Wasserstein space over the Wiener space

Shizan FANG[a,b] Jinghai SHAO[b,c] Karl-Theodor STURM[c]

a: I.M.B, BP 47870, Université de Bourgogne, 21078 Dijon, France
b: School of Mathematics, Beijing Normal University, Beijing 100875, China
c: Institut für Angewandte Mathematik, Universität Bonn, DE-53115 Bonn, Germany

Abstract

The goal of this paper is to study optimal transportation problems and gradient flows of probability measures on the Wiener space, based on and extending fundamental results of Feyel-Üstünel. Carrying out the program of Ambrosio-Gigli-Savaré, we present a complete characterization of the derivative processes for certain class of absolutely continuous curves. We prove existence of the gradient flow curves for the relative entropy w.r.t. the Wiener measure and identify these gradient flow curves with solutions of the Ornstein-Uhlenbeck evolution equation.

Introduction

Let (X, H, μ) be an abstract Wiener space. Consider on X the d_H distance defined as

$$d_H(x,y) = \begin{cases} |x-y|_H & x-y \in H, \\ +\infty & \text{otherwise.} \end{cases} \tag{0.1}$$

It is well-known that $(x,y) \mapsto d_H(x,y)$ is lower semi-continuous over $X \times X$. Denote by $\mathcal{P}(X)$ the space of probability measures on X. For $\nu_1, \nu_2 \in \mathcal{P}(X)$, we define the following Wasserstein distance W_2:

$$W_2(\nu_1,\nu_2) = \inf\Big\{ \int_{X\times X} |x-y|_H^2 \, \pi(dx,dy); \quad \pi \in \mathcal{C}(\nu_1,\nu_2) \Big\}^{1/2} \tag{0.2}$$

where $\mathcal{C}(\nu_1,\nu_2)$ denotes the totality of probability measures on $X\times X$, having ν_1 and ν_2 as marginal laws. The distance $W_2(\nu_1,\nu_2)$ could take the value $+\infty$. Note that it would be more appropriate to attribute the distance W_2 to Kantorovich and Rubinstein, but we keep the name "Wasserstein" (referring to Vasershtein's contribution [Va]) since this terminology is now quite standard.

During recent years, due to the success of constructing Monge optimal transport maps on the Wiener space [FU], there are intensive researches on the transformations of measures on the Wiener space (see [FUZ], [BK], [BKM]). The purpose of this paper is to study the geometrical aspect of the Wasserstein space $(\mathcal{P}(X), W_2)$. Our work is based essentially on the following ones:

1) the lecture note [AS] given by L. Ambrosio and G. Savaré, in which the authors introduced rigorously the tangent spaces of the Wasserstein space $(\mathcal{P}_2(\mathbf{R}^d), W_2)$, where $\mathcal{P}_2(\mathbf{R}^d)$ denotes the space of probability measures with finite second moment, and the structure of gradient flows is systematically studied.

2) the fundamental work [FU] by D. Feyel and A.S. Üstünel about the Monge-Kantorovich optimal transportation problem on the Wiener space.

To emphasize the difference between these two situations, we outline the following two points:

1) the compactness of the closed ball $\{x \in \mathbf{R}^d;\ |x|_{\mathbf{R}^d} \leq R\}$ allows to prove the tightness of a family of probability measures in $\mathcal{P}_2(\mathbf{R}^d)$; while on the Wiener space (X, H, μ), neither $\{x \in X;\ ||x||_X \leq R\}$ (non compact) nor $\{x \in X;\ |x|_H \leq R\}$ (of measure μ zero) does work.

2) for a sequence of probability measures (μ_n) on $\mathbf{R}^d$, converging weakly to μ, there exists a sequence of random variables (Z_n) of law μ_n and Z of law μ such that

$$|Z_n - Z|_{\mathbf{R}^d} \to 0 \quad \text{a.s.},$$

then (see [Ch, chapter 5]) under the uniform integrability of second moment, the weak convergence μ_n to μ implies the convergence

$$W_2(\mu_n, \mu) \to 0 \quad \text{as} \quad n \to +\infty;$$

while on the Wiener space, the convergence with respect to the norm of X does not imply the convergence with respect to the d_H distance, the counterpart does not hold in this latter situation.

Now we describe the content of this work. In a geometric context, the connection between the convexity of the entropy functional (relative to the Riemannian volume or to a reference) and the lower bound of the Ricci curvature has been developed in [LV] and [St]. In section 1, we will clarify this connection in the framework of Wiener space, see Theorem 1.5. Tangent spaces to Wasserstein spaces have been firstly considered at a formal level in [Ot] and rigorously implemented in [AGS]. In Section 2, we will introduce the derivative processes associated to absolutely continuous curves, so that the distance W_2 is expressed as a Riemannian distance, a new interpretation for the Benamou-Brenier's formula, see Theorem 2.6. The gradient flow associated to a general convex functional is defined usually through sub-gradients. For the entropy functional, in section 3, we compute explicitly the directional derivative and prove that the gradient of the entropy functional exists at the minimizers in Jordan-Kinderlehrer-Otto's approximation scheme [JKO]. We will prove that solutions to the Ornstein-Uhlenbeck evolution is the gradient flow associated to the entropy functional, see Theorem 3.10.

1. 1-convexity of the entropy functional

Let (X, H, μ) be an abstract Wiener space, that is, X is a separable Banach space, H is a separable Hilbert space which is densely and continuously embedded in X such that

$$\int_X e^{\sqrt{-1}\ell(x)}\, d\mu(x) = e^{-|i^*(\ell)|_H^2/2} \quad \text{for} \quad \ell \in X^*(\text{dual of } X)$$

where $i : H \to X$ is the injection map and $i^* : X^* \to H$ the dual map. For simplicity, we consider

$$X^* \subset H \subset X.$$

In what follows, we denote by $||\cdot||$ the norm of X and $\mathrm{Ent}(f) = \int_X f \log f\, d\mu$ for any positive measurable function on X such that $\int_X f\, d\mu = 1$. Let W_2 be the Wasserstein distance on the space $\mathcal{P}(X)$ defined in (0.2). Then for any couple of measures (ν_1, ν_2) in $\mathcal{P}(X)$ of finite distance $W_2(\nu_1, \nu_2) < +\infty$, there exists $\pi_o \in \mathcal{C}(\nu_1, \nu_2)$ such that

$$W_2^2(\nu_1, \nu_2) = \int_{X \times X} |x - y|_H^2\, \pi_o(dx, dy). \tag{1.1}$$

Such a π_o is called the optimal coupling plan between ν_1 and ν_2. The following result due to Feyel and Üstunel is our starting point

Theorem FU ([FU, Th. 6.1) *Let $\nu_1 = \rho_1\mu, \nu_2 = \rho_2\mu$ such that $W_2(\nu_1, \nu_2) < +\infty$. Then there exists a unique optimal coupling plan $\pi_o \in \mathcal{C}(\nu_1, \nu_2)$; moreover there exists a unique Borel map $\xi : X \to H$ such that for any bounded Borel function φ on $X \times X$*

$$\int_{X\times X} \varphi(x, y)\pi_o(dx, dy) = \int_X \varphi(x, x + \xi(x))\, d\nu_1(x)$$

and the transformation $T : x \mapsto x + \xi(x)$ is invertible.

It is obvious that T pushes ν_1 forward to ν_2 and

$$W_2^2(\nu_1, \nu_2) = \int_X |\xi(x)|_H^2 \, d\nu_1(x). \tag{1.2}$$

Recall that Talagrand's inequality $W_2^2(\mu, \rho\mu) \le 2\mathrm{Ent}(\rho)$ which was first proven for Gaussian measures on $\mathbf{R}^n$ [Ta] also holds true on the Wiener space [FU] [Gen](see [BGL], [OV] for related topics). It immediately implies $W_2(\nu_1, \nu_2) < +\infty$ whenever $\mathrm{Ent}(\rho_1)$ and $\mathrm{Ent}(\rho_2)$ are finite. Therefore W_2 induces a true distance on the space

$$\mathcal{P}^*(X) = \big\{\nu = \rho\mu;\ \mathrm{Ent}(\rho) < +\infty\big\}. \tag{1.3}$$

For $\nu = \rho\mu \in \mathcal{P}^*(X)$, it is convenient sometimes to use the notation $\mathrm{Ent}(\nu)$ instead of $\mathrm{Ent}(\rho)$. Since the distance d_H is stronger than the norm on X, a sequence of probability measures $(\nu_n)_{n\ge1}$ on X converges to ν with respect to W_2, converges also with respect to the Wasserstein distance defined using the norm of X; therefore ν_n converges weakly to ν (see for example [Vi]). In what follows, we give a direct proof using Theorem FU.

Proposition 1.1 *Let $(\nu_n)_{n\ge1}$ be a sequence in $\mathcal{P}^*(X)$ such that $W_2(\nu_n, \nu) \to 0$ as $n \to +\infty$ for $\nu \in \mathcal{P}^*(X)$. Then ν_n converges weakly to ν.*

Proof. By Theorem FU, there exist $\xi_n : X \to H$ such that $I + \xi_n$ pushes ν forward to ν_n and $W_2^2(\nu_n, \nu) = \int_X |\xi_n|_H^2 d\nu$. Set $\sigma_n = W_2^2(\nu_n, \nu)$. Let $\varphi : X \to \mathbf{R}$ be a bounded continuous function. We have

$$\begin{aligned} &\Big|\int_X \varphi d\nu - \int_X \varphi d\nu_n\Big| \le \int_X |\varphi(x) - \varphi(x + \xi_n(x))|\, d\nu(x) \\ &\le \int_{\{|\xi_n|_H \ge \varepsilon_n\}} |\varphi(x) - \varphi(x + \xi_n(x))|\, d\nu(x) + \int_{\{|\xi_n|_H \le \varepsilon_n\}} |\varphi(x) - \varphi(x + \xi_n(x))|\, d\nu(x), \end{aligned} \tag{1.4}$$

where ε_n are chosen so that $\lim_{n\to+\infty} \frac{\sigma_n}{\varepsilon_n^2} = 0$. The first term on the right hand of (1.4) is dominated by

$$2||\varphi||_\infty \frac{1}{\varepsilon_n^2} \int_X |\xi_n|_H^2 d\nu(x) = 2||\varphi||_\infty \frac{\sigma_n}{\varepsilon_n^2} \to 0 \text{ as } n \to +\infty;$$

for the second term, it is sufficient to notice that $\mathbf{1}_{\{|\xi_n(x)|_H \le \varepsilon_n\}} |\varphi(x) - \varphi(x + \xi_n(x))|$ tends to 0 as $n \to +\infty$ for ν-almost everywhere $x \in X$. Therefore letting $n \to +\infty$ in (1.4) gives the result. ∎

Theorem 1.2 *Let $R > 0$. Then the subset*

$$K_R = \{\nu \in \mathcal{P}^*(X);\ \mathrm{Ent}(\nu) \le R\}$$

is compact in $\mathcal{P}^*(X)$ with respect to the weak topology.

Proof. By the superlinear growth of $s \to s \log s$, K_R is weakly compact in $L^1(X, \mu)$. Combining with the lower semicontinuity of $\nu \mapsto \mathrm{Ent}(\nu)$ (see for example [AGS], [St, p.102], [JKO]), the result follows. ∎

Corollary 1.3 *Let $\nu_0 \in \mathcal{P}^*(X)$ be given. Then the subset*

$$C_R = \{\nu \in \mathcal{P}^*(X);\ W_2^2(\nu_0, \nu) + \mathrm{Ent}(\nu) \le R\}$$

is compact.

