Short time solution to the master equation of a first order mean field
game system

Sergio Mayorga*
School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA

December 4, 2018

Abstract

The goal of this paper is to show existence of short-time classical solutions to the so called Master Equation of first order Mean Field Games, which can be thought of as the limit of the corresponding master equation of a stochastic mean field game as the individual noises approach zero. Despite being the equation of an idealistic model, its study is justified as a way of understanding mean field games in which the individual players’ randomness is negligible; in this sense it can be compared to the study of ideal fluids [23]. We restrict ourselves to potential mean field games but do not impose any monotonicity conditions on the running and initial costs, and we do not require convexity of the Hamiltonian, thus extending the result of [25] to a considerably broader class of Hamiltonians.

MSC: 34A12, 35R06, 35R15, 45K05, 49L99, 49N70, 91A13, 91A23.

Keywords: mean field games; master equation; Hamilton-Jacobi equations; fixed-point method; characteristic equations; Wasserstein gradient.

Contents

1 Introduction 2
2 Preliminaries 5
  2.1 Assumptions for the mean-field game equations 6
    2.1.1 Definitions of classical (strong) solutions 7
  2.2 Differentiability in the Wasserstein space 8
3 Summary of main results 10
4 Solving the MFG system 10
  4.1 System of equations and its solution 10
  4.2 First regularity properties 16
  4.3 Value function $U$ and characteristics 22
  4.4 The full value function $u(s,q,\mu)$ 26

*smayorga3@gatech.edu
The master equation (ME, for short) of first order mean field games, namely,

$$\begin{aligned}
\partial_t u(s, q, \mu) + \int_{\mathbb{T}^d} \nabla_u u(s, q, \mu)(x) \cdot \nabla_p H(q, \nabla_q u(s, x, \mu)) \mu(dx) \\
+ H(q, \nabla_q u(s, q, \mu)) + F(q, \mu) = 0 \\
\text{in } (0, T) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d), \\
u(0, q, \mu) = g(q, \mu), \\
\text{on } \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d)
\end{aligned}$$

(1.1)
is a non-local, infinite-dimensional partial differential equation that arises in mean field game (abbreviated MFG) theory and can be interpreted as either the limit, as $N \to \infty$, of a system of $N$ coupled Hamilton-Jacobi equations that represent the Nash equilibrium of a differential game played by $N$ interacting deterministic particles, or as the limiting case (formally, at least) of the master equation of a stochastic differential game when the viscosity parameter, associated with the intrinsic noise of the infinitesimal particles, tends to zero. In (1.1), $\mathcal{P}(\mathbb{T}^d)$ is the Wasserstein space of Borel probability measures on the $d$-dimensional torus $\mathbb{T}^d$. The objective of this paper is to construct a short-time solution to (1.1) for an arbitrary smooth Hamiltonian $H$. To our knowledge, existence of solutions to (1.1) has only been shown for the particular case [25] of the quadratic Hamiltonian $H(q, p) = \frac{1}{2}|p|^2$.

Mean-field theory in differential games began with the works of P.L. Lions and J.M. Lasry [36] and M. Huang, P.E. Caines, R.P. Malhamé [33], attracting great interest since then for its numerous applications and posing challenging theoretical questions. Equation (1.1) and its higher-order version were first introduced by Lasry and Lions [37], motivated by, among other reasons [18], the need to clarify the connection between games with finitely but many players, and MFGs. Regarding the latter, and considerably more scrutinized so far than (1.1), is the so called first order mean field game system:

$$\begin{aligned}
\partial_t U(t, q) + H(q, \nabla_q U(t, q)) + F(q, \sigma_t) = 0 \\
\text{in } (0, s) \times \mathbb{T}^d, \\
\partial_t \sigma_t + \nabla \cdot (\sigma_t \nabla_p H(q, \nabla_q U)) = 0 \\
\text{in } \mathcal{D}'((0, s) \times \mathbb{T}^d), \\
U(0, \cdot) = g(\cdot, \sigma_0), \\
\sigma_s = \mu,
\end{aligned}$$

(1.2)-(1.5)
which describes a Nash-type equilibrium state of a differential game played by a continuum of players on $\mathbb{T}^d$ who seek to minimize a certain cost function that depends on the collective behavior of all the players; in such a state, $U(t,q)$ represents the value function of a typical player $q$ at time $t$ and $\sigma_t$ is the distribution of all the players at time $t$. The first equation in (1.2-1.5) is a forward Hamilton-Jacobi equation and the second a backward continuity equation. These equations can be derived as the optimality conditions of the aforementioned game (see, e.g. [5, 12, 32]) or as the limit of approximate Nash equilibria in finitely-many player games (e.g. [34]). Alternatively, a solution to (1.1) can be used directly to construct an optimal control for the Nash equilibrium of the mean field game [23]. The now extensive and rapidly growing literature on MFGs includes surveys and books [6, 12, 16, 29] that give comprehensive accounts of the theory, its models and applications, with at least one book [30] devoted entirely to regularity theory of the MFGs.

In the present work, the Hamiltonian $H$ is smooth, and not necessarily convex in $p$ (which is commonly required), while the couplings $F$ and $g$ are smooth and jointly Lipschitz (see Section 2.1 for the assumptions). In the master equation, $F$ and $g$ are, additionally, assumed to have potential form (see (5.1) and (5.2)). For convenience, we refer the reader to Section 3 for a summary of our results. We are aware that, if $H(q, \cdot)$ is not convex, the MFG system completely loses its interpretation as an optimization problem in mean field games, since the Legendre transform of $H$ is not even defined. However, this does not prevent us from looking at (1.1) and (1.2-1.5) as a legitimate problem in PDEs, and this is what we do in this work. Considerable understanding of the relationship between master equations and mean field game systems has been gained since their inception, and, indeed, an established strategy [18], which we also use here, to construct a solution to the master equation is to use solutions to the MFG system as the starting point. However, it is MFG systems that have seen more rigorous and abundant treatment in the literature. Theoretically, this is mostly due to the derivative in measure that appears in (1.1). The main issue with the strategy just mentioned becomes to prove that the solutions to (1.2-1.5) behave nicely enough with respect to the terminal measure $\mu$.

Concerning the system (1.2-1.5), weak solutions in the viscosity sense were initially obtained on $\mathbb{T}^d$ and $\mathbb{R}^d$ for regularizing couplings (i.e., in potential form), Hamiltonians with quadratic growth in the momentum variable and arbitrary time horizons [4, 12, 15, 36]. Uniqueness is usually obtained by imposing monotonicity conditions on the couplings. Several significant modifications and refinements of these results have been obtained since then, e.g., existence and uniqueness of weak solutions in the case of local couplings [13], first-order [14, 35, 40] and higher [31] Sobolev estimates of such solutions, with different growth conditions of the Hamiltonian, whose convexity in the momentum variable is always required, and absolute continuity of the terminal measure $\mu$; all the approaches in these developments work for arbitrary time horizons. Regarding second order MFGs, i.e.,

$$\left\{ \begin{array}{l}
\epsilon \Delta U + \partial_t U(t,q) + H(q, \nabla_q U(t,q)) + F(q, \sigma_t) = 0 \quad \text{in } [0,s] \times \mathbb{T}^d, \\
-\epsilon \Delta U + \partial_t \sigma_t + \nabla \cdot (\sigma_t \nabla_p H(q, \nabla_q U)) = 0 \quad \text{in } \mathcal{D}'([0,s] \times \mathbb{T}^d),
\end{array} \right. \tag{1.6}$$

strong solutions for considerably larger classes of coefficients, in which the interaction of the particles enters the equation directly in $H$ and not necessarily through $F$ as a separate summand, have been obtained by Ambrose [1, 2], with the requirement of certain smallness assumptions on the data.

---

1. Systems such as (1.2-1.5) are often called deterministic because they derive from MFGs where the differential equation that governs the evolution of each player has no stochastic terms, with the resulting MFG system featuring no second order derivatives.

2. Our presentation of the MFG is reversed in time with respect to the most frequent one encountered in literature, i.e., with a terminal condition for $g$ and an initial one for $\sigma$, and with minus signs in front of the time derivatives.
general, the presence of the viscosity term in (1.6) affords solutions to enjoy better regularity properties. For this reason, and because (1.6) accommodates models where players act non-deterministically, second order mean field games have received more attention in the scientific literature [29, 30]. Incidentally, we see that due to the sign of the viscosity term in the first equation of (1.6), the system can be well posed only if an initial condition is prescribed while the opposite sign in the second equation makes it mandatory to prescribe the terminal value of \( \sigma \).

In the course of our work towards the ME, we obtain classical solutions for short times of the system (1.2-1.5), with the conditions on the data \( H, F, g \) mentioned above.

With respect to master equations, most available literature has dealt with the higher-order variants [7, 8, 16, 17]. The recent paper by Cardaliaguet et al. [18] includes rigorous proofs of classical solutions for the master equations of second-order MFGs, such as (1.2-1.5) and their characterization as the convergence of \( N \)-player Nash systems as \( N \to \infty \). As for the first-order master equation (1.1), a major step was achieved by Gangbo-Święch in [25], where short-time strong solutions to both the MFG system (1.2-1.5) and the ME are obtained for quadratic Hamiltonian and regularizing couplings. The same result was later proved by Bessi [9] using different techniques. The present paper keeps the smoothness assumptions of the Hamiltonian and the couplings, but does away with the convexity of \( H \) in the variable \( p \) (and, in particular, no growth condition is imposed). We use similar ideas and techniques to arrive at the ME, but our route to the MFG system is different. Let us explain the differences with [25]. Given \( H, F, g \) as in Section 2.1 and given \( 0 < s < T, \mu \in \mathcal{P}(\mathbb{T}^d) \), we prove that, granted \( T \) is small, there are functions \( \Sigma^1 : [0,T] \times \mathbb{T}^d \to \mathbb{T}^d, \Sigma^2 : [0,T] \times \mathbb{T}^d \to \mathbb{R}^d \) that solve the infinite-dimensional Hamiltonian system

\[
\begin{align*}
\partial_t \Sigma^1 &= \nabla_p H(\Sigma^1, \Sigma^2), \\
\partial_t \Sigma^2 &= -\nabla_q H(\Sigma^1, \Sigma^2) - \nabla_q F(\Sigma^1, \Sigma^1 \# \mu)
\end{align*}
\]

(1.7)

with initial and terminal conditions

\[
\Sigma^2(0,q) = \nabla_q g(\Sigma^1(0,q), \Sigma^1(0,\cdot) \# \mu), \quad \Sigma^1(s,q) = q,
\]

providing us with a path in \( \mathcal{P}(\mathbb{T}^d) \) given by

\[
t \mapsto \sigma_t := \Sigma^1_t \# \mu
\]

and prove that there is a function \( U : [0,T] \times \mathbb{T}^d \to \mathbb{R} \) such that

\[
\nabla_q U(t, \Sigma^1_t) = \Sigma^2_t.
\]

(1.8)

On the other hand, we have that the velocity vector \( v_t \) driving the path \( \sigma_t \) satisfies

\[
v(t, \Sigma^1_t) = \partial_t \Sigma^1_t,
\]

and the first equation in (1.7) implies

\[
\nabla_p H(q, \nabla_q U(t, \Sigma^1_t)) = \partial_t \Sigma^1_t.
\]

(1.9)

Comparing (1.9) and (1.8), we see that they are the same if \( H(q,p) = \frac{1}{2}|p|^2 \), which is the Hamiltonian in [25], and, indeed, in that case, \( \partial_t \Sigma^1 = \Sigma^2 \), with \( \nabla_q U(t, \cdot) \) coinciding with the velocity \( v_t \) \((a \ posteriori\) from (1.8)). Thus, the function \( \Sigma^2 \) is not present in [25], with \( \partial_t \Sigma^1 \) taking its place, while the relationship

\[\text{We follow the convention, common in this field, of using the subindex } t \text{ to mean “at time } t\text{”, and thus a shorthand for } (t, \ldots) \text{ rather than the time derivative.}\]
\[ \nabla \varphi(t, \Sigma^1_t) = \partial_t \Sigma^1_t \] is obtained via the link of the MFG system with a variational problem: if \( L(x, v) := \frac{1}{2} |v|^2 \), and having shown that the pair \((\sigma, v)\) is the unique minimized\(^4\) of

\[
U(s, \mu) = \inf_{\{\sigma, v\}} \left\{ \int_0^s \int_{\mathbb{T}^d} \left( L(q, v_t(q)) - F(\sigma_t) \right) dt + G(\sigma_0) \mid \sigma_s = \mu, \sigma \in AC^2(0, s; \mathcal{P}(\mathbb{T}^d)) \right\}, \tag{1.10}
\]

where \( F, G : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R} \) are functions whose Wasserstein gradients are \( F \) and \( g \), the minimality of the norm of \( v_t \) then follows, leading to the symmetry of \( \nabla \varphi v_t(q) \), which is used to establish \( \nabla \varphi U(t, \Sigma^1_t) = \partial_t \Sigma^1_t \) and, in turn, the Hamilton-Jacobi equation in (1.2-1.5). In the case of the general Hamiltonian, even though it is still true \([38]\) that the pair \((\sigma, v)\) is the unique minimizer of (1.10), it is no longer clear how this approach can give us (1.8). We turn, instead, to a more direct procedure (Lemma 4.22) that also helps to shed further light on how the equations (4.3) are the characteristics of (1.2-1.5). This optic allows us to present a sort of uniqueness counterpart (Theorem 4.25) to the existence result, namely, that if a solution \((\hat{U}, \hat{\sigma})\) is in \( W^{2,3;\infty}(0, T) \times \mathbb{T}^d \times AC^2(0, T; \mathcal{P}(\mathbb{T}^d)) \), then it must coincide, at least for a shorter time \( T \), with the pair \((U, \sigma)\) constructed from \((\Sigma^1, \Sigma^2)\). With this approach we manage to circumvent the specific potential forms for \( F \) and \( g \), assumptions that we add in Section 5 where we work out the differentiability of \( \Sigma \) in \( \mu \) through the same discretization approach used in [25], a short description of which we have included in Section 3. This is followed by the chain rules and Lipschitz estimates of the composite functions that enter the representation formula for \( u \) in (6.1). We should say that the formulas for the Wasserstein gradients have to be defined and their Lipschitz estimates proved; there is no general rigorous rule on composite functions that we can invoke. Finally, and due to the preceding remarks about \( \partial_t \Sigma^1_t \) and \( \Sigma_t^2 \), an extra tool (Lemma 2.4) is needed to complete the chain rule for \( u \) that is really the essence of the master equation (Theorem 6.1).

## 2 Preliminaries

For full details on the theory of optimal transport and the Wasserstein space of probability measures on the \( d \)-dimensional torus \( \mathbb{T}^d := \mathbb{R}^d / \mathbb{Z}^d \), we refer the reader to [26]. In this section we set down the notation for the paper, present a few general results that will be needed and fix our data for the MFG equations. We also fix our meaning of classical solution to the MFG system.

- The set of equivalence classes on \( \mathbb{R}^d \) with respect to the equivalence relation:

\[
x \sim y \quad \text{iff} \quad \text{there exist integers } n_1, \ldots, n_d \text{ such that } x^{(j)} - y^{(j)} = n_j, \ j = 1, \ldots, d
\]

is denoted by \( \mathbb{T}^d \), where \( x^{(j)}, y^{(j)} \) are the \( j \)-th coordinates of \( x, y \). If \( x, y \in \mathbb{T}^d \) then,

\[
|x - y|_{\mathbb{T}^d} := \min\{|x' - y'| \mid x, y \in \mathbb{R}^d, x' \sim x, y' \sim y \}.
\]

- If \( \mu, \nu \) are Borel probability measures on \( \mathbb{R}^d \), \( \Gamma(\mu, \nu) \) denotes the set of those Borel probability measures \( \gamma \) on \( \mathbb{R}^d \times \mathbb{R}^d \) whose marginals are \( \mu \) and \( \nu \): that is, \( \pi_\# \gamma = \mu \) and \( \pi'\# \gamma = \nu \), where \( \pi_1, \pi_2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \) are the first and second coordinate projections, respectively, and the subindex \( \# \) stands for the pushforward operator.

\(^4\)See also [28] for a recent connection between value functionals such as (1.10) and Hopf-Lax formulae on the Wasserstein space.
• We use the standard notation $\mathcal{P}_2(\mathbb{R}^d)$ for the Wasserstein space of Borel probability measures on $\mathbb{R}^d$ whose second moments are finite, with quadratic Wasserstein distance $W_2$.

• For $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, we define

$$W(\mu, \nu) = \left( \inf_{\gamma \in \Gamma(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \gamma(dx, dy) \right)^{1/2}, \quad (2.1)$$

and let $\Gamma_0(\mu, \nu)$ denote the set of optimal transport plans $\gamma$ between $\mu$ and $\nu$, i.e. those for which the infimum in $(2.1)$ is attained. With the equivalence relation

$$\mu \sim \nu \iff \int \phi d\mu = \int \phi d\nu \text{ for all } \phi \in C(\mathbb{T}^d)$$
on $\mathcal{P}_2(\mathbb{R}^d)$, where $C(\mathbb{T}^d)$ are all real-valued continuous functions $\phi$ on $\mathbb{R}^d$ such that $\phi(x) = \phi(x')$ whenever $x \sim x'$, it is true that $W(\mu, \nu) = W(\mu', \nu')$ whenever $\mu \sim \mu'$ and $\nu \sim \nu'$. In this way, $W$ in formula $(2.1)$ is defined on the set of equivalence classes, which we henceforth denote by $\mathcal{P}(\mathbb{T}^d)$. Moreover, $W$ is a metric on $\mathcal{P}(\mathbb{T}^d)$, with respect to which $\mathcal{P}(\mathbb{T}^d)$ is compact.

• By a mapping $F : \mathbb{T}^d \to S$, where $S$ is any set, we mean $F : \mathbb{R}^d \to S$ such that $F(x) = F(x')$ whenever $x \sim x'$. Likewise, a mapping $F : \mathcal{P}(\mathbb{T}^d) \to S$ is a function $F : \mathcal{P}(\mathbb{R}^d) \to S$ that takes constant values on the equivalence classes of $\mathcal{P}(\mathbb{T}^d)$. Furthermore, a function $F : \mathbb{T}^d \to \mathbb{T}^d$ is to be understood as a function $F : \mathbb{R}^d \to \mathbb{R}^d$ such that $F(x) \sim F(y)$ whenever $x \sim y$.

• If $x = (x_1, \ldots, x_n) \in (\mathbb{R}^d)^n$, then $\mu^x \in \mathcal{P}_2(\mathbb{R}^d)$ denotes the measure $\mu^x = \frac{1}{n} \sum_{j=1}^n \delta_{x_j}$. Such measures are called average of Dirac masses.

• If $f, g : \mathbb{R}^d \to \mathbb{R}^d$ are Borelian, and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, then, estimating through $(f \times g)_{\#}\mu$ one obtains

$$W_2(f_{\#}\mu, g_{\#}\mu) \leq \|f - g\|_{L^2(\mathbb{R}^d, \mu)}.$$  

• Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Then $L^2(\mathbb{T}^d, \mu)$ denotes $L^2(\mathbb{R}^d, \mu)$ completion of $C(\mathbb{T}^d)$. At the same time, we define the tangent space to $\mathcal{P}(\mathbb{T}^d)$ at $\mu$, $\mathcal{T}_\mu \mathcal{P}(\mathbb{T}^d)$, to be the $L^2(\mathbb{R}^d, \mu)$ completion of the subspace of $L^2(\mathbb{T}^d, \mu)$ consisting of gradients of smooth periodic functions on $\mathbb{R}^d$. $\mathcal{T}_\mu \mathcal{P}(\mathbb{T}^d) := \overline{\nabla C^\infty(\mathbb{T}^d; \mathbb{R}) L^2(\mathbb{T}^d, \mu)}$.

• Wasserstein distance between average of Dirac masses. If $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ are such that $\mu = \frac{1}{n} \sum_{j=1}^n \delta_{x_j}$ and $\nu = \frac{1}{n} \sum_{j=1}^n \delta_{y_j}$, where $x_j \neq x_k, y_j \neq y_k$ for $j \neq k$, then there is a permutation $p : \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that

$$W^2(\mu, \nu) = \frac{1}{n} \sum_{j=1}^n |y_{p(j)} - x_j|^2.$$  

• We denote by $AC^2(0, T; \mathcal{P}(\mathbb{T}^d))$ the set of paths $\sigma : (0, T) \to \mathcal{P}(\mathbb{T}^d)$ for which there exists $m \in L^2(0, T)$ such that $W(\sigma_{t_1}, \sigma_{t_2}) \leq \int_{t_1}^{t_2} m(\tau) d\tau$ whenever $0 < t_1 \leq t_2 < T$.

