = -\nabla F(x(t)) \tag{3.1.2}$$

Step 1 we want to prove $x_\tau(t)$ converges to $x(t)$ under uniform norm, as $\tau \to \infty$.
First, by (3.1.1), we see that

$$\begin{aligned} F(x_\tau^n) - F(x_\tau^{n+1}) &\geq \frac{1}{2\tau}\langle \nabla^2 F(x_\tau^n) \cdot (x_\tau^{n+1} - x_\tau^n), x_\tau^{n+1} - x_\tau^n \rangle \\ &> \frac{\lambda_1}{2\tau}||x_\tau^{n+1} - x_\tau^n||^2 \end{aligned}$$

then by Cauchy inequality, we can easily get the uniform boundedness:

$$||x_\tau^n - x(0)||^2 < \frac{C}{\lambda_1}\tau|F(x(0) - \inf F(x)|$$

and equicontinuity

$$||x_\tau^n - x_\tau^k||^2 < \frac{C'}{\lambda_1}(n - k)\tau.$$

Then by Arzelà-Ascoli theorem, there exists a subsequence $x_\tau(t)$ uniformly converges to $x(t)$.

Step 2: we will prove $V_\tau(t)$ has a subsequence weakly converging to some $V(t)$ in $L^2(dt)$. Since

$$\frac{||x_\tau^{n+1} - x_\tau^n||^2}{2\tau} \leq \frac{1}{\lambda_1}(F(x_\tau^n) - F(x_\tau^{n+1})),$$

we have

$$\int_0^T V_\tau^2(t)dt \le \frac{C}{\lambda_1}|F(x(0) - \inf F(x)| < +\infty. \tag{3.1.3}$$

By this property, we know that $V_\tau(t)$ is compact with respect to the weak topology of $L^2(dt)$ because of Kakutani's theorem. So we can choose a subsequence , which will be denoted as $V_\tau(t)$ for convenience. And the weak limit point is $V(t)$.

Step 3: we come to prove $\dot{x}(t) = V(t)$ in weak sense, i.e. $\forall f \in \mathcal{C}_c^\infty(\mathbb{T}^d)$, $f(x(T)) - f(x(0)) = \int_0^T \langle \nabla f(x(t)), V(t)\rangle dt$. In fact, by the convergence of $x_\tau(t)$ to $x(t)$ under the uniform norm,we have $\lim_{\tau\to 0} f(x_\tau(T)) - f(x(0)) = f(x(T)) - f(x(0))$. Also,

$$\begin{aligned}
f(x_\tau(T)) - f(x(0)) &= \Sigma_{i=0}^{[\frac{T}{\tau}]} f(x_\tau((i+1)\tau)) - f(x(i\tau)) \\
&= \Sigma_{i=0}^{[\frac{T}{\tau}]} \int_0^1 \langle \nabla f(x_\tau(i\tau + \lambda\tau)), x_\tau^{i+1} - x_\tau^i \rangle d\lambda \\
&= \Sigma_{i=0}^{[\frac{T}{\tau}]} \int_{i\tau}^{(i+1)\tau} \langle \nabla f(x_\tau(t)), V_\tau(t)\rangle dt
\end{aligned}$$

So to prove $\lim_{\tau\to 0} f(x_\tau(T)) - f(x(0)) = \int_0^T \langle \nabla f(x(t)), V(t)\rangle dt$, we only need to prove, as τ goes to 0,

$$\begin{aligned}
&|\int_0^T \langle \nabla f(x(t)), V(t)\rangle dt - \int_0^T \langle \nabla f(x_\tau(t)), V_\tau(t)\rangle dt| \\
&\le \int_0^T |\langle \nabla f(x(t)), V(t) - V_\tau(t)\rangle| dt + \int_0^T |\langle \nabla f(x(t)) - \nabla f(x_\tau(t)), V_\tau(t)\rangle| dt
\end{aligned}$$

The first part on the right side tend to 0 since weak convergence of $V_\tau(t)$, the second part also goes to 0 because Hölder inequality:

$$(\int_0^T |\langle \nabla f(x(t)) - \nabla f(x_\tau(t)), V_\tau(t)\rangle| dt)^2 \le \int_0^T |\nabla f(x(t)) - \nabla f(x_\tau(t))|^2 dt \int_0^T |V_\tau(t)|^2 dt$$

By (3.1.3) and convergence of $x_\tau(t)$, Step 3 finished.

Step 4: prove $V(t)$ satisfies $-\nabla F(x(t)) = \nabla^2 F(x(t)) \cdot V(t)$ in weak sense. We have proved that $\forall f \in \mathcal{C}_c^\infty([0,T] \times \mathbb{R}^d)$, $\lim_{\tau\to 0} \int_0^T \langle \nabla f(x_\tau(t)), V_\tau(t)\rangle dt = \int_0^T \langle \nabla f(x(t)), V(t)\rangle dt$. On the other hand, because of Euler-Lagrange equation, we have

$$-\nabla F(x_\tau(t)) = \nabla^2 F(x_\tau(t-\tau)) \cdot V_\tau(t), \tag{3.1.4}$$

so

$$\langle \nabla f(x_\tau(t)), V_\tau(t)\rangle = \langle \nabla f(x_\tau(t)), -(\nabla^2 F)^{-1}(x_\tau(t-\tau))\nabla F(x_\tau(t))\rangle.$$

Denote $\langle \nabla f(x_\tau(t)), -(\nabla^2 F)^{-1}(x_\tau(t-\tau))\nabla F(x_\tau(t)\rangle$ and $-(\nabla^2 F)^{-1}(x(t))\nabla F(x(t))$ as $h_\tau(t)$ and $\bar{V}(t)$ respectively. Since convergence of $x_\tau(t)$ under uniform norm as τ goes to ∞ , it is easy to see that $h_\tau(t)$ converges to $\langle \nabla f(x(t)), \bar{V}(t)\rangle$ almost surely in t. Next, we use Fatou's lemma to finish the proof.
The crucial point is $h_\tau(t)$ is uniformly bounded from above. In fact,

$$|\langle \nabla f(x_\tau(t)), -(\nabla^2 F)^{-1}(x_\tau(t-\tau))\nabla F(x_\tau(t))\rangle| \le \frac{1}{2}(|\nabla f(x_\tau(t))|^2 + |V_\tau(t)|^2) \le C|x_\tau(t)|^2$$

Therefore, using Fatou's lemma, $\int_0^T \langle \nabla f(x(t)), \bar{V}(t)\rangle dt \le \liminf \int_0^T h_\tau(t)dt$. Then by (3.1.4) and our choice of weak convergent subsequence $V_\tau(t) \to V(t)$, we have

$$\int_0^T \langle \nabla f(x(t)), \bar{V}(t)\rangle dt \le \int_0^T \langle \nabla f(x(t)), V(t)\rangle dt$$

Change f into $-f$, we conclude that $\forall f \in C_c^\infty([0,T]\times\mathbb{R}^d)$,

$$f(x(T)) - f(x(0)) = \int_0^T \langle \nabla f(x(t)), -\nabla^2 F(x(t))\cdot \nabla F(x(t))\rangle dt.$$

Remark 3.1.1. *Actually , we can prove the existence of strong solution of Newton flow equation* (3.1.2) *by classical Peano's existence theorem. However, that proof is based on forward Euler approximation which may not be applicable to Newton flow in infinite dimensional space. Implicit Euler(backward Euler) method guarantees the estimate* (3.1.3) *and tends to have better stability.*

3.2 Newton flow equations on $\mathbb{P}(\mathbb{T}^d)$

In this chapter, if $\mu \in \mathbb{P}_{ac}(\mathbb{T}^d)$ with density ρ, we will use ρ to represent μ to simplify the notation. According to the Theorem 2.1.6 in Chapter 2, for $\mu,\ \nu \in \mathbb{P}_{ac}(\mathbb{T}^d)$, $\exists!$ convex function φ_μ^ν such that $(\nabla\varphi_\mu^\nu)_\sharp\mu = \nu$. And let $T_t = t\nabla\varphi_\mu^\nu + (1-t)Id$, then $\mu_t = (T_t)_\sharp\mu$ is the unique geodesic from μ to ν. The optimal transportation process can be described by

$$\partial_t\rho_t = -\nabla.(\rho_t \cdot \nabla\varphi_\mu^\nu \circ T_t^{-1}).$$

For $u \in \mathbf{T}_\mu$, the geodesic $\{\mu_s\}_{s\in[0,\epsilon)}$, starting from μ with initial velocity v, should satisfy

$$\begin{cases} \partial_s \mu_s = -\nabla.(\mu_s u_s) \\ \partial_s(\mu_s u_s) = -\nabla.(\mu_s u_s \otimes u_s). \end{cases}$$

The initial conditions are $\mu_0 = \mu$ and $u_0 = u$.

Next, we introduce grad operators and Hessian operators. According to [AKR96, LL07, RW21], at point $\mu \in \mathbb{P}(\mathbb{T}^d)$, the directional derivative of a functional F along $u \in \mathbf{T}_\mu$ is defined by

$$D_u F(\mu) := \lim_{\varepsilon \to 0^+} \frac{F((Id + \varepsilon u)_{\#}\mu) - F(\mu)}{\varepsilon}.$$

When $u \to D_u F(\mu)$ is a bounded linear functional on $\mathbf{T}_\mu$, then by Riesz representation, there exists a unique element $v \in \mathbf{T}_\mu$ such that

$$\langle v, u \rangle_{L^2(\mu)} = D_u F(\mu), \quad u \in \mathbf{T}_\mu.$$

We denote $v(\cdot)$ as $gradF(\mu, \cdot)$. We say that F is differentiable at μ if $gradF(\mu, x)$ exists. We write $F \in C^1(\mathbb{P}(\mathbb{T}^d))$ if F is differentiable at any $\mu \in \mathbb{P}(\mathbb{T}^d)$ and $gradF(\mu, x)$ is jointly continuous in $(\mu, x) \in \mathbb{P}(\mathbb{T}^d) \times \mathbb{T}^d$.

If furthermore, for $\forall u, v \in \mathbf{T}_\mu$,

$$D_u(D_v F(\mu))$$

exists, and the following form $H_\mu(u, v)$

$$H_\mu(u, v) = D_u(D_v F(\mu)) - \int_{\mathbb{T}^d} \langle gradF(\mu, x), \nabla v(x) u(x) \rangle \mu(dx)$$

defines a bounded, symmetric quadratic form on $\mathbf{T}_\mu \times \mathbf{T}_\mu$. Then we say F is second differentiable with respect to measure at μ. We denote H_μ as $Hess_\mu$. We say $F \in C^2(\mathbb{P}(\mathbb{T}^d))$ if $F \in C^1(\mathbb{P}(\mathbb{T}^d))$ and for every $\mu \in \mathbb{P}(\mathbb{T}^d)$, F is second order differentiable.

When $0 < \lambda_1 \leq Hess_\mu \leq \lambda_2$, then by Lax-Milgram theorem, we can define a bounded linear operator $\widetilde{Hess}_\mu F$ from $\mathbf{T}_\mu$ to$\mathbf{T}_\mu$, such that for $\forall u, v \in \mathbf{T}_\mu$

$$\langle \widetilde{Hess}_\mu F(u), v \rangle_\mu = Hess_\mu F(u, v).$$

If $d\mu = \rho dx$, $\rho > 0$ and $\rho \in C^2(\mathbb{T}^d)$, then, according to Chapter 4, the projection operator Π_ρ from $L^2(\mu, \mathbb{T}^d)$ to $\mathbf{T}_\mu$ is well defined, and for $\nabla\phi \in \mathbf{T}_\mu$,

$$\widetilde{Hess}_\mu F(\nabla\phi)(x) = \Pi_\mu(\nabla^2 \frac{\delta F}{\delta\rho}(\mu)(x) \cdot \nabla\phi(x) + \int \nabla_x \nabla_y \frac{\delta^2 F}{\delta\rho^2}(\mu, y, x)) \ \cdot \nabla\phi(y)\rho(y)dy) \tag{3.2.1}$$

where $\frac{\delta}{\delta\rho}$ stands for the gradient of the functional of $F(\rho)$ with respect to the $L^2(\mathbf{m}x)$.
In particular , for $F(\rho) = \int \rho V dx + \frac{1}{2}\int W(x-y)\rho(x)\rho(y)dxdy + \int \rho \log \rho dx$,

$$\begin{aligned} gradF(\rho) &= \nabla V + \nabla W * \rho + \nabla \log \rho, \\ Hess_\rho F(u,v) &= \int \langle u, \nabla^2 V v\rangle \rho dx + \int tr(\nabla u \nabla v)\rho dx \\ &\quad + \int \langle \nabla\phi(x) - \nabla\phi(y), \nabla^2 W(x-y)(\nabla\psi(x) - \nabla\psi(y))\rangle \rho dx \end{aligned} \tag{3.2.2}$$

3.2.1 Euler-Lagrange equation

Given time step $\tau > 0$, for $\rho \in \mathbb{P}_{ac}(\mathbb{T}^d)$, we assume that

$$F_{\rho,\tau}(\mu) := F(\mu) + \frac{1}{2\tau} Hess_\rho F(\nabla\phi_\rho^\mu - x, \nabla\phi_\rho^\mu - x)$$

is $(\frac{1}{\tau} + \lambda)-$geodesically convex .
For initial measure $\rho_\tau^0 = \rho_0 \in \mathbb{P}_{ac}(\mathbb{T}^d)$, we will construct discrete solution $\{\rho_\tau^n \in \mathbb{P}_{ac}(\mathbb{T}^d),\ n = 0, 1, \ldots, \frac{T}{\tau}\}$: given ρ_τ^n, we define μ_τ^{n+1} as the solution of the minimization problem $\min F_{\rho_\tau^n,\tau}$. Let $\nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}}$ is the optimal transport map from ρ_τ^n to μ_τ^{n+1}, then $\nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}} - x$ belongs to $L^2(\mathbb{T}^d, \rho_\tau^n)$. Let μ_τ^ϵ is a small perturbation around μ_τ^{n+1}, which satisfies $\mu_\tau^\epsilon = (Id + \epsilon\xi)_\# \mu_\tau^{n+1}$, where $\xi \in \mathbf{T}_{\mu_\tau^{n+1}}$. Suppose that the optimal transport map from ρ_τ^n to μ_τ^{n+1} and ρ_τ^n to μ_τ^ϵ are $\nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}}$ and $\nabla\phi_{\rho_\tau^n}^{\mu_\tau^\epsilon}$ respectively. Then we have the following lemma:

Lemma 3.2.1. $\nabla\phi_{\rho_\tau^n}^{\mu_\tau^\epsilon} = \nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}} + \epsilon\xi(\nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}}) + o(\epsilon)$.

Proof. Since $\mu_\tau^\epsilon = (Id + \epsilon\xi)_\# \mu_\tau^{n+1}$, for any $f \in C^\infty(\mathbb{T}^d)$, we have

$$\int f d\mu_\tau^\epsilon - \int f d\mu_\tau^{n+1} = \epsilon \int \langle \nabla f, \xi\rangle d\mu_\tau^{n+1} + o(\epsilon). \tag{3.2.3}$$

On the other hand,

$$\begin{aligned} &\int f d\mu_\tau^\epsilon - \int f d\mu_\tau^{n+1} \\ &= \int f(\nabla\phi_{\rho_\tau^n}^{\mu_\tau^\epsilon}) d\rho_\tau^n - \int f(\nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}}) d\rho_\tau^n \\ &= \int \langle \nabla f(\nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}}), \nabla\phi_{\rho_\tau^n}^{\mu_\tau^\epsilon} - \nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}}\rangle d\rho_\tau^n + o(||\nabla\phi_{\rho_\tau^n}^{\mu_\tau^\epsilon} - \nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}}||_{L^2(\rho_\tau^n)}). \end{aligned} \tag{3.2.4}$$

By triangle inequality,

$$|\nabla\phi_{\rho_\tau^n}^{\mu_\tau^\epsilon} - \nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}}| \leq |\epsilon\xi(\nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}})|,$$

Thus, $||\nabla\phi_{\rho_\tau^n}^{\mu_\tau^\epsilon} - \nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}}||_{L^2(\rho_\tau^n)} \leq C\epsilon$. As $\epsilon \to 0$, we can prove the lemma by comparing (3.2.3) with (3.2.4).

□

We use this lemma to deal with the following inequality. Because

$$F(\mu_\tau^{n+1}) + \frac{1}{\tau}Hess_{\rho_\tau^n}F(\nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}} - x, \nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}} - x) \leq F(\mu_\tau^\epsilon) + \frac{1}{\tau}Hess_{\rho_\tau^n}F(\nabla\phi_{\rho_\tau^n}^{\mu_\tau^\epsilon} - x, \nabla\phi_{\rho_\tau^n}^{\mu_\tau^\epsilon} - x),$$

we have

$$F(\mu_\tau^\epsilon) - F(\mu_\tau^{n+1}) \geq \epsilon\frac{1}{\tau}Hess_{\rho_\tau^n}F(\nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}} - x, \xi\circ\nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}}) + o(\epsilon). \tag{3.2.5}$$

But μ_τ^{n+1} may not be absolutely continuous, which stops us defining the next step discrete solution. To overcome this difficulty, we pick a mollifier η_τ on $\mathbb{T}^d$, which satisfying $\int_{\mathbb{T}^d}\eta_\tau = 1$ and

$$\int_{\mathbb{T}^d}|x - x_o|^2\eta_\tau(x)dx \leq \tau^6.$$

for some fixed point $x_0 \in \mathbb{T}^d$. Define $\rho_\tau^{n+1} = \mu_\tau^{n+1} * \eta_\tau$. Since $0 \leq x \leq 1$, it holds $x^2 \leq x$. Therefore,

$$W_2^2(\rho_\tau^{n+1}, \mu_\tau^{n+1}) = \inf\int|x-y|^2d\gamma(x,y) \leq \inf\int|x-y|d\gamma(x,y) = W_1(\rho_\tau^{n+1}, \mu_\tau^{n+1}),$$

By Kantorovich-Rubinstein theorem,

$$W_1(\rho_\tau^{n+1}, \mu_\tau^{n+1}) = \sup\left\{\int_{\mathbb{T}^d}\varphi d(\rho_\tau^{n+1} - \mu_\tau^{n+1});\ \varphi \in L^1(d|\rho_\tau^{n+1} - \mu_\tau^{n+1}|),\ ||\varphi||_{Lip} \leq 1\right\}.$$

So,

$$W_1(\rho_\tau^{n+1}, \mu_\tau^{n+1}) \leq \int_{\mathbb{T}^d}|\varphi - \varphi * \eta_\tau|\rho_\tau^{n+1}dx \leq \int\int_{\mathbb{T}^d}|x-y|^2\eta_\tau(x-y)dy\rho_\tau^{n+1}(x)dx \leq \tau^6.$$

Such error is so small that it will never influence the convergence of ρ_τ^{n+1} . We will derive the Euler-Lagrange equation. Due to (3.2.5), let $\xi = \nabla f$, then for $\forall f \in C^\infty$,

$$\langle\nabla f, -gradF(\mu_\tau^{n+1})\rangle_{L^2(\mu_\tau^{n+1})} = \frac{1}{\tau}Hess_{\rho_\tau^n}F(\nabla f, \nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}} - x), \tag{3.2.6}$$

as $\epsilon \to 0$. Note that

$$\nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}} - x = \nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}} - \nabla\phi_{\rho_\tau^n}^{\rho_\tau^{n+1}} + \nabla\phi_{\rho_\tau^n}^{\rho_\tau^{n+1}} - x,$$

Thus, the right hand side of (3.2.6) becomes

$$\frac{1}{\tau}Hess_{\rho_\tau^n}F(\nabla f, \nabla\phi_{\rho_\tau^n}^{\rho_\tau^{n+1}} - x) + \frac{1}{\tau}Hess_{\rho_\tau^n}F(\nabla f, \nabla\phi_{\rho_\tau^n}^{\mu_\tau^{n+1}} - \nabla\phi_{\rho_\tau^n}^{\rho_\tau^{n+1}}).$$

If one wants to derive the Euler-Lagrange equation for $\{\rho_\tau^n\}$, we need Lipschitz conditions on $HessF$ and $gradF$. Next , we introduce the corresponding definitions.

Definition 3.2.2. *We say* $\widetilde{Hess}F$ *is* L^1*-Lips, if* $\forall\nu \in \mathbb{P}_{2,ac}(\mathbb{T}^d)$, $\mu \in \mathbb{P}(\mathbb{T}^d)$ *and* $\forall\xi \in \mathbf{T}_\mu$ *satisfying* $\xi \circ \nabla\varphi \in L^2(\nu)$ *(Here,* $\nabla\varphi$ *is the optimal transport map from* ν *to* μ*),*

$$\frac{||\widetilde{Hess}_\nu F(\xi \circ \nabla\varphi) - \widetilde{Hess}_\mu F(\xi) \circ \nabla\varphi||_{L^1(\nu)}}{W_2(\mu,\nu)} \leq L||\xi||_{L^2(\mu)}. \tag{3.2.7}$$

Proposition 3.2.3. *For* $V, W \in C^3(\mathbb{T}^d)$, $F(\rho) = \int \rho V dx + \frac{1}{2}\int W(x-y)\rho(x)\rho(y)dxdy$ *,* $\widetilde{Hess}_\rho F$ *is* L^1-*Lips .*

Proof. Because $\widetilde{Hess}_\rho F(u) = \nabla^2 V u + \int \nabla^2 W(x-y)(u(x) - u(y))\rho(y)dy$, we have

$$\begin{aligned}
&||\widetilde{Hess}_\nu F(\Pi_\nu(\xi \circ \nabla\varphi)) - \widetilde{Hess}_\mu F(\xi) \circ \nabla\varphi||_{L^1(\nu)} \\
&\leq \int |(\nabla^2 V(x) - \nabla^2 V(\nabla\varphi(x))) \cdot \xi \circ \nabla\varphi(x)|\nu(x)dx + \\
&+ \int |\int (\nabla^2 W(x-y) - \nabla^2 W(\nabla\varphi(x) - \nabla\varphi(y))) \cdot (\xi \circ \nabla\varphi(x) - \xi \circ \nabla\varphi(y))\nu(y)dy|\nu(x)dx \\
&\leq K_1||\xi||_{L^2(\mu)}W_2(\mu,\nu) + \\
&K_2[\int |x - \nabla\varphi(x) - y + \nabla\varphi(y)|^2\nu(x)\nu(y)dxdy]^{\frac{1}{2}} \cdot [\int |\xi \circ \nabla\varphi(x) - \xi \circ \nabla\varphi(y)|^2\nu(x)\nu(y)dxdy]^{\frac{1}{2}} \\
&\leq L||\xi||_{L^2(\mu)}W_2(\mu,\nu),
\end{aligned}$$

where we assume that $|\nabla^2 V|, \ |\nabla^2 W|$ are controlled by $K_1, \ K_2$ on $\mathbb{T}^d$.

$\square$

Let $\nabla\phi_n^{n+1} = \nabla\phi_{\rho_\tau^n}^{\rho_\tau^{n+1}}$. If $\widetilde{Hess}F$ is L^1-Lips, then (3.2.6) becomes

$$\begin{aligned}
&\left|\langle\nabla f, -gradF(\mu_\tau^{n+1})\rangle_{L^2(\mu_\tau^{n+1})} - \frac{1}{\tau}Hess_{\rho_\tau^n}F(\nabla f, \nabla\phi_{\rho_\tau^n}^{\rho_\tau^{n+1}} - x)\right| \\
&\leq \frac{1}{\tau}\left|\langle\nabla f, \widetilde{Hess}_{\rho_\tau^{n+1}}F(\nabla\phi_{\rho_\tau^{n+1}}^{\mu_\tau^{n+1}} - x) \circ \nabla\phi_{\rho_\tau^n}^{\rho_\tau^{n+1}}\rangle_{\rho_\tau^n}\right| \\
&+ \max|\nabla f|\frac{L}{\tau}W_2(\rho_\tau^n, \rho_\tau^{n+1})||\nabla\phi_{\rho_\tau^{n+1}}^{\mu_\tau^{n+1}} - x||_{L^2(\rho_\tau^{n+1})} \\
&\leq \lambda_2 \max|\nabla f|||\frac{1}{\tau}(\nabla\phi_{\rho_\tau^{n+1}}^{\mu_\tau^{n+1}} - x)||_{L^2(\rho_\tau^{n+1})} + \max|\nabla f|L\tau W_2(\rho_\tau^n, \rho_\tau^{n+1}) \\
&\leq C\max|\nabla f|\tau^2(1 + W_2(\rho_\tau^n, \rho_\tau^{n+1})).
\end{aligned}$$

We also give a Lipschitz condition for $gradF$:

Definition 3.2.4. *We say $gradF$ is L^2-Lips , if there exists $K > 0$ such that for all $\nu, \mu \in \mathbb{P}(\mathbb{T}^d)$ and $\pi \in \mathcal{C}_o(\nu, \mu)$,*

$$\int |\nabla\Psi(\mu)(y) - \nabla\Psi(\nu)(x)|^2 d\pi(x, y) \leq KW_2^2(\mu, \nu).$$

In particular, if $\nu \in \mathbb{P}_{ac}(\mathbb{T}^d)$, the condition becomes

$$\int |\nabla\Psi(\mu) \circ \nabla\phi_\nu^\mu - \nabla\Psi(\nu)|^2 d\nu \leq KW_2^2(\mu, \nu).$$

where $\nabla\phi_\nu^\mu$ is the optimal transport map from ν to μ.

Proposition 3.2.5. *For $F(\mu) = \int V d\mu + \frac{1}{2}\int W(x - y)d\mu(y)d\mu(x)$, suppose that $\nabla V, \nabla W$ are differentiable, then $gradF(\rho)$ is L^2-Lips.*

Proof.

$$\begin{aligned}
&\int_{\mathbb{T}^d\times\mathbb{T}^d} |gradF(\nu, y) - gradF(\mu, x)|^2 d\pi(x, y) \\
&= \int_{\mathbb{T}^d\times\mathbb{T}^d} |\nabla V(y) - \nabla V(x) + \nabla W * \mu(y) - \nabla W * \nu(x)|^2 d\pi(x, y) \\
&\leq 2\int_{\mathbb{T}^d\times\mathbb{T}^d} |\nabla V(y) - \nabla V(x)|^2 d\pi(x, y) \\
&+ 2\int_{\mathbb{T}^d\times\mathbb{T}^d} \Big|\int_{\mathbb{T}^d\times\mathbb{T}^d} \nabla W(y - z)d\pi(c, z) - \int_{\mathbb{T}^d\times\mathbb{T}^d} \nabla W(x - c)d\pi(c, z)\Big|^2 d\pi(x, y) \\
&\leq 2K_1W_2^2(\nu, \mu) + 2K_2\int_{\mathbb{T}^d\times\mathbb{T}^d} |y - z - x + c|^2 d\pi(c, z)d\pi(x, y) \\
&\leq KW_2^2(\nu, \mu).
\end{aligned}$$

□

We give some notation. For $t \in [n\tau, (n + 1)\tau)$:

1. $\rho_\tau(t) = \rho_\tau^{n+1}$; $\bar{\rho}_\tau(t) = \rho_\tau^n$
2. For $\forall\mu, \nu \in \mathbb{P}_{2,ac}(\mathbb{T}^d)$, let $\nabla\phi_\mu^\nu$ be the optimal transport map from μ to ν. its inverse $\nabla\phi_\nu^\mu$ is the optimal transport map from ν to μ. Especially, for $t \in [n\tau, (n + 1)\tau)$, $\phi_\tau(t) = \phi_{\rho_\tau^n}^{\rho_\tau^{n+1}}$ is denoted as ϕ_n^{n+1}.
3. We connect the adjacent points of discrete solution ρ_τ by a unique geodesic. We denote this continuous polyline as $\tilde{\rho}_\tau$.
4. $V_\tau(t, x) = \frac{1}{\tau}(x - \nabla\phi_{n+1}^n(t, x)),$ for $t \in [n\tau, (n + 1)\tau)$; $V_\tau^{n+1} = \frac{\nabla\phi_\tau(t) - x}{\tau}$.

For $\forall f \in C^{\infty}$, we have

$$\begin{aligned}
&\left|\langle \nabla f, -gradF(\rho_\tau^{n+1})\rangle_{L^2(\rho_\tau^{n+1})} - \frac{1}{\tau}Hess_{\rho_\tau^{n+1}}F(\nabla f, x-\nabla\phi_{n+1}^n)\right| \\
&\leq \left|\langle \nabla f, -gradF(\mu_\tau^{n+1})\rangle_{L^2(\mu_\tau^{n+1})} - \frac{1}{\tau}Hess_{\rho_\tau^n}F(\nabla f, \nabla\phi_n^{n+1}-x)\right| \\
&+ \left|\langle \nabla f, -gradF(\mu_\tau^{n+1})\rangle_{L^2(\mu_\tau^{n+1})} - \langle \nabla f, -gradF(\rho_\tau^{n+1})\rangle_{L^2(\rho_\tau^{n+1})}\right| \\
&+ \left|\frac{1}{\tau}Hess_{\rho_\tau^n}F(\nabla f, \nabla\phi_n^{n+1}-x) - \frac{1}{\tau}Hess_{\rho_\tau^{n+1}}F(\nabla f, x-\nabla\phi_{n+1}^n)\right| \\
&\leq C\tau^2(1+W_2(\rho_\tau^n,\rho_\tau^{n+1})) + CW_2(\mu_\tau^{n+1},\rho_\tau^{n+1}) + CW_2(\rho_\tau^n,\rho_\tau^{n+1})||\frac{\nabla\phi_n^{n+1}-x}{\tau}||_{L^2(\rho_\tau^n)} \\
&\leq C||V_\tau^{n+1}||^2_{L^2(\rho_\tau^n)}\tau + O(\tau^2).
\end{aligned} \tag{3.2.8}$$

3.2.2 Existence of solutions to the Newton flow equation

Assumptions 1:

1. F is proper, lower semicontinuous(l.s.c), λ_1-geodesically convex and $F \in C^2(\mathbb{P}(\mathbb{T}^d))$.
2. $gradF$ is L^2-Lips (see definition 3.2.4) .
3. $0 < \lambda_1 \leq \widetilde{Hess}_\rho F \leq \lambda_2$, $\widetilde{Hess}$ is L^1-Lips in $\mathbb{P}_{2,ac}(\mathbb{T}^d)$ (see definition 3.2.2) .
4. For any $\rho \in \mathbb{P}_{ac}(\mathbb{T}^d), \tau > 0$, $F_{\rho,\tau}$ is $(\frac{1}{\tau}+\lambda)-$geodesically convex.
5. For any $\mu,\nu \in \mathbb{P}(\mathbb{T}^d)$ and $f \in C^\infty(\mathbb{T}^d)$,

$$|\widetilde{Hess}_\mu F(\nabla f) - \widetilde{Hess}_\nu F(\nabla f)| \leq C_f W_2(\mu,\nu).$$

where C_f is a constant only dependent on f .

Theorem 3.2.6. *Under Assumption 1, suppose that the initial value $\mu_0 = \rho_0 dx \in \mathbb{P}_{ac}(\mathbb{T}^d)$, then there exist a solution $\mu_t \in \mathbb{P}(\mathbb{T}^d)$ to the following Newton flow equation in distributional sense:*

$$\begin{cases} \partial_t \mu = -\nabla.(\mu v) \\ Hess_{\mu_t}(v_t,\nabla f) = \langle -gradF(\mu_t),\nabla f\rangle_{\mu_t}, \quad \forall f \in C_c^\infty([0,T]\times\mathbb{T}^d) \end{cases} \tag{3.2.9}$$

Proof. **Step 1:** We will prove $\{\tilde\rho_\tau(t)\}_\tau$ has a convergent subsequence under $C([0,\infty),\omega^*-\mathbb{P}(\mathbb{T}^d))$. Given $\rho_\tau^n \in \mathbb{P}_{ac}(\mathbb{T}^d)$, since μ_τ^{n+1} is a solution to the following problem

$$\inf_\mu F_{\rho_\tau^n,\tau}(\mu),$$

thus

$$F(\rho_\tau^n) \geq F(\mu_\tau^{n+1}) + \frac{\tau\lambda_1}{2}(\frac{W_2(\rho_\tau^n, \mu_\tau^{n+1})}{\tau})^2.$$

Then for $\forall\, n, m(n < m)$,

$$\begin{aligned} W_2(\rho_\tau^n, \rho_\tau^m) &\leq \tau(\sum_{i=m}^{n-1} \frac{W_2(\rho_\tau^i, \rho_\tau^{i+1})}{\tau}) \\ &\leq \tau(\sum_{i=m}^{n-1}(\frac{W_2(\rho_\tau^i, \rho_\tau^{i+1})}{\tau})^2)^{\frac{1}{2}}(m-n)^{\frac{1}{2}} \\ &\leq \tau(\sum_{i=m}^{n-1}(\frac{W_2(\rho_\tau^i, \mu_\tau^{i+1}) + W_2(\mu_\tau^{i+1}, \rho_\tau^{i+1})}{\tau})^2)^{\frac{1}{2}}(m-n)^{\frac{1}{2}} \\ &\leq C\tau^{\frac{1}{2}}(m-n)^{\frac{1}{2}}(\sum_{i=m}^{n-1} F(\rho_\tau^i) - F(\mu_\tau^{i+1}))^{\frac{1}{2}} + 2(m-n)\tau^3. \end{aligned} \tag{3.2.10}$$

Note that F is $\lambda-$geodesically convex, let $\mu_0 = \rho_\tau^{i+1}$, $\mu_1 = \mu_\tau^{i+1}$ and set

$$g(t) := F(\mu_t) = F((x + t(\nabla\phi_{\rho_\tau^{i+1}}^{\mu_\tau^{i+1}} - x))_{\#}\mu_0).$$

Then, g id $\lambda-$convex, and $g'(1) = 0$. By mean-value theorem,

$$F(\rho_\tau^{i+1}) - F(\mu_\tau^{i+1}) = g(0) - g(1) \leq |g'(0)| = |\langle gradF(\rho_\tau^{i+1}), \nabla\phi_{\rho_\tau^{i+1}}^{\mu_\tau^{i+1}} - x\rangle_{\rho_\tau^{n+1}}|.$$

Because for all $\mu, \nu \in \mathbb{P}(\mathbb{T}^d)$, $W_2(\mu, \nu) \leq d$. Alternatively, $gradF$ is L^2-Lips, we have

$$||gradF(\mu)||_{L^2(\mu)} \leq K.$$

Substituting this inequality to (3.2.10), we get

$$W_2(\rho_\tau^n, \rho_\tau^m) \leq C\tau^{\frac{1}{2}}(m-n)^{\frac{1}{2}}(F(\rho_\tau^0) - \inf F + K(m-n)\tau^3)^{\frac{1}{2}} + 2(m-n)\tau^3,$$

therefore,

$$W_2(\rho_\tau^n, \rho_\tau^m) \leq C(|m-n|\tau)^{\frac{1}{2}} + o(\tau^2).$$

In particular, we have the following energy estimate:

$$W_2^2(\rho_\tau^n, \rho_\tau^{n+1}) \leq C\tau.$$

Due to the construction of ρ_τ , it is easy to see that

$$W_2^2(\rho_\tau(t) - \rho_\tau(s)) \leq C|t-s|.$$

We have proved equi-continuity. Uniform boundedness holds because $F(\rho_\tau^n) \leq F(\rho_0)$. Then, according the compactness theorem, $\{\tilde{\rho}_\tau\}$ has a convergent subsequence under $C([0,T], \omega^* - \mathbb{P}(\mathbb{T}^d))$, converging to $\{\mu_t,\, t \in [0,T]\}$.

Step 2: Let the discrete rescaled optimal plans $\gamma_\tau := (i_x \times V_\tau)_\sharp \rho_\tau$. For every bounded interval $I_T := [0,T]$, denoting by $X_T := X \times I_T$, we can canonically identify $T^{-1}\rho_\tau$ to an element of $\mathbb{P}(\mathbb{T}^d \times I_T)$ and $T^{-1}\gamma_\tau$ to an element in $\mathbb{P}(\mathbb{T}^d \times I_T \times \mathbb{T}^d)$, simply by integrating with respect to the (normalized) Lebesgue measure $T^{-1}dm$ in I_T . Therefore $V_\tau(t)$ can be seen as a vector field in $L^2(\rho_\tau(t))$. By (3.2.10),

$$\int_{[0,T]} \int_{\mathbb{T}^d} V_\tau^2(t,x)\rho_\tau(t,x)dxdt = \sum_{i=1}^{\frac{T}{\tau}} \frac{W_2^2(\rho_\tau^i, \rho_\tau^{i+1})}{\tau^2} < F(\rho_\tau^0) - \inf F + K\frac{T}{\tau}\tau^2 \leq C'. \tag{3.2.11}$$

By ([AGS05]. p.114, lemma 5.1.12), (3.2.11) guarantees that $T^{-1}\gamma_\tau$ is tight with respect to weak* topology in $\mathbb{P}(\mathbb{T}^d \times I_T \times \mathbb{T}^d)$. Therefore we can extract a subsequence γ_{τ_h} weakly converging to γ. Since $\pi_\sharp^{1,2} T^{-1}\gamma_\tau = T^{-1}\rho_\tau$, so $\pi_\sharp^{1,2}\gamma = T^{-1}\mu$. We can define

$$V(x_1,t) \triangleq \int_{\mathbb{T}^d} x_2 d\gamma_{x_1,t}(x_2),$$

where $\gamma_{x_1,t}$ is the disintegration of γ w.r.t. ρ. According to Theorem 5.4.4 in [AGS05], we have

$$\int |V|^2 d\mu \leq \liminf_{h\to\infty} \int |V_{\tau_h}|^2 \rho_{\tau_h} dx \leq C'.$$

For the sake of convenience, we will still use ρ_τ to represent the subsequence ρ_{τ_h} .

Step 3: Next, we will prove $\partial_t \mu = -\nabla.(\mu V)$ holds in distribution, i.e. $\forall f(t,x) \in C_c^\infty(I_T \times \mathbb{T}^d)$,

$$-\int_{I_T \times \mathbb{T}^d} \partial_t f d\mu = \int_{I_T \times \mathbb{T}^d} < \nabla f, V > d\mu.$$

Note that

$$\begin{aligned}
&\int f\rho_\tau^{n+1}dx - \int f\rho_\tau^n dx \\
&= \int (f - f(\nabla\phi_{n+1}^n))\rho_\tau^{n+1}dx \\
&= \int \langle \nabla f(x), x - \nabla\phi_{n+1}^n(x)\rangle \rho_\tau^{n+1}dx + C_f||\nabla\phi_{n+1}^n(x) - x||^2_{L^2(\rho_\tau^{n+1})} \\
&= \tau\int \langle \nabla f(x), y\rangle d(\gamma_\tau(\frac{(n+1)T}{\tau}))(x,y) + C\tau^2||V_\tau(\frac{(n+1)T}{\tau})||^2_{L^2(\rho_\tau^{n+1})},
\end{aligned}$$

thus

$$\begin{aligned}
&-\int_{I_T\times\mathbb{T}^d}\partial_t f d\mu = \lim_{\tau\to 0} -\int_{I_T\times\mathbb{T}^d}\partial_t f\rho_\tau(t,x)dxdt \\
&= \lim_{\tau\to 0} -\frac{1}{\tau}\int_{I_T\times\mathbb{T}^d}(f(t+\tau,x) - f(t,x))\rho_\tau(t,x)dxdt \\
&= \lim_{\tau\to 0}\int_{\mathbb{T}^d\times I_T\times\mathbb{T}^d}\langle\nabla f(x), y\rangle d\bar{\gamma}_\tau(x,t,y) \\
&= \int_{I_T\times\mathbb{T}^d}\langle\nabla f, V\rangle d\mu.
\end{aligned}$$

Step 3 is finished.

Step 4: Finally, we want to prove

$$Hess_\mu F(\nabla f, v) = \langle \nabla f, -gradF(\mu)\rangle_\mu, \quad \text{for any} f \in C_c^\infty(\mathbb{T}^d \times I_T). \tag{3.2.12}$$

Firstly, some definitions and assumptions should be stated.

Definition 3.2.7. *weak convergence and strong convergence If* $(\mu_n) \subset \mathbb{P}_2(\mathbb{T}^d)$ *narrowly converges to* $\mu \in \mathbb{P}_2(\mathbb{T}^d)$. *Let* $v_n \in L^1(\mu_n)$. *We say* v_n *weakly cenverges to* $v \in L^1(\mu)$, *if*

$$\lim_{n\to\infty}\int_{\mathbb{T}^d}\langle\nabla f, v_n\rangle d\mu_n = \int_{\mathbb{T}^d} < \nabla f, v > d\mu, \quad \forall f \in C^\infty(\mathbb{T}^d). \tag{3.2.13}$$

Furthermore, we say v_n *strongly converges to* $v \in L^2$, *if* (3.2.13) *holds and*

$$\limsup_{n\to\infty}||v_n||_{L^2(\mu_n)} \le ||v||_{L^2(\mu)}.$$

We need the following lemma (see [AGS05], Theorem 5.4.4):

Lemma 3.2.8. *If μ_n converges to μ narrowly, $v_n \in L^2(\mu_n)$ satisfy*

$$\sup_{n\in\mathbb{N}} \int_{\mathbb{T}^d} |v_n(x)|^2 d\mu_n(x) < +\infty. \tag{3.2.14}$$

If v_n strongly converge to v, then γ_n narrowly converges to $(i \times v)_\sharp \mu$ and

$$\lim_{n\to\infty} ||v_n||^2_{\mathcal{L}^2(\mu_n)} = \lim_{n\to\infty} \int_{\mathbb{T}^d \times \mathbb{T}^d} |x_2|^2 d\gamma_n = ||v||^2_{\mathcal{L}^2(\mu)}.$$

Proposition 3.2.9. *For any fixed $t \in [0, T]$, there exists subsequence $\{-gradF(\rho_\tau(t))\}$ strongly converges to $-gradF(\mu_t)$.*

Proof. Because

$$\begin{aligned}
&||gradF(\rho_{\tau_j}(t))||_{\mathcal{L}^2(\rho_{\tau_j}(t))} \\
&= \int |gradF(\rho_{\tau_j}(t))|^2 \rho_{\tau_j}(t) dx \\
&= \int |gradF(\mu_t)) \circ \nabla\phi^{\mu_t}_{\rho_{\tau_j}(t)} - gradF(\mu_t) \circ \nabla\phi^{\mu_t}_{\rho_{\tau_j}(t)} + gradF(\rho_{\tau_j}(t))|^2 \rho_{\tau_j}(t) dx \\
&\leq 2 \int |gradF(\mu_t) \circ \nabla\phi^{\rho(t)}_{\rho_{\tau_j}(t)}|^2 \rho_{\tau_j}(t) dx \\
&+ 2 \int |gradF(\mu_t) \circ \nabla\phi^{\mu_t}_{\rho_{\tau_j}(t)} - gradF(\rho_{\tau_j}(t))|^2 \rho_{\tau_j}(t) dx \\
&\leq 2 \int |gradF(\mu_t)|^2 d\mu_t + KW_2^2(\mu_t, \rho_{\tau_j}(t)),
\end{aligned} \tag{3.2.15}$$

then by lemma 3.2.8and $\lambda-$ geodesically convexity, as $n \to \infty$, there exists a subsequence $\{-gradF(\rho_{\tau_j}(t))\}$ weakly converging to $-gradF(\rho(t))$ (see [AGS05],lemma 10.1.3) . And (3.2.15) shows

$$\limsup_{j\to\infty} ||gradF(\rho_{\tau_j}(t))||_{L^2(\rho_{\tau_j}(t))} \leq ||gradF(\rho(t))||_{L^2(\rho(t))},$$

This means $-gradF(\rho_{\tau_j}(t))$ strongly converges to $-gradF(\rho(t))$. □

Next , we assume that $f(t) \in C_0^\infty([0, T]);\ g(x) \in C^\infty(\mathbb{T}^d)$. Let

$$\tilde{V}(t) = \widetilde{Hess}^{-1}_{\mu_t} F(-gradF(\mu_t)).$$

We will prove :

Lemma 3.2.10. *as* $\tau \to 0$*,*

$$Hess_{\rho_\tau(t)} F(\nabla g, V_\tau(t)) + \langle \nabla g, gradF(\mu_t)\rangle_{L^2(\mu_t)} \to 0. \tag{3.2.16}$$

Proof. Note that, because of (3.2.8) and (3.2.10), we have

$$\begin{aligned}
&|\int_0^T f(t)\int \langle \nabla g, -gradF(\rho_\tau(t))\rangle \rho_\tau(t)dxdt \\
&- \int_0^T f(t)\int \langle \nabla g, \widetilde{Hess}_{\rho_\tau(t)} F(V_\tau(t))\rangle \rho_\tau(t)dxdt| \\
&= |\int_0^T f(t)\int \langle \nabla g, -\widetilde{Hess}_{\rho_\tau(t)} F(V_\tau(t)) - gradF(\rho_\tau(t))\rangle \rho_\tau(t)dxdt| \\
&\le \sum_{i=0}^{\frac{T}{\tau}} \int_{i\tau}^{(i+1)\tau} |f(t)|dt \cdot \max|\nabla g| \cdot ||V_\tau^{i+1}||^2_{L^2(\rho_\tau^{i+1})}\tau dt \\
&\le C\tau \int_{[0,T]\times\mathbb{T}^d} |V_\tau|^2 \rho_\tau dx \\
&\le C'\tau.
\end{aligned} \tag{3.2.17}$$

Therefore, according to Proposition 3.2.9, $\{\widetilde{Hess}_{\rho_\tau(t)} F(V_\tau(t))\}$ converges weakly to $-gradF(\mu_t)$. This proposition is proved.

□

Use the convexity, we can get a more accurate estimate on $W_2(\rho_\tau^n, \rho_\tau^{n+1})$:

Proposition 3.2.11. $W_2(\rho_\tau^n, \rho_\tau^{n+1}) \le C\tau.$

Proof. By Assumption 1(4), $F_{\tau,\rho}$ is $(\lambda + \frac{1}{\tau})-$geodesically convex. Set $\mu =$ argmin$F_{\tau,\rho}$. Let the curve $\{\mu_t\}_{t\in[0,1]}$ be the geodesic from ρ to μ, then $F_{\tau,\rho}(\mu_t)$ is convex with respect to t, i.e. for $0 < t < 1$,

$$F_{\tau,\rho}(\mu_t) \le tF_{\tau,\rho}(\mu) + (1-t)F_{\tau,\rho}(\rho) - \frac{\lambda + \frac{1}{\tau}}{2} t(1-t)W_2^2(\rho,\mu).$$

Since $t = 1$ arrive the minimum, the derivative of the right hand side of the above equality at $t = 1$ must be no bigger than 0:

$$F(\mu) - F(\rho) + \frac{1}{\tau} Hess_\rho F(\nabla\phi_\rho^\mu - x, \nabla\phi_\rho^\mu - x) + \frac{\lambda + \frac{1}{\tau}}{2} W_2^2(\rho,\mu) < 0.$$

By the properties of $HessF$,

$$C\frac{W_2^2(\rho,\mu)}{\tau^2} < \frac{F(\rho) - F(\mu)}{W_2(\rho,\mu)} \frac{W_2(\rho,\mu)}{\tau}.$$

It follows that

$$C\frac{W_2(\rho,\mu)}{\tau} < \frac{F(\rho)-F(\mu)}{W_2(\rho,\mu)} \le ||gradF(\rho)||_{L^2(\rho)}, \tag{3.2.18}$$

which means $W_2(\rho_\tau^n, \mu_\tau^{n+1}) < C||gradF(\rho_\tau^n)||_{L^2(\rho_\tau^n)}\tau$.

