

W. Gangbo, J. Haskovec, P. Markowich, J. Sierra

## AN OPTIMAL TRANSPORT APPROACH FOR THE KINETIC BOHMIAN EQUATION

ABSTRACT. We study the existence theory of solutions of the kinetic Bohmian equation, a nonlinear Vlasov-type equation proposed for the phase-space formulation of Bohmian mechanics. Our main idea is to interpret the kinetic Bohmian equation as a Hamiltonian system defined on an appropriate Poisson manifold built on a Wasserstein space. We start by presenting an existence theory for stationary solutions of the kinetic Bohmian equation. Afterwards, we develop an approximative version of our Hamiltonian system in order to study its associated flow. We then prove existence of solutions of our approximative version. Finally, we present some convergence results for the approximative system, the aim being to establish that, in the limit, the approximative solution satisfies the kinetic Bohmian equation in a weak sense.

### §1. INTRODUCTION

In this paper, we study the existence theory of solutions of the kinetic Bohmian equation [15, 16],

$$\partial_t \beta + v \cdot \nabla_x \beta - \nabla_x \left( V - \frac{1}{2} \frac{\Delta_x \sqrt{\varrho}}{\sqrt{\varrho}} \right) \cdot \nabla_v \beta = 0, \quad (1.1)$$

along with the initial value,

$$\beta(t=0, x, v) = \beta_0 \in \mathcal{M}^+(\mathbb{R}^d \times \mathbb{R}^d), \quad (1.2)$$

where  $v, x \in \mathbb{R}^d$ ,  $t \geq 0$ , and  $\mathcal{M}^+(\mathbb{R}^d \times \mathbb{R}^d)$  denotes the set of nonnegative Radon measures defined on *phase space*,  $\mathbb{R}^d \times \mathbb{R}^d$ . Furthermore,  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  is a potential satisfying some regularity assumptions given below, and  $\beta = \beta(t, x, v)$  represents the *generalized* Bohmian measure. Finally,  $\varrho =$

---

*Key words and phrases:* Kinetic equation, Hamiltonian flow, Wasserstein space, Poisson structure, Moreau–Yosida approximation.

The research of W. Gangbo was supported by NSF grant DMS–1160939.

$\varrho(t, x)$  is the position density given by

$$\varrho(t, x) = \int_{\mathbb{R}^d} \beta(t, x, dv).$$

For a comprehensive review of Bohmian mechanics and its role in quantum mechanics, see, e.g., [7, 8].

It was shown in [15, 16] that if the initial condition (1.2) is a mono-kinetic measure, then there exists a connection between the kinetic Bohmian equation and the linear Schrödinger equation that can be used to establish an existence theory for solutions of (1.1). Nevertheless, for the more general situation given by (1.1)–(1.2), such connection is lost. In this case, our analysis relies on interpreting the kinetic Bohmian equation as a Hamiltonian system on a space of probability measures in the following way. Let  $\mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$  stand for the set of Borel probability measures on  $\mathbb{R}^d \times \mathbb{R}^d$  with finite second moments and consider the Hamiltonian  $\mathcal{H} : \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$  given by

$$\begin{aligned} \mathcal{H}(\mu) := & \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} |v|^2 \mu(dx, dv) \\ & + \int_{\mathbb{R}^d \times \mathbb{R}^d} V(x) \mu(dx, dv) + \frac{1}{8} \int_{\mathbb{R}^d} \frac{|\nabla \varrho|^2}{\varrho} \varrho(dx) + \chi_0((\pi_{\#}^1 \mu)_s), \end{aligned}$$

where we have used the Radon–Nikodym decomposition

$$\pi_{\#}^1 \mu = \varrho \mathcal{L}^d + (\pi_{\#}^1 \mu)_s,$$

$\pi^1 : \mathbb{R}^d \times \mathbb{R}^d : (w, z) \rightarrow w$  represents the first projection of  $\mathbb{R}^d \times \mathbb{R}^d$  onto  $\mathbb{R}^d$ , and  $\chi_0 : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \{0, +\infty\}$  assumes the value 0 on null measures and the value  $+\infty$  on probability measures of positive total mass. Formally, at least, if the metric slope of  $\mathcal{H}$  at  $\mu$  is finite, under suitable conditions, the subdifferential of  $\mathcal{H}$  at  $\mu$  is not empty. Its unique element of minimal norm is a Borel vector field,

$$\nabla_{\mu} \mathcal{H} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^d,$$

which is referred to as the Wasserstein gradient of  $\mathcal{H}$  at  $\mu$ .  $\nabla_{\mu} \mathcal{H}$  belongs to the range of the projection map

$$\pi_{\mu} : L^2(\mu) \rightarrow \overline{\nabla C_c^{\infty}(\mathbb{R}^d \times \mathbb{R}^d)}^{L^2(\mu)}$$

and is given by

$$\nabla_{\mu} \mathcal{H}(x, v) = \begin{pmatrix} \nabla_x V(x) - \frac{1}{2} \nabla_x \left( \frac{\Delta_x \sqrt{\varrho}}{\sqrt{\varrho}} \right) (x) \\ v \end{pmatrix}.$$

Using the  $(2d) \times (2d)$  symplectic matrix

$$J = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix},$$

the theory developed in [9] allows us to define a Poisson structure for which  $X_{\mathcal{H}} := \pi_{\mu}(J \nabla_{\mu} \mathcal{H})$  is a Hamiltonian vector field; we have

$$X_{\mathcal{H}}(\mu)(x, v) = \pi_{\mu} \begin{pmatrix} v \\ -\nabla_x V(x) + \frac{1}{2} \nabla_x \left( \frac{\Delta_x \sqrt{\varrho}}{\sqrt{\varrho}} \right) (x) \end{pmatrix}.$$

On the other hand, the path  $t \rightarrow \mu_t \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$  is said to be driven by a velocity vector field,  $\mathbf{v} : (0, 1) \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^d$ , if

$$\partial_t \mu + \nabla \cdot (\mu \mathbf{v}) = 0,$$

in the sense of distributions. According to [2], the path  $t \rightarrow \mu_t$  satisfies the Hamiltonian system (defined in the context of Poisson geometry)

$$\dot{\mu} = X_{\mathcal{H}}(\mu)$$

if  $X_{\mathcal{H}}(\mu)$  is a velocity vector field driving  $t \rightarrow \mu_t$ , namely,

$$\partial_t \mu + \nabla_x \cdot (v \mu) = \nabla_v \cdot \left( \nabla_x V(x) - \frac{1}{2} \nabla_x \left( \frac{\Delta_x \sqrt{\varrho}}{\sqrt{\varrho}} \right) \right),$$

in the sense of distributions. This is exactly (1.1) when  $\mu_t = \beta(t, \cdot, \cdot) \mathcal{L}^{2d}$ .

Therefore, now we can say that one of the main ideas of this paper is to investigate the existence theory of solutions of the kinetic Bohmian equation through the Hamiltonian flow generated by the Hamiltonian vector field  $X_{\mathcal{H}}$ .

To motivate the study of the kinetic Bohmian equation, let us start by reviewing the aforementioned connection with the linear Schrödinger equation,

$$i \partial_t \psi = -\frac{1}{2} \Delta \psi + V \psi, \quad \psi(t=0, \cdot) = \psi_0 \in L^2(\mathbb{R}^d; \mathbb{C}). \quad (1.3)$$

A thorough analysis of this equation can be found in, e.g., [6, 20, 21]. We adopt the normalization of the initial data, i.e.,  $\|\psi_0\|_{L^2} = 1$ . Thus,

$$\|\psi(t)\|_{L^2} = \|\psi_0\|_{L^2} = 1. \quad (1.4)$$

In addition, we assume that  $\psi$  has bounded initial energy. The energy is conserved for all  $t \geq 0$  and is given by

$$E(t) := \frac{1}{2} \int_{\mathbb{R}^d} |\nabla \psi(t, x)|^2 dx + \int_{\mathbb{R}^d} V(x) |\psi(t, x)|^2 dx = E(0).$$

Note that the Schrödinger equation (1.3) has a reduced Planck constant equal to one ( $\hbar = 1$ ).

As a consequence of (1.4), one can define real-valued probability densities from  $\psi(t, x) \in \mathbb{C}$ . These probability densities can be used to compute expectation values of physical observables. In particular, we have the *position* and *current densities* given by

$$\varrho = \varrho(t, x) = |\psi(t, x)|^2, \quad J = J(t, x) = \text{Im}(\overline{\psi}(t, x) \nabla \psi(t, x)). \quad (1.5)$$

**Definition 1.1.** (*Bohmian measure* [15, 16]). For  $\psi \in H^1(\mathbb{R}^d)$ , with associated densities  $\varrho, J$  given by (1.5), the Bohmian measure  $\beta = \beta[\psi] \in \mathcal{M}^+(\mathbb{R}^d \times \mathbb{R}^d)$  is defined by

$$\langle \beta, \varphi \rangle := \int_{\mathbb{R}^d} \varrho(x) \varphi\left(x, \frac{J(x)}{\varrho(x)}\right) dx, \quad \forall \varphi \in C_0(\mathbb{R}^d \times \mathbb{R}^d), \quad (1.6)$$

where  $C_0(\mathbb{R}^d \times \mathbb{R}^d)$  denotes the space of continuous functions vanishing at infinity.

Let

$$\beta_0 = \beta_0(x, v) = \varrho_0(x) \delta(v - u_0(x)), \quad (1.7)$$

where  $\varrho_0 \equiv \varrho(t=0, x)$ ,  $u_0 \equiv u(t=0, x)$ ,  $u = u(t, x) := J/\varrho$ , and  $\delta$  is the delta distribution on  $\mathbb{R}^d$ . It was shown in [15] that if  $\psi(t, x)$  solves the Schrödinger equation (1.3), then the corresponding Bohmian measure (1.6) is the push-forward of (1.7) under the phase space flow

$$\Phi_t : (x, v) \mapsto (X(t, x, v), P(t, x, v)),$$

induced by

$$\begin{cases} \dot{X} = P, \\ \dot{P} = -\nabla V(X) - \nabla V_B(t, X), \end{cases} \quad (1.8)$$

where  $V_B(t, x)$  is the Bohm potential:

$$V_B(t, x) := -\frac{1}{2} \frac{\Delta \sqrt{\varrho(t, x)}}{\sqrt{\varrho(t, x)}}.$$

Note that the specific form of the initial data (1.7) implies that the phase-space flow  $\Phi_t$ , governed by (1.8), is initially projected onto the graph of  $u_0$ , that is,

$$\mathcal{L} := \{(x, v) \in \mathbb{R}^d \times \mathbb{R}^d : v = u_0(x)\}. \quad (1.9)$$

This imposes a big limitation for the application of the theory developed in [15,16]: from the whole phase space, we are restricted to the *Lagrangian submanifold* (1.9) for the initial condition of (1.8).

Furthermore, it was proved in [16] that for  $V \in C_b^1(\mathbb{R}^d; \mathbb{R})$  and  $\psi_0 \in H^3(\mathbb{R}^d)$  with corresponding  $\varrho_0, J_0$  given by (1.5), the Bohmian measure

$$\beta(t, x, v) = \varrho(t, x) \delta(v - u(t, x)),$$

is a weak solution of the kinetic Bohmian equation in  $\mathcal{D}'(\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d)$  and in  $\mathcal{D}'([0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d)$  with initial data (1.7). On the other hand, the uniqueness theory is still an open problem.

As mentioned before, the purpose of this paper is to study the kinetic Bohmian equation with the more general initial data (1.2), which implies that the connection with the Schrödinger equation is lost. Nevertheless, the idea is to use the Wasserstein gradient/Hamiltonian flow techniques to generate rigorous results on (1.1)–(1.2) with the aim of overcoming the limitations mentioned above, in particular, the restriction from the whole phase space to the Lagrangian submanifold (1.9). Moreover, this opens the door for a new interpretation of Bohmian mechanics through optimal transportation.

The remainder of this paper is organized as follows. In Section 2, we present the basic theory and notation used throughout our analysis. In Section 3, we study the existence of stationary solutions of the kinetic Bohmian equation. Sections 4, 5, and 6 are devoted to the development of an approximative version of the kinetic Bohmian equation; in particular, we prove existence of solutions of this approximative version in Section 6. In Section 7, we present some convergence results for the approximative model developed in Sections 4, 5, and 6. Conclusions are drawn in Section 8.

## §2. PRELIMINARIES

Since most of our work is performed inside the framework of probability measures, we present now the basic concepts and notation for this topic. A comprehensive review of this subject can be found in [19]. Furthermore, the theory of optimal transportation is extensively studied in [3, 23, 24].

A Borel measure on a topological space,  $X$ , is any measure defined on the  $\sigma$ -algebra generated by the open sets of  $X$ . The elements of such  $\sigma$ -algebra are called the Borel sets. Furthermore, a map,  $f : X \rightarrow Y$ , between the topological spaces  $X$  and  $Y$ , is called a Borel map if  $f^{-1}(B)$  is a Borel set for any Borel set  $B \subset Y$ .

Suppose that  $\mu$  and  $\nu$  are nonnegative Borel measures on the topological spaces  $X$  and  $Y$ , respectively. We say that the Borel map  $T : X \rightarrow Y$  transports  $\mu$  into  $\nu$ , denoted by  $T_{\#}\mu = \nu$ , if for every Borel set  $B \subset Y$  we have  $\nu[B] = \mu[T^{-1}(B)]$ ; in this case, we also say the  $\nu$  is the pushforward of  $\mu$  through  $T$ . We shall represent by  $\mathcal{J}(\mu, \nu)$  the set of all Borel maps,  $T$ , satisfying  $T_{\#}\mu = \nu$ .

Let  $\pi^1 : X \times Y \rightarrow X$  be the projection of  $X \times Y$  onto  $X$  and let  $\pi^2 : X \times Y \rightarrow Y$  be the projection of  $X \times Y$  onto  $Y$ . A nonnegative Borel measure,  $\gamma$ , on  $X \times Y$  is said to have marginals  $\mu$  and  $\nu$  if  $\mu = \pi^1_{\#}\gamma$  and  $\nu = \pi^2_{\#}\gamma$ ; in this case,  $\gamma$  is called a transport plan between  $\mu$  and  $\nu$ . The set of all transport plans between  $\mu$  and  $\nu$  is denoted by  $\Gamma(\mu, \nu)$ .

Let  $d \geq 1$  be an integer and let  $D \in \{d, 2d\}$ . The  $D$ -dimensional Lebesgue measure on  $\mathbb{R}^D$  is represented by  $\mathcal{L}^D$ .  $\mathcal{P}(\mathbb{R}^D)$  stands for the set of Borel probability measures on  $\mathbb{R}^D$ . The second moment of  $\mu \in \mathcal{P}(\mathbb{R}^D)$  is defined as

$$M_2(\mu) := \int_{\mathbb{R}^D} |z|^2 d\mu(z).$$

Furthermore,

$$\mathcal{P}_2(\mathbb{R}^D) := \{\mu \in \mathcal{P}(\mathbb{R}^D) : M_2(\mu) < +\infty\}.$$

The subspace of  $\mathcal{P}_2(\mathbb{R}^D)$  of absolutely continuous measures with respect to  $\mathcal{L}^D$  is represented by  $\mathcal{P}_2^r(\mathbb{R}^D)$ .

For  $\mu \in \mathcal{P}_2(\mathbb{R}^D)$ , we denote by  $L^2(\mu)$  the set of Borel vector fields,  $\xi : \mathbb{R}^D \rightarrow \mathbb{R}^D$ , which are  $\mu$ -measurable and satisfy

$$\|\xi\|_{\mu}^2 := \int_{\mathbb{R}^D} |\xi(z)|^2 d\mu(z) < +\infty.$$

$\mathcal{P}_2(\mathbb{R}^D)$  is canonically endowed with the Wasserstein distance,  $W_2$ , defined by

$$W_2^2(\mu, \nu) := \min_{\gamma} \left\{ \int_{\mathbb{R}^D \times \mathbb{R}^D} |x - y|^2 d\gamma(x, y) : \gamma \in \Gamma(\mu, \nu) \right\}. \quad (2.1)$$

Any minimizer in (2.1) is called an optimal transport plan between  $\mu$  and  $\nu$ . The set of all such minimizers is indicated by  $\Gamma_o(\mu, \nu)$ .

Suppose now that  $\mu \in \mathcal{P}_2^r(\mathbb{R}^D)$  and  $\nu \in \mathcal{P}_2(\mathbb{R}^D)$ . Then, there exists a unique minimizer,  $\gamma_o$ , in (2.1) which can be represented as  $\gamma_o = (\mathbf{id} \times T_\mu^\nu)_\# \mu$  for some  $T_\mu^\nu : \mathbb{R}^D \rightarrow \mathbb{R}^D$  that coincides  $\mu$ -a.e. with the gradient of a convex function and satisfies  $T_\mu^\nu \# \mu = \nu$ . Hence,  $T_\mu^\nu$  is the unique minimizer of

$$T \rightarrow \int_{\mathbb{R}^D} |z - T(z)|^2 d\mu(z),$$

over  $\mathcal{J}(\mu, \nu)$ .

$(\mathcal{P}_2(\mathbb{R}^D), W_2)$  is a Polish space, namely, a complete and separable metric space (see Section 7.1 in [23] and Proposition 7.1.5 in [3] for details). On the other hand, it is not locally compact. Nevertheless, bounded sets in  $\mathcal{P}_2(\mathbb{R}^D)$  are sequentially relatively compact with respect to the narrow convergence; a sequence  $(\mu_k)_k \subset \mathcal{P}(\mathbb{R}^D)$  converges narrowly to  $\mu \in \mathcal{P}(\mathbb{R}^D)$  as  $k \rightarrow \infty$  if

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^D} g(z) d\mu_k(z) = \int_{\mathbb{R}^D} g(z) d\mu(z),$$

for every  $g \in C_b^0(\mathbb{R}^D)$ , the space of bounded and continuous functions on  $\mathbb{R}^D$ . Moreover, a sequence  $(\beta_k)_k \subset \mathcal{P}_2(\mathbb{R}^D)$  converges to  $\beta \in \mathcal{P}_2(\mathbb{R}^D)$  if and only if  $(\beta_k)_k$  converges narrowly to  $\beta$  and  $M_2(\beta_k) \rightarrow M_2(\beta)$  as  $k \rightarrow \infty$ .

A particularly important subject for our analysis is the differentiable Riemannian structure of  $\mathcal{P}_2(\mathbb{R}^D)$ , which can be derived from its metric structure. For such derivation, we first have to characterize the absolutely continuous curves  $\mu_t : [a, b] \rightarrow \mathcal{P}_2(\mathbb{R}^D)$ . As proved in Theorem 8.3.1 of [3], if  $\mu_t$  solves the continuity equation

$$\frac{d}{dt} \mu_t + \nabla \cdot (w_t \mu_t) = 0, \quad (2.2)$$

in the sense of distributions in  $(a, b) \times \mathbb{R}^D$  for some time-dependent velocity vector field,  $w_t$ , with  $\|w_t\|_{\mu_t} \in L^1(a, b)$ , then

$$W_2(\mu_s, \mu_t) \leq \int_s^t \|w_\tau\|_{\mu_\tau} d\tau \quad \forall a \leq s < t \leq b. \quad (2.3)$$

Therefore, the map  $t \mapsto \mu_t$  is absolutely continuous from  $[a, b]$  to  $\mathcal{P}_2(\mathbb{R}^D)$ . Conversely, for any absolutely continuous curve,  $t \mapsto \mu_t$ , there exists a unique (up to  $\mathcal{L}^1$ -negligible sets in time) velocity vector field,  $v_t$ , for which the continuity equation (2.2) holds, along with asymptotic equality in (2.3):

$$\lim_{h \rightarrow 0} \frac{1}{|h|} W_2(\mu_{t+h}, \mu_t) = \|v_t\|_{\mu_t} \quad \text{for a.e. } t.$$

Proposition 8.4.5 of [3] shows that this minimality property of  $v_t$  is equivalent to the fact that

$$v_t \in \overline{\{\nabla\varphi : \varphi \in C_c^\infty(\mathbb{R}^D)\}}^{L^2(\mu_t)}.$$

This result leads to the identification of  $v_t$  as the ‘‘tangent’’ velocity vector to  $\mu_t$ . Hence, the tangent space to  $\mathcal{P}_2(\mathbb{R}^D)$  at  $\mu$  is defined as

$$T_\mu \mathcal{P}_2(\mathbb{R}^D) := \overline{\{\nabla\varphi : \varphi \in C_c^\infty(\mathbb{R}^D)\}}^{L^2(\mu)}.$$

Furthermore, using a simple duality argument, it has been proved in Lemma 8.4.2 of [3] that

$$[T_\mu \mathcal{P}_2(\mathbb{R}^D)]^\perp = \{w \in L^2(\mu) : \nabla \cdot (w\mu) = 0\}.$$

The following is a useful characterization of the tangent velocity vector,  $v_t$ , given in Proposition 8.4.6 of [3]:

$$\lim_{h \rightarrow 0} \left( w, \frac{z - w}{h} \right)_\# \gamma_h = (\mathbf{id}, v_t)_\# \mu_t \quad \text{in } \mathcal{P}_2(\mathbb{R}^D \times \mathbb{R}^D),$$

for almost every  $t$  and any  $\gamma_h \in \Gamma_o(\mu_t, \mu_{t+h})$ . In addition, if  $\mu_t \in \mathcal{P}_2^r(\mathbb{R}^D)$ , then the last characterization becomes

$$\frac{t_h - \mathbf{id}}{h} \rightarrow v_t \quad \text{in } L^2(\mu_t; \mathbb{R}^D) \quad \text{as } h \rightarrow 0,$$

where  $t_h$  are the optimal transport maps between  $\mu_t$  and  $\mu_{t+h}$ .