Proof. It is sufficient to notice that $\nu \mapsto W_2^2(\nu_0, \nu) + \mathrm{Ent}(\nu)$ is lower semi-continuous for the weak topology. ∎

Let ν_0 and ν_1 in $\mathcal{P}^*(X)$. Let ξ and π_o be given in Theorem FU. We set, for $0 \le t \le 1$,

$$\nu_t = (I + t\xi)_* \nu_0 \tag{1.5}$$

and $\pi_t \in \mathcal{C}(\nu_0, \nu_t)$ defined by

$$\int_{X\times X} \varphi(x, y) \pi_t(dx, dy) = \int_X \varphi(x, x + t\xi(x))\, d\nu_0(x). \tag{1.6}$$

Proposition 1.4 *We have for $0 \le s < t \le 1$,*

$$W_2(\nu_s, \nu_t) = (t - s) W_2(\nu_0, \nu_1). \tag{1.7}$$

Proof. See [AGS] and [FU2]. ∎

The above result says that $t \to \nu_t$ defined in (1.5) is a geodesic with constant speed. Taking $s = 0$ in (1.7), we see that π_t defined in (1.6) is the unique optimal coupling plan in $\mathcal{C}(\nu_0, \nu_t)$, supported by the graph of $T_t := I + t\xi$. The following result strengthen Theorem 7.3 in [FU].

Theorem 1.5 *Let ν_t be defined in* (1.5)*. Then $\nu_t \in \mathcal{P}^*(X)$ and for $0 \le t \le 1$,*

$$\mathrm{Ent}(\nu_t) \le (1 - t)\mathrm{Ent}(\nu_o) + t\mathrm{Ent}(\nu_1) - \frac{t(1-t)}{2} W_2^2(\nu_o, \nu_1). \tag{1.8}$$

Proof. Firstly remark that if ρ_0 and ρ_1 are cylindrical, then (1.8) is reduced to a finite dimensional case: it holds true (see [AS] [AGS]). Secondly for the general case, we consider a sequence of increasing subspaces $V_n \subset X^*$ such that $\cup_n V_n$ is dense in H (with respect to the norm of H). Let $P_n : X \to V_n$ be the projection and denote by $\mathbf{E}^{V_n}$ the conditional expectation with respect to the sub σ-field on X, generated by P_n. Note that $(P_n)_* \mu$ is the standard Gaussian measure γ_n on V_n. Set

$$\rho_0^n = \mathbf{E}^{V_n}(\rho_0), \rho_1^n = \mathbf{E}^{V_n}(\rho_1).$$

Then ρ_0^n, ρ_1^n converge in $L^1(X,\mu)$, respectively to ρ_0 and ρ_1; therefore the measures $\rho_0^n\mu$ (resp. $\rho_1^n\mu$) converges weakly to $\rho_0\mu$ (resp. $\rho_1\mu$) as $n \to +\infty$. Let $\pi_n \in \mathcal{C}(\rho_0^n\mu, \rho_1^n\mu)$ be the optimal coupling plan. Up to a subsequence, π_n converges weakly to $\hat\pi \in \mathcal{C}(\rho_0\mu, \rho_1\mu)$. Then we have

$$
\begin{aligned}
W_2^2(\rho_0\mu, \rho_1\mu) &\le \int_{X\times X} |x-y|_H^2 \hat\pi(dx,dy) \\
&\le \liminf_{n\to+\infty} \int_{X\times X} |x-y|_H^2 \pi_n(dx,dy) = \liminf_{n\to+\infty} W_2^2(\rho_0^n\mu, \rho_1^n\mu).
\end{aligned}
\tag{1.9}
$$

Now we will prove that $\hat\pi$ realizes the minimum:

$$
W_2^2(\nu_0,\nu_1) = \int_{X\times X} |x-y|_H^2 \hat\pi(dx,dy). \tag{1.10}
$$

To this end, introduce the functions $\tilde\rho_i^n : V_n \to \mathbf{R}$ such that $\rho_i^n = \tilde\rho_i^n \circ P_n$ for $i = 0,1$. Define $\hat\pi_n \in \mathcal{C}(\tilde\rho_0^n\gamma_n, \tilde\rho_1^n\gamma_n)$ by

$$
\int_{V_n\times V_n} \psi(z_1,z_2)\hat\pi_n(dz_1,dz_2) = \int_{X\times X} \psi(P_n(x),P_n(y))\pi(dx,dy)
$$

where $\pi \in \mathcal{C}(\nu_0,\nu_1)$ is the optimal coupling plan. We have

$$
\begin{aligned}
W_2^2(\rho_0^n\mu, \rho_1^n\mu) = W_2^2(\tilde\rho_0^n\gamma_n, \tilde\rho_1^n\gamma_n) &\le \int_{V_n\times V_n} |z_1-z_2|^2 \hat\pi_n(dz_1,dz_2) \\
= \int_{X\times X} |P_n(x-y)|^2 \pi(dx,dy) &\le \int_{X\times X} |x-y|_H^2 \pi(dx,dy) = W_2^2(\nu_0,\nu_1).
\end{aligned}
$$

Combining with (1.9), we get the equality (1.10). By uniqueness of optimal coupling plan, we conclude that $\hat\pi = \pi$. Now define

$$
\int_X \varphi\, d\nu_t^n = \int_{X\times X} \varphi((1-t)x+ty)\pi_n(dx,dy). \tag{1.11}
$$

Then for any bounded continuous function $\varphi : X \to \mathbf{R}$,

$$
\lim_{n\to+\infty} \int_X \varphi\, d\nu_t^n = \int_{X\times X} \varphi((1-t)x+ty)\pi(dx,dy). \tag{1.12}
$$

This means that the sequence (ν_t^n) converges weakly to ν_t defined in (1.5), as $n \to +\infty$.
By the first case, we can apply (1.8) to ν_t^n to get

$$
\mathrm{Ent}(\nu_t^n) \le (1-t)\mathrm{Ent}(\nu_0^n) + t\mathrm{Ent}(\nu_1^n) - \frac{t(1-t)}{2} W_2^2(\nu_0^n,\nu_1^n).
$$

For any $\varepsilon > 0$, by (1.9), there exists $n_0 > 0$ such that

$$
W_2^2(\rho_0\mu, \rho_1\mu) - \varepsilon \le W_2^2(\rho_0^n\mu, \rho_1^n\mu), \quad n \ge n_0.
$$

By Jensen inequality $\mathrm{Ent}(\nu_0^n) \le \mathrm{Ent}(\nu_0)$ and $\mathrm{Ent}(\nu_1^n) \le \mathrm{Ent}(\nu_1)$. Then for $n \ge n_0$,

$$
\mathrm{Ent}(\nu_t^n) \le (1-t)\mathrm{Ent}(\nu_0) + t\mathrm{Ent}(\nu_1) - \frac{t(1-t)}{2}\Big(W_2^2(\nu_0,\nu_1) - \varepsilon\Big). \tag{1.13}
$$

By Theorem 1.2, $\nu_t \in \mathcal{P}^*(X)$ and $\mathrm{Ent}(\nu_t)$ is dominated by the right hand of (1.13). Letting $\varepsilon \to 0$ gives (1.8). ∎

Remark: The inequality (1.8) says that the entropy functional is 1-convex along geodesics. The assertion of Theorem 1.5 was already stated in [St, p.125]. Moreover, a sketch of a proof was indicated, based on approximation of X by finite dimensional subspaces equipped with Gaussian measures. However, due the the degeneracy of the metric on X, the proof requires a more careful argumentation since e.g. $W_2(\mu, \gamma_n) = +\infty$.

2. Benamou-Brenier's formula

An absolutely continuous curve $\{c(t); t \in [0,1]\}$ on a Riemannian manifold M admits tangent vectors $c'(t) \in T_{c(t)}M$ for almost everywhere $t \in]0,1[$. In order to understand the tangent spaces of the Wasserstein space $(\mathcal{P}^*(X), W_2)$, it is convenient to consider absolutely continuous curves (ν_t) in $\mathcal{P}^*(X)$.

Definition 2.1 *We say that a curve $(\nu_t)_{t\in[0,1]}$ is in the class* AC_2 *if there exists $m \in L^2([0,1])$ such that*

$$W_2(\nu_{t_1}, \nu_{t_2}) \le \int_{t_1}^{t_2} m(s)ds, \quad t_1 \le t_2. \tag{2.1}$$

For such a curve, for a.e. $t \in [0,1]$,

$$\limsup_{\varepsilon\to 0} \frac{W_2(\nu_{t+\varepsilon}, \nu_t)}{|\varepsilon|} \le m(t). \tag{2.2}$$

For any curve $(\nu_t)_{t\in[0,1]}$ in AC_2, the limit

$$|\nu'|(t) := \lim_{\varepsilon\to 0} \frac{W_2(\nu_{t+\varepsilon}, \nu_t)}{|\varepsilon|}$$

exists for a.e. $t \in [0,1]$, which is called the *metric derivative* of $(\nu_t)_{t\in[0,1]}$(see [AGS, Theorem 1.1.2]). The function $t \mapsto |\nu'|(t)$ belongs to $L^2([0,1])$ and (2.1) holds w.r.t. $|\nu'|(t)$. It is minimal in the sense that for each function m satisfying (2.1), it holds

$$|\nu'|(t) \le m(t), \quad a.e. t \in [0,1].$$

Note that the curve defined in (1.5) is in the class AC_2 due to (1.7). In order to construct another examples, we will recall some elements in Malliavin Calculus (see [Ma] for more details).

A function $F : X \to \mathbf{R}$ is said to be cylindrical if it is written in the form

$$F(x) = f(e_1(x), \dots, e_K(x)), \quad f \in C_c^\infty(\mathbf{R}^K), \tag{2.3}$$

where $\{e_i \in X^*;\ i \ge 1\}$ is a given orthonormal basis of H. We will denote by $\mathrm{Cylin}(X)$ the totality of such cylindrical functions. Note that $\mathrm{Cylin}(X)$ is not a vector space. A cylindrical vector field Z on X is a map $X \to H$ in the form

$$Z = \sum_{j=1}^K F_j h_j, \quad \text{with } F_j \in \mathrm{Cylin}(X), h_j \in X^*. \tag{2.4}$$

For a function $F \in \mathrm{Cylin}(X)$ in the form (2.3), we define

$$\nabla F(x) = \sum_{i=1}^{K} (\partial_i f)(e_1(x), \ldots, e_K(x)) e_i,$$

which is a cylindrical vector field on X, where $\partial_i f$ denotes the derivative with respect to the ith component. Similarly, for Z given above, we define $\nabla Z = \sum_{j=1}^{K} \nabla F_j \otimes h_j$. Now we denote by $\mathbf{D}_1^p(X)$ the Sobolev space which is the closure of $\mathrm{Cylin}(X)$ under the norm $||F||_{1,p}^p = \int_X (|F|^p + |\nabla F|_H^p)\, d\mu$; and $\mathbf{D}_1^p(X;H)$ the closure of cylindrical vector fields under the norm $||Z||_{1,p}^p = \int_X (|Z|_H^p + |\nabla F|_{H\otimes H}^p)\, d\mu$. In the similar way, we define the Sobolev spaces $\mathbf{D}_r^p(X)$ where $r \in \mathbf{N}$ is the order of the derivative. Then for $p > 1$ and $Z \in \mathbf{D}_1^p(X;H)$, the divergence $\delta(Z) \in L^p(X)$ exists such that

$$\int_X F\, \delta(Z)\, d\mu = \int_X \langle \nabla F, Z \rangle_H \, d\mu, \quad F \in \mathrm{Cylin}(X).$$