□

We will prove that, for all $t \in [0,T]$,

$$Hess_{\rho_\tau(t)}F(\nabla g, V_\tau(t)) \to Hess_{\mu_t}F(\nabla g, V_t). \tag{3.2.19}$$

Note that

$$\begin{aligned}
&\left|Hess_{\rho_\tau(t)}F(\nabla g, V_\tau(t)) - Hess_{\mu_t}F(\nabla g, V_t)\right| \\
&= \left|\int \langle \widetilde{Hess}_{\rho_\tau(t)}F(\nabla g), V_\tau(t)\rangle \rho_\tau(t)dx - \int \langle \widetilde{Hess}_{\mu_t}F(\nabla g), V_t\rangle d\mu_t\right| \\
&\le \left|\int \langle \widetilde{Hess}_{\rho_\tau(t)}F(\nabla g), V_\tau(t)\rangle \rho_\tau(t)dx - \int \langle \widetilde{Hess}_{\mu_t}F(\nabla g), V_\tau(t)\rangle \rho_\tau(t)dx\right| \\
&+ \left|\int \langle \widetilde{Hess}_{\mu_t}F(\nabla g), V_\tau(t)\rangle \rho_\tau(t)dx - \int \langle \widetilde{Hess}_{\mu_t}F(\nabla g), V_t\rangle d\mu_t\right| \\
&= (I) + (J).
\end{aligned}$$

For (J), set $u = \widetilde{Hess}_{\mu_t}F(\nabla g)$, we have

$$\begin{aligned}
&\lim_{\tau\to 0}\int_{\mathbb{T}^d\times I_T} \langle u, V_\tau\rangle \rho_\tau(t)dx \\
&= \lim_{\tau\to 0}\int_{\mathbb{T}^d\times \mathbb{T}^d} \langle u(x_1), x_2\rangle d\gamma_\tau(x_1,t,x_2) \\
&= \int_{\mathbb{T}^d\times \mathbb{T}^d} \langle u(x_1), x_2\rangle d\gamma(x_1,t,x_2) \\
&= \int_{\mathbb{T}^d} \langle u, V\rangle d\mu_t.
\end{aligned}$$

It folllows that $(J) \to 0$.
For (I), by Assumption 1(5),

$$\begin{aligned}
&\left|\int \langle \widetilde{Hess}_{\rho_\tau(t)}F(\nabla g) - \widetilde{Hess}_{\mu_t}F(\nabla g), V_\tau(t)\rangle \rho_\tau(t)dx\right| \\
&\le \max|\nabla g| W_2(\rho_\tau(t),\mu_t)||V_\tau||_{L^2(\rho_\tau(t))}.
\end{aligned}$$

Due to Proposition 3.2.11, we get

$$||V_\tau||_{L^2(\rho_\tau(t))} = \frac{1}{\tau} W_2(\rho_\tau(t), \rho_\tau(t+\tau)) \le C.$$

It follows that $(I) \to 0$. (3.2.19) has been proved. This means $Hess_{\mu_t}(v_t, \nabla f)$ conerverges to $\langle -gradF(\mu_t), \nabla f\rangle_{\mu_t}$ for $t - a.e.$. We can prove (3.2.12) by dominated convergence theorem. Therefore, we have proved the existence of solutions to the Newton flow equation.

□

3.2.3 Uniqueness

Next, we state the abstract uniqueness result to finish the well-posedness of Newton flow equation in $\mathbb{P}(\mathbb{T}^d)$.

Theorem 3.2.12. *Under the Assumption 1, if* $\widetilde{Hess}_\mu^{-1} F(gradF(\mu))$ *is* L^2-*Lips , and the solutions to* (3.2.9) *are all abolutely continuous, i.e.* $\mu_t \in \mathbb{P}_{ac}(\mathbb{T}^d)$*, then* (3.2.9) *has a unique solution in sense of* W_2 *metric.*

Proof. Let $\rho_t^1, \rho_t^2 \in \mathbb{P}_{ac}(\mathbb{T}^d)$ are two absolutely continuous solutions to (3.2.9)with the same initial measure ρ_0. Denote $\nabla\phi_t^{1,2}(\nabla\phi_t^{2,1})$ as the optimal transport map from $\rho_t^1(\rho_t^2)$ to $\rho_t^2(\rho_t^1)$, then $\left(\nabla\phi_t^{1,2}\right)^* = \nabla\phi_t^{2,1}$. Let $\nabla\Phi_t^{1,2} = \nabla\phi_t^{1,2} - x$. Note that

$$\nabla\Phi_t^{2,1} \circ \nabla\varphi_t^{1,2} = (\nabla\phi_t^{2,1} - x) \circ \nabla\phi_t^{1,2} = -\nabla\Phi_t^{1,2},$$

Thus,

$$\begin{aligned} &\frac{d}{dt} W_2^2(\rho_t^1, \rho_t^2) \\ &= 2 < \nabla\Phi_t^{1,2}, Hess_{\rho_t^1}^{-1} F(-gradF(\rho_t^1)) >_{\rho_t^1} + 2 < \nabla\Phi_t^{2,1}, Hess_{\rho_t^2}^{-1} F(-gradF(\rho_t^2)) >_{\rho_t^2} \\ &= 2 < \nabla\Phi_t^{1,2}, Hess_{\rho_t^1}^{-1} F(-gradF(\rho_t^1)) - Hess_{\rho_t^2}^{-1} F(-gradF(\rho_t^2)) \circ \nabla\phi_t^{1,2} >_{\rho_t^1} \\ &\le 2KW_2^2(\rho_t^1, \rho_t^2). \end{aligned} \tag{3.2.20}$$

By Gronwall inequality, $W_2^2(\rho_t^1, \rho_t^2) = 0$.

□

Next, let the base space be $\mathbb{R}$. If we consider the Newton flow for the potential functional $F = \int V d\mu$, we will not only give the conditions for uniqueness of the solution to (3.2.9), but also conditions for the uniqueness of Newton flow, i.e. for any initial value, Newton flow converges to the unique minimizer of F.
We consider the absolutely continuous solution ρ to the Newton flow equation for $F = \int V d\mu$.

$$
\begin{cases} \partial_t \rho = -\nabla.(\rho\phi') \\ \displaystyle\int_{\mathbb{R}} f'V''\phi'\rho dx = -\int_{\mathbb{R}} V'\phi' dx, \quad \forall f \in C_c^\infty(\mathbb{R}) \end{cases} \tag{3.2.21}
$$

Because $\mathbb{P}_2(\mathbb{R})$ is flat, $\widetilde{Hess}_\rho F^{-1}(-gradF(\rho)) = -\frac{V'}{V''}$. Thus

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}W_2^2(\rho_t^1, \rho_t^2) \\
&=< \partial_x \Phi_t^{1,2}, \widetilde{Hess}_{\rho_t^1}^{-1} F(-gradF(\rho_t^1)) - \widetilde{Hess}_{\rho_t^2}^{-1} F(-gradF(\rho_t^2)) \circ \partial_x \phi_t^{1,2} >_{\rho_t^1} \\
&= - < \frac{V'}{V''}, \partial_x \Phi_t^{1,2} >_{\rho_t^1} - < \frac{V'}{V''}, \partial_x \Phi_t^{2,1} >_{\rho_t^2} \\
&= - < \frac{V'}{V''} - \frac{V'}{V''} \circ \partial_x \phi_t^{1,2}, \partial_x \Phi_t^{1,2} >_{\rho_t^1} .
\end{aligned}
$$

By mean-value theorem,

$$
\frac{1}{2}\frac{d}{dt}W_2^2(\rho_t^1, \rho_t^2) =< (\frac{-V'}{V''})' \circ \sigma(x)\partial_x \Phi_t^{1,2}, \partial_x \Phi_t^{1,2} >, \tag{3.2.22}
$$

where $\sigma(x)$ is some value in $[x, \partial_x \varphi_t^{1,2}]$. It follows that the absolutely continuous solution is unique if $|(\frac{-V'}{V''})'|$ is bounded. Generally, we have

Theorem 3.2.13. *Assume that $V \in C^3(\mathbb{R})$, $V'' > 0$. Consider the potential functional*

$$
F(\mu) = \int_{\mathbb{R}} V d\mu.
$$

If

$$
|1 - \frac{V'V'''}{(V'')^2}| \leq C, \tag{3.2.23}
$$

Then, There exists a solution to (3.2.21) *.*
If $C \geq 1 - \frac{V'V'''}{(V'')^2} \geq K > 0$, then

$$
W_2^2(\rho_t^1, \rho_t^2) \leq e^{-2Kt} W_2^2(\rho_0^1, \rho_0^2), \tag{3.2.24}
$$

which means, for any initial measure, the absolutely continuous solution to the Newton flow converges to the unique limit point in lifetime.

Proof. Note that $\widetilde{Hess}_\mu F^{-1}(-gradF(\mu)) = -\frac{V'}{V''}$ holds. The assumptions guarantee that $-\frac{V'}{V''}$ is a differentiable vector field, therefore, the following ODE

$$\dot{X} = -\frac{V'}{V''}(X)$$

has a unique solution. $\mu_t = (X_t)_{\#}\mu_0$ is the unique solution to Newton flow equation with initial measure μ_0.
When $1 - \frac{V'V'''}{(V'')^2} \geq K > 0$, (3.2.22) hints

$$\frac{1}{2}\frac{d}{dt}W_2^2(\rho_t^1, \rho_t^2) \leq -KW_2^2(\rho_t^1, \rho_t^2).$$

It follows that $W_2^2(\rho_t^1, \rho_t^2) \leq e^{-2Kt}W_2^2(\rho_0^1, \rho_0^2)$.

□

3.3 Newton flows of several classes of functionals

In this section, we give partial differential equations satisfied by Newton flows of potential functionals, interaction functionals and entropy.
Firstly, there is a natural example which satisfies Assumption 1. For $V, W \in C^\infty$ and $\nabla^2 V \geq \lambda_1$, W being convex, the following functional

$$F(\mu) = \int V d\mu + \int W * \mu d\mu$$

satisfies Assumption 1. We briefly illustrate this. Propositions (3.2.3) and (3.2.5) prove the second and third term in Assumption 1. The first and fourth one in Assumption 1 is already proved in standard textbooks. The last one also can be proved with the representation formula (3.2.2). The corresponding Newton flow equation is

$$\begin{cases} \partial_t \mu = -\nabla.(\mu\nabla\phi) \\ \langle \nabla^2 V \cdot \nabla\phi + \nabla^2 W(x-y)(\nabla\phi(x) - \nabla\phi(y))d\mu(y), \nabla f\rangle_\mu = \langle -\nabla V - \nabla W * \mu, \nabla f\rangle_\mu, \end{cases}$$

$\forall f \in C_c^\infty([0,T] \times \mathbb{T}^d)$.
However, for the functional containing entropy, for example, $F = \int V\rho + \int \rho\log\rho$(for $\rho \in \mathbb{P}_{ac}$), there is no existence of solutions to the Newton flow equation. We can still study its corresponding Newton flow equation in such case.
We consider the following functional:

$$F(\rho) = \begin{cases} \int \rho\log\rho + \int V\rho, & \rho \in \mathbb{P}_{2,ac}(\mathbb{T}^d) \\ +\infty, & otherwise. \end{cases}$$

According to [Vil09], the gradient of F under the Wasserstein metric is

$$gradF = \nabla \log \rho + \nabla V.$$

$HessF$ is

$$Hess_\rho F(\nabla\phi, \nabla\phi) = \int_{\mathbb{T}^d} ||\nabla^2\phi||^2 \rho + \int_{\mathbb{T}^d} \langle \nabla\phi, \nabla^2 V \nabla\phi \rangle \rho.$$

Thus, from aspect of differential equation, we have

Theorem 3.3.1. *For $F(\rho) = \int_{\mathbb{T}^d} \rho \log \rho + \int_{\mathbb{T}^d} V\rho$, the solution to the following equation is the Newton flow of F on $\mathbb{P}_2(\mathbb{T})$:*

$$\begin{cases} \partial_t \rho + \nabla.(\nabla\phi\rho) = 0 \\ \nabla^2 V \nabla\phi - \Delta\nabla\phi - \nabla\log\rho \cdot \nabla^2\phi = -\nabla V - \nabla \log \rho. \end{cases} \tag{3.3.1}$$

Next, we consider Newton flow equations when the base space is a manifold. Generally, for a complete connected compact Riemannian manifold M, let dx be the Riemannian measure on M such that $\int_M dx = 1$. We consider Newton flows of entropy functionals on $\mathbb{P}_2(M)$. According to [Vil09, LL16], Hessian of $E(\rho) = \int_M \rho \log \rho dx$ is

$$Hess_\rho E(\nabla\phi, \nabla\phi) = \int_M \left(||\nabla^2\phi||^2 + \mathrm{Ric}(\nabla\phi, \nabla\phi)\right) \rho dx,$$

where $\rho > 0$. When the base manifold M has a positive Ricci curvature, $Hess_\rho E$ is a positive quadratic form. By theorem 3.2.6, if $Hess_\rho F$ has a Lipschitz property, the solutions to the Newton flow equation exist. We give its corresponding partial differential equations under such case. Denote $\varphi = -\log \rho$, by Bochner's formula, for $\nabla\phi \in \mathbf{T}_\rho$,

$$\widetilde{Hess}_\rho E(\nabla\phi) = -\nabla\Delta\phi + \nabla_{\nabla\varphi}\nabla\phi,$$

and the gradient is

$$gradE(\rho) = \frac{\nabla\rho}{\rho} = \nabla \log \rho.$$

Then we have

Theorem 3.3.2. *For $E(\rho) = \int_M \rho \log \rho$, if M has a positive Ricci curvature, then the solution to the following equation is the Newton flow on $\mathbb{P}_2(M)$ of E*

$$\begin{cases} \partial_t \rho + \nabla.(\rho\nabla\phi) = 0 \\ -\nabla\Delta\phi - \nabla_{\nabla \log \rho} \nabla\phi = -\nabla \log \rho. \end{cases}$$

Remark 3.3.3. *In April 2022, when we were organizing the works in this chapter, we noticed that [LW20] obtained Newton flow equations on $\mathbb{P}_2(\mathbb{R}^d)$, which were similar to some of our results. They formally gave Newton flow equations of relative entropy in Wasserstein spaces, and the convergence rate of the*

Newton's method near the minimum point is analysed.

3.4 Relaxed Newton flow equation and Keller-Segel equation

We consider the following functional:

$$F(\rho) = \begin{cases} \int \rho \log \rho + \int V\rho, & \rho \in \mathbb{P}_{2,ac}(\mathbb{T}^d) \\ +\infty, & otherwise. \end{cases}$$

We will gave the relaxed Newton flow equation. Let $u = \nabla\phi$ and denote $\varphi = -\log\rho$, according to (3.3.1), we give the relaxed Newton flow equation , which no longer requires $u \in \mathbf{T}_\rho$:

$$\begin{cases} \partial_t\rho + \nabla.(\rho u) = 0 \\ \nabla^2 V u - \Delta_\varphi u = -\nabla V + \nabla\varphi, \end{cases} \tag{3.4.1}$$

where $\Delta_\varphi u = \Delta u - \nabla\varphi \cdot \nabla u$. When V is a strictly convex smooth function, then the second equation above has a unique solution, and the operator $\nabla^2 V - \Delta_\varphi$ has a inverse. Then (3.4.1) becomes one equation.

Theorem 3.4.1. *When V is a strictly convex smooth function on $\mathbb{T}^d$, then the solution to the following equation is the Newton flow of F:*

$$\partial_t\rho = -\nabla.\big(\rho(\nabla^2 V - \Delta_\varphi)^{-1}(-\nabla V - \nabla\varphi)\big).$$

Furthermore, according to Bochner formula for 1-form, $\square_\varphi = -\Delta_\varphi + \nabla^2\varphi + \mathrm{Ric}$. We have

$$\begin{cases} \partial_t\rho + \nabla.(\rho u) = 0 \\ \nabla^2(V - \varphi) \cdot u + \square_\varphi u = -\nabla(V - \varphi). \end{cases}$$

When $V = 0$,

$$\begin{cases} \partial_t\rho + \nabla.(\rho u) = 0 \\ -\Delta_\varphi u = \nabla\varphi. \end{cases} \tag{3.4.2}$$

We can see the connection between (3.4.2) and Keller-Segel equation. It is known that Keller-Segel equation is

$$\partial_t\rho = \Delta\rho + \nu\nabla.(\rho\nabla\Delta^{-1}(\rho - 1)). \tag{3.4.3}$$

When the base space is $\mathbb{T}^1$, (3.4.2) becomes

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0 \\ -\partial_x^2 u + \partial_x \varphi \partial_x u = \partial_x \varphi \\ x \in [0,1], u(0) = u(1). \end{cases} \tag{3.4.4}$$

For the second equation, we have

$$\partial_x u(x) = 1 + C\rho(x).$$

In order to make sure u is a function on 1-D Torus , $C = -1$. Thus, $u(x) = x - \int_0^x \rho(s)ds$. The Newton flow equation becomes

$$\partial_t \rho = \partial_x \left(\rho(\int_0^x \rho(s)ds - x) \right).$$

On the other hand, the Keller-Segel equation(take $\nu = 1$) is

$$\partial_t \rho = \partial_x^2 \rho + \partial_x \left(\rho(\int_0^x \rho(s)ds - x) \right).$$

It can be seen as a combination of gradient and Newton flow of entropy functional $S(\rho) = \int_{\mathbb{T}} \rho \log \rho$:

$$\partial_t \rho = -gradS(\rho) + Hess_\rho^{-1} S(-gradS(\rho)).$$

The literature on the Keller-Segel equation is enormous. It is known that in dimensions larger than one, solutions to (3.4.3) can concentrate finite mass in a measure zero region and so blow up in finite time. The well-posedness of (3.4.3) in $d = 2$ and small smooth initial value has been proved by Keller and Segel.

Chapter 4

Geometry and Parallel Transport

In this chapter, based on the Riemannian structure founded by Otto, Sturm, Villani, Lott, etc., we will try to extend the Riemannian geometric computation to a larger probability measure space and larger function space, so that one can introduce parallel translation equation on $\mathbb{P}_2(M)$ as in differential geometry, and study the well-posedness of parallel translation equation.
We will define a formal Riemannian structure on $\mathbb{P}_2(M)$, which is a natural extension of the Riemannian structure on $\mathbb{P}^\infty(M)$ introduced in the former chapter. For the sake of simplicity, we will consider in this paper a connected compact Riemannian manifold M of dimension m. We denote by d_M the Riemannian distance and dx the Riemannian measure on M such that $\int_M dx = 1$. Since the diameter of M is finite, any probability measure μ on M is such that $\int_M d_M^2(x_0, x)\, d\mu(x) < +\infty$, where x_0 is a fixed point of M. As usual, we denote by $\mathbb{P}_2(M)$ the space of probability measures on M, endowed with the Wasserstein distance W_2 defined by

$$W_2^2(\mu_1, \mu_2) = \inf\Big\{\int_{M\times M} d_M^2(x,y)\, \pi(dx, dy), \quad \pi \in \mathcal{C}(\mu_1, \mu_2)\Big\},$$

where $\mathcal{C}(\mu_1, \mu_2)$ is the set of probability measures π on $M \times M$, having μ_1, μ_2 as two marginal laws. It is well known that $\mathbb{P}_2(M)$ endowed with W_2 is a Polish space. In this compact case, the weak convergence for probability measures is metrized by W_2; therefore $(\mathbb{P}_2(M), W_2)$ is a compact Polish space.

4.1 Tangent space of $\mathbb{P}_2(M)$

The introduction of tangent spaces of $\mathbb{P}_2(M)$ can go back to the early work [OV00], as well as in [Ott01]. A more rigorous treatment was given in [AGS05]. In differential geometry, for a smooth curve $\{c(t);\ t \in [0,1]\}$ on a manifold M, the derivative $c'(t)$ with respect to the time t is in the tangent space : $c'(t) \in \mathbf{T}_{c(t)}M$. A classical result says that for an absolutely continuous curve $\{c(t);\ t \in [0,1]\}$ on M, the derivative $c'(t) \in T_{c(t)}M$ exists for almost all $t \in [0,1]$. Following [AGS05], we say that a curve $\{c(t);\ t \in [0,1]\}$ on $\mathbb{P}_2(M)$ is absolutely continuous in L^2 if there exists $k \in L^2([0,1])$ such that

$$W_2(c(t_1), c(t_2)) \le \int_{t_1}^{t_2} k(s)\, ds, \quad t_1 < t_2.$$

The following result is our starting point:

Theorem 4.1.1 (see [AGS05], Theorem 8.3.1)**.** *Let* $\{c_t;\ t \in [0,1]\}$ *be an absolutely continuous curve on* $\mathbb{P}_2(M)$ *in* L^2*, then there exists a Borel vector field* Z_t *on* M *such that*

$$\int_{[0,1]} \Big[\int_M |Z_t(x)|^2_{\mathbf{T}_x M}\, c_t(dx) \Big]\, dt < +\infty$$

and the following continuity equation

$$\frac{dc_t}{dt} + \nabla \cdot (Z_t\, c_t) = 0, \tag{4.1.1}$$

holds in the sense of distribution. Uniqueness to (4.1.1) *holds if moreover* Z_t *is imposed to be in*

$$\overline{\big\{\nabla\psi,\ \psi \in C^\infty(M)\big\}}^{L^2(c_t)}.$$

Then, we can define the tangent space $\mathbf{T}_\mu$ of $\mathbb{P}_2(M)$ at μ by

$$\mathbf{T}_\mu = \overline{\big\{\nabla\psi,\ \psi \in C^\infty(M)\big\}}^{L^2(\mu)}, \tag{4.1.2}$$

the closure of gradients of smooth functions in the space $L^2(\mu)$. Note that here we use the definition of tangent space in [AGS05]. It is isomorphic to the tangent space introduced in Chapter 2, which is the original definition given by Otto. Equation (4.1.1) implies that for almost all $t \in [0,1]$,

$$\frac{d}{dt}\int_M f(x)\, c_t(dx) = \int_M \langle \nabla f(x), Z_t(x)\rangle_{\mathbf{T}_x M}\, c_t(dx), \quad f \in C^1(M). \tag{4.1.3}$$

We will say that Z_t is the intrinsic derivative of c_t and use the notation

$$\frac{d^I c_t}{dt} = Z_t \in \mathbf{T}_{c_t}.$$

In what follows, we will describe the tangent space $\mathbf{T}_\mu$ with the least conditions as possible on the measure μ. Consider the quadratic form defined by

$$\mathcal{E}(\psi) = \int_M |\nabla\psi(x)|^2\, d\mu(x), \quad \psi \in C^1(M).$$

We assume that there is a constant $C_\mu > 0$ such that

$$\int_M (\psi - \langle\psi\rangle)^2\, d\mu \leq C_\mu \int_M |\nabla\psi|^2\, d\mu, \tag{4.1.4}$$

where $\langle\psi\rangle = \int_M \psi(x)\, dx$. The condition (4.1.4) is satisfied if μ admits a positive continuous density $\rho > 0$: $d\mu = \rho\, dx$. In fact, let

$$\beta_1 = \inf_{x\in M} \rho(x) > 0, \quad \beta_2 = \sup_{x\in M} \rho(x) < +\infty.$$

Since M is compact, the following Poincaré inequality holds :

$$\int_M (\psi - \langle\psi\rangle)^2\, dx \leq C \int_M |\nabla\psi|^2\, dx,$$

then

$$\int_M (\psi - \langle\psi\rangle)^2\, d\mu \leq \frac{C\beta_2}{\beta_1} \int_M |\nabla\psi|^2\, d\mu.$$

Remark that Inequality (4.1.4) is not Poincaré inequality, since the mean $\langle\psi\rangle$ is not taken with respect to the measure μ, but to dx.

Now let $Z \in \mathbf{T}_\mu$; there is a sequence of functions $\psi_n \in C^\infty(M)$ such that $Z = \lim_{n\to+\infty} \nabla\psi_n$ in $L^2(\mu)$. By changing ψ_n to $\psi_n - \langle\psi_n\rangle$ and by condition (4.1.4), $\{\psi_n;\ n \geq 1\}$ is a Cauchy sequence in $L^2(\mu)$. If the quadratic form $\mathcal{E}(\psi)$ is closable in $L^2(\mu)$, then there exists a function φ_μ in the Sobolev space $\mathbb{D}_1^2(\mu)$ such that $Z = \nabla\varphi_\mu$, where $\mathbb{D}_1^2(\mu)$ is the closure of $C^\infty(M)$ with respect to the norm

$$||\varphi||^2_{\mathbb{D}_1^2(\mu)} := \int_M |\varphi(x)|^2\, d\mu(x) + \int_M |\nabla\varphi(x)|^2\, d\mu(x).$$

A sufficient condition to ensure the closability for $\mathcal{E}$ is that the formula of integration by parts holds for μ; more precisely, for any C^1 vector field Z on M, there exists a function denoted by $\mathrm{div}_\mu(Z) \in L^2(\mu)$ such that

$$\int_M \langle\nabla f(x), Z(x)\rangle_{\mathbf{T}_x M}\, d\mu(x) = -\int_M f(x)\, \mathrm{div}_\mu(Z)(x)\, d\mu(x), \quad f \in C^1(M). \tag{4.1.5}$$

Definition 4.1.2. *We say that a probability measure μ has divergence if $\mathrm{div}_\mu(Z) \in L^2(\mu)$ exists for all C^1-vector field Z on M . We will use the notation*

$$\mathbb{P}_{\mathsf{div}}(M)$$

to denote the set of probability measures on M having strictly positive continuous density and satisfying conditions (4.1.5).

For example, if $d\mu(x) = \rho(x)\, dx$ for some strictly positive continuous density $\rho \in \mathbb{D}_1^2(dx)$, then $\mu \in \mathbb{P}_{\mathsf{div}}(M)$.

Proposition 4.1.3. *For a measure $\mu \in \mathbb{P}_{\mathsf{div}}(M)$, we have*

$$\mathbf{T}_\mu = \{\nabla\psi;\ \psi \in \mathbb{D}_1^2(\mu)\}.$$

Note that this result is not new, see for example [LL16, LL18]. Here we indicate what are necessary conditions which yield to this result.

The inconvenient for (4.1.3) is the existence of derivative for almost all $t \in [0,1]$. In what follows, we will present two typical classes of absolutely continuous curves in $\mathbb{P}_2(M)$.

4.1.1 Constant vector fields on $\mathbb{P}_2(M)$

For any gradient vector field $\nabla\psi$ on M with $\psi \in C^\infty(M)$, consider the ordinary differential equation (ODE):

$$\frac{d}{dt}U_t(x) = \nabla\psi(U_t(x)), \quad U_0(x) = x \in M.$$

Then $x \to U_t(x)$ is a flow of diffeomorphisms on M. Let $\mu \in \mathbb{P}_2(M)$, consider $c_t = (U_t)_{\#}\mu$. It is easy to see that the curve $\{c_t;\ t \in [0,1]\}$ is absolutely continuous in L^2 and for $f \in C^1(M)$,

$$\frac{d}{dt}\int_M f(x)\, c_t(dx) = \frac{d}{dt}\int_M f(U_t(x))\, d\mu(x) = \int_M \langle \nabla f(U_t(x)), \nabla\psi(U_t(x))\rangle\, d\mu(x),$$

which is equal to, for any $t \in [0,1]$,

$$\int_M \langle \nabla f, \nabla\psi\rangle\, c_t(dx).$$

In other term, c_t is a solution to the following continuity equation:

$$\frac{dc_t}{dt} + \nabla\cdot(\nabla\psi\, c_t) = 0.$$

According to above definition, we see that for each $t \in [0,1]$,

$$\frac{d^I c_t}{dt} = \nabla\psi.$$

It is why we call $\nabla\psi$ a constant vector field on $\mathbb{P}_2(M)$. In order to make clearly different roles played by $\nabla\psi$, we will use notation

$$V_\psi$$

when it is seen as a constant vector field on $\mathbb{P}_2(M)$.

Remark 4.1.4. In section 4.3 below, we will compute Lie brackets of two constant vector fields on $\mathbb{P}_2(M)$ without explicitly using the existence of density of measure, the Lie bracket of two constant vector fields is NOT a constant vector field.

4.1.2 Geodesics with constant speed

It is easy to introduce geodesics with constant speed when the base space is a flat space $\mathbb{R}^m$. A probability measure μ on $\mathbb{R}^m$ is in $\mathbb{P}_2(\mathbb{R}^m)$ if $\int_{\mathbb{R}^m} |x|^2\, d\mu(x) < +\infty$. Let $c_0, c_1 \in \mathbb{P}_2(\mathbb{R}^m)$, there is an optimal coupling plan $\gamma \in \mathcal{C}(c_0, c_1)$ such that

$$W_2^2(c_0, c_1) = \int_{\mathbb{R}^m \times \mathbb{R}^m} |x-y|^2\, d\gamma(x,y).$$

For each $t \in [0,1]$, define $c_t \in \mathbb{P}_2(\mathbb{R}^m)$ by

$$\int_{\mathbb{R}^m} f(x)\, dc_t(x) = \int_{\mathbb{R}^m \times \mathbb{R}^m} f(u_t(x,y))\, d\gamma(x,y),$$

where $u_t(x,y) = (1-t)x + ty$. For $0 \le s < t \le 1$, define $\pi_{s,t} \in \mathcal{C}(c_s, c_t)$ by

$$\int_{\mathbb{R}^m \times \mathbb{R}^m} g(x,y)\, \pi_{s,t}(dx,dy) = \int_{\mathbb{R}^m \times \mathbb{R}^m} g(u_s(x,y), u_t(x,y))\, d\gamma(x,y).$$

Then

$$W_2^2(c_s, c_t) \le \int_{\mathbb{R}^m \times \mathbb{R}^m} |u_t(x,y) - u_s(x,y|^2\, d\gamma(x,y) = (t-s)^2 W_2(c_0,c_1)^2.$$

It follows that $W_2(c_s, c_t) \le (t-s) W_2(c_0, c_1)$. Combing with triangulaire inequality,

$$\begin{aligned} W_2(c_0,c_1) &\le W_2(c_0,c_s) + W_2(c_s,c_t) + W_2(c_t,c_1) \\ &\le sW_2(c_0,c_1) + (t-s)W_2(c_0,c_1) + (1-t)W_2(c_0,c_1) = W_2(c_0,c_1), \end{aligned}$$

we get the property of geodesic with constant speed:

$$W_2(c_s, c_t) = |t-s|\, W_2(c_0, c_1).$$

According to Theorem 4.1.1, there is $Z_t \in \mathbf{T}_{c_t}$ such that, for $f \in C_c^1(\mathbb{R}^d)$,

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^m} f(x) c_t(dx) &= \int_{\mathbb{R}^m} \langle \nabla f(u_t(x,y)), y - x \rangle_{\mathbb{R}^m}\, d\gamma(x,y) \\ &= \int_{\mathbb{R}^d} \langle \nabla f(x), Z_t(x) \rangle_{\mathbb{R}^m}\, c_t(dx) \end{aligned}$$

where $\langle\ ,\ \rangle_{\mathbb{R}^m}$ is the canonical inner product of $\mathbb{R}^m$. We heuristically look for Z_t such that $Z_t(u_t(x,y)) = y - x$. Taking the derivative with respect to t yields

$$(\frac{d}{dt}Z_t)(u_t(x,y)) + \langle \nabla Z_t(u_t(x,y)), y-x\rangle = 0.$$

It follows that

$$(\frac{d}{dt}Z_t) + \nabla Z_t(Z_t) = 0.$$

In the case where $Z_t = \nabla\psi_t$, we have

$$(\frac{d}{dt}\nabla\psi_t) + \nabla^2\psi_t(\nabla\psi_t) = 0.$$

We remark that $\{\nabla\psi_t, t\in]0,1[\}$ satisfies heuristically the equation of Riemannian geodesic obtained in [Lot06] or heuristically obtained in [OV00], in which the authors showed that the convexity of entropy functional along these geodesics is equivalent to Bakry-Emery's curvature condition [BÉ85] (see also [vRS05, Stu06]).

In the case of Riemannian manifold M, it is a bit complicated. We follow the exposition of [Gig11]. Let $\mathbf{T}M$ be the tangent bundle of M and $\pi : \mathbf{T}M \to M$ the natural projection. For each $\mu\in\mathbb{P}(M)$, we consider the set

$$\Gamma_\mu = \Big\{\gamma \text{ probability measure on } \mathbf{T}M;\ \pi_\#\gamma = \mu, \int_{\mathbf{T}M} |v|^2_{\mathbf{T}_xM} d\gamma(x,v) < +\infty\Big\}.$$

The set Γ_μ is obviously non empty. Let $\gamma\in\Gamma_\mu$, we consider $\nu = \exp_\#\gamma$, that is,

$$\int_M f(x)d\nu(x) = \int_{TM} f(\exp_x(v))\, d\gamma(x,v),$$

where $\exp_x : \mathbf{T}_xM \to M$ is the exponential map induced by geodesics on M. The map

$$\mathbf{T}M \to M\times M, \quad (x,v)\to (x, \exp_x(v))$$

sends γ to a coupling plan $\tilde\gamma\in\mathcal{C}(\mu,\nu)$. We have

$$W_2^2(\mu,\nu) \le \int_{\mathbf{T}M} d_M^2(x,\exp_x(v))\, d\gamma(x,v) \le \int_{\mathbf{T}M} |v|^2_{\mathbf{T}_xM}\, d\gamma(x,v).$$

In order to construct geodesics $\{c_t; t\in[0,1]\}$ connecting μ and ν, we need to find $\gamma_0\in\Gamma_\mu$ such that $\nu = \exp_\#\gamma_0$ and

$$W_2^2(\mu,\nu) = \int_{\mathbf{T}M} |v|^2_{\mathbf{T}_xM}\, d\gamma_0(x,v). \tag{4.1.6}$$

As M is connected, let $x\in M$, for each y, there is a minimizing geodesic $\{\xi(t),\ t\in[0,1]\}$ connecting x and y. Let $v_{x,y} = \xi'(0)\in\mathbf{T}_xM$, then

$$y = \exp_x(v_{x,y}) \text{ and } d_M(x,y) = |v_{x,y}|_{\mathbf{T}_x M}.$$

Take a Borel version Ξ of such a map $(x,y) \to (x, v_{x,y})$ from $M \times M$ to TM. Let $\tilde{\gamma}_0 \in \mathcal{C}(\mu,\nu)$ be an optimal coupling plan; define $\gamma_0 \in \Gamma_\mu$ by

$$\int_{\mathbf{T}M} g(x,v)\, d\gamma_0(x,v) = \int_{M\times M} g\big(x, \Xi(x,y)\big)\, d\tilde{\gamma}_0(x,y).$$

Therefore

$$\begin{aligned}\int_{\mathbf{T}M} |v|^2_{\mathbf{T}_x M}\, d\gamma_0(x,v) &= \int_{M\times M} |\Xi(x,y)|^2\, d\tilde{\gamma}_0(x,y)\\ &= \int_{M\times M} d_M(x,y)^2\, d\tilde{\gamma}_0(x,y) = W_2^2(\mu,\nu).\end{aligned}$$

Now we define the curve $\{c_t;\ t \in [0,1]\}$ on $\mathbb{P}_2(M)$ by

$$\int_M f(x) c_t(dx) = \int_{\mathbf{T}M} f(\exp_x(tv))\, d\gamma_0(x,v).$$

Similarly we check that

$$W_2(c_s, c_t) = |t-s|\, W_2(c_0, c_1).$$

The organization of this chapter is as follows. In Section 4.2, we consider ordinary equations on $\mathbb{P}_2(M)$, a Cauchy-Peano's type theorem is established, also Mckean-Vlasov equation involved. In Section 4.3, we emphasize that the suitable class of probability measures for developing the differential geometry is one having divergence and the strictly positive density with certain regularity. The Levi-Civita connection is introduced and the formula for the covariant derivative of a general but smooth enough vector field is obtained. In section 4.4, we precise result on the derivability of the Wasserstein distance on $\mathbb{P}_2(M)$, which enable us to obtain the extension of a vector field along a quite good curve on $\mathbb{P}_2(M)$ in Section 4.5 as in differentiable geometry; the parallel translation along such a good curve on $\mathbb{P}_2(M)$ is naturally and rigorously introduced. And we give the well-posedness results of parallel translation on $\mathbb{P}_2(\mathbb{T})$. In the last section 4.5.1, we give the Lipschitz condition for vector fields and the uniqueness of the solution to ODE.

4.2 Ordinary differential equations on $\mathbb{P}_2(M)$

Let $\varphi \in C^1(M)$, consider the function F_φ on $\mathbb{P}_2(M)$ defined by

$$F_\varphi(\mu) = \int_M \varphi(x)\, d\mu(x). \tag{4.2.1}$$

A function F on $\mathbb{P}_2(M)$ is said to be a polynomial if it is an element of the algebra spanned by all the functions $F = F_{\varphi_1} \cdots F_{\varphi_k}$, where $\varphi_1, \ldots, \varphi_k$ are finite number of functions in $C^1(M)$. Let $Z = V_\psi$ be a constant vector field on $\mathbb{P}_2(M)$ with $\psi \in C^\infty(M)$, and U_t the flow on M associated to $\nabla\psi$. For $\mu_0 \in \mathbb{P}_2(M)$, we set $\mu_t = (U_t)_\# \mu_0$. Then we have seen in section 4.1.1,

$$\Big\{\frac{d}{dt} F_\varphi(\mu_t)\Big\}_{|t=0} = \int_M \langle \nabla\varphi(x), \nabla\psi(x)\rangle\, d\mu_0(x) = \langle V_\varphi, V_\psi\rangle_{\mathbf{T}_{\mu_0}}.$$

The left hand side of above equality is the derivative of F_φ along V_ψ. More generally, for a function F on $\mathbb{P}_2(M)$, we say that F is derivable at μ_0 along V_ψ, if

$$(\bar{D}_{V_\psi} F)(\mu_0) = \Big\{\frac{d}{dt} F(\mu_t)\Big\}_{|t=0} \quad \text{exists.}$$

We say that the gradient $\bar\nabla F(\mu_0) \in \mathbf{T}_{\mu_0}$ exists if for each $\psi \in C^\infty(M)$, $(\bar{D}_{V_\psi} F)(\mu_0)$ exists and

$$\bar{D}_{V_\psi} F(\mu_0) = \langle \bar\nabla F, V_\psi\rangle_{\mathbf{T}_{\mu_0}}. \tag{4.2.2}$$

Note that for $\varphi \in C^1(M)$, there is a sequence of $\psi_n \in C^\infty(M)$ such that $\nabla\psi_n$ converge uniformly to $\nabla\varphi$ so that $V_\varphi \in \mathbf{T}_\mu$ for any $\mu \in \mathbb{P}_2(M)$. It is obvious that $\bar\nabla F_\varphi = V_\varphi$. For the polynomial $F = \prod_{i=1}^k F_{\varphi_i}$, we have

$$\bar\nabla F = \sum_{i=1}^k \Big(\prod_{j\neq i} F_{\varphi_j}\Big)\, V_{\varphi_i}.$$

Note that the family $\{F_\varphi,\ \varphi \in C^1(M)\}$ separates the point of $\mathbb{P}_2(M)$. By Stone-Weierstrauss theorem, the space of polynomials is dense in the space of continuous functions on $\mathbb{P}_2(M)$.

Remark 4.2.1. If the gradient $\nabla\psi$ is replaced by a general C^1-vector field on M, the above definition is also well-settled; in fact this has been done in the early work [AKR96] in another context for other applications. The links among different type of derivatives are recently characterized in [RW21].

Remark 4.2.2. The definition of gradient $gradF$ defined in the former chapter is actually an extension of $\bar\nabla F$. Note that $\bar\nabla F$ is defined by smooth constant fields $\nabla\varphi \in \mathbf{T}_\mu$, $\varphi \in C_c^\infty(M)$, while $gradF$ is defined by $u \in \mathbf{T}_\mu$. The test function space is different. If $gradF$ is well defined, $gradF$ must equal to $\bar\nabla F$. However, if $\bar\nabla F$ is well defined, even if the operator $A_\mu(u) = \langle \bar\nabla F(\mu), u\rangle_{L^2(\mu)}$ is closable in $\mathbf{T}_\mu$, $gradF(\mu)$ may still not exist, not to mention that $gradF = \bar\nabla F$. We give an example to illustrate the difference.
Consider

$$F(\rho) = \begin{cases} F = \int \rho \log \rho + \int V\rho, & \rho \in \mathbb{P}_{2,ac}([0,1]) \\ +\infty, & \text{otherwise.} \end{cases}$$

Suppose that $\rho_0 = \mathbb{1}_{[0,1]}dx$, $u \in L^2(\rho_0)$

$$u(x) = 2^{n+2} - 3 - 2^{n+2}x, \quad x \in [1 - \frac{1}{2^n}, 1 - \frac{1}{2^{n+1}}),\ n = 0, 1, \dots$$

It is obvious to see $F((Id + \frac{1}{2^{k+1}}u)_{\#}\rho_0) = +\infty$ for every $k \in \mathbb{N}$. Thus,

$$\lim_{\epsilon \to 0^+} \frac{F((Id + \epsilon u)_{\#}\rho_0) - F(\rho_0)}{\epsilon}$$

does not exist. It means $gradF(\rho_0)$ does not exist, while $\bar{\nabla}F(\rho_0) = 0$.
In this chapter and later chapters, we use $\bar{\nabla}F$ to represent the gradient of functional F. Similarly, we will use $\bar{\nabla}^2 F$ rather than $HessF$ to represent the Hessian operator in the later chapters.

We will use ∇ to denote the gradient operator on the base space M, and $\bar{\nabla}$ to denote the gradient operator on the Wasserstein space $(\mathbb{P}_2(M), W_2)$. For example, if $(\mu, x) \to \Phi(\mu, x)$ is a function on $\mathbb{P}_2(M) \times M$, then $\nabla\Phi(\mu, x)$ is the gradient with respect to x, while $\bar{\nabla}\Phi(\mu, x)$ is the gradient with respect to μ.

Definition 4.2.3. *We will say that Z is a vector field on $\mathbb{P}_2(M)$ if there exists a Borel map $\Phi : \mathbb{P}_2(M) \times M \to \mathbb{R}$ such that for any $\mu \in \mathbb{P}_2(M)$, $x \to \Phi(\mu, x)$ is C^1 and $Z(\mu) = V_{\Phi(\mu,\cdot)}$.*

A class of test vector fields on $\mathbb{P}_2(M)$ is

$$\chi(\mathbb{P}) = \Big\{ \sum_{finite} \alpha_i V_{\psi_i}, \quad \alpha_i \text{ polynomial}, \ \psi_i \in C^\infty(M) \Big\}. \tag{4.2.3}$$

Let Z be a vector field on $\mathbb{P}_2(M)$, how to construct a solution $\mu_t \in \mathbb{P}_2(M)$ to the following ODE

$$\frac{d^I \mu_t}{dt} = Z(\mu_t)?$$

Theorem 4.2.4. *Let Z be a vector field on $\mathbb{P}_2(M)$ given by Φ. Assume that $(\mu, x) \to \nabla\Phi(\mu, x)$ is continuous, then for any $\mu_0 \in \mathbb{P}_2(M)$, there is an absolutely continuous curve $\{\mu_t;\ t \in [0,1]\}$ on $\mathbb{P}_2(M)$ such that*

$$\frac{d^I \mu_t}{dt} = Z(\mu_t), \quad \mu_{|t=0} = \mu_0. \tag{4.2.4}$$

If moreover, for any $\mu \in \mathbb{P}_2(M)$, $x \to \Phi(\mu, x)$ is C^2 and

$$C_2 := \sup_{\mu\in\mathbb{P}_2(M)} \sup_{x\in M} ||\nabla^2\Phi(\mu,x)|| < +\infty, \tag{4.2.5}$$

then there is a flow of continuous maps $(t,x)\to U_t(x)$ on M, solution to the following Mckean-Vlasov equation

$$\frac{d}{dt}U_t(x) = \nabla\Phi(\mu_t, U_t(x)), \quad \mu_t = (U_t)_{\#}\mu_0. \tag{4.2.6}$$

Proof. We use the Euler approximation to construct a solution. We first note that

$$C_1 := \sup_{(\mu,x)\in\mathbb{P}_2(M)\times M} |\nabla\Phi(\mu,x)| < +\infty. \tag{4.2.7}$$

Let $P_t = e^{t\Delta_M}$ be the heat semi-group associated to the Laplace operator Δ_M on functions, and $T_t = e^{-t\Box}$ the heat semigroup on differential forms, with de Rham-Hodge operator $\Box$. It is well-known that

$$|T_t(\nabla\varphi)| \le e^{-t\kappa/2}P_t|\nabla\varphi|, \quad \varphi\in C^1(M)$$

where κ is lower bound of Ricci tensor on M. Here, $\nabla\varphi$ can be identified by 1-form $d\varphi$. As $t\to 0$, $T_t(\nabla\varphi)$ converges to $\nabla\varphi$ uniformly. For $n\ge 1$, let

$$Z_n(\mu,x) = \big(T_{1/n}\nabla\Phi(\mu,\cdot)\big)(x).$$

According to (4.2.7) and above estimate, for n big enough,

$$\sup_{(\mu,x)\in\mathbb{P}_2(M)\times M} |Z_n(\mu,x)| \le 2C_1. \tag{4.2.8}$$

Now let $t_k = k2^{-n}$ for $k = 1,\dots,2^n$ and

$$[t] = t_k \quad \text{if } t\in[t_k, t_{k+1}[.$$

On the intervall $[t_0,t_1]$, consider the ODE on M:

$$\frac{dU_t^{(n)}}{dt} = Z_n\big(\mu_0, U_t^{(n)}\big), \quad U_0^{(n)}(x) = x, \tag{4.2.9}$$

and $\mu_t^{(n)} = (U_t^{(n)})_{\#}\mu_0$ for $t\in[t_0,t_1]$; inductively, on $[t_k,t_{k+1}]$, we consider

$$\frac{dU_t^{(n)}}{dt} = Z_n\big(\mu_{t_k}^{(n)}, U_t^{(n)}\big), \quad U^{(n)}_{|_{t=t_k}}(x) = U_{t_k}^{(n)}(x), \tag{4.2.10}$$

and for $t \in [t_k, t_{k+1}]$,

$$\mu_t^{(n)} = (U_t^{(n)})_{\#}\mu_{t_k}^{(n)} \tag{4.2.11}$$

and so on, we get a curve $\{\mu_t^{(n)};\ t \in [0,1]\}$ on $\mathbb{P}_2(M)$. We now prove that this family is equicontinuous in $C([0,1], \mathbb{P}_2(M))$. Let $0 \le s < t \le 1$, define $\gamma(\theta) = U_{(1-\theta)s+\theta t}^{(n)}$, then

$$\frac{d\gamma(\theta)}{d\theta} = (t-s) Z_n\big(\mu_{[(1-\theta)s+\theta t]}^{(n)}, U_{(1-\theta)s+\theta t}^{(n)}\big).$$

We have, according to (4.2.8),

$$d_M\big(U_t^{(n)}(x), U_s^{(n)}(x)\big) \le \int_0^1 \Big|\frac{d\gamma(\theta)}{d\theta}\Big|\, d\theta \le 2C_1(t-s).$$

Define a probability measure π on $M \times M$ by

$$\int_{M\times M} g(x,y)\pi(dx,dy) = \int_M g\big(U_t^{(n)}(x), U_s^{(n)}(x)\big)\, d\mu_0(x).$$

Then $\pi \in \mathcal{C}(\mu_t^{(n)}, \mu_s^{(n)})$, we have

$$W_2^2\big(\mu_t^{(n)}, \mu_s^{(n)}\big) \le \int_M d_M^2\big(U_t^{(n)}(x), U_s^{(n)}(x)\big)\, d\mu_0(x) \le 4C_1^2\,(t-s)^2.$$

By Ascoli theorem, up to a subsequence, $\mu_\cdot^{(n)}$ converges in $C([0,1], \mathbb{P}_2(M))$ to a continuous curve $\{\mu_t;\ t \in [0,1]\}$ such that $W_2(\mu_t, \mu_s) \le 2C_1\,(t-s)$.