• Density of average of Dirac masses in $\mathcal{P}(\mathbb{T}^d)$. Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. As it is well known (see, for instance, [10 Ex. 8.1.6]), the set of average of Dirac masses is dense in $\mathcal{P}_2(\mathbb{R}^d)$ with respect to narrow convergence. In $\mathcal{P}(\mathbb{T}^d)$, this convergence coincides with convergence in $W$. Thus, there exists a sequence $\{\mu(n)\}_{n \in \mathbb{N}} \subset \mathcal{P}_2(\mathbb{R}^d)$, with $\mu(n) = \frac{1}{n} \sum_{j=1}^n \delta_{x_j(n)}$, an average of Dirac masses, such that $W(\mu(n), \mu(n)) \to 0$ as $n \to \infty$. Moreover, this sequence can be chosen so that each $x_j(n) \in \text{supp}(\mu(n))$, where $\text{supp}(\mu)$ is the support of the measure $\mu$. 

6
2.1 Assumptions for the mean-field game equations

1. Let \( H \in C^3(\mathbb{T}^d \times \mathbb{R}^d) \), \( H = H(q, p) \). In this manuscript, \( \nabla_q H(\cdot, \cdot) \) will always denote the gradient of \( H \) with respect to \( q \), evaluated at \( \langle \cdot, \cdot \rangle \). Similarly for \( \nabla_p H(\cdot, \cdot) \), and higher-order derivatives.

2. Let \( F = F(q, \mu), q \in \mathbb{T}^d, \mu \in \mathcal{P}(\mathbb{T}^d) \), be continuous in the \( \mu \) variable and of class \( C^3 \) in \( q \), and let \( \kappa > 0 \) be a constant such that

\[
|\nabla_q F(q, \mu)|, |\nabla_{qq}^2 F(q, \mu)|, |\nabla_{qqq}^3 F(q, \mu)| \leq \kappa, \quad q \in \mathbb{T}^d, \mu \in \mathcal{P}(\mathbb{T}^d).
\]

Suppose, further, that \( \nabla_q F \) is \( \kappa \)-Lipschitz on \( \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \), meaning that

\[
|\nabla_q F(q_1, \mu_1) - \nabla_q F(q_2, \mu_2)| \leq \kappa \sqrt{|q_1 - q_2|^2 + \mathcal{W}^2(\mu_1, \mu_2)}, \quad q_1, q_2 \in \mathbb{T}^d, \mu_1, \mu_2 \in \mathcal{P}(\mathbb{T}^d).
\]

3. Let \( g = g(q, \mu), q \in \mathbb{T}^d, \mu \in \mathcal{P}(\mathbb{T}^d) \), and suppose \( g \) satisfies exactly the same conditions asked of \( F \).

We call the triple \((H, F, g)\) the data for the mean-field game equations.

2.1.1 Definitions of classical (strong) solutions

Let \( T > 0 \), and \( F, g : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R} \) be continuous; let \( H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R} \) be continuous and differentiable in \( p \).

**MFG system** Let \( 0 < s < T, \mu \in \mathcal{P}(\mathbb{T}^d) \). We say that the pair of functions \( U : (0, T) \times \mathbb{T}^d \to \mathbb{R}, \sigma : (0, T) \to \mathcal{P}(\mathbb{T}^d) \) is a classical solution to the first-order MFG system \((1.2, 1.3)\) on \( \mathbb{T}^d \) with data \((H, F, g)\) and parameters \( s, \mu \) if the following hold:

- \( U \in C^1((0, T) \times \mathbb{T}^d) \);
- the path \( \sigma \in AC^2(0, T; \mathcal{P}(\mathbb{T}^d)) \) and \((1.3)\) is true in the sense of distributions, i.e., for every \( \varphi \in C^\infty_c((0, T) \times \mathbb{T}^d) \):

\[
\int_0^T \int_{\mathbb{T}^d} \left[ \partial_t \varphi(t, q) + \nabla \varphi(t, q) \cdot \nabla_p H(q, \nabla_q U(t, q)) \right] \sigma_t(dq)dt = 0; \tag{2.3}
\]

- equation \((1.2)\) is satisfied pointwise, along with the condition \((1.4)\) at time \( t = 0 \) for \( U \) and the condition \((1.5)\) at time \( t = s \) for \( \sigma \).

We will often refer to the function \( U \) in \((1.2)\) as the value function.

**Master equation** We say that the function \( u : (0, T) \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \to \mathbb{R} \) is a classical solution the master equation of first-order MFGs \((1.1)\) with data \((H, F, g)\) if:

- \( u \) is differentiable in \( s \), with \( \partial_s u(\cdot, \cdot, \cdot, \mu) \) continuous at every \( \mu \in \mathcal{P}(\mathbb{T}^d) \);
- \( u \) is differentiable in \( q \), with \( \nabla_q u \) continuous in all three variables;
- \( u \) is differentiable in \( \mu \) (see the following section), and \( u \) satisfies \((1.1)\) pointwise.

We will refer to the function \( u \) in \((1.1)\) as the full value function.
2.2 Differentiability in the Wasserstein space

Let $W$ be a real-valued function on $\mathcal{P}(\mathbb{R}^d)$ and let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ be fixed. For $\xi \in L^2(\mathbb{T}^d,\mu)$, $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, $\gamma \in \Gamma(\mu,\nu)$, define

$$e(\nu,\xi,\gamma) := W(\nu) - W(\mu) - \int_{\mathbb{R}^d \times \mathbb{R}^d} \xi(x) \cdot (y - x) \gamma(dx,dy).$$

We have chosen to present this section with a notation similar to the one found in the paper [27], which unifies the different notions of differentiability on $\mathcal{P}_2(\mathbb{R}^d)$ used in the literature. If $r > 0$, set

$$e_r[\xi] = \sup_{\gamma \in \Gamma(\mu,\nu)} \sup_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{|e(\nu,\xi,\gamma)|}{\|\pi^1 - \pi^2\gamma\|} : \|\pi^1 - \pi^2\gamma\| \leq r \right\}$$

and

$$e_0^r[\xi] = \sup_{\gamma \in \Gamma(\mu,\nu)} \sup_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{|e(\nu,\xi,\gamma)|}{\|\pi^1 - \pi^2\gamma\|} : \|\pi^1 - \pi^2\gamma\| \leq r \right\}.$$

**Definition 2.1.** With the preceding notation, we say that $W$ is differentiable at $\mu$ if

$$\lim_{r \to 0^+} e^0_r[\xi] = 0. \quad (2.4)$$

The set of all $\xi \in L^2(\mathbb{T}^d,\mu)$ for which (2.4) holds is denoted $\partial W(\mu)$.

**Lemma 2.2.** If $\xi \in \partial W(\mu)$, then so is its projection $\bar{\xi}$ onto $\mathcal{P}_\mu(\mathbb{T}^d)$, which is then the unique element of minimal norm in $\partial W(\mu)$ and is denoted by

$$\nabla_\mu W(\mu).$$

We will call it the Wasserstein gradient of $W$ at $\mu$.

Its proof can be found in [26].

**Remark 2.3.** The following is an equivalent characterization of a vector field $\xi \in L^2(\mathbb{T}^d,\mu)$ that satisfies (2.4):

$$W(\nu) - W(\mu) - \sup_{\gamma \in \Gamma(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \xi(x) \cdot (y - x) \gamma(dx,dy) = o(\|\mu\|). \quad (2.5)$$

The following lemma will be used in the final section.

**Lemma 2.4.** With the foregoing notation, if $W : \mathcal{P}(\mathbb{T}^d) \to \mathbb{R}$ is differentiable at $\mu$, then

$$\lim_{r \to 0^+} e[\nabla_\mu W(\mu), r] = 0.$$

By Remark 2.3, this is the same as

$$W(\nu) - W(\mu) - \sup_{\gamma \in \Gamma(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \nabla_\mu W(\mu)(x) \cdot (y - x) \gamma(dx,dy) = o(\|\mu\|).$$

8
Proof. We begin by noting the following. Let $\mu, \nu \in \mathcal{P}(\mathbb{T}^d)$, and $\gamma, \tilde{\gamma} \in \Gamma(\mu, \nu)$. Then, for any $\varphi \in C^2(\mathbb{T}^d)$,

$$
\left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \nabla \varphi(x) \cdot (y - x)(\gamma - \tilde{\gamma})(dx, dy) \right| \leq \frac{\| \pi^1 - \pi^2 \|_\gamma^2 + \| \pi^1 - \pi^2 \|_\gamma^2}{2} \| \nabla^2 \varphi \|_\infty.
$$

(2.6)

Indeed, Taylor expansion gives a Borel function $r : \mathbb{T}^d \times \mathbb{T}^d \to [-1, 1]$ such that

$$
\varphi(y) - \varphi(x) - \nabla \varphi(x) \cdot (y - x) = r(x, y)\| \nabla^2 \varphi \|_\infty \frac{|x - y|^2}{2}.
$$

Integrating both sides of this equality over $\mathbb{R}^d \times \mathbb{R}^d$ once with respect to $\gamma$ and then with respect to $\gamma'$, remembering that $\gamma$ and $\gamma'$ have the same marginals $\mu$ and $\nu$, and substracting one of the resulting expressions from the other, yields (2.5). Fix now $\mu \in \mathcal{P}(\mathbb{T}^d)$. Let $\nu \in \mathcal{P}(\mathbb{T}^d)$ and $\gamma \in \Gamma_0(\mu, \nu), \tilde{\gamma} \in \Gamma(\mu, \nu)$. Let $\varphi \in C^\infty(\mathbb{T}^d)$. Write

$$
e(\nu, \nabla_\mu \mathcal{W}(\mu), \gamma) = e(\nu, \nabla \varphi, \gamma) - \int_{\mathbb{R}^d \times \mathbb{R}^d} (\nabla_\mu \mathcal{W}(\mu)(x) - \nabla \varphi(x)) \cdot (y - x)\gamma(dx, dy),
$$

and the same expression which holds with $\tilde{\gamma}$ in place of $\gamma$. Substracting one from the other and taking absolute value gives, using Hölder’s inequality,

$$
|e(\nu, \nabla_\mu \mathcal{W}(\mu), \gamma) - e(\nu, \nabla_\mu \mathcal{W}(\mu), \tilde{\gamma})| \leq |e(\nu, \nabla \varphi, \gamma) - e(\nu, \nabla \varphi, \tilde{\gamma})|
$$

$$
+ \| \nabla_\mu \mathcal{W}(\mu) - \nabla \varphi \|_{L^2(\mu)}(\| \pi^2 - \pi^1 \|_\gamma + \| \pi^2 - \pi^1 \|_\gamma).
$$

Now, $\| \pi^2 - \pi^1 \|_\gamma \leq \| \pi^2 - \pi^1 \|_\gamma$, and, using (2.5),

$$
|e(\nu, \nabla_\mu \mathcal{W}(\mu), \gamma) - e(\nu, \nabla_\mu \mathcal{W}(\mu), \tilde{\gamma})| \leq \| \pi^2 - \pi^1 \|_\gamma(\| \pi^1 - \pi^2 \|_\gamma \| \nabla^2 \varphi \|_\infty + 2\| \nabla_\mu \mathcal{W}(\mu) - \nabla \varphi \|_{L^2(\mu)}).
$$

Dividing by $\| \pi^2 - \pi^1 \|_\gamma$ and once again because $\| \pi^2 - \pi^1 \|_\gamma \leq \| \pi^2 - \pi^1 \|_\gamma$, we obtain

$$
\frac{|e(\nu, \nabla_\mu \mathcal{W}(\mu), \gamma) - e(\nu, \nabla_\mu \mathcal{W}(\mu), \tilde{\gamma})|}{\| \pi^2 - \pi^1 \|_\gamma} \leq \| \pi^1 - \pi^2 \|_\gamma \| \nabla^2 \varphi \|_\infty + 2\| \nabla_\mu \mathcal{W}(\mu) - \nabla \varphi \|_{L^2(\mu)}
$$

This holds for any $\nu \in \mathcal{P}(\mathbb{T}^d)$, $\gamma \in \Gamma_0(\mu, \nu), \tilde{\gamma} \in \Gamma(\mu, \nu), \varphi \in C^\infty(\mathbb{T}^d)$. Fix $r > 0$, and, on the right-hand side, fix $\tilde{\gamma} \in \Gamma(\mu, \nu)$ such that $\| \pi^2 - \pi^1 \|_\gamma < r$. Take then the supremum on the left-hand side over $\nu \in \mathcal{P}(\mathbb{T}^d)$, $\gamma \in \Gamma(\mu, \nu)$ such that $\| \pi^2 - \pi^1 \|_\gamma < r$, to obtain

$$
e[\nabla_\mu \mathcal{W}(\mu), r] \leq \frac{|e(\nu, \nabla_\mu \mathcal{W}(\mu), \gamma) - e(\nu, \nabla_\mu \mathcal{W}(\mu), \tilde{\gamma})|}{\| \pi^2 - \pi^1 \|_\gamma} + (r\| \nabla^2 \varphi \|_\infty + 2\| \nabla_\mu \mathcal{W}(\mu) - \nabla \varphi \|_{L^2(\mu)})
$$

holding for any $\nu \in \mathcal{P}(\mathbb{T}^d)$, $\gamma \in \Gamma_0(\mu, \nu), \varphi \in C^\infty(\mathbb{T}^d)$. Taking now the supremum on the right-hand side over $\nu \in \mathcal{P}(\mathbb{T}^d)$, $\gamma \in \Gamma_0(\mu, \nu)$ such that $\| \pi^2 - \pi^1 \|_\gamma < r$, and then letting $r \to 0^+$ on both sides yields

$$
\lim_{r \to 0^+} e[\nabla_\mu \mathcal{W}(\mu), r] \leq \lim_{r \to 0^+} e_0[\nabla_\mu \mathcal{W}(\mu), r] + 2\| \nabla_\mu \mathcal{W}(\mu) - \nabla \varphi \|_{L^2(\mu)} = 2\| \nabla_\mu \mathcal{W}(\mu) - \nabla \varphi \|_{L^2(\mu)},
$$

by the hypothesis, for any $\varphi \in C^\infty(\mathbb{T}^d)$. By the fact that $\nabla_\mu \mathcal{W}(\mu)$ is an $L^2(\mu)$ limit of gradients of smooth periodic functions $\varphi$, the conclusion follows. \qed
3 Summary of main results

We collect here the main statements of the paper. Let \( H, F, g \) be as in Section 2.1.

Statement 1. (Theorem 4.24) If \( T \) is sufficiently small, in a way that depends only on the data \((H, F, g)\), then, for every \( 0 < s < T \), \( \mu \in \mathcal{P}(\mathbb{T}^d) \), the MFG system (1.2-1.5) admits a classical solution \((U, \sigma)\), in the sense of Section 2.1.1. Moreover, \((U, \sigma) \in W^{2,2;\infty}((0, T) \times \mathbb{T}^d) \times AC^2(0, T; \mathcal{P}(\mathbb{T}^d))\).

Statement 2. (Theorem 4.25) If \((\tilde{U}, \tilde{\sigma}) \in W^{2,2;\infty}((0, T) \times \mathbb{T}^d) \times AC^2(0, T; \mathcal{P}(\mathbb{T}^d))\) is a classical solution to the MFG system (1.2-1.3), then, at least during a possibly shorter interval \([0, T]\) than the one in the previous statement, the pair \((\tilde{U}, \tilde{\sigma})\) must be the pair constructed for Theorem 4.24.

Statement 3. (Theorem 6.6) Let \( F, g \) have the specific convolution forms (5.1), (5.2). If \( T \) is small enough, in a way that depends only on the data \((H, F, g)\), then the master equation (1.1) admits a classical solution in the sense of Section 2.1.1.

The discretization approach Let \( \Sigma^1[s, \mu], \Sigma^2[s, \mu] \) be as in (1.7), i.e., the pair is a solution to (1.3) below. For an average of Dirac masses \( \mu^2 : = \frac{1}{n} \sum_{j=1}^{n} \delta_{x_j}, x \in (\mathbb{T}^d)^n \), let \( M(t, s, q, x) = \Sigma[s, \mu^2](t, q) \) (see Definition 5.1); note that the number of coordinates in the domain of \( M \) depends on \( n \), the number of particles. Some calculations give us (Section 5.1) that
\[
\| \nabla_{x_j}^2 M \|_\infty = O\left( \frac{1}{n} \right), \quad \| \nabla_{x_j x_j}^2 M \|_\infty = O\left( \frac{1}{n} \right), \quad \text{while} \quad \| \nabla_{x_j x_j}^2 M \|_\infty = O\left( \frac{1}{n^2} \right) \text{ if } i \neq j.
\]
This is precisely what is needed to obtain formula (5.27), namely,
\[
n|\nabla_{x_j} M(t, s, q, y) - \nabla_{x_i} M(t, s, q, x)| \leq \sqrt{dC}(|y_j - x_i|_{\mathbb{T}^d} + \| \mu^2 \|_{\mathcal{P}(\mathbb{T}^d)} + \frac{1}{n}), \tag{5.27}
\]
which means, glibly speaking, that \( n \nabla_{x_j(n)} \Sigma[s, \mu^{x(n)}] \) acquires a Lipschitz estimate in the limit as \( n \to \infty \). This allows us to extend \( n \nabla_{x_j(n)} \Sigma[s, \mu^{x(n)}] \) to the full infinite-dimensional manifold \( \mathcal{P}(\mathbb{T}^d) \) (Corollary 5.9), becoming the Wasserstein gradient of \( \Sigma \) in \( \mu \) (Lemma 5.10).

4 Solving the MFG system

In this section, we construct a solution to the first-order MFG system (1.2-1.5). A fixed-point argument will give us existence and uniqueness of solutions to the characteristics of the system. These solutions are functions \( \Sigma : [0, T] \times \mathbb{T}^d \to \mathbb{T}^d \times \mathbb{R}^d \) that depend on \( s \) and \( \mu \) and are the backbone of our work. They incorporate enough regularity that we can construct classical solutions (in the sense defined above) to the MFG system on \( \mathbb{T}^d \).

4.1 System of equations and its solution.

For \( T > 0 \), we will denote by \( \mathcal{M} \) the space of continuous functions
\[
Z = (Q, P) : [0, T] \times \mathbb{T}^d \to \mathbb{T}^d \times \mathbb{R}^d,
\]
endowed with the uniform norm,
\[
\| Z \|_\infty = \max_{0 \leq t \leq T} |Z(t, q)| = \max\{(|Q(t, q)|^2 + |P(t, q)|^2)^{1/2} | t \in [0, T], q \in \mathbb{T}^d\}.
\]
That is,

\[ \mathcal{M} = C([0, T] \times \mathbb{T}^d; \mathbb{T}^d \times \mathbb{R}^d). \]

Similarly, let \( \mathcal{M}^1 := C([0, T] \times \mathbb{T}^d; \mathbb{T}^d), \mathcal{M}^2 := C([0, T] \times \mathbb{T}^d; \mathbb{R}^d). \)

**Definition 4.1.** Let \( \theta \in \mathbb{R}^+ \), and fix \( \mu \in \mathcal{P}(\mathbb{T}^d), s \in [0, T]. \)

1. **(Fixed point operator)** Define the operator \( \tilde{m}^{s,\mu} : \mathcal{M} \to \mathcal{M}, \tilde{m} = ((\tilde{m}^{s,\mu})^1, (\tilde{m}^{s,\mu})^2) \) as follows:

   If \( \tilde{Z} = (\tilde{Q}, \tilde{P}) \in \mathcal{M} \), then \( \tilde{m}^{s,\mu}(\tilde{Z}) = ((\tilde{m}^{s,\mu})^1(\tilde{Z}), (\tilde{m}^{s,\mu})^2(\tilde{Z})) \), where:

   \[
   (\tilde{m}^{s,\mu})^1(\tilde{Z})(t, q) = q + \int_s^t \nabla_p H(\tilde{Q}_\tau(q), \theta \tilde{P}_\tau(q))d\tau,
   \]

   \[
   (\tilde{m}^{s,\mu})^2(\tilde{Z})(t, q) = \frac{1}{\theta} \nabla_q g(\bar{Q}_0(q), \bar{Q}_0(q) - \mu) - \frac{1}{\theta} \int_s^t \nabla_q H(\tilde{Q}_\tau(q), \theta \tilde{P}_\tau(q)) + \nabla_q F(\tilde{Q}_\tau(q), \tilde{Q}^\tau_{\#\mu})d\tau,
   \]

   \(0 \leq t \leq T, q \in \mathbb{R}^d.\) In equalities \( \{4.1\} \) and \( \{4.2\} \), \( \bar{Q}^\tau(q) := Q(\tau, q), \bar{P}^\tau(q) := P(\tau, q), \tau \in [0, T], q \in \mathbb{R}^d.\)

2. **(Data bounds I)** For \( B > 0 \), let

   \[ \bar{l}(B) := \max_{q \in \mathbb{R}^d, |p| \leq B, \mu \in \mathcal{P}(\mathbb{T}^d)} \{ \sqrt{2} |\nabla H(q, \theta p)| + |\nabla_q F(q, \mu)| \}, \]

   \[ \bar{h}(B) := \max_{q \in \mathbb{R}^d, |p| \leq B, \mu \in \mathcal{P}(\mathbb{T}^d)} \{ \sqrt{2} |\nabla^2 H(q, \theta p)| + \sqrt{2} |\nabla^3 H(q, \theta p)| + |\nabla^2_{qq} F(q, \mu)| + |\nabla^3_{qqq} F(q, \mu)| \}, \]

   \[ c := \max\{d, \kappa\}. \]

Thus, for a fixed \( B \), the numbers \( \bar{l}(B), \bar{h}(B), c \) depend only on the data \( (H, F, g) \).

