HAL
archives-ouvertes.fr

Weak KAM theory for potential MFG

Pierre Cardaliaguet, Marco Masoero

▶ **To cite this version:**

Pierre Cardaliaguet, Marco Masoero. Weak KAM theory for potential MFG. 2019. hal-02116125

Weak KAM theory for potential MFG

Pierre Cardaliaguet, Marco Masoero *

April 30, 2019

Abstract

We develop the counterpart of weak KAM theory for potential mean field games. This allows to describe the long time behavior of time-dependent potential mean field game systems. Our main result is the existence of a limit, as time tends to infinity, of the value function of an optimal control problem stated in the space of measures. In addition, we show a mean field limit for the ergodic constant associated with the corresponding Hamilton-Jacobi equation.

Introduction

The theory of mean field games (MFG), introduced simultaneously and independently by Lasry and Lions [27, 28] and Huang, Caines and Malhamé [25], is devoted to the analysis of models where a large number of players interact strategically with each other. Under suitable assumptions, the Nash equilibria of those games can be analyzed through the solutions of the, so-called, MFG system

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d \times [0, T] \\ -\partial_t m + \Delta m + \text{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d \times [0, T] \\ m(0) = m_0, \ u(T, x) = G(x, m(T)) & \text{in } \mathbb{T}^d. \end{cases} \quad (0.1)$$

The first unknown $u(t, x)$ is the value function of an infinitesimal player starting from x at time t while the second one, $m(t)$, describes the distribution of players at time t. The maps $F, G : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ (where $\mathcal{P}(\mathbb{T}^d)$ is the set of Borel probability measures on the torus $\mathbb{T}^d$) describe the interactions between players.

In this paper we investigate the limit behavior, as the horizon T tends to infinity, of this system. This is a very natural question, especially when one looks at those models as dynamical systems.

One natural guess is that the system simplifies in large times and converges to a time independent model, called the ergodic MFG system:

$$\begin{cases} -\bar{\lambda} - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d, \\ \Delta m + \text{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d. \end{cases} \quad (0.2)$$

There is a relatively wide evidence of this phenomenon, starting from [29] and the Mexican wave model in [24] to more recent contributions in [4, 7, 8, 23].

All these papers, however, rely on a structure property, the so-called monotonicity assumption, which is seldom met in practice. More recently, the problem of understanding what happens

---

*PSL Research University, Université Paris-Dauphine, CEREMADE, Place de Lattre de Tassigny, F- 75016 Paris, France

in the non-monotone setting has been addressed in several papers. Gomes and Sedjro [22] found the first example of periodic solutions in the context of one-dimensional first order system with congestion. Cirant in [11] and Cirant and Nurbekyan in [10] forecast and then proved the existence of periodic solutions for a specific class of second order MFG systems (with quadratic Hamiltonian). These periodic trajectories were built through a bifurcation method in a neighborhood of a simple solution. Note that these examples show that the ergodic MFG system is not always the limit of the time-dependent ones. In [30] the second authors gave additional evidence of this phenomena using ideas from *weak KAM theory* [12, 13, 14]. The main interest of the approach is that it allows to study the question for a large class of MFG systems, *potential MFG systems*.

We say that a MFG system like (0.1) is of potential type if it can be derived as optimality condition of the following optimal control problem on the Fokker-Plank equation

$$\mathcal{U}^T(t, m_0) = \inf_{(m,\alpha)} \int_0^T \int_{\mathbb{T}^d} H^*(x, \alpha(s, x))\, dm(s) + \mathcal{F}(m(s))dt + \mathcal{G}(m(T)), \qquad (0.3)$$

where (m, α) verifies the Fokker Plank equation $-\partial_t m + \Delta m + \operatorname{div}(\alpha m) = 0$ with $m(t) = m_0$ and $\mathcal{F}$ and $\mathcal{G}$ are respectively the potentials of the functions F and G that appear in (0.1). Since the very beginning, this class of models has drawn a lot of attention. [28] first explained the mechanism behind the minimizing problem (0.3) and the MFG system (0.1) and, since then, the literature on potential MFG thrived. See for instance [3, 6, 17, 31] for the use of theses techniques to build solutions and analyse their long-time behavior under a monotonicity assumption.

In the present paper, we investigate the behavior, as $T \to +\infty$, of the solutions to the mean field games system (0.1) which are minimizers of (0.3). It is a continuation of [30], which started the analysis of the convergence of the time-dependent, non-monotone, potential MFG systems through weak KAM techniques. We believe that these techniques lead to a more fundamental understanding of long time behavior for potential MFG. When the powerful tools of the weak KAM theory can be deployed, one can look at this problem in a more systematic way. Unlike the PDEs techniques that were so far used, this approach does not depend on the monotonicity of the system. A key point is that the weak KAM theory, exploiting the Hamiltonian structure of potential MFG, gives us a clear understanding of the limit object that the trajectories minimize when the time goes to infinity. We draw fully from both Fathi's seminal papers [12, 13, 14] and his book [15]. Several objects defined along the paper and the very structure of many proofs will sound familiar for who is acquainted with weak KAM theory. Nonetheless, it is not always straightforward to transpose these techniques into the framework of MFG and it often requires more effort than in the standard case. It is worthwhile to mention that infinite dimensional weak KAM theorems are not new, especially in the context of Wasserstein spaces: see for instance [18, 19, 20, 21]. These papers do not address the MFG problem but they surely share the same inspiration.

Let us now present our main results and discuss the strategy of proofs. As we have anticipated, the starting point of this paper are some results proved in [30]. The first one is the existence of the ergodic constant λ, such that

$$\frac{\mathcal{U}^T(0, \cdot)}{T} \longrightarrow -\lambda,$$

where $\mathcal{U}^T$ is defined in (0.3). The second one is the existence of corrector functions. We say that a continuous function χ on $\mathcal{P}(\mathbb{T}^d)$ is a corrector function if it verifies the following dynamic programming principle

$$\chi(m_0) = \inf_{(m,\alpha)} \left(\int_0^t H^*(x, \alpha)dm(s) + \mathcal{F}(m(s))ds + \chi(m(t)) \right) + \lambda t \qquad (0.4)$$

2

where (m, α) solves in the sense of distributions $-\partial_t m + \Delta m + \mathrm{div}(m\alpha) = 0$ with initial condition $m(0) = m_0$. At the heuristic level, this amounts to say that χ solves the ergodic problem

$$\int_{\mathbb{T}^d} (H(y, D_m\chi(m, y)) - \mathrm{div}_y D_m\chi(m, y))m(dy) = \mathcal{F}(m) + \lambda \qquad \text{in } \mathcal{P}(\mathbb{T}^d). \qquad (0.5)$$

(the notion of derivative $D_m\chi$ is described in Section 1 below).

The main results of this paper are Theorem 5.7 and Theorem 5.9. The first one states that $\mathcal{U}^T(0, \cdot) + \lambda T$ uniformly converges to a corrector function while the second one ensures that this convergence does not hold only at the level of minimization problems but also when it comes to optimal trajectories. In particular, Theorem 5.9 says that optimal trajectories for $\mathcal{U}^T(0, \cdot) + \lambda T$ converge to calibrated curves (i.e., roughly speaking, to global minimizers of (0.4)). Let us recall that, in [30], the second author provides examples in which the calibrated curves stay away from the solutions of the MFG ergodic system (0.2). In that framework, our result implies that no solution to the MFG system (0.1) obtained as minimizers of (0.3) converges to a solution of the MFG ergodic system.

The convergence of $\mathcal{U}^T(0, \cdot) + \lambda T$ to a corrector is of course the transposition, in our setting, of Fathi's famous convergence result for Hamilton-Jacobi equations [14]. The basic strategy of proof is roughly the same. Here, the additional difficulty lies in the fact that, in our infinitely dimensional framework, the Hamiltonian in (0.5) is neither first order nor "uniformly elliptic" (cf. the term in divergence in (0.5)).

We overcome this difficulty by introducing two main ideas, that we describe now. As in [14], we start with further characterizations of the limit value λ. Let us set

$$I := \inf_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} \sup_{m \in \mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} (H(y, D_m\Phi(m, y)) - \mathcal{F}(m) - \mathrm{div}_y D_m\Phi(m, y))m(dy).$$

Then, by duality techniques, one can check the following equality (Proposition 2.2)

$$-I = \min_{(\mu, p_1)} \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \left(H^*\left(y, \frac{dp_1}{dm \otimes \mu}\right) + \mathcal{F}(m) \right) m(dy)\mu(dm), \qquad (0.6)$$

where (μ, p_1) are closed measures, in the sense that, for any $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$,

$$-\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m\Phi(m, y) \cdot p_1(dm, dy) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathrm{div}_y D_m\Phi(m, y)m(dy)\mu(dm) = 0.$$

We call Mather measure any couple (μ, p_1) which minimizes the dual problem (on this terminology, see Remark 5.1).

One key step is to show that $I = \lambda$. While it is easy to prove that $I \geqslant \lambda$ (Proposition 2.1), the opposite inequality is trickier. One has to construct a smooth subsolution of the ergodic problem (0.5) in a context where there is no "classical" convolution. The idea is to look at a finite particle system on $(\mathbb{T}^d)^N$. A similar idea was used in [19] for first order problems on the $L^2(0, 1)$-torus. The main difference with [19] is that, for first order problems, the particle system is embedded into the continuous one, which is not the case for problems with diffusion. The argument of proof is therefore completely different. We set $(v^N, \lambda^N) \in C^2((\mathbb{T}^d)^N) \times \mathbb{R}$ solution of

$$-\sum_{i=1}^{N} \Delta_{x_i} v^N(\mathbf{x}) + \frac{1}{N} \sum_{i=1}^{N} H(x_i, ND_{x_i}v^N(\mathbf{x})) = \mathcal{F}(m_{\mathbf{x}}^N) + \lambda^N.$$

Note that, in contrast with [19], the constant λ^N, here, depends on N. Our first main idea is to introduce the smooth function on $\mathcal{P}(\mathbb{T}^d)$

$$W^N(m) := \int_{(\mathbb{T}^d)^N} v^N(x_1, \ldots, x_N) \prod_{i=1}^N m(dx_i)$$

and to show that it satisfies in $\mathcal{P}(\mathbb{T}^d)$

$$-\int_{\mathbb{T}^d} \operatorname{div}_y D_m W^N(m, y) m(dy) + \int_{\mathbb{T}^d} H(y, D_m W^N(m, y), m) m(dy) \leqslant \mathcal{F}(m) + \lambda^N + o(N),$$

which implies that $I \leqslant \liminf_N \lambda^N$. By comparing W^N with true correctors, through viscosity type arguments, we also prove that $\limsup_N \lambda^N \leqslant \lambda$. This shows the equality $I = \lambda$ and the convergence of the constant λ^N associated to the particle system to the ergodic constant λ (Proposition 3.1).

The next difficulty is that the Hamiltonian appearing in (0.5) is singular (because of the divergence term). This prevents us to say, as in the classical setting, that Mather measures are supported by graphs on which the Hamilton-Jacobi is somehow satisfied. To overcome this issue, we introduce our second main idea, the notion of "smooth" Mather measures (measures supported by "smooth" probability measures). We prove that limits of minimizers for $\mathcal{U}^T$ provide indeed "smooth" Mather measures and that, if (μ, p_1) is a "smooth" Mather measure then, if we set

$$q_1(x, m) := D_a H^* \left(y, \frac{dp_1}{dm \otimes \mu}(y, m) \right),$$

we have, for $\mu-$a.e. $m \in \mathcal{P}(\mathbb{T}^d)$,

$$\int_{\mathbb{T}^d} q_1(y, m) \cdot Dm(y) dy + \int_{\mathbb{T}^d} H(y, q_1(y, m)) m(dy) = \mathcal{F}(m) + \lambda. \qquad (0.7)$$

(see Proposition 4.2). Note that (0.7) is a kind of reformulation of the ergodic equation (0.5), in which $q_1 = D_m \chi$ and where the divergence term is integrated by parts. The rest of the proof is more standard and does not bring new difficulties compared to [14].

Let us briefly describe the organization of the paper. In Section 1, we fix the main notation and assumption and collect the results of [30] that we sketched above. Section 2 and 3 focus on further characterizations of the limit value λ. In particular, in Section 2, we prove that (0.6) and $I \geqslant \lambda$ hold, while Section 3 is devoted to the analysis of the particle system and the proof that $I = \lambda$. Section 4 gives a closer look to Mather measures and explains (0.7). Section 5 contains Theorem 5.7 and Theorem 5.9 and their proofs.

1 Assumptions and preliminary results

The aim of this preliminary section is twofold. Firstly, we introduce the notation and the assumptions that we will use throughout the paper. Then, we collect some results from [30] which are the starting point of this work.

1.1 Notation and assumptions

We work on the $d-$dimensional flat torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ to avoid boundary conditions and to set the problem on a compact domain. We denote by $\mathcal{P}(\mathbb{T}^d)$ the set of Borel probability measures on $\mathbb{T}^d$. This is a compact, complete and separable set when endowed with the 1-Wasserstein distance $\mathbf{d}(\cdot,\cdot)$. Let m be a Borel measure over $[t,T] \times \mathbb{T}^d$, with first marginal the Lebesgue measure ds over $[t,T]$, then with $\{m(s)\}_{s\in[t,T]}$ we denote the disintegration of m with respect to dt. We will always consider measures m such that $m(s)$ is a probability measure on $\mathbb{T}^d$ for any $s \in [t,T]$.