For proving that $\{\mu_t;\ t \in [0,1]\}$ is a solution to ODE (4.2.4), we need the following preparation:

Lemma 4.2.5. *Set $\Phi_\mu(x) = \Phi(\mu, x)$, then*

$$\sup_{(\mu,x)\in\mathbb{P}_2(M)\times M} |(T_t\nabla\Phi_\mu)(x) - \nabla\Phi(x)|_{\mathbf{T}_x M} \to 0, \quad \text{as } t \to 0. \tag{4.2.12}$$

Proof. We use $||\cdot||_\infty$ to denote the uniform norm on M. Let $\varepsilon > 0$, for $\mu \in \mathbb{P}_2(M)$, there is $\hat{t}_\mu > 0$ such that

$$\sup_{t\le\hat{t}_\mu} ||T_t\nabla\Phi_\mu - \nabla\Phi_\mu||_\infty < \varepsilon.$$

Since $(\mu, t) \to ||T_t\nabla\Phi_\mu - \nabla\Phi_\mu||_\infty$ is continuous, there is $\delta_\mu > 0$ such that for $t \le \hat{t}_\mu$,

$$W_2(\mu,\nu) < \delta_\mu \ \Rightarrow\ ||T_t\nabla\Phi_\nu - \nabla\Phi_\nu||_\infty < \varepsilon.$$

Let $B(\mu,\delta)$ be the open ball in $(\mathbb{P}_2(M), W_2)$ centered at μ, of radius δ. We have

$$\mathbb{P}_2(M) = \cup_{\mu\in\mathbb{P}_2(M)} B(\mu, \delta_\mu);$$

so there is a finite number of $\{\mu_1, \ldots, \mu_K\}$ such that

$$\mathbb{P}_2(M) = \cup_{i=1}^K B(\mu_i, \delta_{\mu_i}).$$

Let $\hat{t} = \min\{\hat{t}_{\mu_i},\ i = 1, \ldots, K\} > 0$. Then for $0 < t < \hat{t}$,

$$\sup_{\mu\in\mathbb{P}_2(M)} ||T_t \nabla\Phi_\mu - \nabla\Phi_\mu||_\infty \le \varepsilon.$$

So we get (4.2.12). $\square$

End of the proof of theorem : $\{\mu_t^{(n)};\ t \in [0,1]\}$ satisfies the following continuity equation

$$\begin{aligned} &- \int_{[0,1]\times M} \alpha'(t) f(x) \mu_t^{(n)}(dx)dt \\ &= \alpha(0) \int_M f(x) d\mu_0(x) + \int_{[0,1]\times M} \alpha(t) \left\langle \nabla f(x), Z_n\big(\mu_{[t]}^{(n)}, x\big)\right\rangle \mu_t^{(n)}(dx)dt, \end{aligned} \tag{4.2.13}$$

for all $\alpha \in C_c^1([0,1))$ and $f \in C^1(M)$. We have

$$\begin{aligned} &\int_{[0,1]\times M} \alpha(t) \left\langle \nabla f(x), Z_n\big(\mu_{[t]}^{(n)}, x\big)\right\rangle \mu_t^{(n)}(dx)dt - \int_{[0,1]\times M} \alpha(t) \left\langle \nabla f(x), \nabla\Phi\big(\mu_t, x\big)\right\rangle \mu_t(dx)dt \\ &= \int_{[0,1]\times M} \alpha(t) \left\langle \nabla f(x), Z_n\big(\mu_{[t]}^{(n)}, x\big) - \nabla\Phi(\mu_t, x)\right\rangle \mu_t^{(n)}(dx)dt \\ &+ \int_{[0,1]\times M} \alpha(t) \left\langle \nabla f(x), \nabla\Phi\big(\mu_t, x\big)\right\rangle \mu_t^{(n)}(dx)dt - \int_{[0,1]\times M} \alpha(t) \left\langle \nabla f(x), \nabla\Phi\big(\mu_t, x\big)\right\rangle \mu_t(dx)dt. \end{aligned}$$

It is obvious that the sum of two last terms converge to 0 as $n \to +\infty$. Let I_n be the first term on the right side, then

$$|I_n| \le ||\nabla f||_\infty \int_0^1 |\alpha(t)|\, ||T_{1/n} \nabla\Phi_{\mu_{[t]}^{(n)}} - \nabla\Phi_{\mu_t}||_\infty\, dt$$

Note that

$$||T_{1/n} \nabla\Phi_{\mu_{[t]}^{(n)}} - \nabla\Phi_{\mu_t}||_\infty \le ||T_{1/n} \nabla\Phi_{\mu_{[t]}^{(n)}} - \nabla\Phi_{\mu_{[t]}^{(n)}}||_\infty + ||\nabla\Phi_{\mu_{[t]}^{(n)}} - \nabla\Phi_{\mu_t}||_\infty.$$

The term $||T_{1/n}\nabla\Phi_{\mu^{(n)}_{[t]}} - \nabla\Phi_{\mu^{(n)}_{[t]}}||_\infty \to 0$ is due to above lemma. As $n \to +\infty$, $\mu^{(n)}_{[t]}$ converges to μ_t. By continuity of $(\mu, x) \to \nabla\Phi(\mu, x)$, the last term tends to 0. Letting $n \to +\infty$ in (4.2.13) yields

$$\begin{aligned} &- \int_{[0,1]\times M} \alpha'(t) f(x)\mu_t(dx)dt \\ &= \alpha(0) \int_M f(x)\mu_0(dx) + \int_{[0,1]\times M} \alpha(t) \left\langle \nabla f(x), \nabla\Phi\big(\mu_t, x\big)\right\rangle \mu_t(dx)dt, \end{aligned}$$

which is the meaning of Equation (4.2.4) in distribution sense.

For the proof of second part, since $x \to \Phi(\mu, x)$ is C^2, we can directly use $\nabla\Phi(\mu, \cdot)$ instead of Z_n in (4.2.9), (4.2.10), (4.2.11).

On the intervall $[t_0, t_1]$, consider the ODE on M:

$$\frac{dU_t^{(n)}}{dt} = \nabla\Phi\big(\mu_0, U_t^{(n)}\big), \quad U_0^{(n)}(x) = x, \tag{4.2.14}$$

and $\mu_t^{(n)} = (U_t^{(n)})_\# \mu_0$ for $t \in [t_0, t_1]$; inductively, on $[t_k, t_{k+1}]$, we consider

$$\frac{dU_t^{(n)}}{dt} = \nabla\Phi\big(\mu_{t_k}^{(n)}, U_t^{(n)}\big), \quad U_{|t=t_k}^{(n)}(x) = U_{t_k}^{(n)}(x), \tag{4.2.15}$$

and for $t \in [t_k, t_{k+1}]$,

$$\mu_t^{(n)} = (U_t^{(n)})_\# \mu_{t_k}^{(n)}. \tag{4.2.16}$$

By above result, up to a subsequence, $\{\mu_t^{(n)},\ t \in [0,1]\}$ converges to $\{\mu_t, t \in [0,1]\}$ in $C([0,1], \mathbb{P}_2(M))$. We use this subsequence to prove the convergence of $\{U_t^{(n)}(x),\ t \in [0,1]\}$. Now we prove that, under Condition (4.2.7),

$$d_M\Big(U_t^{(n)}(x), U_t^{(n)}(y)\Big) \le e^{C_2 t}\, d_M(x,y), \quad x, y \in M. \tag{4.2.17}$$

For $x, y \in M$ given, there is a minimizing geodesic $\{\xi_s,\ s \in [0,1]\}$ connecting x and y such that $d_M(x,y) = \int_0^1 |\xi_s'|\, ds$. Set

$$\sigma(t,s) = U_t^{(n)}(\xi_s).$$

Since the torsion is free, we have the relation:

$$\frac{D}{ds}\frac{d}{dt}\sigma(t,s) = \frac{D}{dt}\frac{d}{ds}\sigma(t,s), \tag{4.2.18}$$

where $\frac{D}{ds}$ denotes the covariant derivative. We have

$$\frac{d}{dt}U_t^{(n)}(\xi_s) = \nabla\Phi\Big(\mu_{[t]}^{(n)}, U_t^{(n)}(\xi_s)\Big).$$

Taking the derivative with respect to s, we get

$$\frac{D}{ds}\frac{d}{dt}U_t^{(n)}(\xi_s) = \nabla^2\Phi\Big(\mu_{[t]}^{(n)}, U_t^{(n)}(\xi_s)\Big)\cdot\frac{d}{ds}U_t^{(n)}(\xi_s).$$

Combining with (4.2.18) yields

$$\frac{D}{dt}\frac{d}{ds}U_t^{(n)}(\xi_s) = \nabla^2\Phi\Big(\mu_{[t]}^{(n)}, U_t^{(n)}(\xi_s)\Big)\cdot\frac{d}{ds}U_t^{(n)}(\xi_s).$$

Now,

$$\frac{d}{dt}\Big|\frac{d}{ds}U_t^{(n)}(\xi_s)\Big|^2 = 2\Big\langle\nabla^2\Phi\Big(\mu_{[t]}^{(n)}, U_t^{(n)}(\xi_s)\Big)\cdot\frac{d}{ds}U_t^{(n)}(\xi_s),\ \frac{d}{ds}U_t^{(n)}(\xi_s)\Big\rangle,$$

which is, by Condition (4.2.7), less than

$$2C_2\left|\frac{d}{ds}U_t^{(n)}(\xi_s)\right|^2.$$

By Gronwall lemma,

$$\Big|\frac{d}{ds}U_t^{(n)}(\xi_s)\Big| \le e^{C_2t}\,|\xi_s'|,$$

which implies that

$$d_M\Big(U_t^{(n)}(x), U_t^{(n)}(y)\Big) \le e^{C_2t}\,d_M(x,y).$$

Therefore the family $\big\{(t,x)\to U_t^{(n)}(x);\ n\ge 1\big\}$ is equicontinuous in $C([0,1]\times M)$. By Ascoli theorem, up to a subsequence, $U_t^{(n)}(x)$ converges to $U_t(x)$ uniformly in $(t,x)\in[0,1]\times M$. It is obvious to see that (U_t,μ_t) solves Mckean-Vlasov equation (4.2.6). □

Remark 4.2.6. Comparing to [BLPR17], as well to [Wan21], we did not suppose the Lipschitz continuity with respect to μ; in counterpart, we have no uniqueness of solutions of (4.2.6).

Remark 4.2.7. Many interesting PDE can be interpreted as gradient flows on the Wasserstein space $\mathbb{P}_2(M)$ (see [AGS05], [Vil09], [Vil03], [FS11]). The interpolation between geodesic flows and gradient flows were realized using Langevin's deformation in [LL16] and [LL18].

4.3 Levi-Civita connection on $\mathbb{P}_2(M)$

In this section, we will revisit the paper by J. Lott [Lot06]: we try to reformulate conditions given there as weak as possible, also to expose some of them in an intrinsic way, avoiding the use of density. In order to obtain good pictures on the geometry of $\mathbb{P}_2(M)$, the suitable class of probability measures should be the class $\mathbb{P}_{\text{div}}(M)$ of probability measures on M having divergence (see Definition 4.1.2).

For convenience of readers, we will briefly prepare materials needed for our exposition. For a measure $\mu \in \mathbb{P}_2(M)$, for any C^1 vector field A on M, the divergence $\text{div}_\mu(A) \in L^2(M,\mu)$ is such that

$$\int_M \langle \nabla\phi(x), A(x)\rangle_{T_xM}\, d\mu(x) = -\int_M \phi(x)\,\text{div}_\mu(A)(x)\, d\mu(x)$$

for any $\phi \in C^1(M)$. It is easy to see that $\text{div}_\mu(fA) = f\,\text{div}_\mu(A) + \langle\nabla f, A\rangle$ for $f \in C^1(M)$. If $d\mu = \rho\, dx$ has a density $\rho > 0$ in the space $C^1(M)$, we have

$$\int_M \langle\nabla\phi, A\rangle\, d\mu = \int_M \langle\nabla\phi, \rho A\rangle\, dx = -\int_M \phi\,\text{div}(\rho A)\, dx = -\int_M \phi\,\text{div}(\rho A)\,\rho^{-1} d\mu,$$

It follows that

$$\text{div}_\mu(A) = \rho^{-1}\,\text{div}(\rho A) = \text{div}(A) + \langle\nabla(\log\rho), A\rangle. \tag{4.3.1}$$

For $\mu \in \mathbb{P}_{\text{div}}(M)$ and $\phi \in C^2(M)$, we denote $\mathcal{L}^\mu(\phi) \in L^2(\mu)$ such that

$$\int_M \langle\nabla f, \nabla\phi\rangle\, d\mu = -\int_M f\,\mathcal{L}^\mu\phi\, d\mu, \quad \text{for any } f \in C^1(M), \tag{4.3.2}$$

where $\mathcal{L}^\mu\phi = \text{div}_\mu(\nabla\phi)$ is a negative operator.

Let $\psi \in C^3(M)$, consider the ODE

$$\frac{dU_t}{dt} = \nabla\psi(U_t), \quad U_0(x) = x.$$

Proposition 4.3.1. *Let $d\mu = \rho\, dx$ be a probability measure in $\mathbb{P}_{\text{div}}(M)$ with a strictly positive density ρ in $C^1(M)$ and $\psi \in C^3(M)$. Then for each $t \in [0,1]$, $\mu_t := (U_t)_\#\mu \in \mathbb{P}_{div}(M)$.*

Proof. By Kunita [Kun97] (see also [Cru83], [Mal97]), the push-forward measure $(U_t^{-1})_\#\mu$ by inverse map of U_t admits a density $\tilde{K}_t$ with respect to μ, having the following explicit expression

$$\tilde{K}_t = \exp\Big(-\int_0^t \text{div}_\mu(\nabla\psi)(U_s(x))ds\Big).$$

It follows that the density K_t of μ_t with respect to μ has the expression

$$K_t = \exp\Big(\int_0^t \text{div}_\mu(\nabla\psi)(U_{-s}(x))ds\Big).$$

According to (4.3.1), $x \to \mathsf{div}_\mu(\nabla\psi(x))$ is C^1. Therefore the condition in [Cru83]

$$\int_M \exp(\lambda \mathsf{div}_\mu(\nabla\psi(x))\, d\mu(x) < +\infty, \text{ for all } \lambda > 0$$

is automatically satisfied. Again by (4.3.1), $x \to K_t(x)$ is in C^1. Now let A be a C^1 vector field on M and $f \in C^1(M)$, we have

$$\int_M \langle \nabla f(x), A(x)\rangle_{T_xM}\, d\mu_t(x) = \int_M \langle \nabla f, A\rangle_{T_xM}\, K_t(x) d\mu(x) = -\int_M f\, \mathsf{div}_\mu(K_t Z)\, d\mu.$$

It follows that

$$\mathsf{div}_{\mu_t}(A) = \mathsf{div}_\mu(K_t A)\, K_t^{-1}.$$

□

For $\psi_1, \psi_2 \in C^2(M)$, we denote by V_{ψ_1}, V_{ψ_2} the associated constant vector fields on $\mathbb{P}_2(M)$. In what follows, we will compute the Lie bracket $[V_{\psi_1}, V_{\psi_2}]$.

For $f \in C^1(M)$, we set $F_f(\mu) = \int_M f\, d\mu$. According to preparations given at the beginning of Section 4.2,

$$(\bar{D}_{V_{\psi_2}} F_f)(\mu) = \int_M \langle \nabla\psi_2, \nabla f\rangle\, d\mu = F_{\langle \nabla\psi_2, \nabla f\rangle}(\mu).$$

Using again above formula, we have

$$(\bar{D}_{V_{\psi_1}} \bar{D}_{V_{\psi_2}} F_f)(\mu) = \int_M \langle \nabla\psi_1, \nabla\langle \nabla\psi_2, \nabla f\rangle\rangle\, d\mu = -\int_M \mathcal{L}^\mu \psi_1\, \langle \nabla\psi_2, \nabla f\rangle\, d\mu.$$

Therefore

$$\begin{aligned}[V_{\psi_2}, V_{\psi_1}]F_f &= \bar{D}_{V_{\psi_2}} \bar{D}_{V_{\psi_1}} F_f - \bar{D}_{V_{\psi_1}} \bar{D}_{V_{\psi_2}} F_f \\ &= \int_M \langle (\mathcal{L}^\mu\psi_1\, \nabla\psi_2 - \mathcal{L}^\mu\psi_2\, \nabla\psi_1),\ \nabla f\rangle\, d\mu.\end{aligned}$$

Let

$$\mathcal{C}_{\psi_1,\psi_2}(\mu) = \mathcal{L}^\mu\psi_1\, \nabla\psi_2 - \mathcal{L}^\mu\psi_2\, \nabla\psi_1. \tag{4.3.3}$$

Note that $\mathcal{C}_{\psi_1,\psi_2}(\mu)$ is in $L^2(M, \mathbf{T}M; \mu)$, not in $\mathbf{T}_\mu$. Consider the orthogonal projection:

$$\Pi_\mu : L^2(M, \mathbf{T}M; \mu) \to \mathbf{T}_\mu.$$

As $\mu \in \mathbb{P}_{div}(M)$ and by Proposition 4.1.3, there exists $\tilde{\Phi}_\mu \in \mathbb{D}_1^2(\mu)$ such that

$$\Pi_\mu(\mathcal{C}_{\psi_1,\psi_2}(\mu)) = \nabla\tilde{\Phi}_\mu. \tag{4.3.4}$$

Then we have

$$[V_{\psi_2}, V_{\psi_1}]F_f = \int_M \langle \nabla\tilde{\Phi}_\mu,\ \nabla f\rangle\, d\mu = (\bar{D}_{V_{\tilde{\Phi}_\mu}} F_f)(\mu). \tag{4.3.5}$$

Above equality can be extended to the class of polynomials on $\mathbb{P}_2(M)$, that is to say that

$$[V_{\psi_2}, V_{\psi_1}]_\mu = V_{\tilde{\Phi}_\mu} \quad \text{on polynomials,} \tag{4.3.6}$$

We emphasize that Lie bracket of two constant vector fields is no more a constant vector field.

Proposition 4.3.2. *Let $\psi_1, \psi_2 \in C^3(M)$, for $d\mu = \rho\, dx$ with $\rho > 0$ and $\rho \in C^2(M)$, the function $\tilde{\Phi}_\mu$ obtained in* (4.3.4) *has the following expression :*

$$\tilde{\Phi}_\mu = (\mathcal{L}^\mu)^{-1}\, \mathsf{div}_\mu\big(\mathcal{C}_{\psi_1,\psi_2}(\mu)\big). \tag{4.3.7}$$

Proof. By (4.3.1),

$$\mathcal{L}^\mu\psi = \Delta_M\psi + \langle\nabla\log\rho, \nabla\psi\rangle,$$

where Δ_M denotes the Laplace operator on M. It is well-known that $\mathcal{L}^\mu$ has a spectral gap if $\log\rho \in C^2(M)$. In [Lot06], the Lie bracket $[V_{\psi_2}, V_{\psi_1}]$ was expressed using Hodge decomposition for vector fields in $L^2(\mu)$. For a complete study on Hodge decompositions, we refer to the paper [Li09]. For $\psi_1, \psi_2 \in C^3(M)$, we have

$$\mathsf{div}_\mu\big(\mathcal{C}_{\psi_1,\psi_2}(\mu)\big) = \langle\nabla\mathcal{L}^\mu\psi_1,\ \nabla\psi_2\rangle - \langle\nabla\mathcal{L}^\mu\psi_2,\ \nabla\psi_1\rangle.$$

By Hodge decomposition, $\mathcal{C}_{\psi_1,\psi_2}(\mu)$ admits the decomposition

$$\mathcal{C}_{\psi_1,\psi_2}(\mu) = {d_\mu}^*\omega + \nabla f + h,$$

where ω is a differential 2-form on M, ${d_\mu}^*$ is adjoint operator of exterior derivative in $L^2(\mu)$, h is harmonic form : $({d_\mu}^*d + d{d_\mu}^*)h = 0$. Taking the divergence div_μ on the two sides of above equality, we see that f is a solution the following equation

$$\mathcal{L}^\mu f = \mathsf{div}_\mu\big(\mathcal{C}_{\psi_1,\psi_2}(\mu)\big).$$

It follows that $\tilde{\Phi}_\mu$ has the expression (4.3.7).

$\square$

Now we introduce the covariant derivative $\bar{\nabla}_{V_{\psi_1}} V_{\psi_2}$ associated to the Levi-Civita connection on $\mathbb{P}_2(M)$ by

$$\begin{aligned}2\langle \bar{\nabla}_{V_{\psi_1}} V_{\psi_2}, V_{\psi_3}\rangle &= \bar{D}_{V_{\psi_1}} \langle V_{\psi_2}, V_{\psi_3}\rangle + \bar{D}_{V_{\psi_2}} \langle V_{\psi_3}, V_{\psi_1}\rangle - \bar{D}_{V_{\psi_3}} \langle V_{\psi_1}, V_{\psi_2}\rangle \\ &+ \langle V_{\psi_3}, [V_{\psi_1}, V_{\psi_2}]\rangle - \langle V_{\psi_2}, [V_{\psi_1}, V_{\psi_3}]\rangle - \langle V_{\psi_1}, [V_{\psi_2}, V_{\psi_3}]\rangle.\end{aligned}$$

We have $\langle V_{\psi_2}, V_{\psi_3}\rangle = \int_M \langle \nabla\psi_2, \nabla\psi_3\rangle \, d\mu = F_{\langle \nabla\psi_2, \nabla\psi_3\rangle}$. Then

$$\bar{D}_{V_{\psi_1}} \langle V_{\psi_2}, V_{\psi_3}\rangle = \int_M \langle \nabla\psi_1, \nabla \, \langle \nabla\psi_2, \nabla\psi_3\rangle\rangle \, d\mu = -\int_M \langle \mathcal{L}^\mu \psi_1 \, \nabla\psi_2, \ \nabla\psi_3\rangle \, d\mu.$$

Replacing ψ_1 by ψ_2, ψ_2 by ψ_3 and ψ_3 by ψ_1, we get

$$\bar{D}_{V_{\psi_2}} \langle V_{\psi_3}, V_{\psi_1}\rangle = -\int_M \langle \mathcal{L}^\mu \psi_2 \, \nabla\psi_1, \ \nabla\psi_3\rangle \, d\mu.$$

We have, in the same way

$$\bar{D}_{V_{\psi_3}} \langle V_{\psi_1}, V_{\psi_2}\rangle = -\int_M \langle \mathcal{L}^\mu \psi_3 \, \nabla\psi_1, \ \nabla\psi_2\rangle \, d\mu.$$

Now using expression of $[V_{\psi_1}, V_{\psi_2}]$, we have

$$\langle V_{\psi_3}, [V_{\psi_1}, V_{\psi_2}]\rangle = \int_M \langle -\mathcal{L}^\mu \psi_1 \, \nabla\psi_2 + \mathcal{L}^\mu \psi_2 \, \nabla\psi_1, \nabla\psi_3\rangle \, d\mu.$$

In the same way, we get

$$\langle V_{\psi_2}, [V_{\psi_1}, V_{\psi_3}]\rangle = \int_M \langle -\mathcal{L}^\mu \psi_1 \, \nabla\psi_3 + \mathcal{L}^\mu \psi_3 \, \nabla\psi_1, \nabla\psi_2\rangle \, d\mu$$

and

$$\langle V_{\psi_1}, [V_{\psi_2}, V_{\psi_3}]\rangle = \int_M \langle -\mathcal{L}^\mu \psi_2 \, \nabla\psi_3 + \mathcal{L}^\mu \psi_3 \, \nabla\psi_2, \nabla\psi_1\rangle \, d\mu.$$

Combining all these terms, we finally get

$$2\langle \bar{\nabla}_{V_{\psi_1}} V_{\psi_2}, V_{\psi_3}\rangle = \int_M \langle \nabla \langle \nabla\psi_1, \nabla\psi_2\rangle, \ \nabla\psi_3\rangle \, d\mu + \int_M \langle \mathcal{L}^\mu \psi_2 \, \nabla\psi_1 - \mathcal{L}^\mu \psi_1 \, \nabla\psi_2, \ \nabla\psi_3\rangle \, d\mu.$$

Theorem 4.3.3. *(see [Lot06]) For two constant vector fields V_{ψ_1}, V_{ψ_2}, we have*

$$\bar{\nabla}_{V_{\psi_1}} V_{\psi_2} = \frac{1}{2} V_{\langle \nabla\psi_1, \nabla\psi_2\rangle} + \frac{1}{2}[V_{\psi_1}, V_{\psi_2}]. \tag{4.3.8}$$

Moreover, for any constant vector field V_{ψ_3}*,*

$$\langle \bar{\nabla}_{V_{\psi_1}} V_{\psi_2},\ V_{\psi_3}\rangle_{\mathbf{T}_\mu} = \int_M \langle \nabla^2\psi_2,\ \nabla\psi_1 \otimes \nabla\psi_3\rangle\, d\mu. \tag{4.3.9}$$

Proof. It is enough to prove (4.3.9). We have

$$\begin{aligned}
\langle V_{\psi_3}, [V_{\psi_1}, V_{\psi_2}]\rangle_{\mathbf{T}_\mu} &= \int_M \langle -\mathcal{L}^\mu \psi_1\, \nabla\psi_2 + \mathcal{L}^\mu\psi_2\, \nabla\psi_1, \nabla\psi_3\rangle\, d\mu \\
&= \int_M \langle \nabla\psi_1,\ \nabla\langle\nabla\psi_2, \nabla\psi_3\rangle\rangle\, d\mu - \int_M \langle \nabla\psi_2,\ \nabla\langle\nabla\psi_1, \nabla\psi_3\rangle\rangle\, d\mu \\
&= \int_M \Big(\langle \nabla^2\psi_2, \nabla\psi_1 \otimes \nabla\psi_3\rangle + \langle \nabla^2\psi_3, \nabla\psi_1\otimes\nabla\psi_2\rangle\Big) d\mu \\
&\quad - \int_M \Big(\langle \nabla^2\psi_1, \nabla\psi_2 \otimes \nabla\psi_3\rangle + \langle \nabla^2\psi_3, \nabla\psi_2\otimes\nabla\psi_1\rangle\Big) d\mu \\
&= \int_M \Big(\langle \nabla^2\psi_2, \nabla\psi_1 \otimes \nabla\psi_3\rangle - \langle \nabla^2\psi_1, \nabla\psi_2\otimes\nabla\psi_3\rangle\rangle\Big) d\mu,
\end{aligned}$$

due to the symmetry of the Hessian $\nabla^2\psi_3$. On the other hand,

$$\langle V_{\psi_3}, V_{\langle\nabla\psi_1, \nabla\psi_2\rangle}\rangle_{\mathbf{T}_\mu} = \int_M \Big(\langle \nabla^2\psi_2, \nabla\psi_3 \otimes \nabla\psi_1\rangle + \langle \nabla^2\psi_1, \nabla\psi_3\otimes\nabla\psi_2\rangle\rangle\Big) d\mu.$$

Summing these last two equalities yields (4.3.9).

□

Remark 4.3.4. By (4.3.8), for two constant vector fields V_{ψ_1}, V_{ψ_2}, the covariant derivative $\bar{\nabla}_{V_{\psi_1}} V_{\psi_2}$ is not a constant vector field on $\mathbb{P}_2(M)$ if $\psi_1 \neq \psi_2$.

Let $\alpha : \mathbb{P}_2(M) \to \mathbb{R}$ be a differentiable function, we define

$$\bar{\nabla}_{V_{\psi_1}}\big(\alpha\, V_{\psi_2}\big) = \bar{D}_{V_{\psi_1}}\alpha \cdot V_{\psi_2} + \alpha\, \bar{\nabla}_{V_{\psi_1}} V_{\psi_2}. \tag{4.3.10}$$

Proposition 4.3.5. *Let Z be a vector field on $\mathbb{P}_2(M)$ in the test space $\chi(\mathbb{P})$, that is, $Z = \sum_{i=1}^k \alpha_i\, V_{\psi_i}$ with α_i polynomials. Then $\bar{\nabla}_Z Z$ still is in the test space; moreover*

$$\bar{\nabla}_Z Z = V_{\Phi_1} + \frac{1}{2} V_{|\nabla\Phi_2|^2},$$

where

$$\Phi_1 = \sum_{j=1}^k \Big(\sum_{i=1}^k \alpha_i\, \bar{D}_{V_{\psi_i}}\alpha_j\Big)\, \psi_j, \quad \Phi_2 = \sum_{i=1}^k \alpha_i\, \psi_i.$$

Proof. Using the rule concerning covariant derivatives, $\bar{\nabla}_Z Z$ is equal to

$$\sum_{i,j=1}^{k} \alpha_i \left(\bar{D}_{V_{\psi_i}} \alpha_j\right) V_{\psi_j} + \frac{1}{2} \sum_{i,j=1}^{k} \alpha_i \alpha_j V_{\langle \nabla\psi_i, \nabla\psi_j\rangle} + \frac{1}{2} \sum_{i,j=1}^{k} \alpha_i \alpha_j [V_{\psi_i}, V_{\psi_j}].$$

The last sum is equal to 0 due to the skew-symmetry of $[V_{\psi_i}, V_{\psi_j}]$, the first one gives rise to Φ_1 and the second one gives rise to Φ_2.

□

In what follows, we will extend the definition of covariant derivative (4.3.10) for a general vector field Z on $\mathbb{P}_2(M)$. Let Δ be the Laplace operator on M, let $\{\varphi_n,\ n \geq 0\}$ be the eigenfunctions of Δ:

$$-\Delta\varphi_n = \lambda_n\, \varphi_n.$$

We have $\lambda_0 = 0$ and $\varphi_0 = 1$. It is well-known, by Weyl's result, that

$$\lambda_n \sim n^{2/m}, \quad n \to +\infty$$

where m is the dimension of M. The functions $\{\varphi_n;\ n \in \mathbb{N}\}$ are smooth, chosen to form an orthonormal basis of $L^2(M, dx)$. A function f on M is said to be in $H^k(M)$ for $k \in \mathbb{N}$, if

$$||f||_{H^k}^2 = \int_M |(I - \Delta)^{k/2} f|^2\, dx < +\infty.$$

By Sobolev embedding inequality, for $k > \dfrac{m}{2} + q$,

$$||f||_{C^q} \leq C\, ||f||_{H^k}.$$

For $f \in H^k(M)$, put $f = \sum_{n\geq 0} a_n \varphi_n$ which holds in $L^2(M, dx)$ with

$$a_n = \int_M f(x)\varphi_n(x)\, dx.$$

We have :

$$||f||_{H^k}^2 = \sum_{n\geq 0} a_n^2\, (1 + \lambda_n)^k.$$

The system $\left\{ \dfrac{\nabla\varphi_n}{\sqrt{\lambda_n}}; \quad n \geq 1 \right\}$ is orthonormal. Let $V_n = V_{\varphi_n/\sqrt{\lambda_n}}$, then $\{V_n;\ n \geq 1\}$ is an orthonormal basis of $\mathbf{T}_{dx}$.

Let Z be a vector field on $\mathbb{P}_2(M)$ given by $Z(\mu) = V_{\Phi(\mu,\cdot)}$ or $Z(\mu) = \nabla\Phi(\mu, \cdot)$. In the sequel, we denote: $\Phi_\mu(x) = \Phi(\mu, x)$, $\Phi^x(\mu) = \Phi(\mu, x)$. Then, if $x \to \nabla\Phi_\mu(x)$ is continuous,

$$\nabla\Phi_\mu = \sum_{n\geq 1}\Big(\int_M \langle\nabla\Phi_\mu, \frac{\nabla\varphi_n}{\sqrt{\lambda_n}}\rangle\, dx\Big)\,\frac{\nabla\varphi_n}{\sqrt{\lambda_n}} = \sum_{n\geq 1}\Big(\int_M \Phi_\mu\varphi_n dx\Big)\,\nabla\varphi_n,$$

which converges in $L^2(M, dx)$. Let $\mu \in \mathbb{P}_{\text{div}}(M)$, the above series converges also in $\mathbf{T}_\mu$. Let

$$a_n(\mu) = \int_M \Phi_\mu(x)\varphi_n(x)\, dx. \tag{4.3.11}$$

Let V_ψ be a constant vector field on $\mathbb{P}_2(M)$ with $\psi \in C^\infty(M)$. For $q \geq p \geq 1$, set

$$S_{p,q} = \sum_{n=p}^{q}\Big(\bar{D}_{V_\psi} a_n\, V_{\varphi_n} + a_n\,\bar{\nabla}_{V_\psi} V_{\varphi_n}\Big) = S^1_{p,q} + S^2_{p,q} \tag{4.3.12}$$

respectively. Let $\phi \in C^\infty(M)$, according to (4.3.9), we have

$$\langle S^2_{p,q}, V_\phi\rangle_{\mathbf{T}_\mu} = \int_M \Big(\sum_{n=p}^{q} a_n(\mu)\nabla^2\varphi_n\Big)(\nabla\psi(x), \nabla\phi(x))\, d\mu(x).$$

It follows that

$$|\langle S^2_{p,q}, V_\phi\rangle_{\mathbf{T}_\mu}| \leq \Big\|\sum_{n=p}^{q} a_n(\mu)\nabla^2\varphi_n\Big\|_\infty |V_\psi|_{\mathbf{T}_\mu}|V_\phi|_{\mathbf{T}_\mu},$$

therefore

$$|S^2_{p,q}|_{\mathbf{T}_\mu} \leq \Big\|\sum_{n=p}^{q} a_n(\mu)\nabla^2\varphi_n\Big\|_\infty |V_\psi|_{\mathbf{T}_\mu}.$$

We have

$$\begin{aligned}
&||\sum_{n=p}^{q} a_n(\mu)(I-\Delta)^{k/2}\varphi_n||^2_{L^2(dx)} = \sum_{n=p}^{q} a_n(\mu)^2(1+\lambda_n)^k\\
&= \sum_{n=p}^{q}\Big(\int_M (I-\Delta)^{k/2}\Phi_\mu\,\varphi_n\, dx\Big)^2 \to 0
\end{aligned}$$

as $p, q \to +\infty$ if $\Phi_\mu \in H^k(M)$. On the other hand, we have

$$(\bar{D}_{V_\psi} a_n)(\mu) = \int_M (\bar{D}_{V_\psi}\Phi^x)(\mu)\varphi_n(x)\, dx = \int_M \langle\nabla\bar{D}_{V_\psi}\Phi^x, \frac{\nabla\varphi_n}{\sqrt{\lambda_n}}\rangle\,\frac{dx}{\sqrt{\lambda_n}},$$

then

$$S_{p,q}^1 = \sum_{n=p}^q \Big(\int_M \langle \nabla \bar{D}_{V_\psi} \Phi^x, \frac{\nabla \varphi_n}{\sqrt{\lambda_n}}\rangle\, dx\Big) \frac{\nabla \varphi_n}{\sqrt{\lambda_n}}$$

and

$$\int_M |S_{p,q}^1|^2\, dx = \sum_{n=p}^q \Big(\int_M \langle \nabla \bar{D}_{V_\psi} \Phi^x, \frac{\nabla \varphi_n}{\sqrt{\lambda_n}}\rangle\, dx\Big)^2 \to 0$$

as $p, q \to +\infty$ if

$$\int_M |\nabla \bar{D}_{V_\psi} \Phi^x|^2\, dx < +\infty.$$

Therefore for $d\mu = \rho\, dx$ with $\mu \in \mathbb{P}_{\mathrm{div}}(M)$, as $p, q \to \infty$,

$$|S_{p,q}^1|^2_{\mathbf{T}_\mu} \leq ||\rho||_\infty \int_M |S_{p,q}^1|^2\, dx \to 0.$$

We get the following result, which is new.

Theorem 4.3.6. *Let Z be a vector field on $\mathbb{P}_2(M)$ given by $\Phi : \mathbb{P}_2(M) \times M \to \mathbb{R}$. Assume that*

(i) for some number $k > \frac{m}{2} + 2$, $\Phi_\mu \in H^k(M)$ for any $\mu \in \mathbb{P}_2(M)$,

(ii) for any $x \in M$, $\bar{D}_{V_\psi} \Phi^x$ exists and $\nabla \bar{D}_{V_\psi} \Phi^{\cdot} \in L^2(M, dx)$.

Then the covariant derivative $\bar{\nabla}_{V_\psi} Z$ is well defined at $\mu \in \mathbb{P}_{\mathrm{div}}(M)$ and for $\phi \in C^\infty(M)$,

$$\langle \bar{\nabla}_{V_\psi} Z, V_\phi \rangle_{\mathbf{T}_\mu} = \int_M \langle (\nabla \bar{D}_{V_\psi} \Phi^{\cdot}), \nabla \phi \rangle\, d\mu + \int_M \nabla^2 \Phi_\mu \big(\nabla \psi, \nabla \phi\big)\, d\mu. \tag{4.3.13}$$

Proof. Let $Z_q = \sum_{n=1}^q a_n V_{\varphi_n}$. Then

$$\bar{\nabla}_{V_\psi} Z_q = S_{1,q}.$$

Letting $q \to +\infty$ yields the result. □

4.4 Derivability of the square of the Wasserstein distance

Let $\{c_t;\ t \in [0,1]\}$ be an absolutely continuous curve on $\mathbb{P}_2(M)$, for $\sigma \in \mathbb{P}_2(M)$ given, the derivability of $t \to W_2^2(\sigma, c_t)$ was established in Chapter 8 of [AGS05] , as well as in [Vil09] (see pages 636-649); however they hold true only for almost all $t \in [0,1]$. The derivability at $t = 0$ was proved in Theorem 8.13 of [Vil03] if σ and c_0 have a density with respect to dx. When $\{c_t\}$ is a geodesic of constant speed, the derivability at $t = 0$ was given in theorem 4.2 of [Gig11] where the property of semi concavity was used. In what follows, we will use constant vector fields on $\mathbb{P}_2(M)$.

Before stating our result, we recall some well-known facts concerning optimal transport maps (see [Vil09,McC01, BB00]). Let $\sigma \in \mathbb{P}_{2,ac}(M)$ be absolutely continuous with respect to dx and $\mu \in \mathbb{P}_2(M)$, then there is an unique Borel map (up to a constant),$\phi \in \mathbb{D}_1^2(\sigma)$ such that

$$\int_M |\nabla\phi(x)|^2 \, d\sigma(x) = W_2^2(\sigma,\mu)$$

and $x \to T(x) = \exp_x(\nabla\phi(x))$ pushes σ forward to μ. If μ is also in $\mathbb{P}_{2,ac}(M)$, the map $T : M \to M$ is invertible and its inverse map T^{-1} is given by $y \to \exp_y(\nabla\tilde{\phi}(y))$ with some function $\tilde{\phi}$ such that $\int_M |\nabla\tilde{\phi}|^2 d\mu < +\infty$. We need also the following result

Lemma 4.4.1. *Let $x, y \in M$ and $\{\xi(t);\ t \in [0,1]\}$ be a minimizing geodesic connecting x and y, given by $\xi(t) = \exp_x(tu)$ with some $u \in T_xM$. Then*

$$d_M^2(\exp_y(v), x) - d_M^2(y,x) \le 2\langle v, \xi'(1)\rangle_{T_yM} + o(|v|) \quad \text{as } |v| \to 0. \tag{4.4.1}$$

Proof. See [McC01], page 10.

□

Theorem 4.4.2. *Assume that $\sigma \in \mathbb{P}_{2,ac}(M)$ is absolutely continuous with respect to dx, then $\mu \to \chi(\mu) := W_2^2(\sigma,\mu)$ is derivable along each constant vector field V_ψ at any $\mu \in \mathbb{P}_2(M)$. If $\mu \in \mathbb{P}_{2,ac}(M)$, the gradient $\nabla\chi$ exists and admits the expression :*

$$\nabla\chi(\mu) = -2\nabla\tilde{\phi}. \tag{4.4.2}$$

Proof. Remark first that Formula (4.4.2) is well-known in the case where $M = \mathbb{R}^m$ (see for example Theorem 8.13 in [Vil03]). Let $\psi \in C^\infty(M)$ and $(U_t)_{t\in\mathbb{R}}$ be the associated flow of diffeomorphisms of M:

$$\frac{dU_t(x)}{dt} = \nabla\psi(U_t(x)), \quad x \in M. \tag{4.4.3}$$

The inverse map U_t^{-1} of U_t satisfies the ODE

$$\frac{dU_t^{-1}(x)}{dt} = -\nabla\psi(U_t^{-1}(x)), \quad x \in M. \tag{4.4.4}$$

Set $\mu_t = (U_t)_\#\mu$, then $\mu = (U_t^{-1})_\#\mu_t$. Let $\gamma \in \mathcal{C}_o(\sigma,\mu)$ be the optimal coupling plan such that

$$W_2^2(\sigma,\mu) = \int_{M\times M} d_M^2(x,y)\, d\gamma(x,y).$$

The map $(x, y) \to (x, U_t(y))$ pushes γ forword to a coupling plan $\gamma_t \in \mathcal{C}(\sigma, \mu_t)$. Then for $t > 0$,

$$\begin{aligned}
&\frac{1}{t}\Big[W_2^2(\sigma, \mu_t) - W_2^2(\sigma, \mu)\Big] \le \frac{1}{t}\int_{M\times M}\Big(d_M^2(x, U_t(y)) - d_M^2(x, y)\Big)\, d\gamma(x, y)\\
&= \frac{1}{t}\int_{M\times M}\Big(d_M^2(x, U_t(y)) - d_M^2(x, \exp_y(t\nabla\psi(y))\Big)\, d\gamma(x, y)\\
&+ \frac{1}{t}\int_{M\times M}\Big(d_M^2(x, \exp_y(t\nabla\psi(y)) - d_M^2(x, y)\Big)\, d\gamma(x, y) = I_1(t) + I_2(t)
\end{aligned}$$

respectively. Let $\xi(t) = \exp_x(t\nabla\phi(x))$, by [McC01], ξ is a minimizing geodesic connecting x and $y = T(x)$. By Lemma 4.4.1, we have

$$d_M^2\big(x, \exp_y(t\nabla\psi(y)\big) - d_M^2(y, x) \le 2t\langle\nabla\psi(y), \xi'(1)\rangle_{T_yM} + o(|t|) \quad \text{as } t \to 0.$$

On the other hand,

$$\xi'(1) = d\exp_x(\nabla\phi(x)) \cdot \nabla\phi(x) = //_1^\xi \nabla\phi(x),$$

where $//_t^\xi$ denotes the parallel translation along the geodesic ξ. Hence $|\xi'(1)| = |\nabla\phi(x)|$. Therefore

$$I_2(t) \le 2\int_M \langle\nabla\psi(T(x)), d\exp_x(\nabla\phi(x)) \cdot \nabla\phi(x)\rangle\, d\sigma(x) + o(1)$$

To justify the passage of limit throught the integral, we note that for $t > 0$,

$$\begin{aligned}
&\frac{1}{t}\Big|d_M^2\big(x, \exp_y(t\nabla\psi(y))\big) - d_M^2(x, y)\Big|\\
&\le \frac{2}{t}\mathrm{diam}(M)\, d_M\big(y, \exp_y(t\nabla\psi(y))\big) \le 2\,\mathrm{diam}(M)\,|\nabla\psi(y)|.
\end{aligned}$$

Then

$$\overline{\lim_{t\downarrow 0}} I_2(t) \le 2\int_M \langle\nabla\psi(T(x)), d\exp_x(\nabla\phi(x)) \cdot \nabla\phi(x)\rangle\, d\sigma(x).$$

For estimating $I_1(t)$, it is obvious that

$$\lim_{t\downarrow 0}\frac{1}{t}\sup_{y\in M} d_M\big(U_t(y), \exp_y(t\nabla\psi(y))\big) = 0. \tag{4.4.5}$$

Then $\lim_{t\downarrow 0} I_1(t) = 0$. In conclusion:

$$\overline{\lim_{t\downarrow 0}}\frac{1}{t}\Big[W_2^2(\sigma,\mu_t)-W_2^2(\sigma,\mu)\Big]\le 2\int_M \langle\nabla\psi(T(x)), d\exp_x(\nabla\phi(x))\cdot\nabla\phi(x)\rangle\, d\sigma(x). \tag{4.4.6}$$

For obtaining the minoration, we use the fact that $\overline{\lim_{t\downarrow 0}}(-a_t)=-\underline{\lim}_{t\downarrow 0}a_t$.

Let $\tilde\gamma_t\in\mathcal{C}_o(\sigma,\mu_t)$ be the optimal transport plan:

$$W_2^2(\sigma,\mu_t)=\int_{M\times M} d_M^2(x,y)\,\tilde\gamma_t(dx,dy).$$

Let $\eta_t\in\mathcal{C}(\sigma,\mu_t)$ be defined by

$$\int_{M\times M} f(x,y)\eta_t(dx,dy)=\int_{M\times M} f\big(x,U_t^{-1}(y)\big)\,\tilde\gamma_t(dx,dy).$$

Then for $t>0$,

$$\frac{1}{t}\Big[W_2^2(\sigma,\mu)-W_2^2(\sigma,\mu_t)\Big]\le\frac{1}{t}\int_{M\times M}\Big(d_M^2(x,U_t^{-1}(y))-d_M^2(x,y)\Big)\,\tilde\gamma_t(dx,dy).$$

Let $T_t:M\to M$ be the optimal transport map which pushes forward σ to μ_t, with $T_t(x)=\exp_x(\nabla\phi_t(x))$. As $t\downarrow 0$, the map T_t converges in measure to T (see for example [Vil03], page 265). We have

$$\begin{aligned}
&\frac{1}{t}\int_{M\times M}\Big(d_M^2(x,U_t^{-1}(y))-d_M^2(x,y)\Big)\,\tilde\gamma_t(dx,dy)\\
&=\frac{1}{t}\int_M\Big(d_M^2(x,U_t^{-1}(T_t(x)))-d_M^2(x,T_t(x))\Big)\,d\sigma(x)\\
&=\frac{1}{t}\int_M\Big(d_M^2(x,U_t^{-1}(T_t(x)))-d_M^2(x,\exp_{T_t(x)}(-t\nabla\psi(T_t(x))))\Big)\,d\sigma(x)\\
&+\frac{1}{t}\int_M\Big(d_M^2(x,\exp_{T_t(x)}(-t\nabla\psi(T_t(x)))-d_M^2(x,T_t(x))\Big)\,d\sigma(x)=J_1(t)+J_2(t)
\end{aligned}$$

respectively. According to (4.4.5), $\lim_{t\downarrow 0}J_1(t)=0$. Concerning $J_2(t)$, we note as above,

$$\begin{aligned}
&\frac{1}{t}\Big|d_M^2\big(x,\exp_{T_t(x)}(-t\nabla\psi(T_t(x))\big)-d_M^2(x,T_t(x))\Big|\\
&\le\frac{2}{t}\,\text{diam}(M)d_M(T_t(x),\exp_{T_t(x)}(-t\nabla\psi(T_t(x)))\\
&\le 2\,\mathrm{diam}(M)\,|\nabla\psi(T_t(x)))|\le 2\mathrm{diam}(M)\,||\nabla\psi||_\infty.
\end{aligned}$$

Then by Lemma 4.4.1,

$$J_2(t) \le -2\int_M \langle \nabla\psi(T_t(x)), d\exp_x(\nabla\phi_t(x))\cdot\nabla\phi_t(x)\rangle\, d\sigma(x) + o(1)$$

Therefore

$$\overline{\lim_{t\downarrow 0}}\frac{1}{t}\Big[W_2^2(\sigma,\mu) - W_2^2(\sigma,\mu_t)\Big] \le -2\int_M \langle \nabla\psi(T(x)), d\exp_x(\nabla\phi(x))\cdot\nabla\phi(x)\rangle\, d\sigma(x). \tag{4.4.7}$$

Combining (4.4.6) and (4.4.7), we finally get

$$\lim_{t\downarrow 0}\frac{1}{t}\Big[W_2^2(\sigma,\mu_t) - W_2^2(\sigma,\mu)\Big] = 2\int_M \langle \nabla\psi(T(x)), d\exp_x(\nabla\phi(x))\cdot\nabla\phi(x)\rangle\, d\sigma(x). \tag{4.4.8}$$

Now if $\mu \in \mathbb{P}_{2,ac}(M)$ and the map $y \to \exp_y(\nabla\tilde{\phi}(y))$ is the optimal transport map which pushes μ to σ. Consider the minimizing geodesic

$$\xi(t) = \exp_y((1-t)\nabla\tilde{\phi}(y)),$$

which connects x and y. We have $\xi'(1) = -\nabla\tilde{\phi}(y)$. In this case, replacing $d\exp_x(\nabla\phi(x))\cdot\nabla\phi(x)$ in (4.4.8) by $\nabla\tilde{\phi}(y)$, we obtain

$$\begin{aligned}\lim_{t\downarrow 0}\frac{1}{t}\Big[W_2^2(\sigma,\mu_t) - W_2^2(\sigma,\mu)\Big] &= -2\int_M \langle \nabla\psi(T(x)), \nabla\tilde{\phi}(T(x))\rangle\, d\sigma(x)\\ &= -2\int_M \langle \nabla\psi(y), \nabla\tilde{\phi}(y)\rangle\, d\mu(y),\end{aligned} \tag{4.4.9}$$

from which we get (4.4.2). The proof is complete.

□

4.5 Parallel translations

Before introducing parallel translations on the space $\mathbb{P}_{div}(M)$, let's give a brief review on the definition of parallel translations on the manifold M, endowed with an affine connection. Let $\{\gamma(t);\ t \in [0,1]\}$ be a smooth curve on M, and $\{Y_t;\ t\in[0,1]\}$ a family vector fields along γ: $Y_t \in T_{\gamma(t)}M$. Then there exist vector fields X and Y on M such that

$$X(\gamma(t)) = \dot{\gamma}(t), \quad Y(\gamma(t)) = Y_t.$$

Y_t is said to be parallel along $\{\gamma(t);\ t\in[0,1]\}$ if

$$(\nabla_X Y)(\gamma(t)) = 0, \quad t \in [0,1].$$

Now let $\{c_t;\ t \in [0,1]\}$ be a one-to-one absolutely continuous curve on $\mathbb{P}_{div}(M)$ such that

$$\frac{d^I c_t}{dt} = V_{\Phi_t}, \quad \text{with } \Phi_t \in \mathbb{D}_1^2(c_t). \tag{4.5.1}$$

Let $\{Y_t;\ t \in [0,1]\}$ be a vector field along $\{c_t;\ t \in [0,1]\}$, that is, $Y_t \in \mathbf{T}_{c_t}$ given by $Y_t = V_{\Psi_t}$ with $\Psi_t \in \mathbb{D}_1^2(c_t)$.

Theorem 4.5.1. *Assume that $t \to c_t$ is C^1 in the sense that for any $f \in C^1(M)$, $t \to F_f(c_t)$ is C^1 and for $t \in [0,1]$, $x \to \Phi_t(x)$ is C^1. If for each $t \in [0,1]$,*

$$|V_{\Phi_t}|^2_{\mathbf{T}_{c_t}} = \int_M |\nabla \Phi_t(x)|^2\ c_t(dx) > 0, \tag{4.5.2}$$

then there are functions $(\mu, x) \to \tilde{\Phi}(\mu, x)$ and $(\mu, x) \to \tilde{\Psi}(\mu, x)$ on $\mathbb{P}_2(M) \times M$ such that

$$\tilde{\Phi}(c_t, x) = \Phi_t(x), \quad \tilde{\Psi}(c_t, x) = \Psi_t(x); \tag{4.5.3}$$

moreover for $x \in M$, $\mu \to \tilde{\Phi}(\mu, x)$ and $\mu \to \tilde{\Psi}(\mu, x)$ are derivable on $\mathbb{P}_2(M)$ along any constant vector fields V_ψ, their gradients exist on $\mathbb{P}_{2,ac}(M)$.