**Notes.**

1. Since \( \bar{Q}, \bar{P} \) are periodic in \( q \) (i.e., \( q \in \mathbb{T}^d \)), if \( q' \sim q \) then \((\tilde{m}^{s,\mu})^1(\tilde{Z})(t, q) \sim (\tilde{m}^{s,\mu})^1(\tilde{Z})(t, \cdot)\), so \((\tilde{m}^{s,\mu})^1(\tilde{Z})(t, \cdot)\) is indeed a mapping into \( \mathbb{T}^d \), in the sense explained in the Preliminaries.

2. Both the fixed-point operator \( \tilde{m}^{s,\mu} \) and the data bounds depend on the value of \( \theta \).

3. Throughout this text, \( |\nabla H(q, p)|^2 = \sum_{j=1}^{d} \frac{\partial H}{\partial (q^j, p^j)}(q, p)^2 \) and norms of second order derivatives are defined similarly, i.e., we are using quadratic norms. //

Suppose that the operator \( \tilde{m}^{s,\mu} \) has a fixed point \( (\bar{Q}, \bar{P}) \), so on the left-hand side of \( \{4.1\} \) and \( \{4.2\} \) we see \( Q(t, q) \) and \( \bar{P}(t, q) \) respectively. Set \( \bar{Q} := \bar{Q} \) and \( \bar{P} := \theta \bar{P} \). Then \( Z := (Q, P) \) satisfies

\[
\begin{aligned}
\dot{Q}(t, q) &= \nabla_p H(Q(t, q), P(t, q)) \quad \text{in } [0, s] \times \mathbb{T}^d, \\
\dot{P}(t, q) &= -\nabla_q H(Q(t, q), P(t, q)) - \nabla_q F(Q(t, q), Q(t, \cdot) \# \mu) \quad \text{in } [0, s] \times \mathbb{T}^d, \\
Q(s, q) &= q \quad \text{on } \mathbb{T}^d, \\
P(0, q) &= \nabla_q g(Q(0, q), Q(0, \cdot) \# \mu) \quad \text{on } \mathbb{T}^d.
\end{aligned}
\]

We will refer to the system \( \{4.3\} \) as the Hamiltonian ODEs with parameters \( s \) and \( \mu \).
Definition 4.2. If $T > 0$, $A_1, A_2, B, E > 0$, define $M_0(A_1, A_2, B, E, T) \subset M$ to be the subset of those $Z(\cdot, \cdot) = (Q(\cdot, \cdot), P(\cdot, \cdot))$ such that: (i) $Z(\cdot, \cdot)$ belongs to $W^{2,2;\infty}([0,T] \times \mathbb{T}^d; \mathbb{T}^d \times \mathbb{R}^d)$ and (ii)

\[
\begin{align*}
|\dot{c}_t Q|_{\mathcal{M}_1} &\leq A_1, \quad \|\nabla_q Q\|_{C([0,T] \times \mathbb{T}^d \times \mathbb{T}^d)} \leq A_1, \quad \|\nabla^2_{qq} Q\|_{C([0,T] \times \mathbb{T}^d \times \mathbb{T}^d)} \leq A_1; \\
|\dot{c}_t P|_{\mathcal{M}_2} &\leq A_2, \quad \|\nabla_q P\|_{C([0,T] \times \mathbb{T}^d \times \mathbb{R}^d)} \leq A_2, \quad \|\nabla^2_{qq} P\|_{C([0,T] \times \mathbb{T}^d \times \mathbb{R}^d)} \leq A_2 \\
|\dot{P}|_{\mathcal{M}_2} &\leq B; \\
\|\nabla_q Q\|_{C(\mathbb{T}^d \times \mathbb{T}^d)} + \|\nabla^2_{qq} Q\|_{C(\mathbb{T}^d \times \mathbb{T}^d)} &\leq E.
\end{align*}
\]

Here $W^{2,2;\infty}([0,T] \times \mathbb{T}^d; \mathbb{T}^d \times \mathbb{R}^d)$ is the Sobolev space of functions periodic in $q$, taking values in $\mathbb{T}^d \times \mathbb{R}^d$, with essentially bounded second-order weak derivatives in $t$ and second-order weak gradients in $q$. Since functions in $W^{1,1;\infty}$ are Lipschitz, $M_0(A_1, A_2, B, E, T)$ is indeed a subset of $M$. The following is a standard fact, so we will omit its proof.

Proposition 4.3. For any $A_1, A_2, B, E, T > 0$, $M_0(A_1, A_2, B, E, T)$ is closed in $M$.

Lemma 4.4. Let $\theta > 0$ of Definition 4.2 be arbitrary. There exist $A_1, A_2, B, E > 0$, and $T > 0$, such that $m^{s,\mu}$ maps $M_0(A_1, A_2, B, E, T)$ into itself, for any $s \in (0, T)$ and $\mu \in \mathcal{P}(\mathbb{T}^d)$. The numbers $A_1, A_2, B, E, T$ depend only on the data.

Proof. Observe that

\[
|\nabla_q (m^{s,\mu})^2(Z)(t, q)| \leq \frac{1}{\theta} |\nabla_q g(Q_0(q), Q_0, \mu)| + \frac{1}{\theta} \sup_{0 \leq \tau \leq t} \left[ |\nabla_q H(Q_\tau(q), \theta \dot{P}_\tau(q))| + |\nabla_q F(Q_\tau(q), \theta \dot{P}_\tau(q))| \right],
\]

\[
|\dot{c}_t (m^{s,\mu})^1(Z)(t, q)| \leq \frac{1}{\theta} |\nabla_q H(Q_t(q), \theta \dot{P}_t(q))|,
\]

\[
|\dot{c}_t (m^{s,\mu})^2(Z)(t, q)| \leq \frac{1}{\theta} |\nabla_q H(Q_t(q), \theta \dot{P}_t(q))| + \frac{1}{\theta} \left[ |\nabla_q F(Q_t(q), \theta \dot{P}_t(q))| \right];
\]

\[
|\nabla_q (m^{s,\mu})^1(Z)(t, q)| \leq \sqrt{d} + \int_s^t \left[ |\nabla^2_{qq} H(Q_\tau(q), \theta \dot{P}_\tau(q))| \nabla_q Q_\tau(q) + |\nabla^2_{pp} H(Q_\tau(q), \theta \dot{P}_\tau(q))| \nabla_q \dot{P}_\tau(q) \right]|d\tau;
\]

\[
|\nabla_q (m^{s,\mu})^2(Z)(t, q)| \leq \frac{1}{\theta} |\nabla^2_{qq} g(Q_0(q), Q_0, \mu)| \nabla_q Q_0(q) + \frac{1}{\theta} \int_0^t \left[ |\nabla^2_{qq} H(Q_\tau(q), \theta \dot{P}_\tau(q))| \nabla_q Q_\tau(q) + |\nabla^2_{pp} H(Q_\tau(q), \theta \dot{P}_\tau(q))| \nabla_q \dot{P}_\tau(q) \right]|d\tau
\]

The previous lines are inequalities for the moduli of $Q$, $P$, and their derivatives. Let us also compute second-order derivatives to find:

\[
|\nabla^2_{qq} (m^{s,\mu})^1(Z)(t, q)| \leq \int_s^t \left[ \left| (\nabla^3_{qqpp} H(Q_\tau(q), \theta \dot{P}_\tau(q))| \nabla_q Q_\tau(q) + |\nabla^3_{pppp} H(Q_\tau(q), \theta \dot{P}_\tau(q))| \nabla_q \dot{P}_\tau(q) \right)||d\tau
\]

\[
+ \frac{1}{\theta} \int_0^t \left[ |\nabla^2_{qq} H(Q_\tau(q), \theta \dot{P}_\tau(q))| \nabla_q Q_\tau(q) + |\nabla^2_{pp} H(Q_\tau(q), \theta \dot{P}_\tau(q))| \nabla_q \dot{P}_\tau(q) \right]|d\tau
\]

We make a convention here and in the rest of the paper that in the application of the classical chain rule, and only if are concerned solely about estimates, juxtaposition is enough, i.e., we will not pay attention to the order of the factors or whether they are properly transposed.
for the first component of \( \bar{m} \), and, for the second component,

\[
|\nabla_{qq}^2 (\bar{m})^2(\bar{Z})(t, q)| \\
\leq \frac{1}{\theta} |(\nabla_{qq}^3 g(\bar{Q}_0(q), \bar{Q}_0\#\mu)\nabla_q \bar{Q}_0(q))\nabla_q \bar{Q}_0(q) + \nabla_{qq}^2 g(\bar{Q}_0(q), \bar{Q}_0\#\mu)\nabla_{qq}^2 \bar{Q}_0(q)| \\
+ \frac{1}{\theta} \int_0^t |(\nabla_{qq}^3 H(\bar{Q}_r(q), \theta \bar{P}_r(q))\nabla_q \bar{Q}_r(q) + \nabla_{qq}^2 H(\bar{Q}_r(q), \theta \bar{P}_r(q))\theta \nabla_q \bar{P}_r(q))\nabla_q \bar{Q}_r(q) \\
+ \nabla_{qq}^2 H(\bar{Q}_r(q), \theta \bar{P}_r(q))\nabla_q \bar{Q}_r(q) \\
+ (\nabla_{qq}^3 H(\bar{Q}_r(q), \theta \bar{P}_r(q))\nabla_q \bar{Q}_r(q) + \nabla_{qq}^2 H(\bar{Q}_r(q), \theta \bar{P}_r(q))\theta \nabla_q \bar{P}_r(q))\theta \nabla_q \bar{P}_r(q) \\
+ \nabla_{qq}^2 H(\bar{Q}_r(q), \theta \bar{P}_r(q))\theta \nabla_q \bar{P}_r(q))\theta \nabla_q \bar{P}_r(q))\theta \nabla_{qq}^2 \bar{P}_r(q)| \, d\tau \\
+ \frac{1}{\theta} \int_0^t |(\nabla_{qq}^3 F(\bar{Q}_r(q), \bar{Q}_r\#\mu)\nabla_q \bar{Q}_r(q))\nabla_q \bar{Q}_r(q) + \nabla_{qq}^2 F(\bar{Q}_r(q), \bar{Q}_r\#\mu)\nabla_{qq}^2 \bar{Q}_r(q)| \, d\tau. 
\]

Let \( A_1, A_2, B, E, T \) be for the moment arbitrary positive numbers. Suppose that \( \bar{Z} \in \mathcal{M}_0 \), that is, \( \bar{Z} = (\bar{P}, \bar{Q}) \) satisfies (4.4). From the latter inequalities we see that:

(a) \( c/\theta + T\bar{h}(B)/\theta \leq B \) implies \( |(\bar{m})^2(\bar{Z})| \leq B \);
(b1) \( \bar{l}(B) \leq A_1 \) implies \( |\hat{c}_t(\bar{m})^1(\bar{Z})| \leq A_1 \);
(b2) \( \bar{l}(B)/\theta \leq A_2 \) implies \( |\hat{c}_t(\bar{m})^2(\bar{Z})| \leq A_2 \);
(c) \( c + T\bar{h}(B)(A_1 + \theta A_2) \leq A_1 \) implies \( |\nabla_q(\bar{m})^1(\bar{Z})| \leq A_1 \);
(d) \( cE/\theta + \frac{T}{\theta} \bar{h}(B)(A_1 + \theta A_2) \leq A_2 \) implies \( |\nabla_q(\bar{m})^2(\bar{Z})| \leq A_2 \);
(e) \( c + T\bar{h}(B)(A_1 + \theta A_2) \leq E \) implies \( |\nabla_q(\bar{m})^1(\bar{Z})|_{t=0} \leq E \);
(f1) \( T\bar{h}(B)(A_1 + \theta A_2)(A_1 + \theta A_2 + 1) \leq A_1 \) implies \( |\nabla_{qq}^2(\bar{m})^1(\bar{Z})| \leq A_1 \);
(f2) \( T\bar{h}(B)(A_1 + \theta A_2)(A_1 + \theta A_2 + 1) \leq E \) implies \( |\nabla_{qq}^2(\bar{m})^2(\bar{Z})|_{t=0} \leq E \);
(g) \( \frac{1}{\theta}cE(E + 1) + \frac{T}{\theta} \bar{h}(B)(A_1 + \theta A_2)(A_1 + \theta A_2 + 1) + \frac{T}{\theta} \bar{h}(B)A_1(A_1 + 1) \leq A_2 \) implies \( |\nabla_{qq}^2(\bar{m})^2(\bar{Z})| \leq A_2 \).

Thus, we will have the lemma if we show that there exist positive numbers \( A_1, A_2, B, E, T \) such that the above inequalities hold simultaneously. First choose \( B > c/\theta \). The number \( B \) now depends only on the data (through \( c \)), and thus \( \bar{l}(B), \bar{h}(B) \) depend only on the data, through \( B \). Let \( T \) be small enough that \( c/\theta + (T/\theta)\bar{l}(B) \leq B \) (i.e. \( T < (\theta B - c)/\bar{l}(B) \)). This gives (a). Choose \( E \) to be any number such that \( E > c \), and pick \( A_1, A_2 \) such that

\[
A_1 > \max\{\bar{l}(B), E, c\}, \tag{4.5}
\]
\[
A_2 > \max\left\{\frac{\bar{l}(B)}{\theta}, \frac{1}{\theta} cE(E + 1)\right\}. \tag{4.6}
\]
This gives (b1), (b2). Making $T$ possibly smaller by letting

$$
T < R := \min \left\{ \frac{\theta B - c}{l(B)} \cdot \frac{E - c}{h(B)(A_1 + \theta A_2)}, \frac{A_2 - cE(E + 1)/\theta}{(1/\theta)h(B)[(A_1 + \theta A_2)(A_1 + \theta A_2 + 1) + A_1(A_1 + 1)]} \right\},
$$

we make sure that: (e) holds, and, consequently, (c) holds because $A_1 > E$; (g) holds and therefore (d) as well; and (f2) holds, hence, (f1) is true.

\[\square\]

**Proposition 4.5.** (Contraction property) Let $\theta > 2\kappa$. Then there exist positive numbers $A_1, A_2, B, E, T$ such that for any $s \in [0, T]$, $\mu \in \mathcal{P}(\mathbb{T}^d)$, the operator $\bar{m}^{s, \mu}$ maps $\mathcal{M}_0(A, B, E, T)$ into itself and is a contraction.

**Proof.** We run the previous lemma to obtain the numbers $A_1$, $A_2$, $B$, $E$, and $T$, and decrease $T$, if necessary, so that

$$
T < \min \left\{ R, \frac{1 - 2\kappa}{\theta h(B)\sqrt{2(1 + 1/\theta) + 2\kappa/\theta}} \right\},
$$

where $R$ is the number defined in (4.7). Let $\bar{Z} = (\bar{Q}, \bar{P}), \bar{Z}' = (\bar{Q}', \bar{P}') \in \mathcal{M}_0$. Let $s \in [0, T]$, $\mu \in \mathcal{P}(\mathbb{T}^d)$ be arbitrary. We have, for the first component of $\bar{m}^{s, \mu}$, that

$$
|\langle \bar{m}^{s, \mu} \rangle^1(\bar{Z})(t, q) - \langle \bar{m}^{s, \mu} \rangle^1(\bar{Z}')(t, q)| \leq |s - t| \max_{t \leq \tau \leq s} |\nabla_p H(\bar{Q}_\tau(q), \theta \bar{P}_\tau(q)) - \nabla_p H(\bar{Q}'_\tau(q), \theta \bar{P}'_\tau(q))|.
$$

Since $H$ is $C^2$, we can write

$$
|\nabla P H(\bar{Q}_\tau(q), \theta \bar{P}_\tau(q)) - \nabla Q H(\bar{Q}'_\tau(q), \theta \bar{P}'_\tau(q))| \leq M_{1, q}[\bar{Z}(\tau, q) - \bar{Z}'(\tau, q)],
$$

where

$$
M_{1, q} = \max_{0 \leq \lambda \leq 1} |\nabla_p H|[(1 - \lambda)\bar{Q}_\tau(q) + \lambda \bar{Q}'_\tau(q), (1 - \lambda)\theta \bar{P}_\tau(q) + \lambda \theta \bar{P}'_\tau(q)] |
$$

For the second component of $\bar{m}$ we apply (2.2) to get

$$
|\langle \bar{m}^{s, \mu} \rangle^2(\bar{Z})(t, q) - \langle \bar{m}^{s, \mu} \rangle^2(\bar{Z}')(t, q)|
\leq \frac{\kappa}{\theta} \sqrt{Q_0(q) - Q_0'(q)}^2 + \left\| Q_0 - Q_0' \right\|_{L^2(\mu)}^2 + \frac{1}{\theta} \max_{0 \leq \tau \leq t} |\nabla Q H(\bar{Q}_\tau(q), \theta \bar{P}_\tau(q)) - \nabla Q H(\bar{Q}'_\tau(q), \theta \bar{P}'_\tau(q))|
$$

with

$$
|\nabla Q H(\bar{Q}_\tau(q), \theta \bar{P}_\tau(q)) - \nabla Q H(\bar{Q}'_\tau(q), \theta \bar{P}'_\tau(q))| \leq M_{2, q}[\bar{Z}(\tau, q) - \bar{Z}'(\tau, q)],
$$

where

$$
M_{2, q} = \max_{0 \leq \lambda \leq 1} |\nabla_p H|[(1 - \lambda)\bar{Q}_\tau(q) + \lambda \bar{Q}'_\tau(q), (1 - \lambda)\theta \bar{P}_\tau(q) + \lambda \theta \bar{P}'_\tau(q)] |
$$

14
But, since $\mathcal{M}_0$ is a convex subset of $\mathcal{M}$, it is true that $M_{\tau,q}^1, M_{\tau,q}^2 \leq \bar{h}(B), (\tau,q) \in [0,s] \times \mathbb{T}^d$. It follows that
\[
\left\| (\bar{m}^{s,\mu})^1(\bar{Z})(t,q) - (\bar{m}^{s,\mu})^1(\bar{Z}')(t,q) \right\| \leq |s-t|\bar{h}(B)\|\bar{Z} - \bar{Z}'\|_\infty, \quad 0 \leq t \leq s, \ q \in \mathbb{T}^d,
\]
and
\[
\left\| (\bar{m}^{s,\mu})^2(\bar{Z})(t,q) - (\bar{m}^{s,\mu})^2(\bar{Z}')(t,q) \right\| \leq \sqrt{\frac{1}{\theta}}\kappa(1 + t)\|\bar{Z} - \bar{Z}'\|_\infty + \frac{1}{\theta}\bar{h}(B)\|\bar{Z} - \bar{Z}'\|_\infty.
\]
Consequently, since $0 \leq t, s \leq T$, we obtain
\[
\left\| \bar{m}^{s,\mu}(\bar{Z}) - \bar{m}^{s,\mu}(\bar{Z}') \right\|_\infty \leq \sqrt{2}\bar{h}(B) + \sqrt{2}\left(\sqrt{\frac{1}{\theta}}\kappa(1 + T) + \frac{1}{\theta}\bar{h}(B)\right)\|\bar{Z} - \bar{Z}'\|_\infty
\]
\[
= \left[\frac{2\kappa}{\theta} + T(\bar{h}(B)\sqrt{2}(1 + \frac{1}{\theta}) + \frac{2\kappa}{\theta})\right]\|\bar{Z} - \bar{Z}'\|_\infty.
\]
Due to (4.8), the expression inside the square brackets in (4.9) is less than 1. □

It follows now that the operator (4.1) and (4.2) has a unique fixed point in $\mathcal{M}_0(A_1, A_2, B, E, T)$, where $A_1, A_2, B, E, T$ are as above.