We present now some basic results from convex analysis in  $\mathcal{P}_2(\mathbb{R}^D)$  which are extensively used in the sequel.



Let  $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^D)$  and let  $\gamma \in \Gamma_o(\mu_0, \mu_1)$ . Let  $\pi^1 : \mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbb{R}^D : (w, z) \rightarrow w$  and  $\pi^2 : \mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbb{R}^D : (w, z) \rightarrow z$  be the first and second projections of  $\mathbb{R}^D \times \mathbb{R}^D$  onto  $\mathbb{R}^D$ , respectively. Consider the interpolation between the measures  $\mu_0$  and  $\mu_1$  given by

$$\mu_t = ((1-t)\pi^1 + t\pi^2)_\# \gamma, \quad t \in [0, 1].$$

Theorem 7.2.2 of [3] shows that  $t \mapsto \mu_t$  is a constant speed geodesic in  $\mathcal{P}_2(\mathbb{R}^D)$ , i.e.,  $W_2(\mu_s, \mu_t) = |t-s|W_2(\mu_0, \mu_1)$  for all  $s, t \in [0, 1]$ . In addition, any constant speed geodesic has this representation for a suitable optimal transport plan,  $\gamma$ .

Let  $\phi : \mathcal{P}_2(\mathbb{R}^D) \rightarrow [-\infty, +\infty]$ . We define the effective domain of  $\phi$  as

$$D(\phi) := \{z \in \mathcal{P}_2(\mathbb{R}^D) : -\infty < \phi(z) < +\infty\}.$$

**Definition 2.1.** ( *$\lambda$ -convexity*). Let  $\phi : \mathcal{P}_2(\mathbb{R}^D) \rightarrow [-\infty, +\infty]$  be such that  $D(\phi) \neq \emptyset$  and let  $\lambda \in \mathbb{R}$ . We say that  $\phi$  is  $\lambda$ -convex if for every  $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^D)$  and every  $\gamma \in \Gamma_o(\mu_0, \mu_1)$  we have

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

where  $\mu_t = ((1-t)\pi^1 + t\pi^2)_\# \gamma$ . In particular, 0-convexity corresponds to the so-called displacement convexity.

**Definition 2.2.** Let  $\mathcal{G} : \mathcal{P}_2(\mathbb{R}^D) \rightarrow [-\infty, \infty]$  be such that  $D(\mathcal{G}) \neq \emptyset$  and let  $\mu \in D(\mathcal{G})$ .

- (i) We say that  $\xi$  belongs to the subdifferential of  $\mathcal{G}$  at  $\mu$ , and we write  $\xi \in \partial\mathcal{G}$ , if  $\xi \in L^2(\mu)$  and

$$\begin{aligned} \mathcal{G}(\nu) - \mathcal{G}(\mu) \geq \sup_{\gamma \in \Gamma_o(\mu, \nu)} \int_{\mathbb{R}^D \times \mathbb{R}^D} \xi(w) \cdot (z-w) \gamma(dw, dz) \\ + o(W_2(\mu, \nu)), \quad \forall \nu \in D(\mathcal{G}). \end{aligned} \quad (2.4)$$

The unique element of minimal norm in  $\partial\mathcal{G}(\mu)$  belongs to  $T_\mu\mathcal{P}_2(\mathbb{R}^D)$  and is called the gradient of  $\mathcal{G}$  at  $\mu$ ; it is denoted by  $\nabla_\mu\mathcal{G}(\mu)$ .

- (ii) We say that  $\xi$  belongs to the superdifferential of  $\mathcal{G}$  at  $\mu$ , and we write  $\xi \in \bar{\partial}\mathcal{G}(\mu)$ , if  $-\xi \in \partial(-\mathcal{G})(\mu)$ .
- (iii) We say that  $\mathcal{G}$  is differentiable at  $\mu$  if both  $\partial\mathcal{G}(\mu)$  and  $\bar{\partial}\mathcal{G}(\mu)$  are non empty. In that case (see e.g. [10]) both sets coincide and

$$\partial\mathcal{G}(\mu) \cap T_\mu\mathcal{P}_2(\mathbb{R}^D) = \bar{\partial}\mathcal{G}(\mu) \cap T_\mu\mathcal{P}_2(\mathbb{R}^D) = \{\nabla_\mu\mathcal{G}(\mu)\}.$$

Therefore, there is no ambiguity if we define the gradient of  $\mathcal{G}$  at  $\mu$  as the unique element of minimal norm in  $\bar{\partial}\mathcal{G}(\mu)$ ; we denote it by  $\nabla_\mu\mathcal{G}(\mu)$ .

**Remark 2.3.** Here are some remarks.

- (i) We refer the reader to Remark 3.2 of [10] for property (iii) in Definition 2.2.
- (ii) Due to Proposition 8.5.4 of [3], (2.4) holds for  $\xi$  if and only if it holds for any  $\xi_0 \in L^2(\mu)$  such that  $\xi_0 - \xi$  belongs to the orthogonal complement of  $T_\mu\mathcal{P}_2(\mathbb{R}^D)$  in  $L^2(\mu)$ . Rephrasing, if (2.4) holds for  $\xi_0 \in L^2(\mu)$ , then it holds for  $\xi$  defined as the orthogonal projection of  $\xi_0$  onto  $T_\mu\mathcal{P}_2(\mathbb{R}^D)$ . Hence,

$$\nabla_\mu\Phi(\mu) + \{\xi \in L^2(\mu) \mid \operatorname{div}_\mu(\xi) = 0\} \subset \bar{\partial}\Phi(\mu).$$

- (iii) Define  $\psi(\nu) = 1/2W_2^2(\nu, \varrho)$  for  $\nu \in \mathcal{P}_2(\mathbb{R}^D)$ , where  $\varrho \in \mathcal{P}_2(\mathbb{R}^D)$  is absolutely continuous. The proof of Proposition 10.4.12 [3] reveals that if  $\xi \in \bar{\partial}\psi(\nu)$ , since  $\gamma \in \Gamma_o(\nu, \varrho)$  has a unique element, then  $\pi_\nu(\xi) = \mathbf{id} - \bar{\gamma}$ , where  $\bar{\gamma}$  is the barycentric projection of  $\gamma$ . Hence,

$$\bar{\partial}\psi(\nu) = \mathbf{id} - \bar{\gamma} + \{v \in L^2(\nu) \mid \operatorname{div}_\nu(v) = 0\}.$$

We next list some facts about proper functionals,  $\Phi : \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\infty\}$ , for which there exists a functional,  $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\infty\}$ , such that

$$\Phi(\mu) = \phi(\pi_{\#}^1\mu).$$

If  $\xi = (\xi_1, \xi_2) \in \bar{\partial}\Phi(\mu)$ , then  $\bar{\xi}_1 \in \bar{\partial}\phi(\varrho)$ , where

$$\bar{\xi}_1(x) = \int_{\mathbb{R}^d} \xi_1(x, v) \mu_x(dv)$$

and  $(\mu_x)_{x \in \mathbb{R}^d}$  is the disintegration of  $\mu$  with respect to  $\varrho$ . This result holds under the assumption that  $\bar{\partial}\Phi(\mu) \neq \emptyset$ . Moreover, if  $\Phi$  is bounded below and lower semicontinuous for the narrow convergence, we can then draw some conclusions about the functionals  $\Phi_\tau$  defined in 4.1, the Moreau–Yosida approximations of  $\Phi$ . First,  $\bar{\partial}\Phi_\tau(\mu) \neq \emptyset$  and

$$\Phi_\tau(\mu) = \phi_\tau(\varrho).$$

Second, if we further assume that the domain of  $\phi$  is contained in  $\mathcal{P}_2^r(\mathbb{R}^d)$  and  $\varrho \in \mathcal{P}_2^r(\mathbb{R}^d)$ , then  $\bar{\partial}\Phi_\tau(\mu)$  and  $\bar{\partial}\phi_\tau(\varrho)$  are non empty and their elements

of minimal norm, respectively denoted by  $\nabla_\mu \Phi_\tau(\mu)$  and  $\nabla_\varrho \phi_\tau(\varrho)$ , satisfy

$$\nabla_\mu \Phi_\tau(\mu)(x, v) = \begin{pmatrix} \nabla_\varrho \phi_\tau(\varrho)(x) \\ 0 \end{pmatrix}.$$

This is a subtle statement, since (cf. Remark 2.3 (ii))

$$\nabla_\varrho \phi(\varrho) + \{u \in L^2(\varrho) \mid \operatorname{div}_\varrho(u) = 0\} \subset \bar{\partial}\phi(\varrho)$$

and similarly,

$$\nabla_\mu \Phi(\mu) + \{\xi \in L^2(\mu) \mid \operatorname{div}_\mu(\xi) = 0\} \subset \bar{\partial}\Phi(\mu). \quad (2.5)$$

Thus, there are elements,  $\Sigma$ , of  $\bar{\partial}\Phi(\mu)$  which are functions of  $(x, v)$  and have second components that are not null. To see this, it suffices to choose  $\xi$  such that  $\operatorname{div}_\mu(\xi) = 0$  with  $\xi(x, v)$  depending on  $(x, v)$  and  $\pi^2(\xi) \neq 0$ ; then, just set  $\Sigma = \nabla_\mu \Phi(\mu) + \xi$ .

Finally, for simplicity of notation, we define the Fisher information,  $8\mathcal{F}$ , by (see [13, 17]):

$$8\mathcal{F}(\varrho) := \begin{cases} 4 \int_{\mathbb{R}^D} |\nabla \sqrt{\varrho}|^2 dx & \text{if } \sqrt{\varrho} \in W^{1,2}(\mathbb{R}^D) \cap \{\varrho \geq 0\}, \\ +\infty & \text{if } \sqrt{\varrho} \notin W^{1,2}(\mathbb{R}^D) \cap \{\varrho \geq 0\}. \end{cases} \quad (2.6)$$

The Fisher information plays a fundamental role in our subsequent analysis.

### §3. STATIONARY SOLUTIONS ON THE TANGENT BUNDLE

$$TM := \mathbb{R}^d \times \mathbb{R}^d$$

In this section, we start our analysis by exploring special solutions of the kinetic Bohmian equation (1.1). To this end, define the Hamiltonian function

$$H(x, v) := \frac{1}{2} |v|^2 - \frac{1}{2} \frac{\Delta \sqrt{\varrho(x)}}{\sqrt{\varrho(x)}} + V(x),$$

and consider solutions of (1.1) of the form

$$\beta(x, v) = F(H(x, v) - \eta),$$

where  $\eta \in \mathbb{R}$  represents a (quasi) Fermi level and  $F : \mathbb{R} \rightarrow \mathbb{R}^+$  is a continuous strictly decreasing function. In particular, we are interested in functions  $F : \mathbb{R} \rightarrow \mathbb{R}^+$  satisfying

$$A(\alpha) := \int_{\mathbb{R}^D} F\left(\frac{1}{2} |v|^2 + \alpha\right) dv < \infty, \quad (3.1)$$

for any  $\alpha \in \mathbb{R}$ . Furthermore, the condition

$$\int_{\mathbb{R}^D} \int_{\mathbb{R}^D} \beta \, dx \, dv := M \equiv 1,$$

where  $M$  is the (normalized) mass of the system, can be used to compute  $\eta$ .

We have

$$\beta(x, v) = F \left( \frac{1}{2} |v|^2 + V(x) - \eta - \frac{1}{2} \frac{\Delta \sqrt{\varrho(x)}}{\sqrt{\varrho(x)}} \right),$$

and therefore we obtain the following integral equation for  $\varrho$ :

$$\varrho(x) = \int_{\mathbb{R}^D} F \left( \frac{1}{2} |v|^2 + V(x) - \eta - \frac{1}{2} \frac{\Delta \sqrt{\varrho(x)}}{\sqrt{\varrho(x)}} \right) dv.$$

Hence,

$$\varrho(x) = A \left( V(x) - \eta - \frac{1}{2} \frac{\Delta \sqrt{\varrho(x)}}{\sqrt{\varrho(x)}} \right),$$

from which we obtain the equation

$$-\frac{1}{2} \frac{\Delta \sqrt{\varrho(x)}}{\sqrt{\varrho(x)}} + V(x) - A^{-1}(\varrho(x)) = \eta. \quad (3.2)$$

along with

$$\int_{\mathbb{R}^D} \varrho(x) \, dx = 1. \quad (3.3)$$

To proceed further, we now restrict our attention to probability measures. For the rest of this section, and for simplicity of notation, for any probability measure,  $\mu$ , let us define  $\mathcal{F}(\mu)$  as one eighth of the Fisher information, i.e.,

$$\mu \in \mathcal{P}(\mathbb{R}^D) \rightarrow \mathcal{F}(\mu) := \begin{cases} \frac{1}{2} \int_{\mathbb{R}^D} |\nabla(\sqrt{\varrho})|^2 \, dx, & \text{if } \mu = \varrho \mathcal{L}^D, \text{ and } \sqrt{\varrho} \in W^{1,2}(\mathbb{R}^D) \\ \infty, & \text{otherwise.} \end{cases}$$

The properties of  $\mathcal{F}$  can also be studied through the convex lower semi-continuous function  $L: \mathbb{R} \times \mathbb{R}^D \rightarrow [0, +\infty]$  defined by

$$L(\varrho, \xi) := \begin{cases} \frac{|\xi|^2}{2\varrho}, & \text{if } \varrho > 0 \\ 0, & \text{if } \xi = \vec{0} \text{ and } \varrho = 0 \\ \infty, & \text{if } (\xi \neq \vec{0} \text{ and } \varrho = 0) \text{ or } (\varrho < 0) \text{ or } (\varrho = \infty). \end{cases} \quad (3.4)$$

If  $\mu \in \mathcal{P}(\mathbb{R}^D)$  then

$$\mathcal{F}(\mu) := \begin{cases} \frac{1}{4} \int_{\mathbb{R}^D} L(\varrho, \nabla \varrho) dx, & \text{if } \mu = \varrho \mathcal{L}^D, \text{ and } L(\varrho, \nabla \varrho) \in L^1(\mathbb{R}^D) \\ \infty & \text{otherwise.} \end{cases} \quad (3.5)$$

**Remark 3.1.** Since  $F$  is monotone, its set of discontinuity is countable and will be denoted by  $\{t_n\}_{n=1}^\infty$

- (i) The infimum of  $F$  must be 0, otherwise we would have  $A \equiv \infty$ .
- (ii) We exploit (i) and the dominated convergence theorem to obtain

$$\lim_{\alpha \rightarrow \infty} A(\alpha) = \lim_{\alpha \rightarrow \infty} \int_{\mathbb{R}^D} F\left(\frac{1}{2}|v|^2 + \alpha\right) dv = \int_{\mathbb{R}^D} \left( \lim_{\alpha \rightarrow \infty} F\left(\frac{1}{2}|v|^2 + \alpha\right) \right) dv = 0.$$

- (iii) Let  $\bar{\alpha} \in \mathbb{R}$  and denote by  $S_r(0)$  the sphere of radius  $r$  centered at the origin. If  $r_n^2 + 2\bar{\alpha} = 2t_n$ , then the union of  $N(\bar{\alpha}) := \cup_{n=1}^\infty S_{r_n}(0)$  is a set of null Lebesgue measure and

$$\lim_{\alpha \rightarrow \bar{\alpha}} F\left(\frac{|v|^2}{2} + \alpha\right) = F\left(\frac{|v|^2}{2} + \bar{\alpha}\right)$$

for all  $v \notin N(\bar{\alpha})$ . Thus, as above, by the dominated convergence theorem,  $\lim_{\alpha \rightarrow \bar{\alpha}} A(\alpha) = A(\bar{\alpha})$ . In other words,  $A$  is continuous on  $\mathbb{R}$ .

- (iv) Let  $\lambda_0 > 0$  be the supremum of  $F$ . We have

$$\liminf_{\alpha \rightarrow -\infty} \frac{A(\alpha)}{(-\alpha)^D} \geq \frac{|\mathbb{S}^{D-1}|}{2^D}. \quad (3.6)$$

Hence,

$$\lim_{\alpha \rightarrow -\infty} A(\alpha) = \infty. \quad (3.7)$$

Indeed, if  $\alpha < -2$

$$A(\alpha) = |\mathbb{S}^{D-1}| \int_0^\infty r^{D-1} F\left(\frac{r^2}{2} + \alpha\right) dr \geq |\mathbb{S}^{D-1}| \int_{-\frac{\alpha}{2}}^{-\alpha} r^{D-1} F\left(\frac{r^2}{2} + \alpha\right) dr.$$

Since  $F$  decreases, we conclude that

$$A(\alpha) \geq |\mathbb{S}^{D-1}| \left(-\frac{\alpha}{2}\right)^D F\left(\frac{\alpha^2}{2} + \alpha\right) \geq |\mathbb{S}^{D-1}| \left(-\frac{\alpha}{2}\right)^D F(-\alpha),$$

which implies (3.6). Thus, (3.7) holds.

(v) By (i - iv),  $A : \mathbb{R} \rightarrow (0, \infty)$  is a homeomorphism, and

$$\lim_{s \rightarrow \infty} -A^{-1}(s) = \infty, \quad \lim_{s \rightarrow 0} -A^{-1}(s) = -\infty.$$

(vi) Let  $B \in C^1(0, \infty)$  be such that

$$B'(s) = -A^{-1}(s). \quad (3.8)$$

Observe that since  $-A^{-1}$  strictly increases,  $B$  is strictly convex.

(vii) Let  $b(s) = B'(s)$ . Using first (v) and then (iv) we obtain

$$\limsup_{s \rightarrow \infty} \frac{b(s)}{s^{\frac{1}{D}}} = \limsup_{\alpha \rightarrow -\infty} \left(\frac{-\alpha}{A(\alpha)}\right)^{\frac{1}{D}} \leq \frac{2}{|\mathbb{S}^{D-1}|^{\frac{1}{D}}} =: \frac{\bar{\lambda}_1}{2}. \quad (3.9)$$

Therefore, we can choose  $T_1 > 1$  such that

$$0 < b(s) \leq \bar{\lambda}_1 s^{\frac{1}{D}} \quad (3.10)$$

for all  $s \in [T_1, \infty)$ . Since  $b(s)$  increases as  $s$  increases, setting  $\bar{\lambda}_2 := b(T_1) > 0$  we have

$$b(s) \leq \bar{\lambda}_1 s^{\frac{1}{D}} + \bar{\lambda}_2 \quad (3.11)$$

for any  $s \in (0, \infty)$ .