If m is such a measure, then $L_m^2([t,T] \times \mathbb{T}^d)$ is the set of m-measurable functions f such that the integral of $|f|^2 dm(s)$ over $[t,T] \times \mathbb{T}^d$ is finite.

We use throughout the paper the notion of derivative for functions defined on $\mathcal{P}(\mathbb{T}^d)$ introduced in [5]. We say that $\Phi : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ is C^1 if there exists a continuous function $\frac{\delta\Phi}{\delta m} : \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \to \mathbb{R}$ such that

$$\Phi(m_1) - \Phi(m_2) = \int_0^1 \int_{\mathbb{T}^d} \frac{\delta\Phi}{\delta m}((1-t)m_1 + tm_2, x)(m_2 - m_1)(dx)dt, \qquad \forall m_1, m_2 \in \mathcal{P}(\mathbb{T}^d).$$

As this derivative is defined up to an additive constant, we use the standard normalization

$$\int_{\mathbb{T}^d} \frac{\delta\Phi}{\delta m}(m, x)m(dx) = 0. \tag{1.8}$$

We recall that, if $\mu, \nu \in \mathcal{P}(\mathbb{T}^d)$, the 1-Wasserstein distance is defined by

$$\mathbf{d}(\mu,\nu) = \sup\left\{ \int_{\mathbb{T}^d} \phi(x)\, d(\mu-\nu)(x) \,\Big|\, \text{continuous } \phi : \mathbb{T}^d \to \mathbb{R}, \text{Lip}(\phi) \leqslant 1 \right\}.$$

Assumptions: Throughout the paper the following conditions will be in place.

1. $H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ is of class C^2, $p \mapsto D_{pp}H(x,p)$ is Lipschitz continuous, uniformly with respect to x. Moreover, there exists $\bar{C} > 0$ that verifies

$$\bar{C}^{-1}I_d \leqslant D_{pp}H(x,p) \leqslant \bar{C}I_d, \quad \forall(x,p) \in \mathbb{T}^d \times \mathbb{R}^d$$

    and $\theta \in (0,1)$, $C > 0$ such that the following conditions hold true

$$|D_{xx}H(x,p)| \leqslant C(1+|p|)^{1+\theta}, \quad |D_{x,p}H(x,p)| \leqslant C(1+|p|)^\theta, \quad \forall(x,p) \in \mathbb{T}^d \times \mathbb{R}^d.$$

2. $\mathcal{F} : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ is of class C^2. Its derivative $F : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ is twice differentiable in x and $D_{xx}^2 F$ is bounded. Examples of non monotone coupling functions which verify such conditions can be found in [30].

Note that some of the above assumptions will not be used explicitly in this paper but have been used in [30] to prove results that we will assume to hold true. (The only difference is that here we need $\mathcal{F}$ to be of class C^2 while in [30] $\mathcal{F}$ was required only to be C^1).

Very often in the text, we do not need to work explicitly on $\mathcal{F}$, so, in order to have a lighter notation, we incorporate $\mathcal{F}$ in the Hamiltonian defining, for any $(x,p,m) \in \mathbb{T}^d \times \mathbb{R}^d \times \mathcal{P}(\mathbb{T}^d)$,

$$\mathcal{H}(x,p,m) := H(x,p) - \mathcal{F}(m). \tag{1.9}$$

We denote with $\mathcal{H}^*$ the Fenchel conjugate of $\mathcal{H}$ with respect to the second variable. Then,

$$\mathcal{H}^*(x, p, m) = H^*(x, p) + \mathcal{F}(m).$$

We can now introduce the standard minimization problem in potential MFG:

$$\mathcal{U}^T(t, m_0) = \inf_{(m, \alpha)} \int_t^T \int_{\mathbb{T}^d} H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) dt, \quad m_0 \in \mathcal{P}(\mathbb{T}^d)$$

where $m \in C^0([t, T], \mathcal{P}(\mathbb{T}^d))$, $\alpha \in L_m^2([t, T] \times \mathbb{T}^d, \mathbb{R}^d)$ and the following equation is verified in the sense of distributions

$$\begin{cases} -\partial_t m + \Delta m + \operatorname{div}(m\alpha) = 0 & \text{in } [t, T] \times \mathbb{T}^d \\ m(t) = m_0 & \text{in } \mathbb{T}^d. \end{cases} \tag{1.10}$$

1.2 Corrector functions and the limit value λ

Here we collect the results already proved in [30] that we will use. A most important one is the following.

Theorem 1.1. *The function $\frac{1}{T}\mathcal{U}^T(t, \cdot)$ uniformly converges to a limit value $-\lambda$ when T goes to $+\infty$.*

The second result that we will use is the existence of corrector functions.

Definition 1.2. *We say that $\chi : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ is a corrector function if, for any $m_0 \in \mathcal{P}(\mathbb{T}^d)$ and any $t > 0$,*

$$\chi(m_0) = \inf_{(m, \alpha)} \left(\int_0^t H^*(x, \alpha) dm(s) + \mathcal{F}(m(s)) ds + \chi(m(t)) \right) + \lambda t,$$

where $m \in C^0([0, t], \mathcal{P}(\mathbb{T}^d))$, $\alpha \in L_m^2([0, t] \times \mathbb{T}^d, \mathbb{R}^d)$ and the pair (m, α) solves in the sense of distributions $-\partial_t m + \Delta m + \operatorname{div}(m\alpha) = 0$ with initial condition m_0.

Proposition 1.3. *The set of corrector functions is not empty and uniformly Lipschitz continuous.*

A last notion that will come at hand is the one of calibrated curve:

Definition 1.4. *We say that $(\bar{m}, \bar{\alpha})$, which satisfies (1.10) for any $t \in \mathbb{R}$, is a calibrated curve if there exists a corrector function $\chi : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ such that $(\bar{m}, \bar{\alpha})$ is optimal for χ: for any $t_1 < t_2 \in \mathbb{R}$*

$$\chi(\bar{m}(t_1)) = \lambda(t_2 - t_1) + \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(s)) d\bar{m}(s) + \mathcal{F}(\bar{m}(s)) ds + \chi(\bar{m}(t_2)).$$

The set of calibrated curves verifies the following property.

Proposition 1.5. *The set of calibrated curves is not empty. Moreover, if (m, α) is a calibrated curve, then $m \in C^{1,2}(\mathbb{R} \times \mathbb{T}^d)$ and there exists a function $u \in C^{1,2}(\mathbb{R} \times \mathbb{T}^d)$ such that $\alpha = D_p H(x, Du)$ where (u, m) solves*

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{R} \times \mathbb{T}^d, \\ -\partial_t m + \Delta m + \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{R} \times \mathbb{T}^d. \end{cases}$$

2 A dual problem

In this section we introduce the two usual characterizations of the constant λ: λ is expected to be the smallest constant for which there exists a smooth sub-corrector and $-\lambda$ is the smallest value of the Lagrangian when integrated against suitable "closed" measures. The goal of this section is to show that both problems are in duality and have the same value I. We postpone the analysis of the equality $I = \lambda$ to the next section.

We start with the HJ equation, that we write in variational form.

$$I := \inf_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} \sup_{m \in \mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} (\mathcal{H}(y, D_m\Phi(m,y), m) - \operatorname{div}_y D_m\Phi(m,y))m(dy), \qquad (2.11)$$

where by $C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ we mean the set of maps $\Phi : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ such that $D_m\Phi$ and $D_y D_m\Phi$ are continuous. Recall that $\mathcal{H}$ is defined in (1.9).

Let us start with a comparison between I and λ.

Proposition 2.1. *We have $I \geqslant \lambda$.*

Proof. Let $\varepsilon > 0$ and Φ be such that

$$\int_{\mathbb{T}^d} (\mathcal{H}(y, D_m\Phi(m,y), m) - \operatorname{div}_y D_m\Phi(m,y))m(dy) \leqslant I + \varepsilon \qquad \forall m \in \mathcal{P}(\mathbb{T}^d).$$

Let $(\bar{m}, \bar{\alpha})$ be a calibrated curve and χ be a corrector function associated with $(\bar{m}, \bar{\alpha})$. By definition of calibrated curve, $(\bar{m}, \bar{\alpha})$ verifies

$$-\partial_t \bar{m}(t) + \Delta \bar{m}(t) + \operatorname{div}(\bar{\alpha}(t)\bar{m}(t)) = 0, \qquad \forall t \in \mathbb{R}$$

and

$$\chi(\bar{m}(0)) = \int_0^T \int_{\mathbb{T}^d} \mathcal{H}^*(y, \bar{\alpha}(t,y), \bar{m}(t))\bar{m}(t,dy) + \lambda T + \chi(\bar{m}(T)). \qquad (2.12)$$

As Φ is smooth, we get

$$\Phi(\bar{m}(T)) - \Phi(\bar{m}(0)) = \int_0^T \frac{d}{dt}\Phi(\bar{m}(t))dt = \int_0^T \int_{\mathbb{T}^d} \frac{\delta\Phi}{\delta m}(\bar{m}(t), y)\partial_t\bar{m}(t)dydt$$

$$= \int_0^T \int_{\mathbb{T}^d} \operatorname{div}_y D_m\Phi(\bar{m}(t), y)\bar{m}(t,dy)dt - \int_0^T \int_{\mathbb{T}^d} D_m\Phi(\bar{m}(t), y) \cdot \bar{\alpha}(t,y)\bar{m}(t,dy)dt$$

$$\geqslant -(I+\varepsilon)T + \int_0^T \int_{\mathbb{T}^d} (\mathcal{H}(y, D_m\Phi(\bar{m}(t), y), \bar{m}(t)) - D_m\Phi(\bar{m}(t), y) \cdot \bar{\alpha}(t,y))\bar{m}(t,dy)dt$$

$$\geqslant -(I+\varepsilon)T - \int_0^T \int_{\mathbb{T}^d} \mathcal{H}^*(y, \bar{\alpha}(t,y), \bar{m}(t))\bar{m}(t,dy).$$

Using (2.12), we end up with

$$\Phi(\bar{m}(T)) - \Phi(\bar{m}(0)) \geqslant -(I+\varepsilon)T - \int_0^T \int_{\mathbb{T}^d} \mathcal{H}^*(y, \bar{\alpha}(t,y), \bar{m}(t))\bar{m}(t,dy)$$

$$= -(I+\varepsilon)T + \chi(\bar{m}(T)) - \chi(\bar{m}(0)) + \lambda T.$$

Using the fact that χ and Φ are bounded, we divide both sides by T and we conclude that $\lambda \leqslant I$ by letting $T \to +\infty$ and $\varepsilon \to 0$. $\blacksquare$

Next we reformulate I in terms of "closed measures".

Proposition 2.2. *We have*

$$-I = \min_{(\mu, p_1)} \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^*\Big(y, \frac{dp_1}{dm \otimes \mu}, m\Big) m(dy)\mu(dm), \tag{2.13}$$

where the minimum is taken over $\mu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$ and p_1, Borel vector measure on $\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d$, such that p_1 is absolutely continuous with respect to the measure $dm \otimes \mu := m(dy)\mu(dm)$ and such that (μ, p_1) is closed, in the sense that, for any $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$,

$$-\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m\Phi(m, y) \cdot p_1(dm, dy) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m\Phi(m, y)m(dy)\mu(dm) = 0. \tag{2.14}$$

In analogy with weak KAM theory, we call a measure (μ, p_1) satisfying (2.14) a closed measure and a minimum of (2.13) a Mather measure.

Proof. As usual we can rewrite I as

$$I = \inf_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} \sup_{\mu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))} \int_{\mathcal{P}(\mathbb{T}^d)} \left(\int_{\mathbb{T}^d} (\mathcal{H}(y, D_m\Phi(m, y), m) - \operatorname{div}_y D_m\Phi(m, y))m(dy) \right) \mu(dm).$$

We claim that

$$I = \max_{\mu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))} \inf_{\Phi \in C^{1,1}(\mathcal{P})} \int_{\mathcal{P}(\mathbb{T}^d)} \left(\int_{\mathbb{T}^d} (\mathcal{H}(y, D_m\Phi(m, y), m) - \operatorname{div}_y D_m\Phi(m, y))m(dy) \right) \mu(dm). \tag{2.15}$$

Indeed, $\mathcal{P}(\mathcal{P}(\mathbb{T}^d))$ is a compact subspace of $\mathcal{M}(\mathcal{P}(\mathbb{T}^d))$ and, for any fixed $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ the function on $\mathcal{M}(\mathcal{P}(\mathbb{T}^d))$ defined by

$$\mu \mapsto \int_{\mathcal{P}(\mathbb{T}^d)} \left(\int_{\mathbb{T}^d} (\mathcal{H}(y, D_m\Phi(m, y), m) - \operatorname{div}_y D_m\Phi(m, y))m(dy) \right) \mu(dm)$$

is continuous and concave (as it is linear). On the other hand, when we fix $\mu \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$, the function on $C^{1,1}(\mathcal{P}(\mathbb{T}^d))$, defined by

$$\Phi \mapsto \int_{\mathcal{P}(\mathbb{T}^d)} \left(\int_{\mathbb{T}^d} (\mathcal{H}(y, D_m\Phi(m, y), m) - \operatorname{div}_y D_m\Phi(m, y))m(dy) \right) \mu(dm),$$

is continuous with respect to the uniform convergence in $C^{1,1}(\mathcal{P}(\mathbb{T}^d))$ and convex due to the convexity of $\mathcal{H}$. Therefore, the hypothesis of Sion's min-max Theorem are fulfilled and (2.15) holds true.