**Definition 4.6.** Fix $\mu \in \mathcal{P}(\mathbb{T}^d)$, $s \in [0,T]$. Define the operator $m^{s,\mu} : \mathcal{M} \to \mathcal{M}$, $m = ((m^{s,\mu})^1, (m^{s,\mu})^2)$ as follows:

If $Z = (Q,P) \in \mathcal{M}$, then $m^{s,\mu}(Z) = ((m^{s,\mu})^1(Z), (m^{s,\mu})^2(Z))$, where :

\[
(m^{s,\mu})^1(Z)(t,q) = q + \int_s^t \nabla_p H(Q_\tau(q), P_\tau(q))d\tau, \tag{4.10}
\]

\[
(m^{s,\mu})^2(Z)(t,q) = \nabla_q g(Q_0(q), Q_{0,\#}) - \int_0^t \nabla_q H(Q_\tau(q), P_\tau(q)) + \nabla_q F(Q_\tau(q), Q_{\tau,\#})d\tau. \tag{4.11}
\]

$0 \leq t \leq T, \ q \in \mathbb{T}^d$.

**Corollary 4.7.**

(i) For any $T > 0$, $s \in [0,T], \mu \in \mathcal{P}(\mathbb{T}^d), \theta > 0$, the operator $\bar{m}$ has a unique fixed point in $\mathcal{M}_0(A_1, A_2, B, E, T)$ if and only if $m$ has a unique fixed point in $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, T)$.

(ii) With $\theta > 2\kappa$, fix $s \in [0,T], \mu \in \mathcal{P}(\mathbb{T}^d)$. Let $A_1, A_2, B, E, T$ be as obtained in Lemma 4.4. Then $m^{s,\mu}$ maps $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, T)$ into itself and the system (4.3) has a unique solution in $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, T)$.

**Proof.** (i) Suppose $\bar{m}^{s,\mu}$ has a unique fixed point $\Sigma[s, \mu] = (\bar{\Sigma}^1[s, \mu], \bar{\Sigma}^2[s, \mu])$ that satisfies (4.3) with $\bar{Q} = \bar{\Sigma}^1$ and $\bar{P} = \bar{\Sigma}^2$. Define

\[
\Sigma^1[s, \mu] := \bar{\Sigma}^1[s, \mu], \quad \Sigma^2[s, \mu] := \theta \Sigma^2[s, \mu]. \tag{4.12}
\]

Then it is straightforward to check that $\Sigma[s, \mu] = (\Sigma^1[s, \mu], \Sigma^2[s, \mu])$ is the unique fixed point of the operator $m^{s,\mu}$ such that the inequalities (4.4) are true for $\Sigma^1, \Sigma^2$ with the new constants $\theta A_2, \theta B$ in place of $A_2$ and $B$ respectively, that is, such that $Q = \Sigma^1, P = \Sigma^2$ satisfy

\[
\begin{align*}
\left\| \bar{\epsilon}_1 Q \right\|_{\mathcal{M}^1} &\leq A_1, \quad \left\| \nabla_q Q \right\|_{C([0,T] \times \mathbb{T}^d \times \mathbb{T}^d)} \leq A_1, \quad \left\| \nabla_{qq} Q \right\|_{C([0,T] \times \mathbb{T}^d \times \mathbb{T}^d \times \mathbb{T}^d)} \leq A_1; \\
\left\| \bar{\epsilon}_1 P \right\|_{\mathcal{M}^2} &\leq \theta A_2, \quad \left\| \nabla_q P \right\|_{C([0,T] \times \mathbb{T}^d \times \mathbb{R}^d \times \mathbb{R}^d)} \leq \theta A_2, \quad \left\| \nabla_{qq} P \right\|_{C([0,T] \times \mathbb{T}^d \times \mathbb{T}^d \times \mathbb{R}^d \times \mathbb{R}^d)} \leq \theta A_2; \\
\left\| P \right\|_{\mathcal{M}^2} &\leq \theta B; \\
\left\| \nabla_q Q \right\|_{C(\mathbb{T}^d \times \mathbb{T}^d)}, \left\| \nabla_{qq} Q \right\|_{C(\mathbb{T}^d \times \mathbb{T}^d)} &\leq E.
\end{align*}
\]

The sufficiency part of the statement is equally easily verified.

15
(ii) We take \( Z = (Q, P) \in \mathcal{M}_0(A_1, \theta A_2, \theta B, E, T) \), and evaluate \( \bar{m}^{s, \mu} \) at \( \bar{Z} = (\bar{Q}, \bar{P}) \) where \( \bar{Q} = Q \) and \( \bar{P} = P/\theta \). Then, by Lemma 4.4, \( \bar{m}^{s, \mu}(\bar{Z}) \in \mathcal{M}_0(A_1, A_2, B, E, T) \). But \( (\bar{m}^{s, \mu})^1(\bar{Z}) = (m^{s, \mu})^1(Z) \) and \( (\bar{m}^{s, \mu})^2(\bar{Z}) = \frac{1}{\theta}(m^{s, \mu})^1(Z) \), thus, \( m^{s, \mu}(Z) \in \mathcal{M}_0(A_1, \theta A_2, \theta B, E, T) \). Furthermore, Proposition 4.3 provides a unique fixed point \( \bar{\Sigma}[s, \mu] \) of \( \bar{m} \) in \( \mathcal{M}_0(A_1, \theta A_2, \theta B, E, T) \). Defining \( \Sigma[s, \mu] \) as in (4.12), then, by (i), \( \Sigma[s, \mu] \) is the unique fixed point of the operator \( m^{s, \mu} \) on \( \mathcal{M}_0(A_1, \theta A_2, \theta B, E, T) \), so it is the unique solution to (4.3) in \( \mathcal{M}_0(A_1, \theta A_2, \theta B, E, T) \).

We stress that, as long as \( T \) is as in (4.8), we have a solution \( \Sigma[s, \mu] \) to (4.3) for any \( \mu \in \mathcal{P}(\mathbb{T}^d) \), and \( 0 \leq s \leq T \), and moreover, its \( Q \) and \( P \) components \( \Sigma^1[s, \mu] \) and \( \Sigma^2[s, \mu] \) satisfy the bounds (4.4) with \( \theta A_2 \) and \( \theta B \) in place of \( A \) and \( B \) respectively, independently of \( \mu \), with solutions being continuous and differentiable in \( t \) and \( q \).

**Remark 4.8.** The preceding proofs make it clear that \( T \) can be assumed to be smaller if necessary at each following step, without affecting the validity of the previous statements. We choose to refer back to this remark in later stages instead of imposing tighter bounds on \( T \) than (4.8) above that would make their purpose unclear at first reading. We may sometimes just say “\( T \) is small”, having this remark in mind. \( \Box \)

### 4.2 First regularity properties

**Lemma 4.9.** For any fixed \( Z = (Q, P) \in \mathcal{M}, t \in [0, T], q \in \mathbb{R}^d, \) and \( \mu \in \mathcal{P}(\mathbb{T}^d) \), the function

\[
s \mapsto m^{s, \mu}(Z)(t, q)
\]

is continuous. Likewise, for any fixed \( Z \in \mathcal{M}, t \in [0, T], q \in \mathbb{R}^d, \) and \( s \in [0, T] \), the function

\[
\mu \mapsto m^{s, \mu}(Z)(t, q)
\]

is continuous.

**Proof.** Continuity of \( s \mapsto m^{s, \mu}(Z)(t, q) \) for fixed \( t, q, \mu \) is immediate. To show continuity with respect to \( \mu \), let us first look at the definition of \( (m^{s, \mu})^2 \), equation (4.11). For each \( \tau, 0 \leq \tau \leq T \), the function \( q \mapsto Q_\tau(q) \) is Lipschitz. This implies (see, e.g., [25, Remark 3.3]) that the function \( \mu \mapsto Q_\tau \# \mu \) is Lipschitz from \( \mathcal{P}(\mathbb{T}^d) \) into itself, with the same constant. Furthermore, since the mapping \( (\tau, q) \mapsto Q_\tau(q) \) is Lipschitz, the Lipschitz constants of the functions \( q \mapsto Q_\tau(q) \) are bounded with respect to \( \tau, 0 \leq \tau \leq T \). These facts, combined with the Lipschitz continuity of \( \nabla_q F \) and \( \nabla_q g \), show that

\[
\mu_n \rightarrow \mu \text{ in } \mathcal{P}(\mathbb{T}^d) \implies m^{s, \mu_n}(Z)(t, q) \rightarrow m^{s, \mu}(Z)(t, q)
\]

for all \( Z \in \mathcal{M}_0, t, s \in [0, T], q \in \mathbb{R}^d \).

We will need the continuity and differentiability of the fixed point \( \Sigma[s, \mu] \) with respect to \( s \), and its continuity with respect to \( \mu \). This is addressed in Lemmas 4.12 and 4.15 below. Before that, let us name the data bounds that will appear in the calculations.

**Definition 4.10.** (Data bounds II) For \( B > 0 \), let

\[
l(B) := \max_{q \in \mathbb{R}^d, |p| \leq B, \mu \in \mathcal{P}(\mathbb{T}^d)} \sqrt{2} |\nabla H(q, p)| + |\nabla_q F(q, \mu)|,
\]

\[
h(B) := \max_{q \in \mathbb{R}^d, |p| \leq B, \mu \in \mathcal{P}(\mathbb{T}^d)} \sqrt{2} |\nabla^2 H(q, p)| + \sqrt{2} |\nabla^3 H(q, p)| + |\nabla^2_{qq} F(q, \mu)| + |\nabla^3_{qqq} F(q, \mu)|.
\]
Unlike the data bounds $l(B)$ and $h(B)$, here $l(B)$ and $h(B)$ are independent of the number $θ$. However, if, say, $(Q, P) ∈ M_0(A_1, θA_2, θB, E, T)$, then $|∇_p H(Q(t, q), P(t, q))| ≤ l(θB)$.

**Definition 4.11.** For any $D = (D_1, D_2)$, $D_1, D_2 > 0$, define:

1. $\mathcal{M}^*_0, D(A^1, A^2/θB, E, T) ⊂ W^{1,2,∞}(0, T] × [0, T] × T^d; T^d × R^d)$ as the subset of those $Z(\cdot; \cdot, \cdot)$ such that, for each $s ∈ [0, T]$, $Z(s; \cdot, \cdot) ∈ M_0(A^1, A^2/θB, E, T)$, and

   \[
   \|σ_\tau Q(s; \cdot, \cdot)\|_x ≤ D_1, \quad \|σ_\tau P(s; \cdot, \cdot)\|_x ≤ D_2
   \]  

   (4.15)

   wherever $σ_\tau Q, σ_\tau P$ are defined;

2. $Q^*_0, D(A^1, A^2/θB, E, T) ⊂ W^{1,2,∞}(0, T] × [0, T] × T^d; T^d)$ as the subset of those $Q(\cdot; \cdot, \cdot)$ such that, for each $s ∈ [0, T]$, $(Q(s; \cdot, \cdot), 0) ∈ M_0^*(A_1, θA_2, θB, E, T)$, and

   \[
   \|σ_\tau Q(s; \cdot, \cdot)\|_x ≤ D_1
   \]  

   (4.16)

   wherever $σ_\tau Q$ is defined.

By an argument similar to that of Proposition 4.3, the sets $M_0^*$ and $Q_0^*$ just defined are closed subsets for the uniform convergence of $C((0, T] × [0, T] × T^d; T^d × R^d)$ and $C((0, T] × [0, T] × T^d; T^d)$ respectively.

**Lemma 4.12.** For fixed $μ ∈ P(T^d)$, using the same notation of Definition 4.11 and $θ > 2κ$:

1. There exists a pair of positive constants $D = (D_1, D_2)$ such that, if $Z ∈ M_0^*(A_1, θA_2, θB, E, T)$, then the function

   \[
   (s, t, q) \mapsto m^μ(Z(s; \cdot, \cdot))(t, q)
   \]

   belongs to $M_0^*(A_1, θA_2, θB, E, T)$ for any $μ ∈ P(T^d)$.

2. The mapping

   \[
   s \mapsto Σ[s, μ](t, q)
   \]

   is differentiable in $s$ a.e. on the interval $0 < s < T$, for every $μ ∈ P(T^d)$, $t ∈ [0, T]$, and $q ∈ T^d$, and it satisfies

   \[
   \|σ_\tau Σ^1[s, μ](\cdot, \cdot)\|_{C[[0, T] × T^d; T^d]} ≤ D_1, \quad \|σ_\tau Σ^2[s, μ](\cdot, \cdot)\|_{C[[0, T] × T^d; T^d]} ≤ D_2
   \]

   for a.e. $s ∈ [0, T], μ ∈ P(T^d)$.

**Proof.** Given a function $Z ∈ M_0^*(A_1, θA_2, θB, E, T)$, define $m^μ(Z) ∈ C([0, T] × [0, T] × T^d; T^d × R^d)$ to be the first function displayed in the statement.

1. Let us show that $m^μ(Z) ∈ M_0^*(A_1, θA_2, θB, E, T)$ for an appropriate $D$. Indeed, that $m^μ(Z)$ is continuous is evident. For a.e. $s ∈ (0, T)$,

   \[
   \hat{σ}_s Q'(s; t, q) = -∇_p H(Q(s; s, q), P(s; s, q)) + \int_t^s [∇^2_{qq} H(Q(s; t, q), P(s; t, q)) \hat{σ}_s Q(s; t, q)
   \]

   \[
   + ∇^2_{pp} H(Q(s; t, q), P(s; t, q)) \hat{σ}_s P(s; t, q)] dt,
   \]

   (4.17)
so

\[ \| \partial_s Q'(s; \cdot, \cdot) \|_\infty \leq l(\theta B) + Th(\theta B)(D_1 + D_2). \]

Put \( Q(s; \tau, \cdot) =: \sigma^*_s, 0 \leq \tau \leq T. \) Then, for a.e. \( s \in (0, T), \)

\[
\partial_s P'(s; t, q) = (\partial_s)(g(Q(s; 0, q), \sigma^*_0)) - \int_0^t [\nabla^2_{qq} H(Q(s; \tau, q), P(s; \tau, q)) \partial_s Q(s; \tau, q) + \nabla^2_{pq} H(Q(s; \tau, q), P(s; \tau, q)) \partial_s P(s; \tau, q)] + (\partial_s)(g(Q(s; \tau, q), \sigma^*_s))]d\tau. \]

By the joint Lipschitz constants of \( g \) and \( F \) being bounded by \( \kappa \), we obtain

\[ \| \partial_s P'(s; \cdot, \cdot) \|_\infty \leq 2\kappa D_1 + T(2\kappa D_1 + h(\theta B)(D_1 + D_2)). \]

Thus, if \( D_1 > l(\theta B) \), and \( D_2 > 2\kappa D_1 \), we refer to Remark 4.13 and assume that

\[ T < \min \left\{ \frac{D_1 - l(\theta B)}{h(\theta B)(D_1 + D_2)}, \frac{D_2 - 2\kappa D_1}{2\kappa D_1 + h(\theta B)(D_1 + D_2)} \right\}. \]

to obtain \( \| \partial_s Q'(s; \cdot, \cdot) \|_\infty \leq D_1 \) and \( \| \partial_s P'(s; \cdot, \cdot) \|_\infty \leq D_2. \)

The fact that \( m^\mu(Z)(s; \cdot, \cdot) \in M_0(A_1, \theta A_2, \theta B, E, T) \) follows from Corollary 4.17.

(ii) Let \( Z^0 \in M_{0,D}(A_1, \theta A_2, \theta B, E, T) \) be arbitrary, with \( D \) from part (i). Define inductively \( Z^k = m^\mu(Z^{k-1}), k = 1, \ldots \). Then \( \{Z^k\}_{k=0}^\infty \) is a sequence in \( M_{0,D}^*(\cdot, \cdot) \), and for each \( s \in [0, T] \), \( Z^k(s; \cdot, \cdot) = m^\mu(Z^{k-1}(s; \cdot, \cdot)) \), so for each fixed \( s \in [0, T] \),

\[ Z^k(s; \cdot, \cdot) \longrightarrow \Sigma[s, \mu](\cdot, \cdot) \text{ uniformly in } M_0(A_1, \theta A_2, \theta B, E, T), \]

by the fixed point theorem. We thus have pointwise convergence of \( Z^k(\cdot; \cdot, \cdot) \) \( \in Z[\cdot, \cdot, \cdot](\cdot, \cdot, \cdot) \). We call on the equicontinuity, uniform boundedness of the sequence and the periodicity of the functions (i.e., they are defined on \([0, T] \times [0, T] \times \mathbb{T}^d \)) to conclude that this convergence is actually uniform. The closedness of the subspace \( M^*_{0,D}(\cdot, \cdot) \) with respect to uniform convergence now ensures that \( \Sigma[\cdot, \cdot, \cdot](\cdot, \cdot, \cdot) \) belongs to this subspace, so it is differentiable with respect to \( s \) for a.e. \( s \in [0, T] \) and satisfies (4.15).

In the following, the constants \( D_1, D_2 \) will always be as in Lemma 4.12. The next remark will not be used before Section 5.

**Remark 4.13.** If \( Z = (Q, P) \) and \( \tilde{Z} = (\tilde{Q}, \tilde{P}) \) are related as in the proof of Corollary 4.17(ii), that is, \( Q = \tilde{Q}, P = \tilde{P} \), then \( Z \in M^*_{0,D}(A_1, \theta A_2, \theta B, E, T) \) if and only if \( \tilde{Z} \in M^*_{0,D}(A_1, A_2, B, E, T) \), where

\[ \tilde{D} = (\tilde{D}_1, \tilde{D}_2), \quad \tilde{D}_1 := D_1, \quad \tilde{D}_2 := D_2/\theta. \quad \quad (4.17) \]

**Definition 4.14.** *(Master map)* The mapping

\[ \mathcal{M} : [0, T] \times [0, T] \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \longrightarrow [0, T] \times [0, T] \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \]

given by

\[ \mathcal{M}(t, s, q, \mu) = (t, s, \Sigma^1[s, \mu](t, q), \mu) \]

will be called the master map.
Lemma 4.15. Let \( \theta > 2k \). The master map \( M \) is continuous, and for any fixed \( \mu \in \mathcal{P}(\mathbb{T}^d) \), \( M(\cdot,\cdot,\cdot,\mu) \) is a \( C^1 \) diffeomorphism.

Proof. Let \( \{\mu_k\}_{k=1}^\infty \) be a sequence in \( \mathcal{P}(\mathbb{T}^d) \) converging to \( \mu \in \mathcal{P}(\mathbb{T}^d) \). Consider the sequence \( \{\Sigma^1[\cdot,\mu_k](\cdot,\cdot)\}_{j=1}^\infty \) and an arbitrary subsequence \( \{\Sigma^1[\cdot,\mu_{k_j}](\cdot,\cdot)\}_{j=1}^\infty \). Being in \( Q_{0,D}^*(A_1,\theta A_2,\theta B,E,T) \), the latter is equicontinuous and uniformly bounded, so there is a sub-subsequence, which we still index with \( j \), converging to \( S \) for some \( S \in \mathcal{Q}_{0,D}^*(\cdot,\cdot) \) as \( j \to \infty \). For each \( s \in [0,T] \), on one hand, \( \Sigma^1[\cdot,\mu_{k_j}](\cdot,\cdot) \to S(\cdot,\cdot) \).