(viii) Suppose that  $\lim_{s \rightarrow 0^+} B(s)$  exists. Since  $B$  is defined up to additive constant, we can set  $B(0) = 0$  such that

$$B(t) = \int_0^t b(s) ds. \quad (3.12)$$

By (3.11),

$$sb(s) \leq \bar{\lambda}_1 s^{1+\frac{1}{D}} + \bar{\lambda}_2 s \quad (3.13)$$

and

$$B(s) \leq \lambda_1 \left( s^{1+\frac{1}{D}} + s \right), \quad (3.14)$$

for any  $s \in (0, \infty)$ . We have set

$$\lambda_1 := \max\{\bar{\lambda}_1, \bar{\lambda}_2\}.$$

**Lemma 3.2.** *Suppose that  $b$  and  $B$  are as in Remark 3.1 and  $B(0) = 0$ . Then*

- (i) *the infimum of  $B(s)$  is finite.*
- (ii) *The infimum of  $sb(s)$  is finite and for any  $s > 0$  we have*

$$sb_-(s) \leq B_-(s).$$

**Proof.** If  $B^*$  denotes the Legendre transform of  $B$ , then, by the fact that  $B(0) = 0$ , we have

$$B^* \geq 0. \quad (3.15)$$

(i) Since by Remark 3.1  $\lim_{s \rightarrow \infty} b(s) = \infty$ , there exists  $s_0$  such that  $b > 0$  on  $[s_0, \infty)$ . Thus,  $B$  is bounded below on  $[s_0, \infty)$  by  $B(s_0)$ . Since  $B$  is continuous on  $[0, s_0]$  we conclude that it is also bounded below there. Consequently, there exists  $\lambda_b < 0$  such that  $B \geq -\lambda_b$ .

(ii) Let  $s > 0$  and set  $\alpha = B'(s) = b(s)$ . Since

$$sb_+(s) - sb_-(s) = sb(s) = B(s) + B^*(\alpha) = B_+(s) - B_-(s) + B^*(\alpha)$$

we conclude that

$$sb_-(s) + B_+(s) + B^*(\alpha) = sb_+(s) + B_-(s).$$

Since by (3.15)  $B^* \geq 0$ , we conclude the proof.  $\square$

**Example 3.3.** Examples include

$$F(t) = e^{-t}, \quad A(t) = Ce^{-t}, \quad b(s) = \ln\left(\frac{s}{C}\right), \quad B(s) = s \ln\left(\frac{s}{C}\right) - s,$$

where

$$C := |\mathbb{S}^{D-1}| \int_0^\infty r^{D-1} e^{-\frac{r^2}{2}} dr.$$

In general, if  $B$  satisfies (3.8), then, by Remark 3.1 (v), we have

$$\lim_{s \rightarrow \infty} \frac{B(s)}{s} = \infty. \quad (3.16)$$

We shall assume that

$$B(0) := \lim_{s \rightarrow 0^+} B(s) \text{ exists.} \quad (3.17)$$

Define

$$\varrho_\infty := A(V)$$

and assume that

$$\varrho_\infty \mathcal{L}^D \in \mathcal{P}_2(\mathbb{R}^D) \quad \text{and} \quad f^\infty := B(\varrho_\infty) + \varrho_\infty V \in L^1(\mathbb{R}^D). \quad (3.18)$$

**Remark 3.4.** By the convexity of  $B$ ,  $B(s) \geq B(s_0) + b(s_0)(s - s_0)$  and if  $s > 0$ , then  $B'(s)s = B(s) + B^*(B'(s))$ . Hence,

(i) if  $\varrho : \mathbb{R}^D \rightarrow [0, \infty]$  is a Borel function

$$B(\varrho) + V\varrho \geq B(\varrho_\infty) + V\varrho_\infty + (b(\varrho_\infty) + V)(\varrho - \varrho_\infty) = B(\varrho_\infty) + V\varrho_\infty = f^\infty.$$

Consequently, due to (3.18),

$$(B(\varrho) + V\varrho)_- \leq f_-^\infty \quad \text{and} \quad B_-(\varrho) \leq V\varrho - f^\infty. \quad (3.19)$$

Hence, the functional

$$P(\varrho) := \int_{\mathbb{R}^D} (V\varrho + B(\varrho)) \, dx$$

is meaningful and achieves its minimum at  $\varrho_\infty$ .

(ii) We use the first inequality in (i) to conclude that for  $\varrho > 0$  we have

$$B'(\varrho)\varrho + V\varrho = B(\varrho) + B^*(B'(\varrho)) + V\varrho \geq f^\infty + B^*(B'(\varrho)).$$

(iii) In particular, a consequence of (ii) is that, since  $B(0) = 0$  implies  $B^* \geq 0$ , Lemma 3.2 and (3.19) imply

$$\varrho b_-(\varrho) \leq V\varrho - f^\infty.$$

**Lemma 3.5.** Let  $\varrho : \mathbb{R}^D \rightarrow [0, \infty]$  be a Borel function. Then

(i)

$$|B(\varrho)| \leq \lambda_1 \left( \varrho^{1+\frac{1}{D}} + \varrho \right) + V\varrho - f^\infty.$$

(ii) At the point where  $\varrho > 0$ , we have

$$\varrho |b(\varrho)| \leq \bar{\lambda}_1 \varrho^{1+\frac{1}{D}} + \bar{\lambda} \varrho + V\varrho - f^\infty.$$

**Proof.** We combine (3.14) and (3.19) to obtain (i). The proof of (ii) follows by combining (3.11) and Remark 3.4.  $\square$



Now define the functional  $E : \mathcal{P}(\mathbb{R}^D) \rightarrow (-\infty, \infty]$  by

$$E(\mu) := \begin{cases} \mathcal{F}(\mu) + P(\varrho), & \text{if } \mu = \varrho \mathcal{L}^D, \\ \infty, & \text{otherwise.} \end{cases} \quad (3.20)$$

**Lemma 3.6.** *Assume (3.18) holds. On its proper domain, the functional  $E$  defined in (3.20) is strictly convex and bounded below. Furthermore,  $E$  is lower semicontinuous for the narrow convergence on  $\mathcal{P}(\mathbb{R}^D)$ .*

**Proof.** As  $\mathcal{F} \geq 0$ , we use Remark 3.4 to conclude that  $E(\mu) \geq P(\varrho_\infty)$ . Furthermore, we use (3.19) to conclude that the proper domain of  $E$  is the intersection of the proper domains of  $\mathcal{F}$  and  $P$ . The strict convexity of  $B$  implies that of  $P$  on its proper domain.

To show that  $E$  is lower semicontinuous for the narrow convergence on  $\mathcal{P}(\mathbb{R}^D)$  it suffices to show that  $\mathcal{F}$  and  $P$  are both lower semicontinuous. Let  $(\mu_n)_n \subset \mathcal{P}(\mathbb{R}^D)$  be a sequence that converges to  $\mu$  narrowly and assume that

$$\sup_n E(\mu_n) < \infty.$$

By Lemma 2.2 of [17], there exist  $\varrho_n : \mathbb{R}^D \rightarrow [0, \infty]$  and  $\varrho : \mathbb{R}^D \rightarrow [0, \infty]$  such that

$$\begin{aligned} \mu_n = \varrho_n \mathcal{L}^D, \quad \mu = \varrho \mathcal{L}^D, \quad \varrho_n, \varrho \in W_{loc}^{1,1}(\mathbb{R}^D), \quad \frac{|\nabla \varrho_n|}{\sqrt{\varrho_n}}, \frac{|\nabla \varrho|}{\sqrt{\varrho}} \in L^2(\mathbb{R}^D), \\ \liminf_n \mathcal{F}(\mu_n) \geq \mathcal{F}(\mu), \end{aligned} \quad (3.21)$$

$(\sqrt{\varrho_n})_n$  converges to  $\sqrt{\varrho}$ , strongly in  $L^2(\mathbb{R}^D)$  and weakly in  $W^{1,2}(\mathbb{R}^D)$ . Thus, every subsequence of  $(\varrho_n)_n$  admits itself a subsequence which converges almost everywhere to  $\varrho$ . By (3.19),  $B(\varrho_n) + V\varrho_n + f_-^\infty \geq 0$ . Therefore, we can apply Fatou's Lemma to obtain

$$\liminf_{n \rightarrow \infty} \int_{\mathbb{R}^D} (B(\varrho_n) + V\varrho_n + f_-^\infty) dx \geq \int_{\mathbb{R}^D} (B(\varrho) + V\varrho + f_-^\infty) dx.$$

Then,

$$\liminf_n P(\varrho_n) \geq P(\varrho). \quad (3.22)$$

By (3.21) and (3.22),  $E$  is lower semicontinuous.

Convexity of  $\mathcal{F}$  follows from that of  $L$ . Consequently,  $E$  is strictly convex on its proper domain.  $\square$

We shall see that solutions of (3.2) can be obtained by minimizing  $E$ .

**Remark 3.7.** Recall that a set  $\mathcal{K} \subset \mathcal{P}(\mathbb{R}^D)$  is tight if

$$\forall \varepsilon > 0 \quad \exists K_\varepsilon \text{ compact in } \mathbb{R}^D \text{ such that } \mu(\mathbb{R}^D \setminus K_\varepsilon) \leq \varepsilon \quad \forall \mu \in \mathcal{K}. \quad (3.23)$$

Moreover, it can be verified that (3.23) is equivalent to the following integral condition (cf. Remark 5.1.5 in [3]): there exists a function  $\vartheta : \mathbb{R}^D \rightarrow [0, +\infty]$ , whose sublevels  $\{x \in \mathbb{R}^D \mid \vartheta(x) \leq c\}$  are compact in  $\mathbb{R}^D$ , such that

$$\sup_{\mu \in \mathcal{K}} \int_{\mathbb{R}^D} \vartheta(x) d\mu(x) < +\infty.$$

**Lemma 3.8.** Consider a strictly convex function  $B : \mathbb{R} \rightarrow [0, +\infty]$ , with  $B(\infty) = \infty$  and differentiable on  $(0, \infty)$ . Suppose there are strictly positive Borel functions  $\varrho_{\infty, \alpha}$  and  $\varrho_\infty$  such that, on the set where these expressions are positive, we have

$$-B'(\varrho_\infty) = V \quad (3.24)$$

and for some  $0 < \alpha < 1$

$$-B'(\varrho_{\infty, \alpha}) = \alpha V, \quad (3.25)$$

and

$$B(\varrho_{\infty, \alpha}(x)) + \alpha V(x) \varrho_{\infty, \alpha}(x) \in L^1(\mathbb{R}^D). \quad (3.26)$$

Assume  $V : \mathbb{R}^D \rightarrow \mathbb{R}$  is a Borel function which satisfies

$$\lim_{|x| \rightarrow \infty} V(x) = +\infty \quad (3.27)$$

and there exists  $\underline{V} \in \mathbb{R}$  such that  $V(x) \geq \underline{V}$  for almost every  $x \in \mathbb{R}^D$ . For any  $K > 0$  there exists a constant  $\tilde{K} > 0$  such that if  $\varrho \in L^1(\mathbb{R}^D)$  is nonnegative and

$$\int_{\mathbb{R}^D} (B(\varrho(x)) + V(x) \varrho(x)) dx \leq K, \quad (3.28)$$

then,

$$\int_{\mathbb{R}^D} V(x) \varrho(x) dx \leq \tilde{K}.$$

**Proof.** If (3.28) holds, then

$$\begin{aligned} K &\geq \int_{\mathbb{R}^D} (B(\varrho(x)) + \alpha V(x)\varrho(x) + (1-\alpha)V(x)\varrho(x)) dx \\ &= \int_{\mathbb{R}^D} (B(\varrho(x)) - B'(\varrho_{\infty,\alpha}(x))\varrho(x) + (1-\alpha)V(x)\varrho(x)) dx, \end{aligned} \quad (3.29)$$

where we used (3.25) for the last expression. Since

$$B(\varrho) \geq B(\varrho_{\infty,\alpha}) + B'(\varrho_{\infty,\alpha})(\varrho - \varrho_{\infty,\alpha}),$$

(3.29) implies

$$\begin{aligned} K &\geq \int_{\mathbb{R}^D} (B(\varrho_{\infty,\alpha}(x)) - B'(\varrho_{\infty,\alpha}(x))\varrho_{\infty,\alpha}(x) + (1-\alpha)V(x)\varrho(x)) dx \\ &= \int_{\mathbb{R}^D} (B(\varrho_{\infty,\alpha}(x)) + \alpha V(x)\varrho_{\infty,\alpha}(x) + (1-\alpha)V(x)\varrho(x)) dx \\ &= C + (1-\alpha) \int_{\mathbb{R}^D} V(x)\varrho(x) dx, \end{aligned} \quad (3.30)$$

where, due to (3.26), we have set

$$C := \int_{\mathbb{R}^D} (B(\varrho_{\infty,\alpha}(x)) + \alpha V(x)\varrho_{\infty,\alpha}(x)) dx.$$

By (3.30),

$$\int_{\mathbb{R}^D} V(x)\varrho(x) dx \leq \frac{K-C}{1-\alpha} =: \tilde{K}. \quad \square$$

**Remark 3.9.** Let  $F(s) = e^{-s}$ , which implies  $B(s) = s \ln s$ . Then, all the assumptions in Lemma 3.8 are satisfied if we have  $e^{-\alpha V(x)} \in L^1(\mathbb{R}^D)$  for some  $0 < \alpha < 1$ .

**Theorem 3.10.** *Assume  $V : \mathbb{R}^D \rightarrow \mathbb{R}$  is a Borel function, bounded below and satisfying (3.27). Suppose  $F : \mathbb{R} \rightarrow \mathbb{R}^+$  is strictly decreasing and is such that for any  $\alpha \in \mathbb{R}$  the function in (3.1) assumes only finite values. Suppose further that  $B \in C^1(0, \infty) \cap C([0, \infty))$  is such that  $B' = -A^{-1}$ ,  $B(0) = 0$  and (3.17) holds. Finally, assume that  $\lim_{s \rightarrow 0} sB'(s) = 0$ ,*

$B(\varrho_\infty) + V\varrho_\infty \in L^1(\mathbb{R}^D)$ , and (3.26) holds. If  $E \neq \infty$ , then the minimization problem

$$\operatorname{argmin}_{\mu \in \mathcal{P}_2(\mathbb{R}^D)} E(\mu),$$

has a unique solution,  $\mu_s = \varrho_s \mathcal{L}^D$ . Setting

$$\eta_s := \frac{1}{2} \|\nabla \sqrt{\varrho_s}\|_{L^2}^2 + \int_{\mathbb{R}^D} (B'(\varrho_s) + V)\varrho_s \, dx,$$

we have in the weak sense

$$-\frac{1}{2} \Delta \varrho_s + |\nabla \sqrt{\varrho_s}|^2 + 2(B'(\varrho_s) + V)\varrho_s = \eta_s \varrho_s, \quad (3.31)$$

which can be interpreted as (3.2).

**Proof.** Part I: *Existence and uniqueness of a minimizer.* Let  $\{\mu_n\}_{n \in \mathbb{N}}$  be a minimizing sequence of  $E(\mu)$ , i.e.,

$$\lim_{n \rightarrow \infty} E(\mu_n) = \inf_{\mu \in \mathcal{P}_2(\mathbb{R}^D)} E(\mu).$$

Since both  $P$  and  $\mathcal{F}$  are bounded below,

$$\sup_n \mathcal{F}(\mu_n) < \infty.$$

By Lemma 2.2 of [17], there exist  $\varrho_n : \mathbb{R}^D \rightarrow [0, \infty]$  such that  $\mu_n = \varrho_n \mathcal{L}^D$ . We have

$$\sup_n P(\varrho_n) < \infty$$

and hence, Lemma 3.8 implies

$$\sup_n \int_{\mathbb{R}^D} V \varrho_n \, dx < \infty.$$

Thus, by Remark 3.7,  $\{\mu_n\}_n$  is pre-compact for the narrow convergence. Extracting a subsequence if necessary, we assume without loss of generality that  $\{\mu_n\}_n$  converges narrowly to some  $\mu_s \in \mathcal{P}_2(\mathbb{R}^D)$ . By Lemma 2.2 of [17], there exists  $\varrho_s : \mathbb{R}^D \rightarrow [0, \infty]$  such that

$$\mu_n = \varrho_n \mathcal{L}^D, \quad \mu_s = \varrho_s \mathcal{L}^D, \quad \varrho_n, \varrho_s \in W_{loc}^{1,1}(\mathbb{R}^D), \quad \frac{|\nabla \varrho_n|}{\sqrt{\varrho_n}}, \frac{|\nabla \varrho_s|}{\sqrt{\varrho_s}} \in L^2(\mathbb{R}^D).$$

Furthermore,  $(\sqrt{\varrho_n})_n$  converges to  $\sqrt{\varrho_s}$ , strongly in  $L^2(\mathbb{R}^D)$  and weakly in  $W^{1,2}(\mathbb{R}^D)$ . By Lemma 3.6,  $E$  is lower semicontinuous for the narrow convergence and hence,  $\mu_s$  minimizes  $E$  over  $\mathcal{P}_2(\mathbb{R}^D)$ .

Uniqueness of  $\mu_s$  follows from the strict convexity property of  $E$  on its domain (cf. Lemma 3.6).

*Part II: Properties of the minimizer.* Since

$$P(\varrho_s) \leq \inf_{\mu \in \mathcal{P}_2(\mathbb{R}^D)} E(\mu), \quad (3.32)$$

we use the last statement in Lemma 3.8 and the fact that  $V$  is bounded below to deduce that

$$\varrho_s |V| \in L^1(\mathbb{R}^D). \quad (3.33)$$

By Remark 3.4

$$B(\varrho_s)_- \leq \varrho_s V - B(\varrho_\infty) - \varrho_\infty V \in L^1(\mathbb{R}^D). \quad (3.34)$$

Thus, combining (3.32), (3.33) and (3.34) we conclude that

$$B(\varrho_s) \in L^1(\mathbb{R}^D). \quad (3.35)$$

*Part III: The Euler–Lagrange equations.* Let  $v \in C_c^\infty(\mathbb{R})$  and set

$$u_0 = \sqrt{\varrho_s}, \quad u_\epsilon := \frac{u_0 + \epsilon u_0 v}{\|u_0 + \epsilon u_0 v\|_{L^2}}, \quad \mu^\epsilon = u_\epsilon^2 \mathcal{L}^D.$$

We have

$$u_\epsilon^2 = u_0^2 + 2\epsilon u_0^2 a(v) + \epsilon^2 u_0^2 a_\epsilon(v), \quad (3.36)$$

where

$$a(v) := v - \int_{\mathbb{R}^D} u_0^2 v \, dx \quad \text{and} \quad \sup_{0 < |\epsilon| < 1} \|b_\epsilon(v)\|_\infty < \infty.$$

We set

$$S := \left(1 + \|2a(v)\|_{L^\infty} + \sup_{|\epsilon| \leq 1} \|a_\epsilon(v)\|_{L^\infty}\right)^{\frac{1}{2}}.$$

We have

$$\int_{\mathbb{R}^D} V u_\epsilon^2 \, dx - \int_{\mathbb{R}^D} V u_0^2 \, dx = \epsilon \int_{\mathbb{R}^D} V u_0^2 (2a(v) + \epsilon a_\epsilon(v)) \, dx.$$

Therefore, exploiting (3.33) we can apply the dominated convergence theorem to obtain

$$\frac{d}{d\epsilon} \int_{\mathbb{R}^D} V u_\epsilon^2 \, dx \Big|_{\epsilon=0} = \int_{\mathbb{R}^D} 2a(v) V u_0^2 \, dx. \quad (3.37)$$

If  $D \geq 3$ , then

$$\frac{2D}{D-2} \geq 2 + \frac{2}{D}.$$

Since by the Sobolev Embedding theorem  $W^{1,2}(\mathbb{R}^D) \subset L^{\frac{2D}{D-2}}(\mathbb{R}^D)$ , we conclude that  $W^{1,2}(\mathbb{R}^D) \subset L^{2+\frac{2}{D}}(\mathbb{R}^D)$ . The latter inclusion remains true when  $D \in \{1, 2\}$ . Consequently,  $u_0 \in L^{2+\frac{2}{D}}(\mathbb{R}^D)$  and then,  $F^\infty \in L^1(\mathbb{R}^D)$  if we set

$$F^\infty := \lambda_1 \left( u_0^{2+\frac{2}{D}} S^{2+\frac{2}{D}} + u_0^2 S^2 \right) + |V| u_0^2 S^2 - f^\infty.$$

By Lemma 3.5

$$\begin{aligned} |B(u_\epsilon^2)| &\leq \lambda_1 \left( u_\epsilon^{2+\frac{2}{D}} + u_\epsilon^2 \right) + V u_\epsilon^2 - f^\infty \\ &\leq \lambda_1 \left( (u_0 S)^{2+\frac{2}{D}} + (u_0 S)^2 \right) + |V| (u_0 S)^2 - f^\infty = F^\infty. \end{aligned} \quad (3.38)$$