We now define the continuous linear map $\Lambda : C^{1,1}(\mathcal{P}(\mathbb{T}^d)) \to (C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d \times C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)$ by

$$\Lambda(\Phi) = (D_m\Phi, \operatorname{div}_y D_m\Phi).$$

From now on we fix a maximizer μ for (2.15) and we define $E := C^{1,1}(\mathcal{P}(\mathbb{T}^d))$, $F := (C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d \times C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)$ and

$$f(\Phi) = 0, \; g(a, b) = \int_{\mathcal{P}(\mathbb{T}^d)} (\int_{\mathbb{T}^d} (\mathcal{H}(y, a(m, y), m) - b(m, y))m(dy))\mu(dm), \qquad \forall(a, b) \in F.$$

We note that
$$I = \inf_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} \{f(\Phi) + g(\Lambda\Phi)\}.$$

To use the Fenchel-Rockafellar theorem we need to check the transversality conditions. These hypothesis are easily verified, indeed, both f and g are continuous and, therefore, proper functions. The function f is convex because it is linear and so is g, due to the convexity of Hamiltonian $\mathcal{H}$. Moreover, it comes directly from its definition that Λ is a bounded linear functional on E. Then, the Fenchel-Rockafellar Theorem states that

$$I = - \min_{(p_1, p_2) \in F'} \{f^*(-\Lambda^*(p_1, p_2)) + g^*(p_1, p_2)\}.$$

Note that
$$F' = (\mathcal{M}(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d \times \mathcal{M}(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d),$$

that $f^*(q) = 0$ if $q = 0$, $f^*(q) = +\infty$ otherwise. So, for any $(p_1, p_2) \in F'$,

$$f^*(-\Lambda(p_1, p_2)) = \sup_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} -\langle \Lambda^*(p_1, p_2), \Phi \rangle - f(\Phi) = \sup_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} -\langle (p_1, p_2), \Lambda(\Phi) \rangle$$

$$= \sup_{\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))} -\int_{(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m\Phi(m, y) \cdot p_1(dm, dy) - \int_{(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathrm{div}_y D_m\Phi(m, y) p_2(dm, dy)$$

which is 0 if, for any $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$,

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m\Phi(m, y) \cdot p_1(dm, dy) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathrm{div}_y D_m\Phi(m, y) p_2(dm, dy) = 0,$$

and $+\infty$ otherwise. On the other hand,

$$g^*(p_1, p_2) = \sup_{(a,b) \in F} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a(m, y) \cdot p_1(dm, dy) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} b(m, y) p_2(dm, dy)$$

$$- \int_{\mathcal{P}(\mathbb{T}^d)} \left(\int_{\mathbb{T}^d} (\mathcal{H}(y, a(m, y), m) - b(m, y)) m(dy) \right) \mu(dm).$$

So, if $g^*(p_1, p_2)$ is finite, one must have that $p_2(dm, dy) = -m(dy)\mu(dm)$ and that p_1 is absolutely continuous with respect to the measure $dm \otimes \mu := m(dy)\mu(dm)$. Indeed, if p_1 were not absolutely continuous with respect to $dm \otimes \mu$, we could find a sequence of continuous functions $a_n \in (C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d$ such that, a_n is uniformly bounded on the support of $dm \otimes \mu$ and

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a_n(m, y) \cdot p_1(dm, dy) \to +\infty.$$

But then we would have $g^*(p_1, p_2) = +\infty$. So,

$$g^*(p_1, p_2) = \sup_{a \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left(a(m, y) \cdot \frac{dp_1}{dm \otimes \mu}(m, y) - \mathcal{H}(y, a(m, y), m) \right) m(dy)\mu(dm)$$

We now want to prove that

$$\sup_{a \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left(a(m, y) \cdot \frac{dp_1}{dm \otimes \mu}(m, y) - \mathcal{H}(y, a(m, y), m) \right) m(dy)\mu(dm)$$

$$= \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^*\left(y, \frac{dp_1}{dm \otimes \mu}(m, y), m \right) m(dy)\mu(dm).$$

9

We have one inequality by definition of Fenchel's conjugate. Indeed,

$$\sup_{a \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left(a(m,y) \cdot \frac{dp_1}{dm \otimes \mu}(m,y) - \mathcal{H}(y, a(m,y), m) \right) m(dy)\mu(dm)$$

$$\leqslant \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a^*(m,y) \cdot \frac{dp_1}{dm \otimes \mu}(m,y) - \mathcal{H}(y, a^*(m,y), m) m(dy)\mu(dm)$$

$$= \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^*\left(y, \frac{dp_1}{dm \otimes \mu}(m,y), m \right) m(dy)\mu(dm),$$

where

$$a^*(m,y) = D_a \mathcal{H}^*\left(y, \frac{dp_1}{dm \otimes \mu}(m,y), m \right).$$

For the opposite inequality we use a density argument. The function a^* could be not continuous but yet it must be measurable. Moreover, the growth of $\mathcal{H}$ ensures that $a^* \in L^2_\mu(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d; \mathbb{R}^d)$. As $\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d$ is a compact Hausdorff space, the set of continuous functions is dense in $L^2_\mu(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d)$. Let $a_n \in C^0\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d; \mathbb{R}^d)$ be such that $a_n \to a^*$ in L^2_μ. Then,

$$\sup_{a \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d))^d} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a(m,y) \cdot \frac{dp_1}{dm \otimes \mu}(m,y) - \mathcal{H}(y, a(m,y), m) m(dy)\mu(dm) \geqslant$$

$$\lim_{n \to +\infty} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a_n(m,y) \cdot \frac{dp_1}{dm \otimes \mu}(m,y) - \mathcal{H}(y, a_n(m,y), m) m(dy)\mu(dm) =$$

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} a^*(m,y) \cdot \frac{dp_1}{dm \otimes \mu}(m,y) - \mathcal{H}(y, a^*(m,y), m) m(dy)\mu(dm) =$$

$$\int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^*\left(y, \frac{dp_1}{dm \otimes \mu}(m,y), m \right) m(dy)\mu(dm).$$

Therefore, we can conclude that

$$I = -\min_{(\mu,p_1)} \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^*\left(y, \frac{dp_1}{dm \otimes \mu}, m \right) m(dy)\mu(dm),$$

where the minimum is taken over (μ, p_1) satisfying condition (2.14). ∎

3   The $N-$particule problem.

In the previous section, we introduced two problems in duality. These problems have a common value called I and we have checked that $\lambda \leqslant I$. The aim of this section is to show that there is actually an equality: $\lambda = I$. In the standard setting, this equality is proved by smoothing correctors by a convolution; by the convexity of the Hamiltonian, the smoothened corrector is a subsolution to the corrector equation (up to a small error term), thus providing a competitor for problem (2.11). In our framework, there is no exact equivalent of the convolution. We overcome this difficulty by considering the projection of the problem onto the set of empirical measures of size N (thus on $(\mathbb{T}^d)^N$). For the $N-$particle problem, the corrector is smooth. We explain here that a suitable extension of this finite dimensional corrector to the set $\mathcal{P}(\mathbb{T}^d)$ provides a smooth sub-corrector for the problem in $\mathcal{P}(\mathbb{T}^d)$ when N is large. This shows the claimed equality and, in addition, the fact that the ergodic constant associated with the $N-$particle problem converges to λ.

More precisely, we consider $v^N : (\mathbb{T}^d)^N \to \mathbb{R}$ the solution of:

$$-\sum_{i=1}^{N} \Delta_{x_i} v^N(\mathbf{x}) + \frac{1}{N} \sum_{i=1}^{N} \mathcal{H}(x_i, ND_{x_i} v^N(\mathbf{x}), m_{\mathbf{x}}^N) = \lambda^N,$$

where $\mathbf{x} = (x_1, \ldots, x_N) \in (\mathbb{T}^d)^N$ and $m_{\mathbf{x}}^N = N^{-1} \sum_{i=1}^{N} \delta_{x_i}$ and where $\mathcal{H}$ is defined in (1.9). Let us recall that such a corrector exists (it is unique a to additive constants) and is smooth.

Proposition 3.1. *We have*
$$\lambda = I = \lim_{N \to +\infty} \lambda^N.$$

Proof. We define
$$W^N(m) := \int_{(\mathbb{T}^d)^N} v^N(x_1, \ldots, x_N) \prod_{i=1}^{N} m(dx_i).$$

As v^N is smooth, it is clear that W^N is also smooth on $\mathcal{P}(\mathbb{T}^d)$ and we have

$$D_m W^N(m, y) = \sum_{k=1}^{N} \int_{(\mathbb{T}^d)^{N-1}} D_{x_k} v^N(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_N) \prod_{i \neq k} m(dx_i)$$

and

$$\mathrm{div}_y D_m W^N(m, y) = \sum_{k=1}^{N} \int_{(\mathbb{T}^d)^{N-1}} \Delta_{x_k} v^N(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_N) \prod_{i \neq k} m(dx_i).$$

In view of the convexity of $\mathcal{H}$ with respect to p, we obtain, for any $m \in \mathcal{P}(\mathbb{T}^d)$,

$$-\int_{\mathbb{T}^d} \mathrm{div}_y D_m W^N(m, y) m(dy) + \int_{\mathbb{T}^d} \mathcal{H}(y, D_m W^N(m, y), m) m(dy)$$

$$= -\sum_{k=1}^{N} \int_{\mathbb{T}^d} \int_{(\mathbb{T}^d)^{N-1}} \Delta_{x_k} v^N(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_N) \prod_{i \neq k} m(dx_i) m(dy)$$

$$+ \int_{\mathbb{T}^d} \mathcal{H}\Big(y, \sum_{k=1}^{N} \int_{(\mathbb{T}^d)^{N-1}} D_{x_k} v^N(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_N) \prod_{i \neq k} m(dx_i), m\Big) m(dy)$$

$$\leqslant \int_{(\mathbb{T}^d)^N} \Big(-\sum_{k=1}^{N} \Delta_{x_k} v^N(x_1, \ldots, x_N) + \frac{1}{N} \sum_{k=1}^{N} \mathcal{H}\Big(x_k, ND_{x_k} v^N(x_1, \ldots, x_N), m\Big)\Big) \prod_i m(dx_i).$$

Following [9], the Glivenko-Cantelli Theorem states that

$$\int_{(\mathbb{T}^d)^N} \mathbf{d}_2(\mu_{\mathbf{x}}^N, m) \prod_i m(dx_i) \leqslant \varepsilon_N := \begin{cases} N^{-1/2} & \text{if } d < 4, \\ N^{-1/2} \ln(N) & \text{if } d = 4, \\ N^{-2/d} & \text{otherwise.} \end{cases}$$

As $\mathcal{H}$ has a separate form: $\mathcal{H}(x, p, m) = H(x, p) - \mathcal{F}(m)$ where $\mathcal{F}$ is Lipschitz continuous with respect to m, we infer that

$$\int_{(\mathbb{T}^d)^N} \frac{1}{N} \sum_{k=1}^{N} \mathcal{H}\Big(x_k, ND_{x_k} v^N(x_1, \ldots, x_N), m\Big) \prod_i m(dx_i)$$

$$\leqslant \int_{(\mathbb{T}^d)^N} \frac{1}{N} \sum_{k=1}^{N} \mathcal{H}\Big(x_k, ND_{x_k} v^N(x_1, \ldots, x_N), m_{\mathbf{x}}^N\Big) \prod_i m(dx_i) + C\varepsilon_N.$$

11

Recalling the equation satisfied by v^N, we conclude that

$$- \int_{\mathbb{T}^d} \operatorname{div}_y D_m W^N(m, y) m(dy) + \int_{\mathbb{T}^d} \mathcal{H}(y, D_m W^N(m, y), m) m(dy)$$

$$\leqslant \int_{(\mathbb{T}^d)^N} \Big(- \sum_{k=1}^N \Delta_{x_k} v^N(x_1, \ldots, x_N) + \frac{1}{N} \sum_{k=1}^N \mathcal{H}\Big(x_k, N D_{x_k} v^N(x_1, \ldots, x_N), m_{\mathbf{x}}^N \Big) \Big) \prod_i m(dx_i) + C \varepsilon_N$$

$$\leqslant \lambda^N + C \varepsilon_N.$$

As W^N is smooth, this shows that $I \leqslant \liminf_N \lambda^N$.