On the other, since, by Corollary 4.7 and Lemma 4.9, the mapping \( (s,Z,\mu) \mapsto m^s\mu(Z) \) is a continuous mapping of \( [0,T] \times \mathcal{M}_0(A_1,\theta A_2,\theta B,E,T) \times \mathcal{P}(\mathbb{T}^d) \) into \( \mathcal{M}_0(A_1,\theta A_2,\theta B,E,T) \) (because the Lipschitz constant of \( m^s\mu \) is independent of \( s \) and \( \mu \)), we have

\[
\Sigma^1[\cdot,\mu_{k_j}](\cdot,\cdot) = (m^1)^{s,\mu_{k_j}}(\Sigma^1[\cdot,\cdot](\cdot,\cdot)) \quad \to \quad (m^1)^{s,\mu}(S(\cdot,\cdot)).
\]

Therefore, for each \( s \in [0,T] \), \( S(\cdot,\cdot) = (m^1)^{s,\mu}(S(\cdot,\cdot)) \), that is, \( S(\cdot,\cdot) \) is a fixed point of \( (m^1)^{s,\mu} \), so, by uniqueness, \( S(\cdot,\cdot) = \Sigma^1[\cdot,\cdot](\cdot,\cdot) \). Thus, every subsequence of \( \{\Sigma^1[\cdot,\mu_k](\cdot,\cdot)\}_{j=1}^\infty \) has a subsequence that converges to \( \Sigma^1[\cdot,\cdot](\cdot,\cdot) \in \mathcal{Q}_{0,D}^*(A_1,\theta A_2,\theta B,E,T) \). Hence,

\[
\Sigma^1[\cdot,\mu_k](\cdot,\cdot) \to \Sigma^1[\cdot,\cdot](\cdot,\cdot) \quad \text{uniformly},
\]

and this implies that \( M \) is continuous.

The second assertion of the lemma will be an immediate consequence of Lemma 4.16.

Lemma 4.16. Let \( \theta > 2k \), \( 0 \leq s \leq t \leq T \), \( \mu \in \mathcal{P}(\mathbb{T}^d) \). The mapping \( q \mapsto \Sigma^1[s,\mu](t,q) \) is a \( C^1 \) diffeomorphism, for \( 0 \leq t \leq T \), with

\[
\frac{1}{2} < \det \nabla_q \Sigma^1[s,\cdot](t,q), \quad |(\nabla_q \Sigma^1[s,\mu](t,q))^{-1}| < 4(1 + \sqrt{d})^{d-1}, \quad (4.18)
\]

provided \( T \) is sufficiently small.

Proof. We know the mapping is already \( C^1 \) because \( W^{2,\infty} \) mappings are continuously differentiable. To prove invertibility, put \( \Theta(t,q) := \Sigma^1(t,q) - q \). Computing \( \nabla_q \Theta(t,q) \), we have

\[
|\nabla_q \Theta(t,q)| \leq |s-t|(A_1 + \theta A_2)h(\theta B). \quad (4.19)
\]

because \( \Sigma \in \mathcal{M}_0(A_1,\theta A_2,\theta B,E,T) \). By Remark 4.8 this means the function \( q \mapsto \Theta(t,q) \) has Lipschitz constant strictly less than 1. Therefore, the function \( q \mapsto q + \Theta(t,q) = \Sigma^1(t,q) \) is injective, for \( 0 \leq t \leq T \).

To prove that \( q \mapsto \Sigma^1(t,q) \) is onto, note that \( \sup_{q \in \mathbb{R}^d} |\Sigma^1(t,q) - q| \lesssim Tl(\theta B) < 2Tl(\theta B), \) for \( 0 \leq t \leq T \). Let \( y \) be a point in the ball of radius \( R - 2Tl(\theta B) \) in \( \mathbb{R}^d \) centered at the origin, where \( R > 1 > 2Tl(\theta B) \). Then for all \( q \) on the boundary of \( B_R(0) \), the ball of radius \( R \) in \( \mathbb{R}^d \) centered at the origin, we have: \( \Sigma^1(t,q) \neq y \), for \( 0 \leq t \leq T \). Therefore

\[
f(t) := \deg(\Sigma^1, B_R(0), y), \quad 0 \leq t \leq T,
\]

the topological degree of \( \Sigma^1 \) is well defined at \( y \in B_{R-2Tl(\theta B)}(0) \). This counts the number of “signed” solutions (see, e.g., [22]) \( x \) in \( B_R(0) \) of the equation \( \Sigma^1(x) = y \). Since \( f \) is a continuous function taking
on integer values only, we conclude that \( f(t) = f(s) = 1 \). This means that the range of \( \Sigma^1_t \) includes \( B_{R-2T}(\theta B) \). Since \( R > 1 \) is arbitrary, we conclude that the range of \( \Sigma^1_t \) is \( \mathbb{R}^d \).

We will denote the inverse of \( \Sigma^1 \) by \( X \), so

\[
X[s, \mu](t, q) = [\Sigma^1[s, \mu](t)]^{-1}(q),
\]

for \( 0 < s \leq T, 0 \leq t \leq T, q \in \mathbb{T}^d, \mu \in \mathcal{P}(\mathbb{T}^d) \). Next, note that for \( 0 \leq t \leq T, q \in \mathbb{T}^d \), \( |\nabla_q \Sigma^1_t(q) - I_d| \leq T(A_1 + \theta A_2)h(\theta B) \), since \( T \) is small, where \( I_d \) is the \( d \times d \) matrix with 1's in the diagonal and 0's everywhere else. This implies that \( \nabla_q \Sigma^1_t(q) \) is an invertible matrix, for \( 0 \leq t \leq T, q \in \mathbb{T}^d \). By the inverse function theorem, \( X_t \) is differentiable. Moreover, since

\[
\nabla_q X_t(q) = [\nabla_q \Sigma^1_t(X_t(q))]^{-1},
\]

and \( q \rightarrow \nabla_q \Sigma^1_t(q) \) is continuous, continuity of matrix inversion gives that the mapping \( q \rightarrow \nabla_q X_t(q) \) is continuous; this means that \( X \) is \( C^1 \) in \( q \).

To show (4.18), we may use the fact that the determinant function \( \det : \mathbb{R}^{d_2} \rightarrow \mathbb{R} \) has derivative \( \nabla \det \) satisfying

\[
|\nabla \det(\xi)| \leq 2|\xi|^{d-1}, \quad \xi \in \mathbb{R}^d
\]

and the inverse matrix formula

\[
\xi^{-1} = \frac{1}{\det \xi}(\nabla \det \xi)^t,
\]

where the superscript \( t \) denotes transposition. By the mean-value theorem, there is \( \tau \in [0, 1] \) such that

\[
|\det(I_d + \frac{s}{T} \nabla q \Theta) - \det I_d| \leq \frac{s}{T} |\nabla \det(I_d + \tau \frac{s}{T} \nabla q \Theta)||\nabla q \Theta| \leq \frac{s}{T} 2|I_d + \tau \frac{s}{T} \nabla q \Theta|^{d-1}||\nabla q \Theta|
\]

\[
\leq 2(\sqrt{d} + |\nabla q \Theta|)^{d-1}|\nabla q \Theta| \leq 2(1 + \sqrt{d})^{d-1}T(A_1 + A_2)h(\theta B)
\]

by (1.19) and because \( T \) is small enough that

\[
T < \frac{1}{4(1 + \sqrt{d})^{d-1}(A_1 + \theta A_2)h(\theta B)}.
\]

Since \( I_d + \nabla q \Theta(t, q) = \nabla q \Sigma^1(t, q) \) and \( \det I_d = 1 \), we obtain the first inequality in (4.18). Using the inverse matrix formula and the inequality \( |\nabla \det(\xi)| \leq 2|\xi|^{d-1} \) once more, we have

\[
|\nabla_q X(t, q)| = |(I_d + \nabla q \Theta)^{-1}| = |(\det(I_d + \nabla q \Theta))^{-1}[\nabla \det(I_d + \nabla q \Theta)]^t|
\]

\[
\leq \frac{2}{\det(I_d + \nabla q \Theta)}|I_d + \nabla q \Theta|^{d-1} \leq \frac{2}{\det(I_d + \nabla q \Theta)}(1 + \sqrt{d})^{d-1}
\]

and since this holds for any \( t \in [0, T] \) and \( q \in \mathbb{T}^d \), we have obtained the second inequality in (4.18).

\[\text{See, for instance, [25, Remark 3.11].}\]
**Bound on** $\nabla_q X$. Due to the formula $\nabla_q X_t(q) = (\nabla_q \Sigma_t^1)^{-1} \circ X_t(q)$, i.e., formula (4.21), the second inequality in (4.18) implies

$$\|\nabla_q X[s, \mu](t, \cdot)\|_{C(T^d; T^d \times T^d)} < 4(1 + \sqrt{d})^{d-1}.$$ (4.21)

**Definition 4.17.** Let $\theta > 2\kappa$. Given $\mu \in \mathcal{P}(\mathbb{T}^d)$, $s \in [0, T]$, and $\Sigma$ the unique fixed point of $m^s \mu$ in $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, T)$, set

$$\sigma_t = \sigma(t) := \Sigma^1_t[s, \mu](t, \cdot) \# \mu, \quad v[s, \mu](t, q) := \partial_t \Sigma^1_t[s, \mu] \circ X_t[s, \mu](q),$$ (4.22)

for $0 \leq t \leq T$.

It should be kept in mind that the path $\sigma_t$ depends on $s$ and $\mu$. Also, the arguments $s, \mu$ may often be omitted in the notation for $v$, as has been done for $\Sigma$.

**Proposition 4.18.** The path $\sigma$ belongs to $AC^2(0, T; \mathcal{P}(\mathbb{T}^d))$ and $v$ is a velocity associated to $\sigma$, that is, $\partial_t \sigma + \nabla \cdot (\sigma v) = 0$ in the distribution sense, with $v_t = v(t, \cdot) \in L^2(\mathbb{T}^d, \sigma_t), 0 < t < T$.

**Proof.** Using (2.2),

$$W(\sigma_t, \sigma_{t+h}) \leq W^2(\sigma_t, \sigma_{t+h}) \leq \|\Sigma^1_t - \Sigma^1_{t+h}\|_{L^2(\mu)}^2 = \int_{\mathbb{R}^d} |\Sigma^1_t(q) - \Sigma^1_{t+h}(q)|^2 \mu(dq)$$

$$= \int_{\mathbb{R}^d} \left| \int_t^{t+h} \nabla_p H(\Sigma^1(t, q), \Sigma^2(t, q))d\tau \right|^2 \mu(dq) \leq \int_{\mathbb{R}^d} \left( \int_t^{t+h} \left| \nabla_p H(\Sigma^1(t, q), \Sigma^2(t, q)) \right|d\tau \right)^2 \mu(dq)$$

$$\leq l(\theta B)^2 h^2 \leq \left( \int_t^{t+h} l(\theta B)d\tau \right)^2.$$ 

Thus $\sigma \in AC^2(0, T; \mathcal{P}(\mathbb{T}^d))$. Next, if $\varphi \in C_c^\infty(\mathbb{T}^d)$, then

$$\int_{\mathbb{R}^d} \nabla \varphi(q) \cdot v_t(q) \sigma_t(dq) = \int_{\mathbb{R}^d} \nabla \varphi(q) \cdot [\partial_t \Sigma^1_t \circ X_t(q)](\Sigma^1_t, \#) \mu(dq)$$

$$= \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(\Sigma^1_t(q)) \mu(dq) = \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(q) \sigma_t(dq),$$

which shows that $v$ is a velocity vector field for $\sigma$. Clearly, $v_t \in L^2(\mathbb{T}^d, \sigma_t)$ for every $0 < t < T$ because $v$ is smooth on $\mathbb{R}^d$ and periodic.

**Note.** That the definition of the field $v_t$ means that the mappings $t \mapsto \Sigma^1_t(q)$ are the flow lines of $v_t$.

**Proposition 4.19.** Let $\mu_n \rightarrow \mu$ in $\mathcal{P}(\mathbb{T}^d)$. Then

$$v[\cdot, \mu_n](\cdot, \cdot) \rightarrow v[\cdot, \mu](\cdot, \cdot) \quad \text{in} \ C([0, T] \times [0, T] \times \mathbb{T}^d, \mathbb{R}^d).$$

**Proof.** Due to the uniform bound on $\nabla_q X$ (4.21), the sequence $\{v[\cdot, \mu_n](\cdot, \cdot)\}_{n=1}^\infty$ is equicontinuous and uniformly bounded, so an argument entirely analogous to the proof of Lemma 4.15 yields the conclusion.

We conclude this section by formally defining the inverse of the master map introduced above.

**Definition 4.20.** (Inverse master map) The inverse master map, denoted $\mathcal{R}$, is the mapping

$$\mathcal{R} : [0, T] \times [0, T] \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow [0, T] \times [0, T] \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d)$$

$$(t, s, q, \mu) \mapsto (t, s, X[s, \mu](t, q), \mu).$$
4.3 Value function $U$ and characteristics

In what follows, the hypotheses of Corollary 4.7 will be in force, with $\Sigma$ denoting the solution to (4.3). The following statement stems from the fact that $\Sigma^1[s, \mu](\cdot, \cdot) \in W^{2,2;\infty}((0, T) \times \mathbb{T}^d; \mathbb{T}^d)$; we omit its proof.

**Proposition 4.21.** For every $s \in [0, T]$, $\mu \in \mathcal{P}(\mathbb{T}^d)$, the function $X[s, \mu](\cdot, \cdot)$ is in $W^{2,2;\infty}((0, T) \times \mathbb{T}^d; \mathbb{T}^d)$.

Define now

$$V[s, \mu](t, q) := \Sigma^2[s, \mu] \circ X_t[s, \mu](q) = \Sigma^2[s, \mu] \circ (\Sigma^1[s, \mu])^{-1}(q),$$

(4.23)

for $s, t \in [0, T]$, $\mu \in \mathcal{P}(\mathbb{T}^d), q \in \mathbb{T}^d$. Alternatively, we may write $V_t[s, \mu](q)$. We can now proceed to solve the MFG system.

**Lemma 4.22.** Let $T$ be small according to Remark 4.8 $s \in [0, T]$, $\mu \in \mathcal{P}(\mathbb{T}^d)$, $\sigma_t = \Sigma^1[s, \mu](t, \cdot) \# \mu$, and for each $q \in \mathbb{T}^d$, let $U(t, q) = z(t, X[s, \mu](t, q))$, $t \in [0, T]$,

where $z(\cdot, q)$ satisfies

$$\partial_t z(t, q) = \Sigma^2[s, \mu](t, q) \cdot \nabla_H (\Sigma^1[s, \mu](t, q), \Sigma^2[s, \mu](t, q))$$

$$- H(\Sigma^1[s, \mu](t, q), \Sigma^2[s, \mu](t, q)) - F(\Sigma^1[s, \mu](t, q), \sigma_t) \quad \text{in } (0, T),$$

$$z(0, q) = g(\Sigma^1[s, \mu](0, q), \sigma_0).$$

(4.25)

Then $U \in C^1((0, T) \times \mathbb{T}^d)$ and solves the Hamilton-Jacobi equation of the mean-field game system:

$$\partial_t U(t, x) + H(x, \nabla_x U(t, x)) + F(x, \sigma_t) = 0 \quad \text{in } (0, T) \times \mathbb{T}^d,$$

(1.2)

$$U(0, \cdot) = g(\cdot, \sigma_0).$$

(1.4)

**Proof.** Since $s$ and $\mu$ are fixed, we abbreviate $\Sigma_t(s, \mu)(t, q) = \Sigma_t(q)$. Observe that the right-hand side of (4.24) is $C^1$ in $q$, $C^0$ in $t$, so $z$ is $C^1$ in $q$, $C^1$ in $t$. Therefore, $U$ is $C^1$ in both variables $t$ and $q$, because of Proposition 4.21. Moreover, since $\partial_t z(t, q)$ is $C^1$ in $q$, then (e.g. see [39 Thm. 9.41]) $\nabla^2_{qq} z$ exists and is equal to $\nabla^2_{qq} z$. Thus, the calculations below are legitimate. We have $U_t[t, \Sigma^1_t(q)] = z(t, q)$, so $\partial_t z(t, q) = \partial_t(U(t, \Sigma^1_t(q)))$ and

$$\begin{align*}
\partial_t z(t, q) &= \partial_t U(t, \Sigma^1_t(q)) = \partial_t U(t, \Sigma^1_t(q) + \nabla_q U(t, \Sigma^1_t(q)) \cdot \partial_t \Sigma^1_t(q) \\
&= \partial_t U(t, \Sigma^1_t(q)) + \nabla_q U(t, \Sigma^1_t(q)) \cdot \nabla_H (\Sigma^1_t(q), \Sigma^2_t(q))
\end{align*}$$

(4.26)

Now, if

$$\nabla_q U(t, \Sigma^1_t(q)) = \Sigma^2_t(q), \quad t \in (0, T), q \in \mathbb{T}^d,$$

(4.27)

then, comparing (4.26) and (4.24), we get

$$\partial_t U(t, \Sigma^1_t(q)) = -H(\Sigma^1_t(q), \Sigma^2_t(q)) - F(\Sigma^1_t(q), \sigma_t),$$

and the change of variable $x = \Sigma^1_t(q)$ then yields (1.2), (1.4). We set out to prove (4.27) now. Let

$$r_i(t) := \frac{\partial z(t, q)}{\partial q_i} - \sum_{j=1}^d (\Sigma^2_t)^{(j)}(q) \frac{\partial x_{(j)}}{\partial q_i},$$

22
As for \( b \), differentiate \( a \) so and, by the second line in (4.3), namely, 
\[
\dot{B} \text{to } \dot{B}
\]
By (4.27) in the proof of the preceding lemma, we have 
Corollary 4.23. 

Using the first line, \( \dot{a} \Sigma_{1}^{1}(q) = \nabla_{p} H(\Sigma_{1}^{1}(q), \Sigma_{2}^{1}(q)) \), in (4.3), we have

\[
a = 
\sum_{k=1}^{d} \left[ \sum_{k=1}^{d} \partial_{q^{(i)}}(\Sigma_{2}^{1}(k)) \partial_{t} \partial_{q^{(k)}}(\Sigma_{1}^{1}(k)) \right] 
- \sum_{l=1}^{d} [\partial_{q^{(i)}} H(\Sigma_{1}^{1}(q), \Sigma_{2}^{1}(q)) \partial_{q^{(l)}}(\Sigma_{1}^{1}(l)) + \partial_{p^{(i)}} H(\Sigma_{1}^{1}(q), \Sigma_{2}^{1}(q)) \partial_{q^{(l)}}(\Sigma_{2}^{1}(l))] 
- \sum_{l=1}^{d} \partial_{q^{(l)}} F(\Sigma_{1}^{1}(q), \sigma_{l}) \partial_{q^{(l)}}(\Sigma_{1}^{1}(l)),
\]
and, by the second line in (4.3), namely, \( \dot{a} \Sigma_{2}^{1}(q) = -\nabla_{q} H(\Sigma_{1}^{1}(q), \Sigma_{2}^{1}(q)) - \nabla_{q} F(\Sigma_{1}^{1}(q), \sigma_{l}) \), this simplifies to

\[
a = \sum_{k=1}^{d} \left[ \left( \Sigma_{2}^{1}(k) \right) \partial_{t} \partial_{q^{(l)}}(\Sigma_{1}^{1}(k)) + \dot{a} \dot{\Sigma}_{2}^{1}(k) \partial_{q^{(l)}}(\Sigma_{1}^{1}(k)) \right].
\]

As for \( b \),

\[
b = \dot{a} \left( \sum_{j=1}^{d} \left( \Sigma_{2}^{1}(j) \right) \partial_{q^{(l)}}(\Sigma_{1}^{1}(j)) \right)
= \sum_{j=1}^{d} \left( \dot{a} \Sigma_{2}^{1}(j) \partial_{q^{(l)}}(\Sigma_{1}^{1}(j)) + \left( \Sigma_{2}^{1}(j) \right) \partial_{q^{(l)}}(\Sigma_{1}^{1}(j)) \right),
\]

so \( a = b \). Therefore \( \dot{r}(t) \equiv 0 \), and \( r(t) \equiv 0 \) on \((0, T]\), by the uniqueness of (1.24) \([0, T]\). Now we differentiate \( U \), keeping in mind that \( U(t, x) = z(t, q) \); using the fact that \( r_{i}(t) = 0, 0 \leq t \leq T \), we have

\[
\partial_{x^{(i)}} U(t, x) = \sum_{j=1}^{d} \frac{\partial z(t, q)}{\partial q^{(j)}} \frac{\partial q^{(j)}}{\partial x^{(i)}} = \sum_{j=1}^{d} \frac{\partial z(t, q)}{\partial q^{(j)}} \frac{\partial x^{(k)}}{\partial x^{(i)}} = \sum_{k=1}^{d} \frac{\partial z(t, q)}{\partial x^{(i)}} \frac{\partial x^{(k)}}{\partial x^{(i)}} = \left( \Sigma_{1}^{2} \right)^{(i)}(q),
\]
for \( i = 1, \ldots, d \) and \( 0 \leq t \leq T \). This proves (4.27), completing the proof of the lemma.

\[ \square \]

**Corollary 4.23.** By (4.27) in the proof of the preceding lemma, we have

\[
\nabla_{q} U(t, q) = \mathcal{V}[s, \mu](t, q), \quad t \in (0, T), q \in \mathbb{T}^{d}. \tag{4.28}
\]

The dependence of \( U \) on the parameters \( s \) and \( \mu \), made clear by its definition, should not be forgotten. This corollary means that \( \mathcal{V}[s, \mu](t, \cdot) \) is the \( C^{1} \) gradient of a function. Thus,

\[
\nabla_{q} \mathcal{V}[s, \mu](t, q) \text{ is symmetric, for every } t \in (0, T), q \in \mathbb{T}^{d}.
\]
To conclude our statement about the MFG system, observe that
\[
v_t(q) = \partial_t \Sigma_1^1((\Sigma_1^1)^{-1}(q)) = \nabla_p H(q, V(t, q)) = \nabla_p H(q, \nabla_q U(t, q)).
\]
Hence, combining with Proposition 4.18 we have the following.