Let  $\theta_\epsilon : \mathbb{R}^D \rightarrow (0, 1)$  be such that if  $u_0 > 0$  we have the first order expansion

$$B(u_\epsilon^2) - B(u_0^2) = (u_\epsilon^2 - u_0^2) B' \left( u_0^2 + \theta_\epsilon (u_\epsilon^2 - u_0^2) \right).$$

This means that

$$B(u_\epsilon^2) - B(u_0^2) = \epsilon u_0^2 \left( 2a(v) + \epsilon a_\epsilon(v) \right) B' \left( (u_0 \Theta_\epsilon)^2 \right),$$

where

$$\Theta_\epsilon := \left( 1 + \epsilon \theta_\epsilon [2a(v) + \epsilon a_\epsilon(v)] \right)^{\frac{1}{2}}.$$

Reorganizing the expression, we have

$$\frac{B(u_\epsilon^2) - B(u_0^2)}{\epsilon} = \frac{(2a(v) + \epsilon a_\epsilon(v))}{1 + \epsilon \theta_\epsilon [2a(v) + \epsilon a_\epsilon(v)]} (u_0 \Theta_\epsilon)^2 B' \left( (u_0 \Theta_\epsilon)^2 \right). \quad (3.39)$$

This, together with Lemma 3.5, imply

$$\begin{aligned} \left| \frac{B(u_\epsilon^2) - B(u_0^2)}{\epsilon} \right| &\leq \left( \bar{\lambda}_1 (u_0 \Theta_\epsilon)^{2+\frac{2}{D}} + \bar{\lambda}_2 (u_0 \Theta_\epsilon)^2 + |V| (u_0 \Theta_\epsilon)^2 - f^\infty \right) \\ &\quad \times \frac{|2a(v) + \epsilon a_\epsilon(v)|}{1 + \epsilon \theta_\epsilon [2a(v) + \epsilon a_\epsilon(v)]}. \end{aligned}$$

Thus, if  $|\epsilon|$  is small enough so that  $2|\epsilon \theta_\epsilon [2a(v) + \epsilon a_\epsilon(v)]| \leq 1$ , then

$$\begin{aligned} \left| \frac{B(u_\epsilon^2) - B(u_0^2)}{\epsilon} \right| &\leq 2S^2 \left( \bar{\lambda}_1 (u_0 S)^{2+\frac{2}{D}} + \bar{\lambda}_2 (u_0 S)^2 + V (u_0 S)^2 - f^\infty \right) \in L^1(\mathbb{R}^D). \end{aligned} \quad (3.40)$$

Since  $B(0) = 0$  and  $u_\epsilon \equiv 0$  on  $\{u_0 = 0\}$ , we conclude that

$$\int_{\mathbb{R}^D} \frac{B(u_\epsilon^2) - B(u_0^2)}{\epsilon} dx = \int_{\{u_0 > 0\}} \frac{B(u_\epsilon^2) - B(u_0^2)}{\epsilon} dx.$$

Due to (3.40), we can apply the dominated convergence theorem to conclude that

$$\frac{d}{d\epsilon} \int_{\mathbb{R}^D} B(u_\epsilon^2) dx \Big|_{\epsilon=0} = \int_{\{u_0 > 0\}} \lim_{\epsilon \rightarrow 0} \frac{B(u_\epsilon^2) - B(u_0^2)}{\epsilon} dx.$$

We then let  $\epsilon$  go to 0 in (3.39) to deduce that

$$\frac{d}{d\epsilon} \int_{\mathbb{R}^D} B(u_\epsilon^2) dx \Big|_{\epsilon=0} = 2 \int_{\{u_0 > 0\}} a(v) u_0^2 B'(u_0^2) dx.$$

Taking into account the fact that  $\lim_{s \rightarrow 0} sB'(s) = 0$ , we get

$$\frac{d}{d\epsilon} \int_{\mathbb{R}^D} B(u_\epsilon^2) dx \Big|_{\epsilon=0} = 2 \int_{\mathbb{R}^D} a(v) u_0^2 B'(u_0^2) dx. \quad (3.41)$$

Note that

$$|\nabla u_\epsilon|^2 = |\nabla u_0|^2 + 2\epsilon e(v) + \epsilon^2 (u_0^2 + |u_0 \nabla u_0| + |\nabla u_0|^2) e_\epsilon(v), \quad (3.42)$$

where

$$e(v) := \langle \nabla u_0; \nabla(vu_0) \rangle - |\nabla u_0|^2 \int_{\mathbb{R}^D} u_0^2 v dx \quad \text{and} \quad \sup_{0 < |\epsilon| < 1} \|e_\epsilon(v)\|_\infty < \infty.$$

Hence, applying the dominated convergence theorem, we have

$$\frac{d}{d\epsilon} \int_{\mathbb{R}^D} |\nabla u_\epsilon|^2 dx \Big|_{\epsilon=0} = \int_{\mathbb{R}^D} 2e(v) dx. \quad (3.43)$$

We combine (3.37), (3.41) and (3.43) to conclude that

$$\frac{d}{d\epsilon} E(\mu^\epsilon) \Big|_{\epsilon=0} = \int_{\mathbb{R}^D} e(v) dx + 2 \int_{\mathbb{R}^D} (B'(u_0^2) + V) u_0^2 a(v) dx.$$

Using the fact that  $E(\mu^\epsilon)$  achieves its minimum at  $\epsilon = 0$ , we conclude that

$$0 = \frac{d}{d\epsilon} E(\mu^\epsilon) \Big|_{\epsilon=0} = \int_{\mathbb{R}^D} (e(v) + 2(B'(u_0^2) + V) u_0^2 a(v)) dx. \quad (3.44)$$

In other words,

$$\int_{\mathbb{R}^D} \left( \langle \nabla u_0; \nabla(u_0 v) \rangle - l_0^2 u_0^2 v + 2(B'(u_0^2) + V) u_0^2 v - 2u_0^2 l_1 v \right) dx = 0,$$

where

$$l_0 := \|\nabla u_0\|_{L^2}, \quad l_1 := \int_{\mathbb{R}^D} (B'(u_0^2) + V) u_0^2 dx.$$

This implies that for all  $v \in C_c^\infty(\mathbb{R}^D)$

$$\begin{aligned} 0 &= \int_{\mathbb{R}^D} \left( \langle u_0 \nabla u_0; \nabla v \rangle + (|\nabla u_0|^2 - l_0^2 u_0^2 + 2(B'(u_0^2) + V) u_0^2 - 2u_0^2 l_1) v \right) dx \\ &= \int_{\mathbb{R}^D} \left( \frac{1}{2} \langle \nabla \varrho_s; \nabla v \rangle + (|\nabla \sqrt{\varrho_s}|^2 + 2(B'(\varrho_s) + V - l_1 - \frac{l_0^2}{2}) \varrho_s) v \right) dx. \end{aligned} \tag{3.45}$$

This means that (3.31) holds in the distributional sense.  $\square$

**Definition 3.11.** Given  $G : \mathcal{P}(\mathbb{R}^D) \rightarrow (-\infty, \infty]$ , we define  $G^*$  on the set of Borel functions  $W : \mathbb{R}^D \rightarrow (\infty, \infty]$  which is bounded below, by

$$G^*(W) = \sup_{\mu} \left\{ \int_{\mathbb{R}^D} W(x) \mu(dx) - G(\mu) \mid \mu \in \mathcal{P}(\mathbb{R}^D) \right\}.$$

We refer to  $G^*$  as the Legendre transform of  $G$ .

The next result follows immediately from the definition of the Legendre transform.

**Lemma 3.12.** If  $V$  and  $E$  are as in Theorem 3.10 and for any  $\mu \in \mathcal{P}(\mathbb{R}^D)$  we define

$$G(\mu) := \begin{cases} \mathcal{F}(\mu) + \int_{\mathbb{R}^D} B(\varrho) dx, & \text{if } \mu = \varrho \mathcal{L}^D, \\ \infty, & \text{otherwise,} \end{cases}$$

then

$$-G^*(-V) = \inf_{\varrho} E(\varrho \mathcal{L}^D).$$



**Remark 3.13.** The conclusions in Theorem 3.10 remain valid if we replace  $\mathbb{R}^D$  by the torus  $\mathbb{T}^D$ . We keep the same assumptions on  $B$  and  $b$ , but on  $V$  we only assume that  $V : \mathbb{R}^D \rightarrow \mathbb{R}$  is a Borel function bounded below, skipping (3.27).

#### §4. MOREAU–YOSIDA APPROXIMATION

In the remainder of this paper, we develop an approximative version of the kinetic Bohmian equation with the aim of applying the results obtained in [2]. As we shall see, this approximative version allows us to get around one of the main difficulties of the kinetic Bohmian equation when studied in the context of Wasserstein Hamiltonian flows: the lack of  $\lambda$ -convexity of the corresponding Hamiltonian.

We assume throughout this section that  $d$  is an integer with  $d \geq 1$  and  $D \in \{d, 2d\}$ . We also assume that  $\Phi : \mathcal{P}_2(\mathbb{R}^D) \rightarrow [0, \infty]$  is proper and lower semicontinuous with respect to the narrow convergence on bounded subsets of  $\mathcal{P}_2(\mathbb{R}^D)$ . If  $D = 2d$  we assume that

$$\emptyset \neq D(\Phi) \subset \left\{ \mu \in \mathcal{P}_2(\mathbb{R}^{2d}) \mid \pi_{\#}^1 \mu \in \mathcal{P}_2^r(\mathbb{R}^d) \right\}.$$

Finally, when  $D = d$ , we assume that

$$\emptyset \neq D(\Phi) \subset \mathcal{P}_2^r(\mathbb{R}^d).$$

For  $\tau > 0$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^{2d})$ , we define the Moreau–Yosida approximation of  $\Phi$  by

$$\Phi_\tau(\mu) = \inf_{\nu} \left\{ \frac{1}{2\tau} W_2^2(\mu, \nu) + \Phi(\nu) \right\}. \quad (4.1)$$

We shall use the function

$$M_2(\mu) = \frac{1}{2} \int_{\mathbb{R}^D} |z|^2 \mu(dz).$$

We fix  $\nu_* \in D(\Phi)$  and set

$$C_\tau := \frac{2}{\tau} M_2(\nu_*) + \Phi(\nu_*).$$

**Remark 4.1.** Existence of a solution in (4.1) is a standard result due to the fact that  $\Phi$  is lower semicontinuous for the narrow convergence. Moreover, we define the set of minimizers

$$J_\tau^\Phi(\mu) := \left\{ \nu \in \mathcal{P}_2(\mathbb{R}^D) \mid \Phi_\tau(\mu) = \frac{1}{2\tau} W_2^2(\mu, \nu) + \Phi(\nu) \right\}.$$

By abuse of notation, we denote by  $\mu^\tau$  any element of  $J_\tau^\Phi(\mu)$ .

If  $\mu \in \mathcal{P}_2^*(\mathbb{R}^D)$ , then  $W_2^2(\mu, \cdot)$  is strictly convex along geodesics of the  $L^1$ -metric and hence, since in addition  $\Phi$  is convex,  $J_\tau^\Phi(\mu)$  reduces to a single element (cf., e.g., [13] and [23]).

**Lemma 4.2.** *The following hold:*

- (i)  $-\Phi_\tau$  is  $(\frac{-1}{\tau})$ -convex along geodesics of constant speed.
- (ii) If  $\mu \in \mathcal{P}_2(\mathbb{R}^D)$ , then

$$0 \leq \Phi_\tau(\mu) \leq \frac{1}{\tau} M_2(\mu) + C_\tau.$$

- (iii) Let  $\mu_0, \mu \in \mathcal{P}_2(\mathbb{R}^D)$ ; let  $G \in \Gamma_o(\mu_0, \mu_0^\tau)$  and denote by  $G_{\mu_0^\tau}^{\mu_0}$  the barycentric projection of  $G$ . Let  $\bar{G} \in \Gamma_o(\mu_0, \mu)$ . We have

$$\Phi_\tau(\mu) \leq \Phi_\tau(\mu_0) + \int_{\mathbb{R}^D \times \mathbb{R}^D} \left\langle \frac{w - G_{\mu_0^\tau}^{\mu_0}(w)}{\tau}; z - w \right\rangle \bar{G}(dw, dz) + \frac{1}{2\tau} W_2^2(\mu, \mu_0).$$

- (iv) We conclude that

$$\frac{\text{id} - G_{\mu_0^\tau}^{\mu_0}}{\tau} \in \bar{\partial} \Phi_\tau(\mu_0).$$

**Proof.** (i) Let  $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^D)$  and let  $(\mu_t)_t$  be a geodesic of constant speed connecting  $\mu_0$  to  $\mu_1$ . Fix  $t \in (0, 1)$  and let  $\mu_t^\tau \in \mathcal{P}_2(\mathbb{R}^D)$  be such that

$$\Phi_\tau(\mu_t) = \frac{1}{2\tau} W_2^2(\mu_t, \mu_t^\tau) + \Phi(\mu_t^\tau). \quad (4.2)$$

We have

$$\Phi_\tau(\mu_i) \leq \frac{1}{2\tau} W_2^2(\mu_i, \mu_i^\tau) + \Phi(\mu_i^\tau) \quad \forall i \in \{0, 1\}.$$

Thus,

$$(1-t)\Phi_\tau(\mu_0) + t\Phi_\tau(\mu_1) \leq \frac{1-t}{2\tau} W_2^2(\mu_0, \mu_t^\tau) + \frac{t}{2\tau} W_2^2(\mu_1, \mu_t^\tau) + \Phi(\mu_t^\tau). \quad (4.3)$$

Since  $-1/2W_2^2(\cdot, \mu_t^\tau)$  is  $(-1)$ -convex along geodesics of constant speed (cf., e.g., [3]), we conclude that

$$W_2^2(\mu_t, \mu_t^\tau) + t(1-t)W_2^2(\mu_0, \mu_1) \geq (1-t)W_2^2(\mu_0, \mu_t^\tau) + tW_2^2(\mu_1, \mu_t^\tau).$$

This, along with (4.3), yields

$$(1-t)\Phi_\tau(\mu_0) + t\Phi_\tau(\mu_1) \leq \frac{1}{2\tau} W_2^2(\mu_t, \mu_t^\tau) + \frac{1}{2\tau} t(1-t)W_2^2(\mu_0, \mu_1) + \Phi(\mu_t).$$

Therefore, by (4.2)

$$(1-t)\Phi_\tau(\mu_0) + t\Phi_\tau(\mu_1) \leq \frac{1}{2\tau}t(1-t)W_2^2(\mu_0, \mu_1) + \Phi_\tau(\mu_t).$$

This proves (i).

(ii) We have

$$0 \leq \Phi_\tau(\mu) \leq \frac{1}{2\tau}W_2^2(\mu, \nu_*) + \Phi(\nu_*).$$

This, together with the triangle inequality

$$(W_2(\mu, \nu_*))^2 \leq \left(W_2(\mu, \delta_0) + W_2(\delta_0, \nu_*)\right)^2 \leq 4M_2(\mu) + 4M_2(\nu_*),$$

gives (ii).

(iii) Let  $\mu_0, \mu \in \mathcal{P}_2(\mathbb{R}^D)$ . We have

$$\Phi_\tau(\mu) \leq \Phi(\mu_0^\tau) + \frac{W_2^2(\mu, \mu_0^\tau)}{2\tau} = \Phi_\tau(\mu_0) - \frac{W_2^2(\mu_0, \mu_0^\tau)}{2\tau} + \frac{W_2^2(\mu, \mu_0^\tau)}{2\tau}. \quad (4.4)$$

By Theorem 7.3.2 [3],  $\psi := -1/2W_2^2(\cdot, \mu_0^\tau)$  is  $(-1)$ -convex along geodesics.

Since  $\mathbf{id} - G_{\mu_0^\tau}^\mu \in \partial\psi(\mu_0)$ , by Theorem 10.3.6 [3], we have

$$\psi(\mu) \geq \psi(\mu_0) + \int_{\mathbb{R}^D \times \mathbb{R}^D} \langle G_{\mu_0^\tau}^\mu(w) - w; z - w \rangle \bar{G}(dw, dz) - \frac{1}{2}W_2^2(\mu, \mu_0).$$

This, along with (4.4), yields (iii).

We use (i), (iii) and Theorem 10.3.6 [3] to obtain (iv).  $\square$

**Remark 4.3.** Let  $\mu \in \mathcal{P}_2(\mathbb{R}^D)$  and  $G \in \Gamma_o(\mu, \mu^\tau)$ . Furthermore, let  $G_{\mu^\tau}^\mu$  be the barycentric projection of  $G$  based at  $\mu^\tau$ .

(i) We have

$$\frac{G_{\mu^\tau}^\mu - \mathbf{id}}{\tau} \in \partial\Phi(\mu^\tau), \quad \frac{\mathbf{id} - G_{\mu^\tau}^\mu}{\tau} \in \bar{\partial}\Phi_\tau(\mu).$$

(ii) We have

$$\frac{W_2^2(\mu^\tau, \mu)}{\tau^2} \geq \left\| \frac{\mathbf{id} - G_{\mu^\tau}^\mu}{\tau} \right\|_{\mu^\tau}^2, \quad \left\| \frac{\mathbf{id} - G_{\mu^\tau}^\mu}{\tau} \right\|_{\mu}^2.$$

(iii) If  $R > 0$  and  $\mu_1, \mu_0 \in \mathcal{P}_2(\mathbb{R}^D)$  are such that  $W_2(\mu_0, \delta_0) \leq R$  and  $W_2(\mu_1, \delta_0) \leq R$ , then for a constant  $\bar{C}_{\tau, R}$  depending on  $R$  and  $\tau$

$$|\Phi_\tau(\mu_1) - \Phi_\tau(\mu_0)| \leq \bar{C}_{\tau, R}W_2(\mu_1, \mu_0).$$

**Proof.** The first claim in (i) can be derived from Lemma 10.3.4 [3] while the second claim is Lemma 4.2 (iv). The inequalities in (ii) are consequences of Jensen's inequality.

(iii) Assume  $R > 0$  and  $\mu_1, \mu_0 \in \mathcal{P}_2(\mathbb{R}^D)$  are such that  $M_2(\mu_0), M_2(\mu_1) \leq R$ . Without loss of generality, we may assume that  $0 \leq \Phi_\tau(\mu_1) - \Phi_\tau(\mu_0)$ . Let  $\bar{G} \in \Gamma_o(\mu_0, \mu_1)$ , let  $G \in \Gamma_o(\mu_0, \mu_0^\tau)$  and denote by  $G_{\mu_0}^{\mu_0^\tau}$  the barycentric projection of  $G$ . By Lemma 4.2 (iii)

$$|\Phi_\tau(\mu_1) - \Phi_\tau(\mu_0)| \leq \int_{\mathbb{R}^D \times \mathbb{R}^D} \left\langle \frac{w - G_{\mu_0}^{\mu_0^\tau}(w)}{\tau}; z - w \right\rangle \bar{G}(dw, dz) + \frac{1}{2\tau} W_2^2(\mu_1, \mu_0)$$

and hence, by Hölder's inequality

$$|\Phi_\tau(\mu_1) - \Phi_\tau(\mu_0)| \leq \left\| \frac{w - G_{\mu_0}^{\mu_0^\tau}(w)}{\tau} \right\|_{\mu_0} W_2(\mu_1, \mu_0) + \frac{W_2^2(\mu_1, \mu_0)}{2\tau}.$$

We then use (ii) to obtain

$$\begin{aligned} |\Phi_\tau(\mu_1) - \Phi_\tau(\mu_0)| &\leq W_2(\mu_1, \mu_0) \left( \frac{W_2(\mu_0^\tau, \mu_0)}{\tau^2} + \frac{W_2(\mu_1, \mu_0)}{2\tau} \right) \\ &\leq W_2(\mu_1, \mu_0) \left( \sqrt{\frac{\Phi_\tau(\mu_0)}{\tau}} + \frac{W_2(\mu_1, \mu_0)}{2\tau} \right). \end{aligned}$$

We use Lemma 4.2 (ii) to conclude.  $\square$

**Remark 4.4.** Assume  $(\mu_k)_k \subset \mathcal{P}_2(\mathbb{R}^D)$  converges narrowly to  $\mu \in \mathcal{P}_2(\mathbb{R}^D)$ . If there exists  $\eta \in \mathcal{P}_2(\mathbb{R}^D)$  such that

$$\lim_{k \rightarrow \infty} W_2(\mu_k, \eta) = W_2(\mu, \eta) \quad (4.5)$$

then  $(\mu_k)_k$  converges in the Wasserstein metric to  $\mu$ ; this is by now a standard result.