For a vector field Z given by (2.4), the divergence $\delta(Z)$ admits the expression

$$\delta(Z) = \sum_{j=1}^{K} \big(F_j h_j(x) - \langle \nabla F_j(x), h_j \rangle_H \big) \tag{2.5}$$

Note that $\delta(Z)$ is a continuous function of x. Now pick $Z \in \cap_{p>1, r\geq 1} \mathbf{D}_r^p(X;H)$ and assume that (2.6)

$$\int_X e^{\varepsilon_0 |Z|_H^2}\, d\mu < +\infty \text{ for a small } \varepsilon_0 > 0 \quad \text{and} \quad \int_X e^{\lambda_0 (|\delta(Z)| + ||\nabla Z||)}\, d\mu < +\infty \text{ for some } \lambda_0 > 0.$$

Then by Cruzeiro [Cr], there exists a flow of measurable maps $U_t : X \to X$ such that for a.e. $x \in X$,

$$U_t(x) = x + \int_0^t Z(U_s(x))\, ds, \quad t > 0,$$

and $U_{t+s} = U_t \circ U_s$, $(U_t)_* \mu = K_t \mu$ with (see also [Dr] for a detailed proof):

$$K_t = \exp\Big(\int_0^t \delta Z(U_{-s}(x))\, ds \Big), \quad \sup_{0 \leq t \leq T} ||K_t||_{L^2}^2 \leq \int_X e^{4T|\delta(Z)|} d\mu \text{ for } T \leq \lambda_0/4. \tag{2.7}$$

Proposition 2.2 *Let $\nu_0 = \rho_0 \mu \in \mathcal{P}^*(X)$. Define $\nu_t = (U_t)_* \nu_0$. Then under the condition* (2.6), *the curve $(\nu_t)_{t\in[0,1]}$ is in the class* AC_2.

Proof. By definition, $\int_X \varphi\, d\nu_t = \int_X \varphi(U_t) \rho_0\, d\mu = \int_X \varphi \rho_0(U_{-t}) K_t\, d\mu$ holds for any bounded Borel function φ. If we denote $\nu_t = \rho_t \mu$, then $\rho_t = \rho_0(U_{-t}) K_t$, and

$$\mathrm{Ent}(\rho_t) = \mathrm{Ent}(\rho_0) + \int_X (\log K_t(U_t))\, \rho_0 d\mu. \tag{2.8}$$

Using (2.7), $|\log K_t(U_t)| \leq \int_0^t |\delta Z(U_{t-s})|\, ds$ and by Young inequality $uv \leq e^u + v \log v$ for $u, v \geq 0$, we have for any $\eta > 0$

$$|\log K_t(U_t)|\, \rho_0 \leq \int_0^t e^{\eta |\delta Z|(U_{t-s})}\, ds + \frac{\rho_0}{\eta} \log \frac{\rho_0}{\eta}.$$

Then (2.8) yields

$$\mathrm{Ent}(\rho_t) \leq \mathrm{Ent}(\rho_0) + \int_0^t \int_X e^{\eta |\delta Z|(U_{t-s})} d\mu\, ds + \mathrm{Ent}(\frac{\rho_0}{\eta}).$$

Now we will prove that for η small enough

$$\sup_{0\leq t\leq 1} \int_X e^{\eta|\delta Z|(U_t)} d\mu < +\infty. \tag{2.9}$$

First of all, for $T_0 \leq \lambda_0/4$ and $t \in [0, T_0]$, we have,

$$\begin{aligned}
\int_X e^{\eta|\delta Z|(U_t)} d\mu = \int_X e^{\eta|\delta Z|} K_t\, d\mu &\leq \Big(\int_X e^{2\eta|\delta Z|} d\mu\Big)^{1/2} ||K_t||_{L^2} \\
&\leq \Big(\int_X e^{2\eta|\delta Z|} d\mu\Big)^{1/2} \Big(\int_X e^{\lambda_0|\delta(Z)|} d\mu\Big)^{1/2},
\end{aligned}$$

where we used (2.7) for estimating $||K_t||_{L^2}$. Let $A = \Big(\int_X e^{\lambda_0|\delta(Z)|} d\mu\Big)^{1/2}$. Now using the property of flow,

$$\begin{aligned}
\int_X e^{\eta|\delta Z|(U_{T_0+t})} d\mu &= \int_X e^{\eta|\delta Z|(U_{T_0})} K_t\, d\mu \\
&\leq \Big(\int_X e^{2\eta|\delta(Z)|(U_{T_0})} d\mu\Big)^{1/2} \cdot A \\
&\leq \Big(\int_X e^{2^2\eta|\delta(Z)|} d\mu\Big)^{1/2^2} \cdot A^2.
\end{aligned}$$

Let N be the integer such that $N\lambda_0 \geq 1$, then by induction, we have for each $t \in [0, 1]$

$$\int_X e^{\eta|\delta Z|(U_t)} d\mu \leq \Big(\int_X e^{2^N\eta|\delta(Z)|} d\mu\Big)^{1/2^N}.$$

So we get (2.9), which implies that $\mathrm{Ent}(\rho_t) < +\infty$.
Now let $t_1 < t_2$. Define a probability measure π on $X \times X$ by

$$\int_{X\times X} \varphi(x, y)\pi(dx, dy) = \int_X \varphi(U_{t_1}, U_{t_2})\, d\nu_0.$$

Then $\pi \in \mathcal{C}(\nu_{t_1}, \nu_{t_2})$ and

$$W_2^2(\nu_{t_1}, \nu_{t_2}) \leq \int_X |U_{t_1} - U_{t_2}|_H^2\, d\nu_0.$$

But for a.e $x \in X$, $|U_{t_1} - U_{t_2}|_H \leq \int_{t_1}^{t_2} |Z(U_s)|_H\, ds$; therefore

$$W_2(\nu_{t_1}, \nu_{t_2}) \leq \Big\|\int_{t_1}^{t_2} |Z(U_s)|_H\, ds\Big\|_{L^2(\nu_0)} \leq \int_{t_1}^{t_2} ||Z(U_s)||_{L^2(\nu_0)}\, ds. \tag{2.10}$$

Let $m(s) = ||Z(U_s)||_{L^2(\nu_0)}$. We have for any $\varepsilon > 0$

$$m(s)^2 = \int_X |Z(U_s)|_H^2 \rho_0 \, d\mu \le \int_X e^{\varepsilon\, |Z(U_s)|_H^2} \, d\mu + \mathrm{Ent}(\rho_0/\varepsilon).$$

The same procedure as above yields $\int_0^1 m(s)^2 ds < +\infty$. ∎

Theorem 2.3 *Let $(\nu_t)_{t\in[0,1]}$ be a curve in AC_2. Then there exists a Borel vector field $(t,x) \mapsto Z_t(x) \in H$ such that $\int_0^1 ||Z_t||^2_{L^2(\nu_t)}\, dt < +\infty$ and the continuity equation*

$$\frac{\partial \nu_t}{\partial t} + \nabla \cdot (Z_t \nu_t) = 0 \quad \text{in }]0,1[\times X \tag{2.11}$$

holds in the sense (see [Kr]) that

$$\int_0^1 \int_X \big(\alpha'(t)F(x) + \big\langle Z_t(x), \nabla F(x)\big\rangle_H \alpha(t)\big) d\nu_t(x) dt = 0 \tag{2.12}$$

for all $\alpha \in C_c^\infty(]0,1[)$ and $F \in \mathrm{Cylin}(X)$.

Proof. Denote $\Sigma = \{(x,y) \in X \times X;\ x - y \in H\}$. For $s \in]0,1[$ and $\eta > 0$ small enough, we consider the optimal coupling plan $\pi_\eta \in \mathcal{C}(\nu_s, \nu_{s+\eta})$. Then the support of π_η is included in Σ. For $(x,y) \in \Sigma$, we have

$$F(y) - F(x) = \int_0^1 \big\langle (\nabla F)(ty + (1-t)x),\ y - x\big\rangle_H \, dt.$$

Set $H(x,y) = \int_0^1 (\nabla F)(ty + (1-t)x)\, dt$. By expression (2.4), we see that $(x,y) \mapsto H(x,y)$ is a bounded continuous function from $X \times X$ to H. Then

$$\int_X F d\nu_{s+\eta} - \int_X F d\nu_s = \int_\Sigma \big\langle H(x,y), y - x\big\rangle_H \pi_\eta(dx,dy).$$

The Cauchy-Schwarz inequality yields, for $\eta > 0$,

$$\frac{1}{\eta}\Big|\int_X F d\nu_{s+\eta} - \int_X F d\nu_s\Big| \le \frac{W_2(\nu_s, \nu_{s+\eta})}{\eta} \Big(\int_\Sigma |H(x,y)|_H^2 \pi_\eta(dx,dy)\Big)^{1/2}. \tag{2.13}$$

Take a sequence η_n such that $\lim\limits_{n\to+\infty} \frac{1}{\eta_n}\Big|\int_X F d\nu_{s+\eta_n} - \int_X F d\nu_s\Big| = \overline{\lim}_{\eta\to 0} \frac{1}{\eta}\Big|\int_X F d\nu_{s+\eta} - \int_X F d\nu_s\Big|$. As $\nu_{s+\eta_n}$ converges to ν_s with respect to W_2, it converges weakly; therefore the family $\{\pi_{\eta_n}; n \ge 1\}$ is tight. Up to a subsequence, π_{η_n} converges to $\hat\pi \in \mathcal{C}(\nu_s, \nu_s)$. We have

$$\int_{X\times X} |x-y|_H^2 \hat\pi(dx,dy) \le \underline{\lim}_{n\to+\infty} \int_{X\times X} |x-y|_H^2 \pi_{\eta_n}(dx,dy) = \lim_{n\to+\infty} W_2^2(\nu_s, \nu_{s+\eta_n}) = 0,$$

so $\hat\pi$ is supported by the diagonal $D = \{(x,y) \in \Sigma; x = y\ \}$. Hence

$$\lim_{n\to+\infty} \int_\Sigma |H(x,y)|_H^2 \pi_{\eta_n}(dx,dy) = \int_D |H(x,x)|_H^2 \hat\pi(dx,dy) = \int_X |\nabla F|_H^2 \, d\nu_s.$$

According to (2.2) and (2.13), for a.e $s \in]0,1[$,

$$\overline{\lim}_{\eta\downarrow 0} \frac{1}{\eta}\Big|\int_X F d\nu_{s+\eta} - \int_X F d\nu_s\Big| \le m(s)\, ||\nabla F||_{L^2(\nu_s)}. \tag{2.14}$$