**Theorem 4.24.** (Existence of solution to the MFG system) Let \( \mu \in \mathcal{P}(\mathbb{T}^d) \), \( T \) be in accordance with Remark 4.8 and Proposition 4.3, \( 0 < s < T \), and let \( \sigma_t, v_t \) be as in (4.22), where \( (\Sigma_1^1[s, \mu], \Sigma_2^2[s, \mu]) \) is the unique solution to (4.3) with parameters \( s, \mu \). Then the pair \( (U, \sigma) \), where \( U \) is as in Lemma 4.22 is a classical solution to the mean-field game system (1.2-1.5) in the sense explained in Section 2.1.1.

Note that, by Proposition 4.21, the function \( U \) in the pair \( (U, \sigma) \) constructed above is in \( W^{2,2;\infty}((0, T) \times \mathbb{T}^d) \times AC^2(0, T; \mathcal{P}(\mathbb{T}^d)) \). The following is, in a sense, a consistency, or restricted uniqueness, complement to the latter theorem.

**Theorem 4.25.** (The case of \( W^{2,3;\infty}((0, T) \times \mathbb{T}^d) \times AC^2(0, T; \mathcal{P}(\mathbb{T}^d)) \) solutions to the MFG system) Let \( (\tilde{U}, \tilde{\sigma}) \) be a classical solution to the MFG system (1.2-1.5), in the sense explained in Section 2.1.1 such that \( U \in W^{2,3;\infty}((0, T) \times \mathbb{T}^d) \). Then \( (\tilde{U}, \tilde{\sigma}) = (U, \sigma) \), where \( (U, \sigma) \) is the pair constructed for Theorem 4.24.

**Proof.** Let \( (\tilde{U}, \tilde{\sigma}) \) be a solution to the system (1.2-1.5) with parameters \( s \in (0, T) \) and \( \mu \in \mathcal{P}(\mathbb{T}^d) \), according to the definition of Section 2.1.1 and suppose, moreover, that \( \tilde{U} \) is \( W^{3;\infty} \) in \( q \).

1. We will prove that the characteristics of the MFG system satisfied by \( (\tilde{U}, \tilde{\sigma}) \) must solve the Hamiltonian system (4.3). Set
\[
\tilde{v}_t(q) = \tilde{v}(t, q) := \nabla_p H(q, \nabla_q \tilde{U}(t, q)),
\]
0 \( \leq t \leq T \), \( q \in \mathbb{T}^d \). Since \( \tilde{U} \in W^{2,3;\infty}((0, T) \times \mathbb{T}^d) \) and \( H \in C^3 \), we have that \( \text{Lip}(\tilde{v}_t, K) + \sup_{q \in K} |\tilde{v}_t(q)| \) is bounded on \( [0, T] \), where \( K \) is any compact subset of \( \mathbb{R}^d \) and \( \text{Lip}(\tilde{v}_t, K) \) is the Lipschitz constant of \( \tilde{v}_t |_{K} \). By elementary ODE theory (see, e.g., [3, Lemma 8.1.4]), if \( q \in \mathbb{R}^d \), the ODE
\[
\tilde{\Sigma}_1^1(s, q) = q, \quad \frac{\partial}{\partial t} \tilde{\Sigma}_1^1(t, q) = \tilde{v}_t(\tilde{\Sigma}_1^1(t, q))
\]
has a unique maximal solution in a neighborhood \( I(q, s) \subset (0, T) \) of \( s \), but, since \( \tilde{\Sigma}_1^1(t, \cdot) \) is periodic, and therefore bounded for every \( t \in (0, T) \), then \( I(q, s) = (0, T) \). Clearly, the path \( t \mapsto \tilde{\Sigma}_1^1(t, \cdot) \# \mu \) solves the continuity equation with velocity \( \tilde{v}_t \). We can apply Proposition 8.1.7 of [3] to conclude that
\[
\tilde{\sigma}_t = \tilde{\Sigma}_1^1(t, \cdot) \# \mu,
\]
0 \( \leq t \leq T \). Let
\[
\tilde{V}(t, q) := \nabla_q \tilde{U}(t, q), \quad \tilde{\Sigma}_2^2(t, q) := \tilde{V}(t, \tilde{\Sigma}_1^1(t, q)),
\]
0 \( \leq t \leq T \), \( q \in \mathbb{T}^d \). Then
\[
\partial_t \tilde{\Sigma}_1^1(t, q) = \tilde{v}_t(\tilde{\Sigma}_1^1(t, q)) = \nabla_p H(\tilde{\Sigma}_1^1(t, q), \tilde{\Sigma}_2^2(t, q)),
\]
which is the first equation in (4.3). To obtain the second one, observe that
\[
\partial_t \tilde{\Sigma}_2^2(t, q) = \partial_t \tilde{V}(t, \tilde{\Sigma}_1^1(q, t)) + \nabla_q \tilde{V}(t, \tilde{\Sigma}_1^1(t, q)) \partial_t \tilde{\Sigma}_1^1(t, q) = \nabla_{tt}^2 \tilde{U}(t, \tilde{\Sigma}_1^1(t, q)) + \nabla_{qq}^2 \tilde{U}(t, \tilde{\Sigma}_1^1(t, q)) \partial_t \tilde{\Sigma}_1^1(t, q),
\]
24
while, differentiating the Hamilton-Jacobi equation with respect to \( q \) gives
\[
\nabla_q^2 \tilde{U}(t, q) + \nabla_q H(q, \nabla_q \tilde{U}(t, q)) + \nabla_p H(q, \nabla_q \tilde{U}(t, q)) \nabla_{qq} \tilde{U}(t, q) + \nabla_q F(q, \tilde{\sigma}_t) = 0,
\]
which, evaluating at \( \tilde{\Sigma}(t, q) \) in place of \( q \), and using (4.32), (4.33), serves to simplify the former equality to
\[
\partial_t \tilde{\Sigma}^2(t, q) = -\nabla_q H(\tilde{\Sigma}(t, q), \nabla_q \tilde{\Sigma}(t, q)) - \nabla_q F(\tilde{\Sigma}(t, q), \tilde{\sigma}_t),
\]
which is the second equation in (4.3). The condition \( \tilde{\Sigma}^2(0, q) = \nabla_q g(\tilde{\Sigma}^1(0, q), \tilde{\Sigma}^1(0, \cdot \mid \mu) \) follows readily from (4.32) and (1.4).

2. We prove that, for a possibly smaller \( T \), the solutions \( (\tilde{\Sigma}^1, \tilde{\Sigma}^2) \) to (1.3) from the previous paragraph belongs to \( \mathcal{M}_0(A_1, \theta A_2, \theta B, E, T) \), i.e., they satisfy the bounds (1.13). If \( T \) is small enough, since \( |\tilde{\Sigma}^2(0, q)| = |\nabla_q g(\tilde{\Sigma}^1(0, q), \tilde{\sigma}_0)| \leq \kappa \), continuity implies \( |\tilde{\Sigma}^2(t, q)| \leq \kappa + \varepsilon, 0 \leq t \leq T, q \in \mathbb{T}^d \) for an \( \varepsilon > 0 \) such that
\[
|\tilde{\Sigma}^2(t, q)| \leq \theta \frac{\kappa}{\theta} + \varepsilon \leq \theta \frac{1}{\theta} \max\{d, \kappa\} + \varepsilon = \theta c/\theta + \varepsilon \leq \theta B,
\]
because \( B \) in the proof of Lemma 4.4 was chosen as \( B > c/\theta \) (in these lines we are referring back to the proof of Lemma 4.4 in particular (4.5), (4.6) and the paragraph preceding those inequalities). This is the third line in (1.13). Since \( \|\tilde{\Sigma}^2\|_{\mathcal{M}_2} \leq \theta B \), we have \( |\nabla_q \tilde{\Sigma}^1(t, q)| = |\nabla_p H(\tilde{\Sigma}^1(t, q), \tilde{\Sigma}^2(t, q))| \leq \tilde{\ell}(B) \) (see Definition 4.1 “Data bounds \( \Gamma \”) for small enough \( T \) and all \( q \), and since \( A_1 \) in Lemma 4.4 was chosen to be larger than \( \tilde{\ell}(B) \), we obtain the bound for \( \|\nabla_q \tilde{\Sigma}^1\| \) in (1.13). The one for \( \|\nabla_p \tilde{\Sigma}^2\|_{\mathcal{M}_2} \) goes in a similar way, because \( A_2 \) in Lemma 4.4 was chosen larger than \( \tilde{\ell}(B)/\theta \). From (4.30), \( \tilde{\Sigma}^1(t, q) = q + \int_0^t \tau \tilde{\Sigma}^1(\tau, q)\right)\right)d\tau, \)
which makes it clear that, upon taking the gradient in \( q \), if \( T \) is small enough, the norm of \( \nabla_q \tilde{\Sigma}^1(t, q) \) will be only slightly larger than \( \sqrt{\theta} \), making it less than \( A_1 \), because of (1.3). The bound for \( \|\nabla^2_{qq} \tilde{\Sigma}^1\| \), due to \( \nabla_q^2 \tilde{\Sigma}^1(q) = 0 \), can actually be made arbitrarily small by choosing \( T \) small enough. To address \( \nabla_q^2 \tilde{\Sigma}^2 \) and \( \nabla^2_{qq} \tilde{\Sigma}^2 \), since
\[
\tilde{U}(t, q) = g(q, \tilde{\sigma}_0) + \int_0^t [H(q, \nabla_q \tilde{U}(\tau, q)) + F(q, \tilde{\sigma}_\tau)]d\tau,
\]
and \( \nabla_q \tilde{\Sigma}^2(t, q) = \nabla_q \tilde{U}(t, \tilde{\Sigma}^1(t, q)) \nabla_q \tilde{\Sigma}^1(t, q) \), the norm of \( \nabla_q \tilde{\Sigma}^2 \) is the product of a number slightly larger than \( \kappa \) and one slightly larger than \( \sqrt{\theta} \), for small times \( T \). But the constant \( E \) in the proof of Lemma 4.4 is larger than \( c = \max\{d, \kappa\} \), and \( A_2 > \frac{1}{\theta} c E(E + 1) \). This ensures that \( \|\nabla_q^2 \tilde{\Sigma}^1\| \leq \theta A_2 \). Next, given that
\[
\nabla^2_{qq} \tilde{\Sigma}^2 = \nabla^3_{qq} \tilde{U} \nabla_q \tilde{\Sigma}^1 \nabla_q \tilde{\Sigma}^1 + \nabla^2_{qq} \tilde{U} \nabla^2_{qq} \tilde{\Sigma}^1
\]
(because \( \tilde{U} \) is \( W^{3: \infty} \) in \( q \), and \( \nabla^2_{qq} \tilde{\Sigma}^1 \), as already mentioned, can be made as small as needed by reducing \( T \), the same argument shows that \( \|\nabla^2_{qq} \tilde{\Sigma}^2\| \) is also no greater than \( \theta A_2 \), since the norm of \( \nabla^3_{qq} \tilde{U} \nabla^2_{qq} \tilde{\Sigma}^1 \nabla_q \tilde{\Sigma}^1 \) is the product of a number slightly larger than \( \kappa \) and one slightly than \( \sqrt{\theta} \). Finally, \( \|\tilde{\Sigma}^2(0, q)\|, \|\nabla^2_{qq} \tilde{\Sigma}^2(0, q)\| \leq E \) follow from \( E \) having been picked larger than \( d \) (and, therefore, than \( \sqrt{\theta} \)), and taking \( T \) smaller if necessary.

Thus, the mapping \( (\Sigma^1, \Sigma^2) \) constructed in (4.30) and (4.32) from \((\tilde{U}, \tilde{\sigma})\) coincides with the unique solution \( (\Sigma^1[s, \mu], \Sigma^2[s, \mu]) \) of (4.3) in \( \mathcal{M}_0(A_1, \theta A_2, \theta B, E, T) \) during a possibly shorter interval \([0, T]\). Consequently, by (4.31), we further have that \( \tilde{\sigma} = \sigma \). Also, now that \( \tilde{V}(t, q) = \mathcal{V}(t, q), 0 \leq t \leq T, \) we get
\[
\tilde{\Sigma}^2(t, q) = \tilde{\Sigma}^2(0, q) - \int_0^t [H(q, \tilde{V}(\tau, q) - F(x, \tilde{\sigma}_\tau)]d\tau = \tilde{\Sigma}^2(0, q) - \int_0^t [H(x, \mathcal{V}(\tau, q)) - F(q, \sigma_\tau)]d\tau = \tilde{U}(t, q),
\]
for any \( t \in [0, T] \). Thus, \((\tilde{U}, \tilde{\sigma}) = (U, \sigma)\) on the possibly smaller interval \([0, T]\).
4.4 The full value function \( u(s, q, \mu) \)

In this section we begin our study of the dependence of our solution \( U \) to (1.2) on the parameter \( \mu \). First, we present a list of facts that will be used in this and following sections.

**Proposition 4.26.** Let \( 0 \leq s, t_0 \leq T, \mu \in \mathcal{P}(\mathbb{T}^d) \). Set

\[
\sigma_{t_0} = \Sigma^1_{t_0}[s, \mu] \# \mu.
\]

Then,

(i) For every \( 0 \leq t \leq T \):

\[
\Sigma^1_t[t_0, \sigma_{t_0}] \circ \Sigma^1_{t_0}[s, \mu] = \Sigma^1_t[s, \mu].
\] (4.34)

(ii) For every \( 0 \leq t \leq T \):

\[
\Sigma^1_t[t_0, \sigma_{t_0}] \circ \Sigma^1_{t_0}[s, \mu] = \Sigma^1_t[s, \mu];
\] (4.35)

\[
v_t[s, \mu] = v_t[t_0, \sigma_{t_0}],
\] (4.36)

\[
\Sigma^2_t[t_0, \sigma_{t_0}] \circ \Sigma^1_{t_0}[s, \mu] = \Sigma^2_t[s, \mu],
\] (4.37)

\[
\partial_s \Sigma^1_t[s, \mu] = -\nabla q \Sigma^2_t[s, \mu] v_s[t, \sigma_t].
\] (4.38)

(iii) If \( 0 \leq \tau, t \leq T \), then

\[
\Sigma^2_\tau[t, \sigma_t] \circ (\Sigma^1_\tau[t, \sigma_t])^{-1} = \Sigma^2_\tau[s, \mu] \circ \Sigma^1_\tau[s, \mu]^{-1} ;
\] (4.39)

\[
\nu_t[t_0, \sigma_{t_0}] = \nu_t[s, \mu].
\] (4.40)

**Proof.** (i) Let

\[
Q(t, q) = (\Sigma^1_t[s, \mu] \circ \Sigma^1_{t_0}[s, \mu]^{-1})(q),
\]

\[
P(t, q) = (\Sigma^2_t[s, \mu] \circ \Sigma^1_{t_0}[s, \mu]^{-1})(q),
\]

for \( 0 \leq t \leq T, q \in \mathbb{T}^d \). By differentiating, and noting that \( Q(0, \cdot) \# \sigma_{t_0} = \Sigma^1_0[s, \mu] \# \mu \), one verifies that \( Q \) and \( P \) defined this way satisfy the Hamiltonian ODEs (4.3) with

\[
s = t_0, \quad \mu = \sigma_{t_0}.
\]

Since solutions to (4.3) are unique, we conclude that

\[
(Q_t, P_t) = (\Sigma^1_t[t_0, \sigma_{t_0}], \Sigma^2_t[t_0, \sigma_{t_0}]),
\]

yielding (4.34).

(ii) Fact (4.35) follows readily from (i), by setting \( t = s \). For (4.36), see [25, p. 6593]. Formula (4.37) is just the second component of (4.34).

By (4.35), with \( t = s, t_0 = t \), we have

\[
id = \Sigma^1_t[s, \mu] \circ \Sigma^1_s[t, \sigma_t] = \Sigma^1_s[t, \mu](t, \Sigma^1_s[t, \sigma_t]).
\]

By Lemma 4.12 we can differentiate both sides with respect to \( s \):

\[
0 = \partial_s \Sigma^1[s, \mu](t, \Sigma^1_s[t, \sigma_t](q)) + \nabla q \Sigma^1[s, \mu](t, \Sigma^1_s[t, \sigma_t](q)) \partial_s \Sigma^1_s[t, \sigma_t](q), \quad q \in \mathbb{T}^d.
\]
Substituting $\Sigma^1_s[t, \sigma_t](q)$ for $q$, we get
\[ 0 = \partial_s \Sigma^1_s[t, \mu](t, q) + \nabla_q \Sigma^1_s[t, \mu](t, q) \partial_\tau \Sigma^1_s[t, \sigma_t](\Sigma^1_s[t, \sigma_t])^{-1} \]
\[ = \partial_s \Sigma^1_s[t, \mu](t, q) + \nabla_q \Sigma^1_s[t, \mu](t, q) v_0[t, \sigma_t], \]
which gives (4.38).