Let $\lambda^* = \limsup_N \lambda^N$. From now on we work up to a subsequence N (still denoted in the same way) such that $\lim_N \lambda^N = \lambda^*$. We claim that $\lambda^* \leqslant \lambda$. Indeed, let χ be a corrector and, for N large, $\mathbf{x}^N = (x_1^N, \ldots, x_N^N)$ be a minimum point of $\mathbf{y} \to \chi(m_{\mathbf{y}}^N) - v^N(\mathbf{y})$. We know from Lemma 6.3 that there exists a map $\Phi^N : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ of class $C^2(\mathcal{P}(\mathbb{T}^d))$ (and bounded in $C^2(\mathcal{P}(\mathbb{T}^d))$ independently of N) such that $\Phi^N > \chi$ on $\mathcal{P}(\mathbb{T}^d) \backslash \{m_{\mathbf{x}^N}^N\}$ and $\Phi^N(m_{\mathbf{x}^N}^N) = \chi(m_{\mathbf{x}^N}^N)$. Moreover,

$$- \int_{\mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m_{\mathbf{x}^N}^N, y) m_{\mathbf{x}^N}^N(dy) + \int_{\mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m_{\mathbf{x}^N}^N, y), m_{\mathbf{x}^N}^N) m_{\mathbf{x}^N}^N(dy) \leqslant \lambda.$$

As $\mathbf{x}^N$ is also a minimum point of $\mathbf{y} \to \Phi^N(m_{\mathbf{y}}^N) - v^N(\mathbf{y})$, we have, following [9] Section 5.4.4,

$$D_{x_i} v^N(\mathbf{x}^N) = \frac{1}{N} D_m \Phi^N(m_{\mathbf{x}^N}^N, x_i^N),$$

$$D_{x_i x_i}^2 v^N(\mathbf{x}^N) \geqslant \frac{1}{N} D_{ym}^2 \Phi^N(m_{\mathbf{x}^N}^N, x_i^N) + \frac{1}{N^2} D_{mm}^2 \Phi^N(m_{\mathbf{x}^N}^N, x_i^N, x_i^N).$$

Hence

$$\lambda^N = - \sum_{i=1}^N \Delta_{x_i} v^N(\mathbf{x}^N) + \frac{1}{N} \sum_{i=1}^N \mathcal{H}(x_i^N, N D_{x_i} v^N(\mathbf{x}^N), m_{\mathbf{x}^N}^N)$$

$$\leqslant - \sum_{i=1}^N \Big(\frac{1}{N} \operatorname{div}_y D_m \Phi^N(m_{\mathbf{x}^N}^N, x_i^N) + \frac{1}{N^2} \operatorname{Tr} D_{mm}^2 \Phi^N(m_{\mathbf{x}^N}^N, x_i^N, x_i^N) \Big)$$

$$\quad + \frac{1}{N} \sum_{i=1}^N \mathcal{H}(x_i^N, D_m \Phi^N(m_{\mathbf{x}^N}^N, x_i^N), m_{\mathbf{x}^N}^N)$$

$$\leqslant - \int_{\mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m_{\mathbf{x}^N}^N, y) m_{\mathbf{x}^N}^N(dy) + C N^{-1} + \int_{\mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m_{\mathbf{x}^N}^N, y), m_{\mathbf{x}^N}^N) m_{\mathbf{x}^N}^N(dy)$$

$$\leqslant \lambda + C N^{-1},$$

where, in the last lines, we used the bound on $D_{mm}^2 \Phi^N$. Letting $N \to +\infty$, we find therefore $\lambda^* \leqslant \lambda$. As $\lambda \leqslant I \leqslant \liminf_N \lambda^N \leqslant \lambda^* \leqslant \lambda$, this shows that $\lambda = \lim_N \lambda^N = I$. ∎

4 On the support of the Mather measures

In this section we take a closer look at Mather measures and the properties of their support points.

Definition 4.1. *We say that the closed measure (μ, p_1) is smooth if there exists a constant $C > 0$ such that, for $\mu-$a.e. $m \in \mathcal{P}(\mathbb{T}^d)$, m has a positive density and*

$$\|D\ln(m)\|_{L^2_m(\mathbb{T}^d)} \leqslant C.$$

The aim of this section is to prove the following property of smooth Mather measures.

Proposition 4.2. *Let (μ, p_1) be a smooth Mather measure. Let us set*

$$q_1(x, m) := D_a \mathcal{H}^* \left(y, \frac{dp_1}{dm \otimes \mu}(y, m), m \right).$$

Then we have, for $\mu-$a.e. $m \in \mathcal{P}(\mathbb{T}^d)$,

$$\int_{\mathbb{T}^d} q_1(y, m) \cdot Dm(y) dy + \int_{\mathbb{T}^d} \mathcal{H}(y, q_1(y, m), m) m(dy) = \lambda.$$

In order to prove the proposition, let us start with a preliminary step. Let (μ, p_1) be optimal in problem (2.13) (where we recall that $I = \lambda$) and let Φ^N be any minimizing sequence in (2.11).

Lemma 4.3. *The sequence $(D_p \mathcal{H}(y, D_m \Phi^N(m, y)))$ converges to $\frac{dp_1}{dm \otimes \mu}$ in $L^2(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d, dm \otimes \mu)$. Moreover,*

$$\lim_{N \to +\infty} - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm)$$

$$= \lambda. \tag{4.16}$$

Proof. Recall that Φ^N satisfies

$$- \int_{\mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) + \int_{\mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \leqslant \lambda + o_N(1). \tag{4.17}$$

We integrate equation (4.17) against μ and add the problem for (μ, p_1) to find:

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm)$$

$$+ \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}^* \left(y, \frac{dp_1}{dm \otimes \mu}, m \right) m(dy) \mu(dm) \leqslant o_N(1).$$

Using the uniform convexity of $\mathcal{H}$, this implies that

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \frac{dp_1}{dm \otimes \mu} \cdot D_m \Phi^N(m, y) m(dy) \mu(dm)$$

$$+ C^{-1} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left| \frac{dp_1}{dm \otimes \mu} - D_p \mathcal{H}(y, D_m \Phi^N(m, y), m) \right|^2 m(dy) \mu(dm) \leqslant o_N(1).$$

Then, (2.14) implies that the first line vanishes and so

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \left| \frac{dp_1}{dm \otimes \mu} - D_p \mathcal{H}(y, D_m \Phi^N(m, y), m) \right|^2 m(dy) \mu(dm) \leqslant o_N(1),$$

which proves the first statement of the lemma. We now turn to (4.16). First of all, as Φ^N is a minimizing sequence for (2.11), we have that

$$\limsup_{N \to +\infty} - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm)$$

$$\leqslant \lambda.$$

To prove the other inequality we start with

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm).$$

We add and subtract the same quantity to get

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \frac{dp_1}{dm \otimes \mu} \cdot D_m \Phi^N(m, y) m(dy) \mu(dm) +$$

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm) - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \frac{dp_1}{dm \otimes \mu} \cdot D_m \Phi^N(m, y) m(dy) \mu(dm).$$

As (μ, p_1) verifies (2.14), the first line above vanishes. Then, using the Fenchel's inequality, we find that

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm) \geqslant$$

$$- \int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathcal{H}^* \left(y, \frac{dp_1}{dm \otimes \mu}, m \right) m(dy) \mu(dm).$$

By hypothesis, (μ, p_1) is a minimizer for (2.13), so, for any $N \in \mathbb{N}$,

$$- \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm) \geqslant \lambda.$$

Therefore,

$$\liminf_{N \to +\infty} - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, D_m \Phi^N(m, y), m) m(dy) \mu(dm)$$

$$\geqslant \lambda.$$

and the result follows. ∎

Proof of Proposition 4.2. From our assumption on (μ, p_1), we have

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) = - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m \Phi^N(m, y) \cdot Dm(y) dy \mu(dm).$$

As, by Lemma 4.3, the sequence $(D_p \mathcal{H}(y, D_m \Phi^N(m, y)), m)$ converges to $\frac{dp_1}{dm \otimes \mu}$ in $L^2(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d, dm \otimes \mu)$, we also have that $(D_m \Phi^N(m, y))$ converges to q_1 in $L^2(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d, dm \otimes \mu)$ by regularity and invertibility of $D_p \mathcal{H}$ and $D_a \mathcal{H}^*$. On the other hand, $\frac{Dm(y)}{m(y)}$ is bounded in $L_m^2(\mathbb{T}^d)$ for μ−a.e. $m \in \mathcal{P}(\mathbb{T}^d)$. Therefore

$$\lim_{N \to +\infty} \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \operatorname{div}_y D_m \Phi^N(m, y) m(dy) \mu(dm) = - \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} q_1(m, y) \cdot Dm(y) dy \mu(dm).$$

We conclude thanks to (4.16) that

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} q_1(m,y) \cdot Dm(y) dy \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \mathcal{H}(y, q_1(m,y), m) m(dy) \mu(dm) = \lambda.$$

On the other hand, extracting a subsequence if necessary, the sequence $(D_m \Phi^N(m,y))$ converges μ−a.e. to q_1. So, by (4.17), we have, for μ−a.e. $m \in \mathcal{P}(\mathbb{T}^d)$,

$$\int_{\mathbb{T}^d} q_1(y,m) \cdot Dm(y) dy + \int_{\mathbb{T}^d} \mathcal{H}(y, q_1(y,m), m) m(dy) \leqslant \lambda.$$

Putting together the previous inequality with the previous equality gives the result. ∎

5   The long time behavior of potential MFG

In this section, we prove the two main results of the paper: the first one is the convergence, as $T \to +\infty$, of $\mathcal{U}^T(0, \cdot) + \lambda T$. The second one states that limits of time-dependent minimizing mean field games equilibria, as the horizon tends to infinity, are calibrated curves.

5.1   Convergence of $\mathcal{U}^T(0, \cdot) + \lambda T$

We recall that $\mathcal{U}^T(t, m_0)$ is defined by

$$\mathcal{U}^T(t, m_0) = \inf_{(m,\alpha)} \int_t^T \left(\int_{\mathbb{T}^d} H^*(y, \alpha(s,y)) \, m(s,y) dy + \mathcal{F}(m(s)) \right) \, ds,$$

where (m, α) verifies the usual constraint

$$\partial_t m - \Delta m - \operatorname{div}(m\alpha) = 0 \text{ in } (t, T) \times \mathbb{T}^d, \qquad m(t) = m_0.$$

Let (m^T, α^T) be a minimizer of the problem, then, $\alpha^T(s,x) = D_p H(x, Du^T(s,x))$, where (u^T, m^T) solves

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x,m) & \text{in } \mathbb{T}^d \times [t, T] \\ -\partial_t m + \Delta m + \operatorname{div}(mD_p H(x, Du)) = 0 & \text{in } \mathbb{T}^d \times [t, T] \\ m(t) = m_0, \ u(T,x) = 0 & \text{in } \mathbb{T}^d, \end{cases} \tag{5.18}$$

(see for instance [3] for details).

We also take from [30, Lemma 1.3] some uniform estimates on the solutions of (5.18) which will be useful in the next propositions.

Lemma 5.1. *There exists $C > 0$ independent of m_0 and T such that, if (u, m) is a classical solution of (5.18), then*

- $\|Du\|_{L^\infty([0,T] \times \mathbb{T}^d)} + \|D^2 u\|_{L^\infty([0,T] \times \mathbb{T}^d)} \leqslant C$

- $\mathbf{d}(m(s), m(l)) \leqslant C|l - s|^{1/2}$ *for any $l, s \in [0, T]$*

Consequently, we also have that $|\partial_t u(s, \cdot)| \leqslant C$ for any $s \in [0, T]$.

Lemma 5.2. *For any (u, m) solution of the MFG system (5.18), there exists $c(u, m) \in \mathbb{R}$ such that, for any $t \in [0, T]$,*

$$\int_{\mathbb{T}^d} \Big(H(x, Du^T(t, x)) - \Delta u^T(t, x) \Big) m^T(t, dx) - \mathcal{F}(m^T(t)) = c(m, u).$$

Proof. As for any $t > 0$ both m^T and u^T are smooth in time and space, the integral

$$\int_{\mathbb{T}^d} H(x, Du(t, x)) m^T(t, x) + Du(t, x) \cdot Dm(t, x) dx - \mathcal{F}(m(t))$$

is well defined and we can derive it in time. Then,

$$\int_{\mathbb{T}^d} D_p H(x, Du(t, x)) \partial_t Du(t, x) m(t, x) + H(x, Du(t, x)) \partial_t m(t, x) +$$

$$\int_{\mathbb{T}^d} \partial_t Du(t, x) \cdot Dm(t, x) dx + Du(t, x) \cdot \partial_t Dm(t, x) - F(x, m(t)) \partial_t m(t, x).$$

Integrating by parts and rearranging the terms we get that the above expression is equal to

$$\int_{\mathbb{T}^d} (-\Delta m(t, x) - \operatorname{div}(m(t, x) D_p H(x, Du^T(t, x)))) \partial_t u^T(t, x) dx +$$

$$(-\Delta u^T(t, x) + H(x, Du^T(t, x) - F(x, m^T(t, x)))) \partial_t m^T(t, x) dx.$$

If we plug into the last equality the equations verified by u^T and m^T, we get

$$\frac{d}{dt} \left(\int_{\mathbb{T}^d} H(x, Du^T(t, x)) m^T(t, x) + Du^T(t, x) \cdot Dm^T(t, x) dx - \mathcal{F}(m^T(t)) \right) =$$

$$\int_{\mathbb{T}^d} -\partial_t m^T(t, x) \partial_t u^T(t, x) + \partial_t m^T(t, x) \partial_t u^T(t, x) = 0.$$

Integrating by parts the term $Du^T(t, x) \cdot Dm^T(t, x)$ and using the continuity of $t \to u(t, \cdot)$ in $C^2(\mathbb{T}^d)$ and the continuity of $t \to m(t)$ in $\mathcal{P}(\mathbb{T}^d)$ we conclude that the result holds. ∎

Proposition 5.3. *Let (m^T, α^T) be optimal for $\mathcal{U}(T, m_0)$ and (u^T, m^T) be a solution of (5.18) associated to (m^T, α^T). Then, $c(u^T, m^T) \to \lambda$ as $T \to +\infty$. Moreover, this limit is uniform with respect to the initial condition m_0 and the choice of the minimizer (m^T, α^T).*

Proof. We prove it by contradiction. Let us suppose that there exist a sequence $T_i \to +\infty$ and a sequence (m^i, α^i), minimizing $\mathcal{U}^{T_i}(0, m_0^i)$, such that, for any $i \in \mathbb{N}$ and a some $\varepsilon > 0$,

$$|c(u^i, m^i) - \lambda| \geqslant \varepsilon, \tag{5.19}$$

where, as usual, $\alpha^i(s, x) = D_p H(x, Du^i(s, x))$ and (u^i, m^i) solves (5.18). Thanks to Lemma 5.1 we know that there exists $C > 0$, independent of i such that

$$\sup_{T_i > 0} \sup_{t \in [0, T_i]} \|\alpha^i(t)\|_\infty + \|D\alpha^i(t)\|_\infty \leqslant C.$$