Now take $\delta > 0$ such that $\mathrm{supp}(\alpha)+]-\delta,\delta[\subset]0,1[$. Then for $0<\eta<\delta$,

$$\int_0^1\int_X \alpha(s)F(x)\,d\nu_{s+\eta}(x)\,ds = \int_0^1\int_X \alpha(s-\eta)F(x)\,d\nu_s(x)ds,$$

and

$$\begin{aligned}&\int_0^1 \frac{1}{\eta}\Big[\int_X \alpha(s)F(x)\,d\nu_s(x) - \int_X \alpha(s)F(x)\,d\nu_{s+\eta}(x)\Big]\,ds\\ &=\int_0^1\int_X \frac{1}{\eta}\big[\alpha(s)-\alpha(s-\eta)\big]F(x)\,d\nu_s(x)ds.\end{aligned}\tag{2.15}$$

It is obvious that as $\eta\to 0$, the right hand side of (2.15) tends to $\int_0^1\int_X \alpha'(s)F(x)\,d\nu_s(x)ds$. By (2.1), $\frac{1}{\eta}W_2(\nu_s,\nu_{s+\eta})\le\frac{1}{\eta}\int_s^{s+\eta}|\nu'|(u)du$ and the fact that $s\mapsto\sup_{\eta>0}\big(\frac{1}{\eta}\int_s^{s+\eta}|\nu'|(u)du\big)$ is integrable over $[0,1]$. Now we can use (2.14) to get that

$$\begin{aligned}\Big|\int_0^1\int_X \alpha'(s)F(x)\,d\nu_s(x)ds\Big| &\le \int_0^1 m(s)\,||\alpha(s)\nabla F||_{L^2(\nu_s)}\,ds\\ &\le\Big(\int_0^1|\nu'|^2(s)ds\Big)^{1/2}\Big(\int_0^1\int_X|\alpha(s)\nabla F(x)|_H^2\,d\nu(x)ds\Big)^{1/2}.\end{aligned}\tag{2.16}$$

Let P_ν be the probability measure on $[0,1]\times X$ defined by

$$\int_{[0,1]\times X}\varphi(s,x)dP_\nu(s,x) = \int_0^1\int_X \varphi(s,x)d\nu_s(x)ds.$$

Introduce the vector space

$$V=\Big\{\sum_{i=1}^K \alpha_i(s)\nabla F_i(x);\ \alpha_i\in C_c^\infty(]0,1[,F_i\in \mathrm{Cylin}(X),K\in\mathbf{N}\Big\}.$$

Let $\overline{V}$ be the closure of V in $L^2(P_\nu)$. Define for $\psi=\sum_{i=1}^K\alpha_i(s)\nabla F_i(x)\in V$,

$$L(\psi) = -\sum_{i=1}^K\int_0^1\int_X \alpha_i'(s)F_i(x)\,d\nu_s(x)ds.\tag{2.17}$$

By linearity of the two sides of (2.15), the inequality (2.16) holds for ψ, that is

$$|L(\psi)|\le\sqrt{\int_0^1|\nu'|^2(s)ds}\cdot||\psi||_{L^2(P_\nu)}.\tag{2.18}$$

It follows that L is well defined and is a bounded linear operator on V. Therefore there exists $Z\in\overline{V}$ such that

$$L(\psi)=\int_0^1\int_X\langle Z,\psi\rangle_H\,d\nu_s ds,\quad \psi\in V.$$

Now take $\psi = \alpha \nabla F$ and according (2.17), we get (2.12). Moreover,

$$\|Z\|^2_{L^2(P_\nu)} = \int_0^1 \int_X |Z(t,x)|^2_H \, d\nu_s(x) ds \leq \int_0^1 |\nu'|^2(s)\, ds. \tag{2.19}$$

∎

Following [AS] and [AGS], we define, for any $\nu \in \mathcal{P}^*(X)$,

$$\mathcal{E} = \Big\{ \sum_{i=1}^K \nabla F_i;\ F_i \in \mathrm{Cylin}(X) \Big\}, \quad T_\nu = \text{closure of } \mathcal{E} \text{ in } L^2(X, H, \nu). \tag{2.20}$$

Proposition 2.4 *Let Z be constructed as in Theorem 2.3. Then for a.e. $t \in]0,1[$, $Z(t,\cdot) \in T_{\nu_t}$. The solution to* (2.11) *satisfying this property is unique. Moreover, it holds that*

$$W_2^2(\nu_0, \nu_1) \leq \int_0^1 \int_X |Z(s,x)|^2_H \, d\nu_s(x) ds, \quad \textit{and } \|Z\|^2_{L^2(P_\nu)} = \int_0^1 |\nu'|^2(s)\, ds. \tag{2.21}$$

Proof. Let $\psi_n \in V$ such that $||Z - \psi_n||_{L^2(P_\nu)} \to 0$. Or

$$\lim_{n\to+\infty} \int_0^1 \Big(\int_X |Z(t,x) - \psi_n(t,x)|^2_H d\nu_t(x) \Big) dt = 0.$$

Then up to a subsequence, for a.e. $t_o \in]0,1[$,

$$\lim_{n\to+\infty} \int_X |Z(t_o,x) - \psi_n(t_o,x)|^2_H d\nu_{t_o}(x) = 0.$$

This means that $Z(t_o,\cdot) \in T_{\nu_{t_o}}$. Now let $\hat{Z}$ be another solution to (2.12) such that $\hat{Z}(t,\cdot) \in T_{\nu_t}$ for a.e $t \in]0,1[$. Then we have

$$\int_0^1 \alpha(t) \Big(\int_X \big\langle Z(t,x) - \hat{Z}(t,x), \nabla F(x) \big\rangle_H d\nu_t(x) \Big)\, dt = 0.$$

It follows that $\int_X \big\langle Z(t,x) - \hat{Z}(t,x), \nabla F(x) \big\rangle_H d\nu_t(x) = 0$ holds for t in a full measure subset $\Omega_F \subset]0,1[$. For each $K \geq 1$, let $\mathcal{D}_K \subset C_c^\infty(\mathbf{R}^K)$ be a dense countable subset. Set

$$\mathcal{D} = \Big\{ \sum_{i=1}^m f_i \circ P_{K_i};\ f_i \in \mathcal{D}_{K_i}, m \in \mathbf{N} \Big\}, \tag{$*$}$$

where $P_K : X \to V_K = \mathrm{span}\{e_1, \cdots, e_K\}$. For each $\nabla F \in \mathcal{E}$, there exists a finite number of $K_1, \cdots, K_q$ such that $F = \sum_{i=1}^q f_i \circ P_{K_i}$ with $f_i \in C_c^\infty(\mathbf{R}^{K_i})$. We have $\nabla F = \sum_{i=1}^q (\nabla_{\mathbf{R}^{K_i}} f_i) \circ \mathcal{P}_{K_i}$. Therefore there exists $F_n \in \mathcal{D}$ such that

$$\sup_{x\in X} |\nabla F_n(x) - \nabla F(x)|_H \to 0.$$

Define $\Omega_Z = \cap_{F\in\mathcal{D}}\Omega_F$. Then for $t \in \Omega_Z$, $\int_X \big\langle Z(t,x) - \hat{Z}(t,x), \nabla F(x)\big\rangle_H d\nu_t(x) = 0$ holds for all $\nabla F \in \mathcal{E}$. Therefore $Z(t,\cdot) = \hat{Z}(t,\cdot)$ ν_t-a.e. For proving (2.21), we consider a sequence of increasing subspaces $V_n \subset X^*$ such that $\cup_n V_n$ is dense in H. Define $\nu_t^{(n)} = (P_n)_*\nu_t$. Since $W_2(\nu_t^{(n)}, \nu_s^{(n)}) \le W_2(\nu_t, \nu_s)$, $t \to \nu_t^{(n)}$ is also an absolutely continuous curve in AC_2. Therefore, according to the result on finite dimensional spaces (see [AS][AGS]), there exists $Z_t^{(n)}$ such that $\int_0^1\int_{V_n} |Z_t^{(n)}|^2 d\nu_t^{(n)} dt < +\infty$ and the continuity equation

$$\frac{d\nu_t^{(n)}}{dt} + \nabla\cdot(Z_t^{(n)}\nu_t^{(n)}) = 0$$

holds in the distribution sense:

$$\int_0^1\int_{V_n} (\alpha'(t) f + \langle Z_t^{(n)}, \nabla f\rangle \alpha(t))\, d\nu_t^{(n)} dt = 0,$$

or

$$\int_0^1\int_X (\alpha'(t) f\circ P_n + \big\langle Z_t^{(n)}\circ P_n, \nabla f\circ P_n\big\rangle_H \alpha(t))\, d\nu_t dt = 0.$$

In the continuity equation (2.12), take $F = f\circ P_n$ with $f \in C_c^\infty(V_n)$, we get

$$\int_0^1\int_X \big(\alpha'(t)\, f\circ P_n + \big\langle Z_t, \nabla f\circ P_n\big\rangle_H\, \alpha(t)\big)\, d\nu_t dt = 0.$$

¿From the above two equations, we deduce that for a.e $t \in]0,1[$, $P_n Z_t - Z_t^{(n)}\circ P_n$ is orthogonal in $L^2(\nu_t)$ to the space $\overline{\{\nabla f\circ P_n; f\in C_c(V_n)\}}^{L^2(\nu_t)}$, which contains $Z_t^{(n)}\circ P_n$. It follows that

$$||Z_t^{(n)}||_{L^2(\nu_t^{(n)})} \le ||P_n Z_t||_{L^2(\nu_t)} \le ||Z_t||_{L^2(\nu_t)}.$$

In the finite dimensional case, it holds that (see [AGS, Theorem 8.3.1])

$$(*) \qquad W_2(\nu_t^{(n)}, \nu_s^{(n)}) \le \int_s^t ||Z_u^{(n)}||_{L^2(\nu_u^{(n)})} du.$$

For reader's convenience, we will give a sketch of the proof of $(*)$ as in [AGS]. To this end, we omit (n).

(i) If Z_t is a good vector field on $\mathbf{R}^d$, more precisely, assume that

$$\int_0^1 \Big(\sup_{x\in B}|Z_t(x)| + \mathrm{Lip}(Z_t, B)\Big)\, dt < +\infty \text{ for all ball } B\subset\mathbf{R}^d,$$

and $\int_0^1\int_{R^d} |Z_t| d\nu_t dt < +\infty$,where $\mathrm{Lip}(Z_t, B)$ is the Lipschitz constant of $x \to Z_t(x)$ on the ball B, such that

$$\frac{d\nu_t}{dt} + \nabla\cdot(Z_t\nu_t) = 0 \quad \text{on} \quad (0,1)\times\mathbf{R}^d,$$

then for ν_0-a.s $x\in\mathbf{R}^d$, the differential equation

$$X_t(x) = x + \int_0^t Z_s(X_s(x))\, ds$$

admits a unique solution $X_t(x)$ for $t \in [0,1]$ and $\nu_t = (X_t)_*\nu_0$. In this case, for the coupling measure $\pi \in \mathcal{C}(\nu_{t_1}, \nu_{t_2})$, defined by $\pi = (X_{t_1}, X_{t_2})_*\nu_0$, we have

$$W_2(\nu_{t_1}, \nu_{t_2}) \le \Big(\int_{\mathbf{R}^d} |X_{t_1} - X_{t_2}|^2 \, d\nu_0\Big)^{1/2} \le \int_{t_1}^{t_2} ||Z_s||_{L^2(\nu_s)} \, ds.$$

ii) For the general case, we regularize ν_t and Z_t by convolution product with the Gauss kernel $\rho_\varepsilon(x) = (2\pi\varepsilon)^{-d/2} e^{-|x|^2/2}$ by setting

$$\nu_t^\varepsilon = \nu_t * \rho_\varepsilon, \quad Z_t^\varepsilon = (Z_t \nu_t) * \rho_\varepsilon / \nu_t^\varepsilon.$$