(iii) For (4.39), simply use (4.37) with $\tau$ in place of $t$ and $t$ in place of $t_0$:
\[ \Sigma^2_\tau[t, \sigma_t] \circ (\Sigma^1_\tau[t, \sigma_t])^{-1} = \Sigma^2_\tau[s, \mu] \circ \Sigma_\tau[s, \mu]^{-1} \circ \Sigma^1_\tau[t, \sigma_t]^{-1} \]
\[ = \Sigma^2_\tau \circ (\Sigma^1_\tau[t, \sigma_t] \circ \Sigma_\tau[s, \mu])^{-1} = \Sigma^2_\tau[s, \mu] \circ \Sigma^1_\tau[s, \mu]^{-1}. \]

To get the last formula, (4.37) gives $\Sigma^2_{t}[t_0, \sigma_{t_0}] = \Sigma^2_t[s, \mu] \circ \Sigma_{t_0}[s, \mu]^{-1}$, so we obtain
\[ \mathcal{V}_{t}[t_0, \sigma_{t_0}] = \Sigma^2_t[t_0, \sigma_{t_0}] \circ \Sigma^1_t[t_0, \sigma_{t_0}]^{-1} = \Sigma^2_t[s, \mu] \circ \Sigma_{t_0}[s, \mu]^{-1} \circ \Sigma^1_t[t_0, \sigma_{t_0}]^{-1} \]
\[ = \Sigma^2_t[s, \mu] \circ (\Sigma^1_t[t_0, \sigma_{t_0}] \circ \Sigma_{t_0}[s, \mu])^{-1} \]
\[ = \Sigma^2_t[s, \mu] \circ \Sigma^1_t[s, \mu] = \mathcal{V}_t[s, \mu]. \]

\[ \square \]

The following is the definition of the function that will turn out to be the solution to the ME. It is nothing but the representation formula for $U$ furnished by (1.2), including the parameter $\mu$ explicitly. Given $s \in [0, T]$, $q \in \mathbb{T}^d$, $\mu \in \mathcal{P}(\mathbb{T}^d)$, define
\[ u(s, q, \mu) = g(q, \Sigma^1_s[s, \mu](0, \cdot | \mu) - \int_0^s [H(q, \mathcal{V}[s, \mu](\tau, q)) + F(q, \Sigma^1_s[s, \mu](\tau, \cdot | \mu)]d\tau, \] (4.41)
and, as before, $\sigma_t = \Sigma^1_t[s, \mu] \# \mu$, $0 \leq t \leq T$. Note that (4.39) reads now as
\[ \mathcal{V}_{\tau}[t, \sigma_t] = \mathcal{V}_t[s, \mu]. \] (4.42)

Coupled with the fact that, by (4.34), $\Sigma^1_\tau[t, \sigma_t] \circ \Sigma^1_t[s, \mu] = \Sigma^1_\tau[s, \mu]$, $0 \leq \tau \leq T$, we have
\[ u(t, q, \sigma_t) = g(q, \Sigma^1_0[t_0, \sigma_{t_0}] \# \sigma_t) - \int_0^t [H(q, \mathcal{V}[t, \sigma_t](\tau, q)) + F(q, \Sigma^1_t[t, \sigma_t](\tau, \cdot | \mu))]d\tau \]
\[ = g(q, \sigma_0) - \int_0^t [H(q, \mathcal{V}[s, \mu](\tau, q)) + F(q, \Sigma^1_\tau[s, \mu] \# \mu)]d\tau. \] (4.43)

Since
\[ \mathcal{V}[s, \mu](t, \Sigma^1_s[s, \mu](t, q)) = \Sigma^2[s, \mu](t, q), \]
and $\Sigma^2$ satisfies the second of the Hamiltonian ODEs (1.3), it follows, by taking the total time derivative of $\mathcal{V}[s, \mu](t, \Sigma^1_s[s, \mu](t, q))$, and then changing variable from $\Sigma^1(t, q)$ to $q$, that $\mathcal{V}(t, q) = \mathcal{V}[s, \mu](t, q)$ satisfies the equation
\[ \partial_t \mathcal{V}(t, q) + \nabla_q \mathcal{V}(t, q) \nabla_\mu H(q, \mathcal{V}(t, q)) = - \nabla_q H(q, \mathcal{V}(t, q)) - \nabla_q F(q, \Sigma^1_\# \mu), \] (4.44)
\[ \mathcal{V}(0, q) = \nabla_q g(q, \sigma_0). \] (4.45)
If we differentiate \( u(t, q, \sigma_t) \) with respect to \( q \) in (4.43), and use (4.44) and (4.45), we get

\[
\nabla_q u(t, q, \sigma_t) = \nabla_q g(q, \sigma_t) + \int_0^t \left[ \partial_t \mathcal{V}[s, \mu](\tau, q) + \nabla_q \mathcal{V}[s, \mu](\tau, q) \nabla_p H(q, \mathcal{V}[s, \mu](\tau, q)) \right. \\
\left. - \nabla_p H(q, \mathcal{V}[s, \mu](\tau, q)) \nabla_q \mathcal{V}[s, \mu](\tau, q) \right] \, d\tau.
\]

Since \( \nabla_q \mathcal{V} \) is a symmetric matrix for \( \tau \in [0, T] \), only the term \( \partial_t \mathcal{V} \) survives in the integral. Hence

\[
\nabla_q u(t, q, \sigma_t) = \mathcal{V}[s, \mu](t, q), \quad 0 \leq t \leq T, \quad q \in \mathbb{T}^d.
\]

Differentiating now with respect to \( t \) in (4.43), and substituting (4.46) into it, we conclude that

\[
\partial_t (u(t, q, \sigma_t)) + H(q, \nabla_q u(t, q, \sigma_t)) + F(q, \sigma_t) = 0.
\]

Thus, we have shown:

**Lemma 4.27.** For \( s \in [0, T] \), \( \mu \in \mathcal{P}(\mathbb{T}^d) \), we have:

(i) For any \((t, q) \in [0, T] \times \mathbb{T}^d, \)

\[
u(t, q) = g(\cdot, \mu), \quad \nabla_q u(t, q, \sigma_t) = \mathcal{V}[s, \mu](t, q).
\]

(ii) The function \( t \mapsto u(t, q, \sigma_t) \) is continuously differentiable and

\[
\partial_t (u(t, q, \sigma_t)) + H(q, \nabla_q u(t, q, \sigma_t)) + F(q, \sigma_t) = 0, \quad (t, q) \in [0, T] \times \mathbb{T}^d.
\]

### 5 Regularity of \( \Sigma[s, \cdot](t, q) \)

From here on, suppose that we are given \( U^0, U^1, \phi \) in \( C^3(\mathbb{T}^d) \), with the latter two being even, and

\[
\|\phi\|_{C^3(\mathbb{T}^d)}, 2\|U^0\|_{C^3(\mathbb{T}^d)}, 2\|U^1\|_{C^3(\mathbb{T}^d)} \leq \kappa,
\]

and that for any \( q \in \mathbb{T}^d \), any \( \mu \in \mathcal{P}(\mathbb{T}^d) \),

\[
F(q, \mu) = \phi \ast \mu(q), \quad \text{so that} \quad \nabla_q F(\cdot, \mu) = \nabla \phi \ast \mu(\cdot);
\]

and

\[
g(q, \mu) = U^0(q) + U^1 \ast \mu(q), \quad \text{so that} \quad \nabla_q g(q, \mu) = \nabla U^0(q) + \nabla U^1 \ast \mu(q).
\]

#### 5.1 The discretized map \( M \)

For the remainder of the paper, let

\[
\theta > \max\{1, 5\sqrt{2}\kappa\},
\]

and \( A_1, A_2, B, E, T, D \) be as in Proposition 4.13 and Corollary 4.7 with \( T \) being subject to Remark 4.8. The functions \( \Sigma[s, \mu], \mu \in \mathcal{P}(\mathbb{T}^d) \), as before, denote the fixed points of the operators \( \mathfrak{m}^s,\mu \), while \( \Sigma[s, \mu] = (\Sigma^1[s, \mu], \Sigma^2[s, \mu]) \) are the fixed points of the operators \( \bar{\mathfrak{m}}^s,\mu \); recall (4.12) from the proof of Corollary 4.7. The master map was defined in Definition 4.14 as \( M(t, s, q, \mu) = (t, s, \Sigma^1[s, \mu](t, q), \mu) \). Let \( M = (M_1, M_2) \) be the map \( \Sigma = (\Sigma^1, \Sigma^2) \) restricted to average of Dirac masses. Namely:
Definition 5.1. For any $s, t \in [0, T]$, $q \in \mathbb{T}^d$, $x \in (\mathbb{T}^d)^n$, let

\[ M = (M_1, M_2) : [0, T] \times [0, T] \times \mathbb{T}^d \times (\mathbb{T}^d)^n \rightarrow \mathbb{T}^d \times \mathbb{R}^d \]

\[ (t, s, q, x) \rightarrow \Sigma[s, \mu^x](t, q) = (\Sigma^1[s, \mu^x](t, q), \Sigma^2[s, \mu^x](t, q)). \]

Note. The domain of the mapping $M$ depends on $n \in \mathbb{Z}^+$. \hfill \Box

This section deals with the regularity of $\Sigma$ in the measure variable $\mu$. See Section 3 or the first paragraph of Section 4 for the general plan. Several preliminary smoothness estimates are set down first, before the main result is presented in Section 5. The proofs of said estimates are rather lengthy, but do not involve more than repeated applications of the chain rule and the bounds from the fixed point theory of Section 4.

Definition 5.2. (i) For $n \in \mathbb{N}$, let $\bar{M}^k$, $k = 0, 1, \ldots$ be the sequence of $\mathbb{T}^d \times \mathbb{R}^d$-valued functions on $[0, T] \times [0, T] \times \mathbb{T}^d \times (\mathbb{T}^d)^n$ defined by

\[ \bar{M}^0 = (q, 0), \quad \bar{M}^{k+1} = \bar{m}^{s, \mu^x}(\bar{M}^k(\cdot, s, \cdot, x)), \]

where $\mu^x = \frac{1}{n} \sum_{j=1}^{n} \delta_{x_j}, x = (x_1, \ldots, x_n) \in (\mathbb{T}^d)^n$. That is,

\[ \bar{M}^{k+1}_1(t, s, q, x) = q + \int_s^t \nabla_p H(\bar{M}^k_1(\tau, s, q, x), \theta \bar{M}^k_2(\tau, s, q, x))d\tau \] \hfill (5.4)

and

\[ \bar{M}^{k+1}_2(t, s, q, x) = \frac{1}{\theta} [\nabla U^0_0(\bar{M}^k_1(0, s, q, x)) + \frac{1}{n} \sum_{j=1}^{n} \nabla U^1(\bar{M}^k_1(0, s, q, x) - \bar{M}^k_1(0, s, x_j, x))] \] \hfill (5.5)

\[ - \frac{1}{\theta} \int_0^t \left[ \nabla q H(\bar{M}^k_1(\tau, s, q, x), \theta \bar{M}^k_2(\tau, s, q, x)) ight. \]

\[ + \left. \frac{1}{n} \sum_{j=1}^{n} \nabla \phi(\bar{M}^k_1(\tau, s, q, x) - \bar{M}^k_1(\tau, s, x_j, x)) \right] d\tau. \] \hfill (5.6)

(ii) Let $\bar{M}^k, k = 0, 1, \ldots$ be the sequence of $\mathbb{T}^d \times \mathbb{R}^d$-valued functions on $[0, T] \times [0, T] \times \mathbb{T}^d \times (\mathbb{T}^d)^n$ defined by

\[ M^0 = (q, 0), \quad M^{k+1} = m^{s, \mu^x}(M^k(\cdot, s, \cdot, x)). \]

Remark 5.3. It follows that, for every $k = 0, 1, \ldots$

\[ M^k_1(t, s, q, x) = \bar{M}^k_1(t, s, q, x); \quad M^k_2(t, s, q, x) = \theta \bar{M}^k_2(t, s, q, x), \]

\[ t, s \in [0, T], q \in \mathbb{T}^d, x \in (\mathbb{T}^d)^n. \] \hfill \Box

We begin with the $x_j$-derivative of the $(k+1)$-th iteration $(j = 1, \ldots, n)$:

\[ \nabla_j \bar{M}^{k+1}_1(t, s, q, x) = \int_s^t \left[ \nabla^2_{pp} H(\bar{M}^k_1(\tau, s, q, x), \theta \bar{M}^k_2(\tau, s, q, x)) \nabla_j \bar{M}^k_1(\tau, s, q, x) \right. \]

\[ + \left. \nabla^2_{pp} H(\bar{M}^k_1(\tau, s, q, x), \theta \bar{M}^k_2(\tau, s, q, x)) \theta \nabla_j \bar{M}^k_2(\tau, s, q, x) \right] \nabla_j \bar{M}^k_1(\tau, s, q, x) \] \hfill (5.7)

\[ + \nabla^2_{pp} H(\bar{M}^k_1(\tau, s, q, x), \theta \bar{M}^k_2(\tau, s, q, x)) \theta \nabla_j \bar{M}^k_2(\tau, s, q, x) \] \hfill (5.8)
\[ \nabla_{x_j} M^{k+1}_2(t, s, q, x) = \frac{1}{\theta} \nabla^2 U^0(\bar{M}^{k}_1(0, s, q, x)) \nabla_{x_j} \bar{M}^{k}_1(0, s, q, x) \]
\[ - \frac{1}{\theta n} \nabla^2 U^1(\bar{M}^{k}_1(0, s, q, x) - \bar{M}^{k}_1(0, s, x_j, x)) \nabla_q \bar{M}^{k}_1(0, s, x_j, x) \] 
\[ + \frac{1}{\theta n} \sum_{i=1}^n \nabla^2 U^1(\bar{M}^{k}_1(0, s, q, x) - \bar{M}^{k}_1(0, s, x_i, x)) \Delta^s_j(0, q, x_i, x) \]
\[ - \frac{1}{\theta} \int_0^t \left[ \nabla_{q,q} H(\bar{M}^{k}_1(\tau, s, q, x), \theta \bar{M}^{k}_2(\tau, s, q, x)) \nabla_{x_j} \bar{M}^{k}_1(\tau, s, q, x) \right. 
\[ + \nabla^2_{p,q} H(\bar{M}^{k}_1(\tau, s, q, x), \theta \bar{M}^{k}_2(\tau, s, q, x)) \theta \nabla_{x_j} \bar{M}^{k}_2(\tau, s, q, x) \]
\[ - \frac{1}{n} \nabla^2 \phi(\bar{M}^{k}_1(\tau, s, q, x) - \bar{M}^{k}_1(\tau, s, x_j, x)) \nabla_q \bar{M}^{k}_1(\tau, s, x_j, x) \]
\[ + \frac{1}{n} \sum_{i=1}^n \nabla^2 \phi(\bar{M}^{k}_1(\tau, s, q, x) - \bar{M}^{k}_1(\tau, s, x_i, x)) \Delta^s_j(\tau, q, x_i, x) \] 
\[ d\tau, \]
where
\[ \Delta^s_j(\tau, q, x_i, x) = \nabla_{x_j} \bar{M}^{k}_1(\tau, s, q, x) - \nabla_{x_j} \bar{M}^{k}_1(\tau, s, x_i, x). \]

For the next lemma, we remind the reader of Remark 4.13

**Lemma 5.4.** Fix \( n \in \mathbb{Z}^+ \). Using the terminology of Definition 5.2, the following hold:

(i) For each \( k = 0, 1, \ldots \) and each \( x \in (\mathbb{T}^d)^n, M(\cdot, \cdot, x) \in \mathcal{M}^*_0(A_1, \theta A_2, \theta B, E, T). \)

(ii) There is a constant \( C > 0 \), independent of \( k \) such that for any \( j = 1, \ldots, n \):

\[ \| \nabla_{x_j} M^k \|_\infty \leq \frac{C}{n}. \]  

**Proof.** (i) Clearly, \( \bar{M}^0(\cdot, \cdot, x) \in \mathcal{M}^*_0(A_1, A_2, B, E, T) \), and by Lemma 4.12 each \( \bar{M}^k(\cdot, \cdot, x) \in \mathcal{M}^*_0(A_1, A_2, B, E, T) \). By Remark 4.13 each \( M^k(\cdot, \cdot, x) \in \mathcal{M}^*_0(A_1, \theta A_2, \theta B, E, T) \).

(ii) Using the data bounds, we obtain, after simplifying:

\[ \| \nabla_{x_j} \bar{M}^{k+1}_1 \|_\infty \leq T \tilde{h}(B)[\| \nabla_{x_j} \bar{M}^k_1 \|_\infty + \theta \| \nabla_{x_j} \bar{M}^k_2 \|_\infty], \]
\[ \| \nabla_{x_j} \bar{M}^{k+1}_2 \|_\infty \leq \frac{1}{n} \left( \frac{\kappa}{\theta} A_1 + \frac{T\kappa}{\theta} A_1 \right) + \| \nabla_{x_j} \bar{M}^k_1 \|_\infty \left( \frac{3\kappa}{\theta} + \frac{T}{\theta} \tilde{h}(B)(1 + \theta) + \frac{2\kappa T}{\theta} \right). \]

Combining the two, we get
\[ \| \nabla_{x_j} \bar{M}^{k+1} \|_\infty \leq \frac{\sqrt{2} \kappa}{n} A_1(1 + \theta) + \frac{\sqrt{2}}{n} \| \nabla_{x_j} \bar{M}^k_1 \|_\infty \left( \frac{3\kappa}{\theta} + T(1 + \theta)\tilde{h}(B)(1 + \frac{1}{\theta}) + \frac{2\kappa T}{\theta} \right). \]

By (5.3), we can invoke Remark 4.8 to obtain that the latter is an expression of the form
\[ \| \nabla_{x_j} \bar{M}^{k+1} \|_\infty \leq \frac{a}{n} + b \| \nabla_{x_j} \bar{M}^k_1 \|_\infty \]

30
with positive constants \( a, b \) in which \( b < 1 \). This holds for every \( k = 0, 1, \ldots \) Applying this inequality recursively (see [25, Remark 8.1]), we find a constant \( C > 0 \) such that \( \| \nabla_x M^k \|_\infty \leq C/n \), for every \( k \in \mathbb{Z}^+ \). Thus,

\[
\left( \| \nabla_x M^1_1 \|_\infty^2 + \| \nabla_x M^2_1 \|_\infty^2 \right)^{1/2} = \left( \| \nabla_x M^1_1 \|_\infty^2 + \| \nabla_x M^2_1 \|_\infty^2 \right)^{1/2} \leq \frac{C}{n}.
\]

(5.11)

Multiplying by \( \theta \) on both sides we get, since \( \theta > 1 \), that \( \left( \| \nabla_x M^k_1 \|_\infty^2 + \| \nabla_x M^k_2 \|_\infty^2 \right)^{1/2} \leq \frac{\theta C}{n} \), which is (5.10) for a larger constant \( C \).

5.1.1 Regularity of \( M \) in \( x \) and \( q \)

**Corollary 5.5.** The sequence \( \{M^k\}_1^\infty \) of Definition 5.2(ii) converges uniformly to the function \( M \) of Definition 5.4 with \( M(\cdot, \cdot, x) \in \mathcal{M}_{n,D}(A_1, \theta A_2, \theta B, E, T) \) for every \( x \in (\mathbb{T}^d)^n \) and \( M(t, s, q, x) = M(t, s, q, \bar{x}) \) whenever \( \bar{x} \) is a permutation of \( x \). Moreover, there is a constant \( C > 0 \) such that

\[
\| \nabla^2_{q_j x_j} M \|_\infty \leq \frac{C}{n}, \quad j = 1, \ldots, n,
\]

(5.12)

\[
\| \nabla^2_{x_i x_j} M \|_\infty \leq \frac{C}{n}, \quad i \neq j, \quad i, j \in \{1, \ldots, n\},
\]

(5.13)

\[
\| \nabla^2_{x_i x_j} M \|_\infty \leq \frac{C}{n}, \quad j = 1, \ldots, n.
\]

(5.14)

**Note.** Since on \( W^{2,\infty}(\mathbb{T}^d \times \mathbb{T}^d) \) the mixed partial derivatives \( \nabla^2_{x_i x_j} \) and \( \nabla^2_{q_i q_j} \) are equal, estimate (5.12) holds for \( \nabla^2_{q_j q_j} M \) too.

**Proof.** The proof of Lemma 4.12 shows that for every \( x \in (\mathbb{T}^d)^n \), \( M^k(\cdot, \cdot, x) \) converges uniformly to \( \Sigma[\cdot, \mu^x](\cdot, \cdot) \). Formula (5.10) means that the sequence \( M^k \) is equicontinuous, and uniformly bounded, on \( [0, T] \times [0, T] \times \mathbb{T}^d \times (\mathbb{T}^d)^n \). It follows, by Ascoli’s theorem, that the convergence of \( M^k \) to \( M \) is uniform, and \( M \) also satisfies (5.10); to be more precise:

\[
\| \nabla_x M \|_\infty \leq \frac{C}{n}.
\]

(5.15)

If \( \bar{x} \) is a permutation of \( x \) then \( \mu^x = \mu^{\bar{x}} \) and so \( M(t, s, q, x) = M(t, s, q, \bar{x}) \). The three estimates above will be true of the limit function \( M \) if they hold for every \( M^k \). Like before, we are unable to obtain them for \( M^k \) directly due to the size of the constant \( \kappa \), so we again do it for \( M^k \) first. We differentiate (5.8) and (5.9) with respect to \( q \), dropping arguments of \( M \) in the writing, where possible, to save space. We get

\[
\nabla^2_{q_j x_j} M_{1}^{k+1} = \int_s^t \left[ \begin{array}{c}
\nabla^3_{qpp} H(M_1^k, \theta M_2^k) \nabla_q M_1^k + \nabla^3_{ppp} H(M_1^k, \theta M_2^k) \theta \nabla_q M_2^k \\
\theta \nabla_x \nabla^2_{q_j x_j} M_1^k
\end{array} \right] d\tau,
\]

\[
+ \nabla^3_{qpp} H(M_1^k, \theta M_2^k) \nabla^2_{q_j x_j} M_1^k \\
+ \nabla^3_{ppp} H(M_1^k, \theta M_2^k) \theta \nabla_q M_2^k \\
+ \nabla^2_{pp} H(M_1^k, \theta M_2^k) \theta \nabla^2_{q_j x_j} M_2^k \right] d\tau,
\]
\[ \nabla^2_{q_x} M_{2}^{k+1}(q) = \frac{1}{\theta} \nabla^3 U^0(M^k) \nabla_q M^k \nabla_q M^k + \frac{1}{\theta} \nabla^2 U^0(M^k) \nabla^2_{q_x} M^k \\
- \frac{1}{\theta n} \nabla^3 U^1(M^k(q) - M^k_x(q)) \nabla_q M^k \\
+ \frac{1}{\theta n} \sum_{i=1}^{n} \left( \nabla^3 U(M^k(q) - M^k_x(q)) \nabla_q M^k \Delta^j_x(q_x) + \frac{1}{\theta} \nabla^2 U^1(M^k(q) - M^k_x(q)) \nabla^2_{q_x} M^k(q) \right) \\
- \frac{1}{\theta} \int_0^t \left[ \left( \nabla^3 \phi(M^k(q) - M^k_x(q)) \nabla_q M^k \Delta^j_x(q_x) \right) \nabla^2_{q_x} M^k(q) \right] \, dr. \]

Using (5.10) and the bounds on the data, we have:

\[ \| \nabla^2_{q_x} M^k_1 \|_\infty \leq T[(\tilde{h}(B) A_1 + \tilde{h}(B) \theta A_2) \frac{C}{n} (1 + \theta) + \tilde{h}(B) (\| \nabla^2_{q_x} M^k_1 \|_\infty + \theta \| \nabla^2_{q_x} M^k_2 \|_\infty )], \]

\[ \| \nabla^2_{q_x} M^k_2 \|_\infty \leq \frac{\kappa}{\theta} A_1 \frac{C}{n} + \frac{\kappa}{\theta} \| \nabla^2_{q_x} M^k \|_\infty + \frac{1}{n \theta} \kappa A_1 + \frac{2}{\theta} A_1 \frac{C}{n} + \frac{\kappa}{\theta} \| \nabla^2_{q_x} M^k \|_\infty \\
+ \frac{T}{\theta} \left[ (\tilde{h}(B) A_1 + \tilde{h}(B) \theta A_2) \frac{C}{n} (1 + \theta) + \tilde{h}(B) (\| \nabla^2_{q_x} M^k \|_\infty + \theta \| \nabla^2_{q_x} M^k \|_\infty ) \\
+ \frac{\kappa}{\theta} \| \nabla^2_{q_x} M^k \|_\infty \right]. \]

Simplifying and combining gives

\[ \| \nabla^2_{q_x} \tilde{M}^{k+1} \|_\infty \leq \frac{1}{n} [\sqrt{2} T C \tilde{h}(B) (A_1 + \theta A_2) (1 + \theta) (1 + \frac{1}{\theta}) + \sqrt{2} \frac{\kappa A_1}{\theta} (1 + 3C + \frac{T}{\theta} (1 + 2C)) ] \\
+ \sqrt{2} \| \nabla^2_{q_x} M^k \|_\infty [T \tilde{h}(B) \left( 1 + \frac{1 + \theta}{\theta} \right) + \frac{2 \sqrt{2}}{\theta} \kappa + \frac{T}{\theta^2} \kappa], \]

which is an inequality of the form

\[ \| \nabla^2_{q_x} \tilde{M}^{k+1} \|_\infty \leq \frac{a}{n} + b \| \nabla^2_{q_x} M^k \|_\infty \]

for constants \( a, b \) with \( b < 1 \), because \( \theta > 4 \) and \( T \) is small. By induction again, increasing \( C \) and switching back to \( M^k \) in similar fashion to (5.11), we obtain (5.12) in the limit as \( k \to \infty \).