Let E be the set

$$E := \{\alpha \in W^{1, \infty}(\mathbb{T}^d, \mathbb{R}^d), \|\alpha\|_\infty + \|D\alpha\|_\infty \leqslant C\}.$$

16

Then E, endowed with the topology of the uniform convergence, is compact. Moreover, $\alpha^i(t) \in E$ for any $t \in [0, T_i]$. Let us introduce the probability measure ν^i on $\mathcal{P}(\mathbb{T}^d) \times E$ by

$$\int_{\mathcal{P}(\mathbb{T}^d) \times E} f(m, \alpha) \nu^i(dm, d\alpha) = \frac{1}{T_i - 1} \int_1^{T_i} f(m^i(t), \alpha^i(t)) dt.$$

Then ν^i converges, up to a subsequence denoted in the same way, to some probability measure ν on $\mathcal{P}(\mathbb{T}^d) \times E$. Note that

$$\frac{1}{T_i} \mathcal{U}(0, m_0^i) = \frac{1}{T_i} \int_0^1 \left(\int_{\mathbb{T}^d} H^*(y, \alpha^i(s, y)) m^i(s, y) dy \right) + \mathcal{F}(m^i(s))) \ ds$$
$$+ \frac{T_i - 1}{T_i} \int_{\mathcal{P}(\mathbb{T}^d) \times E} \left(\int_{\mathbb{T}^d} H^*(y, \alpha(y)) m(dy) + \mathcal{F}(m) \right) \ \nu^i(dm, d\alpha)$$

Hence, as the left-hand side converges, uniformly with respect to m_0^i, to $-\lambda$ (see [30]), we obtain

$$\int_{\mathcal{P}(\mathbb{T}^d) \times E} \left(\int_{\mathbb{T}^d} H^*(y, \alpha(y)) m(dy) + \mathcal{F}(m) \right) \ \nu(dm, d\alpha) = -\lambda. \tag{5.20}$$

Now we make the link between ν and the measure (μ, p_1) of Section 4. Let μ be the first marginal of ν and let us define the vector measure p_1 on $\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d$ as

$$\int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} \phi(m, y) \cdot p_1(dm, dy) = \int_{\mathcal{P}(\mathbb{T}^d) \times E} \int_{\mathbb{T}^d} \phi(m, y) \cdot \alpha(y) m(dy) \nu(dm, d\alpha)$$

for any test function $\phi \in C^0(\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d, \mathbb{R}^d)$. We note that p_1 is absolutely continuous with respect to μ, since, if we disintegrate ν with respect to μ: $\nu = \nu_m(d\alpha) \mu(dm)$, then

$$p_1(dm, dy) = \int_E \alpha(y) m(dy) \nu_m(d\alpha) \mu(dm).$$

Therefore,

$$\frac{dp_1}{dm \otimes \mu}(m, y) = \int_E \alpha(y) \nu_m(d\alpha).$$

Let us check that (μ, p_1) is closed. Indeed, for any map $\Phi \in C^{1,1}(\mathcal{P}(\mathbb{T}^d))$, we have

$$\frac{d}{dt} \Phi(m^i(t)) = \int_{\mathbb{T}^d} \text{div}(D_m \Phi(m^i(t), y)) m^i(t, dy) - \int_{\mathbb{T}^d} D_m \Phi(m^i(t), y) \cdot H_p(y, Du^i(t, y)) m^i(dy)$$
$$= \int_{\mathbb{T}^d} \text{div}(D_m \Phi(m^i(t), y)) m^i(t, dy) + \int_{\mathbb{T}^d} D_m \Phi(m^i(t), y) \cdot \alpha^i(t, y) m^i(dy).$$

So,

$$\int_{\mathcal{P}(\mathbb{T}^d) \times E} \int_{\mathbb{T}^d} \text{div}(D_m \Phi(m, y)) + D_m \Phi(m, y) \cdot \alpha(y) m(dy) \ \nu^i(dm, d\alpha)$$
$$= \frac{1}{T_i - 1} \int_1^{T_i} \int_{\mathbb{T}^d} \text{div}(D_m \Phi(m^i(t), y)) m^i(t, dy) + \int_{\mathbb{T}^d} D_m \Phi(m^i(t), y) \cdot \alpha^i(t, y) m^i(dy) dt$$
$$= \frac{1}{T_i - 1} \left[\Phi(m^i(T_i)) - \Phi(m^i(1)) \right].$$

17

Letting $i \to +\infty$ gives

$$\int_{\mathcal{P}(\mathbb{T}^d) \times E} \int_{\mathbb{T}^d} \mathrm{div}(D_m \Phi(m, y)) + D_m \Phi(m, y) \cdot \alpha(y) m(dy) \, \nu(dm, d\alpha) = 0,$$

which can be rewritten, in view of the definition of p_1, as

$$\int_{\mathcal{P}(\mathbb{T}^d)} \int_{\mathbb{T}^d} \mathrm{div}(D_m \Phi(m, y)) m(dy) \mu(dm) + \int_{\mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d} D_m \Phi(m, y) \cdot p_1(dm, dy) = 0.$$

This proves that (μ, p_1) is closed. Next we come back to (5.20): using the convexity of H^*, we also have

$$-\lambda = \int_{\mathcal{P}(\mathbb{T}^d) \times E} \left[\int_{\mathbb{T}^d} H^*(y, \alpha(y)) m(dy) + \mathcal{F}(m) \right] \nu(dm, d\alpha) =$$

$$\int_{\mathcal{P}(\mathbb{T}^d)} \int_E \left[\int_{\mathbb{T}^d} H^*(y, \alpha(y)) m(dy) + \mathcal{F}(m) \right] \nu_m(d\alpha) \mu(dm) \geqslant$$

$$\int_{\mathcal{P}(\mathbb{T}^d)} \left[\int_{\mathbb{T}^d} H^* \left(y, \int_E \alpha(y) \nu_m(d\alpha) \right) m(dy) + \mathcal{F}(m) \right] \mu(dm) =$$

$$\int_{\mathcal{P}(\mathbb{T}^d)} \left[\int_{\mathbb{T}^d} H^* \left(y, \frac{dp_1}{dm \otimes \mu}(m, y) \right) m(dy) + \mathcal{F}(m) \right] \mu(dm).$$

Therefore,

$$\int_{\mathcal{P}(\mathbb{T}^d)} \left[\int_{\mathbb{T}^d} H^* \left(y, \frac{dp_1}{dm \otimes \mu}(m, y) \right) m(dy) + \mathcal{F}(m) \right] \mu(dm) \leqslant -\lambda, \qquad (5.21)$$

which proves the minimality of (μ, p_1). By the uniform convexity of H, relation (5.21) shows also that, for $\mu-$a.e. $m \in \mathcal{P}(\mathbb{T}^d)$ and for ν_m-a.e. α, one has

$$\frac{dp_1}{dm \otimes \mu}(m, y) = \alpha(y). \qquad (5.22)$$

Note also that, $m^i(t)$ has a positive density for any $t \in [1, T_i]$ and there exists a constant $C > 0$ independent of i such that

$$\sup_{t \in [1, T_i]} \|1/m^i(t, \cdot)\|_\infty + \|Dm^i(t, \cdot)\|_\infty \leqslant C. \qquad (5.23)$$

The bounds on Dm^i are standard and we refer to [26, Ch4, Theorem 5.1]. While, for the estimates on $1/m^i$, we used the Harnack's inequality in [2, Theorem 8.1.3]. In our setting, this theorem states that, for any $x, y \in \mathbb{T}^d$ and for any $0 < s < t < T_i$, there exists a constant C_{t-s}, depending only on $|t - s|$, such that

$$m^i(t, x) \geqslant C_{t-s} m^i(s, y).$$

As we work on the torus, for any $s > 0$, there exists a point $y_s \in \mathbb{T}^d$ such that $m^i(s, y_s) \geqslant 1$. Then, we can chose $s = t - 1$ and we get that for any $t > 1$

$$m^i(t, x) \geqslant C_1 m^i(t - 1, y_{t-1}) \geqslant C_1,$$

which proves (5.23).

The estimates in (5.23) ensure that the pair (μ, p_1) is smooth in the sense of Definition 4.1. In particular, we know by Proposition 4.2 that, for $\mu-$a.e. $m \in \mathcal{P}(\mathbb{T}^d)$,

$$-\int_{\mathbb{T}^d} q_1(y, m) \cdot Dm(y) dy + \int_{\mathbb{T}^d} H(y, q_1(y, m)) m(dy) - \mathcal{F}(m) = \lambda,$$

where

$$q_1(x, m) := D_a H^* \left(y, \frac{dp_1}{dm \otimes \mu}(y, m) \right).$$

By the convergence of ν^i to ν, there exists (up to a subsequence again) $t_i \in [1, T_i]$ such that $(m^i(t_i), \alpha^i(t_i))$ converges to an element $(m, \alpha) \in \mathcal{P}(\mathbb{T}^d) \times E$ which belongs to the support of μ. Then by (5.22), $\alpha = \frac{dp_1}{dm \otimes \mu}$. Thus $Du^i(t_i) = D_a H^*(y, \alpha^i(t_i))$ converges uniformly to $q_1(\cdot, m)$. This shows that

$$\lim_{i \to +\infty} c(u^i, m^i)$$

$$= \lim_{i \to +\infty} -\int_{\mathbb{T}^d} Du^i(t_i, y) \cdot Dm^i(t_i, y) dy + \int_{\mathbb{T}^d} H(y, Du^i(t_i, y)) m^T(t_i, dy) - \mathcal{F}(m^i(t_i))$$

$$= -\int_{\mathbb{T}^d} q_1(y, m) \cdot Dm(y) dy + \int_{\mathbb{T}^d} H(y, q_1(y, m)) m(dy) - \mathcal{F}(m) = \lambda,$$

which is in contradiction with (5.19). ∎

The next step towards Theorem 5.7 is to prove that the map $(s, m) \to \mathcal{U}^T(s, m) + \lambda(T - s)$ has a limit. In the next proposition we prove that $(s, m) \to \mathcal{U}^T(s, m) + \lambda(T - s)$ is bounded and equicontinuous on $[0, T] \times \mathcal{P}(\mathbb{T}^d)$ and so that there exists a subsequence $(\mathcal{U}^{T_n} + \lambda(T_n - \cdot))$ which, locally in time, converges uniformly to a continuous function ξ.

Proposition 5.4. *The maps $(s, m) \to \mathcal{U}^T(s, m) + \lambda(T - s)$ are uniformly bounded and uniformly continuous.*

Proof. We first prove that $(s, m) \to \mathcal{U}^T(s, m) + \lambda(T - s)$ is bounded, uniformly in T. Let χ be a corrector function. As χ is a continuous function on the compact set $\mathcal{P}(\mathbb{T}^d)$, there exists a constant $C > 0$ such that $0 \leq \chi(m) + C$ for any $m \in \mathcal{P}(\mathbb{T}^d)$. If $(m(t), w(t))$ is an admissible trajectory for the minimization problem of $\mathcal{U}^T(s, m)$, then

$$\int_s^T \int_{\mathbb{T}^d} H^*(x, \alpha(t, x)) dm(t) + \mathcal{F}(m(t)) dt + \lambda(T - s)$$

$$\leq \int_s^T \int_{\mathbb{T}^d} H^*(x, \alpha(t, x)) dm(t) + \mathcal{F}(m(t)) dt + \chi(m(T)) + \lambda(T - s) + C.$$

Taking the infimum over all the possible (m, α), the definition of $\mathcal{U}^T(s, m)$ and the dynamic programming principle verified by χ lead to

$$\mathcal{U}^T(s, m) + \lambda(T - s) \leq \chi(m) + C.$$

As χ is bounded, we get an upper bound independent of T, m and s. The lower bound is analogous.

We turn to the equicontinuity. For what concern the continuity in the m variable, one can adapt the proof of [30, Theorem 1.5] with minor adjustments, to show that, if $T - s \geq \varepsilon > 0$

19

for given $\varepsilon > 0$, then there exits a constant K independent of T and s such that $\mathcal{U}^T(s, \cdot)$ is K-Lipschitz continuous.