Applying (i) gives

$$W_2(\nu_{t_1}^\varepsilon, \nu_{t_2}^\varepsilon) \le \int_{t_1}^{t_2} \Big(\int_{\mathbf{R}^d} |Z_s^\varepsilon|^2 d\nu_s^\varepsilon\Big)^{1/2} ds.$$

But by Jensen inequality

$$|Z_s^\varepsilon(x)|^2 \le \int_{\mathbf{R}^d} |Z_s(y)|^2 \frac{\rho_\varepsilon(x-y) d\nu_s(y)}{\nu_s^\varepsilon(x)},$$

which implies that $\int_{\mathbf{R}^d} |Z_s^\varepsilon|^2 d\nu_s^\varepsilon \le \int_{\mathbf{R}^d} |Z_s|^2 d\nu_s$. Using the lower semi-continuity of $(\mu, \nu) \to W_2(\mu, \nu)$, we get the desired result by letting $\varepsilon \downarrow 0$ in

$$W_2(\nu_{t_1}^\varepsilon, \nu_{t_2}^\varepsilon) \le \int_{t_1}^{t_2} ||Z_s||_{L^2(\nu_s)} \, ds.$$

Now we return to our situation. By $(*)$, we have $W_2(\nu_t^{(n)}, \nu_s^{(n)}) \le \int_s^t ||Z_u||_{L^2(\nu_u)} du$. Noting that $W_2(\nu_t, \nu_s) = \lim_{n \to +\infty} W_2(\nu_t^{(n)}, \nu_s^{(n)})$ and letting $n \to +\infty$, we get

$$W_2(\nu_t, \nu_s) \le \int_s^t ||Z_u||_{L^2(\nu_u)} du.$$

Hence,

$$|\nu'|(s) = \lim_{t \to s} \frac{W_2(\nu_t, \nu_s)}{|t-s|} \le \|Z_s\|_{L^2(\nu_s)},$$

$$\int_0^1 |\nu'|^2(s) \, ds \le \int_0^1 \int_X |Z(s,x)|_H^2 \, d\nu_s(x) ds.$$

Combining this with (2.19), we get the last inequality in (2.21) and the argument is complete now. ∎

Definition 2.5 *Let $\{\nu_t;\ t \in [0,1]\}$ be a family of probability measures in $\mathcal{P}^*(X)$. We will say that $t \to Z_t \in T_{\nu_t}$ is the derivative process of $t \mapsto \nu_t$ in the sense of Otto-Ambrosio-Savaré if $\int_0^1 \int_X |Z_t(x)|_H^2 d\nu_t(x) dt < +\infty$ and the continuity equation* (2.12) *holds. We denote Z_t by $\dfrac{d^o \nu_t}{dt}$.*

Using $\dfrac{d^o \nu_t}{dt}$, the result obtained in [AS, p.30] (for previous versions, see [BB], [Ot]) can be expressed exactly as a Riemannian distance. Namely, in our setting,

Theorem 2.6 *Let $\nu_0, \nu_1 \in \mathcal{P}^*(X)$ be given. Then*

$$W_2^2(\nu_0, \nu_1) = \inf\Big\{\int_0^1 \Big\|\frac{d^o\nu_t}{dt}\Big\|_{T_{\nu_t}}^2 dt;\ \nu_t \in \mathrm{AC}_2 \text{ connecting } \nu_0, \nu_1\Big\}. \tag{2.22}$$

Proof. Let ν_t be defined in (1.5). By (1.7), $W_2(\nu_s, \nu_t) = (t-s)W_2(\nu_0, \nu_1)$. Then taking $m(s) = W_2(\nu_0, \nu_1)$ in (2.18), we get

$$|L(\psi)| \leq ||\psi||_{L^2(P_\nu)} \cdot W_2(\nu_0, \nu_1).$$

Let $Z = \dfrac{d^o\nu_t}{dt}$ be given in Theorem 2.3. Then

$$\Big|\int_0 \int_X \langle Z, \psi\rangle_H \, d\nu_s ds\Big| \leq W_2(\nu_0, \nu_1) \cdot ||\psi||_{L^2(P_\nu)}, \quad \psi \in V.$$

It follows that $||Z||_{L^2(P_\nu)} \leq W_2(\nu_0, \nu_1)$. The equality is realized for $\dfrac{d^o\nu_t}{dt}$, according to (2.21). ∎

Corollary 2.7 *Let $\nu_0, \nu_1 \in \mathcal{P}^*(X)$ and ξ be given in Theorem FU. Define $T_t = I + t\xi$, $\nu_t = (T_t)_*\nu_0$ and $W_t = \xi(T_t^{-1})$. Then for a.e. $t \in]0,1[$, $W_t \in T_{\nu_t}$.*

Proof. By 1-convex inequality (1.8), $\nu_t \in \mathcal{P}^*(X)$, so T_t^{-1} exists for each $t \in [0,1]$. Let $F \in \mathrm{Cylin}(X)$. We have

$$\frac{d}{dt}\int_X F\, d\nu_t = \frac{d}{dt}\int_X F(x + t\xi(x))d\nu_0(x) = \int_X \langle \nabla F(T_t), \xi\rangle_H \, d\nu_0 = \int_X \langle \nabla F, W_t\rangle_H d\nu_t.$$

On the other hand, let $Z(t,x) = \dfrac{d^o\nu_t}{dt}$. The equation (2.12) implies that for a.e $t \in]0,1[$,

$$\frac{d}{dt}\int_X F\, d\nu_t = \int_X \langle \nabla F, Z_t\rangle_H d\nu_t.$$

In the same way as in the proof of Proposition 2.4, there exists a full measure subset $\Omega \subset]0,1[$ such that for $t \in \Omega$,

$$\int_X \langle \nabla F, W_t - Z_t\rangle_H \, d\nu_t = 0, \quad F \in \mathrm{Cylin}(X).$$

It follows that there exists $\eta_t \in L^2(X, H, \nu_t)$ orthogonal to all ∇F such that $W_t = Z_t + \eta_t$. Then

$$\int_X |\xi|_H^2 d\nu_0 = \int_X |W_t|_H^2 d\nu_t = \int_X |Z_t|_H^2 d\nu_t + \int_X |\eta_t|_H^2 d\nu_t.$$

From this equality, we see that $t \to \int_X |\eta_t|_H^2 \, d\nu_t$ is measurable; integrating the two sides over $[0,1]$, we get

$$W_2^2(\nu_0, \nu_1) = \int_0^1\int_X |Z_t|_H^2 d\nu_t dt + \int_0^1\int_X |\eta_t|_H^2 d\nu_t \, dt.$$

But by (2.22), we deduce that $\int_0^1\int_X |\eta_t|_H^2 d\nu_t \, dt = 0$. Therefore for a.e. $t \in]0,1[$, $\eta_t = 0$ for ν_t-a.e. It follows that $W_t = Z_t \in T_{\nu_t}$. ∎

3. Gradient flow associated to the entropy functional

Let $\nabla F \in \mathcal{E}$. Let $(U_t)_{t\in\mathbf{R}}$ be the quasi-invariant flow associated to ∇F.

Proposition 3.1 *Let $\nu_0 \in \mathcal{P}^*(X)$ be given and denote $\nu_t = (U_t)_*\nu_0$. Then*

$$\frac{d}{dt}|_{t=0}\mathrm{Ent}(\nu_t) = \int_X LF\, d\nu_0, \tag{3.1}$$

where $LF = \delta(\nabla F)$.

Proof. By expression (2.5), LF admits the expression

$$LF = -\sum_{i,j=1}^N (\partial_j\partial_i f)\left\langle e_j, e_i\right\rangle_H + \sum_{i=1}^N (\partial_i f)e_i(x).$$

Note that $x \to LF(x)$ is a continuous function and for a small $\varepsilon_0 > 0$,

$$\int_X e^{2\varepsilon_0|LF|^2} d\mu < +\infty.$$

Set $u_t = \frac{1}{t}\int_0^t (LF)(U_{t-s})ds$. By Jensen inequality,

$$\begin{aligned}
\int_X e^{\varepsilon_0|u_t|^2} d\mu &\le \int_X \Big(\frac{1}{t}\int_0^t e^{\varepsilon_0|(LF)|^2(U_{t-s})}\, ds\Big) d\mu \\
&= \frac{1}{t}\int_0^t \Big(\int_X e^{\varepsilon_0|LF|^2}\cdot K_{t-s} d\mu\Big) ds \\
&\le \Big(\int_X e^{2\varepsilon_0|LF|^2} d\mu\Big)^{1/2}\Big(\int_X e^{4|LF|} d\mu\Big)^{1/2}
\end{aligned}$$

where we used (2.7) for estimating $||K_t||_{L^2(\mu)}$. By Young inequality,

$$\int_X |u_t|^2\rho_0 d\mu \le \int_X e^{\varepsilon_0|u_t|^2} d\mu + \mathrm{Ent}(\rho_0/\varepsilon_0).$$

Therefore $\sup_{0<t\le 1}\Big(\int_X |u_t|^2\rho_0 d\mu\Big) < +\infty$. Now remarking that

$$\frac{1}{t}\log K_t(U_t) = u_t \to LF \text{ as } t \to 0,$$

and using (2.8), we get (3.1). ∎

Definition 3.2 *Let $Z = \nabla F \in \mathcal{E}$, we denote $(\partial_Z \mathrm{Ent})(\nu_0) = \dfrac{d}{dt}|_{t=0}\mathrm{Ent}(\nu_t)$.*

Corollary 3.3 *Let $\rho_0 \ge \varepsilon > 0$ be given in the form (2.3) but with $f \in C_b^1$. Then there exists a unique $v \in T_{\nu_0}$ such that*

$$(\partial_Z \mathrm{Ent})(\nu_0) = \left\langle v, Z\right\rangle_{T_{\nu_0}}, \quad \text{for all } Z = \nabla F \in \mathcal{E}. \tag{3.2}$$

Proof. Rewrite (3.1) in the form

$$(\partial_Z \mathrm{Ent})(\nu_0) = \int_X \delta(\nabla F)\rho_0\, d\mu = \int_X \langle \nabla F, \nabla \rho_0 \rangle_H d\mu = \int_X \langle \nabla F, v \rangle_H \, d\nu_0,$$

with $v = \nabla \log \rho_0$. Take a sequence of $F_n \in \mathrm{Cylin}(X)$ such that $F_n \to \log \rho_0$ in $\mathbf{D}_1^2(X)$. Then

$$\int_X |\nabla F_n - \nabla \log \rho_0|_H^2 \, \rho_0 d\mu \le ||\rho_0||_\infty \int_X |\nabla F_n - \nabla \log \rho_0|_H^2 \, d\mu \to 0,$$

as $n \to +\infty$. It follows that $\nabla \log \rho_0 \in T_{\nu_0}$. ∎

Definition 3.4 *We will say that the gradient* $\nabla\mathrm{Ent}$ *exists at* $\nu_0 \in \mathcal{P}^*(X)$*, if there exists* $v \in T_{\nu_0}$ *such that for all* $Z = \nabla F \in \mathcal{E}$*,*

$$\langle v, Z \rangle_{T_{\nu_0}} = (\partial_Z \mathrm{Ent})(\nu_0). \tag{3.3}$$

and we denote v *by* $(\nabla\mathrm{Ent})(\nu_0)$*.*

The Corollary 3.3 says that the gradient $(\nabla\mathrm{Ent})(\nu_0)$ exists for a good measure ν_0. The following result plays an important role for our understanding of the gradient flow associated to the entropy functional.