Proof. Fix $t_0 \in [0,1]$; consider $\alpha(t) = F_{\Phi_{t_0}}(c_t)$. Then

$$\alpha'(t) = \frac{d}{dt} F_{\Phi_{t_0}}(c_t) = \int_M \langle \nabla \Phi_{t_0}, \nabla \Phi_t \rangle\ c_t(dx),$$

which is > 0 at $t = t_0$. Therefore there is an open interval $I(t_0)$ of t_0 such that $t \to \alpha(t)$ is a C^1 diffeomorphism from $I(t_0)$ onto an interval $J(t_0)$ containing $\alpha(t_0)$. Let $\beta : J(t_0) \to I(t_0)$ be the inverse map of α. We have

$$F_{\Phi_{t_0}}(c_t) \in J(t_0) \quad \text{for } t \in I(t_0).$$

Let

$$U(t_0) = \big\{\mu \in \mathbb{P}_2(M);\ F_{\Phi_{t_0}}(\mu) \in J(t_0)\big\},$$

which is an open set in $\mathbb{P}_2(M)$. Let $r > 0$ and $\nu \in \mathbb{P}_2(M)$, we denote by $B(\nu, r)$ the open ball in $\mathbb{P}_2(M)$ centered at ν of radius r. Take $r_0 > 0$ small enough such that

$$B(c_{t_0}, r_0) \subset U(t_0).$$

We define, for $\mu \in B(c_{t_0}, r_0)$,

$$\tilde{\Phi}_{t_0}(\mu) = \Phi_{\beta(F_{\Phi_{t_0}}(\mu))}, \quad \tilde{\Psi}_{t_0}(\mu) = \Psi_{\beta(F_{\Phi_{t_0}}(\mu))}. \tag{4.5.4}$$

We remark that for $t \in [0,1]$ such that $c_t \in U(t_0)$, we have: $\beta(F_{\Phi_{t_0}}(c_t)) = t$. Note that $\{c_t;\ t \in [0,1]\}$ is a compact set of $\mathbb{P}_2(M)$ and

$$\big\{c_t;\ t \in [0,1]\big\} \subset \cup_{t_0 \in [0,1]} B(c_{t_0}, r_0).$$

There exists a finite number of $t_1, \ldots, t_k \in [0,1]$ such that

$$\big\{c_t;\ t \in [0,1]\big\} \subset \cup_{i=1}^k B(c_{t_i}, r_i).$$

Set $U = \cup_{i=1}^k B(c_{t_i}, r_i)$. Let $\mu \in U$, then $\mu \in B(c_{t_i}, r_i)$; according to (4.5.4), we define,

$$\tilde{\Phi}_{t_i}(\mu) = \Phi_{\beta_i(F_{\Phi_{t_i}}(\mu))}, \quad \tilde{\Psi}_{t_i}(\mu) = \Psi_{\beta_i(F_{\Phi_{t_i}}(\mu))}.$$

Then for $t \in [0,1]$ such that $c_t \in B(c_{t_i}, r_i)$, $\tilde{\Phi}_{t_i}(c_t) = \Phi_t$ and $\tilde{\Psi}_{t_i}(c_t) = \Psi_t$. Now for $r > 0$ and $\nu \in \mathbb{P}_2(M)$, we define

$$g_{r,\nu}(\mu) = \exp\Big(\frac{1}{W_2^2(\nu,\mu) - r^2}\Big), \quad \text{if } W_2(\nu,\mu) < r, \tag{4.5.5}$$

and $g_{r,\nu}(\mu) = 0$ otherwise. Then $g_{r,\nu}(\mu) > 0$ if and only if $\mu \in B(\nu, r)$. By Theorem 4.4.2, if $\nu \in \mathbb{P}_{\text{div}}$, $\mu \to g_{r,\nu}(\mu)$ is derivable along any constant vector field V_ψ. Remark that

$$\sum_{i=1}^k g_{r_i, c_{t_i}} > 0 \quad \text{on } U.$$

Let

$$\alpha_i = \frac{g_{r_i,c_{t_i}}}{\sum_{i=1}^k g_{r_i,c_{t_i}}} \quad \text{for } \mu \in U, \quad \text{and } \alpha_i = 0 \text{ otherwise.} \tag{4.5.6}$$

Now define

$$\tilde{\Phi}(\mu) = \sum_{i=1}^k \alpha_i(\mu)\tilde{\Phi}_{t_i}(\mu), \quad \tilde{\Psi}(\mu) = \sum_{i=1}^k \alpha_i(\mu)\tilde{\Psi}_{t_i}(\mu). \tag{4.5.7}$$

We have

$$\tilde{\Phi}(c_t) = \sum_{i=1}^k \alpha_i(c_t)\tilde{\Phi}_{t_i}(c_t).$$

Note that $\alpha_i(c_t) > 0$ if and only if $c_t \in B(c_{t_i}, r_i)$, which implies that $\tilde{\Phi}_{t_i}(c_t) = \Phi_t$ and $\tilde{\Phi}(c_t) = \sum_{i=1}^{k} \alpha_i(c_t)\Phi_t = \Phi_t$. It is the same for $\tilde{\Psi}$. The proof is completed.

□

Notice that for such a curve $\{c_t;\ t \in [0,1]\}$ given in Theorem 4.5.1, and $\{Y_t; t \in [0,1]\}$ a vector field along $\{c_t;\ t \in [0,1]\}$ given by Ψ_t. If furthermore for any $t \in [0,1]$, $\Psi_t \in H^k(M)$ with $k > \frac{m}{2} + 2$, then the extension obtained $\tilde{\Psi}$ obtained in Theorem 4.5.1 satisfies conditions in Theorem 4.3.6.

Definition 4.5.2. *We say that* $\{Y_t; t \in [0,1]\}$ *is parallel along* $\{c_t;\ t \in [0,1]\}$ *if*

$$(\bar{\nabla}_{\frac{dI_{c_t}}{dt}} V_{\tilde{\Psi}})(c_t) = 0, \quad t \in [0,1].$$

Using this definition, we re-discover the following formula, originally due to [Lot06].

Theorem 4.5.3. *Keeping the same notation in Theorem 4.5.1, if* $\{Y_t;\ t \in [0,1]\}$ *is parallel along* $\{c_t, t \in [0,1]\}$*, the following equation holds*

$$\int_M \Big\langle \nabla\Big(\frac{d\Psi_t}{dt}\Big) + \nabla_{\nabla\Phi_t}\nabla\Psi_t,\ \nabla\phi \Big\rangle\, c_t(dx) = 0, \quad \phi \in C^\infty(M). \tag{4.5.8}$$

Proof. Note that

$$(\bar{D}_{\frac{dI_{c_t}}{dt}}\tilde{\Psi})(c_t) = \frac{d}{dt}\tilde{\Psi}(c_t) = \frac{d\Psi_t}{dt} \text{ and } \nabla\tilde{\Psi}(c_t, \cdot) = \nabla\Psi_t.$$

Then (4.5.8) follows from (4.3.13). □

When $\nabla\big(\frac{d\Psi_t}{dt}\big) = \frac{d\nabla\Psi_t}{dt}$, it is more convenient to put Equation (4.5.8) in the following form :

$$\Pi_{c_t}\Big(\frac{d}{dt}\nabla\Psi_t + \nabla_{\nabla\Phi_t}\nabla\Psi_t\Big) = 0, \tag{4.5.9}$$

or

$$\frac{d}{dt}\nabla\Psi_t + \Pi_{c_t}\Big(\nabla_{\nabla\Phi_t}\nabla\Psi_t\Big) = 0, \tag{4.5.10}$$

where Π_{c_t} the orthogonal projection from $L^2(M, \mathbf{T}M, c_t)$ onto $\mathbf{T}_{c_t}$. By arguments in the proof of Proposition 4.3.2, when $dc_t = \rho_t\, dx$ with $\rho_t \in C^2(M)$ and $\rho_t > 0$, Π_{c_t} admits the expression

$$\Pi_{c_t} u = (\nabla\mathcal{L}_{c_t}^{-1}\mathrm{div}_{c_t})(u), \quad u \in L^2(M, \mathbf{T}M, c_t).$$

The price for this pointwise formulation of (4.5.9) as well as of (4.5.10) is the involement of second order derivative of Ψ.

Remark 4.5.4. Let $s \to \xi(s)$ is a smooth curve on M such that $\xi(0) = x$ and $\xi'(0) = \nabla\Phi_t(x)$, then

$$\frac{d}{dt}\nabla\Psi_t + \nabla_{\nabla\Phi_t}\nabla\Psi_t = \lim_{\varepsilon\to 0}\frac{\tau_\varepsilon^{-1}\nabla\Psi_{t+\varepsilon}(\xi(\varepsilon)) - \nabla\Psi_t(x)}{\varepsilon}, \tag{4.5.11}$$

where τ_s is the parallel translation along $s \to \xi(s)$. We refind the similar expression of parallel translations given in [AG08].

Proposition 4.5.5. *Assume that the curve $\{c_t;\ t \in [0,1]\}$ is induced by a flow of diffeomorphisms Φ_t, that is, there is a $C^{1,2}$ function $(t,x) \to \Phi_t(x)$ such that*

$$\begin{cases} \frac{dU_{s,t}(x)}{dt} &= \nabla\Phi_t(U_{s,t}(x)), \quad U_{s,s}(x) = x, \\ c_t &= (U_{0,t})_\# c_0. \end{cases}$$

Then for any $u_0 = \nabla\Psi_0 \in \mathbf{T}_{c_0}$, there is a unique vector field $u_t = \nabla\Psi_t \in \mathbf{T}_{c_t}$ along $\{c_t; t \in [0,1]\}$ such that

$$\Pi_{c_t}\Big(\lim_{\varepsilon\to 0}\frac{\tau_\varepsilon^{-1}\nabla\Psi_{t+\varepsilon}(U_{t,t+\varepsilon}(x)) - \nabla\Psi_t(x)}{\varepsilon}\Big) = 0 \tag{4.5.12}$$

holds in $L^2(c_t)$, where τ_ε is the parallel translation along $\{s \to U_{t,t+s}(x), s \in [0,\varepsilon]\}$.

Proof. Following Section 5 of [AG08], for $s \le t$, we define

$$\mathcal{P}_{t,s} : \mathbf{T}_{c_s} \to \mathbf{T}_{c_t}, \quad u_s \to \Pi_{c_t}\big(\tau_{t-s}u_s \circ U_{s,t}^{-1}\big).$$

For a subdivision $\mathcal{D} = \{0 = t_0 < t_1 < \ldots < t_n = 1\}$ of $[0,1]$, we define

$$\mathcal{P}_{\mathcal{D}} : \mathbf{T}_{c_0} \to \mathbf{T}_{c_1}, \quad u_0 \to (\mathcal{P}_{1,t_{n-1}} \circ \cdots \circ \mathcal{P}_{t_1,0})(u_0).$$

Under the assumption of Theorem, we have the uniform bound

$$\sup_{(t,x)\in[0,1]\times M} ||\nabla^2\Phi_t(x)|| < +\infty,$$

which allows us to mimic the construction of section 5 in [AG08], so that we get that $\mathcal{P}_{\mathcal{D}}$ converges as $\mathcal{D}$ becomes finer and finer, with $|\mathcal{D}| = \max_i |t_i - t_{i-1}| \to 0$. □

As a result of (4.5.12), we have as in [AG08] the following property:

Proposition 4.5.6. *Let $\{\nabla\Psi_t; t \in [0,1]\}$ be given in Proposition 4.5.5, then*

$$\frac{d}{dt}||\nabla\Psi_t||_{c_t}^2 = 0. \tag{4.5.13}$$

Proof. We have $c_{t+\varepsilon} = (U_{t,t+\varepsilon})_{\#} c_t$, and

$$\int_M |\nabla\Psi_{t+\varepsilon}(x)|^2\, c_{t+\varepsilon}(dx) = \int_M |\nabla\Psi_{t+\varepsilon}(U_{t,t+\varepsilon}(x))|^2\, c_t(dx).$$

Therefore

$$\begin{aligned} ||u_{t+\varepsilon}||^2_{\mathbf{T}_{t+\varepsilon}} - ||u_t||^2_{\mathbf{T}_{c_t}} &= \int_M \Big[|\tau_\varepsilon^{-1}\nabla\Psi_{t+\varepsilon}(U_{t,t+\varepsilon}(x))|^2 - |\nabla\Psi_t(x)|^2\Big]\, c_t(dx) \\ &= \int_M \Big\langle \tau_\varepsilon^{-1}\nabla\Psi_{t+\varepsilon}(U_{t,t+\varepsilon}(x)) - \nabla\Psi_t(x), \tau_\varepsilon^{-1}\nabla\Psi_{t+\varepsilon}(U_{t,t+\varepsilon}(x))\Big\rangle\, c_t(dx) \\ &+ \int_M \Big\langle \nabla\Psi_t(x),\ \tau_\varepsilon^{-1}\nabla\Psi_{t+\varepsilon}(U_{t,t+\varepsilon}(x)) - \nabla\Psi_t(x)\Big\rangle\, c_t(dx). \end{aligned}$$

It follows that

$$\frac{d}{dt}||\nabla\Phi_t||^2_{c_t} = 2\int_M \Big\langle \lim_{\varepsilon\to 0} \frac{\tau_\varepsilon^{-1}\nabla\Psi_{t+\varepsilon}(U_{t,t+\varepsilon}(x)) - \nabla\Psi_t(x)}{\varepsilon}, \nabla\Psi_t(x)\Big\rangle\, c_t(dx) = 0.$$

□

4.5.1 The case when $M = \mathbb{T}$

In this section, we give well-posedness of parallel translation on $\mathbb{P}(\mathbb{T})$. A function v on $\mathbb{T}$ is the derivative of a function ϕ if and only if $\displaystyle\int_{\mathbb{T}} v(x)dx = 0$. Let $d\mu = \rho\, dx$ be a probability measure on $\mathbb{T}$ with $\rho > 0$. Set $\phi' = \Pi_\mu(v)$; then

$$\int_{\mathbb{T}} f'(x)v(x)\rho(x)\, dx = \int_{\mathbb{T}} f'(x)\phi'(x)\rho(x)\, dx \quad \text{for any} \quad f \in C^\infty(\mathbb{T}),$$

which implies that $(v\rho)' = (\phi'\rho)'$; so there is a constante $K \in \mathbb{R}$ such that

$$v\rho = \phi'\,\rho + K, \quad \text{or} \quad v = \phi' + \frac{K}{\rho};$$

integrating the two sides over $\mathbb{T}$ yields

$$K = -\frac{\int_{\mathbb{T}} v(x)\, dx}{\int_{\mathbb{T}} \frac{dx}{\rho}}.$$

It follows that

$$\Pi_\mu(v) = v - \Big(\frac{\int_{\mathbb{T}} v(x)\, dx}{\int_{\mathbb{T}} \frac{dx}{\rho}}\Big)\,\frac{1}{\rho}. \tag{4.5.14}$$

In particular, $\Pi_\mu(1) = 1 - \frac{1}{(\int_{\mathbb{T}} \frac{dx}{\rho})\rho}$. In what follows, we denote $\hat{\rho} = \frac{1}{(\int_{\mathbb{T}} \frac{dx}{\rho})\,\rho}$. It is obvious that $\int_{\mathcal{T}} \hat{\rho}\, dx = 1$. In order to make clear the dependence of the density $\rho = \frac{d\mu}{dx}$, we write the projection Π_μ in the form:

$$\Pi_\rho(v) = v - \Big(\int_{\mathbb{T}} v(x)\, dx\Big)\, \hat{\rho}. \tag{4.5.15}$$

Theorem 4.5.7. *Assume that the initial vector $\partial_x \Psi_0 \in \mathcal{C}^\infty$, the initial measure density $\rho_o > 0,\ \rho_0 \in \mathcal{C}^\infty$, and $\phi \in \mathcal{C}^\infty$. Let the flow $\{X_t,\ t \in [0,1]\}$ is induced by the following ODE:*

$$d_t X_t = \partial_x \phi(X_t) dt.$$

Denote $\Xi_t = (X_t)^{-1}$ and the image measure $\rho_t = (X_t)_{\#}\rho_0$, then the parallel translation equation (4.5.10) *has a unique smooth solution g_t satisfies*

$$g_t = \partial_x \Psi_0(\Xi_t) + \int_0^t \frac{\int_{\mathbb{T}} g_s \partial_x^2 \phi dx}{\int_{\mathbb{T}} \frac{1}{\rho_s} dx} \frac{1}{\rho_s \circ \Xi_{t-s}} ds. \tag{4.5.16}$$

Proof. If g_t solves (4.5.10) , i.e.

$$\partial_t g_t = -\Pi_{\rho_t}(\partial_x g_t \cdot \partial_x \phi), \tag{4.5.17}$$

then, by (4.5.14) , we have

$$\partial_t g_t = -\partial_x g_t \cdot \partial_x \phi + \frac{1}{\rho_t} K_t^g$$

where

$$K_t^g = -\frac{\int_{\mathbb{T}} g_t \partial_x^2 \phi dx}{\int_{\mathbb{T}} \frac{1}{\rho_t} dx}. \tag{4.5.18}$$

This is a transport-type integral differential equation. By taking integration on both sides, we can see

$$\partial_t \int_{\mathbb{T}} g_t dx = -\int_{\mathbb{T}} \partial_x g_t \partial_x \phi dx + K_t^g \int_{\mathbb{T}} \frac{1}{\rho_t} dx = 0.$$

Assume that $f_t = g_t(X_t)$, then

$$\frac{d}{dt} f_t = \frac{1}{\rho_t \circ X_t} K_t^g \tag{4.5.19}$$

We can use Euler approximation to prove the existence of solution. Given $N-$piece partition of $[0,1]$,

$g_0^N = \partial_x \Psi_0$, $f_0^N = \partial_x \Psi_0$, then for the next step, let

$$g_t^N = g_0^N, \qquad t \in [0, \frac{1}{N}), \tag{4.5.20}$$

$$f_t^N(x) = f_0^N(x) + \int_0^t K_s^{g^N} \frac{1}{\rho_s \circ X_s} ds, \qquad t \in (0, \frac{1}{N}]. \tag{4.5.21}$$

Define $g^N_{\frac{1}{N}} = f^N_{\frac{1}{N}} \circ \Xi_{\frac{1}{N}}$, then we can continue this construction for g^N and f^N . For the k-th step, let

$$g_t^N = g^N_{\frac{k}{N}}, \qquad t \in [\frac{k}{N}, \frac{k+1}{N}), \tag{4.5.22}$$

$$f_t^N(x) = f_0^N(x) + \int_0^t K_s^{g^N} \frac{1}{\rho_s \circ X_s} ds, \qquad t \in (\frac{k}{N}, \frac{k+1}{N}]. \tag{4.5.23}$$

Define $g^N_{\frac{k+1}{N}} = f^N_{\frac{k+1}{N}} \circ \Xi_{\frac{k+1}{N}}$. Set $M = \max_{[0,1]\times\mathbb{T}} \rho$, $m = \min_{[0,1]\times\mathbb{T}} \rho$ and $M' = \max_{\mathbb{T}} \partial_x^2 \phi$. Note that,

$$\begin{aligned} |K_s^{g^N}| &\leq MM' \int_{\mathbb{T}} |g_s^N| dx \\ &\leq C \int_{\mathbb{T}} \left| f^N_{\frac{[Ns]}{N}} \circ \Xi_{\frac{[Ns]}{N}} \right|^2 dx \\ &\leq \frac{C}{m} \int_{\mathbb{T}} \left| f^N_{\frac{[Ns]}{N}} \circ \Xi_{\frac{[Ns]}{N}} \right|^2 \rho_{\frac{[Ns]}{N}} dx \\ &\leq C ||f^N_{\frac{[Ns]}{N}}||_{L^2}. \end{aligned} \tag{4.5.24}$$

Thus, by (4.5.23), when N is large enough,

$$\frac{d}{dt} ||f_t^N||^2_{L^2} \leq \frac{C}{m} ||f_t^N||^2_{L^2}.$$

So, by Gronwell inequality,

$$||f_t^N||^2_{L^2} \leq ||\partial_x \Psi_0||^2_{L^2} \exp Ct. \tag{4.5.25}$$

L^2-uniform boundedness has been proved. Moreover, we can prove uniform boundedness of $\{f_t^N\}$ in $\mathbb{D}_1^2$ so that $\{f_t^N\}$ is compact in L^2 for each t . In fact,

$$||\partial_x f_t^N||_{L^2} \leq ||\partial_x \partial_x \Psi_0||_{L^2} + \max_{s\in[0,t]} |K_s^{g^N}| \max_{[0,t]\times\mathcal{T}} |\partial_x (\frac{1}{\rho_s \circ X_s})| \leq C \, ||\partial_x^2 \Psi_0||_{L^2}. \tag{4.5.26}$$

The last inequality needs estimates (4.5.31) and (4.5.32) below. For the equicontinuity, through (4.5.24) and (4.5.25), we see that

$$||f_t^N - f_s^N||_{L^2} \leq C||\partial_x \Psi_0||_{L^2}|t-s|.$$

Therefore, with compactness and equicontinuity, we know that, according to Arzelà-Ascoli theorem, $\{f_t^N, t \in [0,1]\}$ has a convergent subsequence $\{f_t^n, t \in [0,1]\}$ in $\mathcal{C}([0,1], L^2(\mathbb{T}))$. Denote f_t as the convergent limit . Note that, since

$$g_t^n = f^n_{\frac{[nt]}{n}} \circ \Xi_{\frac{[nt]}{n}},$$

thus,

$$\int_{\mathbb{T}} |g_t^n|^2 dx \leq \frac{1}{m}\int_{\mathbb{T}} |g_t^n|^2 \rho_{\frac{[nt]}{n}} dx \leq C \int_{\mathbb{T}} \left|f^n_{\frac{[nt]}{n}}\right|^2 dx.$$

The last inequality is due to (4.5.24). This , combined with (4.5.25), gives uniform boundedness of $\{g_t^n, t \in [0,1]\}$. Similarly, we can also prove uniform boundedness of $\{g_t^n\}$ in $\mathbb{D}_1^2$. Actually,

$$||\partial_x g_t^n|| = ||(\partial_x f^n_{\frac{[nt]}{n}}) \circ \Xi_{\frac{[nt]}{n}} \cdot \partial_x \Xi_{\frac{[nt]}{n}}||$$

should be uniformly bound because of (4.5.26) and (4.5.31). Next , according to lemma 4.5.8 below, $||g_t^n - g_s^n||_{L^2} \leq C|t-s|$. We proved the equicontinuity of $\{g_t^n,\, t \in [0,1]\}$. Thus , again by Arzelà-Ascoli theorem, we have a subsequence (f^{n_k}, g^{n_k}) such that $f_t^{n_k}$ and $g_t^{n_k}$ converge to f_t and g_t respectively under $\mathcal{C}([0,1], L^2)$. Note that , by (4.5.23) and L^2 convergence of $g_t^{n_k}$, we can easily check that

$$\max_{\mathcal{T}} |f_t^{n_k} - f_t| \to 0, \tag{4.5.27}$$

Thus, again , taking pointwise limit of (4.5.23) , we get

$$f_t = \partial_x \Psi_0 + \int_0^t K_s^g \frac{1}{\rho_s \circ X_s} ds.$$

Since $\rho_s \in \mathcal{C}^\infty$, we see that $f_t \in \mathcal{C}^\infty$. Let $\bar{g}_t = f_t \circ \Xi_t$, then for each $x \in \mathcal{T}$, $\bar{g}_t(x) = \lim_{k\to\infty} g_t^{n_k}(x)$ due to (4.5.27). By dominated convergence theorem, $||g_t - \bar{g}_t||_{L^2} = 0$. Next, we will prove $\bar{g}_t$ is a gradient of some function on Torus and solves (4.5.17) . In fact,

$$\begin{aligned}
\frac{d}{dt}\bar{g}_t &= \left(\frac{d}{dt} f_t\right) \circ \Xi_t + \partial_x f_t(\Xi_t) \cdot \frac{d}{dt}\Xi_t \\
&= \left(K_t^g \frac{1}{\rho_t \circ \Psi_t}\right) \circ \Xi_t - \partial_x f_t(\Xi_t) \partial_x \Xi_t \partial_x \phi \\
&= K_t^g \frac{1}{\rho_t} - \partial_x(\bar{g}_t)\partial_x \phi.
\end{aligned}$$

It is easy to check that $|K_t^g - K_t^{\bar{g}_t}| \le C||g - \bar{g}_t||_{L^2} = 0$. Therefore, $\bar{g}_t$ solves (4.5.17) . Also, by Fubini theorem,

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{T}}\bar{g}_t dx &= \int_{\mathcal{T}}\frac{d}{dt}\bar{g}_t dx\\ &= \int_{\mathbb{T}} K_t^g \frac{1}{\rho_t} - \partial_x(\bar{g}_t)\partial_x\phi dx\\ &= \int_{\mathbb{T}}(\bar{g}_t - g_t)\partial_x^2\phi dx.\end{aligned}$$

Thus, $\left|\frac{d}{dt}\int_{\mathbb{T}}\bar{g}_t dx\right| \le C||g_t - \bar{g}_t||_{L^2} = 0$. Note that $\int_{\mathbb{T}}\bar{g}_0 dx = \int_{\mathbb{T}}\partial_x\Psi_0 dx = 0$. So we proved $\bar{g}_t$ is a gradient of some function on Torus. We finished the proof. ☐

Lemma 4.5.8. *For $t > s$, when n is large enough,*

$$\left|\left|f^n_{\frac{[nt]}{n}} \circ \Xi_{\frac{[nt]}{n}} - f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[ns]}{n}}\right|\right|_{L^2} \le C|t-s|.$$

Proof. Since $||f_t^n - f_s^n||_{L^2} \le K|t-s|$, thus

$$\left|\left|f^n_{\frac{[nt]}{n}} \circ \Xi_{\frac{[nt]}{n}} - f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[nt]}{n}}\right|\right|_{L^2} \le \frac{1}{m}\left(\int (f^n_{\frac{[nt]}{n}} - f^n_{\frac{[ns]}{n}})^2 dx\right)^{\frac{1}{2}} \le C_1|t-s|. \tag{4.5.28}$$

Also, because of

$$X_t = x + \int_0^t \partial_x\phi(X_s)ds, \tag{4.5.29}$$

$|\Xi_t - x| = |X(\Xi_t(x)) - \Xi_t(x)| \le M't$. Therefore, when n is large enough,

$$\begin{aligned}&\left|\left|f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[nt]}{n}} - f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[ns]}{n}}\right|\right|_{L^2}\\ &\le \frac{1}{m}\left(\int (f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[nt]}{n}} - f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[ns]}{n}})^2 \rho_{\frac{[ns]}{n}} dx\right)^{\frac{1}{2}}\\ &= \frac{1}{m}\left(\int (f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[nt]}{n} - \frac{[ns]}{n}} - f^n_{\frac{[ns]}{n}})^2 dx\right)^{\frac{1}{2}}\\ &\le \frac{M'}{m}\max_{\mathcal{T}}\left|\partial_x f^n_{\frac{[ns]}{n}}\right| \cdot |t-s|.\end{aligned} \tag{4.5.30}$$

Due to (4.5.23),

$$\left|\partial_x f^n_{\frac{[ns]}{n}}\right| \le \max_{\mathbb{T}}|\partial_x\Psi_0| + \int_0^t |K_s^{g^n}| \cdot \left|\frac{\partial_x\rho_s}{\rho_s^2} \circ X_s\right| \cdot |\partial_x X_s| ds.$$

Note that , by (4.5.24) and (4.5.25),

$$|K_s^{g^n}| \leq C||\partial_x \Psi_0||_{L^2},$$

Furthermore, by (4.5.29) , we can get

$$\partial_x X_t = 1 + \int_0^t \partial_x^2 \phi(X_s) \partial_s X_s ds,$$

which means $|\partial_x X_t| \leq \exp\{\max_{\mathbb{T}} |\partial_x^2 \phi| t\}$. Similarly, by the standard argument, when $\phi \in \mathcal{C}^\infty$,

$$|\partial_x^k X_s| \leq C, \quad \text{for } s \in [0,1]. \tag{4.5.31}$$

These estimates also hold for the inverse map Ξ_t , which satisfies

$$\Xi_t = x - \int_0^t \partial_x \phi(\Xi_{t-s}) ds.$$

On the other hand, by the property of push-forward measure $\rho_t = (X_t)_\# \rho_0$:

$$\rho_t(X_t) = \rho_0 |\partial_x X_t|,$$

therefore, it is easy to deduce that , when $\rho_0 \in \mathcal{C}^\infty$,

$$|\partial_x \rho_s| \leq C. \tag{4.5.32}$$

In fact, estimates (4.5.32) and (4.5.31) are standard results on diffeomorphism induced by smooth vector fields.
Finally, we come to estimate (4.5.30) and get

$$\left|\left| f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[nt]}{n}} - f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[ns]}{n}} \right|\right|_{L^2} \leq C_2 |t-s|. \tag{4.5.33}$$

Then, combining (4.5.28) and (4.5.33) , we have

$$\begin{aligned} &\left|\left| f^n_{\frac{[nt]}{n}} \circ \Xi_{\frac{[nt]}{n}} - f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[ns]}{n}} \right|\right|_{L^2} \\ &\leq \left|\left| f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[nt]}{n}} - f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[ns]}{n}} \right|\right|_{L^2} + \left|\left| f^n_{\frac{[nt]}{n}} \circ \Xi_{\frac{[nt]}{n}} - f^n_{\frac{[ns]}{n}} \circ \Xi_{\frac{[nt]}{n}} \right|\right|_{L^2} \\ &\leq (C_1 + C_2)|t-s|. \end{aligned} \tag{4.5.34}$$

$\square$

4.6 Lipschitz condition for vector fields and uniqueness of solution to ODE

In what follows, we will say a few words on the Lipschitz condition on vector fields Z on $\mathbb{P}_{2,ac}(M)$. Let $\mu, \nu \in \mathbb{P}_{2,ac}(M)$. Recall that there is a unique optimal transport map $T_{\mu,\nu} : M \to M$ which pushes μ to ν such that

$$T_{\mu,\nu}(x) = \exp_x(\nabla\phi(x)).$$

Let $\xi_x^{\mu,\nu}(t) = \exp_x\big(t\nabla\phi(x)\big)$ and $//_t^{\xi_x}$ be the parallel translation along $\{\xi_x(t);\ t \in [0,1]\}$.

Definition 4.6.1. *We say that a vector field Z on $\mathbb{P}_{2,ac}(M)$ given by Φ (see Definition 4.2.3) is Lipschitzian if there exists a constant $\kappa > 0$ such that*

$$\int_M \Big| //_1^{\xi_x^{\mu,\nu}} \nabla\Phi(\mu, x) - \nabla\Phi(\nu, T_{\mu,\nu}(x)) \Big|^2 d\mu(x) \le \kappa^2\, W_2^2(\mu,\nu) \tag{4.6.1}$$

for any couple $(\mu,\nu) \in \mathbb{P}_{2,ac}(M) \times \mathbb{P}_{2,ac}(M)$.

Remark that the quantity defined by the left hand side of (4.6.1) is symmetric with respect to (μ,ν), using the inverse map $T_{\nu,\mu}$ of $T_{\mu,\nu}$.

Proposition 4.6.2. *Assume that for each $\mu \in \mathbb{P}_2(M)$, $x \to \nabla^2\Phi(\mu, x)$ exists and is continuous such that*

$$C_1 = \sup_{\mu\in\mathbb{P}_2(M)} ||\nabla^2\Phi(\mu,\cdot)||_\infty < +\infty, \tag{4.6.2}$$

and there is a constant $C_2 > 0$ such that

$$|\nabla\Phi(\mu,x) - \nabla\Phi(\nu,x)| \le C_2\, W_2(\mu,\nu), \quad x \in M; \tag{4.6.3}$$

then the Lipschitz condition (4.6.1) *holds with $\kappa^2 \le 2(C_1^2 + C_2^2)$.*

Proof. We have

$$\begin{aligned} &\Big| //_1^{\xi_x^{\mu,\nu}} \nabla\Phi(\mu,x) - \nabla\Phi(\nu, T_{\mu,\nu}(x)) \Big| \\ \le &\Big| //_1^{\xi_x^{\mu,\nu}} \nabla\Phi(\mu,x) - //_1^{\xi_x^{\mu,\nu}} \nabla\Phi(\nu,x) \Big| + \Big| //_1^{\xi_x^{\mu,\nu}} \nabla\Phi(\nu,x) - \nabla\Phi(\nu, T_{\mu,\nu}(x)) \Big| \\ \le &\Big| \nabla\Phi(\mu,x) - \nabla\Phi(\nu,x) \Big| + C_1\, d_M(x, T_{\mu,\nu}(x)) \end{aligned}$$

where the second inequality is deduced from the fact for $x, y \in M$ and $\{\eta_t; t \in [0,1]\}$ a minimizing geodesic connecting x and y, then for $\varphi \in C^2(M)$,

$$\Big|//_1^\eta \nabla\varphi(x) - \nabla\varphi(y)\Big| \leq ||\nabla^2\varphi||_\infty \, d_M(x,y). \tag{4.6.4}$$

In fact, set $z(t) = //_t^\eta \nabla\varphi(x) - \nabla\varphi(\eta_t)$. Then the covariant derivative $\dfrac{D}{dt}z$ of $z(t)$ along η has the expression

$$\frac{D}{dt}z(t) = \nabla_{\eta_t'}\nabla\varphi(\eta_t).$$

It follows that $\Big|\dfrac{D}{dt}z(t)\Big| \leq |\eta_t'|\,||\nabla^2\varphi||_\infty$; therefore

$$\Big|//_1^\eta \nabla\varphi(x) - \nabla\varphi(y)\Big| \leq ||\nabla^2\varphi||_\infty \int_0^1 |\eta_t'|\,dt = ||\nabla^2\varphi||_\infty \, d_M(x,y).$$

Using conditions (4.6.2) and (4.6.3), we get

$$\begin{aligned}
&\int_M \Big|//_1^{\xi_x^{\mu,\nu}} \nabla\Phi(\mu,x) - \nabla\Phi(\nu, T_{\mu,\nu}(x))\Big|^2 d\mu(x) \\
&\leq 2\,\Big[C_2^2 W_2^2(\mu,\nu) + C_1^2 \int_M d_M^2(x, T_{\mu,\nu}(x))\,d\mu(x)\Big].
\end{aligned}$$

The result follows. □

Theorem 4.6.3. *Let Z be a vector field on $\mathbb{P}_2(M)$ satisfying the Lipschitz condition* (4.6.1)*, then the ODE*

$$\frac{d^I\mu_t}{dt} = Z(\mu_t), \quad \mu_{|t=0} = \mu_0$$

admits unique solution on the space $\mathbb{P}_{2,ac}(M)$.

Proof. Let μ_t^1, μ_t^2 be two solutions in $\mathbb{P}_{2,ac}(M)$ to above ODE. For fixed t, denote by $T_t^{1,2} : M \to M$ the optimal transport map which pushes μ_t^1 to μ_t^2, with

$$T_t^{1,2}(x) = \exp_x\Big(\nabla\phi^{1,2}(x)\Big).$$

Let

$$T_t^{2,1}(y) = \exp_x\Big(\nabla\phi^{2,1}(y)\Big)$$

be the inverse map of $T_t^{1,2}$. Let

$$\eta_s^{1,2}(x) = \exp_x\Big(s\nabla\phi^{1,2}(x)\Big).$$

It is well known (see [Vil09]) that $\phi^{1,2}$ and $\phi^{2,1}$ are linked by the following relation

$$//_1^{\eta^{1,2}(x)}\nabla\phi^{1,2}(x) = -\nabla\phi^{2,1}(T_t^{1,2}(x)), \quad x \in M. \tag{4.6.5}$$

According to Theorem 23.9 in [Vil09], for almost all $t \in (0,1)$,

$$\begin{aligned}
\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t^1,\mu_t^2) &= -\langle\nabla\phi^{1,2}, \frac{d^I\mu_t^1}{dt}\rangle_{\mathbf{T}_{\mu_t^1}} - \langle\nabla\phi^{2,1}, \frac{d^I\mu_t^2}{dt}\rangle_{\mathbf{T}_{\mu_t^2}} \\
&= -\int_M \langle\nabla\phi^{1,2}(x), \nabla\Phi(\mu_t^1,x)\rangle\,\mu_t^1(dx) - \int_M \langle\nabla\phi^{2,1}(y), \nabla\Phi(\mu_t^2,y)\rangle\,\mu_t^2(dy).
\end{aligned}$$

The second term on the right hand side is equal to

$$-\int_M \langle\nabla\phi^{2,1}(T_t^{1,2}(x)), \nabla\Phi(\mu_t^2, T_t^{1,2}(x))\rangle\,\mu_t^1(dx),$$

which is equal to, by (4.6.5),

$$\int_M \langle //_1^{\eta^{1,2}(x)}\nabla\phi^{1,2}(x), \nabla\Phi(\mu_t^2, T_t^{1,2}(x))\rangle\,\mu_t^1(dx).$$

Therefore

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t^1,\mu_t^2) = \int_M \langle //_1^{\eta^{1,2}(x)}\nabla\phi^{1,2}(x),\ \nabla\Phi(\mu_t^2, T_t^{1,2}(x)) - //_1^{\eta^{1,2}(x)}\nabla\Phi(\mu_t^1,x)\rangle\,\mu_t^1(dx),$$

which is dominated, using Cauchy-Schwarz inequality by

$$\Big(\int_M |\nabla\phi^{1,2}(x)|\mu_t^1(dx)\Big)^{1/2}\ \Big(\int_M \Big|\nabla\Phi(\mu_t^2, T_t^{1,2}(x)) - //_1^{\eta^{1,2}(x)}\nabla\Phi(\mu_t^1,x)\Big|^2 \mu_t^1(dx)\Big)^{1/2},$$

which is again dominated, using Lipschitz condition (4.6.1), by

$$\kappa\, W_2^2(\mu_t^1,\mu_t^2).$$

Now using Gronwall lemma, we complete the proof.

□

Chapter 5

Stochastic Parallel Transport and $Q-$Wiener Process

Generally, one needs to construct stochastic parallel translation if one wants to intrinsically construct Brownian motion on a Riemannian manifold. Therefore, we will study stochastic parallel translation problem on the Wasserstein space in this chapter. First, we review some differential calculus on the Wasserstein space. Let M be a connected compact Riemannian manifold. For any gradient vector field $\nabla\psi$ on M with $\psi \in C^\infty(M)$, we consider the ordinary differential equation (ODE):

$$\frac{d}{dt}U_t(x) = \nabla\psi(U_t(x)), \quad U_0(x) = x \in M.$$

Then $x \to U_t(x)$ is a flow of diffeomorphisms on M. Let $\mu \in \mathbb{P}_2(M)$, and $\mu_t = (U_t)_\#\mu$. It is obvious that for $f \in C^1(M)$ and any $t \in [0,1]$,

$$\frac{d}{dt}\int_M f(x)\,\mu_t(dx) = \frac{d}{dt}\int_M f(U_t(x))\,d\mu(x) = \int_M \langle\nabla f, \nabla\psi\rangle\,\mu_t(dx).$$

We say that the intrinsic derivatives of $\{\mu_t; t \in [0,1]\}$ at the time t is $\nabla\psi$. In order to make clearly different roles played by $\nabla\psi$, we will use notation V_ψ as in [Lot06] when it is seen as a constant vector field on $\mathbb{P}_2(M)$. Namely we denote

$$\frac{d^I\mu_t}{dt} = V_\psi \in \mathbf{T}_{\mu_t}, \quad t \in [0,1].$$

For a functionnal F on $\mathbb{P}_2(M)$, we say that F is derivable at μ along V_ψ, if the directional derivative

$$(\bar{D}_{V_\psi}F)(\mu) = \left\{\frac{d}{dt}F(\mu_t)\right\}_{|t=0} \quad \text{exists.}$$

We say that the gradient $\bar{\nabla}F(\mu)$ exists in in $\mathbf{T}_\mu$ if for each $\psi \in C^\infty(M)$, $(\bar{D}_{V_\psi}F)(\mu)$ exists and

$$\bar{D}_{V_\psi}F(\mu) = \langle \bar{\nabla}F, V_\psi \rangle_{\mathbf{T}_\mu}. \tag{5.0.1}$$

The main purpose of this work is to develop Itô stochastic calculus on $\mathbb{P}_2(M)$; to this end, we will need the differential calculus of order 2. Following J. Lott [Lot06], the covariant derivative $\bar{\nabla}_{V_{\psi_1}}V_{\psi_2}$ associated to the Levi-Civita connection on $\mathbb{P}_2(M)$ is defined by

$$\begin{aligned} 2\langle \bar{\nabla}_{V_{\psi_1}}V_{\psi_2}, V_{\psi_3}\rangle_{\mathbf{T}_\mu} &= \bar{D}_{V_{\psi_1}}\langle V_{\psi_2}, V_{\psi_3}\rangle_{\mathbf{T}_\mu} + \bar{D}_{V_{\psi_2}}\langle V_{\psi_3}, V_{\psi_1}\rangle_{\mathbf{T}_\mu} - \bar{D}_{V_{\psi_3}}\langle V_{\psi_1}, V_{\psi_2}\rangle_{\mathbf{T}_\mu} \\ &\quad + \langle V_{\psi_3}, [V_{\psi_1}, V_{\psi_2}]\rangle_{\mathbf{T}_\mu} - \langle V_{\psi_2}, [V_{\psi_1}, V_{\psi_3}]\rangle_{\mathbf{T}_\mu} - \langle V_{\psi_1}, [V_{\psi_2}, V_{\psi_3}]\rangle_{\mathbf{T}_\mu}. \end{aligned}$$

A few computation yields the formula (see [Lot06] and [DF21])

$$\langle \bar{\nabla}_{V_{\psi_1}}V_{\psi_2}, V_{\psi_3}\rangle_{\mathbf{T}_\mu} = \int_M \langle \nabla^2\psi_2, \nabla\psi_1 \otimes \nabla\psi_3\rangle\, d\mu, \tag{5.0.2}$$

or

$$(\bar{\nabla}_{V_{\psi_1}}V_{\psi_2})(\mu) = \Pi_\mu(\nabla_{\nabla\psi_1}\nabla\psi_2), \tag{5.0.3}$$

where $\Pi_\mu : L^2(M, \mathbf{T}M; \mu) \to \mathbf{T}_\mu$ is the orthogonal projection.

For a functional F on $\mathbb{P}_2(M)$, we say that the Hessian $\bar{\nabla}^2F(\mu) \in \mathbf{T}_\mu \otimes \mathbf{T}_\mu$ exists if for any $\psi_1 \in C^\infty(M)$, $\bar{\nabla}_{V_{\psi_1}}\bar{\nabla}F$ exists and

$$\langle \bar{\nabla}_{V_{\psi_1}}\bar{\nabla}F,\ V_{\psi_2}\rangle_{\mathbf{T}_\mu} = \langle \bar{\nabla}^2F, V_{\psi_1} \otimes V_{\psi_2}\rangle_{\mathbf{T}_\mu\otimes\mathbf{T}_\mu}, \quad \text{for any } \psi_2 \in C^\infty(M).$$

The following three examples of functionals on $\mathbb{P}_2(M)$ will play the role of test functions.

Example 1. Let $\varphi \in C^1(M)$ andF_φ defined by

$$F_\varphi(\mu) = \int_M \varphi(x)\, d\mu(x). \tag{5.0.4}$$

We have

$$\left\{\frac{d}{dt}F_\varphi(\mu_t)\right\}_{|t=0} = \int_M \langle \nabla\varphi(x), \nabla\psi(x)\rangle\, d\mu(x) = \langle V_\varphi, V_\psi\rangle_{\mathbf{T}_\mu}.$$

Therefore the gradient $\bar{\nabla}F_\varphi$ of F_φ is equal to V_φ. According to (5.0.2), we have

$$\langle \bar{\nabla}^2F_\varphi, V_{\psi_1} \otimes V_{\psi_2}\rangle_{\mathbf{T}_\mu\otimes\mathbf{T}_\mu} = \int_M \langle \nabla^2\varphi, \nabla\psi_1 \otimes \nabla\psi_2\rangle\, d\mu, \quad \psi_1, \psi_2 \in C^\infty(M).$$

Example 2. The entropy functional $F(\mu) = \mathrm{Ent}(\mu) = \int_M \rho \ln(\rho)\, dx$ for $d\mu = \rho\, dx$.

Let $d\mu_0 = \rho_0(x)\, dx$ and define $\mu_t = (U_t)_\#\mu_0$. Then $d\mu_t = \rho_t(x)\, dx$ with $\rho_t = \rho_0(U_{-t})K_t$ where

$$K_t = \exp\Big(-\int_0^t \mathrm{div}(\nabla\psi)(U_{-s})\, ds\Big).$$

We have

$$K_t(U_t) = \exp\Big(-\int_0^t (\Delta\psi)(U_{t-s})\, ds\Big).$$

It follows that, if $\rho_0 \in C^1(M)$ with $\rho_0 > 0$,

$$\langle \bar\nabla \mathrm{Ent},\ V_\psi\rangle_{\mathbf{T}_{\mu_0}} = -\int_M \Delta\psi\, \rho_0\, dx = \int_M \langle \nabla\psi, \nabla \ln(\rho_0)\rangle\, \mu_0(dx). \tag{5.0.5}$$

Therefore at such a measure μ_0, the gradient $\bar\nabla\mathrm{Ent}$ of Ent exists and

$$\bar\nabla \mathrm{Ent}(\mu_0) = V_{\ln(\rho_0)}.$$

The Hessian of Ent was first heuristically computed in [OV00], it is profoundly related to the Ricci curvature of M. We have, by (5.0.5),

$$\langle \bar\nabla \mathrm{Ent},\ V_\psi\rangle_{\mathbf{T}_{\mu_t}} = -\int_M \Delta\psi\, \rho_t\, dx = -\int_M \Delta\psi(U_t)\rho_0\, dx.$$

Taking the derivative with respect to t, at $t = 0$, we get the following expression for the Lie derivative of order 2:

$$(\bar D_{V_\psi}\bar D_{V_\psi}\mathrm{Ent})(\mu_0) = \frac{d}{dt}_{|t=0} \langle \bar\nabla \mathrm{Ent},\ V_\psi\rangle_{\mathbf{T}_{\mu_t}} = -\int_M \langle \nabla\Delta\psi(x), \nabla\psi(x)\rangle \mu_0(dx). \tag{5.0.6}$$

Next example comes from the framework of particle system (see [LWZ21]).

Example 3.

$$F_3(\mu) = \int_{M\times M} W(x, y)\mu(dx)\mu(dy),$$

where $W \in C^2(M \times M)$.