We now need to estimate $|\mathcal{U}^T(t_2, m_0) - \mathcal{U}^T(t_1, m_0)|$. We suppose $t_2 > t_1$ and we fix $(\bar{m}(s), \bar{\alpha}(s))$ an optimal trajectory for $\mathcal{U}^T(t_1, m_0)$, then

$$|\mathcal{U}^T(t_2, m_0) - \mathcal{U}^T(t_1, m_0)| \leqslant |\mathcal{U}^T(t_2, m_0) - \mathcal{U}^T(t_2, \bar{m}(t_2))| + |\mathcal{U}^T(t_2, \bar{m}(t_2)) - \mathcal{U}^T(t_1, m_0)|. \quad (5.24)$$

We can estimate the first term on the right hand-side using at first the uniform Lipschitz continuity we discussed before and then the estimates on the solution of the MFG system in Lemma 5.1. So,

$$|\mathcal{U}^T(t_2, m_0) - \mathcal{U}^T(t_2, \bar{m}(t_2))| \leqslant K\mathbf{d}(m_0, \bar{m}(t_2)) \leqslant KC|t_1 - t_2|^{\frac{1}{2}}. \quad (5.25)$$

To estimate the second term in the right hand-side of (5.24), we just need to use that

$$\mathcal{U}^T(t_1, m_1) = \inf_{(m,\alpha)} \left\{ \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*\left(x, \alpha(t, x)\right) dm(t) + \mathcal{F}(m(t))dt + \mathcal{U}^T(t_2, m(t_2)) \right\}.$$

As $(\bar{m}, \bar{\alpha})$ is optimal for $\mathcal{U}^T(t_1, m_1)$, we get

$$\mathcal{U}^T(t_1, m_1) - \mathcal{U}^T(t_2, \bar{m}(t_2)) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*\left(x, \bar{\alpha}(t, x)\right) d\bar{m}(t) + \mathcal{F}(\bar{m}(t))dt.$$

Note that, $\bar{\alpha}(t, x) = D_p H(x, D\bar{u}(x, t))$ where $(\bar{m}, \bar{u})$ solves the MFG system (5.18) and, according to Lemma 5.1, we have uniform estimates on $\bar{u}$. Therefore,

$$|\mathcal{U}^T(t_2, \bar{m}(t_2)) - \mathcal{U}^T(t_1, m_0)| \leqslant \int_{t_1}^{t_2} \left| \int_{\mathbb{T}^d} H^*\left(x, \bar{\alpha}(t, x)\right) d\bar{m}(t) + \mathcal{F}(\bar{m}(t)) \right| dt \leqslant C(t_2 - t_1).$$

Putting together the last inequality with (5.25) we have that, for a possibly different constant $C > 0$, independent of T, m_0 and m_1,

$$|\mathcal{U}^T(t_2, m_0) - \mathcal{U}^T(t_1, m_0)| \leqslant C(|t_1 - t_2|^{\frac{1}{2}} + |t_1 - t_2|),$$

which, in turn, implies the uniform continuity in time. $\blacksquare$

From now on we fix a continuous map $\xi : [0, +\infty) \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$, limit of a subsequence (denoted in the same way) of the sequence $(\mathcal{U}^T + \lambda(T - \cdot))$ as $T \to +\infty$.

Proposition 5.5. *The map* $\xi_0(\cdot) := \xi(0, \cdot)$ *satisfies*

$$\xi_0(m) \leqslant \inf_{(m,\alpha)} \left\{ \int_0^t \left(\int_{\mathbb{T}^d} H^*\left(y, \alpha(s, y)\right) m(s, y) dy + \mathcal{F}(m(s)) \right) ds + \xi_0(m(t)) \right\} + \lambda t.$$

Proof. We first claim that ξ is a viscosity solution to

$$-\partial_t \xi \geqslant 0 \qquad \text{in } [0, +\infty) \times \mathcal{P}(\mathbb{T}^d). \quad (5.26)$$

Let $\Phi = \Phi(t, m)$ be a smooth test function such that $\xi \geqslant \Phi$ with an equality only at (t_0, m_0). Then there exists a subsequence $(t_n, \bar{m}_n)$ converging to (t_0, m_0) and such that $\mathcal{U}^{T_n} + \lambda(T_n - s) - \Phi$ has a minimum at $(t_n, \bar{m}_n)$. Let (m_n, α_n) be a minimizer for $\mathcal{U}^{T_n}(t_n, \bar{m}_n)$. We consider

$u_n \in C^{1,2}([t_n, T_n] \times \mathbb{T}^d)$ such that (u_n, m_n) is a solution to the MFG system (5.18) and $\alpha_n = D_p H(x, Du_n)$. Then, by Lemma 6.1, we have

$$-\partial_t \Phi(t_n, \bar{m}_n) - \lambda + \int_{\mathbb{T}^d} (H(y, Du_n(t_n, y)) - \Delta u_n(t_n, y)) \bar{m}_n(dy) - \mathcal{F}(\bar{m}_n) \geqslant 0.$$

By Lemma 5.2 and Proposition 5.3, we have, given $\varepsilon > 0$,

$$\int_{\mathbb{T}^d} (H(y, Du_n(t_n, y)) - \Delta u_n(t_n, y)) \bar{m}_n(dy) - \mathcal{F}(\bar{m}_n) \leqslant \lambda + \varepsilon,$$

for n large enough. So

$$-\partial_t \Phi(t_n, \bar{m}_n) \geqslant -\varepsilon.$$

We obtain therefore, after letting $n \to +\infty$ and then $\varepsilon \to 0$,

$$-\partial_t \Phi(t_0, m_0) \geqslant 0.$$

This shows that ξ satisfies (5.26) holds in the viscosity solution sense.

We now prove that (5.26) implies that ξ is nonincreasing in time. Fix $m_0 \in \mathcal{P}(\mathbb{T}^d)$ and assume on the contrary that there exists $0 \leqslant t_1 < t_2$ such that $\xi(t_1, m_0) < \xi(t_2, m_0)$. Let $\Psi = \Phi(m)$ be a smooth test function such that $\Psi > 0$ on $\mathcal{P}(\mathbb{T}^d) \backslash \{m_0\}$ with $\Phi(m_0) = 0$. Then, we can find $\eta > 0$ small such that, if m_1 and m_2 are such that $\xi(t_1, \cdot) + \eta^{-1}\Psi$ has a minimum at m_1 and $\xi(t_2, \cdot) + \eta^{-1}\Psi$ has a minimum at m_2, then $\xi(t_1, m_1) < \xi(t_2, m_2)$. Note that this implies that

$$\min_{m \in \mathcal{P}(\mathbb{T}^d)} \xi(t_1, m) + \eta^{-1}\Psi(m) = \xi(t_1, m_1) + \eta^{-1}\Psi(m_1)$$
$$< \xi(t_2, m_2) + \eta^{-1}\Psi(m_2) = \min_{m \in \mathcal{P}(\mathbb{T}^d)} \xi(t_2, m) + \eta^{-1}\Psi(m).$$

Recalling that ξ is bounded, this implies that we can find $\varepsilon > 0$ small such that the map $(t, m) \to \xi(t, m) + \eta^{-1}\Psi(m) + \varepsilon t$ has an interior minimum on $[t_1, +\infty) \times \mathcal{P}(\mathbb{T}^d)$ at some point $(t_3, m_3) \in (t_1, +\infty) \times \mathcal{P}(\mathbb{T}^d)$. This contradicts (5.26).

Now that we have proved the monotonicity in time of ξ we can finally show the statement of the proposition. We have that $\mathcal{U}^T$ verifies the following dynamic programming principle

$$\mathcal{U}^T(0, m_0) = \inf_{(m, \alpha)} \left\{ \int_0^t \int_{\mathbb{T}^d} H^*(x, \alpha(t, x)) \, dm(t) + \mathcal{F}(m(t)) dt + \mathcal{U}^T(t, m(t)) \right\}. \tag{5.27}$$

Then, adding on both sides λT and passing to the limit $T \to +\infty$, one easily checks that ξ satisfies

$$\xi(0, m) = \inf_{(m, \alpha)} \left\{ \int_0^t \left(\int_{\mathbb{T}^d} H^*(y, \alpha(s, y)) \, m(s, y) dy + \mathcal{F}(m(s)) \right) ds + \xi(t, m(t)) \right\} + \lambda t.$$

Using the fact that ξ is nonincreasing in time we get the desired result. ∎

Before we can prove that ξ_0 is a corrector function we need to state some standard properties of $\tau_h : C^0(\mathcal{P}(\mathbb{T}^d)) \to C^0(\mathcal{P}(\mathbb{T}^d))$ which is defined as follows. For $h > 0$ and $\Phi \in C^0 \mathbb{T}^d)$ we set

$$\tau_h \Phi(m_0) = \inf_{(m, \alpha)} \left\{ \int_0^h \left(\int_{\mathbb{T}^d} H^*(y, \alpha(s, y)) \, m(s, y) dy + \mathcal{F}(m(s)) \right) ds + \Phi(m(t)) \right\} + \lambda h,$$

where (m, α) solves in the sense of distribution

$$\begin{cases} -\partial_t m + \Delta m + \text{div}(m\alpha) = 0 & \text{in } [0, h] \times \mathbb{T}^d \\ m(0) = m_0. \end{cases}$$

Lemma 5.6. *The function τ_h verifies the following properties*

1. *For any h_1, $h_2 > 0$, $\tau_{h_1} \circ \tau_{h_2} = \tau_{h_1 + h_2}$*

2. *For any $h > 0$, τ_h is not expansive, i.e. for any Φ, $\Psi \in C^0\mathcal{P}(\mathbb{T}^d))$*

$$\|\tau_h \Phi - \tau_h \Psi\|_\infty \leqslant \|\Phi - \Psi\|_\infty.$$

3. *For any $h > 0$, τ_h is order preserving, i.e. for any Φ, $\Psi \in C^0\mathcal{P}(\mathbb{T}^d))$ such that $\Phi \leqslant \Psi$,*

$$\tau_h \Phi \leqslant \tau_h \Psi.$$

4. *Let $\Phi \in C^0\mathcal{P}(\mathbb{T}^d))$ be such that, for any $h > 0$, $\Phi \leqslant \tau_h \Phi$. Then, for any $0 < h_1 < h_2$,*

$$\Phi \leqslant \tau_{h_1} \Phi \leqslant \tau_{h_2} \Phi.$$

Proof. The proof is standard, see for instance, in a closely related context, [15]. ∎

Theorem 5.7. *ξ_0 is a corrector and $\mathcal{U}^T(0) + \lambda T$ converges uniformly to ξ_0 on $\mathcal{P}(\mathbb{T}^d)$.*

Proof. The proof follows closely the one of [15, Theorem 6.3.1]. We define $\widetilde{\mathcal{U}}^T(t, m) = \mathcal{U}^T(t, m) + \lambda(T - t)$. Let $T_n \to +\infty$ be a sequence such that $\widetilde{\mathcal{U}}^{T_n}$ converges locally uniformly to ξ on $[0, +\infty) \times \mathcal{P}(\mathbb{T}^d)$. We can suppose that, if we define $s_n = T_{n+1} - T_n$, then $s_n \to +\infty$. Note that $\widetilde{\mathcal{U}}^{T_{n+1}}(s_n, m) = \widetilde{\mathcal{U}}^{T_n}(0, m)$. Then, using (5.27), we get

$$\widetilde{\mathcal{U}}^{T_{n+1}}(0, m) = \tau_{s_n} \widetilde{\mathcal{U}}^{T_{n+1}}(s_n, m) = \tau_{s_n} \widetilde{\mathcal{U}}^{T_n}(0, m).$$

We also know from Lemma 5.6 that τ_h is a contraction and that it verifies the semigroup property. Therefore,

$$\|\tau_{s_n} \xi_0 - \xi_0\|_\infty \leqslant \|\tau_{s_n} \xi_0 - \tau_{s_n} \widetilde{\mathcal{U}}^{T_n}(0)\|_\infty + \|\tau_{s_n} \widetilde{\mathcal{U}}^{T_n}(0) - \xi_0\|_\infty$$
$$\leqslant \|\xi_0 - \widetilde{\mathcal{U}}^{T_n}(0)\|_\infty + \|\widetilde{\mathcal{U}}^{T_{n+1}}(0) - \xi_0\|_\infty \to 0,$$

Moreover, Proposition 5.5 and Lemma 5.6 prove that $\tau_s \xi_0$ is monotone in s. Then, for any $s > 0$, we have that, for a sufficiently large $n \in \mathbb{N}$,

$$\xi_0 \leqslant \tau_s \xi_0 \leqslant \tau_{s_n} \xi_0 \to \xi_0,$$

which proves that $\tau_t \xi_0$ is constant in t and so ξ_0 is corrector function. It remains to check that the whole sequence $\widetilde{\mathcal{U}}^T(0)$ converges to ξ_0. Let $T > T_n$, then

$$\|\widetilde{\mathcal{U}}^T(0) - \xi_0\|_\infty = \|\tau_{T - T_n} \widetilde{\mathcal{U}}^{T_n}(0) - \tau_{T - T_n} \xi_0\|_\infty \leqslant \|\widetilde{\mathcal{U}}^{T_n}(0) - \xi_0\|_\infty \to 0$$

and the result follows. ∎

5.2 Convergence of optimal trajectories

Now that we have proved the convergence of $\mathcal{U}^T(0, \cdot) + \lambda T$ to a corrector function χ, we can properly define the limit trajectories for time dependent MFG and the set where these trajectories lay. We will show that this set is a subset of the projected Mather set $\mathcal{M}$ as was suggested in [30]. We recall the definition of $\mathcal{M}$.