**Lemma 4.5.** Suppose  $(\mu_n)_n$  is a bounded sequence in  $\mathcal{P}_2(\mathbb{R}^D)$  that converges narrowly to  $\mu \in \mathcal{P}_2(\mathbb{R}^D)$ . Let  $\mu_n^\tau \in J_\tau^\Phi(\mu_n)$  and let  $G_n \in \Gamma_o(\mu_n, \mu_n^\tau)$ .

- (i) Up to a subsequence,  $(\mu_n^\tau)_n \subset \mathcal{P}_2(\mathbb{R}^D)$  converges in the Wasserstein metric to some  $\mu^\tau$ . Furthermore, a subsequence of  $(G_n)_n$  obtained from a second extraction has itself a subsequence which converges narrowly to some  $G \in \Gamma_o(\mu, \mu^\tau)$ .
- (ii) If  $J_\tau^\Phi(\mu) = \{\mu^\tau\}$ , then the whole sequence  $(\mu_n^\tau)_n \subset \mathcal{P}_2(\mathbb{R}^D)$  converges in the Wasserstein metric to  $\mu^\tau$ .

(iii) If  $J_\tau^\Phi(\mu) = \{\mu^\tau\}$  and  $\Gamma_o(\mu, \mu^\tau)$  has a unique element  $G$ , then the whole sequence  $(G_n)_n$  converges narrowly to  $G$ .

**Proof.** (i) Assume  $(\mu_n)_n \subset \mathcal{P}_2(\mathbb{R}^D)$  narrowly converges to  $\mu$ . Since  $\Phi \geq 0$ , we use Lemma 4.2 to conclude that

$$\sup_n W_2(\mu_n, \mu_n^\tau) < \infty \quad \text{and} \quad \sup_n \Phi(\mu_n^\tau) < \infty. \quad (4.6)$$

This, together with the fact that  $(\mu_n)_n$  is bounded in  $\mathcal{P}_2(\mathbb{R}^D)$ , implies that  $(\mu_n^\tau)_n$  is bounded in  $\mathcal{P}_2(\mathbb{R}^D)$ . Consider a subsequence  $(\mu_{n_k}^\tau)_k \subset \mathcal{P}_2(\mathbb{R}^D)$ . Since bounded subsets of  $\mathcal{P}_2(\mathbb{R}^D)$  are tight (cf., e.g., Remark 5.1.5 [3]) we may assume without loss of generality that  $(\mu_{n_k}^\tau)_k \subset \mathcal{P}_2(\mathbb{R}^D)$  converges narrowly to some  $\bar{\mu} \in \mathcal{P}_2(\mathbb{R}^D)$ . Because  $(G_{n_k}^\tau)_k \subset \mathcal{P}(\mathbb{R}^D \times \mathbb{R}^D)$  is tight, extracting a subsequence if necessary, we may assume that  $(G_{n_k}^\tau)_k$  converges narrowly to some  $G$ . By the stability of optimal transport plans for the narrow convergence (cf., e.g., Proposition 7.1.3 [3]),  $G \in \Gamma_o(\mu, \bar{\mu})$  and

$$\liminf_{k \rightarrow \infty} W_2(\mu_{n_k}, \mu_{n_k}^\tau) \geq W_2(\mu, \bar{\mu}). \quad (4.7)$$

The lower semicontinuity of  $\Phi$  for the narrow convergence and the second inequality in (4.6) allow us to assert that

$$\infty > \liminf_{k \rightarrow \infty} \Phi(\mu_{n_k}^\tau) \geq \Phi(\bar{\mu}). \quad (4.8)$$

If  $\nu \in \mathcal{P}_2(\mathbb{R}^D)$  then

$$\Phi(\nu) + \frac{W_2^2(\mu_{n_k}, \nu)}{2\tau} \geq \Phi(\mu_{n_k}^\tau) + \frac{W_2^2(\mu_{n_k}, \mu_{n_k}^\tau)}{2\tau}.$$

Therefore, by (4.7) and (4.8)

$$\Phi(\nu) + \frac{W_2^2(\mu, \nu)}{2\tau} \geq \Phi(\bar{\mu}) + \frac{W_2^2(\mu, \bar{\mu})}{2\tau}. \quad (4.9)$$

Hence,  $\bar{\mu} \in J_\tau^\Phi(\mu)$ . Would the inequality in (4.7) be strict, so would be the one in (4.9), yielding a contradiction. Thus,

$$\lim_{k \rightarrow \infty} W_2(\mu_{n_k}, \mu_{n_k}^\tau) = W_2(\mu, \bar{\mu}).$$

The identities

$$\begin{aligned} & |W_2(\mu, \mu_{n_k}^\tau) - W_2(\mu, \bar{\mu})| \\ &= \left| (W_2(\mu, \mu_{n_k}^\tau) - W_2(\mu_{n_k}^\tau, \mu_{n_k})) + (W_2(\mu_{n_k}^\tau, \mu_{n_k}) - W_2(\mu, \bar{\mu})) \right| \\ &\leq W_2(\mu, \mu_{n_k}) + |W_2(\mu_{n_k}^\tau, \mu_{n_k}) - W_2(\mu, \bar{\mu})| \end{aligned}$$

yield

$$\lim_{k \rightarrow \infty} W_2(\mu, \mu_{n_k}^\tau) = W_2(\mu, \bar{\mu}).$$

We apply Remark 4.4 to conclude that  $(\mu_{n_k}^\tau)_n \subset \mathcal{P}_2(\mathbb{R}^D)$  converges in the Wasserstein metric to  $\bar{\mu}$ .

(ii) By (i), if  $\mu^\tau$  is unique, every subsequence of  $(\mu_n)_n$  admits itself a subsequence converging to  $\mu^\tau$ . Hence, the whole sequence must converge to  $\mu^\tau$ .

(iii) As in (ii), we use (i) to conclude that if  $\mu^\tau$  is unique and  $G$  is the unique element of  $\Gamma_o(\mu, \mu^\tau)$ , then the whole sequence  $(G_n)_n$  must converge to  $G$ .  $\square$

### §5. FUNCTIONS ON $\mathcal{P}_2(\mathbb{R}^{2d})$ DEPENDING ONLY ON FIRST MARGINALS

To emphasize the difference between the spatial and velocity variables, we set

$$M := \mathbb{R}^d, \quad TM := M \times \mathbb{R}^d,$$

and use notation such as  $x \in M$ ,  $(x, a) \in M \times M$ ,  $(x, v) \in TM$ , and so forth.

Suppose

$$\Phi : \mathcal{P}_2(TM) \rightarrow (-\infty, \infty], \quad \phi : \mathcal{P}_2(M) \rightarrow (-\infty, \infty]$$

are lower semicontinuous for the narrow convergence and

$$\Phi(\mu) = \phi(\pi_{\#}^1 \mu) \quad \forall \mu \in \mathcal{P}_2(TM).$$

In this section we study the relation between the superdifferential of the Moreau–Yosida approximations  $\Phi_\tau$  at  $\mu \in \mathcal{P}_2(\mathbb{R}^{2D})$  and that of  $\phi_\tau$  at  $\pi_{\#}^1 \mu \in \mathcal{P}_2(\mathbb{R}^D)$ . The set

$$S := \{(x, v, a, b) \in TM \times TM \mid v = b\}$$

plays an important role in our study.

**Definition 5.1.** Let  $\mu \in \mathcal{P}_2(TM)$ ,  $\eta \in \mathcal{P}_2(M)$ ,  $\pi_{\#}^1 \mu = \eta$  and let  $\gamma \in \Gamma_o(\eta, \mu)$ . Let  $(\mu_x)_x$  be the disintegration of  $\mu$  with respect to  $\eta$  in the sense that

$$\int_{TM} l(x, v) \mu(dx, dv) = \int_M \eta(dx) \int_{\mathbb{R}^d} l(x, v) \mu_x(dv) \quad \forall l \in C_b(TM).$$

(i) We define the Borel measure  $G := G^{\mu, \gamma}$  on  $TM \times TM$  by

$$\begin{aligned} & \int_{TM \times TM} g(x, v, a, b) G(dx, dv, da, db) \\ &= \int_{M \times M} \gamma(dx, da) \int_{\mathbb{R}^d} g(x, v, a, v) \mu_x(dv) \quad \forall g \in C_c(TM \times TM). \end{aligned} \quad (5.1)$$

(ii) We define the Borel measure  $m^{\mu, \gamma}$  by

$$\begin{aligned} & \int_{TM} g(a, b) m^{\mu, \gamma}(da, db) \\ &= \int_{M \times M} \gamma(dx, da) \int_{\mathbb{R}^d} g(a, b) \mu_x(db) \quad \forall g \in C_c(TM). \end{aligned} \quad (5.2)$$

**Remark 5.2.** Using the above notation, the following hold:

- (i)  $G^{\mu, \gamma}$  is supported by the closed set  $S$ .
- (ii)  $G^{\mu, \gamma} \in \Gamma_o(\mu, m^{\mu, \gamma})$ .
- (iii)  $\pi_{\#}^1 m^{\mu, \gamma} = \eta$ .
- (iv)  $W_2(\varrho, \eta) = W_2(\mu, m^{\mu, \gamma})$ .

**Proof.** (i) Observe that

$$\int_{TM \times TM} |v - b|^2 G^{\mu, \gamma}(dx, dv, da, db) = \int_{M \times M} \gamma(dx, da) \int_{\mathbb{R}^d} 0 \mu_x(dv) = 0,$$

which proves that  $G^{\mu, \gamma}$  is supported by the closed set  $S$ .

(ii) Let  $g \in C_c(TM)$ . We have

$$\begin{aligned} \int_{TM \times TM} g(x, v) G^{\mu, \gamma}(dx, dv, da, db) &= \int_{M \times M} \gamma(dx, da) \int_{\mathbb{R}^d} g(x, v) \mu_x(dv) \\ &= \int_M \varrho(dx) \int_{\mathbb{R}^d} g(x, v) \mu_x(dv) \\ &= \int_{TM} \varrho(dx) g(x, v) \mu(dx, dv). \end{aligned} \quad (5.3)$$

Similarly,

$$\begin{aligned}
\int_{TM \times TM} g(a, b) G^{\mu, \gamma}(dx, dv, da, db) &= \int_{M \times M} \gamma(dx, da) \int_{\mathbb{R}^d} g(a, v) \mu_x(dv) \\
&= \int_{M \times M} \gamma(dx, da) \int_{\mathbb{R}^d} g(a, b) \mu_x(db) \quad (5.4) \\
&= \int_{TM} g(a, b) m^{\mu, \gamma}(da, db).
\end{aligned}$$

By (5.3) and (5.4),  $G^{\mu, \gamma} \in \Gamma(\mu, m^{\mu, \gamma})$ .

To conclude that  $G \in \Gamma_o(\varrho, \eta_0)$ , it suffices to show that the support of  $G$  is cyclically monotone (cf. e.g. Section 6.2.3 [3]). Let  $\{(x_i, v_i, a_i, b_i)\}_{i=1}^n \subset \text{spt } G$  and let  $\sigma$  be a permutation of  $n$  letters. By (i),  $b_i = v_i$ , and therefore, using the fact that  $\{(x_i, a_i)\}_{i=1}^n \subset \text{spt } \gamma$  and  $\gamma \in \Gamma_o(\varrho, \nu_0)$  we conclude that

$$\sum_{i=1}^n |(x_i, v_i) - (a_i, b_i)|^2 = \sum_{i=1}^n |x_i - a_i|^2 \leq \sum_{i=1}^n |x_i - a_{\sigma(i)}|^2 + \sum_{i=1}^n |v_i - b_{\sigma(i)}|^2.$$

Equivalently, this means

$$\sum_{i=1}^n |(x_i, v_i) - (a_i, b_i)|^2 \leq \sum_{i=1}^n |(x, v_i) - (a_{\sigma(i)}, b_{\sigma(i)})|^2.$$

Thus, the support of  $G^{\mu, \eta, \gamma}$  is cyclically monotone, which concludes the proof of (ii).

(iii) Let  $g \in C_c(M)$ . We have

$$\begin{aligned}
\int_{TM \times TM} g(a) \bar{m}(da, db) &= \int_{M \times M} \bar{\gamma}(dx, da) \int_{\mathbb{R}^d} g(a) \mu_x(db) \\
&= \int_{M \times M} g(a) \bar{\gamma}(dx, da) = \int_M g(a) \eta(da).
\end{aligned}$$

Thus  $\pi_{\#}^1 \bar{m} = \eta$ .



(iv) Using the fact that by (i)  $G^{\mu,\gamma}$  is supported by  $S$  and by (ii) it is optimal, we have

$$\begin{aligned} W_2^2(\mu, m^{\mu,\gamma}) &= \int_{TM \times TM} |(x, v) - (a, b)|^2 G^{\mu,\gamma}(dx, dv, da, db) \\ &= \int_{TM \times TM} |x - a|^2 G(dx, dv, da, db). \end{aligned}$$

Since

$$\pi_{\#}^{1,3} G^{\mu,\gamma} = \gamma \in \Gamma_o(\varrho, \eta),$$

the previous identity becomes  $W_2^2(\mu, m^{\mu,\gamma}) = W_2^2(\varrho, \eta)$ .  $\square$

**Lemma 5.3.** *Let  $\mu \in \mathcal{P}_2(TM)$  and let  $\varrho, \eta \in \mathcal{P}_2(M)$  be such that  $\pi_{\#}^1 \mu = \varrho$ .*

(i) *We have*

$$\inf_{m \in \mathcal{P}_2(TM)} \left\{ W_2^2(\mu, m) \mid m \in \mathcal{P}_2(TM), \pi_{\#}^1 m = \eta \right\} = W_2^2(\varrho, \eta). \quad (5.5)$$

(ii) *If  $\gamma \in \Gamma_o(\varrho, \eta)$ , then  $m^{\mu,\gamma}$  minimizes (5.5).*

(iii) *If  $\bar{m}$  minimizes (5.5) and  $\bar{G} \in \Gamma_o(\mu, \bar{m})$ , then  $\bar{\gamma} := \pi_{\#}^{1,3} \bar{G} \in \Gamma_o(\varrho, \eta)$  and  $\bar{G}$  is supported by  $S$ .*

(iv) *If  $\varrho \ll \mathcal{L}^D$ , then  $\bar{m} = m^{\mu,\gamma}$  is the unique minimizer in (5.5) and  $\Gamma_o(\mu, \bar{m}) = \{G^{\mu,\gamma}\}$ .*

**Proof.** Let  $m \in \mathcal{P}_2(TM)$  be such that  $\pi_{\#}^1 m = \eta$  and let  $G \in \Gamma_o(\mu, m)$ . Set  $\bar{\gamma} := \pi_{\#}^{1,3} G \in \Gamma(\varrho, \eta)$ . We have

$$\begin{aligned} W_2^2(\mu, m) &= \int_{TM \times TM} |(x, v) - (a, b)|^2 G(dx, dv, da, db) \\ &= \int_{TM \times TM} (|x - a|^2 + |v - b|^2) G(dx, dv, da, db) \\ &= \int_{M \times M} |x - a|^2 \bar{\gamma}(dx, da) + \int_{TM \times TM} |v - b|^2 G(dx, dv, da, db) \\ &\geq W_2^2(\varrho, \eta). \end{aligned} \quad (5.6)$$

Observe that the inequality in (5.6) is strict unless  $\bar{\gamma} \in \Gamma_o(\varrho, \eta)$  and  $G$  is supported by  $S$ . In light of Remark 5.2 and (5.6)

$$W_2^2(\mu, m) \geq W_2^2(\varrho, \eta) = W_2^2(\mu, m^{\mu,\eta,\gamma}).$$

Hence, we have established (i) and (ii).

(iii) From the previous result, if  $\bar{m}$  is another minimizer in (5.5) and  $\bar{G} \in \Gamma_o(\mu, m)$ , then  $\bar{G}$  must be supported by  $S$  and we must have  $\bar{\gamma} := \pi_{\#}^{1,3} \bar{G} \in \Gamma_o(\varrho, \eta)$ , otherwise the inequality in (5.6) would be strict.

(iv) Assume now that  $\varrho \ll \mathcal{L}^D$  and let  $u : \mathbb{R}^D \rightarrow (-\infty, \infty]$  be a lower semicontinuous convex function such that  $(\mathbf{id} \times \nabla u)_{\#} \varrho = \bar{\gamma}$ . The first of the following identities is due to (iii). If  $g \in C_c^\infty(TM \times TM)$ , then

$$\begin{aligned}
& \int_{TM \times TM} g(x, v, a, b) \bar{G}(dx, dv, da, db) \\
&= \int_{TM \times TM} g(x, v, \nabla u(x), v) \bar{G}(dx, dv, da, db) \\
&= \int_{TM} g(x, v, \nabla u(x), v) \mu(dx, dv) \\
&= \int_M \varrho(dx) \int_{\mathbb{R}^D} g(x, v, \nabla u(x), v) \mu_x(dv) \\
&= \int_{TM} \gamma(dx, da) \int_{\mathbb{R}^D} g(x, v, \nabla u(x), v) \mu_x(dv) \\
&= \int_{TM} \gamma(dx, da) \int_{\mathbb{R}^D} g(x, v, a, v) \mu_x(dv) \\
&= \int_{TM \times TM} g(x, v, a, b) G^{\mu, \gamma}(dx, dv, da, db). \quad \square
\end{aligned} \tag{5.7}$$

**Definition 5.4.** Let  $(\mathcal{S}, \text{dist})$  be a metric space and let  $\phi : \mathcal{S} \rightarrow [-\infty, \infty]$ . If  $v \in D(\phi)$ , we define the global (metric) slope of  $\phi$  at  $v$  to be

$$|\partial\phi|(v) = \limsup_{w \rightarrow v} \frac{(\phi(v) - \phi(w))^+}{\text{dist}(w, v)}.$$

**Lemma 5.5.** Let  $\mu \in \mathcal{P}_2(TM)$  and let  $\pi_{\#}^1 \mu = \varrho$ . We have

$$|\partial\Phi|(\mu) = |\partial\phi|(\varrho).$$

**Proof.** Lemma 5.3 implies not only the straightforward inequality  $|\partial\Phi|(\mu) \leq |\partial\phi|(\varrho)$ , but in fact, it implies that  $|\partial\Phi|(\mu) = |\partial\phi|(\varrho)$ .  $\square$

**Lemma 5.6.** *Let  $\mu \in \mathcal{P}_2(TM)$ , let  $\pi_{\#}^1 \mu = \varrho$  and let  $(\mu_x)_x$  be the disintegration of  $\mu$  with respect to  $\varrho$ .*

(i) *We have  $\bar{\xi}_1 \in \bar{\partial}\phi(\varrho)$  if*

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \in \bar{\partial}\Phi(\mu) \quad \text{and} \quad \bar{\xi}_1(x) = \int_{\mathbb{R}^d} \xi_1(x, v) \mu_x(dv).$$

(ii) *We have*

$$\|\xi\|_{\mu} \geq \|\xi_1\|_{\varrho}$$

*and the inequality is strict unless  $\xi_2 = 0$   $\mu$ -a.e. and  $\xi_1(x, v)$  is independent on  $v$ .*

(iii) *If  $\bar{\partial}\Phi(\mu) \neq \emptyset$ , then  $\|\nabla_{\mu}\Phi(\mu)\|_{\mu} \geq \|\nabla_{\varrho}\phi(\varrho)\|_{\varrho}$ .*

**Proof.** (i) Let  $\eta \in \mathcal{P}_2(M)$  and let  $\gamma \in \Gamma_o(\varrho, \eta)$ . Suppose  $\xi$  and  $\xi_1$  are as above. By Remark 5.2,  $G^{\mu, \gamma} \in \Gamma_o(\mu, m^{\mu, \gamma})$  and  $\pi_{\#}^1 m^{\mu, \gamma} = \eta$ . Thus, (setting  $w = (x, v)$  and  $z = (a, b)$ )

$$\begin{aligned} \phi(\eta) - \phi(\varrho) &= \Phi(m^{\mu, \gamma}) - \Phi(\mu) \\ &\leq \int_{TM \times TM} \langle \xi(w); z - w \rangle G^{\mu, \eta, \gamma} + o(W_2(\mu, m^{\mu, \eta, \gamma})). \end{aligned} \quad (5.8)$$

By Remark 5.2 (iv),

$$W_2(\mu, m^{\mu, \gamma}) = W_2(\varrho, \eta). \quad (5.9)$$

But

$$\begin{aligned} &\int_{TM \times TM} \langle \xi(w); z - w \rangle G^{\mu, \eta, \gamma} \\ &= \int_{M \times M} \gamma(dx, da) \int_{\mathbb{R}^d} \langle \xi(w); \begin{pmatrix} a - x \\ 0 \end{pmatrix} \rangle \mu_x(dv). \end{aligned} \quad (5.10)$$

We combine (5.8), (5.9) and (5.10) to conclude that

$$\phi(\eta) - \phi(\varrho) \leq \int_{M \times M} \langle \bar{\xi}_1(x); a - x \rangle \gamma(dx, da) + o(W_2(\varrho, \eta)),$$

which proves (i).