Let $\mu_t = (U_t)_\#\mu_0$. We have

$$F_3(\mu_t) = \int_{M\times M} W(U_t(x), U_t(y))\mu(dx)\mu(dy).$$

Taking the derivative with respect to t, at $t = 0$, we get

$$\frac{d}{dt}_{|t=0} F_3(\mu_t) = \int_{M\times M} \Big(\langle \nabla_1 W(x, y), \nabla\psi(x)\rangle + \langle \nabla_2 W(x, y), \nabla\psi(y)\rangle\Big)\mu(dx)\mu(dy), \tag{5.0.7}$$

where ∇_1 denotes the partial gradient with respect to the first component, while ∇_2 for the second component. Let $\Phi(x,\mu)=\int_M \Big(W(x,y)+W(y,x)\Big)\,\mu(dy)$; then we have

$$\bar{D}_{V_\psi}F_3(\mu)=\int_M \langle\nabla\Phi(x,\mu),\nabla\psi(x)\rangle\,\mu(dx).$$

Therfore the gradient $\bar{\nabla}F_3(\mu)$ exists and

$$\bar{\nabla}F_3(\mu)=V_{\Phi_\mu},\quad \Phi_\mu(x)=\Phi(x,\mu).$$

We will compute the Hessian $\bar{\nabla}^2F_3$ of F_3. Denote

$$\tilde{W}(x,y)=\langle\nabla_1W(x,y),\nabla\psi(x)\rangle+\langle\nabla_2W(x,y),\nabla\psi(y)\rangle.$$

Then $\bar{D}_{V_\psi}F_3(\mu)=\displaystyle\int_{M\times M}\tilde{W}(x,y)\mu(dx)\mu(dy)$. Using (5.0.7), we have

$$\bar{D}_{V_\psi}\bar{D}_{V_\psi}F_3(\mu)=\int_{M\times M}\Big(\langle\nabla_1\tilde{W}(x,y),\nabla\psi(x)\rangle+\langle\nabla_2\tilde{W}(x,y),\nabla\psi(y)\rangle\Big)\mu(dx)\mu(dy).$$

We have

$$\begin{aligned}\langle\nabla_1\tilde{W}(x,y),\nabla\psi(x)\rangle&=\langle\nabla_1^2W(x,y),\nabla\psi(x)\otimes\nabla\psi(x)\rangle\\&\quad+\langle\nabla_1W(x,y),\nabla_{\nabla\psi(x)}\nabla\psi(x)\rangle+\langle\nabla_1\nabla_2W(x,y),\nabla\psi(x)\otimes\nabla\psi(y)\rangle,\end{aligned}$$

and

$$\begin{aligned}\langle\nabla_2\tilde{W}(x,y),\nabla\psi(y)\rangle&=\langle\nabla_2^2W(x,y),\nabla\psi(y)\otimes\nabla\psi(y)\rangle\\&\quad+\langle\nabla_2W(x,y),\nabla_{\nabla\psi(y)}\nabla\psi(y)\rangle+\langle\nabla_2\nabla_1W(x,y),\nabla\psi(x)\otimes\nabla\psi(y)\rangle.\end{aligned}$$

Combing these two terms, we get

$$\begin{aligned}&\int_{M\times M}\langle\nabla_1\tilde{W}(x,y),\nabla\psi(x)\rangle+\langle\nabla_2\tilde{W}(x,y),\nabla\psi(y)\rangle\mu(dx)\mu(dy)\\&=\int_{M\times M}\mathrm{Hess}_{x,y}W(\nabla\psi(x),\nabla\psi(y))\mu(dx)\mu(dy)\\&+\int_{M\times M}\langle\nabla_1W(x,y)+\nabla_2W(y,x),\nabla_{\nabla\psi(x)}\nabla\psi(x)\rangle\mu(dx)\mu(dy).\end{aligned}$$

Note that

$$\nabla\Phi(x,\mu)=\int_M\Big(\nabla_1W(x,y)+\nabla_2W(y,x)\Big)\,\mu(dy).$$

By (5.0.2), we have

$$\langle \bar{\nabla} F_3, \bar{\nabla}_{V_\psi} V_\psi \rangle = \int_M \langle \nabla \Phi(x,\mu), \nabla_{\nabla\psi(x)} \nabla\psi(x) \rangle \mu(dx).$$

Proposition 5.0.1. *We have*

$$\langle \bar{\nabla}^2 F_3, V_\psi \otimes V_\psi \rangle = \int_{M\times M} \textit{Hess}_{x,y} W(\nabla\psi(x), \nabla\psi(y)) \mu(dx)\mu(dy). \tag{5.0.8}$$

In Chapter 4, some elements of differential geometry of the Wassertein space $\mathbb{P}_2(M)$ were revisited in order to construct the parallel translation in an intrinsic way; namely, a vector field along a regular curve in $\mathbb{P}_2(M)$ was enlarged into a vector field defined on the whole space, so that the parallel translation was introduced as in the classical differential geometry. We have to note that the equation for parallel translations was stated in [Lot06], but no existence result was provided. In [AG08], the authors considered regular curves $\{\mu_t; t \in [0,1]\}$ generated by a flow of Lipschitz maps and proved the existence of parallel translations $\{V_{\Psi_t}; t \in [0,1]\}$ along such a regular curve in L^2. The method used in [AG08] is extrinsic and solutions to Lott's equation for parallel translations is in a weak sense. In the paper [Lot17], Lott proposed an intrinsic construction for parallel translation along geodesics in $\mathbb{P}_2(M)$, also a weak result of existence was obtained. To our knowledge, the existence of strong solutions to Lott's equation remains unsolved.

In this chapter, we will consider stochastic regular curves in $\mathbb{P}_2(M)$, which are generated by stochastic flows of diffeomorphisms; the main purpose is to construct stochastic parallel translations along them. The involvement of the Brownian motion arises a basic difficulty, that is, the path of diffusion process is only Hölder of exponent less than $1/2$: the method in [AG08] does not work. On the other hand, the limit theorem developed in [Bis81, Mal97, IW81] provides a powerful tool in stochastic analysis on Riemannian manifolds, we will do some tentatives in this direction. Let's now explain a bit the content of this chapter. In section 5.1, we first state main results obtained in the literature. Since the orthogonal projection plays a fundamental role in our work, we will make a brief study on it: a representation formula is obtained, and its evolution along an absolutely continuous curve in $\mathbb{P}_2(M)$ is studied. In Section 5.2, we will establish an intrinsic formalism for Itô stochastic calculus on $\mathbb{P}_2(M)$: Itô formula is proved throughout three functionals; it takes the form as on a Riemannian manifold, much simpler than those previously obtained in [BLPR17, Wan21]; stochastic differential equations on $\mathbb{P}_2(M)$ with a finite number of Brownian motions are also considered. Section 5.3 is devoted to find, in more or less formal way, a suitable weak form and a strong form of stochastic partial differential equations for parallel translations along stochastic regular curves in $\mathbb{P}_2(M)$; concerning the strong solution, the preservation of norms is proved. The purpose of Section 5.4 is to introduce an infinite numbers of noises in order to construct nondegenerated diffusion processes in $\mathbb{P}_2(M)$; to this end, we will use eigenfunctions of the Laplace operator on M. Finally, in Section 6, we deal with the case of $\mathbb{P}_2(\mathbb{T})$, the Wasserstein space over the torus: we prove the existence of strong solutions to J. Lott's equation for parallel translations, as well as the existence of strong stochastic parallel translations.

5.1 Regular curves and parallel translations on $\mathbb{P}_2(M)$

Let's first show the state of art for parallel translations in the Wasserstein space $\mathbb{P}_2(M)$. Let $\{c_t;\ t \in [0,1]\}$ be an absolutely continuous curve in $\mathbb{P}_2(M)$ and $\{Y_t;\ t \in [0,1]\}$ a family of vector fields along $\{c_t;\ t \in [0,1]\}$, that is $Y_t \in \mathbf{T}_{c_t}$. Suppose there are smooth functions $(t,x) \to \Phi_t(x)$ and $(t,x) \to \Psi_t(x)$ such that

$$\frac{d^I c_t}{dt} = V_{\Phi_t}, \quad Y_t = V_{\Psi_t},$$

Lott obtained formally in [Lot06] that if $\{Y_t;\ t \in [0,1]\}$ is parallel along $\{c_t;\ t \in [0,1]\}$, then $\{\nabla\Psi_t;\ t \in [0,1]\}$ is a solution to the following linear partial differential equation:

$$\frac{d}{dt}\nabla\Psi_t + \Pi_{c_t}\Big(\nabla_{\nabla\Phi_t}\nabla\Psi_t\Big) = 0, \tag{5.1.1}$$

where Π_{c_t} is the orthogonal projection to $\mathbf{T}_{c_t}$. Up to now, only two classes of absolutely continuous curves have been considered in the literature: regular curves generated by a flow of Lipschitz maps in [AG08], geodesics of $\mathbb{P}_2(M)$ in [Lot17].

To introduce regular curves, we consider the flow of diffeomorphisms defined by the following ODE

$$dX_{t,s} = \nabla\phi(t, X_{t,s})\, dt, \quad t \geq s, X_s(x) = x,$$

where $(t,x) \to \phi(t,x)$ is a smooth enough function. Let $c_t = (X_{t,0})_\# c_0$ with $dc_0(x) = \rho_0\, dx$ and $\rho_0 > 0$. The following result mimics section 5 in [AG08] and was proved in [DF21].

Theorem 5.1.1. *For any $\nabla\Psi_0 \in L^2(c_0)$, there is a unique weak solution $\{\nabla\Psi_t, t \in [0,1]\}$ in the sense that $V_{\Psi_t} \in \mathbf{T}_{c_t}$ and*

$$\Pi_{c_t}\Big(\lim_{\varepsilon\downarrow 0} \frac{\tau_\varepsilon^{-1}\,\nabla\Psi_{t+\varepsilon}(X_{t+\varepsilon,t}) - \nabla\Psi_t}{\varepsilon}\Big) = 0 \tag{5.1.2}$$

holds in $L^2(c_t)$ for almost all $t \in [0,1]$, where τ_ε is the parallel translation along $\{s \to X_{t+s,t}, s \in [0,\varepsilon]\}$, that is equivalent to say that $t \to \nabla\Psi_t$ is absolutely continuous and

$$\frac{d}{dt}\int_M \langle\nabla f, \nabla\Psi_t\rangle\, c_t(dx) = \int_M \langle\nabla^2 f, \nabla\phi(t,\cdot)\otimes\nabla\Psi_t\rangle\, c_t(dx), \quad f \in C^\infty(M). \tag{5.1.3}$$

Even in this case, the well-posedness of (5.1.1) is not yet established to our knowledge, the implication of (5.1.1) as well as (5.1.2) to (5.1.3) is obvious. However, for the case of geodesics, it requires some investigation for this implication, see [Lot17]. In [DF21], it was proved if for any t, $\Psi_t \in H^k(M)$ with $k > \frac{\dim(M)}{2} + 2$, then Ψ_t admits an extension $(\mu, x) \to \tilde{\Psi}(\mu, x)$ defined on $\mathbb{P}_2(M) \times M$ such that for any μ, $\tilde{\Psi}(\mu,\cdot) \in H^k(M)$ and $\Big(\bar{\nabla}_{V_{\phi(t,\cdot)}} V_{\tilde{\Psi}}\Big)(c_t) = 0$, that is the classical definition for parallel translation in differential geometry.

Since the projection $\Pi_\mu : L^2(M, \mathbf{T}M; \mu) \to \mathbf{T}_\mu$ is basically involved in our work, it will be useful to make a study on it. Let $d\mu = \rho\, dx$ with a smooth density $\rho > 0$, recall that for a vector field ζ on M,

$$\mathrm{div}_\mu(\zeta) = \mathrm{div}(\zeta) + \langle \nabla \log \rho, \zeta \rangle$$

and for a function $f \in C^2$, $\mathcal{L}^\mu f = \mathrm{div}_\mu(\nabla f)$ has the expression

$$\mathcal{L}^\mu f = \Delta f + \langle \nabla \log \rho, \nabla f \rangle.$$

It is well-known that $\mathcal{L}^\mu$ has discrete spectrum of eigenvalue $\lambda_n^\mu \sim n^{2/\dim(M)}$. Consider the equation, for a given g such that $\int_M g\, \mu(dx) = 0$,

$$\Delta f + \langle \nabla \log \rho, \nabla f \rangle = g.$$

By Shauder estimate for elliptic operators, if $\nabla \log \rho$ is in $C^{q,\alpha}$, then for $g \in C^{q,\alpha}$, the solution f to $\mathcal{L}^\mu f = g$ is in the class $C^{q+2,\alpha}$. For a regular vector field ζ on M, by Hodge decomposition (see for example [Li09]), there exists a function β and a vector field B of $\mathrm{div}_\mu(B) = 0$ such that $\zeta = \nabla\beta + B$; therefore $\mathrm{div}_\mu(\zeta) = \mathcal{L}^\mu(\beta)$ and

$$\Pi_\mu(\zeta) = \nabla\, (\mathcal{L}^\mu)^{-1}\big(\mathrm{div}_\mu(\zeta)\big). \tag{5.1.4}$$

We will get a representation formula for Π_μ. Let $T_s^\mu = e^{s\mathcal{L}^\mu}$ be the semi-group associated to $\mathcal{L}^\mu$, then $(\mathcal{L}^\mu)^{-1} = \int_0^{+\infty} T_s^\mu\, ds$ and (5.1.4) becomes

$$\Pi_\mu(\zeta) = \int_0^{+\infty} \nabla\, T_s^\mu\big(\mathrm{div}_\mu(\zeta)\big)\, ds. \tag{5.1.5}$$

To insure the convergence in (5.1.5), we have to introduce a modified De Rham-Hodge operator $\square^\mu$ on differential 1-forms. As usual, for a vector field A on M, we denote by $A^\flat$ the associated differential form and for a differential 1-form ω, we denote by $\omega^\#$ the associated vector field. Define $\delta_\mu(\omega) = -\mathrm{div}_\mu(\omega^\#)$ and d_μ^* the dual operator of exterior derivative d, that is

$$\int_M \langle d_\mu^* \sigma, \omega \rangle_{\Lambda^1}\, d\mu = \int_M \langle \sigma, d\omega \rangle_{\Lambda^2}\, d\mu.$$

Let $\square^\mu = d\delta_\mu + d_\mu^* d$. Then the following commutation formula holds: $d\, e^{s\mathcal{L}^\mu} f = e^{-s\square^\mu}(df)$. Note now

$$\square^\mu(df) = d\delta_\mu(df) = \square(df) + i_{\nabla V}(df),$$

where we denote for a moment $V = \log \rho$ and $i_{\nabla V}$ denotes the inner product by ∇V. By Cartan formula: $\mathcal{L}_{\nabla V} = i_{\nabla V} d + d i_{\nabla V}$, we get

$$i_{\nabla V}(df) = \mathcal{L}_{\nabla V}(df) = \nabla_{\nabla V}(df) + \langle \nabla^2 V, df\rangle.$$

Therefore, $\omega_s = dT^\mu_s f$ is a solution to the following heat equation:

$$\frac{d\omega_t}{dt} = -\square\omega_t - \langle \nabla^2 V, \omega_t\rangle.$$

Let $\{A_1, \ldots, A_m\}$ be a family of vector fields on M such that $\sum_{i=1}^m \mathcal{L}^2_{A_i} = \Delta$ and $\sum_{i=1}^m \nabla_{A_i} A_i = 0$. Let Y_s be the solution to the following SDE on M

$$dY^\rho_s = \sqrt{2}\sum_{i=1}^m A_i(Y^\rho_s)\circ dW^i_s + \nabla\log(\rho)(Y^\rho_s)\,ds, \tag{5.1.6}$$

where $s \to (W^1_s, \ldots, W^m_s)$ is a standard Brownian motion on a probability space $(\Omega, \mathbb{P})$. Then $T^\mu_s f(x) = \mathbb{E}\big(f(Y^\rho_s(x))\big)$. Let

$$\mathrm{Ric}^\mu = \mathrm{Ric} - \nabla^2(\log\rho), \tag{5.1.7}$$

and Q^μ_s be the resolvent defined by

$$\frac{dQ^\mu_s}{ds} = \mathrm{Ric}^\mu_{Y^\rho_s}\, Q^\mu_s.$$

It is well-known that the following representation formula holds

$$\langle e^{-s\square^\mu} df, A\rangle = \mathbb{E}\Big(\langle df(Y^\rho_s), Q^\mu_s A\rangle\Big), \quad A \in \chi(M).$$

Proposition 5.1.2. *We have*

$$\Pi_\mu(\zeta) = \int_0^{+\infty} \mathbb{E}\Big((Q^\mu_s)^*(\nabla \mathsf{div}_\mu(\zeta)_{Y^\rho_s})\Big)\,ds. \tag{5.1.8}$$

Hence the dependence $\mu \to \Pi_\mu$ is good in the class of probability measures having C^2 positive density.

Theorem 5.1.3. *For a smooth vector field ζ on M, $t \to \Pi_{c_t}(\zeta)$ is absolutely continuous and*

$$\frac{d}{dt}\Pi_{c_t}(\zeta) = -\Pi_{c_t}\Big(\mathcal{L}^{c_t}(\phi(t,\cdot))\big(\zeta - \Pi_{c_t}(\zeta)\big)\Big). \tag{5.1.9}$$

Proof. The density ρ_t of c_t with respect to c_0 admits the expression (see [Cru83, Kun97])

$$\rho_t(x) = \exp\Big[\int_0^t \mathsf{div}_{c_0}(\nabla\phi)\big(s, X_{s,t}(x)\big)\,ds\Big].$$

Under the condition that the density ρ_0 of c_0 is in class C^3, it is easy to see that $t \to \log \rho_t$ is continuous from $[0,1]$ to $C^2(M)$. Now replacing $\nabla \log \rho$ by $\nabla \log \rho_t$ in (5.1.6) and using the dependence of SDE, combining with definition Ric^{c_t} in (5.1.7), we get the absolute continuity of $t \to \Pi_{c_t}(\zeta)$. We will use the following equation for ρ_t

$$\frac{d}{dt}\rho_t = -\mathrm{div}_{c_t}(\nabla\phi(t,\cdot))\rho_t = -\mathcal{L}^{c_t}(\phi(t,\cdot))\,\rho_t. \tag{5.1.10}$$

Let $f \in C^\infty(M)$, we have $\int_M \langle \nabla f, \zeta\rangle\, c_t(dx) = \int_M \langle \nabla f, \Pi_{c_t}(\zeta)\rangle\, c_t(dx)$ or

$$\int_M \langle \nabla f, \zeta\rangle\, \rho_t\, c_0(dx) = \int_M \langle \nabla f, \Pi_{c_t}(\zeta)\rangle\, \rho_t\, c_0(dx).$$

Taking the derivative with respect to t and using (5.1.10), we get

$$\begin{aligned}-\int_M \langle \nabla f, \zeta\rangle\, \mathcal{L}^{c_t}(\phi(t,\cdot))\rho_t\, c_0(dx) = &-\int_M \langle \nabla f, \Pi_{c_t}(\zeta)\rangle\, \mathcal{L}^{c_t}(\phi(t,\cdot))\rho_t\, c_0(dx)\\ &+\int_M \langle \nabla f, \frac{d}{dt}\Pi_{c_t}(\zeta)\rangle\, \rho_t\, c_0(dx).\end{aligned}$$

The result (5.1.9) follows.

□

Proposition 5.1.4. *Let ζ be a smooth vector field on M, $\{\Psi_t;\ t\in[0,1]\}$ be a parallel translation along $\{c_t;\ t\in[0,1]\}$ given in Theorem 5.1.1, then*

$$\begin{aligned}\frac{d}{dt}\int_M \langle \zeta, \nabla\Psi_t\rangle\, c_t(dx) = &-\int_M \langle \mathcal{L}^{c_t}(\phi(t,\cdot))\Pi_{c_t}^{\perp}(\zeta),\ \nabla\Psi_t\rangle\, c_t(dx)\\ &+\int_M \langle \nabla_{\nabla\phi(t,\cdot)}\big(\Pi_{c_t}(\zeta)\big), \nabla\Psi_t\rangle\, c_t(dx),\end{aligned} \tag{5.1.11}$$

where $\Pi_{c_t}^{\perp}(\zeta) = \zeta - \Pi_{c_t}(\zeta)$.

Proof. Let $I_t = \int_M \langle \Pi_{c_t}(\zeta),\ \nabla\Psi_t\rangle\, c_t(dx)$. We have, for $\varepsilon > 0$,

$$I_{t+\varepsilon} = \int_M \langle \Pi_{c_{t+\varepsilon}}(\zeta),\ \nabla\Psi_{t+\varepsilon}\rangle\, c_{t+\varepsilon}(dx) = \int_M \langle \tau_\varepsilon^{-1}\Pi_{c_{t+\varepsilon}}(\zeta),\ \tau_\varepsilon^{-1}\nabla\Psi_{t+\varepsilon}\rangle(X_{t+\varepsilon,t})\, c_t(dx).$$

Then

$$\begin{aligned}I_{t+\varepsilon}-I_t &= \int_M \langle \tau_\varepsilon^{-1}\Pi_{c_{t+\varepsilon}}(\zeta)(X_{t+\varepsilon,t})-\Pi_{c_t}(\zeta)(x),\ \tau_\varepsilon^{-1}\nabla\Psi_{t+\varepsilon}(X_{t+\varepsilon,t})\rangle\, c_t(dx)\\ &\quad+\int_M \langle \Pi_{c_t}(\zeta),\ \tau_\varepsilon^{-1}\nabla\Psi_{t+\varepsilon}(X_{t+\varepsilon,t})-\nabla\Psi_t(x)\rangle\, c_t(dx)=J_\varepsilon^1+J_\varepsilon^2\end{aligned}$$

respectively. As $\varepsilon\to 0$, $\tau_\varepsilon^{-1}\nabla\Psi_{t+\varepsilon}(X_{t+\varepsilon,t})$ converges to $\nabla\Psi_t(x)$ and the term $J_\varepsilon^2/\varepsilon$ converges to 0 according to (5.1.2). For J_ε^1, note that

$$\begin{aligned}&\frac{1}{\varepsilon}\Big(\tau_\varepsilon^{-1}\Pi_{c_{t+\varepsilon}}(\zeta)(X_{t+\varepsilon,t})-\Pi_{c_t}(\zeta)(x)\Big)\\ &=\frac{1}{\varepsilon}\Big(\tau_\varepsilon^{-1}\Pi_{c_{t+\varepsilon}}(\zeta)(X_{t+\varepsilon,t})-\tau_\varepsilon^{-1}\Pi_{c_t}(\zeta)(X_{t+\varepsilon,t})\Big)+\frac{1}{\varepsilon}\Big(\tau_\varepsilon^{-1}\Pi_{c_t}(\zeta)(X_{t+\varepsilon,t})-\Pi_{c_t}(\zeta)(x)\Big).\end{aligned}$$

As $\varepsilon\to 0$, the last term converges to $\nabla_{\phi(t,\cdot)}\Pi_{c_t}(\zeta)$, while

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\Big(\Pi_{c_{t+\varepsilon}}(\zeta)(X_{t+\varepsilon,t})-\Pi_{c_t}(\zeta)(X_{t+\varepsilon,t})\Big)=\frac{d}{dt}\Pi_{c_t}(\zeta)(x).$$

Now using (5.1.9), we obtain (5.1.11). □

5.2 Itô stochastic calculus on $\mathbb{P}_2(M)$

We will introduce stochastic regular curves $\{\mu_t;\ t\in[0,1]\}$ on $\mathbb{P}_2(M)$ and establish Itô formula for them. Let $\{X_{t,s}, t\geq s\}$ be a stochastic flow of diffeomorphisms defined by the following Stratanovich stochastic differential equation (SDE) on M:

$$dX_{t,s}=\sum_{i=0}^N \nabla\phi_i(t,X_{t,s})\circ dB_t^i,\quad t\geq s;\quad X_{s,s}(x)=x, \tag{5.2.1}$$

where $dB_t^0=dt$, $(B_t^1,\ldots,B_t^N)$ is a Standard Brownian motion on $\mathbb{R}^N$ and $(t,x)\to\phi_i(t,x)$ is smooth enough for $i=0,1,\ldots,N$. Let $\mu_t(\omega)=(X_{t,0})_{\#}\mu$. Then for $F_\varphi(\mu)=\int_M\varphi\,d\mu$ with $\varphi\in C^2(M)$, $t\to F_\varphi(\mu_t)$ is a real valued semi-martingale. The Itô differential $\circ d_t\, F_\varphi(\mu_t)$ admits the expression:

$$\begin{aligned}\circ d_t\, F_\varphi(\mu_t) &= d_t\int_M \varphi(X_{t,0})\,d\mu=\sum_{i=0}^N\Big(\int_M\langle\nabla\varphi,\nabla\phi_i(t,\cdot)\rangle\,d\mu_t\Big)\circ dB_t^i\\ &=\sum_{i=0}^N\langle V_\varphi,V_{\phi_i(t,\cdot)}\rangle_{\mathbf{T}_{\mu_t}}\circ dB_t^i.\end{aligned}$$

Definition 5.2.1. *We will say that the intrinsic Itô stochastic differential of μ_t, denoted by $\circ d_t^I \mu_t$, admits the following expression*

$$\circ d_t^I \mu_t = \sum_{i=0}^{N} V_{\phi_i(t,\cdot)} \ \circ dB_t^i. \tag{5.2.2}$$

Then using this notation, $\circ d_t\, F_f(\mu_t)$ can be rewritten in the form:

$$\circ d_t\, F_\varphi(\mu_t) = \langle \bar{\nabla} F_\varphi,\ \circ d_t^I \mu_t \rangle_{\mathbf{T}_{\mu_t}},$$

the last term can be symbolically read as inner product in $\mathbf{T}_{\mu_t}$. We will establish Itô formula for such a stochastic process $\{\mu_t;\ t \in [0,1]\}$ on $\mathbb{P}_2(M)$. The Itô form of SDE (5.2.1) is the following

$$dX_{t,s} = \sum_{i=0}^{N} \nabla \phi_i(t, X_{t,s})\, dB_t^i + \frac{1}{2} \sum_{i=1}^{N} \Big(\nabla_{\nabla \phi_i(t,\cdot)} \nabla \phi_i(t,\cdot) \Big)(X_{t,s})\, dt. \tag{5.2.3}$$

First of all, we consider the functional $F_\varphi(\mu) = \int_M \varphi\, d\mu$. By Itô formula,

$$\begin{aligned} d_t\varphi(X_{t,0}) = &\sum_{i=0}^{N} \langle \nabla\varphi, \nabla\phi_i(t,\cdot)\rangle(X_{t,0})\, dB_t^i + \frac{1}{2}\sum_{i=1}^{N} \langle \nabla\varphi, \nabla_{\nabla\phi_i(t,\cdot)}\nabla\phi_i(t,\cdot)\rangle(X_{t,0})\, dt \\ &+ \frac{1}{2}\sum_{i=1}^{N} \langle \nabla^2\varphi, \nabla\phi_i(t,\cdot)\otimes\nabla\phi_i(t,\cdot)\rangle(X_{t,0})\, dt. \end{aligned}$$

Then

$$d_t F_\varphi(\mu_t) = \sum_{i=0}^{N} \Big(\int_M \langle \nabla\varphi, \nabla\phi_i(t,\cdot)\rangle\, d\mu_t \Big)\, dB_t^i + \frac{1}{2}\sum_{i=1}^{N} \Big(\int_M \mathcal{L}_{\nabla\phi_i(t,\cdot)} \mathcal{L}_{\nabla\phi_i(t,\cdot)} \varphi\, d\mu_t \Big)\, dt. \tag{5.2.4}$$

According to [Lot06] or (5.0.2) or (5.0.3), we have

$$\int_M \langle \nabla\varphi, \nabla_{\nabla\phi_i(t,\cdot)}\nabla\phi_i(t,\cdot)\rangle\, d\mu_t = \langle \bar{\nabla} F_\varphi, \bar{\nabla}_{V_{\phi_i}(t,\cdot)} V_{\phi_i(t,\cdot)} \rangle_{\mathbf{T}_{\mu_t}},$$

and

$$\int_M \langle \nabla^2\varphi, \nabla\phi_i(t,\cdot)\otimes\nabla\phi_i(t,\cdot)\rangle\, d\mu_t = \langle \bar{\nabla}^2 F_\varphi, V_{\phi_i(t,\cdot)} \otimes V_{\phi_i(t,\cdot)} \rangle_{\mathbf{T}_{\mu_t}\otimes\mathbf{T}_{\mu_t}}.$$

In other words,

$$d_t F_\varphi(\mu_t) = \sum_{i=0}^{N} \langle \bar{\nabla} F_\varphi, V_{\phi_i(t,\cdot)} \rangle_{\mathbf{T}_{\mu_t}} \, dB_t^i + \frac{1}{2} \sum_{i=1}^{N} \langle \bar{\nabla} F_\varphi, \bar{\nabla}_{V_{\phi_i(t,\cdot)}} V_{\phi_i(t,\cdot)} \rangle_{\mathbf{T}_{\mu_t}} \, dt$$
$$+ \frac{1}{2} \sum_{i=1}^{N} \langle \bar{\nabla}^2 F_\varphi, V_{\phi_i(t,\cdot)} \otimes V_{\phi_i(t,\cdot)} \rangle_{\mathbf{T}_{\mu_t} \otimes \mathbf{T}_{\mu_t}} \, dt.$$

Remark that

$$\bar{D}_{V_{\phi_i(t,\cdot)}} \bar{D}_{V_{\phi_i(t,\cdot)}} F_\varphi = \langle \bar{\nabla} F_\varphi, \bar{\nabla}_{V_{\phi_i(t,\cdot)}} V_{\phi_i(t,\cdot)} \rangle_{\mathbf{T}_{\mu_t}} + \langle \bar{\nabla}^2 F_\varphi, V_{\phi_i(t,\cdot)} \otimes V_{\phi_i(t,\cdot)} \rangle_{\mathbf{T}_{\mu_t} \otimes \mathbf{T}_{\mu_t}}.$$

So we get the following Itô formula:

$$d_t F_\varphi(\mu_t) = \sum_{i=0}^{N} \langle \bar{\nabla} F_\varphi, V_{\phi_i(t,\cdot)} \rangle_{\mathbf{T}_{\mu_t}} \, dB_t^i + \frac{1}{2} \sum_{i=1}^{N} (\bar{D}_{V_{\phi_i(t,\cdot)}} \bar{D}_{V_{\phi_i(t,\cdot)}} F_\varphi)(\mu_t) \, dt.$$

Proposition 5.2.2. *Let F be a polynomial on $\mathbb{P}_2(M)$, we have*

$$d_t F(\mu_t) = \sum_{i=0}^{N} \langle \bar{\nabla} F, V_{\phi_i(t,\cdot)} \rangle_{\mathbf{T}_{\mu_t}} \, dB_t^i + \frac{1}{2} \sum_{i=1}^{N} (\bar{D}_{V_{\phi_i(t,\cdot)}} \bar{D}_{V_{\phi_i(t,\cdot)}} F)(\mu_t) \, dt. \tag{5.2.5}$$

Proof. For two functionals F and G satisfying Formula (5.2.5), by Itô formula, $d_t(FG)(\mu_t) = d_t F(\mu_t)\, G(\mu_t) + F(\mu_t)\, d_t G(\mu_t) + d_t F(\mu_t) \cdot d_t G(\mu_t)$. Notice that

$$\bar{D}_{V_{\phi_i(t,\cdot)}} \bar{D}_{V_{\phi_i(t,\cdot)}} (FG) = G \bar{D}_{V_{\phi_i(t,\cdot)}} \bar{D}_{V_{\phi_i(t,\cdot)}} F + F \bar{D}_{V_{\phi_i(t,\cdot)}} \bar{D}_{V_{\phi_i(t,\cdot)}} G + 2 \langle \bar{\nabla} F, V_{\phi_i(t,\cdot)} \rangle \cdot \langle \bar{\nabla} G, V_{\phi_i(t,\cdot)} \rangle,$$

and $d_t F(\mu_t) \cdot d_t G(\mu_t) = \sum_{i=1}^{N} \langle \bar{\nabla} F, V_{\phi_i(t,\cdot)} \rangle \cdot \langle \bar{\nabla} G, V_{\phi_i(t,\cdot)} \rangle \, dt$; so Formula (5.2.5) holds true for FG. A polynomial F on $\mathbb{P}_2(M)$ is a finite sum of $F_{\varphi_1} \cdots F_{\varphi_k}$, therefore Formula (5.2.5) remains true. We complete the proof. □

Secondly we deal with the entropy functional in example 3, which is defined for probability measures having positive density. Note that if $d\mu(x) = \rho(x)\, dx$ with $\rho > 0$, the measure μ_t induced by SDE (5.2.3) has a density $\rho_t > 0$ with respect to μ.

Proposition 5.2.3. *The stochastic process $\{\rho_t,\ t \geq 0\}$ satisfies the following SPDE:*

$$d\rho_t = -\sum_{i=0}^{N} \mathrm{div}_\mu(\rho_t \nabla \phi_i(t,\cdot)) \, dB_t^i + \frac{1}{2} \sum_{i=1}^{N} \mathrm{div}_\mu\big(\mathrm{div}_\mu(\rho_t \nabla \phi_i(t,\cdot)) \nabla \phi_i(t,\cdot)\big). \tag{5.2.6}$$

Proof. . We have

$$\int_M \langle \nabla\varphi,\ \nabla\phi_i(t,\cdot)\rangle\ d\mu_t = \int_M \langle \nabla\varphi,\ \rho_t\nabla\phi_i(t,\cdot)\rangle\ d\mu$$
$$= -\int_M \varphi\ \mathsf{div}_\mu\big(\rho_t\nabla\phi_i(t,\cdot)\big)\, d\mu.$$

In the same way, we have

$$\int_M \mathcal{L}_{\nabla\phi_i(t,\cdot)}\mathcal{L}_{\nabla\phi_i(t,\cdot)}\varphi\, d\mu_t = \int_M \varphi\,\mathsf{div}_\mu\big(\mathsf{div}_\mu(\rho_t\nabla\phi_i(t,\cdot))\nabla\phi_i(t,\cdot)\big)\ d\mu.$$

Using $F(\mu_t) = \int_M \varphi\,\rho_t\, d\mu$ and (5.2.4), combined with above equalities, we get (5.2.6). □

Proposition 5.2.4. *We have*

$$d_t\mathrm{Ent}(\mu_t) = \sum_{i=0}^{N} \langle \bar\nabla \mathrm{Ent}, V_{\phi_i(t,\cdot)}\rangle_{\mathbf{T}_{\mu_t}}\ dB_t^i + \frac{1}{2}\sum_{i=1}^{N} (\bar D_{V_{\phi_i(t,\cdot)}} \bar D_{V_{\phi_i(t,\cdot)}}\mathrm{Ent})(\mu_t)\ dt. \tag{5.2.7}$$

Proof. For the functional Ent, we have to take the density ρ_t of μ_t with respect to the Riemannian measure dx; in this case, we use div for the usual divergence. Therefore ρ_t satisfies the relation

$$d\rho_t = -\sum_{i=0}^{N} \mathsf{div}(\rho_t\nabla\phi_i(t,\cdot))\ dB_t^i + \frac{1}{2}\sum_{i=1}^{N} \mathsf{div}\big(\mathsf{div}(\rho_t\nabla\phi_i(t,\cdot))\nabla\phi_i(t,\cdot)\big).$$

It follows that $d\rho_t\cdot d\rho_t = \sum_{i=1}^{N}\Big[\mathsf{div}(\rho_t\nabla\phi_i(t,\cdot))\Big]^2$. By Itô formula, we have

$$\begin{aligned} d_t\big(\rho_t\ln\rho_t\big) &= (\ln\rho_t + 1)d\rho_t + \frac{1}{2}\frac{1}{\rho_t}d\rho_t\cdot d\rho_t = -(\ln\rho_t+1)\sum_{i=0}^{N}\mathsf{div}(\rho_t\nabla\phi_i(t,\cdot))\ dB_t^i \\ &\quad + \frac{1}{2}(\ln\rho_t+1)\sum_{i=1}^{N}\mathsf{div}\big(\mathsf{div}(\rho_t\nabla\phi_i(t,\cdot))\nabla\phi_i(t,\cdot)\big)\ dt + \frac{1}{2\rho_t}\sum_{i=1}^{N}\Big[\mathsf{div}(\rho_t\nabla\phi_i(t,\cdot))\Big]^2\ dt. \end{aligned} \tag{5.2.8}$$

We have

$$\int_M (\ln\rho_t+1)\mathsf{div}\big(\mathsf{div}(\rho_t\nabla\phi_i(t,\cdot))\nabla\phi_i(t,\cdot)\big)\ dx = -\int_M \frac{\langle\nabla\rho_t, \nabla\phi_i(t,\cdot)\rangle}{\rho_t}\,\mathsf{div}(\rho_t\nabla\phi_i(t,\cdot))\, dx$$

Then integrating over M with respect to dx the sum of last two terms in (5.2.8), we get the quantity which is equal to

$$\int_M \frac{1}{2\rho_t}\Big[\mathrm{div}(\rho_t\nabla\phi_i(t,\cdot))\Big(\mathrm{div}(\rho_t\nabla\phi_i(t,\cdot)) - \langle\nabla\rho_t, \nabla\phi_i(t,\cdot)\rangle\Big)\Big]\,dx$$
$$= \int_M \frac{1}{2\rho_t}\mathrm{div}(\rho_t\nabla\phi_i(t,\cdot))\,\rho_t\Delta\phi_i(t,\cdot)\,dx = -\frac{1}{2}\int_M \langle\nabla\phi_i(t,\cdot)), \nabla\Delta\phi_i(t,\cdot))\rangle\,\rho_t dx,$$

which is $(\bar{D}_{V_{\phi_i(t,\cdot)}}\bar{D}_{V_{\phi_i(t,\cdot)}}\mathrm{Ent})(\mu_t)$ by (5.0.6). For the martingale term, we note that

$$-\int_M (\ln\rho_t + 1)\mathrm{div}(\rho_t\nabla\phi_i(t,\cdot))\,dx = \int_M \langle\frac{\nabla\rho_t}{\rho_t}, \nabla\phi_i(t,\cdot))\rangle\,\rho_t\,dx,$$

which is equal to $\langle\bar{\nabla}\mathrm{Ent}, V_{\phi_i(t,\cdot)}\rangle_{\mathbf{T}_{\mu_t}}$ according to (5.0.5). Therefore we get Equality (5.2.7). □

Proposition 5.2.5. *Itô formula* (5.2.5) *remains true for the functional F_3 considered in Section 1, that is, $F_3(\mu) = \int_{M\times M} W(x,y)\mu(dx)\mu(dy)$.*

Definition 5.2.6. *Let $\{\mu_t,\ t\geq 0\}$ be a stochastic process on $\mathbb{P}_2(M)$; we say that it solves the following SDE :*

$$\circ d_t^I\mu_t = \sum_{i=0}^N V_{\phi_i(t,\cdot)}(\mu_t)\ \circ dB_t^i, \quad \mu_0 = \mu. \tag{5.2.9}$$

if for each F of three functionals considered in Section 1, the following Itô formula holds:

$$d_tF(\mu_t) = \sum_{i=0}^N \langle\bar{\nabla}F, V_{\phi_i(t,\cdot)}\rangle_{\mathbf{T}_{\mu_t}}\ dB_t^i + \frac{1}{2}\sum_{i=1}^N(\bar{D}_{V_{\phi_i(t,\cdot)}}\bar{D}_{V_{\phi_i(t,\cdot)}}F)(\mu_t)\ dt.$$

In what follows, we will add an interesting drift term to SDE (5.2.9). For the sake of simplicity, we suppose that $W(x,y) = W(y,x)$ in Example 3; recall that $\Phi(x,\mu) = \int_M W(x,y)\mu(dy)$, then $\nabla\Phi(x,\mu) = 2\int_M (\nabla_1 W)(x,y)\mu(dy)$, where ∇_1 denotes the partial gradient with respect to the first component. We have

$$\nabla^2\Phi(x,\mu) = 2\int_M \nabla_1^2 W(x,y)\mu(dy).$$

It is obvious that $(x,\mu)\to\nabla\Phi(x,\mu)$ is continuous and $\sup_{(x,\mu)\in M\times\mathbb{P}_2(M)} |\nabla^2\Phi(x,\mu)|^2 < +\infty$. Let $\pi\in\mathcal{C}(\mu,\nu)$, we have

$$\begin{aligned}\nabla\Phi(x,\mu) - \nabla\Phi(x,\nu) =& 2\Big(\int_M \nabla_1 W(x,y)\mu(dy) - \int_M \nabla_1 W(x,y)\nu(dy)\Big)\\ &= 2\int_{M\times M}\Big(\nabla_1 W(x,y) - \nabla_1 W(x,z)\Big)\pi(dy,dz).\end{aligned}$$

Hence

$$
\begin{aligned}
|\nabla\Phi(x,\mu)-\nabla\Phi(x,\nu)| &\le 2\int_{M\times} ||\nabla_2\nabla_1 W||_\infty d_M(y,z)\,\pi(dy,dz) \\
&\le ||\nabla_2\nabla_1 W||_\infty\, W_2(\mu,\nu).
\end{aligned} \tag{5.2.10}
$$

We prove that $\mu \to \bar{\nabla} F_3(\mu)$ satisfies the Lipschitz condition introduced in [DF21].

Stochastic Mckean-Vlasov equations have been recently considered in [Wan21, BLPR17], the following proposition is highly related to [Wan21].

Proposition 5.2.7. *There is a solution (X_t,μ_t) to the following Mckean-Vlasov SDE:*

$$
dX_t = \sum_{i=0}^N \nabla\phi_i(X_t)\circ dB_t^i + \nabla\Phi(X_t,\mu_t)\,dt,\ \mu_t = (X_t)_\#\mu, \tag{5.2.11}
$$

where $\Phi(x,\mu) = \int_M W(x,y)\mu(dy)$.

Proof. Let $(U_t)_{t\ge 0}$ be the stochastic flow associated to the folllowing SDE

$$
dU_t = \sum_{i=0}^N \nabla\phi_i(U_t)\circ dB_t^i.
$$

Define the stochastic measure dependent vector fields $V_t(\omega,x,\mu)$ on M by

$$
V_t(\omega,x,\mu) = \big(U_t^{-1}(\omega,\cdot)\big)_*\nabla\Phi(x,(U_t)_\#\mu) = (U_t^{-1})'(\omega,U_t(x))\nabla\Phi\big(U_t(x),(U_t)_\#\mu\big),
$$

where the prime denotes the differential with respect to x. Since the manifold M is compact, we have

$$
|V_t(\omega,x,\mu)-V_t(\omega,x,\nu)| \le ||(U_t^{-1})'||_\infty\ |\Phi\big(U_t(x),(U_t)_\#\mu\big)-\Phi\big(U_t(x),(U_t)_\#\nu\big)|.
$$

Now according to (5.2.10), we get

$$
|V_t(\omega,x,\mu)-V_t(\omega,x,\nu)| \le ||(U_t^{-1})'||_\infty\ |||\nabla_2\nabla_1 W||_\infty\, W_2\big((U_t)_\#\mu,(U_t)_\#\nu\big),
$$

which is dominated by

$$
||(U_t^{-1})'||_\infty\ |||\nabla_2\nabla_1 W||_\infty\ ||U_t'||_\infty\ W_2(\mu,\nu).
$$

So there is a unique solution (Y_t,ν_t) to

$$\frac{d}{dt} Y_t = V_t(Y_t, \nu_t), \quad \nu_t = (Y_t)_{\#}\mu.$$

Let $\tilde{X}_t = U_t(Y_t)$. By Itô-Ventzell formula,

$$d\tilde{X}_t = \sum_{i=0}^{N} \nabla\phi_i(U_t(Y_t)) \circ dB_t^i + U_t'(Y_t)\, V_t(Y_t, \nu_t),$$

the last term in above equality is

$$\nabla\Phi\big(\tilde{X}_t, (U_t)_{\#}\nu_t\big).$$

Note that $(\tilde{X}_t)_{\#}\mu = (U_t)_{\#}\,(Y_t)_{\#}\mu = (U_t)_{\#}\nu_t$; therefore $(\tilde{X}_t, (U_t)_{\#}\nu_t)$ is a solution to Equation (5.2.11). For the uniqueness of solutions, see [Wan21].

□

Theorem 5.2.8. *Let F_3 be the functional in Example 3, and $d\mu = \rho\, dx$ with $\rho > 0$ in C^1; then there is a unique solution $\{\mu_t;\ t \geq 0\}$ to the following SDE on $\mathbb{P}_2(M)$:*

$$\circ d_t^I \mu_t = \sum_{i=0}^{N} V_{\phi_i}(\mu_t)\ \circ dB_t^i + \bar{\nabla} F_3(\mu_t)\, dt, \quad \mu_0 = \mu. \tag{5.2.12}$$

Proof. Let (X_t, μ_t) be the unique solution to the Mckean-Vlasov SDE (5.2.11), then for any polynomial F on $\mathbb{P}_2(M)$, we have

$$d_t F(\mu_t) = \sum_{i=0}^{N} \langle \bar{\nabla} F, V_{\phi_i}\rangle_{\mathbf{T}_{\mu_t}}\ dB_t^i + \frac{1}{2}\sum_{i=1}^{N} (\bar{D}_{V_{\phi_i}} \bar{D}_{V_{\phi_i}} F)(\mu_t)\ dt + \langle \bar{\nabla} F, \bar{\nabla} F_3\rangle_{\mathbf{T}_{\mu_t}}\, dt.$$

We check also this is true for two other examples in Section 1. The uniqueness comes from Lipschitz continuity of coefficients in (5.2.12).

□

5.3 Towards stochastic parallel translations in $\mathbb{P}_2(M)$

For the reason of simplicity, we consider the following SDE on M

$$dX_t = \sum_{i=0}^{N} \nabla\phi_i(X_t) \circ dB_t^i, \quad X_0(x) = x, \tag{5.3.1}$$

where $\{\phi_0, \phi_1, \ldots, \phi_N\}$ are smooth enough and independent of the time t. We know that SDE (5.3.1) defines a stochastic flow of C^r-diffeomorphisms. The main purpose of this scetion is to deal with the stochastic parallel translation along stochastic regular curves $\{\mu_t;\ t \geq 0\}$ in $\mathbb{P}_2(M)$ defined by

$$\circ d_t^I \mu_t = \sum_{i=0}^{N} V_{\phi_i}(\mu_t)\ \circ dB_t^i, \quad \mu_0 = \mu. \tag{5.3.2}$$

For almost surely ω, $t \to \mu_t(\omega, dx)$ is not a regular curve of $\mathbb{P}_2(M)$ in the sense of [AG08]. In fact, denoting

$$D(t,s) = \mathrm{Lip}\big(X_t \circ X_s^{-1} - \mathrm{Id}\big),$$

then the condition

$$\lim_{t\to s} \frac{D^2(t,s)}{|t-s|} = 0$$

in [AG08] fails to hold, since for a Brownian motion $\{B_t\}$, $\lim_{t\to s} \dfrac{|B_t - B_s|^2}{|t-s|} \neq 0$. Therefore the method in [AG08] does not work directly for stochastic regular curved defined by (5.3.2). On the other hand, divers limit theorems from ODE to SDE provide powerful tools in stochastic analysis, see for example [Bis81, Mal97, IW81]. In what follows, we will show what happens in this direction.

We consider the regularized Brownian motion $\{B_t^n, t \in [0,1]\}$ which is piecewise linear. More precisely, for $n \geq 1$, denote

$$t_n = \frac{[2^n t]}{2^n},\ t_n^+ = \frac{[2^n t]+1}{2^n}, \quad \text{and} \quad \dot{B}_n(t) = 2^n\,\big(B_{t_n^+} - B_{t_n}\big),$$

where $[x]$ denotes the integral part of real number x. Let X_t^n be the solution to the ODE

$$dX_t^n = \sum_{i=0}^{N} \nabla\phi_i(X_t^n)\,\dot{B}_n^i(t)\,dt, \quad X_0^n(x) = x. \tag{5.3.3}$$

It is well-known ([Bis81, Mal97, IW81]) that for almost surely $\omega \in \Omega$, as $n \to +\infty$, $X_t^n(x,\omega)$ converges to $X_t(x,\omega)$ in a C^r topology uniformly with respect to $t \in [0,1]$. Let μ be a probability measure on M having a positive density $\rho > 0$ in C^2, put

$$\mu_t^n(\omega) = \big(X_t^n(\cdot,\omega)\big)_{\#}\mu.$$

It is clear that for almost surely ω, as $n \to \infty$, μ_t^n converges to μ_t uniformly in $t \in [0,1]$. By Lemma 4.3.1 in [Kun97], the measure $(X_t^{-1})_{\#}\mu$ relative to μ admits a positive density $\tilde{K}_t(x)$, which has the following expression, for almost surely ω, all $t \geq 0$ and $x \in M$:

$$\tilde{K}_t(x) = \exp\Big[\sum_{i=0}^{N} \int_0^t \mathrm{div}_\mu(\nabla\phi_i)(X_s(x)) \circ dB_s^i\Big]. \tag{5.3.4}$$

The density ρ_t of $(X_t)_{\#}\mu$ relative to μ is given by

$$\rho_t(x) = \frac{1}{\tilde{K}_t\big(X_t^{-1}(x)\big)}.$$

The SDE for writing X_t^{-1} is much more complicated than ODE. On the other hand, for a C^1-diffeomorphism $\Xi : M \to M$, the differential $d\Xi(x)$ sends T_xM into $T_{\Xi(x)}M$, its dual map $(d\Xi(x))^*$ sends $\mathbf{T}_{\Xi(x)}M$ into $\mathbf{T}_xM$. Denoting $\sigma_\Xi(x) = (d\Xi(x))^* \circ d\Xi(x)$, the density k of $\Xi_{\#}\mu$ relative to dx has the expression

$$k = \frac{\rho}{\sqrt{\det(\sigma_\Xi(x))}} \circ \Xi^{-1}. \tag{5.3.5}$$

Let $\rho_t^n = \dfrac{d\mu_t^n}{dx}$; then, according to above formula,

$$\rho_t^n = \frac{\rho}{\sqrt{\det(\sigma_{t,n})}} \circ (X_t^n)^{-1}, \quad \sigma_{t,n} = (dX_t^n(x))^* d(X_t^n(x)). \tag{5.3.6}$$

For the convergence of k_t^n, we prepare the following lemma

Lemma 5.3.1. *Let Ξ_n and Ξ be C^1-diffeomorphism of M such that Ξ_n and $\nabla\Xi_n$ converge to Ξ and $\nabla\Xi$ uniformly as $n \to +\infty$, then Ξ_n^{-1} converges to Ξ^{-1} uniformly as $n \to +\infty$.*

Proof. Let γ be a geodesic curve which connects $\Xi_n(\Xi_n^{-1}(x))$ and $\Xi_n(\Xi^{-1}(x))$. Let $\tilde{\gamma}(s) = \Xi_n^{-1}(\gamma(s))$; then $\tilde{\gamma}$ connects $\Xi_n^{-1}(x)$ and $\Xi^{-1}(x)$. We have

$$\gamma(s) = \Xi_n\big(\tilde{\gamma}(s)\big), \quad \gamma'(s) = d\Xi_n\big(\tilde{\gamma}(s)\big)\,\tilde{\gamma}'(s).$$

There is a constant $c > 0$ such that $\langle \sigma_\Xi(x)u, u\rangle_{T_xM} \geq c|u|^2_{T_xM}$ for all $x \in M$. Since

$$\lim_{n\to+\infty} \sup_{x\in M} |\sigma_{\Xi_n}(x) - \sigma_\Xi(x)| = 0$$

for big enough n,

$$\langle \sigma_{\Xi_n}(x)u, u\rangle_{T_xM} \geq c|u|^2_{T_xM}/2,$$

which implies that $|d\Xi_n(\tilde{\gamma}(s))\,\tilde{\gamma}'(s)| \geq \sqrt{\dfrac{c}{2}}\,|\tilde{\gamma}'(s)|$. It follows that

$$\int_0^1 |\gamma'(s)|\,ds \geq \sqrt{\frac{c}{2}} \int_0^1 |\tilde{\gamma}'(s)|\,ds \geq \sqrt{\frac{c}{2}}\, d_M\big(\Xi_n^{-1}(x), \Xi^{-1}(x)\big).$$

Hence

$$d_M\big(\Xi_n^{-1}(x),\Xi^{-1}(x)\big) \leq \sqrt{\frac{2}{c}}\, d_M\big(\Xi(\Xi^{-1}(x),(\Xi_n(\Xi^{-1}(x))\big) \leq \sqrt{\frac{2}{c}}\,\sup_{y\in M} d_M\big(\Xi(y),\Xi_n(y)\big).$$

The result follows.