Proposition 3.5 *Fix* $\nu_0 \in \mathcal{P}^*(X)$*. Then for any* $\eta > 0$*, there exists a unique* $\hat\nu \in \mathcal{P}^*(X)$ *such that*

$$\frac{1}{2} W_2^2(\nu_0, \hat\nu) + \eta \mathrm{Ent}(\hat\nu) = \inf\Big\{ \frac{1}{2} W_2^2(\nu_0, \nu) + \eta\mathrm{Ent}(\nu);\ \nu \in \mathcal{P}^*(X) \Big\}. \tag{3.4}$$

Moreover the gradient $(\nabla\mathrm{Ent})(\hat\nu)$ *at* $\hat\nu$ *exists.*

Proof. By Corollary 1.2 and the fact that $\nu \to \frac{1}{2}W_2^2(\nu_0,\nu) + \eta\mathrm{Ent}(\nu)$ is semi-lower continuous with respect to the weak convergence, such a $\hat\nu$ does exist. The uniqueness comes from the strict convexity of the entropy functional.

Now let $Z = \nabla F \in \mathcal{E}$ and $(U_t)_{t\in\mathbf{R}}$ be the associated quasi-invariant flow of X. Let $\pi \in \mathcal{C}(\nu_0, \hat\nu)$ be the optimal coupling plan. We define $\pi_t \in \mathcal{C}(\nu_0, (U_t)_*\hat\nu)$ by

$$\int_{X\times X} \psi(x,y)\pi_t(dx,dy) = \int_{X\times X} \psi(x, U_t(y))\pi(dx,dy).$$

Then we have

$$W_2^2(\nu_0, (U_t)_*\hat\nu) - W_2^2(\nu_0,\hat\nu) \le \int_{X\times X} \Big\{ |x - U_t(y)|_H^2 - |x-y|_H^2 \Big\} \pi(dx,dy).$$

It follows that

$$\overline{\lim}_{t\downarrow 0} \frac{1}{2t}\Big[W_2^2(\nu_0, (U_t)_*\hat\nu) - W_2^2(\nu_0,\hat\nu)\Big] \le -\int_{X\times X} \langle Z(y), x-y \rangle_H \pi(dx,dy). \tag{3.5}$$

By construction of $\hat\nu$, for $t > 0$,

$$\frac{\eta}{t}\Big[\mathrm{Ent}((U_t)_*\hat\nu) - \mathrm{Ent}(\hat\nu)\Big] + \frac{1}{2t}\Big[W_2^2(\nu_0, (U_t)_*\hat\nu) - W_2^2(\nu_0,\hat\nu)\Big] \ge 0. \tag{3.6}$$

By Proposition 3.1, as $t \downarrow 0$, the first term in (3.6) tends to $(\partial_Z \mathrm{Ent})(\hat{\nu})$. Combining with (3.5), we get

$$\eta(\partial_Z \mathrm{Ent})(\hat{\nu}) - \int_{X\times X} \big\langle Z(y), x-y\big\rangle_H \pi(dx,dy) \geq 0.$$

Changing Z into $-Z$, we get another inequality, so that

$$\eta(\partial_Z \mathrm{Ent})(\hat{\nu}) = \int_{X\times X} \big\langle Z(y), x-y\big\rangle_H \pi(dx,dy). \tag{3.7}$$

Now by Theorem FU, there exists $\xi : X \to H$ such that $T_1 = I + \xi$ pushes ν_0 forward to $\hat{\nu}$ and $W_2^2(\nu_0, \hat{\nu}) = \int_X |\xi|_H^2 \, d\nu_0$. Rewriting (3.7), we get

$$(\partial_Z \mathrm{Ent})(\hat{\nu}) = \frac{1}{\eta}\int_X \big\langle Z(T_1), -\xi\big\rangle_H \, d\nu_0 = -\int_X \big\langle Z, \xi(T_1^{-1})/\eta\big\rangle_H d\hat{\nu}. \tag{3.8}$$

Note that $\int_X |\xi(T_1^{-1})|_H^2 d\hat{\nu} = \int_X |\xi|_H^2 d\nu_0 < +\infty$; So the gradient $(\nabla \mathrm{Ent})(\hat{\nu}) \in T_{\hat{\nu}}$ exists, which is the orthogonal projection of $-\xi(T_1^{-1})/\eta$ on $T_{\hat{\nu}}$. ∎

Denote by $\mathrm{Dom}(\nabla \mathrm{Ent})$ the set of $\nu \in \mathcal{P}^*(X)$ such that $(\nabla \mathrm{Ent})(\nu) \in T_\nu$ exists. In what follows, we will develop De Giorgi's "minimizing movement" approximation scheme, avoiding the use of the space $\mathcal{P}_2(\mathbf{R}^d)$ done in [AS]

We denote by $\nu^{(1)}$ the element $\hat{\nu}$ obtained in Proposition 3.5. By induction, define step by step $\nu^{(n)}$ which realizes the minimum of

$$\nu \mapsto \frac{1}{2} W_2^2(\nu^{(n-1)}, \nu) + \eta \mathrm{Ent}(\nu).$$

So we get a sequence of probability measures $\{\nu^{(n)};\ n \geq 0\}$ with $\nu^{(0)} = \nu_0$. Let N be an integer such that $N\eta \leq 1$. Define

$$\nu_\eta(t, dx) = \sum_{k=1}^{N+1} \nu^{(k)}(dx) \mathbf{1}_{](k-1)\eta, k\eta]}(t). \tag{3.9}$$

By Proposition 3.5, for $t > 0$, $\nu_\eta(t, \cdot) \in \mathrm{Dom}(\nabla \mathrm{Ent})$.

Proposition 3.6 *The family of measures $\{\nu_\eta(t,dx)dt;\ \eta > 0\}$ over $[0,1] \times X$ is tight.*

Proof. By construction of $\{\nu^{(k)}; k \geq 1\}$, we have

$$\frac{1}{2} W_2^2(\nu^{(k-1)}, \nu^{(k)}) + \eta \mathrm{Ent}(\nu^{(k)}) \leq \eta \mathrm{Ent}(\nu^{(k-1)}).$$

For any $1 \leq q \leq N$, summing the above inequality from $k = 1$ to q gives

$$\frac{1}{2} \sum_{k=1}^{q} W_2^2(\nu^{(k-1)}, \nu^{(k)}) + \eta \mathrm{Ent}(\nu^{(q)}) \leq \eta \mathrm{Ent}(\nu^{(0)}). \tag{3.10}$$

But for each $1 \leq q \leq N$, $W_2^2(\nu^{(0)}, \nu^{(q)}) \leq N \sum_{k=1}^N W_2^2(\nu^{(k-1)}, \nu^{(k)}) \leq 2N\eta \mathrm{Ent}(\nu^{(0)})$. It follows that

$$W_2^2(\nu^{(0)}, \nu^{(q)}) + \mathrm{Ent}(\nu^{(q)}) \leq (2N+1)\eta \mathrm{Ent}(\nu^{(0)}) \leq 3\mathrm{Ent}(\nu^{(0)}).$$

By Corollary 1.2, for any $\varepsilon > 0$, there exists a compact $K \subset X$ such that $\nu^{(q)}(K^c) \leq \varepsilon$. Then

$$\int_{[0,1]\times K^c} \nu_\eta(t,dx)dt \leq \sum_{q=1}^{N+1} \eta\nu^{(q)}(K^c) \leq N\eta\varepsilon \leq \varepsilon,$$

the result follows. ∎

By Prokhorov theorem, there is a sequence $\eta \downarrow 0$ such that $\nu_\eta(t,dx)dt$ converges weakly to $\nu(dt,dx)$. Set $\nu^{(k)}(dx) = \rho^{(k)}(x)d\mu(x)$. Then

$$\nu_\eta(t,dx)dt = \Big(\sum_{k=1}^{N+1} \rho^{(k)}\mathbf{1}_{](k-1)\eta,k\eta]}(t)\Big)\, d\mu(x)dt = \rho_\eta(x,t)d\mu(x)dt.$$

We have

$$\begin{aligned}&\int_{[0,1]\times X} \rho_\eta(x,t)\log\rho_\eta(x,t)d\mu(x)dt\\ &= \sum_{k=1}^{N+1}\int_{(k-1)\eta}^{k\eta\wedge 1}\Big(\int_X \rho^{(k)}\log\rho^{(k)}d\mu\Big)dt \leq \sum_{k=1}^{N+1}\eta\mathrm{Ent}(\nu^{(k)}),\end{aligned}$$

which is less than, again by (3.10), $\sum_{k=0}^{N}\eta\mathrm{Ent}(\nu^{(0)}) \leq \mathrm{Ent}(\nu^{(0)}) < +\infty$. Therefore $\nu(dx,dt)$ admits a density with respect to $d\mu dt$: $\nu(dx,dt) = \rho(x,t)\, d\mu(x)dt$, with

$$\int_{[0,1]\times X} \rho(x,t)\log\rho(x,t)\, d\mu(x)dt \leq \mathrm{Ent}(\nu^{(0)}). \tag{3.11}$$

It follows that for a.e. $t_0 \in [0,1]$, $\mathrm{Ent}(\rho(t_0,\cdot)) < +\infty$. Now we denote:

$$\nu_t(dx) = \rho(x,t)d\mu(x). \tag{3.12}$$

Then for a.e. $t \in [0,1]$, $\nu_t \in \mathcal{P}^*(X)$.

Theorem 3.7 *The curve $\{\nu_t; t\in[0,1]\}$ solves the following Fokker-Planck equation:*

$$-\int_{[0,1]\times X}\alpha'(t)Fd\nu_t dt + \int_{[0,1]\times X}\alpha(t)\, LF\, d\nu_t dt = \alpha(0)\int_X Fd\nu_0, \tag{3.13}$$

for all $\alpha \in C_c^\infty([0,1[), F \in \mathrm{Cylin}(X)$.