(ii) Note that

$$\begin{aligned} \|\xi\|_\mu^2 &= \int_{TM} (|\xi_1|^2 + |\xi_2|^2) \mu(dx, dv) \\ &\geq \int_{TM} |\xi_1|^2 \mu(dx, dv) = \int_M \varrho(dx) \int_{\mathbb{R}^d} |\xi_1|^2 \mu_x(dv), \end{aligned}$$

and equality holds if and only if  $\|\xi_2\|_\mu = 0$ . Hence, by Jensen's inequality

$$\|\xi\|_\mu^2 \geq \int_M \varrho(dx) \left| \int_{\mathbb{R}^d} \xi_1 \mu_x(dv) \right|^2 = \|\xi_1\|_\varrho^2.$$

The inequality is strict unless for  $\rho$  a.e.  $x$  we have  $\xi_1(x, v) = \bar{\xi}_1(x)$  for a.e.  $v$ .

(iii) Follows from (i) and (ii).  $\square$

**Remark 5.7.** Let  $\mu \in \mathcal{P}_2(TM)$  and let  $\varrho = \pi_{\#}^1 \mu$ . Let  $G \in \Gamma_o(\mu, \mu^\tau)$  and recall that  $G_\mu^{\mu^\tau}$  is its barycentric projection onto  $\mu$ .

(i) We have  $\Phi_\tau(\mu) = \phi_\tau(\varrho)$  and  $\varrho^\tau := \pi_{\#}^1 \mu^\tau$ .

(ii) We have  $\pi^2(G_\mu^{\mu^\tau}(x, v)) \equiv v$   $\mu$  a.e., with  $\pi^2$  defined by  $\pi^2(x, v) = v$  for  $(x, v) \in TM$ .

**Proof.** (i) follows from Lemma 5.6.

Assume  $G \in \Gamma_o(\mu, \mu^\tau)$  and let  $A \in C_c(TM, \mathbb{R}^d)$  be arbitrary. We exploit Lemma 5.6 (iii) which asserts that  $G$  is supported by  $S$  to obtain

$$\begin{aligned} &\int_{TM} \langle v - \pi^2(G_\mu^{\mu^\tau}(x, v)); A(x, v) \rangle \mu(dx, dv) \\ &= \int_{TM \times TM} \langle v - b; A(x, v) \rangle G(dx, dv, da, db) = 0. \end{aligned} \quad \square$$

**Proposition 5.8.** Assume that  $D(\phi) \subset \mathcal{P}_2^r(M)$  and that  $\phi$  is convex for the  $L^1$ -metric. Let  $\mu \in \mathcal{P}_2(TM)$  and  $\varrho = \pi_{\#}^1 \mu$  be such that  $J_\tau^\phi(\varrho)$  contains a unique element,  $\varrho^\tau$ , which then belongs to  $\mathcal{P}_2^r(M)$ . Denote by  $\gamma$  the unique element of  $\Gamma_o(\varrho, \varrho^\tau)$  and let  $G \in \Gamma_o(\mu, \mu^\tau)$ .

(i) If  $\xi \in \bar{\partial}\phi_\tau(\varrho)$  and  $\gamma_\varrho^{\varrho^\tau}$  denotes the barycentric projection of  $\gamma$  onto  $\varrho$ , then

$$\pi_\varrho(\xi) = \frac{\mathbf{id} - \gamma_\varrho^{\varrho^\tau}}{\tau}.$$

- (ii) Further assume that  $\varrho \ll \mathcal{L}^d$  and let  $u : M \rightarrow (-\infty, \infty]$  be a lower semicontinuous convex function such that  $(\nabla u)_{\#} \varrho = \varrho^\tau$ . If  $X \in \bar{\partial} \Phi_\tau(\mu)$ , then

$$\pi_\mu(X) = \frac{\mathbf{id} - G_\mu^{\mu^\tau}}{\tau} = \begin{pmatrix} \frac{\mathbf{id} - \nabla u}{\tau} \\ 0 \end{pmatrix}.$$

- (iii) As a consequence, if  $\varrho \ll \mathcal{L}^D$ , then

$$\nabla_\varrho \phi_\tau(\varrho) = \frac{\mathbf{id} - \nabla u}{\tau}, \quad \text{and} \quad \nabla_\mu \Phi_\tau(\mu) = \begin{pmatrix} \nabla_\varrho \phi_\tau(\varrho) \\ 0 \end{pmatrix}.$$

Furthermore,  $J_\tau^\Phi(\mu) = \{m^{\mu, \gamma}\}$ .

**Proof.** (i) Applying Lemma 4.2 to  $\phi_\tau$ , we have  $\bar{\partial} \phi_\tau(\varrho) \neq \emptyset$ . For  $U \in C_c^\infty(M)$  and for  $s \in \mathbb{R}$ , we define

$$\mathbf{g}_s := \mathbf{id} + s \nabla U, \quad \text{and} \quad \varrho_s := \mathbf{g}_s \# \varrho.$$

Observe that for  $|s|$  small enough,  $\mathbf{g}_s$  is the gradient of a convex function and therefore, it is optimal among the maps that push  $\varrho$  forward to  $\varrho_s$ , where optimality is measured against the cost  $c(x, a) = |x - a|^2$  where  $x, a \in M$ . Hence,

$$\beta_s := (\mathbf{id} \times \mathbf{g}_s)_{\#} \varrho \in \Gamma_o(\varrho, \varrho_s).$$

By the fact that  $\varrho_s^\tau \in J_\tau^\phi(\varrho_s)$  we have

$$\phi_\tau(\varrho_s) - \phi_\tau(\varrho) \geq \frac{1}{2\tau} \left( W_2^2(\varrho_s, \varrho_s^\tau) - W_2^2(\varrho, \varrho_s^\tau) \right). \quad (5.11)$$

By the fact that  $\xi \in \bar{\partial} \phi_\tau(\varrho)$ , there exists a function  $\bar{\epsilon} : \mathbb{R} \rightarrow \mathbb{R}$  such that  $\lim_{t \rightarrow 0} \bar{\epsilon}(t) = 0$  and

$$\phi_\tau(\varrho_s) - \phi_\tau(\varrho) \leq W_2(\varrho, \varrho_s) \bar{\epsilon}(W_2(\varrho, \varrho_s)) + \int_{M \times M} \langle \xi(x); a - x \rangle \beta_s(dx, da).$$

This, together with (5.11), implies

$$\begin{aligned} & \frac{1}{2\tau} \left( W_2^2(\varrho_s, \varrho_s^\tau) - W_2^2(\varrho, \varrho_s^\tau) \right) \\ & \leq W_2(\varrho, \varrho_s) \bar{\epsilon}(W_2(\varrho, \varrho_s)) + \int_{M \times M} \langle \xi(x); a - x \rangle \beta_s(dx, da). \end{aligned} \quad (5.12)$$

Let  $\gamma_s \in \Gamma_o(\varrho_s, \varrho_s^\tau)$  and define on  $M \times M$  the Borel probability measure  $\bar{\gamma}_s$  by

$$\int_{M \times M} l(x, a) \bar{\gamma}_s(dx, da) = \int_{M \times M} F(\mathbf{g}_s^{-1}(a), y) \gamma_s(da, dy) \quad \forall l \in C_b(M \times M).$$

We have

$$\mathbf{g}_s^{-1}(a) = a - s \nabla U(a) + \frac{s^2}{2} \nabla^2 U(a) \nabla U(a) + o(s^2)$$

and  $\bar{\gamma}_s \in \Gamma(\varrho, \varrho_s^\tau)$ . Thus,

$$\begin{aligned} & W_2^2(\varrho_s, \varrho_s^\tau) - W_2^2(\varrho, \varrho_s^\tau) \\ & \geq \int_{M \times M} |a - y|^2 \gamma_s(dy, da) - \int_{M \times M} |a - x|^2 \bar{\gamma}_s(dx, da) \\ & = \int_{M \times M} (|a - y|^2 - |a - \mathbf{g}_s^{-1}(y)|^2) \gamma_s(dy, da) \\ & = 2s \int_{M \times M} \langle y - a; \nabla U(y) \rangle \gamma_s(dx, da) + o(s). \end{aligned} \quad (5.13)$$

Recall that for  $|s|$  small enough,  $\beta_s \in \Gamma_o(\varrho, \varrho_s)$  and hence,

$$W_2^2(\varrho, \varrho_s) = \int_{M \times M} |x - y|^2 \beta_s(dx, dy) = \|s \nabla U\|_{\varrho}^2. \quad (5.14)$$

We combine (5.12), (5.13) and (5.14) to obtain

$$\begin{aligned} & \frac{o(s)}{s} + \int_{M \times M} \left\langle \frac{y - a}{\tau}; \nabla U(y) \right\rangle \gamma_s(da, dy) \\ & \leq \|\nabla U\|_{\mu_0} \bar{\epsilon} (\|s \nabla U\|_{\mu_0}) + \int_M \langle \xi(x); \nabla U(x) \rangle \varrho(dx). \end{aligned}$$

Letting  $s \rightarrow 0$  we conclude that

$$\liminf_{s \rightarrow 0^+} \int_{M \times M} \left\langle \frac{y - a}{\tau}; \nabla U(y) \right\rangle \gamma_s(da, dy) \leq \int_M \langle \xi(x); \nabla U(x) \rangle \varrho(dx). \quad (5.15)$$

Observe that

$$\sup_{|s| \leq 1} W_2^2(\varrho_s, \delta_0) \leq \sup_{|s| \leq 1} \int_M |x + s \nabla U(x)|^2 \varrho(dx) < \infty.$$

This, together with Lemma 4.2 (ii), implies

$$\sup_{|s| \leq 1} W_2^2(\varrho_s^\tau, \delta_0) < \infty.$$

Thus,

$$\sup_{|s| \leq 1} W_2^2(\gamma_s, \delta_{(0,0)}) < \infty. \quad (5.16)$$

By Lemma 4.5, as  $s$  tends to 0,  $(\gamma_s)_s$  converges narrowly to the unique element  $\gamma \in \Gamma_o(\varrho, \varrho^\tau)$ . Since (5.16) holds and  $|\frac{a-x}{\tau}; \nabla U(x)|$  grows at most linearly as  $|x|$  and  $|a|$  tend to  $\infty$ , we conclude that

$$\liminf_{s \rightarrow 0} \int_{M \times M} \left\langle \frac{y-a}{\tau}; \nabla U(y) \right\rangle \gamma_s(da, dy) = \int_{M \times M} \left\langle \frac{y-a}{\tau}; \nabla U(x) \right\rangle \gamma(da, dy).$$

This, together with (5.15), yields

$$\int_{M \times M} \left\langle \frac{y-a}{\tau}; \nabla U(y) \right\rangle \gamma(da, dy) \leq \int_M \langle \xi(x); \nabla U(x) \rangle \varrho(dx).$$

Replacing  $U$  by  $-U$  we conclude that

$$\begin{aligned} \int_M \langle \xi(x); \nabla U(x) \rangle \varrho(dx) &= \int_{M \times M} \left\langle \frac{y-a}{\tau}; \nabla U(y) \right\rangle \gamma(da, dy) \\ &= \int_M \left\langle \frac{y - \gamma_\varrho^{\varrho^\tau}(y)}{\tau}; \nabla U(y) \right\rangle \varrho(dy). \end{aligned}$$

As a consequence,

$$\pi_\varrho(\xi) = \pi_\varrho \left( \frac{\mathbf{id} - \gamma_\varrho^{\varrho^\tau}}{\tau} \right) = \frac{\mathbf{id} - \gamma_\varrho^{\varrho^\tau}}{\tau},$$

since by Theorems 8.5.5 and 12.4.4 [3], we know that  $\gamma_\varrho^{\varrho^\tau} - \mathbf{id} \in T_\varrho \mathcal{P}_2(M)$ .

(ii) Further assume that  $\varrho \ll \mathcal{L}^d$ . Then as observed in Remark 4.1,  $J_\tau^\varrho = \{\varrho^\tau\}$  reduces to a single point such that  $\varrho^\tau \ll \mathcal{L}^d$ . Thus,  $\Gamma_o(\varrho, \varrho^\tau) = \{\gamma\}$  also reduces to a single point and  $\gamma = (\mathbf{id} \times \nabla u)_\# \varrho$  for a lower semi-continuous convex function,  $u : M \rightarrow (-\infty, \infty]$ . By Lemma 5.3

$$J_\tau^\Phi(\mu) = \{m^{\mu, \gamma}\}.$$

That uniqueness result is all we need to repeat the same arguments as in (i) to conclude the first identity in (ii). Remark 5.7 asserts that

$$\pi^2(G_\mu^{\mu^\tau}(x, v)) \equiv v$$

while by Lemma 5.3

$$\Gamma_o(\mu, \bar{m}) = \{G^{\mu, \gamma}\}.$$

Thus, if  $A \in C_c(TM)$  is arbitrary, denoting by  $G_\mu^{\mu^\tau}$  the barycentric projection of  $G^{\mu, \gamma}$  onto  $\mu$ , we have

$$\begin{aligned} & \int_{TM} \langle A(x, v); \pi^1(G_\mu^{\mu^\tau}(x, v)) \rangle \mu(dx, dv) \\ &= \int_{TM \times TM} \langle A(x, v); a \rangle G^{\mu, \gamma}(dx, dv, da, dv). \end{aligned}$$

Using the fact that  $\gamma = (\mathbf{id} \times \nabla u)_\# \varrho$ , we conclude that

$$\begin{aligned} & \int_{TM} \langle A(x, v); \pi^1(G_\mu^{\mu^\tau}(x, v)) \rangle \mu(dx, dv) \\ &= \int_{TM \times TM} \langle A(x, v); \nabla u(x) \rangle G^{\mu, \gamma}(dx, dv, da, dv) \quad (5.17) \\ &= \int_{TM} \langle A(x, v); \nabla u(x) \rangle \mu(dx, dv). \end{aligned}$$

Therefore,

$$\pi^1(G_\mu^{\mu^\tau}(x, v)) = \nabla u(x) = \gamma_\varrho^{\varrho^\tau} \quad \mu \text{ a.e.}$$

In light of (i),  $(\mathbf{id} - \gamma_\varrho^{\varrho^\tau})/\tau$  is the element of minimal norm in  $\bar{\partial}\phi_\tau(\varrho)$ ; hence, the first identity in (iii) holds. Similarly, we use (ii) to obtain the second identity in (iii). Since  $J_\tau^\varphi(\varrho)$  contains only  $\varrho^\tau$ , we use Lemma 5.3 (iv) to conclude that  $J_\tau^\Phi(\varrho)$  contains only  $m^{\mu, \gamma}$ .  $\square$

## §6. SOLUTIONS TO AN APPROXIMATE HAMILTONIAN SYSTEMS IN THE PERIODIC SETTING

To avoid technical issues, in this section, we shall study an approximative version of the kinetic Bohmian equation (1.1) on  $\mathbb{T}^d \times \mathbb{R}^d$  instead of  $\mathbb{R}^d \times \mathbb{R}^d$ . In the sequel, we set

$$M := \mathbb{T}^D,$$



and fix a function  $V \in C^2(M)$ . The function  $\mathcal{F}$ , defined in (2.6) (or equivalently in (3.5)) as  $1/8$  times the Fisher information, will be used in this section. For  $\mu \in \mathcal{P}_2(TM)$ , we define the function

$$\mathcal{H}(\mu) = M_2^1(\mu) + \Phi(\mu) + \mathcal{V}(\mu)$$

where

$$\begin{aligned} \mathcal{V}(\mu) &\equiv \mathcal{V}(\pi_{\#}^1 \mu) := \int_{TM} V(x) \mu(dx, dv), \\ \Phi(\mu) &:= \phi(\pi_{\#}^1 \mu), \\ M_2^1(\mu) &= \int_{TM} \frac{|v|^2}{2} \mu(dx, dv), \end{aligned}$$

and

$$\phi := \mathcal{F}.$$

Fix  $\tau > 0$  and recall that if  $\varrho \ll \mathcal{L}^d$  we denote by  $\varrho^\tau$  the unique measure satisfying

$$\phi_\tau(\varrho) = \phi(\varrho^\tau) + \frac{W_2^2(\varrho, \varrho^\tau)}{2\tau}.$$

Similarly, Lemma 5.3 ensures that there is a unique  $\mu^\tau \in \mathcal{P}_2(TM)$  such that

$$\Phi_\tau(\mu) = \Phi(\mu^\tau) + \frac{W_2^2(\mu, \mu^\tau)}{2\tau}.$$

We set

$$\mathcal{H}_\tau(\mu) = M_2^1(\mu) + \Phi_\tau(\mu) + \mathcal{V}(\mu).$$

**Lemma 6.1.** *Let  $\mu \in \mathcal{P}_2^r(TM)$  and assume that  $\varrho := \pi_{\#}^1 \mu \ll \mathcal{L}^D$ . Then,*

$$\nabla_\mu \mathcal{H}_\tau(\mu)(x, v) = \begin{pmatrix} \nabla V(x) + \frac{\mathbf{t}_\varrho^{\varrho^\tau}(x) - x}{\tau} \\ v \end{pmatrix} =: \mathbf{H}(x, v), \quad (6.1)$$

where  $\mathbf{t}_\varrho^{\varrho^\tau}$  is the optimal map that pushes  $\varrho$  forward to  $\varrho^\tau$

**Proof.** By Proposition 5.8 ,

$$\nabla_\mu \Phi_\tau(\mu)(x, v) = \begin{pmatrix} \frac{\mathbf{t}_\varrho^{\varrho^\tau}(x) - x}{\tau} \\ 0 \end{pmatrix}.$$

Since

$$\nabla_{\mu} M_2^1(\mu) \equiv \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad \nabla_{\mu} \mathcal{V}(\mu) \equiv \begin{pmatrix} \nabla V(x) \\ 0 \end{pmatrix} \quad \forall (x, v) \in TM,$$

and

$$\nabla_{\mu} M_2^1(\mu) \in \bar{\partial} M_2^1(\mu) \cap \partial M_2^1(\mu) \quad \text{and} \quad \nabla_{\mu} \mathcal{V}(\mu) \in \bar{\partial} \mathcal{V}(\mu) \cap \partial \mathcal{V}(\mu)$$

we conclude that if  $Z \in \bar{\partial} \mathcal{H}_{\tau}(\mu)$ , then

$$Z - \nabla_{\mu} M_2^1(\mu) - \nabla_{\mu} \mathcal{V} \in \bar{\partial} \Phi_{\tau}(\mu).$$

Furthermore, by Proposition 5.8

$$\nabla_{\mu} \Phi_{\tau}(\mu) = \pi_{\mu}(Z - \nabla_{\mu} M_2^1(\mu) - \nabla_{\mu} \mathcal{V}) = \pi_{\mu}(Z) - \nabla_{\mu} M_2^1(\mu) - \nabla_{\mu} \mathcal{V}.$$

In particular, setting  $Z := \nabla_{\mu} \mathcal{H}_{\tau}(\mu)$ , we conclude the proof.  $\square$

**Theorem 6.2.** *Let  $\mu_0 = f_0 \mathcal{L}^{2D} \in \mathcal{P}_2^r(TM)$  and let  $\tau > 0$ .*

- (i) *There exists a path  $t \rightarrow \bar{\mu}_t^{\tau}$  such that for each  $T > 0$  we have  $\bar{\mu}^{\tau} \in AC_2(0, T; \mathcal{P}_2(TM))$  and*

$$\partial_t \bar{\mu}^{\tau} + \nabla \cdot \left( \bar{\mu}^{\tau} J \nabla_{\mu} \mathcal{H}_{\tau}(\bar{\mu}^{\tau}) \right) = 0 \quad \mathcal{D}'((0, T) \times TM).$$

- (ii) *We have  $\bar{\mu}_t^{\tau} \ll \mathcal{L}^{2D}$  for all  $t > 0$ .*  
 (iii) *Given  $r \rightarrow M_r \in (0, \infty)$  there exists  $r \rightarrow L_r \in (0, \infty)$  such that*

$$f_0 \leq M_r \text{ on } B_r(0) \quad \implies \quad \frac{d\bar{\mu}_t^{\tau}}{d\mathcal{L}^{2D}} \leq L_r \text{ on } B_r(0)$$

- (iv) *Given  $r \rightarrow m_r \in (0, \infty)$  there exists  $r \rightarrow l_r \in (0, \infty)$  (depending on  $\tau$ ) such that*

$$f_0 \geq m_r \text{ on } B_r(0) \quad \implies \quad \frac{d\bar{\mu}_t^{\tau}}{d\mathcal{L}^{2D}} \geq l_r \text{ on } B_r(0)$$

- (v) *We have  $\mathcal{H}_{\tau}(\bar{\mu}_t^{\tau}) = \mathcal{H}_{\tau}(\mu_0)$ .*

**Proof. 1.** Let  $\mu_0 \in \mathcal{P}_2^r(TM)$  and set  $\varrho_0 := \pi_{\#}^1 \mu_0$ . Similarly, for any arbitrary  $\mu \in \mathcal{P}_2^r(TM)$  we set  $\varrho := \pi_{\#}^1 \mu$ . Recall that  $\mathbf{t}_{\varrho}^{\varrho^{\tau}}$  is the optimal map that pushes  $\varrho$  forward to  $\varrho^{\tau}$ . Since  $\mathbf{t}_{\varrho}^{\varrho^{\tau}} : M \rightarrow M$  and  $M$  is a bounded set, Lemma 6.1 supplies us with a constant  $C$  depending on  $\tau$ , but independent of  $\mu$ , such that

$$\left| \nabla_{\mu} \mathcal{H}_{\tau}(\mu)(x, v) \right| \leq C(|(x, v)| + 1), \quad \forall (x, v) \in TM. \quad (6.2)$$

This is referred to as assumption (H1) in [2].