□

Proposition 5.3.2. *Almost surely,*

$$\lim_{n\to+\infty}\sup_{(t,x)\in[0,1]\times M} |\rho_t^n(x)-\rho_t(x)| = 0. \tag{5.3.7}$$

Furthermore

$$\lim_{n\to+\infty}\sup_{(t,x)\in[0,1]\times M} |\nabla\log(\rho_t^n(x))-\nabla\log(\rho_t(x))| = 0. \tag{5.3.8}$$

Proof. By formula (5.3.6) and above Lemma, we get the result (5.3.7). For (5.3.8), we note that for a diffeomorphism Ξ, $\nabla\Xi^{-1} = \big(\nabla\Xi(\Xi^{-1})\big)^{-1}$. Taking the derivative with respect to x in formula (5.3.6), we have

$$\nabla k_t^n = \nabla\Big(\frac{k}{\sqrt{\det(\sigma_{t,n})}}\Big)\circ(X_t^n)^{-1}\cdot\Big(\nabla X_t^n\circ(X_t^n)^{-1}\Big)^{-1}.$$

Again by Lemma 5.3.1, we get (5.3.8). □

Now by Theorem 5.1.1, for any $\nabla\Psi_0\in L^2(\mu)$ given, there is a unique

$$\nabla\Psi_t^n(\omega,\cdot)\in\mathbf{T}_{\mu_t^n(\omega)},$$

which is the parallel translation along $\{\mu_t^n(\omega); t\in[0,1]\}$. Then for almost $\omega\in\Omega$, $n\geq 1$,

$$\int_M |\nabla\Psi_t^n(\omega,x)|^2\,\mu_t^n(\omega,dx) = \int_M |\nabla\Psi_0(x)|^2\,\mu(dx),$$

or using the density ρ_t^n of μ_t^n,

$$\int_M |\nabla\Psi_t^n(\omega,x)|^2\,\rho_t^n(\omega,x)\,dx = \int_M |\nabla\Psi_0(x)|^2\,\mu(dx). \tag{5.3.9}$$

This result impies that for each $(t,\omega)\in[0,1]\times\Omega$, the sequence $\{\nabla\Psi_t^n(\cdot,x)\sqrt{\rho_t^n(\cdot,\omega)};\ n\geq 1\}$ is bounded in L^2 by $||\nabla\Psi_0||_{L^2(\mu)}$. There is a limit point, but unfortunately, the subsequence is dependent of (t,ω). We have to consider the integration in the space $[0,1]\times\Omega\times M$. For any $n\geq 1$,

$$\int_{[0,1]\times\Omega}\Big[\int_M |\nabla\Psi_t^n(\omega,x)|^2\,\rho_t^n(\omega,x)\,dx\Big]\,dt\,P(d\omega) = \int_M |\nabla\Psi_0(x)|^2\,\mu(dx);$$

there exists then a Random time-dependent vector field $v_t(\omega,x)$ satisfying

$$\int_{[0,1]\times\Omega}\Big[\int_M |v_t(\omega,x)|^2 dx\Big]\,dt\,P(d\omega) \le \int_M |\nabla\Psi_0(x)|^2\,\mu(dx),$$

such that, up to a subsequence, the sequence $\{\nabla\Psi_t^n(\omega,x)\,\sqrt{\rho_t^n(\omega,x)}; n\ge 1\}$ converges weakly to $v_t(\omega,x)$ in L^2. We note that for any bounded Ramdon variable $\xi:\Omega\to\mathbb{R}$ and any bounded function $\alpha:[0,1]\to\mathbb{R}$,

$$\int_{[0,1]\times\Omega}\Big[\int_M |\nabla f(x)|^2\,\rho_t\,dx\Big]\,\alpha(t)\xi(\omega)\,dt\,P(d\omega) < +\infty.$$

Therefore

$$\begin{aligned}
&\int_{[0,1]\times\Omega}\Big[\int_M \langle\nabla f(x), v_t(\omega,x)\rangle\,\sqrt{\rho_t}\,dx\Big]\alpha(t)\xi(\omega)\,dt\,P(d\omega)\\
&=\lim_{n\to+\infty}\int_{[0,1]\times\Omega}\Big[\int_M \langle\nabla f(x),\nabla\Psi_t^n\rangle\,\sqrt{\rho_t^n(\omega,x)}\,\sqrt{\rho_t}\,dx\Big]\,\alpha(t)\xi(\omega)\,dt\,P(d\omega)\\
&=\lim_{n\to+\infty}\int_{[0,1]\times\Omega}\Big[\int_M \langle\nabla f(x),\nabla\Psi_t^n\rangle\,\rho_t^n(\omega,x)\,dx\Big]\,\alpha(t)\xi(\omega)\,dt\,P(d\omega),
\end{aligned}$$

Since $v_t(\omega,\cdot)\,\rho_t^{-1/2}\in L^2(\mu_t)$ for almost surely (t,ω), there exists $\Psi_t(\omega,\cdot)\in H^1(\mu_t)$ such that for any $f\in C^2(M)$,

$$\int_M \langle\nabla f(x),\ v_t(\omega,x)\,\rho_t^{-1/2}\rangle\mu_t(dx) = \int_M \langle\nabla f(x),\ \nabla\Psi_t(\omega,x)\,\rangle\mu_t(dx).$$

We obtain the following result:

Proposition 5.3.3. *There exists $\nabla\Psi_\cdot$ such that $\int_{[0,1]\times\Omega}\Big[\int_M |\nabla\Psi_t(\omega,x)|^2\mu_t(dx)\Big]\,dtP(d\omega)$ is finite and*

$$\begin{aligned}
&\int_{[0,1]\times\Omega}\Big[\int_M \langle\nabla f(x),\nabla\Psi_t\rangle\,\mu_t(\omega,dx)\Big]\,\alpha(t)\xi(\omega)\,dt\,P(d\omega)\\
&=\lim_{n\to+\infty}\int_{[0,1]\times\Omega}\Big[\int_M \langle\nabla f(x),\nabla\Psi_t^n\rangle\,\mu_t^n(\omega,dx)\Big]\,\alpha(t)\xi(\omega)\,dt\,P(d\omega).
\end{aligned}\tag{5.3.10}$$

This convergence is too weak to yield interesting informations on $\{\Psi_t;\ t\in[0,1]\}$.

In what follows, we will try to get a weak form of SPDE for stochastic parallel translations.

Let $f\in C^2(M)$; by (5.1.3), for any $n\ge 1$, almost surely ω,

$$\begin{aligned}\frac{d}{dt}\int_M \langle \nabla f, \nabla \Psi_t^n\rangle\, \mu_t^n(dx) &= \int_M \langle \nabla f^2\,, \sum_{i=0}^N \nabla\phi_i \dot{B}_n^i(t)\otimes \nabla\Psi_t^n\rangle\, \mu_t^n(dx)\\ &= \sum_{i=0}^N \Big(\int_M \langle \nabla_{\nabla\phi_i}\nabla f,\ \nabla\Psi_t^n\rangle\, \mu_t^n(dx)\Big)\, \dot{B}_n^i(t).\end{aligned} \tag{5.3.11}$$

For a C^1 vector field ζ on M, set

$$z_\zeta^n(\omega,t) = \int_M \langle \zeta, \nabla\Psi_t^n\rangle\, \mu_t^n(dx).$$

By (5.1.11), we have

$$\begin{aligned}\frac{d}{dt}\int_M \langle \zeta, \nabla\Psi_t^n\rangle\, \mu_t^n(dx) =& \sum_{i=0}^N \Big(\int_M \langle \nabla(\Pi_{\mu_t^n}(\zeta)),\ \nabla\phi_i\otimes\nabla\Psi_t^n\rangle \mu_t^n(dx)\Big)\, \dot{B}_t^n\\ &- \sum_{i=0}^N \Big(\int_M \langle \mathcal{L}^{\mu_t^n}(\phi_i)\Pi_{\mu_t^n}^{\perp}(\zeta), \nabla\Psi_t^n\rangle \mu_t^n(dx)\Big)\, \dot{B}_t^n,\end{aligned}$$

or for $s<t$,

$$\begin{aligned}z_\zeta^n(t)-z_\zeta^n(s) =& \sum_{i=0}^N \int_s^t \Big(\int_M \langle \nabla(\Pi_{\mu_\tau^n}(\zeta)),\ \nabla\phi_i\otimes\nabla\Psi_\tau^n\rangle \mu_\tau^n(dx)\Big)\, \dot{B}_\tau^n\, d\tau\\ &- \sum_{i=0}^N \int_s^t \Big(\int_M \langle \mathcal{L}^{\mu_\tau^n}(\phi_i)\Pi_{\mu_\tau^n}^{\perp}(\zeta), \nabla\Psi_\tau^n\rangle \mu_\tau^n(dx)\Big)\, \dot{B}_\tau^n\, d\tau\end{aligned} \tag{5.3.12}$$

Therefore there is a constant $C>0$ independent of n such that

$$\mathbb{E}(|z_\zeta^n(t)-z_\zeta^n(s)|^p) \le C\,|t-s|^{p/2}.$$

By Kolmogorov's modification theorem, there exist $M_n\in L^p(\Omega)$, bounded in $L^p(\Omega)$ such that

$$|z_\zeta^n(\omega,t)-z_\zeta^n(\omega,s)| \le M_n(\omega)\,|t-s|^\alpha,\quad \alpha>0, \tag{5.3.13}$$

Remark that

$$||z_\zeta^n||_\infty \le ||\zeta||_\infty\, ||\nabla\Psi_0||_{L^2(\mu)}. \tag{5.3.14}$$

For simplicity, denote for the moment, $\zeta_i = \nabla_{\nabla\phi_i}(\nabla f)$. Consider the following family of $\mathbb{R}^{2N+2}$ valued stochastic process

$$t \to \Lambda_f^n(t) = \big(z_{\nabla f}^n(t), z_{\zeta_0}^n(t), \cdots, z_{\zeta_N}^n(t), B^1(t), \cdots, B^N(t)\big).$$

Let $R > 0$, define $K_R = \big\{z \in C([0,1], \mathbb{R}^{2N+2});\ ||z(0)|| \le R,\ ||z(t)-z(s)|| \le R\,|t-s|^\alpha\big\}$. By Ascoli theorem, K_R is a compact subset of $C([0,1], \mathbb{R}^{2N+2})$. Let ν_f^n be the law of $\omega \to \Lambda_f^n(\omega, \cdot)$ in $C([0,1], \mathbb{R}^{2N+2})$. Then

$$\nu_f^n(K_R^c) \le \nu_f^n([||z(0)|| > R) + \nu_f^n\big(\{\exists t \ne s, ||z(t) - z(s)|| > R|t-s|^\alpha\}\big).$$

But

$$\begin{aligned}\nu_f^n\big(\{\exists t \ne s, ||z(t) - z(s)|| > R|t-s|^\alpha\}\big) &\sum_{i=1}^{2N+2} \mathbb{P}\big(\{\exists t \ne s, ||z_{\zeta_i}(t) - z_{\zeta_i}(s)|| > R|t-s|^\alpha\}\big)\\ &\le C_1\,\mathbb{P}(M_n \ge R) \le \frac{C_1\,||M_n||_{L^p}^p}{R^p} \le \frac{C}{R^p},\end{aligned}$$

for a constant $C > 0$ independent of n. Therefore the family $\{\nu_f^n;\ n \ge 1\}$ is tight. Up to a subsequence, $\{\nu_f^n;\ n \ge 1\}$ converges weakly to a probability measure ν_f on $C([0,1], \mathbb{R}^{2N+2})$.

Now by Skorohod representation theorem, there is a probability space $(\Omega_f, \mathbb{P}_f)$ and a sequence of Random variables $\hat{\Lambda}_f^n : \Omega_f \to C([0,1], \mathbb{R}^{2N+2})$ and $\hat{\Lambda}_f : \Omega_f \to C([0,1], \mathbb{R}^{2N+2})$ such that the law of $\hat{\Lambda}_f^n$ is ν_f^n, that of $\hat{\Lambda}_f$ is ν_f, and

$$\hat{\Lambda}_f^n \text{ converges almost surely to } \hat{\Lambda}_f, \quad \text{as } n \to +\infty.$$

Furthermore let

$$\hat{\Lambda}_f^n(t) = \big(\hat{Z}_{\nabla f}^n(t), \hat{Z}_{\zeta_0}^n(t), \cdots, \hat{Z}_{\zeta_N}^n(t), \hat{B}^1(t), \cdots, \hat{B}^N(t)\big),$$

and

$$\hat{\Lambda}_f(t) = \big(\hat{Z}_{\nabla f}(t), \hat{Z}_{\zeta_0}(t), \cdots, \hat{Z}_{\zeta_N}(t), \hat{B}^1(t), \cdots, \hat{B}^N(t)\big).$$

As marginal laws, $\big(z_{\nabla f}^n(t), z_{\zeta_0}^n(t), \cdots, z_{\zeta_N}^n(t)\big)$ and $\big(\hat{Z}_{\nabla f}^n(t), \hat{Z}_{\zeta_0}^n(t), \cdots, \hat{Z}_{\zeta_N}^n(t)\big)$ have the same law, and $\hat{B}_t = \big(\hat{B}^1(t), \cdots, \hat{B}^N(t)\big)$ is a $\mathbb{R}^N$-valued standard Brownian motion on $(\Omega_f, \mathbb{P}_f)$. By (5.3.11), we have, for $s < t$,

$$z_{\nabla f}^n(t) - z_{\nabla f}^n(s) = \sum_{i=0}^{N} \int_s^t z_{\zeta_i}^n(\tau)\, \dot{B}_n^i(\tau)\, d\tau,$$

or

$$z^n_{\nabla f}(t) - z^n_{\nabla f}(s) - \sum_{i=0}^{N} \int_s^t z^n_{\zeta_i}(\tau)\, \dot{B}^i_n(\tau)\, d\tau = 0.$$

We can express the left hand side of above equality as a function $J(\Lambda^n_f)$ of Λ^n_f. Let $G : \mathbb{R} \to \mathbb{R}_+$ be the bounded continuous function defined by $G(\xi) = |\xi|^2 \wedge M$. We have

$$\begin{aligned}&\hat{\mathbb{E}}\Big(G\big(\hat{Z}^n_{\nabla f}(t) - \hat{Z}^n_{\nabla f}(s) - \sum_{i=0}^{N} \int_s^t \hat{Z}^n_{\zeta_i}(\tau)\, \dot{\hat{B}}^i_n(\tau)\, d\tau\big)\Big) \\ &= \mathbb{E}\Big(G\big(z^n_{\nabla f}(t) - z^n_{\nabla f}(s) - \sum_{i=0}^{N} \int_s^t z^n_{\zeta_i}(\tau)\, \dot{B_n}^i(\tau)\, d\tau\big)\Big) = 0.\end{aligned}$$

Now letting $n \to +\infty$,

$$\hat{Z}^n_{\nabla f}(t) - \hat{Z}^n_{\nabla f}(s) \to \hat{Z}_{\nabla f}(t) - \hat{Z}_{\nabla f}(s),$$

and

$$\sum_{i=0}^{N} \int_s^t \hat{Z}^n_{\zeta_i}(\tau)\, \dot{\hat{B}}^i_n(\tau)\, d\tau \to \sum_{i=0}^{N} \int_s^t \hat{Z}_{\zeta_i}(\tau) \circ d\hat{B}^i_\tau.$$

Therefore we obtain

$$\hat{Z}_{\nabla f}(t) - \hat{Z}_{\nabla f}(s) = \sum_{i=0}^{N} \int_s^t \hat{Z}_{\zeta_i}(\tau) \circ d\hat{B}^i_\tau \quad \text{almost surely.}$$

Using the separability of $C^2(M)$ and diagonal method, we can get the common subsequence for all $f \in C^2(M)$. We state the above result as follows

Theorem 5.3.4. *There is a probability space $(\hat{\Omega}, \hat{\mathbb{P}})$ such that there is a subsequence n_k, for each of them and each $f \in C^2(M)$, the $C([0,1], \mathbb{R}^{N+2})$ valued Random variable*

$$\big(z^n_{\nabla f}, z^n_{\nabla_{\nabla\phi_0}(\nabla f)}, \cdots, z^n_{\nabla_{\nabla\phi_N}(\nabla f)}\big),$$

has a version $\hat{\Lambda}^n_f$ defined on $(\hat{\Omega}, \hat{\mathbb{P}})$, says,

$$\hat{\Lambda}^n_f = \big(\hat{Z}^n_{\nabla f}, \hat{Z}^n_{\nabla_{\nabla\phi_0}(\nabla f)}, \cdots, \hat{Z}^n_{\nabla_{\nabla\phi_N}(\nabla f)}\big),$$

which converges almost surely to

$$\Lambda_f = \big(\hat{Z}_{\nabla f}, \hat{Z}_{\nabla_{\nabla\phi_0}(\nabla f)}, \cdots, \hat{Z}_{\nabla_{\nabla\phi_N}(\nabla f)}\big).$$

Furthermore for $s < t$,

$$\hat{Z}_{\nabla f}(t) - \hat{Z}_{\nabla f}(s) = \sum_{i=0}^{N} \int_s^t \hat{Z}_{\nabla_{\nabla\phi_i}(\nabla f)}(\tau) \circ d\hat{B}^i_\tau \quad \textit{almost surely.}$$

Now we look for the strong form of SPDE for stochastic parallel translations. To this end, we suppose that there is a continuous process $\big\{\nabla\Psi_t \in \mathbf{T}_{\mu_t};\ t \in [0,1]\big\}$ such that, up to a subsequence, almost surely, for any C^1 vector field ζ on M,

$$\int_M \langle \zeta(x), \nabla\Psi^n_t(x)\rangle\, \mu^n_t(dx)$$

converge uniformly in $t \in [0,1]$, as $n \to +\infty$, to

$$\int_M \langle \zeta(x), \nabla\Psi_t(x)\rangle\, \mu_t(dx).$$

In the spirit of Wong-Zakai approximation, the term

$$\sum_{i=1}^{N} \int_0^t \Big(\int_M \langle \nabla_{\nabla\phi_i}(\nabla f),\ \nabla\Psi^n_\tau\rangle\, \mu^n_\tau(dx)\Big)\, \dot{B}^i_n(\tau)\, d\tau.$$

converges, as $n \to +\infty$, to the following Stratanovich stochastic integral:

$$\sum_{i=1}^{N} \int_0^t \Big(\int_M \langle \nabla_{\nabla\phi_i}(\nabla f),\ \nabla\Psi_\tau\rangle\, \mu_\tau(dx)\Big) \circ dB^i(\tau).$$

We have to compute the Itô stochastic contraction:

$$\frac{1}{2}\sum_{i=1}^{N} d_t \int_0^t \Big(\int_M \langle \nabla_{\nabla\phi_i}(\nabla f),\ \nabla\Psi_t\rangle\, \mu_t(dx)\Big) \cdot dB^i(t).$$

Using formally the equality (5.1.11), we have

$$\begin{aligned} d_t \int_M \langle \nabla_{\nabla\phi_i}(\nabla f),\ \nabla\Psi_t\rangle\, \mu_t(dx) = &-\sum_{j=1}^{N} \Big(\int_M \langle \mathcal{L}^{\mu_t}(\phi_j)\Pi^{\perp}_{\mu_t}\big(\nabla_{\nabla\phi_i}(\nabla f)\big),\ \nabla\Psi_t\rangle \mu_t(dx)\Big) \circ dB^j(t) \\ &+ \sum_{j=1}^{N} \Big(\int_M \langle \nabla_{\nabla\phi_j}\Pi_{\mu_t}\big(\nabla_{\nabla\phi_i}(\nabla f)\big),\ \nabla\Psi_t\rangle \mu_t(dx)\Big) \circ dB^j(t). \end{aligned}$$

Let's introduce the following notation:

$$R_t^f = \frac{1}{2}\sum_{i=1}^{N} \Pi_{\mu_t}\Big(\nabla_{\nabla\phi_i}\Pi_{\mu_t}\big(\nabla_{\nabla\phi_i}(\nabla f)\big)\Big), \tag{5.3.15}$$

and

$$S_t^f = \frac{1}{2}\sum_{i=1}^{N} \Pi_{\mu_t}\Big(\mathcal{L}^{\mu_t}(\phi_i)\Pi_{\mu_t}^{\perp}\big(\nabla_{\nabla\phi_i}(\nabla f)\big)\Big). \tag{5.3.16}$$

The term R_t^f has an intrinsic expression using covariant derivatives on $\mathbb{P}_2(M)$, due to (5.0.3), that is,

$$R_t^f = \frac{1}{2}\sum_{i=1}^{N}(\bar{\nabla}_{V_{\phi_i}}\bar{\nabla}_{V_{\phi_i}}V_f)(\mu_t). \tag{5.3.17}$$

Hence for any $f \in C^3(M)$, we obtain the following Itô form of weak SPDE,

$$\begin{aligned}\int_M \langle \nabla f,\ \nabla\Psi_t\rangle\, \mu_t(dx) &= \int_M \langle \nabla f,\ \nabla\Psi_0\rangle\, \mu(dx)\\ &+ \sum_{i=1}^{N}\int_0^t \Big(\int_M \langle \nabla_{\nabla\phi_i}(\nabla f)\,,\ \nabla\Psi_\tau\rangle \mu_\tau(dx)\Big)\, dB^i(\tau)\\ &+ \int_0^t \Big(\int_M \langle \nabla_{\nabla\phi_0}(\nabla f) + R_\tau^f + S_\tau^f\ ,\ \nabla\Psi_\tau\rangle\mu_\tau(dx)\Big)\, d\tau,\end{aligned} \tag{5.3.18}$$

or more intrinsically

$$\begin{aligned}\langle V_f, V_{\Psi_t}\rangle_{\mathbf{T}_{\mu_t}} &= \langle V_f, V_{\Psi_0}\rangle_{\mathbf{T}_\mu} + \sum_{i=1}^{N}\int_0^t \langle \bar{\nabla}_{V_{\phi_i}}V_f, V_{\Psi_\tau}\rangle_{\mathbf{T}_{\mu_\tau}}\, dB^i(\tau) + \int_0^t \langle \bar{\nabla}_{V_{\phi_0}}V_f, V_{\Psi_\tau}\rangle_{\mathbf{T}_{\mu_\tau}}\, d\tau\\ &+ \frac{1}{2}\sum_{i=1}^{N}\int_0^t \langle \bar{\nabla}_{V_{\phi_i}}\bar{\nabla}_{V_{\phi_i}}V_f, V_{\Psi_\tau}\rangle_{\mathbf{T}_{\mu_\tau}}\, d\tau + \int_0^t \langle S_\tau^f, V_{\Psi_\tau}\rangle_{\mathbf{T}_{\mu_\tau}}\, d\tau.\end{aligned} \tag{5.3.19}$$

The last term in above equality is novel. If furthermore, for $t \in [0,1]$, $x \to \nabla\Psi_t(x)$ is regular enough, we have the following strong SPDE:

Theorem 5.3.5. *Let $\{\nabla\Psi_t;\ t \in [0,1]\}$ be a solution to* (5.3.18) *such that $x \to \Psi_t(x)$ is C^3, then*

$$d_t\nabla\Psi_t = -\sum_{i=1}^{N}\Pi_{\mu_t}(\nabla_{\nabla\phi_i}\nabla\Psi_t)\, dB_t^i + \Pi_{\mu_t}\Big(-\nabla_{\nabla\phi_0}\nabla\Psi_t + R_t^{\Psi_t} + S_t^{\Psi_t}\Big)\, dt, \tag{5.3.20}$$

or in Stratanovich form:

$$\circ d_t \nabla \Psi_t = -\sum_{i=0}^{N} \Pi_{\mu_t} \big(\nabla_{\nabla \phi_i} \nabla \Psi_t \big) \circ dB_t^i, \tag{5.3.21}$$

or intrinsically

$$\circ d_t \, V_{\Psi_t} = -\sum_{i=0}^{N} \bar{\nabla}_{V_{\phi_i}} V_{\Psi_t} \circ dB_t^i. \tag{5.3.22}$$

Proof. Let $\rho_t = \dfrac{d\mu_t}{d\mu}$ be the density of μ_t with respect to the initial measure μ, then $\{\rho_t;\ t \in [0,1]\}$ satisfies the following SPDE:

$$\circ d_t \rho_t = -\sum_{i=0}^{N} \Big(\mathsf{div}_{\mu_t} (\nabla \phi_i) \, \rho_t \Big) \circ dB_t^i. \tag{5.3.23}$$

Using ρ_t, the left hand side of (5.3.18) is equal to $\int_M \langle \nabla f(x), \nabla \Psi_t(x) \rangle \, \rho_t \, \mu(dx)$, so the Stratanovich stochastic differential of this term is

$$\int_M \langle \nabla f(x), \circ d_t \nabla \Psi_t(x) \rangle \, \rho_t \, \mu(dx) + \int_M \langle \nabla f(x), \nabla \Psi_t(x) \rangle \ \circ d_t \rho_t \, \mu(dx) = J_1(t) + J_2(t)$$

respectively. By (5.3.23),

$$\begin{aligned} J_2(t) &= -\sum_{i=0}^{N} \Big[\int_M \langle \nabla f(x), \nabla \Psi_t(x) \rangle \, \mathsf{div}_{\mu_t} (\nabla \phi_i) \, \mu_t(dx) \Big] \circ dB_t^i \\ &= \sum_{i=0}^{N} \Big[\int_M \big(\langle \nabla_{\nabla \phi_i} \nabla f(x), \nabla \Psi_t(x) \rangle \ + \langle \nabla f(x), \nabla_{\nabla \phi_i} \nabla \Psi_t(x) \rangle \, \big) \, \mu_t(dx) \Big] \circ dB_t^i. \end{aligned}$$

In Stratanovich form, the right hand side of (5.3.18) is

$$\sum_{i=0}^{N} \Big[\int_M \langle \nabla_{\nabla \phi_i} \nabla f(x), \ \nabla \Psi_t(x) \rangle \, \big) \, \mu_t(dx) \Big] \circ dB_t^i.$$

Combing these equalities, we obtain, for any $f \in C^2(M)$,

$$\int_M \langle \nabla f(x), \circ d_t \nabla \Psi_t(x) \rangle \, \mu_t(dx) + \sum_{i=0}^{N} \Big[\int_M \big(\langle \nabla f(x), \nabla_{\nabla \phi_i} \nabla \Psi_t(x) \rangle \Big] \circ dB_t^i = 0,$$

or (5.3.21) holds. Now transforming Stratanovich stochastic calculus to Itô stochastic calculus yields the equation (5.3.20). $\square$

Proposition 5.3.6. *For such a solution to* (5.3.20)*, we have* $||V_{\Psi_t}||_{\mathbf{T}_{\mu_t}} = ||V_{\Psi_0}||_{\mathbf{T}_\mu}$ *for all* $t \in [0,1]$.

Proof. Using (5.3.23), we have formally,

$$
\begin{aligned}
&d_t \int_M \langle \nabla\Psi_t, \nabla\Psi_t\rangle \, \mu_t(dx) = d_t \int_M \langle \nabla\Psi_t, \nabla\Psi_t\rangle \, \rho_t \, \mu(dx) \\
&= 2\int_M \langle \nabla\Psi_t, \circ d_t \nabla\Psi_t\rangle \, \rho_t \, \mu(dx) - \sum_{i=0}^N \int_M \langle \nabla\Psi_t, \nabla\Psi_t\rangle \, \mathsf{div}_{\mu_t}(\nabla\phi_i) \, \rho_t \, \mu(dx) \circ dB_t^i \\
&= 2\int_M \langle \nabla\Psi_t, \ \circ d\nabla\Psi_t + \sum_{i=0}^N \nabla_{\nabla\phi_i}\nabla\Psi_t \circ dB_t^i\rangle \, \mu_t(dx) = 0
\end{aligned}
$$

due to (5.3.21). □

We will give a rigorous proof of above result in the case where $M = \mathbb{T}^d$, a d-dimensional torus. First we recall the following Kunita-Itô-Wenzell formula [dLHLT20]:

Theorem 5.3.7. *Let* $t \to K(t,\cdot) \in C^2(\mathbb{T}^d)$ *be a continuous adapted semimartingale, given by*

$$
K(t,x) = K(0,x) + \int_0^t G(s,x)ds + \sum_{j=1}^N \int_0^t H_j(s,x)dB_s^j, \quad t \in [0,T]
$$

where $(B_t^1, \ldots, B_t^N)$ *is a standard Brownian motion on* $\mathbb{R}^N$*, and* $G \in L^1([0,T], C^2(\mathbb{T}^d))$*,* $H \in L^2([0,T], C^2(\mathbb{T}^d))$ *are adapted semimartingales. Let* X_t *be the solution of the following Stratanovich SDE:*

$$
dX_t = b(t, X_t)dt + \sum_{j=1}^N \xi_j(t, X_t) \circ dB_t^j, \quad X_0(x) = x
$$

which is assumed to be a C^1 *diffeomorphism,* $b(t,\cdot) \in W^{1,1}(\mathbb{T}^d, \mathbb{R}^d)$ *,* $\xi_j(t,\cdot) \in C^2(\mathbb{T}^d, \mathbb{R}^d)$ *and*

$$
\int_0^T \Big[|b(s, X_s(x)) + \frac{1}{2}\sum_{j=1}^N \xi_j \cdot \nabla\xi_j(s, X_s(x))| + \sum_{j=1}^N |\xi_j(s, X_s(x))|^2\Big] ds < \infty, \quad x \in \mathcal{T}^d.
$$

Then the following formula holds:

$$
\begin{aligned}
K(t, X_t(x)) = K(0, x) &+ \int_0^t G(s, X_s(x))ds + \sum_{j=1}^{N} \int_0^t H_j(s, X_s(x))dB_s^j \\
&+ \int_0^t \langle \nabla K, b\rangle(s, X_s(x))ds + \sum_{j=1}^{N} \int_0^t \langle \nabla K, \xi_j\rangle(s, X_s(x))dB_s^j \\
&+ \frac{1}{2}\sum_{j=1}^{N} \int_0^t \langle \nabla\langle \nabla K, \xi_j\rangle, \xi_j\rangle(s, X_s(x))ds + \sum_{j=1}^{N} \int_0^t \langle \nabla H_j, \xi_j\rangle(s, X_s(x))ds.
\end{aligned}
$$

Using this theorem, we can prove the conservation of norm.

Theorem 5.3.8. *If $\{\Psi_t,\ t \in [0,1]\}$ is a $\mathcal{L}^2\left([0,1]\times\Omega, C^3(\mathbb{T}^d)\right)$ solution of strong S.P.T equation* (5.3.20)*, then Ψ_t is a solution to weak S.P.T equation. Furthermore, for $t \in [0,1]$,*

$$
\int_{\mathbb{T}^d} |\nabla\Psi_t|^2 \mu_t(dx) = \int_{\mathbb{T}^d} |\nabla\Psi_0|^2\ \mu(dx). \tag{5.3.24}
$$

Proof. Let $F_t(x) = \langle \nabla f, \nabla\Psi_t(x)\rangle$. We have

$$
\begin{aligned}
F_t(x) =& F_0(x) + \int_0^t \left\langle \nabla f, \Pi_{\mu_s}\left(-\nabla_{\nabla\phi_0}\Psi_s + R_s^{\Psi_s} + S_s^{\Psi_s}\right)\right\rangle(x)ds \\
&- \sum_{j=1}^{N} \int_0^t \langle \nabla f, \Pi_{\mu_s}\left(\nabla_{\nabla\phi_j}\nabla\Psi_s\right)\rangle(x)dB_s^j.
\end{aligned}
$$

Let $\mathbb{L}$ be the infinitesimal generator corresponding to diffusion (5.3.1), which satisfies, for $\forall f \in C^2$,

$$
\mathbb{L}f = \frac{1}{2}\sum_{j=1}^{N} \langle \nabla\langle \nabla f, \nabla\phi_j\rangle, \nabla\phi_j\rangle.
$$

Then, by Kunita-Ito-Wenzell formula,

$$
\begin{aligned}
F_t(X_t) = F_0(x) &+ \int_0^t \left\langle \nabla f, \Pi_{\rho_s}\left(-\nabla^2\Psi_s\nabla\phi_0 + \frac{1}{2}R_s^{\Psi_s} + \frac{1}{2}S_s^{\Psi_s}\right)\right\rangle (X_s)ds \\
&- \sum_{j=1}^N \int_0^t \left\langle \nabla f, \Pi_{\rho_s}\left(\nabla^2\Psi_s\nabla\phi_j\right)\right\rangle (X_s)dB_s^j + \int_0^t \left\langle \nabla F_s, \nabla\phi_0\right\rangle (X_s)ds \\
&+ \sum_{j=1}^N \int_0^t \left\langle \nabla F_s, \nabla\phi_j\right\rangle (X_s)dB_s^j + \int_0^t \mathbb{L}F_s(X_s)ds \\
&- \sum_{j=1}^N \int_0^t \left\langle \nabla\left\langle \nabla f, \Pi_{\rho_s}\left(\nabla^2\Psi_s\nabla\phi_j\right)\right\rangle, \nabla\phi_j\right\rangle (X_s)ds.
\end{aligned}
$$

Denote

$$
\begin{aligned}
A_s =& \left\langle \nabla f, \Pi_{\rho_s}\left(-\nabla^2\Psi_s\nabla\phi_0 + \frac{1}{2}R_s^{\Psi_s} + \frac{1}{2}S_s^{\Psi_s}\right)\right\rangle (X_s) + \left\langle \nabla F_s, \nabla\phi_0\right\rangle (X_s) \\
&+ \mathbb{L}F_s(X_s) - \left\langle \nabla\left\langle \nabla f, \Pi_{\rho_s}\left(\nabla^2\Psi_s\nabla\phi_j\right)\right\rangle, \nabla\phi_j\right\rangle (X_s); \\
M_s =& -\sum_{j=1}^N \left\langle \nabla f, \Pi_{\rho_s}\left(\nabla^2\Psi_s\nabla\phi_j\right)\right\rangle (X_s) + \sum_{j=1}^N \left\langle \nabla F_s, \nabla\phi_j\right\rangle (X_s).
\end{aligned}
$$

Since $\Psi_t \in \mathcal{L}^2\left([0,1]\times\Omega, C^3(\mathbb{T}^d)\right)$, $\phi_j \in C^\infty$,

$$||\left\langle \nabla F_s, \nabla\phi_0\right\rangle + \mathbb{L}F_s|| \le K_1||\Psi_s||_{C^3}.$$

The boundedness of the left two terms in A_s need a uniform estimate on $\Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)$. In fact, it is known that $\rho_t \in \mathcal{C}^2$ and $\rho_t,\ \nabla\rho_t$ are continuous functions on $[0,1]\times\mathbb{T}^d$ for almost surely ω. Thus, for the elliptic operators defined by $L_{\rho_t}u = \rho_t\Delta u + \langle\nabla\rho_t, \nabla u\rangle$, we have the uniform bound on the coefficients:

$$\min_{[0,1]\times\mathbb{T}^d} \rho > \lambda(\omega); \quad \max_{[0,1]\times\mathbb{T}^d}\{\rho, |\partial_x\rho|\} \le \Lambda(\omega).$$

For the unique classical solution u of elliptic equation $L_{\rho_t}u = f$, we have, by Shauder estimate,

$$||u||_{C^2} \le C_1(d,\lambda,\Lambda)(||u||_C + ||f||_C).$$

On the other hand, it can be proved that for $\forall V \in C^k(\mathcal{T}^d;\mathbb{R}^d)$, $m \le k-1$,

$$||\nabla.(\rho_t V)||_{C^m} \le C_2(\lambda,\Lambda)||V||_{C^{m+1}}.$$

Therefore, by (5.1.4),

$$||\Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)||_{C^1} = ||\nabla L_{\rho_t}^{-1}\nabla.(\rho_t\nabla^2\Psi_t\nabla\phi_j)||_{C^1} \leq C||\Psi_t||_{C^3} \tag{5.3.25}$$

where C is not dependent on t. Thus,

$$|A_s| \leq K_2||\Psi_s||_{C^3}. \tag{5.3.26}$$

Again, by applying (5.3.25) , we also find

$$|M_s| \leq K_3||\Psi_s||_{C^2}. \tag{5.3.27}$$

Combined with (5.3.26) and (5.3.27), we prove that , for almost surely $\omega \in \Omega$,

$$\int_{\mathcal{T}^d}\int_0^t |A_s| ds\rho_0(x)dx < \infty \ ; \quad \int_{\mathcal{T}^d}\left(\int_0^t |M_s|^2 ds\right)^{\frac{1}{2}}\rho_0 dx < \infty.$$

Thus, by applying stochastic Fubini's theorem, we get

$$\int_{\mathbb{T}^d}\langle\nabla f, \nabla\phi_t\rangle\rho_t dx = \int_0^t\int_{\mathbb{T}^d} A_s\rho_0 dxds + \int_0^t\Big(\int_{\mathbb{T}^d} M_s\rho_0 dx\Big)dB_s^j.$$

By direct substitution and integration by part, we proved ϕ_t is a solution to weak S.P.T. equation .

The conservation of norm can be proved by the same method by defining $G_t(x) = |\nabla\Psi_t|^2$. By Ito formula, we have

$$\begin{aligned}
d_tG_t(x) &= 2\langle\nabla\Psi_t(x), d_t\nabla\Psi_t(x)\rangle + d_t < \nabla\Psi_t(x) > \\
&= 2\left\langle\nabla\Psi_t(x), \Pi_{\rho_t}\left(-\nabla^2\Psi_t\nabla\phi_0 + \frac{1}{2}R_t^{\Psi_t} + \frac{1}{2}S_t^{\Psi_t}\right)(x)\right\rangle dt \\
&+ \sum_{j=1}^N\left\langle\Pi_{\rho_t}\left(\nabla^2\phi_t\nabla\phi_j\right)(x), \Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)(x)\right\rangle dt \\
&- \sum_{j=1}^N 2\left\langle\nabla\Psi_t(x), \Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)(x)\right\rangle dB_t^j.
\end{aligned}$$

Based on estimates above, we can again apply two major tools : Kunita-Ito-Wenzell formula and stochastic Fubini theorem . In fact, we find

$$d_t||\nabla\Psi_t||^2_{\rho_t} = \langle\nabla\Psi_t, -2\nabla^2\Psi_t\nabla\phi_0\rangle_{\rho_t}dt - 2\sum_{j=1}^N\langle\nabla\Psi_t, \nabla^2\Psi_t\nabla\phi_j\rangle_{\rho_t}dB_t^j \tag{5.3.28}$$

$$+\sum_{j=1}^N||\Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)||_{\rho_t}dt + \langle\nabla\Psi_t, R_t^{\Psi_t} + S_t^{\Psi_t}\rangle_{\rho_t}dt \tag{5.3.29}$$

$$+\langle\nabla G_t, \nabla\phi_0\rangle_{\rho_t}dt + \sum_{j=1}^N\langle\nabla G_t, \nabla\phi_j\rangle_{\rho_t}dB_t^j \tag{5.3.30}$$

$$+\sum_{j=1}^N\frac{1}{2}\langle\nabla\langle\nabla G_t, \nabla\phi_j\rangle, \nabla\phi_j\rangle_{\rho_t} - 2\langle\nabla\langle\nabla\Psi_t, \Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)\rangle, \nabla\phi_j\rangle_{\rho_t}dt. \tag{5.3.31}$$

We have (5.3.28) $+$ (5.3.30) $= 0$. By integration by parts, we have

$$\text{(5.3.29)} = \sum_{j=1}^N||\Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)||_{\rho_t} + \sum_{j=1}^N\langle\nabla\Psi_t, \nabla\left(\Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)\right)\nabla\phi_j\rangle_{\rho_t}$$
$$+\sum_{j=1}^N\int_{\mathbb{T}^d}\left\langle\nabla\Psi_t, \nabla^2\Psi_t\nabla\phi_j - \Pi_{\rho_t}(\nabla^2\Psi_t\nabla\phi_j)\right\rangle\nabla.\left(\nabla\phi_j\rho_t\right)dx,$$

which is equal to

$$\sum_{j=1}^N||\Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)||_{\rho_t} + \sum_{j=1}^N\langle\nabla\Psi_t, \nabla\left(\Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)\right)\nabla\phi_j\rangle_{\rho_t}$$
$$+\sum_{j=1}^N\langle\nabla\langle\nabla\Psi_t, \Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)\rangle, \nabla\phi_j\rangle_{\rho_t} - \langle\nabla\langle\frac{1}{2}\nabla G_t, \nabla\phi_j\rangle, \nabla\phi_j\rangle_{\rho_t}.$$

Thus,

$$\text{(5.3.29)} + \text{(5.3.31)} = \sum_{j=1}^N||\Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)||_{\rho_t} - \langle\nabla^2\Psi_t\nabla\phi_j, \Pi_{\rho_t}\left(\nabla^2\Psi_t\nabla\phi_j\right)\rangle_{\rho_t} = 0.$$

The proof is complete. □

5.4 Q-Wiener process on $\mathbb{P}_2(M)$

Now we will construct a non-degenerated diffusion process on $\mathbb{P}_2(M)$. Let $\{\varphi_n;\ n \geq 0\}$ be the eigenfunctions of the Laplace Δ on M:

$$-\Delta\varphi_n = \lambda_n\varphi_n.$$

We have $\lambda_0 = 0$ and $\varphi_0 = 1$. It is well known that

$$\lambda_n \sim n^{2/\dim(M)} \quad \text{as } n \to +\infty. \tag{5.4.1}$$

The functions φ_n are smooth, and $\{\varphi_n;\ n \geq 0\}$ forms an orthonormal basis of $L^2(M, dx)$: $\int_M \varphi_n\varphi_m\, dx = \delta_{nm}$. The system $\big\{\dfrac{\nabla\varphi_n}{\sqrt{\lambda_n}};\ n \geq 1\big\}$ is orthonormal, so that $\big\{V_{\varphi_n/\sqrt{\lambda_n}}; n \geq 1\big\}$ is an orthonormal basis of $\mathbf{T}_{dx}$. A function f on M is in Sobolev space $H^k(M)$ if

$$||f||^2_{H^k} = \int_M |(I-\Delta)^{k/2}f|^2\, dx < +\infty;$$

it is obvious that $||f||^2_{H^k} = \sum_{n\geq 0}(1+\lambda_n)^k\Big(\int_M f(x)\varphi_n(x)\, dx\Big)^2$. By the Sobolev embedding inequality, for $k > \dfrac{\dim(M)}{2} + q$,

$$||f||_{C^q} \leq C\, ||f||_{H^k}. \tag{5.4.2}$$

In particular, $||f||_\infty \leq C\, ||f||_{H^k}$ for $k > \dim(M)/2$.

Lemma 5.4.1. *There is a universal constant $C > 0$, independent of $i \in \mathbb{N}^*$ such that, for $k > \dim(M)/2$, $t \in [0, 1]$, for almost surely ω,*

$$\int_M |\varphi_i|^2\, \mu_t(x) \leq C\,(1+\lambda_i)^k. \tag{5.4.3}$$

Proof. We have $\int_M |\varphi_i(x)|^2\, \mu_t(dx) \leq ||\varphi_i||^2_\infty$, which is dominated, according to (5.4.2), by

$$C\int_M |(I-\Delta)^{k/2}\varphi_i|^2\, dx = C\,(1+\lambda_i)^k \int_M \varphi_i^2\, dx = C\,(1+\lambda_i)^k.$$

The result (5.4.3) follows. □

In this section, we are given a sequence of strictly positive real numbers $\{a_n;\ n \geq 1\}$. Consider the following SDE on M:

$$dX_t^N = \sum_{i=1}^N a_i \nabla\varphi_i(X_t^N)\circ dB_t^i, \tag{5.4.4}$$

where $\{B_t^i;\ i\geq 1\}$ is a sequence of independent standard Brownian motions on $\mathbb{R}$. For a given probability measure $d\mu=\rho\,dx$ with $\rho\in C^2$ and $\rho>0$, we consider $\mu_t^N=(X_t^N)_{\#}\mu$. It has been shown in Section 2 that $\{\mu_t^N; t\geq 0\}$ solves the following SDE on $\mathbb{P}_2(M)$:

$$\circ d_t^I\mu_t^N = \sum_{i=1}^N a_i V_{\varphi_i}(\mu_t^N)\ \circ dB_t^i,\quad \mu_0=\mu. \tag{5.4.5}$$

Let Ent be the entropy functional on $\mathbb{P}_2(M)$. By Proposition 5.2.4, we have

$$d_t\mathrm{Ent}(\mu_t^N) = \sum_{i=1}^N a_i\,\langle\bar\nabla\mathrm{Ent}, V_{\varphi_i}\rangle_{\mathbf{T}_{\mu_t^N}} dB_t^i + \sum_{i=1}^N \frac{a_i^2}{2}\big(\bar D_{V_{\varphi_i}}\bar D_{V_{\varphi_i}}\mathrm{Ent}\big)(\mu_t^N)\,dt.$$

It follows that for any $t\in[0,1]$,

$$\mathbb{E}\big(\mathrm{Ent}(\mu_t^N)\big) = \mathrm{Ent}(\mu) + \sum_{i=1}^N\frac{a_i^2}{2}\int_0^t \mathbb{E}\big(\bar D_{V_{\varphi_i}}\bar D_{V_{\varphi_i}}\mathrm{Ent}\big)(\mu_s^N)\big)\,ds. \tag{5.4.6}$$

Lemma 5.4.2. *For $k>\dim(M)/2+1$, there is a universal constant $C>0$ such that, for any $i\geq 1$, $t\in[0,1]$, almost surely ω, such that*

$$|\big(\bar D_{V_{\varphi_i}}\bar D_{V_{\varphi_i}}\mathrm{Ent}\big)(\mu_t^N)|\leq C\,\lambda_i(1+\lambda_i)^k. \tag{5.4.7}$$

Proof. By Formula (5.0.6), we have

$$\begin{aligned}\big(\bar D_{V_{\varphi_i}}\bar D_{V_{\varphi_i}}\mathrm{Ent}\big)(\mu_t) &= -\int_M\langle\nabla\Delta\varphi_i,\ \nabla\varphi_i\rangle\,\mu_t(dx) = \lambda_i\int_M|\nabla\varphi_i(x)|^2\,\mu_t(x)\\ &\leq \lambda_i\,||\nabla\varphi_i||_\infty^2\leq\lambda_i\,C\,||\varphi_i||_{H^k}^2 = C\,\lambda_i(1+\lambda_i)^k.\end{aligned}$$

The result (5.4.7) follows. □

Theorem 5.4.3. *For an integer $k>\dim(M)/2+1$ given, if*

$$\sum_{i\geq 1} a_i^2\,\lambda_i(1+\lambda_i)^k<+\infty, \tag{5.4.8}$$

then the family $\{\mu_t^N; N\geq 1\}$ is tight.

Proof. Let $\rho_t^N(\omega,x)$ be the density of μ_t^N with respect to Riemannian measure dx, then for any $N\geq 1$, according to (5.4.6) and (5.4.7),

$$\int_{[0,1]\times\Omega\times M}\rho_t^N(\omega,x)\log(\rho_t^N(\omega,x))\,dtP(d\omega)\,dx=\int_{[0,1]}\mathbb{E}\big(\mathrm{Ent}(\mu_t^N)\big)\,dt$$
$$\leq \mathrm{Ent}(\mu)+\frac{C}{2}\sum_{i\geq 1}a_i^2\,\lambda_i(1+\lambda_i)^2,$$

which is finite under Condition (5.4.8). The result follows. □

In fact, we have a stronger result, which says that the sequence $\{\rho_t^N;\geq 1\}$ is in a weakly compact subset in $L^1([0,1]\times\Omega\times M)$. Therefore there is $\rho\in L^1$ and up to a subsequence, for any $\alpha\in L^\infty([0,1])$, $\xi\in L^\infty(\Omega)$ and $g\in L^\infty(M)$, such that,

$$\lim_{N\to+\infty}\int_{[0,1]\times\Omega\times M}\alpha(t)\xi(\omega)g(x)\,\rho_t^N(\omega,x)\,dtP(d\omega)\,dx$$
$$=\int_{[0,1]\times\Omega\times M}\alpha(t)\xi(\omega)g(x)\,\rho_t(\omega,x)\,dtP(d\omega)\,dx.$$

It is obvious that for almost all (t,ω), $\rho_t(\omega,x)\geq 0$ and $\int_M\rho_t(\omega,x)dx=1$.