Proof. The proof is similar to [JKO], but for the reader's convenience and the difference with finite dimensional spaces that we emphasized in the introduction, we will give a full proof. We have

$$\begin{aligned}&\int_{[0,1]\times X}\alpha'(t)F(x)\nu_\eta(t,dx)dt\\ &= \sum_{k=1}^{N+1}(\alpha(k\eta)-\alpha((k-1)\eta))\int_X F(x)\rho^{(k)}(x)d\mu(x)\\ &= \sum_{k=1}^{N}\alpha(k\eta)\Big[\int_X F(x)(\rho^{(k)}(x)-\rho^{(k+1)}(x))\, d\mu(x)\Big] - \alpha(0)\int_X Fd\nu^{(1)}.\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\int_{[0,1]\times X} \alpha(t) LF(x)\nu_\eta(t,dx)dt\\
&= \sum_{k=1}^{N+1} \Big(\int_{(k-1)\eta}^{k\eta} \alpha(t)dt\Big) \int_X LF(x)\rho^{(k)}d\mu(x)\\
&= \sum_{k=0}^{N} \Big(\frac{1}{\eta}\int_{k\eta}^{(k+1)\eta} \alpha(t)dt\Big)\cdot \eta \int_X LF(x)\rho^{(k+1)}d\mu(x).
\end{aligned}$$

Let $\pi^{(k)} \in \mathcal{C}(\nu^{(k)}, \nu^{(k+1)})$ be the optimal coupling plan and set

$$I_k = \int_X F(x)(\rho^{(k)}(x) - \rho^{(k+1)}(x))d\mu(x) - \int_{X\times X} \big\langle x-y, (\nabla F)(y)\big\rangle_H \pi^{(k)}(dx,dy).$$

Then

$$I_k = \int_X \Big(F(x) - F(y) - \big\langle x-y, (\nabla F)(y)\big\rangle_H\Big)\, \pi^{(k)}(dx,dy).$$

But

$$\Big| F(x) - F(y) - \big\langle x-y, (\nabla F)(y)\big\rangle_H \Big| \leq C\,|x-y|_H^2,$$

where C is a constant governing $\frac{1}{2}|\nabla^2 F|_{H\otimes H}$. It follows that $|I_k| \leq C\, W_2^2(\nu^{(k)}, \nu^{(k-1)})$. By (3.7) and (3.1),

$$\int_{X\times X} \big\langle \nabla F(y), x-y\big\rangle_H \pi^{(k)}(dx,dy) = \eta\,(\partial_Z \mathrm{Ent})(\nu^{(k+1)}) = \eta \int_X LF d\nu^{(k+1)}.$$

Therefore, noting $\beta_k = \alpha(k\eta) - \frac{1}{\eta}\int_{k\eta}^{(k+1)\eta} \alpha(t)dt$,

$$\begin{aligned}
&\int_{[0,1]\times X} \alpha'(t)F(x)\nu_\eta(t,dx)dt - \int_{[0,1]\times X} \alpha(t)LF(x)\nu_\eta(t,dx)dt\\
&= \sum_{k=1}^{n} \alpha(k\eta) I_k + \sum_{k=1}^{N} \beta_k \int_{X\times X} \big\langle \nabla F(x), x-y\big\rangle_H \pi^{(k)}(dx,dy) \qquad (3.14)\\
&\quad - \alpha(0)\int_X F d\nu^{(1)} - \Big(\int_0^\eta \alpha(t)dt\Big)\cdot \int_X LF d\nu^{(1)}.
\end{aligned}$$

The first term on the right hand of (3.14) is dominated, according to (3.10), by

$$C\,||\alpha||_\infty \sum_{k=1}^{N} W_2^2(\nu^{(k)}, \nu^{(k+1)}) \leq \eta\, C\,||\alpha||_\infty \mathrm{Ent}(\nu_0) \to 0 \text{ as } \eta \to 0;$$

The second term is dominated by

$$\begin{aligned}
&||\nabla F||_{L^\infty} ||\alpha'||_\infty\, \eta \sum_{k=1}^{n} \int_{X\times X} |x-y|_H \pi^{(k)}(dx,dy)\\
&\leq ||\nabla F||_{L^\infty} ||\alpha'||_\infty\, \eta\,\sqrt{N} \Big(\sum_{k=1}^{N} W_2^2(\nu^{(k)}, \nu^{(k+1)})\Big)^{1/2}\\
&\leq \sqrt{\eta} ||\nabla F||_{L^\infty} ||\alpha'||_\infty\, \sqrt{\mathrm{Ent}(\nu_0)} \to 0 \text{ as } \eta \to 0
\end{aligned}$$

Note that $W_2^2(\nu_0, \nu^{(1)}) \leq \eta \text{Ent}(\nu_0) \to 0$ as $\eta \to 0$. By Proposition 3.6, as $\eta \to 0$, the first term on the left hand of (3.14) tends to $\int_{[0,1]\times X} \alpha'(t) F(x) d\nu_t dt$. Since LF is not bounded, for the convergence of the second term, we have to use the cut-off function. By the expression of LF, $LF = G_1 + G_2$, where G_1 is a bounded continuous function and $|G_2(x)| \leq C||x||_K$ with $||x||_K^2 = \sum_{i=1}^K e_i^2(x)$. Let $\chi_R \in C_b(\mathbf{R})$ be a cut-off function such that $0 \leq \chi_R \leq 1$ and $\chi_R = 1$ over $[0, R]$ and $\chi_R = 0$ over $[2R, +\infty[$. We have

$$\begin{aligned}
&\int_{[0,1]\times X} \alpha(t) G_2 \big(1 - \chi_R(\sum_{i=1}^K e_i^2(x)\big)\nu_\eta(t, dx) dt \\
&= \sum_{k=1}^{N+1} \Big(\int_{(k-1)\eta}^{k\eta} \alpha(t) dt\Big) \cdot \int_X G_2(x) \big(1 - \chi_R(\sum_{i=1}^K e_i^2(x)\big) \rho^{(k)} d\mu. \\
&\leq C\,||\alpha||_\infty \eta \sum_{k=1}^{N+1} \int_{\{||x||_K^2 \geq R\}} ||x||_K \rho^{(k)} d\mu.
\end{aligned}$$

But

$$\begin{aligned}
&\int_{\{||x||_K^2 \geq R\}} ||x||_K \rho^{(k)} d\mu \leq \frac{1}{\sqrt{R}} \int_X ||x||_K^2 \rho^{(k)} d\mu \\
&\leq C\,||\alpha||_\infty \frac{1}{\sqrt{R}} \Big(\int_X e^{\varepsilon_0 ||x||_K^2} d\mu + \frac{1}{\varepsilon_0} \text{Ent}(\nu^{(k)}) + \frac{1}{\varepsilon_0} \log \frac{1}{\varepsilon_0}\Big).
\end{aligned}$$

Note that $\text{Ent}(\nu^{(k)}) \leq \text{Ent}(\nu^{(0)})$. Then the term $\int_{[0,1]\times X} \alpha(t) G_2 \big(1 - \chi_R(\sum_{i=1}^K e_i^2(x)\big)\nu_\eta(t, dx) dt$ can be arbitrarily small (independent of $\eta > 0$) as R is big enough. So the second term on the left hand of (3.14) tends to $\int_{[0,1]\times X} \alpha(t)\, LF\, d\nu_t dt$, as $\eta \to 0$. The proof is completed. ∎

Remark: The Fokker-Planck equations and related topics on a Hilbert space were studied recently in [ASZ].

We will prove the existence of the derivative process $\dfrac{d^o \nu_t}{dt}$ in the sense of Otto-Ambrosio-Savaré of $(\nu_t)_{t\in[0,1]}$ (see Definition 2.5). Define

$$Z_\eta(x, t) = \sum_{k=1}^{N+1} Z^{(k)} \mathbf{1}_{](k-1)\eta, k\eta]}(t), \quad Z^{(k)} = (\nabla \text{Ent})(\nu^{(k)}). \tag{3.15}$$

Denote by $T^{(k)} = I + \xi_k$ which pushes $\nu^{(k-1)}$ forward $\nu^{(k)}$. We have, according to (3.8)

$$\begin{aligned}
&\int_0^1 \int_X |Z_\eta(x, t)|_H^2 \nu_\eta(t, dx) dt \leq \sum_{k=1}^{N+1} \eta \int_X |Z^{(k)}|_H^2 d\nu^{(k)} \\
&\leq \eta \sum_{k=1}^{N+1} \int_X \frac{1}{\eta^2} |\xi_k((T^{(k)})^{-1})|_H^2 d\nu^{(k)} = \frac{1}{\eta} \sum_{k=1}^{N+1} W_2^2(\nu^{(k-1)}, \nu^{(k)}) \leq 2\text{Ent}(\nu^{(0)}).
\end{aligned} \tag{3.16}$$

Lemma 3.8 *There exists a sequence $\eta \downarrow 0$ and $Z \in L^2(X, H, P_\nu)$ such that*

$$\lim_{\eta \to 0} \int_0^1 \int_X \alpha(t) \big\langle \nabla F(x), Z_\eta(x, t)\big\rangle_H \nu_\eta(t, dx) dt = \int_0^1 \int_X \alpha(t) \big\langle \nabla F(x), Z(x, t)\big\rangle_H \nu_t(dx) dt, \tag{3.17}$$

for any $\alpha \in C_c^\infty([0,1[), F \in \mathrm{Cylin}(X)$.

Proof. Define a probability measure on $[0,1]\times X\times X$ by

$$\int_{[0,1]\times X^2} \psi(t,x,y)d\Gamma_\eta(t,x,y) = \int_{[0,1]\times X} \psi(t,x,Z_\eta(t,x))\nu_\eta(t,dx)dt. \tag{3.18}$$

Let $\pi^{1,2}$ be the projection $(t,x,y)\to(t,x)$ and π^3 the projection $(t,x,y)\to y$. Then

$$(\pi^{1,2})_*\Gamma_\eta = P_{\nu_\eta}, \quad (\pi^3)_*\Gamma_\eta = (Z_\eta)_*(P_{\nu_\eta}).$$

Note that $(\pi^3)_*\Gamma_\eta$ is a measure on X, supported by H. Recall that $B_H(R)=\{x\in X; |x|_H\le R\}$ is a compact subset of X. We have

$$\begin{aligned}\big[(\pi^3)_*\Gamma_\eta\big](B_H(R)^c) &= \int_{[0,1]\times X} \mathbf{1}_{B_H(R)^c}(Z_\eta(t,x))\nu_\eta(t,dx)dt\\ &\le \frac{1}{R^2}\int_{[0,1]\times X} |Z_\eta(t,x)|_H^2\nu_\eta(t,dx)dt \le \frac{2}{R^2}\mathrm{Ent}(\nu_0),\end{aligned}$$

this last inequality was deduced from (3.16). It follows that $\{(\pi^3)_*\Gamma_\eta, \eta>0\}$ is tight. Combining with Proposition 3.6, the family $\{\Gamma_\eta,\ \eta>0\}$ is tight. Up to a sequence, we get the weak convergence of

$$(\pi^3)_*\Gamma_\eta \to w(dx), \quad \Gamma_\eta\to\Gamma.$$

We have

$$(\pi^{1,2})_*\Gamma = \rho(t,x)d\mu dt, \quad (\pi^3)_*\Gamma = w(dx).$$

By semi-lower continuity of $x\to|x|_H$, we have

$$\int_X |x|_H^2 w(dx) \le \underline{\lim}_{\eta\to 0}\int_{[0,1]\times X} |Z_\eta(t,x)|_H^2\nu_\eta(t,dx)dt \le 2\mathrm{Ent}(\nu_0). \tag{3.19}$$

Therefore the measure w is supported by H. Let $\Gamma(dy|\pi^{1,2}=(t,x))$ be the conditional probability given $\pi^{1,2}=(t,x)$. By (3.19),

$$\int_{[0,1]\times X}\Big(\int_X |y|_H^2\Gamma(dy|\pi^{1,2}=(t,x))\Big)\rho(t,x)d\mu(x)dt < +\infty.$$

Then for a.e. $(t,x)\in[0,1]\times X$, $y\to y$ is Bochner integrable with respect to $\Gamma(dy|\pi^{1,2}=(t,x))$. Define

$$Z(t,x) = \int_X y\,\Gamma(dy|\pi^{1,2}=(t,x)). \tag{3.20}$$

We have

$$\begin{aligned}&\int_{[0,1]\times X} |Z(t,x)|_H^2\rho(t,x)d\mu(x)dt\\ &\le \int_{[0,1]\times X}\Big(\int |y|_H^2\Gamma(dy|\pi^{1,2}=(t,x))\Big)\rho(t,x)d\mu(x)dt\\ &= \int_{[0,1]\times X^2} |y|_H^2\, d\Gamma(t,x,y) = \int_X |y|_H^2 w(dy) < +\infty.\end{aligned} \tag{3.21}$$