Assume  $(\mu_n)_n \subset \mathcal{P}_2(TM)$  is a sequence of absolutely continuous measures which converges narrowly to  $\mu \ll \mathcal{L}^{2d}$ . Then  $(\mu_n)_n$  is bounded in  $\mathcal{P}_2(TM)$  for the Wasserstein metric and  $(\varrho_n) := (\pi_{\#}^1 \mu_n)_n$  is a sequence of absolutely continuous measures that converges narrowly to  $\varrho \ll \mathcal{L}^d$ . Let  $u_n : \mathbb{R}^d \rightarrow \mathbb{R}$  be convex functions such that  $x \rightarrow u(x) - |x|^2/2$  is convex,  $u_n(0) = 0$  and  $\nabla u_n = \mathbf{t}_{\varrho_n}^{\tau}$ . By Remark 4.1 both  $J_{\tau}^{\Phi}(\varrho) = \{\varrho^{\tau}\}$  and  $J_{\tau}^{\Phi}(\mu) = \{\mu^{\tau}\}$  are of cardinality 1. By Lemma 4.5,  $(\varrho_n^{\tau})_n$  converges to  $\varrho^{\tau}$ . Since  $M$  is a compact set,  $(\nabla u_n)_n$  is uniformly bounded on  $M$ . We use the convexity of  $u_n$  to conclude that  $(\nabla u_n)_n$  is pre-compact in  $L^p(M)$  for any  $1 \leq p < \infty$ . Any point of accumulation of  $(\nabla u_n)_n$  in  $L^p(M)$ ,  $\mathbf{t}$ , is an optimal map for the Wasserstein metric,  $W_2$ , among the maps that push  $\varrho$  forward to  $\varrho^{\tau}$ . Since such an optimal map is unique, we conclude that the whole sequence  $(\nabla u_n)_n$  converges to  $\mathbf{t} = \mathbf{t}_{\varrho^{\tau}}^{\tau}$ . Using the expression of  $\nabla_{\mu} \mathcal{H}_{\tau}(\mu_n)$  provided by Lemma 6.1 we conclude that  $(\nabla_{\mu} \mathcal{H}_{\tau}(\mu_n))_n$  converges almost everywhere to  $\nabla_{\mu} \mathcal{H}_{\tau}(\mu)$ . This is referred to as assumption (H2) in [2]. By (H1) and (H2) we obtain (i)–(iv).

**2.** For the conservation of the Hamiltonian, [2] requires the Hamiltonian to be  $\lambda$ -convex. We now check that  $\lambda$ -concavity is sufficient as well.

By Remark 4.3,  $\Phi_{\tau}$  is Lipschitz on bounded subsets of  $\mathcal{P}_2(TM)$ . Since  $\mathcal{V}$  and  $M_2^1$  are also Lipschitz on bounded subsets of  $\mathcal{P}_2(TM)$ , so is  $\mathcal{H}_{\tau} = \Phi_{\tau} + M_2^1 + \mathcal{V}$ . Fix  $T > 0$ . Since  $\mu \in AC_2(0, T; \mathcal{P}_2(TM))$ , we conclude that  $t \rightarrow \mathcal{H}(\bar{\mu}_t^{\tau})$  is Lipschitz on  $[0, T]$ . To show that  $\mathcal{H}(\bar{\mu}_t^{\tau})$  is time independent, it suffices to show that its derivative vanishes almost everywhere.

Let  $W$  be the velocity of minimal norm for the path  $t \rightarrow \bar{\mu}_t^{\tau}$  provided by Theorem 8.3.1 [3]. Since both  $W$  and  $J\nabla_{\mu} \mathcal{H}(\bar{\mu}^{\tau})$  are velocities for  $t \rightarrow \bar{\mu}_t^{\tau}$ , we have

$$\nabla \cdot \left( W - J\nabla_{\mu} \mathcal{H}(\bar{\mu}^{\tau}) \right) = 0 \quad \mathcal{D}' \left( (0, T) \times TM \right).$$

In other words

$$\int_0^T dt \int_{TM} \langle W - J\nabla_{\mu} \mathcal{H}(\bar{\mu}_t^{\tau}); \nabla F \rangle \bar{\mu}_t^{\tau}(dx, dv) = 0 \quad \forall F \in C_0^1((0, T) \times TM).$$

Choosing  $F$  in the form  $F(t, x, v) = A(t)B(x, v)$  and using a density argument, we conclude that for almost every  $t \in (0, T)$  we have

$$\int_{TM} \langle W - J\nabla_{\mu} \mathcal{H}(\bar{\mu}_t^{\tau}); \nabla B \rangle \bar{\mu}_t^{\tau}(dx, dv) = 0 \quad \forall B \in C_0^1(TM).$$

Thus, for almost every  $t \in (0, T)$ ,  $W_t$  is the orthogonal projection of  $J\nabla_\mu \mathcal{H}(\bar{\mu}_t^\tau)$  onto the tangent space  $T_{\bar{\mu}_t^\tau} \mathcal{P}_2(TM)$ :

$$W_t := \pi_{\bar{\mu}_t^\tau} \left( J\nabla_\mu \mathcal{H}(\bar{\mu}_t^\tau) \right).$$

By (8.4.6) [3], for almost every  $t \in (0, T)$ , if  $t + h \in (0, T)$  and  $G_h \in \Gamma_o(\bar{\mu}_t^\tau, \bar{\mu}_{t+h}^\tau)$ , then we have the following convergence in the  $W_2$ -metric:

$$\lim_{h \rightarrow 0} \left( \bar{\pi}^1, \frac{\bar{\pi}^2 - \bar{\pi}^1}{h} \right)_\# G_h = (\mathbf{id} \times W_t)_\# \bar{\mu}_t^\tau. \quad (6.3)$$

Here,

$$\bar{\pi}^1(w, z) = w, \quad \bar{\pi}^2(w, z) = z \quad \forall w := (x, v), z := (a, b) \in TM.$$

Denote by  $|(\bar{\mu}_t^\tau)'|$  the metric derivative of  $t \rightarrow \bar{\mu}_t^\tau$  (cf. e.g. Definition 1.1.1 [3]). By definition

$$\lim_{h \rightarrow 0} \frac{W_2(\bar{\mu}_t^\tau, \bar{\mu}_{t+h}^\tau)}{h} = |(\bar{\mu}_t^\tau)'|(t)$$

for almost every  $t \in (0, T)$ . Hence, for these  $t$ ,

$$\frac{W_2^2(\bar{\mu}_t^\tau, \bar{\mu}_{t+h}^\tau)}{h} = o(h), \quad (6.4)$$

where  $o(h)$  depends on  $t$ . Note that by Lemma 4.2 (ii),  $\Phi_\tau$  is  $\tau^{-1}$ -concave. Since the second derivatives of  $(x, v) \rightarrow V(x)$  and that of  $(x, v) \rightarrow |v|^2$  are bounded, we conclude that there exists a constant  $\bar{C}_\tau$  such that  $\mathcal{H}_\tau$  is  $\bar{C}_\tau$ -concave. Thus,

$$\begin{aligned} & \mathcal{H}_\tau(\bar{\mu}_{t+h}^\tau) - \mathcal{H}_\tau(\bar{\mu}_t^\tau) \\ & \leq \int_{TM \times TM} \langle \nabla_\mu \mathcal{H}(\bar{\mu}_t^\tau)(w); z - w \rangle G_h(dw, dz) + \bar{C}_\tau W_2^2(\bar{\mu}_t^\tau, \bar{\mu}_{t+h}^\tau). \end{aligned}$$

If  $t$  is such that (6.3) holds, since  $|\langle \nabla_\mu \mathcal{H}_\tau(\bar{\mu}_t^\tau); z - w \rangle|$  grows at most quadratically, we conclude that

$$\begin{aligned} & \mathcal{H}_\tau(\bar{\mu}_{t+h}^\tau) - \mathcal{H}_\tau(\bar{\mu}_t^\tau) \\ & \leq \int_{TM} h \langle \nabla_\mu \mathcal{H}_\tau(\bar{\mu}_t^\tau)(w); W_t(w) \rangle \bar{\mu}_t^\tau(dw) + \bar{C}_\tau W_2^2(\bar{\mu}_t^\tau, \bar{\mu}_{t+h}^\tau) + o(h). \quad (6.5) \end{aligned}$$

We use the fact that  $W_t$  is the projection of  $J\nabla_\mu \mathcal{H}(\bar{\mu}_t^\tau)$  onto  $T_{\bar{\mu}_t^\tau} \mathcal{P}_2(TM)$  to conclude that

$$\begin{aligned} & \int_{TM} \langle \nabla_\mu \mathcal{H}(\bar{\mu}_t^\tau)(w); W_t(w) \rangle \bar{\mu}_t^\tau(dw) \\ &= \int_{TM} \langle \nabla_\mu \mathcal{H}(\bar{\mu}_t^\tau)(w); J\nabla_\mu \mathcal{H}(\bar{\mu}_t^\tau)(w) \rangle \bar{\mu}_t^\tau(dw) = 0. \end{aligned}$$

This, together with (6.4) and (6.5), implies

$$\mathcal{H}(\bar{\mu}_{t+h}^\tau) - \mathcal{H}(\bar{\mu}_t^\tau) \leq o(h). \quad (6.6)$$

The map  $t \rightarrow \mathcal{H}(\bar{\mu}_t^\tau)$  is Lipschitz on  $[0, T]$ . Therefore, it is differentiable almost everywhere. If  $t$  is a point of differentiability, using alternatively  $h > 0$  and  $h < 0$  in (6.6), we conclude that

$$\frac{d}{ds} \mathcal{H}(\bar{\mu}_s^\tau)|_{s=t} = 0.$$

Since the derivative of the Lipschitz function  $t \rightarrow \mathcal{H}(\bar{\mu}_t^\tau)$  vanishes almost everywhere, the function must be constant.  $\square$

**Remark 6.3.** If we replace  $\mathbb{T}^d$  by  $\mathbb{R}^d$  then, because of Remark 4.3, (H1') of [2] holds. [2] ensures that if (H2') also holds, then there is a solution to our Hamiltonian system. The proof of (H2') requires some effort and this is why we worked on  $\mathbb{T}^d$ . Note that the above arguments go through if we replace  $\mathbb{T}^d$  by any open bounded set.

## §7. INGREDIENTS TOWARD A CONVERGENCE ANALYSIS IN THE PERIODIC SETTING

Let  $\mu_0 = f_0 \mathcal{L}^{2d} \in \mathcal{P}_2^f(TM)$  and let  $T > 0$ . For  $\tau > 0$  we define  $t \rightarrow \bar{\mu}_t^\tau \in \mathcal{P}_2^f(TM)$  as in Theorem 6.2. Write

$$\begin{aligned} \bar{\mu}_t^\tau &= \bar{f}_t^\tau \mathcal{L}^{2d}, \quad \pi_{1\#} \bar{\mu}_t^\tau = \bar{g}_t^\tau \mathcal{L}^d, \\ \bar{f}_t^\tau(x, v) &= \bar{g}_t^\tau(x) \bar{F}_t^\tau(x, v), \quad \text{with} \quad \int_{\mathbb{R}^d} \bar{F}_t^\tau(x, v) dv = 1. \end{aligned}$$

**7.1. Continuity equation.** Since  $f_0 \in L^1(TM)$ , we apply de la Vallée Poussin Theorem to  $\{f_0\}$ , a compact subset of  $L^1(TM)$ , to conclude that there exists a super linear convex function  $\theta : [0, \infty) \rightarrow [0, \infty)$  such that  $\theta(f_0) \in L^1(TM)$ . We use Lemma 6.2 [2] to conclude that

$$\sup_{t \in [0, T]} \int_{TM} \theta(\bar{f}_t) dx dv \leq \int_{TM} \theta(\bar{f}_0) dx dv < \infty. \quad (7.1)$$

We apply again de la Vallée Poussin Theorem to conclude that  $\{\bar{f}^\tau \mid \tau > 0\}$  is a compact subset of  $L^1((0, T) \times TM)$ .

Recall that since  $\pi_{1\#} \bar{\mu}_t^\tau \ll \mathcal{L}^d$ ,  $J_\tau^\phi(\bar{\varrho}_t)$  reduces to a single element  $\varrho_t^\tau \mathcal{L}^d$ . We have

$$\phi_\tau(\bar{\varrho}_t^\tau) = \phi(\varrho_t^\tau) + \frac{W_2^2(\varrho_t^\tau, \bar{\varrho}_t^\tau)}{2\tau}. \quad (7.2)$$

By Theorem 6.2 (v)

$$\phi_\tau(\bar{\varrho}_t^\tau) + \int_M V(x) \bar{\varrho}_t^\tau(x) dx + \frac{1}{2} \int_{TM} |v|^2 \bar{\mu}_t^\tau(dx, dv) = \mathcal{H}_\tau(\mu_0) \leq \mathcal{H}(\mu_0). \quad (7.3)$$

By Proposition 5.8

$$\|\nabla_\varrho \phi_\tau(\bar{\varrho}_t^\tau)\|_{\bar{\varrho}_t^\tau} = \frac{W_2(\varrho_t^\tau, \bar{\varrho}_t^\tau)}{\tau}. \quad (7.4)$$

This, together with (7.3), yields

$$\frac{\tau}{2} \|\nabla_\varrho \phi_\tau(\bar{\varrho}_t^\tau)\|_{\bar{\varrho}_t^\tau}^2 + \phi(\varrho_t^\tau) \leq \mathcal{H}(\mu_0) + \|V\|_\infty. \quad (7.5)$$

Define

$$\bar{\mathbf{u}}_t^\tau(x) := \int_{\mathbb{R}^d} v \bar{F}_t^\tau(x, v) dv.$$

We use (7.1) to deduce that up to a subsequence,  $(\bar{f}^\tau)_\tau$  converges weakly to some  $\bar{f}$  in  $L^1((0, 1) \times TM)$ .

**Proposition 7.1.** *The following hold:*

(i)  $\bar{\varrho}^\tau \in AC_2(0, T; \mathcal{P}(M))$ .

(ii)

$$\frac{1}{2} \int_M |\bar{\mathbf{u}}_t^\tau(x)|^2 \bar{\varrho}_t^\tau(x) dx \leq \mathcal{H}(\mu_0) + \|V\|_\infty.$$

(iii)

$$\partial_t \bar{\varrho}^\tau + \nabla \cdot (\bar{\varrho}^\tau \bar{\mathbf{u}}^\tau) = 0 \quad \mathcal{D}'((0, T) \times M).$$

**Proof.** (i) We use that  $\pi_1$  is a contraction of  $(\mathcal{P}_2(TM), W_2)$  into  $(\mathcal{P}_2(M), W_2)$  and use the fact that  $\bar{\mu}^\tau \in AC_2(0, T; \mathcal{P}_2(TM))$  to conclude the proof of (i).

(ii) We use Jensen's inequality to deduce that

$$\int_M |\bar{\mathbf{u}}_t^\tau(x)|^2 \bar{\varrho}_t^\tau(x) dx \leq \int_{TM} |v|^2 \bar{\mu}_t^\tau(dx, dv),$$

which, together with (7.3), yields (ii).

(iii) The differential equation in Theorem 6.2 yields (iii).  $\square$

**7.2. Convergence in  $\mathcal{P}(M)$ .** The goal of this subsection is to establish some convergence results. We prove that the paths  $\bar{\varrho}^{\tau_n}$  and  $\varrho^{\tau_n}$  converge to the same limit  $\bar{\varrho}$ . Setting

$$\bar{\nu}^{\tau_n} := \mathcal{L}_{(0,T)}^1 \otimes \bar{\mu}^{\tau_n}$$

we show that for the narrow convergence topology,  $(\nu^{\tau_n})_n$  contains points of accumulation of the form  $\mathcal{L}_{(0,T)}^1 \otimes \bar{\mu}_t$  where  $\bar{\varrho}_t \mathcal{L}^d$  is the projection of  $\bar{\mu}_t$  onto  $M$ .

**Proposition 7.2.** *There exists a sequence  $(\tau_n)_n$  decreasing to 0 such that the following hold:*

- (i) *For any  $t \in (0, T)$ ,  $(\bar{\varrho}_t^{\tau_n})_n$  converges in  $\mathcal{P}(M)$  to  $\bar{\varrho}_t$ .*
- (ii) *For any  $t \in (0, T)$ ,  $(\varrho_t^{\tau_n})_n$  converges in  $\mathcal{P}(M)$  to  $\bar{\varrho}_t$ .*
- (iii) *We have  $\sup_{t \in (0, T)} \phi(\varrho_t) < \infty$ .*
- (iv)  *$(\nu^{\tau_n})_n$  converges narrowly on  $[0, T] \times TM$  to some  $\nu = \mathcal{L}_{(0,T)}^1 \otimes \bar{\mu}_t$ .*
- (v) *We have  $\bar{\mu}_t(TM) = 1$  for  $\mathcal{L}^1$ -a.e.  $t \in (0, T)$ .*
- (vi) *We have  $\pi_{\#}^1 \bar{\mu}_t = \bar{\varrho}_t \mathcal{L}^d$  for  $\mathcal{L}^1$ -a.e.  $t \in (0, T)$ .*

**Proof.** Recall that  $\mathcal{H}_\tau \leq \mathcal{H}$ . Therefore, using Theorem 6.2 (v) we have

$$\mathcal{H}_\tau(\bar{\mu}_t^\tau) = \mathcal{H}_\tau(\mu_0) \leq \mathcal{H}(\mu_0). \tag{7.6}$$

(i) By Proposition 7.1,  $|(\bar{\varrho}^\tau)'|^2 \leq 2\mathcal{H}_\tau(\mu_0) + 2\|V\|_\infty$  except maybe on a set of null measure. Thus,

$$W_2(\bar{\varrho}_t^\tau, \bar{\varrho}_s^\tau) \leq \int_s^t |(\bar{\varrho}^\tau)'(l)| dl \leq |t-s| \sqrt{2(\mathcal{H}(\mu_0) + \|V\|_\infty)}.$$

Now we can apply the Ascoli-Arzelà theorem (see Proposition 3.3.1 [3]) to get (i).