In order to obtain stronger results, we have to deal with the convergence of diffusion processes $\{X_t^N;\ n\geq 1\}$ appeared in (5.4.4). First of all, we consider the following Random series

$$\sum_{i=1}^{+\infty}a_i\nabla\phi_i\,B_t^i. \tag{5.4.9}$$

Note that for any smooth function f on M, $-\nabla\Delta f=\square\nabla f$, so that for any $k\geq 1$,

$$\nabla(I-\Delta)^{k/2}f=(I+\square)^{k/2}\nabla f.$$

Let $q\geq p$ be two integers,

$$\sum_{i=p}^{q}(I+\square)^{k/2}\Big(a_i\nabla\varphi_iB_t^i\Big)=\sum_{i=p}^{q}a_i(1+\lambda_i)^{k/2}\,\nabla\varphi_iB_t^i.$$

Then

$$\mathbb{E}\Big[\int_M\Big|\sum_{i=p}^{q}(I+\square)^{k/2}\Big(a_i\nabla\varphi_iB_t^i\Big)\Big|^2\,dx\Big]=\sum_{i=p}^{q}a_i^2\lambda_i(1+\lambda_i)^k\,t.$$

Under Condition (5.4.8), almost surely the Random series (5.4.9) converges in $H^k(M)$ uniformly in $t\in[0,1]$; let $W_t(\omega,x)$ be the sum of this series, which gives rise to a continuous martingale taking values in $H^k(TM)$. When

$k > \dim(M)/2 + 2$, the vector field $x \to W_t(\omega, x)$ is of the class $C^{2,\alpha}$. By the classical theory of stochastic flow [Kun97, Mal97, Elw92], there is a C^1-diffeomorphisms $X_t(\omega, \cdot)$ of M, solving the SDE on $\mathrm{Diff}^1(M)$:

$$dX_t = \circ dW_t(X_t)$$

or more explicitly

$$dX_t = \sum_{i=1}^{+\infty} a_i \nabla\varphi_i(X_t) \circ dB_t^i, \quad X_0(\omega, x) = x. \tag{5.4.10}$$

Proposition 5.4.4. *Assume that, for $k > \dim(M)/2 + 3$,*

$$\beta := \sum_{i=1}^{+\infty} a_i^2 (1 + \lambda_i)^k < +\infty. \tag{5.4.11}$$

Then almost surely, $X_t^N(x)$ converges to $X_t(x)$ uniformly in $(t, x) \in [0, 1] \times m$, as $N \to +\infty$.

Proof. Put

$$A_N = \frac{1}{2} \sum_{i=1}^{N} a_i^2\, \nabla_{\nabla\varphi_i}(\nabla\varphi_i).$$

Using (5.4.2), there is a constant $C > 0$ such that for $k > \dim(M)/2 + 3$,

$$||A_N||_\infty \le C \sum_{i=1}^{+\infty} a_i^2 (1 + \lambda_i)^k \text{ and } ||\nabla A_N||_\infty \le C \sum_{i=1}^{+\infty} a_i^2 (1 + \lambda_i)^k.$$

Again

$$\sum_{i=1}^{N} a_i^2\, ||\nabla^2\varphi_i||_\infty \le C \sum_{i=1}^{+\infty} a_i^2 (1 + \lambda_i)^k.$$

These uniform estimates allow us to conclude. □

Theorem 5.4.5. *Let $d\mu = \rho\, dx$ be a probability measure on M with a strictly positive C^2 density ρ and $\mu_t = (X_t)_{\#}\mu$. Then under Condition* (5.4.11)*, $\{\mu_t;\ t \in [0, 1]\}$ is a solution to the following SDE on $\mathbb{P}_2(M)$:*

$$\circ d_t^I \mu_t = \sum_{i=1}^{+\infty} a_i V_{\varphi_i}(\mu_t)\ \circ dB_t^i, \quad \mu_0 = \mu. \tag{5.4.12}$$

Proof. Note first that

$$\sup_{t\in[0,1]} W_2^2(\mu_t,\mu_t^N) \le \int_M \sup_{t\in[0,1]} d_M^2\big(X_t(x),X_t^N(x)\big)\,\mu(dx);$$

then Proposition 5.4.4 implies that almost surely, μ_t^N converges to μ_t uniformly in $t\in[0,1]$ as $N\to+\infty$. Let F be a polynomial on $\mathbb{P}_2(M)$, by Proposition 5.2.2, we have

$$F(\mu_t^N) = F(\mu) + \sum_{i=1}^N \int_0^t (a_i\bar{D}_{V_{\varphi_i}}F)(\mu_s^N)\,dB_s^i + \frac{1}{2}\sum_{i=1}^N \int_0^t a_i^2(\bar{D}_{V_{\varphi_i}}\bar{D}_{V_{\varphi_i}}F)(\mu_s^N)\,ds.$$

Letting $N\to+\infty$ yields

$$F(\mu_t) = F(\mu) + \sum_{i=1}^{+\infty} \int_0^t (a_i\bar{D}_{V_{\varphi_i}}F)(\mu_s)\,dB_s^i + \frac{1}{2}\sum_{i=1}^{+\infty} \int_0^t a_i^2(\bar{D}_{V_{\varphi_i}}\bar{D}_{V_{\varphi_i}}F)(\mu_s)\,ds.$$

The entropy functional $\mu\to\mathrm{Ent}(\mu)$ is not continuous. However, if we denote by ρ_t^N the density of μ_t^N with respect to dx, then $\rho_t^N\log(\rho_t^N)$ converges to $\rho_t\log(\rho_t)$ almost surely, and according to [FLT10], the family $\{\rho_t^N\log(\rho_t^N); N\ge 1\}$ is uniformly integrable, so that we have

$$\lim_{N\to+\infty}\mathrm{Ent}(\mu_t^N) = \mathrm{Ent}(\mu_t).$$

By Proposition 5.2.3, we have

$$\mathrm{Ent}(\mu_t^N) = \mathrm{Ent}(\mu) + \sum_{i=1}^N \int_0^t a_i\,\langle\bar{\nabla}\mathrm{Ent}, V_{\varphi_i}\rangle_{\mathbf{T}_{\mu_s^N}}\,dB_s^i + \sum_{i=1}^N \frac{a_i^2}{2}\int_0^t \big(\bar{D}_{V_{\varphi_i}}\bar{D}_{V_{\varphi_i}}\mathrm{Ent}\big)(\mu_s^N)\,ds.$$

Letting $N\to+\infty$ yields

$$\mathrm{Ent}(\mu_t) = \mathrm{Ent}(\mu) + \sum_{i=1}^{+\infty} \int_0^t a_i\,\langle\bar{\nabla}\mathrm{Ent}, V_{\varphi_i}\rangle_{\mathbf{T}_{\mu_s}}\,dB_s^i + \sum_{i=1}^{+\infty} \frac{a_i^2}{2}\int_0^t \big(\bar{D}_{V_{\varphi_i}}\bar{D}_{V_{\varphi_i}}\mathrm{Ent}\big)(\mu_s)\,ds.$$

□

Let $F_3(\mu)=\int_{M\times M} W(x,y)\mu(dx)\mu(dy)$ be the Example 3.

Theorem 5.4.6. *Under condition* (5.4.11)*, there is a unique solution* (X_t,μ_t) *to the following Mckean-Vlasov equation:*

$$dX_t = \sum_{i=1}^{+\infty} a_i\nabla\varphi_i(X_t)\circ dB_t^i + \nabla\Phi(X_t,\mu_t)\,dt,\ \mu_t=(X_t)_{\#}\mu, \tag{5.4.13}$$

where $\Phi(x,\mu)=\displaystyle\int_M W(x,y)\mu(dy)$*. Moreover,* $\{\mu_t; t\in[0,1]\}$ *is a solution to*

$$\circ d_t^I \mu_t = \sum_{i=1}^{+\infty} a_i V_{\varphi_i}(\mu_t) \circ dB_t^i + \bar{\nabla} F_3(\mu_t)\, dt, \quad \mu_0 = \mu. \tag{5.4.14}$$

Remark 5.4.7. Let $\mu_t^{\mathbb{P}}$ be the law of μ_t in the Wasserstein space $\mathbb{P}_2(M)$. By the Bakry-Emery's Γ_2 theory, the asymptotic behavior of $\mu_t^{\mathbb{P}}$ as $t \to +\infty$ is dependent of

$$\mathrm{Ric}^{\mathbb{P}} + \bar{\nabla}^2 F_3,$$

where $\mathrm{Ric}^{\mathbb{P}}$ is the "Ricci tensor" associated to the Q-Brownian motion.

Remark 5.4.8. Since $\mathbb{P}_2(M)$ is compact, it is hopeful that for some constant $\kappa \in \mathbb{R}$

$$\langle \mathrm{Ric}^{\mathbb{P}} V_\phi, V_\phi \rangle_{\mathbf{T}_\mu} \geq \kappa |V_\phi|_{\mathbf{T}_\mu}, \quad \phi \in C^\infty(M),\ \mu \in \mathbb{P}_2(M).$$

Now by Proposition 5.0.1, if the function W is such that

$$\int_{M\times M} Hess_{x,y} W\big(\nabla\phi(x), \nabla\phi(y)\big)\mu(dx)\mu(dy) \geq \kappa_1\, |V_\phi|^2_{\mathbf{T}_\mu}, \quad \phi \in C^\infty(M),\ \mu \in \mathbb{P}_2(M) \tag{5.4.15}$$

with $\kappa + \kappa_1 > 0$, then as $t \to +\infty$, $\mu_t^{\mathbb{P}}$ converges to a Gaussian like probability measure γ_∞ on $\mathbb{P}_2(M)$.

5.5 Stochastic parallel translation on $\mathbb{P}(\mathbb{T})$

For simplicity, we consider the following SDE on $\mathbb{T}$:

$$dX_t = \nabla\phi(X_t) \circ dB_t.$$

Let $\mu_t = (X_t)_\#(dx)$ and $d\mu_t = \rho_t\, dx$, that is to say that the initial measure μ_0 is the Haar measure dx. Suppose there is a solution $\{\partial_x \Psi_t;\ t \in [0,1]\}$ to the equation of strong parallel translations:

$$d_t \partial_x \Psi_t = \Pi_{\rho_t}\Big(R_t^{\Psi_t} + S_t^{\Psi_t}\Big)\, dt - \Pi_{\rho_t}\Big(\partial_x^2 \Psi_t\, \partial_x \phi\Big)\, dB_t. \tag{5.5.1}$$

Let $f_t = \partial_x \Psi_t(X_t)$. Then by Kunita-Itô-Wentzell formula,

$$d_t f_t = \frac{1}{\rho_t(X_t)}\, K_t^{\partial_x \Psi_t}\, dB_t + \frac{1}{2} K_t^{\partial_x \Psi_t} \frac{\partial_x^2 \phi}{\rho_t}(X_t)\, dt - \frac{1}{2} H_t^{\partial_x \Psi_t}\, \frac{1}{\rho_t}(X_t)\, dt,$$

where

$$K_t^{\partial_x \Psi_t} = -\frac{\int_{\mathbb{T}} \partial_x \Psi_t\, \partial_x^2 \phi\, dx}{\int_{\mathbb{T}} \frac{dx}{\rho_t}},$$

and

$$H_t^{\partial_x \Psi_t} = \frac{\int_{\mathbb{T}} \Big[\partial_x \Psi_t\, \partial_x(\partial_x^2\phi\, \partial_x\phi) + 3\frac{K_t^{\partial_x\Psi_t}}{\rho_t}\partial_x^2\phi\Big]\, dx}{\int_{\mathbb{T}} \frac{dx}{\rho_t}}.$$

Using the notation

$$\hat{\rho} = \frac{1}{\rho\, \int_{\mathbb{T}} \frac{dx}{\rho}},$$

we will simplify expression for K_t as well for H_t. We have

$$\frac{1}{\rho_t}\, K_t^{\partial_x\Psi_t} = -\Big(\int_{\mathbb{T}} \partial_x \Psi_t\, \phi''\, dx\Big)\, \hat{\rho}_t,$$

and

$$\frac{1}{\rho_t}\, H_t^{\partial_x\Psi_t} = \Big(\int_{\mathbb{T}} \Big[\partial_x\Psi_t\, \partial_x(\partial_x^2\phi\, \partial_x\phi) - 3\hat{\rho}_t\, \partial_x^2\phi\, \int_{\mathbb{T}} \partial_x\Psi_t\, \partial_x^2\phi\, dx\Big]\, dx\Big)\, \hat{\rho}_t.$$

Now remark that

$$\int_{\mathbb{T}} \partial_x \Psi_t\, \phi''\, dx = \int_{\mathbb{T}} (\partial_x\Psi_t)(X_t)\, \partial_x^2\phi(X_t)\, \frac{1}{\rho_t(X_t)}\, dx = \int_{\mathbb{T}} f_t\, \frac{\partial_x^2\phi}{\rho_t}(X_t)\, dx.$$

In the same way,

$$\int_{\mathbb{T}} \Big[\partial_x\Psi_t\, \partial_x(\partial_x^2\phi\, \partial_x\phi)\Big]\, dx = \int_{\mathbb{T}} f_t\, \frac{\partial_x(\partial_x^2\phi\, \partial_x\phi)}{\rho_t}(X_t)\, dx.$$

Set

$$a_t = \frac{\partial_x^2\phi}{\rho_t}(X_t), \quad b_t = \frac{\partial_x(\partial_x^2\phi\, \partial_x\phi)}{\rho_t}(X_t).$$

Then we get the following equation for $\{f_t;\ t \in [0,1]\}$:

$$\begin{aligned} d_t f_t = &-\Big(\int_{\mathbb{T}} f_t\, a_t\, dx\Big)\, \hat{\rho}_t(X_t)\, dB_t - \frac{1}{2}\Big(\int_{\mathbb{T}} f_t\, a_t\, dx\Big)\, \big(\hat{\rho}_t\partial_x^2\phi\big)(X_t)\, dt \\ &+ \frac{1}{2}\Big(\int_{\mathbb{T}} f_t\, b_t\, dx\Big)\, \hat{\rho}_t(X_t)\, dt + \frac{3}{2}\Big(\int_{\mathbb{T}} f_t\, a_t\, dx\Big)\, \Big(\int_{\mathbb{T}} \partial_x^2\phi\, \hat{\rho}_t\, dx\Big)\, \hat{\rho}_t(X_t)\, dt. \end{aligned} \tag{5.5.2}$$

We have

$$\int_{\mathbb{T}} |a_t|^2\, dx = \int_{\mathbb{T}} \Big(\frac{\partial_x^2\phi}{\rho_t}\Big)^2\, \rho_t\, dx \leq ||\partial_x^2\phi||_\infty^2\, \int_{\mathbb{T}} \frac{dx}{\rho_t},$$

and

$$\int_{\mathbb{T}} \hat{\rho}_t(X_t)^2\, dx = \frac{1}{\left(\int_{\mathbb{T}} \frac{dx}{\rho_t}\right)^2} \int_{\mathbb{T}} \Big(\frac{1}{\rho_t}\Big)^2 \rho_t\, dx = 1/\int_{\mathbb{T}} \frac{dx}{\rho_t},$$

We get the following key estimate:

$$\Big(\int_{\mathbb{T}} \hat{\rho}_t(X_t)^2\, dx\Big)\Big(\int_{\mathbb{T}} |a_t|^2\, dx\Big) \le ||\partial_x^2\phi||_\infty^2, \tag{5.5.3}$$

and

$$\Big(\int_{\mathbb{T}} \hat{\rho}_t(X_t)^2\, dx\Big)\Big(\int_{\mathbb{T}} |b_t|^2\, dx\Big) \le ||\partial_x(\partial_x^2\phi\, \partial_x\phi)||_\infty^2. \tag{5.5.4}$$

Theorem 5.5.1. *There is a unique strong solution* $\{f_t; t \in [0,1]\}$ *to the equation* (5.5.2) *such that* $f_0 = \partial_x \Psi_0$.

Proof. The estimate (5.5.3) allows us to use the Picard iteration. Let $f_t^0 = \partial_x\Psi_0$, and

$$\begin{aligned} f_t^{n+1} = \partial_x\Psi_0 &- \int_0^t \Big(\int_{\mathbb{T}} f_s^n\, a_s\, dx\Big)\hat{\rho}_s(X_s)\, dB_s - \frac{1}{2}\int_0^t \Big(\int_{\mathbb{T}} f_s^n\, a_s\, dx\Big)\big(\hat{\rho}_s\partial_x^2\phi)(X_s)\, ds \\ &+ \frac{1}{2}\int_0^t \Big(\int_{\mathbb{T}} f_s^n\, b_s\, dx\Big)\hat{\rho}_s(X_s)\, ds + \frac{3}{2}\int_0^t \Big(\int_{\mathbb{T}} f_s^n\, a_s\, dx\Big)\Big(\int_{\mathbb{T}} \partial_x^2\phi\, \hat{\rho}_s\, dx\Big)\hat{\rho}_s(X_s)\, ds. \end{aligned} \tag{5.5.5}$$

Set

$$M_t(x) = \int_0^t \Big(\int_{\mathbb{T}} (f_s^n - f_s^{n-1})\, a_s\, dx\Big)\hat{\rho}_s(X_s)(x)\, dB_s.$$

We have

$$\begin{aligned} &\mathbb{E}\Big[\sup_{0\le t\le t}\int_{\mathbb{T}} M_s^2\, dx\Big] \le \int_{\mathbb{T}} \mathbb{E}\Big(\sup_{0\le s\le t} M_s^2\Big)\, dx \\ &\le 4\int_{\mathcal{T}} \mathbb{E}\Big[\int_0^t \Big(\int_{\mathbb{T}} (f_s^n - f_s^{n-1})\, a_s\, dx\Big)^2 \hat{\rho}_s(X_s)^2\, ds\Big]\, dx \\ &= 4\int_0^t \mathbb{E}\Big[\int_{\mathbb{T}}\Big(\int_{\mathbb{T}} (f_s^n - f_s^{n-1})\, a_s\, dx\Big)^2 \hat{\rho}_s(X_s)^2\, dx\Big]\, ds \\ &\le 4\, ||\partial_x^2\phi||_\infty^2 \int_0^t \mathbb{E}\Big(\int_{\mathbb{T}} |f_s^n - f_s^{n-1}|^2\, dx\Big)\, ds, \end{aligned}$$

due to Cauchy-Schwarz inequality and (5.5.3). In the last term of (5.5.5), with respect to previous ones, there is an extra term:

$$\Big(\int_{\mathbb{T}} \partial_x^2\phi\, \hat{\rho}_s\, dx\Big)$$

which is dominated by $||\partial_x^2\phi|||_\infty$. Finally, there is a constant $C_\phi > 0$ such that

$$\mathbb{E}\Big[\sup_{0\leq s\leq t}\int_{\mathbb{T}} |f_t^{n+1} - f_t^n|^2\, dx\Big] \leq C_\phi \int_0^t \mathbb{E}\Big[\sup_{0\leq \tau\leq s}\int_{\mathbb{T}} |f_\tau^n - f_\tau^{n-1}|^2\, dx\Big]\, ds.$$

Now standard Picard iteration yields the result. □

Proposition 5.5.2. *Let $g_t = f_t(X_t^{-1})$. Then for any $t \in [0,1]$, $\int_{\mathbb{T}} g_t(x)\, dx = 0$.*

Proof. Let $\tilde{K}_t$ be the density of X_t^{-1}. We have $\int_{\mathbb{T}} g_t(x)\, dx = \int_{\mathbb{T}} f_t(x)\, \tilde{K}_t(x)\, dx$. By (5.3.4),

$$\tilde{K}_t = \exp\Big[\int_0^t (\partial_x^2\phi)(X_s) \circ dB_s\Big].$$

Let's see first the martingale part of $d_t \int_{\mathbb{T}} f_t(x)\, \tilde{K}_t(x)\, dx$. Using Itô formula, the martingale part of $f_t\, \tilde{K}_t$ is

$$-\Big(\int_{\mathbb{T}} f_t\, a_t\, dx\Big)\hat{\rho}_t(X_t)\, \tilde{K}_t\, dB_t + f_t\, \partial_x^2\phi(X_t)\, \tilde{K}_t\, dB_t.$$

We have $\int_{\mathbb{T}} \hat{\rho}_t(X_t)\, \tilde{K}_t\, dx = \int_{\mathbb{T}} \hat{\rho}_t(x)\, dx = 1$; on the other hand, by the relation $\tilde{K}_t = \dfrac{1}{\rho_t(X_t)}$, we see that

$$\int_{\mathbb{T}} f_t\, \partial_x^2\phi(X_t)\, \tilde{K}_t\, dx = \int_{\mathbb{T}} f_t\, a_t\, dx.$$

Therefore the martingale part of $d_t \int_{\mathbb{T}} f_t(x)\, \tilde{K}_t(x)\, dx$ is equal to 0. Futhermore we get $d_t \int_{\mathbb{T}} f_t(x)\, \tilde{K}_t(x)\, dx = 0$. It follows that

$$\int_{\mathbb{T}} g_t(x)\, dx = \int_{\mathbb{T}} g_0(x)\, dx = 0.$$

We complete the proof. □

Chapter 6

Diffusive Dean-Kawasaki Equation

Dean-Kawasaki equation is a class of nonlinear SPDEs arising in fluctuating hydrodynamics theory([Kaw98], [Dea96], [Eyi90]). As a prototype, one may consider the following diffusive Dean-Kawasaki equation

$$\partial_t \mu = \alpha \Delta \mu - \nabla.(\sqrt{\mu}\dot{\xi}), \tag{6.0.1}$$

for space-time white noise $\dot{\xi}$ and $\alpha > 0$. In general, we say a continuous measure-valued process $\{\mu_t, t \in [0,T]\}$ is a solution to the diffusive Dean-Kawasaki martingale problem $(MP)^{\alpha}_{\mu_0}$ of (6.0.1) with initial condition μ_0 if there exists a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ such that for all $\phi \in C^2(\mathbb{T}^d)$,

$$M_t(\phi) := \langle \mu_t, \phi \rangle - \langle \mu_0, \phi \rangle - \alpha \int_0^t \langle \mu_s, \Delta\phi \rangle ds$$

is a $\mathcal{F}_t-$ adapted martingale, whose quadratic variation is given by

$$\langle M_t(\phi) \rangle = \int_0^t ||\nabla \phi||^2_{L^2(\mu_s)} ds.$$

The well-posedness of (6.0.1) is challenging. The noise coefficient $\sqrt{\mu}$ causes nonlinearity and possible lack of Lipschitz continuity, also the noise term in the form of a stochastic conservation law causes irregularity. Actually, according to the regularity theory [Hai14], (6.0.1) is a supercritical equation due to the irregularity of space-time white noise. And in [vRLK19], it is proved that a unique measure-valued martingale solution to (6.0.1) exists if and only if $2\alpha \in \mathbb{N}^+$, and in this case, the solution is trivial, i.e. $\mu_t = \frac{1}{N}\sum_{i=1}^N \delta_{W_t^i}$, where $\{W_t^i\}_{i=1,\ldots,N}$ are N independent Brownian particles starting at different sites.

In order to get nontrivial solution, many works give regularization methods in various settings, along with some particle approximations. Sturm, Von Renesse, Konarovskyi and their collaborators ([vRS09, KvR17, KvR15, AvR10]) prove that the Wasserstein diffusion, which can be seen as a infinite dimensional counterpart of Brownian motion in probability measure space equipped with Otto's formal Riemannian metric, is a solution to the Dean-Kawasaki equation with a modified drift term. And they also give several related particle models in case of 1-D Torus. Cornalba, Shardlow and Zimmer ([CSZ19], [CSZ20]) regularize the model from second order Langevin dynamic derivation and get well-posedness for a regularised undamped equivalent of (6.0.1). Other works ([Mar10, FG21]) deal with the case when the noise is spatially regularized. For example, Fehrman and Gess prove a general well-posedness result on a class of Dean-Kawasaki type equations in Stratonovich form of multiplicative noise in [FG21]. Besides, Marx ([Mar18]) gives a particle approximation to a diffusion process on $\mathbb{P}_2(\mathbb{R})$, which has similar properties of Wasserstein diffusion but have better regularity on the measure.

According to the literature we know, there is existence of nontrivial solutions of (6.0.1) only when the spatial correlated intensity is larger than $\frac{3}{2}$. Also, only under such conditions on noise, can a particle approximation model, whose limit measure has a good spatial regularity, be constructed. The main contribution of this chapter is that, inspired by the idea of $Q-$Wiener process on $\mathbb{P}(\mathbb{T})$, we give a new particle approximation to the solution of diffusive Dean-Kawasaki regularised martingale problem $(RMP)^{\alpha,\beta}_{\mathbb{1}_{\mathbb{T}}dx}$ on 1-D Torus in sense of definition 6.2.1, with colored noise $\dot{\xi}^{\beta}_{\mu}$(see (6.2.3)) , whose spatial correlated intensity is larger than 1(see definition 6.1.1), thus proving the existence of solution in this case. We also prove that such solution $\{\mu_t,\, t \in [0,T]\}$, approximated by the interacting particle model, is nonatomic for all $t \in [0,T]$ almost surely. Next, we will introduce the motivation of the particle model's construction.

6.1 From $Q-$Wiener process to the Dean-Kawasaki equation

Generally, let Q be a nonnegative definite symmetric trace-class on a separable Hilbert space K , $\{f_j\}_{j=1}^{\infty}$ be an O.N.B. in K diagonalizing Q , and the corresponding eigenvalues be $\{\lambda_j\}_{j=1}^{\infty}$. Then, in general, we say the following process

$$W_t = \sum_{j=1}^{\infty} \lambda_j f_j W_t^j$$

is a $Q-$Wiener process in K. its derivative with respect t in distributional sense, which denoted as $\dot{W}_t$, are called a colored Guassian noise.

Definition 6.1.1. *We say the spatial correlated intensity of W_t is larger than β if*

$$\sum_{j=1}^{\infty} j^{\beta-1}\lambda_j < \infty.$$

Especially, for $K = L^2(\mathbb{T})$, we realize the 1-D Torus as the interval $[0,1]$ in this paper, and set

$$\begin{aligned}
&e_k = \sqrt{2}\sin(2k\pi x), \quad k = 1, 2...;\\
&e_0 = 1;\\
&e_k = \sqrt{2}\cos(2k\pi x), \quad k = -1, -2, ...
\end{aligned}$$

We denote

$$K_2^\beta = \frac{1}{2} + \sum_{j=1}^{\infty} \frac{1}{j^{2\beta}},$$

where $\beta > 1$ is a constant such that $K_2^\beta < \infty$. Let $\{W^k\}_{k\in\mathbb{N}}$ is a sequence of independent standard Brownian motions on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. Then the usual $Q-$Wiener process on $L^2([0,1])$ with spatial correlated intensity β can be defined as

$$\xi^\beta(t,x) := \sum_{k=-\infty}^{+\infty} \frac{1}{|k|^\beta} e_k(x) W_t^k$$

and it satisfies

$$\mathbb{E}[\xi^\beta(t,x)] = 0; \quad \mathbb{E}[\xi^\beta(t,x)\xi^\beta(s,y)] = t \wedge s \cdot \left(1 + \sum_{k=1}^{+\infty} \frac{2}{|k|^{2\beta}} \cos(2k(x-y))\right).$$

It is obvious that $\xi^\beta(t) \in L^2([0,1]),\ (t,\omega) - a.s.$. The kernel

$$\bar{Q}^\beta(x,y) = 1 + \sum_{k=1}^{+\infty} \frac{2}{|k|^{2\beta}} \cos(2k(x-y))$$

determines the distribution of ξ^β , and of course , its spatial correlated intensity. Generally, for a spatially correlated noise with such kernel, we denote it as $(\bar{Q}^\beta)^{\frac{1}{2}}-$Wiener process.

$Q-$Wiener process can be naturally seen as a infinite dimensional counterpart of Bownian motion in K. On the other hand, it is known (see [vRS09], [AvR10], [Wan21]) that the solution of (6.0.1) or its regularised form can be seen as a Wasserstein diffusion. To introduce the motivation of the particle model in section 6.3 , we start from the viewpoint of $Q-$Wiener process on Wasserstein space. Firstly, we will briefly show the connection between $Q-$Wiener process on Wasserstein space and the solution to the diffusive Dean-Kawasaki equation.

In [DFL21], they construct a $Q-$Wiener process extrinsiclly on Wasserstein space on general connected compact Riemannian manifold M. When it applies to the case $M = \mathbb{T}$, we can choose the orthonormal system as the standard Fourier base $\{e_k\}_{k\in\mathbb{N}}$ on $[0,1]$, then

$$dX_t^Q = \sum_{k=-\infty}^{\infty} a_k e_k(X_t^Q) dW_t^k \tag{6.1.1}$$

induce a stochastic C^1-diffeomorphic flow when $a_k = \dfrac{1}{|k|^4}$. Suppose that $\mu_0 = \mathbb{1}_{[0,1]}$, let $\mu_t^Q = (X_t^Q)_{\#}\mu_0$, and denote

$$C = \sum_{k=0}^{\infty} \frac{1}{|k|^8}.$$

By applying Itô formula on $\langle f, \mu_t^Q\rangle$ for $f \in C^2(M)$, we get

$$d\langle f, \mu_t^Q\rangle = \sum_{k=-\infty}^{\infty} \frac{1}{|k|^4} \langle f', e_k\rangle_{\mu_t^Q} dW_t^k + C\langle f'', \mu_t^Q\rangle dt. \tag{6.1.2}$$

Rewrite (6.1.2) in form of SPDE on μ_t^Q:

$$\partial_t \mu^Q = C\partial_x^2 \mu^Q - \partial_x(\mu^Q \dot{\xi}^\beta) \tag{6.1.3}$$

for $\beta = 4$. We see that the drift term coincides with the drift term in the diffusive Dean-Kawasaki equation. Following this idea, we want to construct a solution as a image measure process induced by a process X_t, which is in form of (6.1.1). In fact, from the point of fluid dynamic, if we see the diffusive Dean-Kawasaki equation as a Eulerian discription of some stochastically moving fluid, then, our construction can be seen as a corresponding Lagrangian's discription.

However, μ_t^Q will never be a candidate for the solution of martingale problem associated with the diffusive Dean-Kawasaki equation because their quadratic variation process are not consistent. In fact, if we assume $a_k = 1$ for all $k \in \mathbb{N}$ in (6.1.1), and formally write the flow equation as

$$dX_t' = \sum_{k=-\infty}^{\infty} e_k(X_t') dW_t^k.$$

We denote $\mu_t' = (X_t')_{\#}\mu_0$, and formally compute the quadratic variation of the martingale part of $\langle f, \mu_t'\rangle$ without consideration of regularity of the flow, we find that $d < \langle f, \mu_t'\rangle >= \sum_{i=1}^{\infty} \langle f', e_k\rangle^2_{\mu_t'} dt$, while for the solution μ_t of $(MP)^C_{\mu_0}$, $d < \langle f, \mu_t\rangle >= ||f'||^2_{\mathcal{L}^2(\mu_t)} dt$. This is not surprising because if one wants to construct a Brownian motion on a manifold, the 'velocity' should be stochastically parallel translated along the path, while in (6.1.1), the vector fields $\{e_k,\, k \in \mathbb{N}\}$ are just fixed. Here, as an experimental attempt, let

$$dX_t = \sum_{k=-\infty}^{\infty} e_k(X_0) dW_t^k, \tag{6.1.4}$$

then formally we have

$$\begin{aligned} d < \langle f, \mu_t \rangle > &= \sum_{i=1}^{\infty} \langle f', e_i(X_t^{-1}) \rangle_{\mu_t}^2 dt \\ &= \sum_{i=1}^{\infty} \langle f'(X_t), e_i \rangle_{\mu_0}^2 dt \\ &= ||f'(X_t)||_{\mathcal{L}^2(\mu_0)}^2 dt = ||f'||_{\mathcal{L}^2(\mu_t)}^2 dt \end{aligned}$$

Although the computation above is not strict, we still get a direct insight: we can construct a solution to diffusive Dean-Kawasaki equation on Torus by constructing a image process induced by a diffeomorphic, or at least one-to-one continuous map flow X_t satisfying

$$dX_t = \sum_{i=-\infty}^{\infty} a_i e_i(t, X_t) dW_t^i,$$

where $e_i(t, x)$ is a stochastically moving frame in form of $e_i(X_t^{-1}(x))$. We will construct a new particle approximation in section 3 by following this idea.

We briefly introduce the main contents of this chapter. In section 6.2 , we give the definition of the noise term ξ_μ^β and regularised martingale problem $(RMP)_{\mu_0}^{\alpha,\beta}$ for initial measure $\mu_0 = \mathbb{1}_{\mathcal{T}} dx$, and show its consistency with usual martingale problem to (6.0.1). In section 6.3, we will construct a particle model. Theorem 6.3.1 shows the well-posedness of this discrete model for any $\beta > 1$. In section 6.4, we will prove that, as the particle number goes to infinity, the distribution induced by the empirical measure process in $\mathcal{C}([0, T], \mathbb{P}(\mathbb{T}))$ is tight so that we can pick a weakly convergent limit process. We will also prove that any weakly convergent limiting process $\{p_t,\ t \in [0, T]\}$ is a solution to $(RMP)_{\mathbb{1}_{\mathbb{T}} dx}^{K_2^\beta,\beta}$. Thus we can prove the existence of solution to $(RMP)_{\mathbb{1}_{\mathbb{T}} dx}^{K_2^\beta,\beta}$ (see theorem 6.4.1). As a necessary step in the proof, we find that p_t is non-atomic for all $t \in [0, T]$ almost surely (see lemma 6.4.2).

6.2 Introduction of the regularised martingale problem and the noise

we firstly give the definition of regularised martingale problem $(RMP)_{\mu_0}^{\alpha,\beta}$ for $\mu_0(dx) = \mathbb{1}_{\mathbb{T}} dx$:

Definition 6.2.1. *We say a continuous $\mathbb{P}([0,1])$-valued process $\{\mu_t, t \in [0,T]\}$ is a solution to the regularised martingale problem $(RMP)^{\alpha,\beta}_{\mu_0}$, if there exists a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ such that for all $\phi \in C^2([0,1])$,*

$$M_t(\phi) := \langle \mu_t, \phi \rangle - \langle \mu_0, \phi \rangle - \alpha \int_0^t \langle \mu_s, \phi'' \rangle ds$$

is a $\mathcal{F}_t-$ adapted martingale, whose quadratic variation process is given by

$$< M_t(\phi) >= \int_0^t Q^\beta_{\mu_s}(\phi,\phi) ds.$$

The quadratic form $Q^\beta_{\mu_s}(\phi,\phi)$ is defined as

$$Q^\beta_{\mu_s}(\phi,\phi) := \int_{[0,1]} \int_{[0,1]} \phi'(x)\phi'(y)\Big(1 + \sum_{k=1}^{\infty} \frac{2}{k^{2\beta}} \cos(2\pi k(F_{\mu_s}(x) - F_{\mu_s}(y)))\Big)\mu_s(dx)\mu_s(dy),$$

where F_{μ_s} is the distribution function of μ_s, satisfying $F_{\mu_s}(0) = 0,\ F_{\mu_s}(1) = 1$ and

$$F_{\mu_s}(x) = \int_0^x \mathbb{1}_{(0,x]}(y)\mu_s(dy), \quad \text{for } 0 \le x \le 1.$$

In particular, we denote such regularised martingale problem, with initial condition $d\mu_0 = \mathbb{1}_{\mathcal{T}} dx$, as $(RMP)^{\alpha,\beta}_{\mathbb{1}_{\mathbb{T}}dx}$.

Note that, due to $dx = (F)_\# d\mu$, we have

$$Q^\beta_{\mu_s}(\phi,\phi) = \int_{[0,1]} \int_{[0,1]} \phi'(G_{\mu_s}(x))\phi'(G_{\mu_s}(y))\Big(1 + \sum_{k=1}^{\infty} \frac{2}{k^{2\beta}} \cos(2\pi k(x-y))\Big) dxdy,$$

where G_{μ_s} is the quantile function of μ_s . Because $|\sum_{k=1}^{\infty} \frac{2}{k^{2\beta}} \cos(2\pi k(x-y))| < 2K^\beta_2$, we get

$$\begin{aligned}
Q^\beta_{\mu_s}(\phi,\phi) &= \int_{[0,1]} \int_{[0,1]} \phi'(G_{\mu_s}(x))\phi'(G_{\mu_s}(y))\Big(\sum_{k=-\infty}^{\infty} \frac{1}{|k|^{2\beta}} e_k(x)e_k(y)\Big) dxdy \\
&= \sum_{k=-\infty}^{+\infty} \frac{1}{|k|^{2\beta}} \int_{[0,1]} \phi'(G_{\mu_s}(x))e_k(x)dx \int_{[0,1]} \phi'(G_{\mu_s}(y))e_k(y)dy \qquad (6.2.1) \\
&= \sum_{k=-\infty}^{+\infty} \frac{1}{k^{2\beta}} |\widehat{\phi'(G_{\mu_s})}_k|^2,
\end{aligned}$$

where the fourier coefficient is defined as

$$\widehat{f}_k = 2\int_0^1 f(x)\sin(2\pi kx)dx, \quad k = 1, 2, \dots;$$
$$\widehat{f}_0 = \int_0^1 f(x)dx;$$
$$\widehat{f}_k = 2\int_0^1 f(x)\cos(2\pi kx)dx, \quad k = -1, -2, \dots.$$

Remark 6.2.2. *In fact,* (6.2.1) *shows that the spatial correlated intensity of our noise is β , which we will only require $\beta > 1$ in existence theorem 6.4.1. Especially, when $\beta = 0$, the quadratic variation above becomes*

$$< M_t(\phi) >= \int_0^t ||\phi'(G_{\mu_s}(x))||^2_{L^2[0,1]}ds = \int_0^t ||\phi'||^2_{L^2(\mu_s)}ds. \tag{6.2.2}$$

Although this is just a formal computation, since we can not prove the existence of μ_s a priori, it still shows that our definition of regularised martingale problem is consistent with the definition of general martingale problem(see [vRLK19]) .

Next, we introduce the colored noise $\dot{\xi}^\beta_\mu$. Note that, given $\phi \in C^2([0,1])$, the kernel

$$\bar{Q}^\beta_\mu(x, y) := 1 + \sum_{k=1}^{\infty} \frac{2}{k^{2\beta}} \cos(2\pi k(F_\mu(x) - F_\mu(y)))$$

determines the martingale $M_t(\phi)$ in distribution. Although $L^2(\mu)$ may not be separable, we still can define a $(\bar{Q}^\beta_\mu)^{\frac{1}{2}}-$Wiener process in the tangent space $L^2(\mu)$ with orthonormal eigenfunctions $\{e_k(\mu)\}_{k\in\mathbb{N}}$ in $L^2(\mu)$, which are defined as

$$e_k(\mu, x) = e_k(F_\mu(x)), \qquad k \in \mathbb{N}.$$

This is because $dx = (F)_{\#}d\mu$,

$$\begin{aligned}\int_{[0,1]} \bar{Q}^\beta_\mu(x, y)e_k(F(y))\mu(dy) &= \int_{[0,1]} \Big(\sum_{i=-\infty}^{\infty} \frac{1}{|i|^{2\beta}} e_i(F(x))e_i(F(y))\Big) e_k(F(y))\mu(dy) \\ &= \int_{[0,1]} \Big(\sum_{i=-\infty}^{\infty} \frac{1}{|i|^{2\beta}} e_i(F(x))e_i(y)\Big) e_k(y)dy = \frac{1}{|i|^{2\beta}} e_k(F(x)).\end{aligned}$$

Therefore , for general $\mu \in \mathbb{P}(\mathbb{T})$, we still can define a generalized $(\bar{Q}^\beta_\mu)^{\frac{1}{2}}-$Wiener process in $L^2(\mu)$ as

$$\xi_\mu^\beta(t,x) = \sum_{k=-\infty}^{+\infty} \frac{1}{|k|^\beta} e_k(F_\mu) W_t^k \tag{6.2.3}$$

where $\{e_k(\mu,\cdot)\}_{k\in\mathbb{N}}$ is a family of orthonormal vectors in $L^2(\mu)$ .And ξ_μ^β satisfies :

$$\mathbb{E}[\xi_\mu^\beta(t,x)] = 0; \quad \mathbb{E}[\xi_\mu^\beta(t,x)\xi_\mu^\beta(s,y)] = (t\wedge s)\cdot\left(1+\sum_{k=1}^{\infty}\frac{2}{k^{2\beta}}\cos\Big(2\pi k(F_\mu(x)-F_\mu(y))\Big)\right),$$

We denote its time derivative, in distribution, as $\dot{\xi}_\mu^\beta$. Still, it can be proved by Doob's inequality, that $\dot{\xi}_\mu^\beta(t,\cdot) \in L^2(\mu_t) - a.s.$.

Remark 6.2.3. *If the solution μ_t to $(RMP)_{\mu_0}^{\alpha,\beta}$ is absolutely continuous with respect to Lebesgue measure, i.e. $d\mu_t = \rho_t dx$, then it is easy to see that $\{\rho_t,\, t\in[0,T]\}$ is a martingale solution to the following SPDE*

$$\partial_t\rho = \alpha\partial_x^2\rho - \partial_x(\rho\dot{\xi}_\rho^\beta).$$

Comparing with the original form of (6.0.1)*, we actually change the bad term $\sqrt{\mu}$ into μ by transferring nonlinearity to the noise. Luckily, in case of 1-D Torus, the noise ξ_μ^β has the form of* (6.2.3) *so that we can analyse it.*

6.3 Construction of the particle model on $\mathbb{T}$

Following the idea introduced in section 6.1 and the definition of $L^2(\mu)-$Wiener process ξ_μ^β, $e_k(\mu,x)$ is the stochastically moving frame, and we want to construct a solution to $(RMP)_{\mu_0}^\alpha$ as a image measure process $\mu_t = (X_t)_\#\mu_0$, induced by the process X_t satisfying

$$dX_t = \sum_{k=-\infty}^{+\infty}\frac{1}{|k|^\beta}e_k(X_0)dW_t^k.$$

The main difficulty is we can not guarantee X_t is a diffeomorphism, or even a one-to-one C^α map, when β is only larger than 1. Although in this paper we will not analyse X_t directly since we only need to construct the particle approximation of X_t, the similar difficulty still appears in the construction of the particle model. In detail, given N particles $\{X_N^i(t)\}_{i=1,\dots,N}$, if we use a direct idea for the construction of a particle approximation to X_t, we usually want $X_N^i(t)$ to satisfy

$$dX_N^i(t) = \sum_{k=-N}^{+N}\frac{1}{|k|^\beta}e_k(X_N^i(0))dW_t^k, \quad i=1,\dots,N.$$

However, we can not guarantee that $\{X_N^i(t)\}_{i=1,...,N}$ do not collide for $t \in [0,T]$, i.e. $\exists\, i,j$ and $T > t > 0$ such that $X_N^i(t) = X_N^{i+1}(t)$. This collision phenomenon shows the problem of concentration of mass, which is one of the main obstacle to avoid triviality of the solution to the martingale problem of (6.0.1). Inspired by mean-field background ([LLX20], [RS93]), We will construct a interacting particle model without collision by adding a replusive interaction between $\{X_N^i(t)\}$, and make sure that the interaction term is so small that its influence can be neglected when the empirical measure of $\{X_N(t)\}$ weakly converges to a solution to $(RMP)^{\alpha,\beta}_{\mathbb{1}_{\mathbb{T}}dx}$. In this section, we will construct the particle model.

For each $N > 0$, we define the following process

$$dX_N^i(t) = \frac{1}{2N^{\alpha+1}} \sum_{j=1, j\neq i}^{N} \cot\left(\pi(X_N^i(t) - X_N^j(t))\right)dt + \sum_{k=-N}^{N} \frac{1}{|k|^\beta} e_k(x^i) dW_t^k, \tag{6.3.1}$$

where the initial value is $X_0^i = x^i$. Note that in this case , the diffusion coefficient is fixed since $e_k(x^i)$ is independent of $\{X_N^i(t)\}_{i=1,...,N}$. α is some positive constant which will be chosen later.
Define $\Delta_N = \{(x_i)_{1\leq i\leq N} \in \mathbb{R}^N : x_1 < x_2 < ... < x_N, \text{ and } |x_1 - x_N| < 1\}$ and $X_N(t) = (X_N^i(t))_{1\leq i \leq N}$. We denote

$$K_1^N = \sum_{j=1}^{N} \frac{4\pi^2}{j^{2\beta-2}}; \quad K_2^N = \frac{1}{2} + \sum_{j=1}^{N} \frac{1}{j^{2\beta}}.$$

where $\beta > 1$ is a constant such that $K_2^N < \infty$. It is obvious that $K_1^N \leq O(N^{3-2\beta})$ for $1 < \beta < \frac{3}{2}$,$K_1^N \leq O(\log N)$ for $\beta = \frac{3}{2}$ and $K_1^N \leq C$ for $\beta > \frac{3}{2}$.

Theorem 6.3.1. *For any $\beta > 1$ and initial condition $X_N^i(0) = \frac{i}{N}$, we choose $0 < \alpha < (2\beta - 2) \wedge 1$. Then there exists a unique strong solution $(X_N(t))_{t\in[0,T]}$, which takes value in Δ_N , to SDE* (6.3.1) *when N is large enough.*

Proof. We follow the method stated in [RS93] and [LLX20]. We firstly construct the truncated process. Let $\phi_R(x)$ be a $C^2(\mathbb{R})$ function which satisfies $\phi_R(x) = \cot(\pi x)$ for $x \in (-1 + \frac{1}{R}, -\frac{1}{R}) \cup (\frac{1}{R}, 1 - \frac{1}{R})$. Then the following SDE

$$dX_{R,N}^i(t) = \frac{1}{2N^{\alpha+1}} \sum_{j=1, j\neq i}^{N} \phi_R(X_{R,N}^i(t) - X_{R,N}^j(t))dt + \sum_{k=N}^{N} \frac{1}{|k|^\beta} e_k(\frac{i}{N}) dW_t^k,$$

with initial value $X_{R,N}^i(0) = \frac{i}{N}$ for $1 \leq i \leq N$, has a unique strong solution $X_{R,N}(t)$. Let

$$\tau_R := \inf\left\{t : \min_{l\neq j} |e^{2\pi i X_{R,N}^l(t)} - e^{2\pi i X_{R,N}^j(t)}| \leq R^{-1}\right\}.$$

Then τ_R is monotone increasing in R and $X_{R,N}(t) = X_{R',N}(t)$ for all $t \leq \tau_R$ and $R < R'$.
Let $X_N(t) = X_{R,N}(t)$ on $t \in [0, \tau_R)$. Then we need to prove: $(X_N(t))_{t\in[0,T]}$ does not explode, never collide and $|X_N(t) - X_1(t)| < 1$. For abbreviation of notation, we denote $X_N^i(t)$ as X_t^i without confusion .