Definition 5.8. *We say that $m_0 \in \mathcal{P}(\mathbb{T}^d)$ belongs to the projected Mather set $\mathcal{M} \subset \mathcal{P}(\mathbb{T}^d)$ if there exists a calibrated curve $(m(t), \alpha(t))$ such that $m(0) = m_0$.*

Remark 5.1. *Note that the notion of projected Mather set that we use here is consistent with the one that was already introduced in [30]. On the other hand, Definition 5.8 is not the transposition of the definition of projected Mather set that is generally used in standard Weak KAM theory. In this latter case the projected Mather set is the union of the projection of the supports of Mather measures on the torus. What we call here projected Mather set would be rather the projected Aubry set or the projected Mané set (we refer to [16] and [15] for these definitions). We decided to use Definition 5.8 mostly to be consistent with the terminology in [30]. Moreover, it is worthwhile to mention that in the standard theory the Mather set and the Aubry set are deeply connected while in this framework such a relation is no longer clear. In particular, in standard Weak KAM theory the Mather set is contained in the Aubry set (where the latter is defined as the intersection of graphs of calibrated curves). One can check this inclusion defining a calibrated curve, starting from any point of the Mather set, through the Lagrangian flow. In the MFG setting, the lack of uniqueness of solutions and the forward/backward structure of the system prevent from defining any sensible notion of flow. Moreover, an other important difference, that highlights how the connection between Mather set and Aubry set is not clear in the MFG framework, is that, on the one hand, we know that calibrated curves lay on smooth probability measure but, on the other, we know nothing about the regularity of Mather measures' support points (reason why we introduced the notion of "smooth" Mather measure).*

We recall also that a couple $(\bar{m}, \bar{\alpha})$, which satisfies $-\partial_t \bar{m}(t) + \Delta \bar{m}(t) + \mathrm{div}(\bar{\alpha}(t)\bar{m}(t)) = 0$ for any $t \in \mathbb{R}$, is a calibrated curve, if there exists a corrector function $\chi : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ such that, for any $t_1 < t_2 \in \mathbb{R}$,

$$\chi(\bar{m}(t_1)) = \lambda(t_2 - t_1) + \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(s))\, d\bar{m}(s) + \mathcal{F}(\bar{m}(s))ds + \chi(\bar{m}(t_2)).$$

We fix (m^T, α^T) a minimizer for $\mathcal{U}^T(-T, m_0)$. As usual $\alpha^T = D_p H(x, D\bar{u}^T)$ where $(\bar{u}^T, m^T)$ solves (5.18) on $[-T, T] \times \mathbb{T}^d$. We define $u^T(t, x) = \bar{u}^T(t, x) - u^T(0, \bar{x})$, for a fixed $\bar{x} \in \mathbb{T}^d$. We know from Lemma 5.1 that $D^2 \bar{u}^T$ and $\partial_t \bar{u}^T$ are uniformly bounded. This means that u^T and Du^T are uniformly bounded and uniformly continuous on any compact set of $\mathbb{R} \times \mathbb{T}^d$. Therefore, we have that up to subsequence u^T converges to a function $u \in C^{1,2}(\mathbb{R} \times \mathbb{T}^d)$. The convergence, up to subsequence, of m^T to a function $m \in C^0(\mathbb{R}, \mathcal{P}(\mathbb{T}^d))$ is ensured again by Lemma 5.1 and the uniform $C^{\frac{1}{2}}([0, T], \mathcal{P}(\mathbb{T}^d))$ bounds on m^T therein. It is standard that the couple (u, m) solves in classical sense

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{R} \times \mathbb{T}^d, \\ -\partial_t m + \Delta m + \mathrm{div}(m D_p H(x, Du)) = 0 & \text{in } \mathbb{R} \times \mathbb{T}^d. \end{cases}$$

As by product we have that (m^T, α^T) uniformly converges on compact sets to the couple $(m, D_p H(x, Du))$. Moreover, if we define $\alpha = D_p H(x, Du)$, then (m, α) solves $-\partial_t m + \Delta m + \mathrm{div}(m\alpha) = 0$. We can now prove that (m, α) is a calibrated curve and therefore that $\mathcal{M}$ contains the uniform limits of optimal trajectories.

Theorem 5.9. *Let (m^T, α^T) be an optimal trajectory for $\mathcal{U}^T(-T, m_0)$. Then, (m^T, α^T) converges, up to subsequence, to a calibrated curve (m, α). Consequently, $m(t) \in \mathcal{M}$ for any $t \in \mathbb{R}$.*

Proof. As we have already discussed the convergence of (m^T, α^T) to (m, α) we just need to check that (m, α) is a calibrated curve. We fix $t_1 < t_2 \in \mathbb{R}$, then, by dynamic programming principle,

$$\mathcal{U}^T(t_1, m_0) + \lambda(T - t_1) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*\left(x, \alpha^T(s)\right) dm^T(s) + \mathcal{F}(m^T(s))ds + \mathcal{U}^T(t_2, m^T(t_2)) + \lambda(T - t_1).$$

We recall that $\mathcal{U}^T(t, m_0) = \mathcal{U}^{T-t}(0, m_0)$. Given the continuity of $\mathcal{U}^T(0, \cdot)$, the uniform convergence of (m^T, α^T) on compact subsets and the uniform convergence of $\mathcal{U}^T(0, \cdot) + \lambda T$ to $\chi(\cdot)$, we can pass to the limit in T and we get that, for any interval $[t_1, t_2]$, the couple (m, α) verifies

$$\chi(m(t_1)) = \int_{t_1}^{t_2} \int_{\mathbb{T}^d} H^*\left(x, \alpha(s)\right) dm(s) + \mathcal{F}(m(s))ds + \chi(m(t_2)) + \lambda(t_2 - t_1).$$

So (m, α) is a calibrated curve. ∎

6 Appendix

6.1 A viscosity solution property

Lemma 6.1. *Let $\Phi = \Phi(t, m)$ be a smooth test function such that $\mathcal{U}^T - \Phi$ has a minimum at a point $(t_0, m_0) \in [0, T) \times \mathcal{P}(\mathbb{T}^d)$. Let (u, m) be a solution of the MFG system (5.18) starting from (t_0, m_0) and such that $(m, D_pH(x, Du(t, x))$ is optimal for $\mathcal{U}(t_0, m_0)$. Then*

$$D_m\Phi(t_0, m_0, x) = Du(t_0, x) \text{ for } m_0-\text{a.e. } x \in \mathbb{T}^d \tag{6.28}$$

and

$$-\partial_t\Phi(t_0, m_0) + \int_{\mathbb{T}^d} (H(x, Du(t_0, x) - \Delta u(t_0, x))m_0(dx) - \mathcal{F}(m_0) \geqslant 0. \tag{6.29}$$

Proof. Without loss of generality we assume that $\Phi(t_0, m_0) = \mathcal{U}^T(t_0, m_0)$. Let $m'(t_0) \in \mathcal{P}(\mathbb{T}^d)$ and $m' = m'(t)$ be the solution to

$$\partial_t m' - \Delta m' - \text{div}(m' D_pH(x, Du(t, x)) = 0.$$

We set $\mu(t) = m'(t) - m(t)$ and, for $h \in (0, 1]$ and note that the pair $(m + h\mu, D_pH(x, Du(t, x))$ is a solution to $-\partial_t m + \Delta m + \text{div}(m\alpha) = 0$ with initial condition $m_0^h := (1 - h)m_0 + hm'(t_0)$. Hence, by the definition of Φ and $\mathcal{U}^T$ we have

$$\Phi(t_0, m_0^h) \leqslant \mathcal{U}^T(t_0, m_0^h)$$

$$\leqslant \int_{t_0}^T (\int_{\mathbb{T}^d} H^*(x, D_pH(x, Du(t, x))))(m + h\mu)(t, dx) + \mathcal{F}((m + h\mu)(t))dt$$

$$\leqslant \Phi(t_0, m_0) + h\Big(\int_{t_0}^T \int_{\mathbb{T}^d} (H^*(x, D_pH(x, Du(t, x))) + F(x, m(t)))\mu(t, dx)dt + o_h(1)\Big).$$

Next we use the equation for u and then for μ:

$$\int_{t_0}^T \int_{\mathbb{T}^d} (H^*(x, D_pH(x, Du(t, x))) + F(x, m(t)))\mu(t, dx)dt$$

$$= \int_{t_0}^T \int_{\mathbb{T}^d} (H^*(x, D_pH(x, Du(t, x))) - \partial_t u - \Delta u + H(x, Du(t, x)))\mu(t, dx)dt$$

$$= \int_{\mathbb{T}^d} u(t_0, x)\mu(t_0, dx) = \int_{\mathbb{T}^d} u(t_0, x)(m'(t_0) - m_0)(dx).$$

Plugging this into the estimate of $\Phi(t_0, m_0^h)$ above, we obtain, dividing by h and letting $h \to 0$,

$$\int_{\mathbb{T}^d} \frac{\delta\Phi(t_0, m_0, x)}{\delta m}(m'(t_0) - m_0)(dx) \leqslant \int_{\mathbb{T}^d} u(t_0, x)(m'(t_0) - m_0)(dx).$$

Recalling the convention (1.8) on the derivative and the arbitrariness of $m'(t_0)$, we infer, by choosing Dirac masses for $m'(t_0)$, that

$$\frac{\delta\Phi(t_0, m_0, x)}{\delta m} \leqslant u(t_0, x) - \int_{\mathbb{T}^d} u(t_0, y)m_0(dy) \qquad \forall x \in \mathbb{T}^d,$$

while

$$\int_{\mathbb{T}^d} \frac{\delta\Phi(t_0, m_0, x)}{\delta m} m_0(dx) = \int_{\mathbb{T}^d} \left(u(t_0, x) - \int_{\mathbb{T}^d} u(t_0, y)m_0(dy) \right) m_0(dx) = 0.$$

Therefore
$$\frac{\delta\Phi(t_0, m_0, x)}{\delta m} = u(t_0, x) - \int_{\mathbb{T}^d} u(t_0, y)m_0(dy), \qquad m_0-\text{a.e. } x \in \mathbb{T}^d.$$

This shows that the map $x \to \frac{\delta\Phi(t_0, m_0, x)}{\delta m} - u(t_0, x)$ has a maximum on $\mathbb{T}^d$ at m_0-a.e. $x \in \mathbb{T}^d$ and thus (6.28) holds.

As $\mathcal{U}^T$ satisfies a dynamic programming principle and $\mathcal{U}^T - \Phi$ has a minimum at (t_0, m_0), it is standard that Φ also satisfies

$$-\partial_t\Phi(t_0, m_0) + \int_{\mathbb{T}^d} H(x, D_m\Phi(t_0, m_0, x))m_0(dx) - \int_{\mathbb{T}^d} \text{div}_y D_m\Phi(t_0, m_0, x)m_0(dx) - \mathcal{F}(m_0) \geqslant 0.$$

Using (6.28) one then infers that (6.29) holds. $\blacksquare$

6.2 Smooth test functions

Here we fix a corrector χ and construct a smooth function that touches χ from above. We fix $m_0 \in \mathcal{P}(\mathbb{T}^d)$ and $\tau > 0$. We know from [30, Appendix] that, if $(\bar{m}, \bar{\alpha})$ verifies

$$\chi(m_0) = \int_0^{2\tau} \int_{\mathbb{T}^d} H^*(x, \bar{\alpha})d\bar{m}(s) + \mathcal{F}(\bar{m}(s))ds + 2\lambda\tau + \chi(\bar{m}_{2\tau}), \qquad (6.30)$$

then there exists a couple $(\bar{u}, \bar{m})$ which solves the MFG system

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m) & \text{in } \mathbb{T}^d \times [0, 2\tau], \\ -\partial_t m + \Delta m + \text{div}(mD_pH(x, Du)) = 0 & \text{in } \mathbb{T}^d \times [0, 2\tau], \\ m(0) = m_0, \end{cases} \qquad (6.31)$$

such that $\bar{\alpha} = D_pH(x, D\bar{u})$.

For any $m_1 \in \mathcal{P}(\mathbb{T}^d)$ we define $m(t)$ and $\alpha(t)$ as follows. We first consider $\tilde{m}$ solution of

$$\begin{cases} -\partial_t\tilde{m} + \Delta\tilde{m} + \text{div}(\tilde{m}D_pH(x, D\bar{u})) = 0 & \text{in } [0, \tau] \times \mathbb{T}^d \\ \tilde{m}(0) = m_1 \end{cases} \qquad (6.32)$$

and then we set

$$m(s, x) = \begin{cases} \tilde{m}(s, x), & \text{if } s \in (0, \tau] \\ \frac{2\tau - s}{\tau}\tilde{m}(s, x) + \frac{s - \tau}{\tau}\bar{m}(s, x), & \text{if } s \in [\tau, 2\tau]. \end{cases}$$

Let $\zeta : [0, 2\tau] \times \mathbb{T}^d \to \mathbb{R}$ be the solution to $\Delta\zeta(t) = \bar{m}(t) - \tilde{m}(t)$ so that $\int_{\mathbb{T}^d} \zeta(s, x) = 0$. Then, the drift α will be $\alpha(t, x) = D_pH(x, D\bar{u}(t, x)) + \frac{D\zeta(t,x)}{\tau\, m(t,x)}$ in $[\tau, 2\tau]$ and $\alpha(t, x) = D_pH(x, D\bar{u}(t, x))$ elsewhere.

We define the function $\Psi(m_1)$ as

$$\Psi(m_1) = \int_0^{2\tau} \int_{\mathbb{T}^d} H^*(x, \alpha(t))dm(t) + \mathcal{F}(m(t))dt + 2\lambda\tau + \chi(\bar{m}(2\tau)). \tag{6.33}$$

Proposition 6.2. *The function $\Psi : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ defined in (6.33) is twice differentiable with respect to m with C^2 continuous derivatives in space and with derivatives bounded independently of χ.*

Proof. We first introduce $\Gamma : \mathbb{R}^+ \times \mathbb{T}^d \times \mathbb{R}^+ \times \mathbb{T}^d \to \mathbb{R}$, the fundamental solution of (6.32), i.e. $\Gamma(\cdot, \cdot; s, x)$ is the solution of (6.32) starting at time s with initial condition $\Gamma(s, y; s, x) = \delta_x(y)$. Then, by superposition, the solution $\bar{m}(t)$ of (6.32) is given by $\bar{m}(t, x) = \int_{\mathbb{T}^d} \Gamma(t, x; 0, y)m_1(dy)$ (for $t > 0$).