(ii) We exploit (7.6) to get

$$\|V\|_\infty + \mathcal{H}(\mu_0) \geq \phi_\tau(\bar{\varrho}_t^\tau) = \frac{W_2^2(\bar{\varrho}_t^\tau, \varrho_t^\tau)}{2\tau} + \phi(\varrho_t^\tau). \quad (7.7)$$

Hence,

$$W_2^2(\bar{\varrho}_t^\tau, \varrho_t^\tau) \leq 2\tau(\mathcal{H}(\mu_0) + \|V\|_\infty),$$

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

(iii) We use (7.7) and the fact that  $\phi$  is lower semicontinuous for the narrow convergence to conclude that

$$\phi(\varrho_t) \leq \|V\|_\infty + \mathcal{H}(\mu_0).$$

(iv) By (7.3)

$$\begin{aligned} & \int_0^T dt \int_{TM} (1 + |x|^2 + |v|^2) \bar{\mu}_t^\tau(dx, dv) \\ &= \int_0^T \left( 1 + \int_M |x|^2 \bar{\varrho}_t^\tau(dx) \right) dt + \int_0^1 \int_{TM} |v|^2 \bar{\mu}_t^\tau(dx, dv) \\ &\leq 1 + (\text{diam}M)^2 + 2(\mathcal{H}(\mu_0) + \|V\|_\infty). \end{aligned}$$

Hence,  $(\bar{\nu}^{\tau_n})_n$  is pre-compact for the narrow convergence. Extracting a subsequence if necessary, we obtain a Borel measure  $\nu$  on  $[0, 1] \times TM$  such that  $(\bar{\nu}^{\tau_n})_n$  converges narrowly on  $[0, T] \times TM$  to  $\nu$ . Since the projection of  $\mathcal{L}_{(0,T)}^1 \bar{\mu}_t^{\tau_n}$  onto  $[0, T]$  is less than 1, the same is true for the projection of  $\nu$  (cf. e.g. Theorem 2.28 [1]). This concludes the proof of (iv).

(v) Let  $\varphi \in C_b([0, 1])$ . Note that

$$\begin{aligned} \int_0^1 \varphi(t) dt &= \lim_n \int_0^1 \varphi(t) dt \int_{TM} \mu_t^{\tau_n}(dx, dv) \\ &= \lim_n \int_{[0,T] \times TM} \varphi(t) \nu^{\tau_n}(dt, dx, dv). \end{aligned}$$

We use (iv) to deduce that

$$\int_0^1 \varphi(t) dt = \int_0^1 \varphi(t) dt \int_{TM} \bar{\mu}_t(dx, dv).$$



Since  $t \rightarrow \int_{TM} \bar{\mu}_t(dx, dv)$  belongs to  $L^1(0, 1)$ , (v) follows.

(vi) Let  $\varphi \in C_b([0, 1])$  and  $\psi \in C_b(M)$ . We first use (i) and then use (v) to obtain

$$\begin{aligned} \int_0^T \varphi(t) dt \int_M \bar{\varrho}_t(x) \psi(x) dx &= \lim_{\tau_n \rightarrow 0} \int_0^T \varphi(t) dt \int_M \bar{\varrho}_t^{\tau_n}(x) \psi(x) dx \\ &= \lim_n \int_{[0, T] \times TM} \varphi(t) \nu^{\tau_n}(dt, dx, dv). \end{aligned}$$

Thus by (iv),

$$\int_0^T \varphi(t) dt \int_M \bar{\varrho}_t(x) \psi(x) dx = \int_0^T \varphi(t) dt \int_{TM} \psi(x) \bar{\mu}_t(dx, dv),$$

which means that

$$\int_{TM} \psi(x) \bar{\mu}_t(dx, dv) = \int_M \bar{\varrho}_t(x) \psi(x) dx.$$

Since  $\psi \in C_b(M)$  is arbitrary, we conclude the proof of (vi). □

**7.3. Momentum equations for approximate solutions.** Recall that according to Section 6, if  $\mathbf{t}^\tau$  is the unique gradient of a lower semicontinuous convex function such that  $\mathbf{t}_{\#}^\tau \varrho_t^\tau = \bar{\varrho}_t^\tau$ , then

$$\xi^\tau := \frac{\mathbf{t}^\tau - \mathbf{id}}{\tau} \in \partial\phi(\varrho_t^\tau) \quad \text{and} \quad \frac{\mathbf{id} - (\mathbf{t}^\tau)^{-1}}{\tau} \in \bar{\partial}\phi(\bar{\varrho}_t^\tau).$$

Thus, by Proposition 5.8, the Wasserstein gradient of  $\phi_\tau$ ,  $\bar{\varrho}_t^\tau$ , and  $\xi^\tau$  satisfy the relation

$$\xi^\tau = \nabla_\varrho \phi_\tau(\bar{\varrho}_t^\tau) \circ \mathbf{t}^\tau. \tag{7.8}$$

Using  $\bar{F}^\tau$  as introduced at the beginning of the current section, we define the averages

$$\widehat{v \otimes v}^{\bar{\mu}^\tau}(t, x) = \int_{\mathbb{R}^d} v \otimes v \bar{F}_t^\tau(x, v) dv.$$

**Definition 7.3.** Let  $\varrho \in AC_2(0, T; \mathcal{P}(M))$ . Moreover, let

$$\mu \in AC_2(0, T; \mathcal{P}(M))$$

be such that  $\varrho_t$  is the projection of  $\mu_t$  on  $M$  and set

$$\widehat{v \otimes v}^\mu := \int_{\mathbb{R}^d} v \otimes v F_t(x, dv),$$

where  $(F_t(x, \cdot))_x$  is the disintegration of  $\mu_t$ . Assume that  $\xi : (0, T) \times M \rightarrow \mathbb{R}^d$  is a Borel vector field such that  $\xi_t \in L^2(\varrho_t)$  for  $\mathcal{L}^1$ -a.e.  $t \in (0, 1)$ . We say that  $(\varrho, u, \widehat{v \otimes v}^\mu, \xi)$  satisfies the momentum equation

$$\partial_t(\varrho u) + \nabla \cdot (\varrho \widehat{v \otimes v}^\mu) = -\varrho(\nabla V + \xi) \tag{7.9}$$

in the sense of distribution if

$$\int_0^T dt \int_M (\partial_t A + \widehat{v \otimes v}^\mu \nabla A) \varrho_t(dx) = \int_0^T dt \int_M \langle A; \nabla V + \xi \rangle \varrho_t(dx),$$

for all  $A \in C_c^\infty((0, T) \times M; \mathbb{R}^d)$ .

**Remark 7.4.** The following hold:

- (i) If  $\varrho$  belongs to the appropriate Sobolev space, then it is smooth enough such that we can write the Wasserstein gradient of  $\phi$  at  $\varrho$  as

$$\nabla_\varrho \phi(\varrho) = -\frac{1}{2} \nabla \left( \frac{\Delta \sqrt{\varrho}}{\sqrt{\varrho}} \right).$$

Therefore (cf., e.g., [15]),

$$\varrho \nabla_\varrho \phi(\varrho) = \frac{1}{2} \nabla(\Delta \varrho) - \operatorname{div}(\nabla \sqrt{\varrho} \otimes \nabla \sqrt{\varrho}). \tag{7.10}$$

- (ii) Since  $\phi$  is the Fisher information up to a multiplicative constant and  $J_\tau^\phi(\varrho_t^\tau) = \{\varrho_t^\tau\}$ , by Lemma 10.1.2 [3],  $\xi^\tau$  is in the strong sub-differential of  $\phi$ . By Corollary 5.8 [13]

$$\sqrt{\varrho_t^\tau} \in W^{2,2}(M). \tag{7.11}$$

- (iii) If  $A \in C^1((0, T) \times M, \mathbb{R}^d)$ , then we can apply Corollary 5.8 [13] to deduce that (7.10) holds for  $\varrho = \varrho^\tau$  in the sense that

$$\begin{aligned} & \int_0^T dt \int_M \langle A; \xi^\tau \rangle \varrho_t^\tau(x) dx \\ &= \int_0^T dt \int_M \left( -\frac{1}{2} (\nabla \cdot A) \Delta \varrho_t^\tau + \langle \nabla A; \nabla \sqrt{\varrho_t^\tau} \otimes \nabla \sqrt{\varrho_t^\tau} \rangle \right) dx. \end{aligned} \tag{7.12}$$

- (iv) Observe that Proposition 7.2 (iii) alone ensures that, for the limiting measures, we have  $\sqrt{\bar{\varrho}_t^\tau} \in W^{1,2}$  and therefore, the expression on the right-hand side of (7.12) continues to make sense for the limiting densities  $\bar{\varrho}$  obtained in Proposition 7.2; it can be written as

$$\int_0^T dt \int_M \left( -\frac{1}{2} \Delta(\nabla \cdot A) \bar{\varrho}_t + \langle \nabla A; \nabla \sqrt{\bar{\varrho}_t} \otimes \nabla \sqrt{\bar{\varrho}_t} \rangle \right) dx.$$

For any vector valued Borel field,  $\xi$ , on  $M$  of null average, we define the norm

$$\|\xi\|_{-1} = \sup_{A \in C_c^\infty(M; \mathbb{R}^d)} \left\{ \int_M \langle A; \xi(dx) \rangle \mid \|\nabla A\|_\infty \leq 1 \right\}.$$

**Theorem 7.5.** *Using the notation of Subsection 7.1, the following hold:*

- (i)  $(\bar{\varrho}^\tau, \bar{u}^\tau, \widehat{v \otimes v}^{\bar{\mu}^\tau}, \nabla_\varrho \phi_\tau(\bar{\varrho}^\tau))$  satisfies the momentum equation (7.9) in the sense of distributions.  
(ii) In the sense of distributions, as given by Definition 7.3 and (7.12),

$$\begin{aligned} & \partial_t(\bar{\varrho}^\tau \bar{u}^\tau) + \nabla \cdot (\bar{\varrho}^\tau \widehat{v \otimes v}^{\bar{\mu}^\tau}) \\ &= -\bar{\varrho}^\tau \nabla V + \nabla \cdot \left( \frac{1}{2} \nabla(\Delta \varrho^\tau) - \operatorname{div}(\nabla \sqrt{\bar{\varrho}^\tau} \otimes \nabla \sqrt{\bar{\varrho}^\tau}) \right) + \vec{0}^\tau, \end{aligned}$$

where

$$\vec{0}^\tau := \bar{\varrho}^\tau \nabla_\varrho \phi(\bar{\varrho}^\tau) - \varrho^\tau \xi^\tau.$$

- (iii) Further assume that there exists a sequence  $(\tau_n)_n$  decreasing to 0 such that for  $\mathcal{L}^1$  a.e.  $t \in (0, T)$  we have

$$\lim_{n \rightarrow \infty} \phi_{\tau_n}(\bar{\varrho}_t^{\tau_n}) - \phi(\varrho_t^{\tau_n}) = 0. \quad (7.13)$$

Then, for any  $p \in [1, \infty)$  we have

$$\lim_{n \rightarrow \infty} \int_0^T \|\vec{0}_t^{\tau_n}\|_{-1}^p dt = 0.$$

**Proof.** (i) By Theorem 6.2 for any  $L \in C_c^\infty((0, T) \times TM)$  we have

$$\begin{aligned} & \int_0^T dt \int_{TM} (\partial_t L + \langle v; \nabla_x L \rangle) \bar{\mu}_t^\tau(dx, dv) \\ &= \int_0^T dt \int_{TM} \langle \nabla_v L; V + \nabla_\varrho \phi_\tau(\bar{\varrho}^\tau) \rangle \bar{\mu}_t^\tau(dx, dv) = 0. \end{aligned} \quad (7.14)$$

The uniform bound in (7.3) implies that

$$\sup_{t, \tau} \int_{TM} |v|^2 \bar{\mu}_t^\tau(dx, dv) < \infty.$$

Thus, if  $A \in C_c^\infty((0, T) \times M)$ ,  $B_i(v) \equiv v_i$ , since  $B_i$  grows slower than  $|v|^2$  at infinity, by a standard approximation argument, we can use  $L(t, x, v) := B_i(v)A(t, x)$  in (7.14) and read off the proof of (i).

(ii) Applying Remark 7.4 (iii), we obtain in the sense of distributions

$$\nabla \left( \frac{1}{2} \nabla(\Delta \varrho^\tau) - \operatorname{div}(\nabla \sqrt{\varrho}^\tau \otimes \nabla \sqrt{\varrho}^\tau) \right) + \varrho^\tau \xi^\tau = 0.$$

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

(iii) For any  $A \in C_c^\infty(M)$  such that  $\|\nabla A\|_\infty \leq 1$ , we have

$$\int_M \langle \nabla_\varrho \phi_\tau(\bar{\varrho}_t^\tau); A \rangle \bar{\varrho}_t^\tau(dx) = \int_M \langle \nabla_\varrho \phi_\tau(\bar{\varrho}_t^\tau) \circ \mathbf{t}^\tau; A(\mathbf{t}^\tau) \rangle \varrho_t^\tau(x) dx.$$

Thus, using (7.8) we conclude that

$$\begin{aligned} \left| \int_M \langle \bar{\mathbf{0}}_t^\tau; A \rangle dx \right| &= \left| \int_M \langle \xi^\tau; A(\mathbf{t}^\tau) - A(\mathbf{id}) \rangle \varrho_t^\tau(x) dx \right| \\ &\leq \|\xi^\tau\|_{\varrho_t^\tau} \|\mathbf{t}^\tau - \mathbf{id}\|_{\varrho_t^\tau} = \|\xi^\tau\|_{\varrho_t^\tau} W_2(\varrho_t^\tau, \bar{\varrho}_t^\tau). \end{aligned}$$

Since by Remark 4.3 (ii)

$$W_2(\varrho_t^\tau, \bar{\varrho}_t^\tau) \|\xi^\tau\|_{\varrho_t^\tau} \leq \frac{W_2^2(\varrho_t^\tau, \bar{\varrho}_t^\tau)}{\tau} = 2(\phi_\tau(\bar{\varrho}_t^\tau) - \phi(\varrho_t^\tau)),$$

we obtain

$$\left| \int_M \langle \bar{\mathbf{0}}_t^\tau; A \rangle dx \right| \leq 2(\phi_\tau(\bar{\varrho}_t^\tau) - \phi(\varrho_t^\tau)).$$

Hence,

$$\|\bar{\mathbf{0}}_t^\tau\|_{-1} \leq 2(\phi_\tau(\bar{\varrho}_t^\tau) - \phi(\varrho_t^\tau)). \quad (7.15)$$

We use the fact that  $\phi \geq 0$  and (7.7) to obtain for any  $t \in (0, T)$  and  $\tau \in (0, 1)$

$$\phi_\tau(\bar{\varrho}_t^\tau) - \phi(\varrho_t^\tau) \leq \phi_\tau(\bar{\varrho}_t^\tau) \leq \mathcal{H}(\mu_0) + \|V\|_\infty.$$

We can use (7.13) and the Lebesgue dominated convergence theorem to conclude that for any  $p \geq 1$

$$\lim_{n \rightarrow \infty} \int_0^T \|\bar{\varrho}_t^{\tau_n}\|_{-1}^p dt = 0. \quad \square$$

## §8. CONCLUDING REMARKS

It is important to mention that the previous results require the initial condition to be absolutely continuous with respect to the Lebesgue measure, and therefore, the mono-kinetic case presented in the introduction is not covered. It remains an interesting question to determine if our results may be extended to an arbitrary initial measure if we consider the second method proposed in [2].

On the other hand, the convergence analysis needs to be improved in order to verify that the limit of the approximative scheme satisfies the kinetic Bohmian equation in a weak sense. We leave it as an open question for now to investigate if the flow exchange technique introduced in [17] for the analysis of Wasserstein gradient flows may be extended to our problem, giving us the additional estimates that we need to pass to the limit in our approximative scheme.

## REFERENCES

1. L. Ambrosio, N. Fusco, D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*. — Clarendon Press Oxford, **254** (2000).
2. L. Ambrosio, W. Gangbo, *Hamiltonian ODEs in the Wasserstein space of probability measures*. — Commun. Pure Appl. Math., **61**, No. 1 (2008), 18–53.
3. L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science Business Media (2008).
4. L. Ambrosio, S. Lisini, G. Savaré, *Stability of flows associated to gradient vector fields and convergence of iterated transport maps*. — Manuscripta Math., **121**, No. 1 (2006), 1–50.
5. P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, A. Porretta, *Long time average of mean field games with a nonlocal coupling*. — Control and Optimization, SIAM, **51**, No. 5 (2013), 3558–3591.
6. T. Cazenave, *Semilinear Schrödinger Equations*, **10**. — Courant Lecture Notes, Vol. 10, Amer. Math. Soc. (2003).
7. J. Cushing, S. Goldstein, A. Fine (eds.), *Bohmian Mechanics and Quantum Theory: an appraisal*. — Springer Science Business Media, **184** (2013).

8. D. Dèurr, S. Teufel, *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory*. Springer (2009).
9. W. Gangbo, H-K. Kim, T. Pacini, *Differential forms on Wasserstein space and infinite-dimensional Hamiltonian systems*. — *Memoirs of the AMS*, **211** (2011), 1–77.
10. W. Gangbo, T. Nguyen, A. Tudorascu, *Hamilton–Jacobi equations in the Wasserstein space*. — *Methods Appl. Anal.*, **51**, No. 15 (2008), 155–183.
11. W. Gangbo, M. Westdickenberg, *Optimal transport for the system of isentropic Euler equations*. — *Commun. Partial Diff. Eqs.*, **34**, No. 9 (2009), 1041–1073.
12. I. Gasser, P. Markowich, *Quantum hydrodynamics, Wigner transforms, the classical limit*. — *Asympt. Anal.*, **14**, No. 2 (1997), 97–116.
13. U. Gianazza, G. Savaré, G. Toscani, *The Wasserstein gradient flow of the Fisher information and the quantum drift–diffusion equation*. — *Archive for Rational Mechanics and Analysis*, **194**, No. 1 (2009), 133–220.
14. E. Lieb, R. Seiringer, J. Yngvason, *Bosons in a trap: A rigorous derivation of the Gross-Pitaevskii energy functional, The Stability of Matter: From Atoms to Stars*. Springer (2001), 685–697.
15. P. Markowich, T. Paul, C. Sparber, *Bohmian measures and their classical limit*. — *J. Funct. Anal.*, **259**, No. 6 (2010), 1542–1576.
16. P. Markowich, T. Paul, C. Sparber, *On the dynamics of Bohmian measures*. — *Archive for Rational Mechanics and Analysis*, **205**, No. 3 (2012), 1031–1054.
17. D. Matthes, R. McCann, G. Savaré, *A family of nonlinear fourth order equations of gradient flow type*. — *Commun. Partial Diff. Eqs.*, **34**, No. 11 (2009), 1352–1397.
18. F. Otto, *The Geometry of Dissipative Evolution Equations: the Porous Medium Equation*. Taylor and Francis (2001).
19. K. Parthasarathy, *Probability Measures on Metric Spaces*. — *Amer. Math. Soc.*, **352** (1967).
20. C. Sulem, P. Sulem, *The Nonlinear Schrödinger Equation: Self-Focusing and Wave Collapse*. Ser. Applied Mathematical Sciences, Vol. 139, Springer Science Business Media (1999).
21. T. Tao, *Nonlinear Dispersive Equations: Local and Global Analysis*. CBMS Regional Conference Series in Mathematics, Vol. 106, Amer. Math. Soc. (2006).
22. A. W. van der Vaart, J. Wellner, *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer (1996).
23. C. Villani, *Topics in Optimal Transportation*. Ser. Graduate Studies in Math., Vol. 58, Amer. Math. Soc. (2003).
24. F. Santambrogio, *Optimal Transport for Applied Mathematicians*. Birkhäuser, NY Springer (2015).

University of California at Los Angeles,  
Los Angeles, CA 90095, U.S.A.

Поступило 6 марта 2017 г.

*E-mail*: [wgangbo@math.ucla.edu](mailto:wgangbo@math.ucla.edu)

CEMSE Division, King Abdullah  
University of Science and Technology,  
Saudi Arabia