Now for $\alpha \in C_c^\infty([0,1[)$ and $F \in \mathrm{Cylin}(X)$. By expression (2.4),

$$(t,x,y) \to \alpha(t)\big\langle \nabla F(x), y\big\rangle_H = \alpha(t)\sum_{i=1}^K (\partial_i f)\, e_i(y)$$

is continuous from $[0,1]\times X \times X$ to $\mathbf{R}$. Let $R>0$, consider

$$\psi_R(t,x,y) = \alpha(t)\big\langle \nabla F(x), y\big\rangle_H \cdot \chi_R\big(\sum_{i=1}^K e_i(y)^2\big),$$

where $\chi_R \in C_b(\mathbf{R})$ is the cut-off function considered in the proof of Theorem 3.7. Then $(t,x,y)\to \psi_R(t,x,y)$ is a bounded continuous function; therefore

$$\int \psi_R(t,x,y)d\Gamma(t,x,y) = \lim_{\eta\to 0}\int \psi_R(t,x,y)d\Gamma_\eta(t,x,y).$$

Since

$$\begin{aligned}
&\int |\alpha(t)\big\langle \nabla F(x), y\big\rangle_H|\Big[1-\chi_R\big(\sum_{i=1}^K e_i(y)^2\big)\Big]d\Gamma_\eta(t,x,y)\\
&\leq ||\alpha||_\infty ||\nabla F||_\infty \int |Z_\eta|_H\Big[1-\chi_R\big(\sum_{i=1}^K \big\langle e_i, Z_\eta(t,x)\big\rangle^2\big)\Big]\nu_\eta(t,dx)dt\\
&\leq ||\alpha||_\infty ||\nabla F||_\infty \int_{\sum_{i=1}^K \langle e_i, Z_\eta(t,x)\rangle^2 \geq R} |Z_\eta|_H \nu_\eta(t,dx)dt\\
&\leq \frac{||\alpha||_\infty ||\nabla F||_\infty}{\sqrt{R}}\int |Z_\eta(t,x)|_H^2 \nu_\eta(t,dx)dt \leq \frac{2||\alpha||_\infty ||\nabla F||_\infty}{\sqrt{R}}\mathrm{Ent}(\nu_0),
\end{aligned}$$

which is arbitrarily small as R is big enough. Hence

$$\int \alpha(t)\big\langle \nabla F(x), y\big\rangle_H d\Gamma(t,x,y) = \lim_{\eta\to 0}\int \alpha(t)\big\langle \nabla F(x), y\big\rangle_H d\Gamma_\eta(t,x,y),$$

or (3.17) holds. ∎

Proposition 3.9 $\{\nu_t; t\in[0,1]\}$ *and* $Z(t,x)$ *are linked by the following continuity equation*

$$\int_{[0,1]\times X} \alpha(t)\big\langle \nabla F(x), Z(t,x)\big\rangle_H d\nu_t(x)dt + \int_{[0,1]\times X} \alpha'(t)F(x)d\nu_t(x)dt = 0, \tag{3.22}$$

for all $F \in \mathrm{Cylin}(X)$ *and* $\alpha \in C_c^\infty(]0,1[)$.

Proof. Let $I_\eta^1 = \int_{[0,1]\times X}\alpha(t)\big\langle \nabla F(x), Z_\eta(t,x)\big\rangle_H \nu_\eta(t,dx)dt$. Then I_η^1 admits the expression

$$I_\eta^1 = \sum_{k=1}^{N+1}\Big(\frac{1}{\eta}\int_{(k-1)\eta}^{k\eta}\alpha(t)dt\Big)\cdot \int_X \big\langle \nabla F(x+\xi_k), \xi_k\big\rangle_H d\nu^{(k-1)}.$$

Changing the index and using the optimal coupling plan $\pi^{(k)} \in C(\nu^{(k)}, \nu^{(k+1)})$, we get

$$I_\eta^1 = \sum_{k=0}^{N}\Big(\frac{1}{\eta}\int_{k\eta}^{(k+1)\eta}\alpha(t)dt\Big)\int_{X\times X}\big\langle \nabla F(y), y-x\big\rangle_H \pi^{(k)}(dx,dy).$$

On the other hand, let $I_\eta^2 = \int_{[0,1]\times X} \alpha'(t)F(x)\nu_\eta(t,dx)dt$. Then I_η^2 admits the expression

$$I_\eta^2 = -\sum_{k=1}^{N} \alpha(k\eta) \int_{X\times X} (F(y) - F(x))\pi^{(k)}(dx, dy).$$

The same quantities appeared already in the proof of Theorem 3.7, we see that $\lim_{\eta\to 0}(I_\eta^1 + I_\eta^2) = 0$. But by Lemma 3.8, I_η^1 tends to $\int_{[0,1]\times X} \alpha(t)\langle\nabla F(x), Z(t,x)\rangle_H d\nu_t(x)dt$, while the term I_η^2 tends to $\int_{[0,1]\times X} \alpha'(t)F(x)d\nu_t(x)dt$. So we get (3.22). ∎

Theorem 3.10 *Let* $(\nu_t)_{t\in[0,1]}$ *be the solution to the Fokker-Planck equation* (3.13)*. Then for a.e.* $t \in [0,1], \nu_t \in \mathrm{Dom}(\nabla\mathrm{Ent})$ *and*

$$\frac{d^o\nu_t}{dt} = -(\nabla\mathrm{Ent})(\nu_t). \tag{3.23}$$

Proof. By (3.13) and (3.22), we have

$$\int_{[0,1]\times X} \alpha(t)\langle\nabla F(x), Z(t,x)\rangle_H d\nu_t(x)dt = -\int_{[0,1]\times X} \alpha(t)LF(x)d\nu_t(x)dt. \tag{3.24}$$

Let V be the vector space generated by $\{\alpha\nabla F;\ \alpha \in C_c^\infty(]0,1[), F \in \mathrm{Cylin}(X)\}$ and $\bar{V}$ the closure of V in $L^2([0,1]\times X, H; P_\nu)$. Let $\hat{Z}$ be the orthogonal projection of Z onto $\bar{V}$. Then for a.e.$t \in]0,1[$, $\hat{Z}_t \in T_{\nu_t}$. By (3.24), there exists a full subset $\Omega_F \subset]0,1[$ such that for $t \in \Omega_F$,

$$\int_X \langle\nabla F(x), \hat{Z}(t,x)\rangle_H d\nu_t(x) = -\int_X LF(x)d\nu_t(x).$$

Again by density arguments, there exists a full measure subset $\Omega \subset]0,1[$ such that for $t \in \Omega$ the above equality holds for all $\nabla F \in \mathcal{E}$. Now by (3.1), the right hand side is equal to $-(\partial_{\nabla F}\mathrm{Ent})(\nu_t)$. Therefore $\nabla\mathrm{Ent}$ exists at ν_t and

$$(\nabla\mathrm{Ent})(\nu_t) = -\hat{Z}_t,$$

this last term was denoted as $\frac{d^o\nu_t}{dt}$; therefore we get (3.23). ∎

Acknowledgment: The first named author is grateful to Professors Anton Thalmaier and Luigi Ambrosio for their invitations of visit to Luxembourg University in September 2006 and to De Giorgi center in November 2006 respectively, from which the author has much benefited. The second named author is supported in part by Creative Research Group Fund of the National Natural Science Foundation of China (No. 10721091) and the 973-Project(No.2006CB805901). The three authors thank the two referees for their useful remarks and suggestions.

References

[AGS] L. Ambrosio, N. Gigli and G. Savaré: *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005.

[AS] L. Ambrosio and G. Savaré: Gradient flows of probability measures: *Handbook of differential equations*, Evolutionary equations, vol. 3, ed. by C.M. Dafermos and E. Feireisl, 2007 Elsevier.
[ASZ] L. Ambrosio, G. Savaré and L. Zambotti: Existence and stability for Fokker-Planck equations with log-concave reference measure, to appear in Prob. Th. Related Fields.
[BB] J.D. Benamou and Y. Brenier: A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.*, **84** (2000), 375-393.
[BGL] S. Bobkov, I. Gentil and M. Ledoux: Hypercontractivity of Hamilton-Jacobi equations, *J. Math. Pure Appl.*, bf 80 (2001), 669-696.
[BK] V. Bogachev and A. Kolesnikov: On the Monge-Ampère equation on Wiener space, *Doklady Mathematics*, **73** (2006), 1-5.
[BKM] V. Bogachev, A. Kolesnikov and K.V. Medvedev: Triangular transformations of measures, *Mat. Sb.*, **196** (2005), 3-30.
[Ch] MuFa Chen: *From Markov chains to non-Equilibrium particle systems*, World Scientific, 2nd edition 2004.
[Cr] A.B. Cruzeiro: Equations différentielles sur l'espace de Wiener et formules de Cameron-Martin non linéaires, J. Funct. Analysis, **54** (1983), 206-227.
[Dr] B. Driver: Integration by parts and quasi-invariance for heat measures on loop groups, J. Funct. Anal. **149** (1997), 470-547.
[FU] D. Feyel and A.S. Üstünel: Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space. Probab. Theory Related Fields **128** (2004), 347-385.
[FU2] D. Feyel and A.S. Üstünel: Monge-Kantorovitch measure transportation, Monge-Ampère equation and the Itô calculus: *Advanced Studies in Pure Math.*, Math. Soc. of Japan. **41** (2004), 32-49.
[FUZ] D. Feyel, A.S. Üstünel and Zakai: The realization of positive random variables via absolutely continuous transformation of measures on Wiener space, *Prob. Survey*, **3** (2006), 170-205.
[Gen] I. Gentil: Inégalités de Sobolev logarithmiques et hypercontractivité en mécanique statistiqueet en EDP. Thèse de Doctorat de l'Université Paul Sabatier, Toulouse, 2001.
[JKO] R. Jordan, D. Kinderlehrer and F. Otto: The variational formulation of the Fokker-Planck equation, J. Math. Anal. **29** (1998), 1-17.
[Kr] P. Krée: Introduction aux théories des distributions en dimension infinie, *Mémoires de la société Mathématique de France*, **46** (1976), 143-162.
[LV] J. Lott and C. Villani: Ricci curvature for metric-measure spaces via optimal transport, Ann of Math.
[Ma] P. Malliavin, *Stochastic Analysis*, vol. 313, Grund. Math. Wissen., Springer, 1997.
[Ot] F. Otto: The geometry of dissipative evolution equations: The porous medium equation, Comm. partial Diff. equations, **26** (2001), 101-174.
[OV] F. Otto and C. Villani: Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal., **173** (2000), 361-400.
[St] K.Th. Sturm: On the geometry of measures spaces, Acta Math. Vol. **196** (2006), 65-131.
[Ta] M. Talagrand, Transportation cost for Gaussian and other product measures, Geom. Funct. Anal. **6**(1996), 587-600.
[Va] L. N. Vasershtein: Markov processes over denumerable products of spaces describing large systems of automata, Problemy Peredaci Informacii bf 5 (1969), 64-72.
[Vi] C. Villani, *Topics in Mass Transportation.* American Mathematical Society, 2003.