We consider separately the following two integrals:

$$I_1(m_1) = \int_0^\tau \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(t))d\bar{m}(t) + \mathcal{F}(\bar{m}(t))dt \tag{6.34}$$

and

$$I_2(m_1) = \int_\tau^{2\tau} \int_{\mathbb{T}^d} H^*(x, \alpha(t))dm(t) + \mathcal{F}(m(t))dt.$$

Note that $\Psi = I_1 + I_2 + 2\lambda\tau + \chi(\bar{m}(2\tau))$.

If we plug Γ into (6.34), then

$$I_1(m_1) = \int_0^\tau \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(t))\Gamma(t, x, 0, y)m_1(dy)dxdt + \mathcal{F}\left(\int_{\mathbb{T}^d} \Gamma(t, \cdot, 0, y)m_1(dy)\right) dt$$

We can now derive I_1 with respect to m_1 andwe get

$$\frac{\delta I_1}{\delta m}(m_1, y) = \int_0^\tau \int_{\mathbb{T}^d} H^*(x, \bar{\alpha}(t))\Gamma(t, x, 0, y)dxdt + \int_0^\tau \int_{\mathbb{T}^d} F(x, \bar{m}(t))\Gamma(t, x, 0, y)dxdt.$$

As the functions $\bar{\alpha}$ and $\bar{m}$ are smooth, standar results in parabolic equation ensures that $\delta I_1/\delta m(m_1, \cdot)$ is smooth (see for instance Chapter 4§14 in [26]).

We now focus on I_2. We fix $G : \mathbb{T}^d \to \mathbb{R}$ the kernel associated to the integral representation of the solution of the Poisson equation (see for instance [1, Theorem 4.13]). Explicitly, if $\Delta\zeta = f$, then

$$\zeta(x) = \int_{\mathbb{T}^d} G(x - y)f(y)dy$$

We first analyse the integral $\int_\tau^{2\tau} \int_{\mathbb{T}^d} H^*(x, \alpha(t))dm(t)$, which explicitly becomes

$$\int_\tau^{2\tau} \int_{\mathbb{T}^d} H^*\left(x, D_pH(x, D\bar{u}) + \frac{D\zeta}{(2\tau - s)\tilde{m} + (s - \tau)\bar{m}}\right)\left(\frac{2\tau - s}{\tau}\tilde{m}(s, x) + \frac{s - \tau}{\tau}\bar{m}(s, x)\right).$$

If we derive the above expresion we get

$$\frac{1}{\tau}\int_\tau^{2\tau} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} -D_pH^*(\alpha)\frac{G(x - z)D_z\Gamma_0(t, z; y)m(t, x) + (\tau - s)\Gamma_0(t, x; y)D\zeta(t, x)}{m(t, x)}dxdz \\ + \frac{1}{\tau}\int_\tau^{2\tau} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} H^*(\alpha)(2\tau - s)\Gamma_0(t, x; y)dxdz. \tag{6.35}$$

Therefore, $\frac{\delta I_2}{\delta m}(m_1, y)$ is equal to

$$\frac{1}{\tau}\int_\tau^{2\tau}\int_{\mathbb{T}^d}\int_{\mathbb{T}^d} -D_pH^*(\alpha)\frac{G(x-z)D_z\Gamma_0(t,z;y)m(t,x)+(\tau-s)\Gamma_0(t,x;y)D\zeta(t,x)}{m(t,x)}dxdzdt+$$

$$\frac{1}{\tau}\int_\tau^{2\tau}\int_{\mathbb{T}^d}\int_{\mathbb{T}^d} H^*(x,\alpha)(2\tau-s)\Gamma_0(t,x;y)dxdz + \frac{1}{\tau}\int_\tau^{2\tau}\int_{\mathbb{T}^d}(2\tau-s)F(x,m(s,x))\Gamma_0(t,x;y)dxdt$$

and

$$\frac{\delta\Psi}{\delta m}(m_1,y) = \frac{\delta I_1}{\delta m}(m_1,y) + \frac{\delta I_2}{\delta m}(m_1,y)$$

Note that, as in the above expression we are looking at a time interval bounded away from zero, the parabolic regularity ensures that all the functions therein are smooth with respect to the state variable. This implies that also $D_m\Psi$ is well defined.

We omit the proof for second order derivatives. It does not present any further difficulties. Indeed, the parabolic regularity, enjoyed by the solutions of the MFG system at any time $t > 0$, ensures that we can deploy the same kind of computations that we used in (6.35) and so that both $D^2_{mm}\Psi$ and $D^2_{my}\Psi$ are well defined and bounded. ∎

Lemma 6.3. *For any $m_0 \in \mathcal{P}(\mathbb{T}^d)$, there exists a function $\bar\Psi \in C^2(\mathcal{P}(\mathbb{T}^d))$ such that $\bar\Psi(m) > \chi(m)$ for any $m \neq m_0$ and $\bar\Psi(m_0) = \chi(m_0)$.*

Moreover, we can choose $\bar\Psi$ such that $D^2_{mm}\bar\Psi$ and $D^2_{ym}\bar\Psi$ are bounded independently of χ and with $D_m\bar\Psi(m_0, x) = D\bar u(0, x)$ where $\bar u$ is defined in (6.30) and (6.31).

Proof. Let $\{\phi_n\}_n$ be a countable collection of $C^\infty(\mathbb{T}^d)$ maps such that $\{\phi_n\}_n$ is dense in the set of $\text{Lip}_1(\mathbb{T}^d)$, which is the set 1-Lipschitz function on $\mathbb{T}^d$. Then,

$$\mathbf{d}(m, m_0) = \sup_{f\in\text{Lip}_1(\mathbb{T}^d)}\int_{\mathbb{T}^d} f(x)(m-m_0)(dx) = \sup_{n\in\mathbb{N}}\int_{\mathbb{T}^d}\phi_n(x)(m-m_0)(dx).$$

We define $Q : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ as follows

$$Q(m) = \sum_{n\in\mathbb{N}}\frac{\left(\int_{\mathbb{T}^d}\phi_n(x)(m-m_0)(dx)\right)^2}{(n+1)^2(\|\phi_n\|_\infty + \|D^2\phi_n\|_\infty + 1)}.$$

The denominator in the above fraction ensures that Q is well defined for any $m \in \mathcal{P}(\mathbb{T}^d)$. Note that $Q(m) = 0$ if and only if, for any $n \in \mathbb{N}$, $\int_{\mathbb{T}^d}\phi_n(x)(m-m_0)(dx) = 0$. In this case, $\mathbf{d}(m, m_0) = 0$ and so $m = m_0$. One easily checks that Q is smooth and that its derivatives

$$D_mQ(m,y) = 2\sum_{n\in\mathbb{N}}\frac{\left(\int_{\mathbb{T}^d}\phi_n(x)(m-m_0)(dx)\right)D\phi_n(y)}{(n+1)^2(\|\phi_n\|_\infty + \|D^2\phi_n\|_\infty + 1)},$$

$$D^2_{my}Q(m,y) = 2\sum_{n\in\mathbb{N}}\frac{\left(\int_{\mathbb{T}^d}\phi_n(x)(m-m_0)(dx)\right)D^2\phi_n(y)}{(n+1)^2(\|\phi_n\|_\infty + \|D^2\phi_n\|_\infty + 1)},$$

and

$$D^2_{mm}Q(m,y,z) = \sum_{n\in\mathbb{N}}\frac{D\phi_n(y)\otimes D\phi_n(z)}{(n+1)^2(\|\phi_n\|_\infty + \|D^2\phi_n\|_\infty + 1)}$$

are bounded. Note also that $D_mQ(m_0, y) = 0$ for any $y \in \mathcal{P}(\mathbb{T}^d)$. We can now define

$$\bar\Psi(m) = \Psi(m) + Q(m),$$

where Ψ is the function defined in (6.33). By construction, $\bar{\bar{\Psi}}$ is such that $\bar{\bar{\Psi}}(m) > \chi(m)$ for any $m \neq m_0$ and $\bar{\bar{\Psi}}(m_0) = \chi(m_0)$. Moreover,

$$D_m\bar{\bar{\Psi}}(m_0, y) = D_m\Psi(m_0, y) + D_mQ(m_0, y) = D_m\Psi(m_0, y) = D\bar{u}(0, y).$$

The boundedness of the derivatives comes from Proposition 6.2 and the properties of Q that we discussed above. ∎

References

[1] Thierry Aubin, *Some nonlinear problems in riemannian geometry*, Springer Science & Business Media, 2013.

[2] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov, *Fokker-planck-kolmogorov equations*, vol. 207, American Mathematical Soc., 2015.

[3] Ariela Briani and Pierre Cardaliaguet, *Stable solutions in potential mean field game systems*, Nonlinear Differential Equations and Applications NoDEA **25** (2018), no. 1, 1.

[4] P. Cardaliaguet, *Long time average of first order mean field games and weak kam theory*, Dynamic Games and Applications **3** (2013), no. 4, 473–488.

[5] P. Cardaliaguet, François D., J.-M. Lasry, and P.-L. Lions, *The master equation and the convergence problem in mean field games*, arXiv preprint arXiv:1509.02505 (2015).

[6] P. Cardaliaguet, P. J. Graber, A. Porretta, and D. Tonon, *Second order mean field games with degenerate diffusion and local coupling*, Nonlinear Differential Equations and Applications NoDEA **22** (2015), no. 5, 1287–1317.

[7] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, and A. Porretta, *Long time average of mean field games with a nonlocal coupling*, SIAM Journal on Control and Optimization **51** (2013), no. 5, 3558–3591.

[8] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, A. Porretta, et al., *Long time average of mean field games.*, NHM **7** (2012), no. 2, 279–301.

[9] René Carmona and François Delarue, *Probabilistic theory of mean field games with applications i-ii*, Springer, 2018.

[10] M. Cirant and L. Nurbekyan, *The variational structure and time-periodic solutions for mean-field games systems*, arXiv preprint arXiv:1804.08943 (2018).

[11] Marco Cirant, *On the existence of oscillating solutions in non-monotone mean-field games*, Journal of Differential Equations (2019).

[12] A. Fathi, *Solutions kam faibles conjuguées et barrieres de peierls*, Comptes Rendus de l'Académie des Sciences-Series I-Mathematics **325** (1997), no. 6, 649–652.

[13] _____, *Théoreme kam faible et théorie de mather sur les systemes lagrangiens*, Comptes Rendus de l'Académie des Sciences-Series I-Mathematics **324** (1997), no. 9, 1043–1046.

[14] _____, *Sur la convergence du semi-groupe de lax-oleinik*, Comptes Rendus de l'Académie des Sciences-Series I-Mathematics **327** (1998), no. 3, 267–270.

[15] _____ , *Weak kam theorem in lagrangian dynamics preliminary version number 10*, by CUP (2008).

[16] Albert Fathi, *Weak kam theory: the connection between aubry-mather theory and viscosity solutions of the hamilton-jacobi equation*, Proceedings of the International Congress of Mathematicians, vol. 3, 2014, pp. 597–621.

[17] Rita Ferreira and Diogo A Gomes, *On the convergence of finite state mean-field games through γ-convergence*, Journal of Mathematical Analysis and Applications **418** (2014), no. 1, 211–230.

[18] Wilfrid Gangbo and Adrian Tudorascu, *Lagrangian dynamics on an infinite-dimensional torus; a weak kam theorem*, Advances in Mathematics **224** (2010), no. 1, 260–292.

[19] _____ , *Weak kam theory on the wasserstein torus with multidimensional underlying space*, Communications on Pure and Applied Mathematics **67** (2014), no. 3, 408–463.

[20] Diogo Gomes and Levon Nurbekyan, *On the minimizers of calculus of variations problems in hilbert spaces*, Calculus of Variations and Partial Differential Equations **52** (2015), no. 1-2, 65–93.

[21] _____ , *An infinite-dimensional weak kam theory via random variables*, Discrete & Continuous Dynamical Systems-A **36** (2016), no. 11, 6167–6185.

[22] Diogo Gomes and Marc Sedjro, *One-dimensional, forward-forward mean-field games with congestion*, Discrete and Continuous Dynamical Systems - Series S **11** (2017), 901.

[23] Diogo A Gomes, Joana Mohr, and Rafael Rigao Souza, *Discrete time, finite state space mean field games*, Journal de mathématiques pures et appliquées **93** (2010), no. 3, 308–328.

[24] Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions, *Mean field games and applications*, Paris-Princeton lectures on mathematical finance 2010, Springer, 2011, pp. 205–266.

[25] M. Huang, R. P. Malhamé, P. E. Caines, et al., *Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle*, Communications in Information & Systems **6** (2006), no. 3, 221–252.

[26] O. A. Ladyzhenskaia, V. A. Solonnikov, and N. N Ural'tseva, *Linear and quasi-linear equations of parabolic type*, vol. 23, American Mathematical Soc., 1988.

[27] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. i–le cas stationnaire*, Comptes Rendus Mathématique **343** (2006), no. 9, 619–625.

[28] _____ , *Jeux à champ moyen. ii–horizon fini et contrôle optimal*, Comptes Rendus Mathématique **343** (2006), no. 10, 679–684.

[29] P.-L. Lions, *Mean-field games. cours au collège de france (2007–2008)*.

[30] Marco Masoero, *On the long time convergence of potential mfg*, Nonlinear Differential Equations and Applications NoDEA **26** (2019), no. 2, 15.

[31] A. R. Mészáros and F. J. Silva, *On the variational formulation of some stationary second-order mean field games systems*, SIAM Journal on Mathematical Analysis **50** (2018), no. 1, 1255–1277.