Firstly, we prove non-explosion. Let $R_t^N := \frac{1}{2N}\sum_{i=1}^N (X_t^i)^2$, then by Ito formula,

$$dR_t^N = (\frac{N-1}{4N^{1+\alpha}} + K_2^N)dt + \frac{1}{N}\sum_{i=1}^N X_t^i \left(\sum_{k=-N}^N \frac{1}{|k|^\beta} e_k(\frac{i}{N})dW_t^k\right)$$

Computing the quadratic variation process of R_t^N, we find

$$\begin{aligned}
\frac{d}{dt} < R_t^N > &= \frac{1}{N^2}\sum_{k=-N}^N \frac{1}{|k|^{2\beta}}\left(\sum_{i=1}^N X_t^i e_k(\frac{i}{N})\right)^2 \\
&= \frac{1}{N^2}\sum_{k=-N}^N \frac{1}{|k|^{2\beta}}\left(\sum_{i=1}^N (X_t^i)^2 e_k^2(\frac{i}{N}) + \sum_{j=1}^N\sum_{i=1,i\neq j}^N X_t^i X_t^j e_k(\frac{i}{N})e_k(\frac{j}{N})\right)^2 \\
&= \frac{2}{N^2}K_2^N\sum_{i=1}^N (X_t^i)^2 + \frac{1}{N^2}\sum_{j=1}^N\sum_{i=1,i\neq j}^N X_t^i X_t^j \sum_{k=-N}^N \frac{1}{|k|^{2\beta}} e_k(\frac{i}{N})e_k(\frac{j}{N}).
\end{aligned}$$

Note that

$$|\sum_{k=-N}^N \frac{1}{|k|^{2\beta}} e_k(\frac{i}{N})e_k(\frac{j}{N})| = |1 + \sum_{k=1}^N \frac{2}{k^{2\beta}}\cos(\frac{i-j}{N}2k\pi)| < 2K_2^\beta,$$

thus,

$$\frac{d}{dt}\langle R_t^N\rangle < (\frac{C_1}{N} + C_2)R_t^N.$$

Then, by B.D.G. inequality, we have

$$\begin{aligned}
\mathbb{E}[|\max_{s\in[0,t]} R_s^N|^2] &\leq Ct^2 + \mathbb{E}\left[< \int_0^t \frac{1}{N}\sum_{i=1}^N X_s^i \left(\sum_{k=-N}^N \frac{1}{|k|^\beta} e_k(\frac{i}{N})dW_s^k\right) >\right] \\
&\leq Ct^2 + C\int_0^t \mathbb{E}[R_s^N]ds \\
&\leq Ct^2 + C\int_0^t \mathbb{E}[\max_{q\in[0,s]} R_q^N]ds.
\end{aligned} \tag{6.3.2}$$

On the other hand, by Cauchy inequality, we have

$$\mathbb{E}[|\max_{s\in[0,t]} R_s^N|^2] \geq \left(\mathbb{E}[\max_{s\in[0,t]} R_s^N]\right)^2. \tag{6.3.3}$$

Denote $r(t) := \left(\mathbb{E}[\max_{s\in[0,t]} R_s^N]\right)^2$. By (6.3.2) and (6.3.3) , we finally get

$$r(t) \leq Ct^2 + C\int_0^t \sqrt{r(s)}ds.$$

According to Gronwall type inequality and monotonicity of $r(t)$, we prove non-explosion of r . It follows that , if we set $\zeta = \lim_{K\to\infty}\zeta_K$, where

$$\zeta_K \ := \inf\{t \geq 0 : |X_t^j| \geq K, \text{for some } j = 1, \ldots, N\}, \tag{6.3.4}$$

then $\zeta \geq T$ for any T . R_t^N will not explode in finite time almost surely. Thus the process $\{e^{2\pi i X_N^j(t)}\}_{j=1,\ldots,N}$ is well defined on $[0,T]$.

We secondly prove the non-collision. Consider the Lyapunov function $F(x_1, ...x_N) = -\frac{1}{N^2}\sum_{l\neq j}\log|e^{2\pi i x_l} - e^{2\pi i x_j}|$, by Itô formula,

$$\begin{aligned}
&d_t F(X_t^1, ...X_t^N)\\
&= -\frac{1}{2N^2}\sum_{l=1}^N\sum_{j=1,j\neq l}^N \cot(\pi(X_t^l - X_t^j))dX_t^l + \frac{1}{2N^2}\left(\sum_{l=1}^N\sum_{j=1,j\neq l}^N \frac{\pi}{\sin^2(\pi(X_t^l - X_t^j))}d_t\langle X_t^l\rangle\right)\\
&- \frac{1}{2N^2}\left(\sum_{l=1}^N\sum_{j=1,j\neq l}^N \frac{\pi}{\sin^2(\pi(X_t^l - X_t^j))}d_t\langle X_t^l, X_t^j\rangle\right)
\end{aligned}$$

Note that for the above three terms (denoted as $\mathbf{A}$,$\mathbf{B}$ and $\mathbf{C}$) , we have

$$\begin{aligned}
\mathbf{A} &= M_N(t) - \frac{1}{4N^{3+\alpha}}\sum_{i=1}^N\sum_{j=1,j\neq i}^N \cot^2(\pi(X_t^l - X_t^j))dt\\
\mathbf{B} &= \frac{1}{2N^2}\sum_{l=1}^N\sum_{j=1,j\neq l}^N\left(\frac{\pi}{\sin^2(\pi(X_t^l - X_t^j))}\sum_{k=-N}^N \frac{1}{|k|^{2\beta}}|e_k(\frac{l}{N})|^2\right)dt\\
\mathbf{C} &= -\frac{1}{2N^2}\sum_{l=1}^N\sum_{j=1,j\neq l}^N\left(\frac{\pi}{\sin^2(\pi(X_t^l - X_t^j))}\sum_{k=-N}^N \frac{1}{|k|^{2\beta}}e_k(\frac{l}{N})e_k(\frac{j}{N})\right)dt,
\end{aligned}$$

where $M_N(t)$ is a local martingale. Thus,

$$d_tF(X_t^1,...X_t^N)$$
$$= \frac{1}{2N^2}\left(-\frac{1}{2N^{1+\alpha}}\sum_{1\le l<j\le N}\cot^2(\pi(X_t^l-X_t^j))+\pi\sum_{1\le l<j\le N}\sum_{k=-N}^{N}\frac{1}{|k|^{2\beta}}|\frac{e_k(\frac{l}{N})-e_k(\frac{j}{N})}{\sin(\pi(X_t^l-X_t^j))}|^2\right)dt \quad (6.3.5)$$
$$+\, d_tM_N(t).$$

]Next, we are going to estimate $\sum_{1\le l<j\le N}\sum_{k=-N}^{N}\frac{1}{|k|^{2\beta}}|\frac{e_k(\frac{l}{N})-e_k(\frac{j}{N})}{\sin(\pi(X_t^l-X_t^j))}|^2$. We divide it into three parts:

$$(A)=\sum_{M=1}^{M_1-1}\sum_{i=1}^{N-M}\sum_{k=-N}^{N}\frac{1}{|k|^{2\beta}}\left|\frac{e_k(\frac{i}{N})-e_k(\frac{i+M}{N})}{\sin(\pi(X_t^i-X_t^{i+M}))}\right|^2$$
$$(B)=\sum_{M=M_1}^{M_2-1}\sum_{i=1}^{N-M}\sum_{k=-N}^{N}\frac{1}{|k|^{2\beta}}\left|\frac{e_k(\frac{i}{N})-e_k(\frac{i+M}{N})}{\sin(\pi(X_t^i-X_t^{i+M}))}\right|^2$$
$$(C)=\sum_{M=M_2}^{N-1}\sum_{i=1}^{N-M}\sum_{k=-N}^{N}\frac{1}{|k|^{2\beta}}\left|\frac{e_k(\frac{i}{N})-e_k(\frac{i+M}{N})}{\sin(\pi(X_t^i-X_t^{i+M}))}\right|^2.$$

We denote

$$a_m=\sum_{i=1}^{N-m}\left|\frac{1}{\sin(\pi(X_t^i-X_t^{i+m}))}\right|^2;$$
$$b_m=\sum_{i=1}^{N-m}\left|\frac{1}{\pi(X_t^i-X_t^{i+m})}\right|^2;$$
$$c_m=\sum_{i=1}^{N-m}\left|\frac{1}{\tan(\pi(X_t^i-X_t^{i+m}))}\right|^2;$$
$$Q_N=\frac{1}{2N^{1+\alpha}}\sum_{1\le l<j\le N}\frac{1}{|\tan(\pi(X_t^i-X_t^{i+m}))|^2}.$$

For (A) , Note that

$$|e_k(\frac{i}{N})-e_k(\frac{i+M}{N})|\le\frac{1}{|k|^{-1}}\frac{2\sqrt{2}\pi M}{N}. \quad (6.3.6)$$

Thus,

$$
\begin{aligned}
(A) \leq & \sum_{M=1}^{M_1-1} \sum_{i=1}^{N-M} \sum_{k=-N}^{N} \frac{1}{|k|^{2\beta-2}} \frac{8\pi^2 M^2}{N^2} \cdot \frac{1}{\sin^2(\pi(X_t^i - X_t^{i+M}))} \\
\leq & \frac{CK_1^N}{N^2} \sum_{M=1}^{M_1-1} M^2 a_M \\
< & \frac{M_1^2}{N^{2-\epsilon}} \sum_{M=1}^{M_1-1} a_M .
\end{aligned}
$$

where $\epsilon := (3-2\beta) \vee 0$. We pick $\alpha' > \alpha + \epsilon$ and choose M_1 such that

$$
M_1^2 \leq N^{1-\alpha'}, \tag{6.3.7}
$$

then we have

$$
(A) \leq \frac{1}{N^{1+\alpha'-\epsilon}} \sum_{M=1}^{M_1-1} a_M < \frac{1}{6N^{1+\alpha}} \sum_{1 \leq l < j \leq N} \frac{1}{\sin^2(\pi(X_t^l - X_t^j))} = \frac{1}{3} Q_N + \frac{N-1}{12N^\alpha}.
$$

When M is large, (6.3.6) is not enough to estimate (B) and (C) . Note that

$$
\frac{1}{\sin^2 x} = \frac{1}{\tan^2 x} + 1 \leq \frac{1}{x^2} + 1 \leq \frac{1}{\sin^2 x} + 1 = \frac{1}{\tan^2 x} + 2. \tag{6.3.8}
$$

Because of convexity of the function $\frac{1}{x^2}$, we have, for each $1 \leq k \leq [\frac{M}{2}]$,

$$
\begin{aligned}
& \frac{1}{|X_t^i - X_t^{i+M}|^2} \\
= & \frac{1}{|\sum_{l=0}^{M-k}(X_t^{i+l} - X_t^{i+l+k}) + \sum_{n=1}^{k-1}(-X_t^{i+n} + X_t^{i+M-k+n})|^2} \\
\leq & \frac{1}{M^3} \left(\sum_{l=0}^{M-k} \frac{1}{|X_t^{i+l} - X_t^{i+l+k}|^2} + \sum_{n=1}^{k-1} \frac{1}{|X_t^{n+i} - X_t^{i+n+M-k}|^2} \right)
\end{aligned} \tag{6.3.9}
$$

Thus,

$$
\begin{aligned}
b_M \leq & \sum_{i=1}^{N-M} \frac{1}{\pi^2[\frac{M}{2}]} \sum_{k=1}^{[\frac{M}{2}]} \frac{1}{M^3} \left(\sum_{l=0}^{M-k} \frac{1}{|X_t^{i+l} - X_t^{i+l+k}|^2} + \sum_{n=1}^{k-1} \frac{1}{|X_t^{n+i} - X_t^{i+n+M-k}|^2} \right) \\
\leq & \frac{2}{\pi^2 M^4} \sum_{k=1}^{[\frac{M}{2}]} \left(\sum_{l=0}^{M-k} \sum_{i=1}^{N-M} \frac{1}{|X_t^{i+l} - X_t^{i+l+k}|^2} + \sum_{n=1}^{k-1} \sum_{i=1}^{N-M} \frac{1}{|X_t^{n+i} - X_t^{i+n+M-k}|^2} \right) \\
\leq & \frac{M}{M^4} \sum_{k=1}^{[\frac{M}{2}]} (b_k + b_{M-k}) .
\end{aligned}
\tag{6.3.10}
$$

We denote $S_n = \sum_{i=1}^{n} b_i$, then

$$
b_M < \frac{1}{M^3} S_{M-1}. \tag{6.3.11}
$$

Therefore, we see that, for $m > n$,

$$
\frac{S_m}{S_n} < \prod_{j=n}^{m-1} (1 + \frac{1}{j^3}). \tag{6.3.12}
$$

And when N goes to infinity and n is large enough ,

$$
\log \frac{S_\infty}{S_n} < \sum_{j=n}^{\infty} \frac{1}{j^3} \leq \frac{1}{n^2}. \tag{6.3.13}
$$

For (C) , by (6.3.8), we find that

$$
\begin{aligned}
(C) < & \sum_{M=M_2}^{N-1} \sum_{i=1}^{N-M} \frac{K_2^\beta}{|\sin(\pi(X_t^i - X_t^{i+M}))|^2} \\
\leq & K_2^\beta \sum_{M=M_2}^{N-1} (b_M + N - M) \\
= & K_2^\beta (S_{N-1} - S_{M_2}) + K_2^\beta \sum_{M=M_2}^{N-1} (N - M).
\end{aligned}
$$

Combined with (6.3.12), (6.3.13) and choose M_2 such that

$$
N^2 \gg M_2^2 \geq N^{1+\eta} > N^{1+\alpha}, \tag{6.3.14}
$$

for some constant $\eta > \alpha$, then, using (6.3.8), we get

$$
\begin{aligned}
(C) &\leq \frac{1}{N^{1+\eta}} S_{M_2} + K_2^{\beta} \sum_{M=M_2}^{N-1} (N-M) \\
&< \frac{1}{3} Q_N + \sum_{M=1}^{N-1} (N-M) < \frac{1}{3} Q_N + K_2^{\beta} N^2.
\end{aligned}
\tag{6.3.15}
$$

Based on the estimates above , we deal with the part (B) .

$$
\begin{aligned}
(B) &\leq \sum_{M=M_1}^{M_2-1} \sum_{i=1}^{N-M} \frac{8\pi^2 M^2}{N^2 \sin^2(\pi(X_t^i - X_t^{i+M}))} \sum_{k=-N}^{N} \frac{1}{|k|^{2\beta-2}} \\
&\leq C K_1^N \sum_{M=M_1}^{M_2} \frac{M^2}{N^2} (b_M + N - M).
\end{aligned}
$$

Based on (6.3.7) and (6.3.14) and taking use of (6.3.11), we have, for N is large enough,

$$
\begin{aligned}
(B) &< \frac{C K_1^N}{N^2} \sum_{M=M_1}^{M_2} \frac{1}{M} S_{M_2} + \frac{K_1^N}{N^2} \sum_{M_1}^{M_2} (N-M) M^2 \\
&< N^{\epsilon} S_{M_2} \frac{\log M_2 - \log M_1}{N^2} + \frac{K_1^N}{3N} M_2^3.
\end{aligned}
$$

It follows that

$$
(B) < S_{M_2} \frac{\log N}{N^{2-\epsilon}} + N^{\frac{1}{2}+\epsilon} < \frac{1}{3} Q_N + \frac{\sum_{M=1}^{M_2} (N-M)}{N^{2-\epsilon}} + N^{\frac{1}{2}+\epsilon}.
$$

We conclude that when N is large enough,

$$
(A) + (B) + (C) < Q_N + K_2^{\beta} N^2.
$$

Therefore, $F(X^1_{t\wedge\tau_R}, \ldots X^N_{t\wedge\tau_R}) - K_2^{\beta} t \wedge \tau_R$ is a super-martingale. Since the diffusion process $\{e^{2\pi i X_N^j(t\wedge\tau_R)}\}_{j=1,\ldots,N}$ on the torus is well defined alomstly surely, then, following the standard argument(see [RS93]) , we denote

$$
S = \{\tau_R \leq T\},
$$

then

$$
\begin{aligned}
F(X_0) + K_2^\beta \tau_R \wedge T &\geq \mathbb{E}[F(X_{\tau_R \wedge T}]\\
&= \mathbb{E}[F(X_{\tau_R}\mathbb{1}_S] + \mathbb{E}[F(X_T \mathbb{1}_{S^c}]\\
&\geq -\frac{1}{N^2}\log(\frac{1}{R})\mathbb{P}(S) - \frac{1}{2N^2}(N^2 - N - 2)\log 2 \cdot \mathbb{P}(S)\\
&- \frac{1}{2N^2}(N^2 - N)\log 2 \cdot \mathbb{P}(S^c)\\
&= \frac{1}{N^2}(\log(2N) + \log 2)\mathbb{P}(S) - \frac{N-1}{2N}\log 2.
\end{aligned}
$$

Therefore,

$$
\mathbb{P}(\tau_R \leq T) \leq \frac{N^2(F(X_0) + K_2^\beta T + \log 2)}{\log R + \log 2}
$$

For fixed T, Letting $R \to \infty$, it follows that $\{(e^{2\pi i X_N^j(t\wedge T)})\}_{j=1,\ldots,N}$ never collide. Then letting $T \to \infty$, since $\mathbb{P}(\tau_\infty \leq T) = 0$ always holds, so there is no collision of the particles $\{e^{2\pi i X_N^j(t)}\}_{j=1,\ldots,N}$ in torus for all $t \in [0, +\infty)$. Furthermore, coming back to the original process , this means $\{X_N^j(t)\}_{j=1,\ldots,N}$ never collides and $|X_N^1(t) - X_N^N(t)| < 1$.
Finally, by continuity of the trajectories of $X_N(t)$, we have $X_N(t) \in \Delta_N$ for all $t \geq 0$. We finished the proof . □

Remark 6.3.2. *We give a short comparison between the common noise and the stochastically moving noise above . Generally, if we apply the same computation on the Lyapunov function for the case of the common noise, i.e.* $\int_0^t \sum_{k=-N}^{N} e_k(X_s^l)dW_s^k$ *, the last term in* (6.3.5) *becomes*

$$
\sum_{1\leq l<j\leq N}\sum_{k=-N}^{N}\frac{1}{|k|^{2\beta}}\left|\frac{e_k(X_t^l) - e_k(X_t^j)}{X_t^l - X_t^j}\right|^2 . \tag{6.3.16}
$$

We can bound it by $N^2K_1^N$ *. It is obvious that* β *should be larger than* $\frac{3}{2}$ *in order to get non-collision of particles in the case of common noise. However, if we use the stochastically moving noise, we can prove the non-collision of particles for each* $\beta > 1$ *.*

6.4 Construction of a solution to $(RMP)_{\mathbb{1}_\mathbb{T}dx}^{K_2^\beta,\beta}$

In this section, we will construct a solution to $(RMP)_{\mathbb{1}_\mathbb{T}dx}^{K_2^\beta,\beta}$ as a weakly convergent subsequence limit of the empirical measure process of the interacting particle model introduced in section 3.
Let the integer function $[\cdot] : \mathbb{R} \to \mathbb{N}$ be defined as

$$\begin{cases} [x] = x-1, & x \in \mathbb{N}; \\ [x] = \max\{n \in \mathbb{N} \,|\, n < x\}, & \text{otherwise}. \end{cases}$$

And $\{x\} := x - [x]$. Then , we define the empirical measure on $[0,1]$:

$$L_N(t) := \frac{1}{N}\sum_{i=1}^{N} \delta_{\{X_N^i(t)\}}.$$

The distribution function F_t^N of $L_N(t)$, defined on $[0,1]$, satisfies $F_t^N(0) = 0$ and

$$F_t^N(x) = \int_0^x L_N(t, dy).$$

We also denote the corresponding quantile function $G_t^N(x)$, which satisfies

$$G_t^N(F_t^N(x)) = x, \quad -a.s.$$

Theorem 6.4.1. *Under the assumption in Theorem 6.3.1,* $\{L_N(t),\ t \in [0,T]\}$ *is tight in* $\mathcal{C}([0,T], \mathbb{P}([0,1]))$ *, and the limit of any weakly convergent subsequence of* $\{L_N(t),\ t \in [0,T]\}$ *is a solution to* $(RMP)_{\mathbb{1}_{\mathbb{T}}dx}^{K_2^\beta}$ *.*

Proof. Denote P_N as the distribution of $\{L_N(t),\ t \in [0,T]\}$ in $C([0,T]; \mathbb{P}(\mathbb{T}))$, and P_N^ϕ as the distribution of $\{< L_N(t), \phi >,\ t \in [0,T]\}$ in $C([0,T]; \mathbb{R})$ for $\phi \in C^\infty(\mathbb{T})$. Then, by [Daw93] (Theorem 3.7.1), P_N is tight if and only if $P_{\mathbb{I}}^\phi$ is tight for each $\phi \in C^\infty(\mathbb{T})$. Here, for $\phi \in C^\infty(\mathbb{T})$, we actually means $\phi \in C^\infty([0,1])$ so that we can extend it as a period function on $\mathbb{R}$. For sake of convenience, we still denote the extended function as ϕ . Note that, by Theorem 6.3.1, there is no collision and no explosion for the particles $(X_N^i(t))$ for all $t \in [0,T]$. Therefore, we can apply Itô formula to get, $\forall\, \phi \in C^\infty(\mathbb{T})$,

$$\begin{aligned} \langle L_N(t), \phi\rangle &= \langle L_N(0), \phi\rangle + \frac{1}{N}\sum_{i=1}^{N}\int_0^t \phi'(X_s^i)\cdot dX_s^i + \frac{1}{2N}\sum_{i=1}^{N}\int_0^t \sum_{k=-N}^{+N}\frac{1}{|k|^{2\beta}}\phi''(X_s^i)e_k^2(\frac{i}{N})ds \\ &= \langle L_N(0), \phi\rangle + \frac{1}{2N^{2+\alpha}}\sum_{i=1}^{N}\int_0^t \phi'(X_s^i)\cdot\left(\sum_{j=1, j\neq i}^{N}\cot\left(\pi(X_N^i(t) - X_N^j(t))\right)\right) ds \\ &+ \frac{1}{2N}\sum_{i=1}^{N}\int_0^t \sum_{k=-N}^{N}\frac{1}{|k|^{2\beta}}\phi''(X_s^i)e_k^2(\frac{i}{N})ds + M_N^\phi(t) \\ &= (\mathbb{K}) + (\mathbb{I}) + (\mathbb{J}) + M_N^\phi(t), \end{aligned} \tag{6.4.1}$$

where

$$M_N^{\phi}(t) = \frac{1}{N}\sum_{i=1}^{N}\int_0^t \sum_{k=-N}^{N} \frac{1}{|k|^{\beta}}\phi'(X_s^i)e_k(\frac{i}{N})dW_s^k$$

Note that

$$\begin{aligned}(\mathbb{I}) &= \frac{1}{4N^{\alpha}}\int_0^t \sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} \frac{1}{N^2}\frac{\phi'(X_N^i(s)) - \phi'(X_N^j(s))}{\tan(\pi(X_N^i(s) - X_N^j(s)))}ds \\ &\leq \frac{1}{4\pi N^{\alpha}}\int_0^t\int_{[0,1]^2} \frac{\phi'(x)-\phi'(y)}{x-y}L_N(s,dx)L_N(s,dy)ds = O(N^{-\alpha}).\end{aligned} \tag{6.4.2}$$

Here, the inequality above is because ϕ is a function on torus, we can choose a shorter interval between $X_N^i(s)$ and $X_N^j(s)$ such that we can make sure $X_N^i(s) - X_N^j(s) \in (0, \frac{1}{2}]$ or $[-\frac{1}{2}, 0)$, then, applying the mean value theorem, we have

$$\left|\frac{\phi'(X_N^i(s)) - \phi'(X_N^j(s))}{\tan(\pi(X_N^i(s) - X_N^j(s)))}\right| \leq ||\phi''||_{\infty}\left|\frac{X_N^i(s)) - X_N^j(s)}{\tan(\pi(X_N^i(s) - X_N^j(s)))}\right| \leq \frac{||\phi''||_{\infty}}{\pi}.$$

On the other hand, we have

$$(\mathbb{J}) = \frac{K_2^N}{N}\int_0^t \phi''(X_s^i)\mathbf{m}s = K_2^N\int_0^t\int_{[0,1]}\phi''(x)L_N(s,dx)ds.$$

For the martingale part, by Cauchy inequality and boundness of $|\phi'|$, we have

$$\begin{aligned}< M_N^{\phi}(t) > &= \frac{1}{N^2}\int_0^t \sum_{k=-N}^{N}\frac{1}{|k|^{2\beta}}\big(\sum_{i=1}^{N}\phi'(X_s^i)e_k(\frac{i}{N})\big)^2 ds \\ &\leq \frac{1}{N}\int_0^t\sum_{i=1}^{N}(\phi'(X_s^i))^2\Big(\sum_{k=-N}^{N}\frac{1}{|k|^{2\beta}}e_k^2(\frac{i}{N})\Big)ds \\ &\leq C\int_0^t \langle(\phi')^2, L_N(s)\rangle ds \leq Ct.\end{aligned}$$

Therefore, by B.D.G inequality,

$$\mathbb{E}[|\langle L_N(t),\phi\rangle - \langle L_N(s),\phi\rangle|^{2m}] \leq O(N^{-\alpha})|t-s|^m + C'|t-s|^m.$$

Also, $(\mathbb{K}) \to \int_{\mathbb{T}} \phi dx$. According to ([KS12] p.63 Theorem 4.10) , we have proved tightness of $\{P_N^\phi\}$. Thus , P_N is tight. Due to separability, we can apply Prohorov theorem, so we proved the relative compactness of the distribution $\mathbb{P}^N$ on $C([0,T],\mathbb{P}(\mathbb{T}))$. Therefore , we have a subsequence, still denoted as P_N for convenience, weakly converges to some P in $C([0,T],\mathbb{P}(\mathbb{T}))$. By Skorohod representation theorem, we can find a new probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ and a sequence of random variable $\{p^n\},\ p$ defined on it , which takes value in $C([0,T],\mathbb{P}(\mathcal{T}))$ and satisf ies $Law(p^n)=P_N,\, Law(p)=P$, such that p^n converges to p weakly almost surely.

Next, we will show that the limiting process $\{p_t(\omega,x), t\in[0,T]\}$, associated with $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}})$, is a solution to $(RMP)^{K_2^\beta}_{\mathbb{1}_{\mathbb{T}}dx}$. Note that for a solution μ_t to $(RMP)^{K_2^\beta}_{\mathbb{1}_{\mathbb{T}}dx}$, the generator $\mathbb{L}$ associated with $\langle\mu_t,\phi\rangle$ is

$$\mathbb{L}f = K_2^\beta\langle\mu_t,\phi\rangle f' + \frac{1}{2}Q_{\mu_t}(\phi)f'', \qquad \forall\, f\in\mathcal{C}^2(\mathbb{R}).$$

Thus, according to the equivalent description of $\mathbb{P}(\mathbb{T})-$valued process, see [Daw93] lemma 7.2.1, we only need to prove that, for $\forall\, G\in\mathcal{D}:=\{G:\ G(\mu)=g(\langle\mu,\phi\rangle),\ \phi\in C^2(\mathbb{T}),\ g\in C^2(\mathbb{R})\}$,

$$M_t^G(p) := G(p_t) - G(p_0) - \int_0^t \mathrm{D}G(p_s)ds,$$

where $\mathrm{D}G(\mu) = K_2^\beta g'(\langle\phi,\mu\rangle)\langle\mu,\phi''\rangle + \frac{1}{2}g''(\langle\mu,\phi\rangle)Q_\mu(\phi)$, is a $\tilde{\mathbb{P}}-$local martingale. This suffices to prove that , for every $s<t\in[0,T]$, and any continuous function $H: C([0,T];\mathbb{P}(\mathbb{T}))\to\mathbb{R}$,

$$\tilde{\mathbb{E}}\big[\big(G(p_t)-G(p_s)-\int_s^t \mathrm{D}G(p_r)dr\big)\cdot H(p\big|_{[0,s]})\big]=0. \tag{6.4.3}$$

In fact, when $k>0$, $e_k(x)=-e_k(-x)$, thus

$$Q_\mu(\phi)=\int_{[0,1]}\int_{[0,1]}\phi'(x)\phi'(y)\Big(1+\sum_{k=1}^{\infty}\frac{2}{k^{2\beta}}\cos\big(2\pi k(\int_{[0,1]}\mathbb{1}_{(x\wedge y,x\vee y]}(z)\mu(dz))\big)\Big)\mu(dx)\mu(dy).$$

Since we can prove that, for $\tilde{\mathbb{P}}-$a.s. $p_t(\omega)$ is non-atomic for all $t\in[0,T]$ (lemma 6.4.2 below), thus , for $\tilde{\mathbb{P}}-$almost surely,

$$\int_{[0,1]}\mathbb{1}_{(a,b]}dp_t^n \to \int_{[0,1]}\mathbb{1}_{(a,b]}dp_t, \quad \forall\, t\in[0,T].$$

It follows that $M_t^G(p^n)H(p^n\big|_{[0,s]})$ converges to $M_t^G(p)H(p\big|_{[0,s]})$ almost surely. Note that

$$\tilde{\mathbb{E}}\big[\big|M_t^G(p^n)-M_s^G(p^n)\big|\cdot\big|H(p\big|_{[0,s]})\big|\big]<\infty,$$

then, by dominated convergence theorem,

$$
\begin{aligned}
&\tilde{\mathbb{E}}\big[\big(G(p_t) - G(p_s) - \int_s^t \mathrm{D}G(p_r)dr\big) \cdot H(p\big|_{[0,s]})\big] \\
&= \lim_{n\to\infty} \tilde{\mathbb{E}}\big[\big(G(p_t^n) - G(p_s^n) - \int_s^t \mathrm{D}G(p_r^n)dr\big) \cdot H(p^n\big|_{[0,s]})\big].
\end{aligned}
$$

Also, we define

$$
Q_\mu^n(\phi) = \int_{[0,1]}\int_{[0,1]} \phi'(x)\phi'(y)\Big(1 + \sum_{k=1}^n \frac{2}{k^{2\beta}} \cos\big(2\pi k(\int_{[0,1]} \mathbb{1}_{(x\wedge y, x\vee y]}(z)\mu(dz))\big)\Big)\mu(dx)\mu(dy).
$$

Since

$$
\left|\sum_{k=n}^{\infty} \frac{2}{k^{2\beta}} \cos\big(2\pi k(\int_{[0,1]} \mathbb{1}_{(x\wedge y, x\vee y]}(z)\mu(dz))\big)\right| < \sum_{k=n}^{+\infty} \frac{2}{k^{2\beta}} \to 0, \text{ as } n\to\infty,
$$

and

$$
(\sum_{k=n}^{+\infty} + \sum_{k=-\infty}^{-n})\frac{2}{k^{2\beta}} \to 0, \quad \text{as } n\to\infty,
$$

thus, by denoting $\mathrm{D}^n G(\mu) = K_2^n g'(\langle\phi,\mu\rangle)\langle\mu,\phi''\rangle + g''(\langle\mu,\phi\rangle)Q_\mu^n(\phi)$, we have

$$
\begin{aligned}
&\tilde{\mathbb{E}}\big[\big(G(p_t) - G(p_s) - \int_s^t \mathrm{D}G(p_r)dr\big) \cdot H(p\big|_{[0,s]})\big] \\
&= \lim_{n\to\infty} \tilde{\mathbb{E}}\big[\big(G(p_t^n) - G(p_s^n) - \int_s^t \mathrm{D}^n G(p_r^n)dr\big) \cdot H(p^n\big|_{[0,s]})\big].
\end{aligned}
$$

Because $Law(p^n) = P_n$,

$$
\begin{aligned}
&\tilde{\mathbb{E}}\big[\big(G(p_t^n) - G(p_s^n) - \int_s^t \mathrm{D}^n G(p_r^n)dr\big) \cdot H(p^n\big|_{[0,s]})\big] \\
&= \mathbb{E}\big[\big(G(L_n(t)) - G(L_n(s)) - \int_s^t \mathrm{D}^n G(L_n(r))dr\big) \cdot H(L_n\big|_{[0,s]})\big].
\end{aligned}
$$

Therefore, to prove (6.4.3) , we only need to prove

$$
\lim_{n\to\infty} \mathbb{E}\big[G(L_n(t)) - G(L_n(0)) - \int_0^t \mathrm{D}^n G(L_n(s))ds\big] = 0 \tag{6.4.4}
$$

In fact,

$$G(L_n(t)) = g(\frac{1}{n}\sum_{i=1}^{n}\phi(X_n^i(t))).$$

Then, by Itô formula ,

$$\begin{aligned}
&dG(L_n(t)) \\
&= g'(\langle \phi, L_n(t)\rangle)\frac{1}{2n^{2+\alpha}}\sum_{i=1}^{n}\phi'(X_n^i(t))\cdot\left(\sum_{j=1,j\neq i}^{n}\cot\left(\pi(X_n^i(t)-X_n^j(t))\right)\right)dt \\
&+ g'(\langle \phi, L_n(t)\rangle)\frac{1}{2n}\sum_{i=1}^{n}\phi''(X_n^i(t))\sum_{k=-n}^{n}\frac{1}{|k|^{2\beta}}e_k^2(\frac{i}{n})dt \\
&+ g''(\langle \phi, L_n(t)\rangle)\left(\frac{1}{n^2}\sum_{i,j=1}^{n}\phi'(X_n^i(t))\phi'(X_n^j(t))\sum_{k=0}^{n}\frac{2}{k^{2\beta}}\cos(2\pi k\frac{i-j}{n})\right)dt + dM_n^{g,\phi}(t) \\
&= (I)dt + (J)dt + (K)dt + dM_n^{g,\phi}(t),
\end{aligned}$$

where $M_n^{g,\phi}(t)$ is a $\mathbb{P}-$local martingale .
Note that

$$(J) = K_2^n g'(\langle \phi, L_n(t)\rangle)\langle L_n(t), \phi''\rangle, \tag{6.4.5}$$

and

$$\begin{aligned}
(K) &= g''(\langle L_n(t), \phi\rangle)\int_{[0,1]}\int_{[0,1]}\phi'(x)\phi'(y)\sum_{k=0}^{n}\frac{2}{k^{2\beta}}\cos(2\pi k(F_t^n(x)-F_t^n(y)))L_n(t,dx)L_n(t,dy) \\
&= g''(\langle L_n(t), \phi\rangle)Q^n_{L_n(t)}(\phi).
\end{aligned} \tag{6.4.6}$$

For the last part (I) , we have

$$|(I)| \le \frac{C}{n^\alpha}||g'||_\infty\cdot\frac{n(n-1)}{n^2}||\phi''||_\infty \to 0, \quad \text{as}\, n\to\infty. \tag{6.4.7}$$

So, combining (6.4.5), (6.4.6) and (6.4.7) , we proved (6.4.4) . We finished the proof.

□

Lemma 6.4.2. *For $\tilde{\mathbb{P}}-$a.s., p_t is non-atomic for all $t\in[0,T]$.*

Proof. Let $U = \{\omega:\, \exists t, \text{ such that } p_t(\omega)\text{is atomic}\}$. If the measurable set U has positive measure, i.e. $\tilde{\mathbb{P}}(U) = C > 0$. We define

$$U_i = \{\omega \in U \ : \exists t, x, \text{ such that } p_t(\omega, dx) = \eta \delta_x \text{ with } \eta > \frac{1}{i}\},$$

then it is obvious that $U_i \subset U_{i+1}$ and $\bigcup_{i=1}^{\infty} U_i = U$. Thus , we can find some U_k such that $\tilde{\mathbb{P}}(U_k) > \frac{C}{2}$.

We define $E_x^N = (x - \frac{1}{2N}, x + \frac{1}{2N})$ .Note that, for $\tilde{\mathbb{P}}-$a.s. ω , $p_t^n(\omega, \cdot)$ weakly converges to $p_t(\omega, \cdot)$ uniformly in $t \in [0, T]$. If $p_t(\omega, dx) = \eta \delta_x dx$, then, for each N, there exists $n(N, \omega, t, x,)$ such that $\forall\, n \geq n(N, \omega, t, x)$,

$$\int_{E_x^N} p_t^n(\omega, dy) > \frac{\eta}{2}.$$

Based on this observation, we define

$$U_k^{n,N} = \{\omega : \exists\, (t, x), \text{ such that for} \,\forall\, j \geq n, \int_{E_x^N} p_t^j(\omega, dy) > \frac{1}{2k}\},$$

then we must have $U_k^{n,N} \subset U_k^{n+1,N}$ and $U_k \subset \bigcup_{n=1}^{\infty} U_k^{n,N}$. Therefore, for each N, we can find m_N such that $\tilde{\mathbb{P}}(U_k^{m_N,N}) > \frac{C}{3} = C'$. Now, let

$$\bar{U}_k^{m_N,N} = \{\omega : \exists\, (t, x), \text{ such that for} \,\forall\, j \geq m_N, \int_{E_x^N} L_j(\omega, t, y) dy > \frac{1}{2k}\}.$$

Remember that L_n has the same distribution with p^n . We must have

$$\mathbb{P}(\bar{U}_k^{m_N,N}) = \tilde{\mathbb{P}}(U_k^{m_N,N}) = C' > 0.$$

On the other hand, we define a stopping time

$$\tau_{m_N}^k := \inf\{t : \min_j |e^{2\pi i X_{m_N}^j(t)} - e^{2\pi i X_{m_N}^{j+\frac{m_N}{2k}}(t)}| \leq \frac{1}{2N}\}.$$

We have proved $F(X_{m_N}(t)) + K_2^{\beta} t$ is a super-martingale. Denote

$$A = \{\tau_{m_N}^k \leq T\}.$$

Then we have

$$
\begin{aligned}
&F(X_{m_N}(0)) + K_2^\beta \tau^k_{m_N} \wedge T \\
&\geq \mathbb{E}[F(X_{m_N}(\tau^k_{m_N} \wedge T)] \\
&= \mathbb{E}[F(X_{m_N}(\tau^k_{m_N})\mathbb{1}_A] + \mathbb{E}[F(X_{m_N}(T)\mathbb{1}_{A^c}] \\
&\geq -\frac{1}{2m_N^2}\frac{m_N}{2k}(\frac{m_N}{2k} - 1)\log(\frac{1}{2N})\mathbb{P}(A) \\
&- \frac{1}{2m_N^2}(m_N^2 - m_N - \frac{m_N}{2k}(\frac{m_N}{2k} - 1))\log 2\mathbb{P}(A) \\
&- \frac{1}{2m_N^2}(m_N^2 - m_N)\log 2\mathbb{P}(A^c) \\
&= \frac{1}{2m_N^2}\frac{m_N}{2k}(\frac{m_N}{2k} - 1)(\log(2N) + \log 2)\mathbb{P}(A) - \frac{m_N - 1}{2m_N}\log 2.
\end{aligned}
$$

It follows that

$$
\mathbb{P}(\tau^k_{m_N} \leq T) \leq \frac{F(X_{m_N}(0)) + K_2^\beta T + \log 2}{\frac{\frac{m_N}{2k} - 1}{4km_N}(\log(2N) + \log 2)} \leq \frac{10k^2(C + 2F(X_{m_N}(0)))}{\log(2N) + \log 2}. \tag{6.4.8}
$$

Note that

$$
-\int_{[0,1]\times[0,1]} \log|x - y| dxdy = \frac{3}{2},
$$

thus , when N is large enough,

$$
F(X_{m_N}(0)) = -\frac{1}{N^2}\sum_{i \neq j}\log(\frac{|i - j|}{N}) < 2.
$$

Now, by (6.4.8), we can choose N so that $\mathbb{P}(\tau^k_{m_N} \leq T) < \frac{1}{2}C'$. However, for each $\omega \in \bar{U}^{m_N,N}_k$, there must be at least $\frac{m_N}{2k}$ particles included in some interval $(x - \frac{1}{2N}, x + \frac{1}{2N})$, which means $\tau^k_{m_N}(\omega) \leq T$. Therefore , $\mathbb{P}(\tau^k_{m_N} \leq T) \geq \mathbb{P}(\bar{U}^{m_N,N}_k) = C'$. Contradiction! We finished the proof.

□

Remark 6.4.3. *Actually, we can extend the initial measure into any absolutely continuous measure on $\mathbb{T}$ because the only difference is that when we construct the particle model, we need to set the initial distribution as*

$$
\mu_0^N(x) = \int_{\frac{i-1}{N}}^{\frac{i}{N}} \mu_0(dx) \cdot \mathbb{1}_{(\frac{i-1}{N}, \frac{i}{N}]}(x)
$$

and the O.N.B of $L^2(\mu_0)$, $\{\bar{e}_k\}_{k\in\mathbb{N}}$, should be

$$\bar{e}_k = e_k \circ F_{\mu_0}.$$

The proof is the similar.

Bibliography

[AG08] Luigi Ambrosio and Nicola Gigli. Construction of the parallel transport in the Wasserstein space. *Methods and Applications of Analysis*, 15(1):1–30, 2008.

[AGS05] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2005.

[AKR96] Sergio Albeverio, Yuri G. Kondratiev, and Michael Röckner. Differential geometry of Poisson spaces. 1996.

[AvR10] Sebastian Andres and Max-Konstantin von Renesse. Particle approximation of the Wasserstein diffusion. *Journal of Functional Analysis*, 258(11):3879–3905, 2010.

[BB00] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[BÉ85] Dominique Bakry and Michel Émery. Diffusions hypercontractives. In *Seminaire de Probabilités XIX 1983/84*, pages 177–206. Springer, 1985.

[Bis81] Jean-Michel Bismut. Mécanique aléatoire. lect. *Notes. Maths*, 966:19881, 1981.

[BLPR17] Rainer Buckdahn, Juan Li, Shige Peng, and Catherine Rainer. Mean-field stochastic differential equations and associated pdes. *The Annals of Probability*, 45(2):824–878, 2017.

[Bre91] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on Pure and Applied Mathematics*, 44(4):375–417, 1991.

[Cru83] Ana Bela Cruzeiro. Équations différentielles sur l'espace de Wiener et formules de Cameron-Martin non-linéaires. *Journal of Functional Analysis*, 54(2):206–227, 1983.

[CSZ19] Federico Cornalba, Tony Shardlow, and Johannes Zimmer. A regularized Dean–Kawasaki model: derivation and analysis. *SIAM Journal on Mathematical Analysis*, 51(2):1137–1187, 2019.

[CSZ20] Federico Cornalba, Tony Shardlow, and Johannes Zimmer. From weakly interacting particles to a regularised Dean-Kawasaki model. *Nonlinearity*, 33(2):864, 2020.

[Daw93] Donald Dawson. Measure-valued Markov processes. In *École d'été de probabilités de Saint-Flour XXI-1991*, pages 1–260. Springer, 1993.

[Dea96] David Dean. Langevin equation for the density of a system of interacting Langevin processes. *Journal of Physics A: Mathematical and General*, 29(24):L613, 1996.

[DF21] Hao Ding and Shizan Fang. Geometry on the Wasserstein space over a compact Riemannian manifold. *Acta Mathematica Scientia*, 41(6):1959–1984, 2021.

[DFL21] Hao Ding, shizan Fang, and Xiangdong Li. Towards stochastic analysis in the Wasserstein space. *preprint*, 2021.

[dLHLT20] Aythami Bethencourt de Leon, Darryl D Holm, Erwin Luesink, and So Takao. Implications of Kunita-Itô-Wentzell formula for k-forms in stochastic fluid dynamics. *Journal of Nonlinear Science*, 30(4):1421–1454, 2020.

[Elw92] Kenneth David Elworthy. Stochastic flows on Riemannian manifolds. In *Diffusion processes and related problems in analysis, volume II*, pages 37–72. Springer, 1992.

[Eyi90] Gregory Lawrence Eyink. Dissipation and large thermodynamic fluctuations. *Journal of Statistical Physics*, 61(3):533–572, 1990.

[FG21] Benjamin Fehrman and Benjamin Gess. Well-posedness of the Dean-Kawasaki and the nonlinear Dawson-Watanabe equation with correlated noise. *arXiv preprint arXiv:2108.08858*, 2021.

[FLT10] Shizan Fang, Dejun Luo, and Anton Thalmaier. Stochastic differential equations with coefficients in Sobolev spaces. *Journal of Functional Analysis*, 259(5):1129–1168, 2010.

[FS11] Shizan Fang and Jinghai Shao. Fokker–Planck equation with respect to heat measures on loop groups. *Bulletin des Sciences Mathematiques*, 135(6-7):775–794, 2011.

[Gig11] Nicola Gigli. On the inverse implication of Brenier-McCann theorems and the structure of (P 2 (M), W 2). *Methods and Applications of Analysis*, 18(2):127–158, 2011.

[Hai14] Martin Hairer. A theory of regularity structures. *Inventiones Mathematicae*, 198(2):269–504, 2014.

[IW81] Nobuyuki Ikeda and Shinzo Watanabe. Stochastic differential equations and diffusion processes, 1981.

[JKO98] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

[KA82] Leonid Vitaliyevich Kantorovich and Gleb Pavlovich Akilov. Functional Analysis (seconde édition). *Pergamon Press, London*, 326:429, 1982.

[Kaw98] Kyozi Kawasaki. Microscopic analyses of the dynamical density functional equation of dense fluids. *Journal of Statistical Physics*, 93(3):527–546, 1998.

[KM97] Andreas Kriegl and Peter W. Michor. *The convenient setting of global analysis*, volume 53. American Mathematical Soc., 1997.

[Kon11] Vitalii Konarovskii. On infinite system of diffusing particles with coalescing. *Theory of Probability & Its Applications*, 55(1):134–144, 2011.

[Kon17] Vitalii Konarovskyi. A system of coalescing heavy diffusion particles on the real line. *The Annals of Probability*, 45(5):3293–3335, 2017.

[KS12] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.

[Kun97] Hiroshi Kunita. *Stochastic flows and stochastic differential equations*, volume 24. Cambridge university press, 1997.

[KvR15] Vitalii Konarovskyi and Max von Renesse. Modified Arratia flow and Wasserstein diffusion. *Communications on Pure and Applied Mathematics*, 2015.

[KvR17] Vitalii Konarovskyi and Max von Renesse. Reversible coalescing-fragmentating Wasserstein dynamics on the real line. *arXiv preprint arXiv:1709.02839*, 2017.

[Li09] Xiang-Dong Li. On the strong lp-Hodge decomposition over complete Riemannian manifolds. *Journal of Functional Analysis*, 257(11):3617–3646, 2009.

[LL06a] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. i–le cas stationnaire. *Comptes Rendus Mathématique*, 343(9):619–625, 2006.

[LL06b] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. ii–Horizon fini et contrôle optimal. *Comptes Rendus Mathématique*, 343(10):679–684, 2006.

[LL07] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

[LL16] Songzi Li and Xiang-Dong Li. W-entropy formulas and Langevin deformation of flows on Wasserstein space over Riemannian manifolds. *arXiv preprint arXiv:1604.02596*, 2016.

[LL18] Songzi Li and Xiang-Dong Li. W-entropy formulas on super Ricci flows and Langevin deformation on Wasserstein space over Riemannian manifolds. *Science China Mathematics*, 61(8):1385–1406, 2018.

[LLX20] Songzi Li, Xiang-Dong Li, and Yong-Xiao Xie. On the law of large numbers for the empirical measure process of generalized Dyson Brownian motion. *Journal of Statistical Physics*, 181(4):1277–1305, 2020.

[Lot06] John Lott. Some Geometric Calculations on Wasserstein space. *Communications in Mathematical Physics*, 277:423–437, 2006.

[Lot17] John Lott. An intrinsic parallel transport in Wasserstein space. *Proceedings of the American Mathematical Society*, 145(12):5329–5340, 2017.

[LV09] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Annals of Mathematics*, pages 903–991, 2009.

[LW20] Wuchen Li and Yifei Wang. Information Newton's flow: second-order optimization method in probability space. *arXiv preprint arXiv:2001.04341*, 2020.

[LWZ21] Wei Liu, Liming Wu, and Chaoen Zhang. Long-time behaviors of mean-field interacting particle systems related to McKean–Vlasov equations. *Communications in Mathematical Physics*, 387(1):179–214, 2021.

[Mal97] Paul Malliavin. *Stochastic analysis*, volume 313 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, 1997.

[Mar10] Mauro Mariani. Large deviations principles for stochastic scalar conservation laws. *Probability Theory and Related Fields*, 147(3):607–648, 2010.

[Mar18] Victor Marx. A new approach for the construction of a Wasserstein diffusion. *Electronic Journal of Probability*, 23:1–54, 2018.

[McC97] Robert John McCann. A convexity principle for interacting gases. *Advances in Mathematics*, 128(1):153–179, 1997.

[McC01] Robert John McCann. Polar factorization of maps on Riemannian manifolds. *Geometric & Functional Analysis GAFA*, 11(3):589–608, 2001.

[Ott01] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. 2001.

[OV00] Felix Otto and Cédric Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *Journal of Functional Analysis*, 173(2):361–400, 2000.

[RS93] Leonard Christopher Gordon Rogers and Zhan Shi. Interacting Brownian particles and the Wigner law. *Probability Theory and Related Fields*, 95(4):555–570, 1993.

[RW21] Panpan Ren and Feng-Yu Wang. Derivative formulas in measure on Riemannian manifolds. *Bulletin of the London Mathematical Society*, 2021.

[Stu06] Karl-Theodor Sturm. On the geometry of metric measure spaces. *Acta Mathematica*, 196(1):65–131, 2006.

[Stu14] Karl-Theodor Sturm. A monotone approximation to the Wasserstein diffusion. In *Singular Phenomena and Scaling in Mathematical Models*, pages 25–48. Springer, 2014.

[Vil03] Cédric Villani. Graduate studies in mathematics: Vol. 58. Topics in optimal transportation, 2003.

[Vil09] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.

[vRLK19] Vitalii von Renesse, Tobias Lehmann, and Max-Konstantin Konarovskyi. Dean-Kawasaki dynamics: ill-posedness vs. triviality. *Electronic Communications in Probability*, 24:1–9, 2019.

[vRS05] Max-K von Renesse and Karl-Theodor Sturm. Transport inequalities, gradient estimates, entropy and Ricci curvature. *Communications on Pure and Applied Mathematics*, 58(7):923–940, 2005.

[vRS09] Max-Konstantin von Renesse and Karl-Theodor Sturm. Entropic measure and Wasserstein diffusion. *The Annals of Probability*, 37(3):1114–1191, 2009.

[Wan21] Feng-Yu Wang. Image-dependent conditional McKean–Vlasov SDEs for measure-valued diffusion processes. *Journal of Evolution Equations*, 21(2):2009–2045, 2021.