We do not present here the full calculations that lead to (5.13) and (5.14), but describe what goes on in them.

Case \( i \neq j \): In the estimation, when \( i \neq j \), \( \nabla^2_{x,x_j} \) of the \( (k+1) \)th iteration’s first component, \( \tilde{M}_1^{k+1} \), has a common factor \( T \) outside (from the integral), data bounds\(^7\) multiplied by the norm of the previous
Thus, we get

\[ \theta \]

Thanks to the presence of \( \phi \), except the latter will be multiplied by \( T \) in the estimation, so we can focus our attention on the terms coming from \( g \). For \( U^1 \), we have data bounds multiplied by \( \frac{C}{n^2} \), and by the norm of the previous iteration’s derivative in another summand. The segment involving \( U^1 \) is more delicate, so we display it:

\[
- \frac{1}{\theta n} \nabla^3 U^1(\bar{M}_1^k(q) - \bar{M}_1^k(x_i))(\nabla_{x_i} \bar{M}_1^k(q) - \nabla_{x_i} \bar{M}_1^k(x_i)) \nabla_q \bar{M}_1^k \nabla^2 U^1(\bar{M}_1^k(q) - \bar{M}_1^k(x_j)) \nabla^2_{x_i q} \bar{M}_1^k \\
+ \frac{1}{\theta n} \sum_{l \neq i} \nabla^3 U^1(\bar{M}_1^k(q) - \bar{M}_1^k(x_i))(\nabla_{x_i} \bar{M}_1^k(q) - \nabla_{x_i} \bar{M}_1^k(x_i)) \Delta_\phi^2(x_i) \\
+ \frac{1}{\theta n} \nabla^3 U(\bar{M}_1^k(q) - \bar{M}_1^k(x_i))(\nabla_{x_i} \bar{M}_1^k(q) - \nabla_{x_i} \bar{M}_1^k(x_i) - \nabla_{x_i} \bar{M}_1^k(x_i)) \Delta_\phi^k(x_i) \\
+ \frac{1}{\theta n} \nabla^2 U^1(\bar{M}_1^k(q) - \bar{M}_1^k(x_i))(\nabla^2_{x_i x_j} \bar{M}_1^k(q) - \nabla^2_{x_i x_j} \bar{M}_1^k(x_i) - \nabla^2_{x_i x_j} \bar{M}_1^k(x_i)) ;
\]

it contributes:

\[
\frac{1}{\theta n} \kappa \left( \frac{C}{\theta n} + \frac{C}{\theta n} \right) A_1 + \frac{1}{\theta n} \kappa \left( \frac{C}{\theta n} + \frac{n - 1}{\theta n} \kappa \frac{2C}{\theta n} \frac{2C}{\theta n} + \frac{1}{\theta n} \kappa \left( \frac{2C}{\theta n} + A_1 \right) \frac{2C}{\theta n} \\
+ \frac{n - 1}{\theta n} 2\kappa \| \nabla^2_{x_i x_j} \bar{M}^k \|_\infty + \frac{1}{\theta n} \kappa \left( 2 \| \nabla^2_{x_i x_j} \bar{M}^k \|_\infty + \frac{C}{\theta n} \right) \right)
\]

Thus, we get

\[
\| \nabla_{x_i x_j} \bar{M}_1^{k+1} \|_\infty \leq T \left[ \frac{C}{n^2} (data \ bounds) + \| \bar{M}^k \|_\infty (data \ bounds) \right], \\
\| \nabla_{x_i x_j} \bar{M}_2^{k+1} \|_\infty \leq \frac{C}{n^2} (data \ bounds) + \| \bar{M}^k \|_\infty (data \ bounds) \\
+ T \left[ \frac{C}{n^2} (data \ bounds) + \| \bar{M}^k \|_\infty (data \ bounds) \\
+ \frac{C}{n^2} (data \ bounds) + \| \bar{M}^k \|_\infty (data \ bounds) \right].
\]

Thanks to the presence of \( \theta \) in the denominators of the expressions *data bounds* and (5.13), we obtain

\[
\| \nabla^2_{x_i x_j} \bar{M}^k \|_\infty \leq \frac{a}{n^2} + b \| \nabla^2_{x_i x_j} \bar{M}^k \|_\infty
\]

for constants \( a, b \) with \( b < 1 \), holding for all \( k \in \mathbb{Z}^+ \). From this (5.13) follows.

Case \( i = j \): The same remarks are valid for the estimation of \( \| \nabla^2_{x_i x_j} \bar{M}^{k+1} \| \). We only display the
The third and sixth lines contribute terms of order $1/n$. This time, though, $1/n^2$ does not factor out, only $1/n$ does, leading to

$$
\| \nabla^2_{x,j} \tilde{M}^{k+1} \|_{\infty} \leq \frac{a}{n} + b \| \nabla^2_{x,j} \tilde{M}^k \|_{\infty}
$$

with $b < 1$, for all $k \in \mathbb{Z}^+$. Therefore (5.14) holds.

### 5.1.2 Regularity of $M$ in $x$ and $s$

Identity (4.38) depends on fact (4.35), which holds for $\Sigma^1$, not for $\Sigma^2$. In order to obtain an estimate of $\nabla^2_{x,j} M_1$, this forces us to work with the sequence $\{ \tilde{M}^k \}_{k=1}^{\infty}$ once more to derive a bound for $\nabla^2_{x,j} \tilde{M}$. Going through the same process once more, this time with the operator $\nabla^2_{x,j}$, ends up in

$$
\| \nabla^2_{x,j} \tilde{M}^{k+1} \|_{\infty} \leq a/n + b \| \nabla^2_{x,j} \tilde{M}^k \|_{\infty},
$$

with $b < 1$, from which the estimate

$$
\| \nabla^2_{x,j} M \|_{\infty} \leq \frac{C}{n}
$$

(5.16) follows.

### 5.1.3 Regularity of $M$ in $x$ and $t$ and first order Taylor estimate

Our next step is to get an appropriate bound on the remainder of a first-order Taylor approximation of $M(\cdot, s, \cdot, \cdot)$ around $(t, s, q, x)$. Since

$$
\partial_t M_1(t, s, q, x) = \nabla_p H(M_1(t, s, q, x), M_2(t, s, q, x)),
$$

$$
\partial_t M_2(t, s, q, x) = \nabla_q H(M_1(t, s, q, x), M_2(t, s, q, x)) + \frac{1}{n} \sum_{j=1}^{n} \nabla \phi(M_1(t, s, q, x) - M_1(t, s, x, j)),
$$

34
The constant \(C\) is a constant, thus, there is a larger constant, still denoted by \(C\).

Therefore, bringing in the estimates obtained in the foregoing paragraphs, we get

\[
|M' - M - \partial_t M(t' - t) - \nabla_q M \cdot (q' - q) - \nabla_x M \cdot (x' - x)| \leq C(|t' - t|^2 + |q - q'|^2 + |x - x'|^2).
\]

The constant \(C\) does not depend on \(n\).  

**Note.** The norm in the left-hand side of the latter inequality is the Euclidean norm on \(\mathbb{R}^d\). ∎

**Proof.** Let \(i \in \{1, 2\}, j \in \{1, \ldots, d\}\). Denoting \(t' - t = \Delta t, q' - q = \Delta q, x_i'(x_i - x_i) = \Delta x_i, \text{and } |\Delta x| = (\sum_{l=1}^n |\Delta x_i|^2)^{1/2}\), the mean-value theorem implies that

\[
|\mathcal{M}_i^{(j)} - \mathcal{M}_i^{(j)}(t' - t) - \nabla_q \mathcal{M}_i^{(j)}(q' - q) - \nabla_x \mathcal{M}_i^{(j)}(x' - x)|
\leq \|
\mathcal{M}_i^{(j)}
\|_{\infty} |\Delta t|^2 + 2 \|
\nabla_q \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta t| |\Delta q| + 2 \sum_{l=1}^n \|
\nabla_{x_l} \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta t| |\Delta x_l| + \|
\nabla_{qq} \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta q|^2
\]

\[
+ 2 \sum_{l=1}^n \|
\nabla_{x_l} \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta x_l|
+ \sum_{l,m=1}^n \|
\nabla_{x_l x_m} \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta x_l| |\Delta x_m|
+ \sum_{l=1}^n \|
\nabla_{x_l} \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta x_l|^2.
\]

Therefore, bringing in the estimates obtained in the foregoing paragraphs, we get

\[
|M' - M - \partial_t M(t' - t) - \nabla_q M \cdot (q' - q) - \nabla_x M \cdot (x' - x)|
\leq \|
\mathcal{M}_i^{(j)}
\|_{\infty} |\Delta t|^2 + 2 \|
\nabla_q \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta t| |\Delta q| + 2 \sum_{l=1}^n \|
\nabla_{x_l} \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta t| |\Delta x_l| + \|
\nabla_{qq} \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta q|^2
\]

\[
+ 2 \sum_{l=1}^n \|
\nabla_{x_l} \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta x_l|
+ \sum_{l,m=1}^n \|
\nabla_{x_l x_m} \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta x_l| |\Delta x_m|
+ \sum_{l=1}^n \|
\nabla_{x_l} \mathcal{M}_i^{(j)}
\|_{\infty} |\Delta x_l|^2
\]

\[
\leq 2 \sqrt{2}(A_1 + A_2)(\kappa + h(\theta B)) |\Delta t|^2 + 2 \Delta t |\Delta q| + 4 \sqrt{2} \sqrt{n} \frac{C}{n} (2h(\theta B) + \kappa) |\Delta x||\Delta t|
\]

\[
+ \sqrt{2}(A_1 + A_2)|\Delta q|^2 + 2n \frac{C}{n} \sqrt{2}|\Delta x||\Delta q| + n(n-1) \frac{C}{n^2} \sqrt{2}|\Delta x|^2 + n \frac{C}{n} |\Delta x|^2.
\]

Thus, there is a larger constant, still denoted by \(C\), and not depending on \(n\), such that inequality in the corollary’s statement holds. ∎

Given \(x, x' \in (\mathbb{T}^d)^n\), we can reorder and shift the coordinates of \(x' = (x'_1, \ldots, x'_n)\) so that \(|x - x'|^2 = \mathcal{W}^2(\mu^x, \mu^{x'})\). Thus, the inequality of Corollary 5.6 reads

\[
|M' - M - \partial_t M(t' - t) - \nabla_q M \cdot (q' - q) - \nabla_x M \cdot (x' - x)|
\leq C(|t' - t|^2 + |q - q'|^2 + \mathcal{W}^2(\mu^x, \mu^{x'})).
\]
5.2 The discretized map \( N (= M^{-1}) \)

Let us define

\[
N : [0, T] \times [0, T] \times \mathbb{T}^d \times (\mathbb{T}^d)^n \longrightarrow \mathbb{T}^d \\
(t, s, q, x) \longmapsto X[s, \mu^x](t, q); 
\]

(5.19)

recall that \( X[s, \mu](t, \cdot) \) is the inverse of \( \Sigma^1[s, \mu](t, \cdot) \). The function \( N \) takes values in \( \mathbb{T}^d \), so it has only one component, unlike \( M = (M_1, M_2) \). Thus, \( M_1 \) and \( N \) are related by

\[
M^1(t, s, N(t, s, q, x), x) = q, \quad t, s \in [0, T], q \in \mathbb{T}^d, x \in (\mathbb{T}^d)^n. 
\]

Let us define also

\[
\mathfrak{N} : [0, T] \times [0, T] \times \mathbb{T}^d \times (\mathbb{T}^d)^n \longrightarrow [0, T] \times [0, T] \times \mathbb{T}^d \times (\mathbb{T}^d)^n \\
(t, s, q, x) \longmapsto (t, s, M_1(t, s, q, x), x), 
\]

(5.20)

where the superindex \( n \) now makes it explicit that \( n \) is the number of particles. The map \( \mathfrak{N} \) is a "discretized" version of the master map \( \mathfrak{M} \) of Definition 4.14. By Lemma 4.16 \( \mathfrak{N} \) is a diffeomorphism, with inverse

\[
\mathfrak{N}^{-1}(t, s, q, x) := (t, s, X[s, \mu^x](t, q), x) = (t, s, N(t, s, q, x), x) 
\]

for \( t, s \in [0, T], q \in \mathbb{T}^d, x \in (\mathbb{T}^d)^n; \mathfrak{N}^{-1} \) is thus the "discretized" version of the inverse master map (Definition 4.20). By Corollary 5.5 and (5.17), for each fixed \( s \in [0, T], \mathfrak{N}^{-1}(\cdot, s, \cdot, \cdot) \in W^{2,2,2,\infty}([0, T] \times \mathbb{T}^d \times (\mathbb{T}^d)^n; [0, T] \times \mathbb{T}^d \times (\mathbb{T}^d)^n). \) We are going to derive now the Lipschitz property of \( X[s, \cdot](\cdot, \cdot) \) before addressing the full regularity of the master map in the next section.

Recall estimate (4.21):

\[
\|\nabla q X[s, \mu](t, \cdot)\|_{\infty} < 4(1 + \sqrt{d})^{d-1}, \quad s, t \in [0, T], \mu \in \mathcal{P}(\mathbb{T}^d). 
\]

Differentiating the identity \( q = X[s, \mu](t, \Sigma^1[s, \mu](t, q)) \) with respect to \( t \), we have

\[
0 = \partial_t X[s, \mu](t, \Sigma^1[s, \mu](t, q)) + \nabla q X[s, \mu](t, \Sigma^1[s, \mu](t, q)) \partial_t \Sigma^1[s, \mu](t, q), 
\]

(5.21)

from which

\[
\partial_t X[s, \mu](t, q) = -\nabla q X[s, \mu](t, q)v[s, \mu](t, q), 
\]

(5.22)

at any \( s, t \in [0, T], q \in \mathbb{T}^d, \mu \in \mathcal{P}(\mathbb{T}^d). \) Therefore

\[
\|\partial_t X[s, \mu]\|_{\infty} \leq \|\nabla q X[s, \mu]\|_{\infty}\|v[s, \mu]\|_{\infty} \leq 4A_1(1 + \sqrt{d})^{d-1}. 
\]

(5.23)

For the regularity with respect to \( x = (x_1, \ldots, x_n) \in (\mathbb{T}^d)^n \), we use the identity

\[
M_1(t, s, N(t, s, q, x), x) = q, 
\]

which holds by definition. Taking the derivative with respect to \( x_j, j = 1, \ldots, n \), gives

\[
-\nabla x_j N(t, s, q, x) = [\nabla q M_1(t, s, N(t, s, q, x), x)]^{-1} \nabla x_j M_1(t, s, q, x) = \nabla q N(t, s, q, x) \nabla x_j M_1(t, s, q, x). 
\]

Thus, \( \|\nabla x_j N\|_{\infty} \leq 4(1 + \sqrt{d})^{d-1}\frac{C}{n}, \) which, increasing the value of \( C \), gives

\[
\|\nabla x_j N\|_{\infty} \leq \frac{C}{n}, 
\]

(5.24)
Corollary 5.7. Let $t, t', s \in [0, T]$, $q', q \in \mathbb{T}^d$, $\mu, \nu \in \mathcal{P}(\mathbb{T}^d)$. Then there is a constant $C > 0$ such that

$$|X[s, \nu](t', q') - X[s, \mu](t, q)| \leq C(|t' - t| + |q' - q| + \mathcal{W}(\mu, \nu)).$$

Proof. Let $x, x' \in (\mathbb{T}^d)^n$, and $N = N(t, s, q, x)$, $N' = N(t', s, q', x')$, where $N$ is defined in (3.19). By the bounds (5.29), (5.21), (5.24), and relabeling the sequence $x_1, \ldots, x_n$ and shifting the points so that $\mathcal{W}^2(\mu^x, \mu^{x'}) = \sum_{j=1}^n |x_j - x'_j|^2$,

$$N(t', s, q', x') - N(t, s, q, x) \leq \|Ct\|_\mathcal{X}|t' - t| + \|N q N\|_{\mathcal{X}}|q' - q| + \sum_{j=1}^n \|\nabla x_j N\|_{\mathcal{X}}|x'_j - x_j|$$

$$\leq 4A_1(1 + \sqrt{d})^d|t' - t| + 4(1 + \sqrt{d})^{d-2}|q' - q| + \sum_{j=1}^n C_n|x'_j - x_j|;$$

therefore, since $\sum \frac{C_n}{n}|x'_j - x_j| \leq C(\sum 1/n)^{1/2}(\sum |x'_j - x_j|^2/n)^{1/2}$, we get, by increasing $C$,

$$|N' - N| \leq C(|t' - t| + |q - q'| + \mathcal{W}(\mu^x, \mu^{x'})).$$

The constant $C$ does not depend on $n$. Applying the last fact in the list of Section 2 we now extend this to the arbitrary measure case: let $\mu, \nu \in \mathcal{P}(\mathbb{T}^d)$, and $(x(n))_{n=1}^\infty, (x'(n))_{n=1}^\infty$, with $x(n), x'(n) \in (\mathbb{T}^d)^n$, sequences such that

$$\lim_{n \to \infty} \mathcal{W}(\mu^{x(n)}, \mu) = 0, \quad \lim_{n \to \infty} \mathcal{W}(\mu^{x'(n)}, \nu) = 0.$$

Since, by definition, $N(t, s, q, x) = X[s, \mu^x](t, q)$, the latter estimate means

$$|X[s, \mu^{x(n)}](t', q') - X[s, \mu^{x(n)}](t, q)| \leq C(|t' - t| + |q - q'| + \mathcal{W}(\mu^{x(n)}, \mu^{x'(n)})),$$

for every $n \in \mathbb{Z}^n$. Letting $n \to \infty$, the continuity of $X$ in all its variables finalizes the proof. \hfill \square

5.3 Regularity properties of the master map

We follow [25] closely here. The idea, roughly speaking, begins with introducing a Lipschitz extension of the “discretized derivative” $\nabla_{x_j} M_1$, $j = 1, \ldots, n$, that is defined at every measure $\mu$, through an argument reminiscent of Moreau-Yosida approximation, the extension being closer to $n \nabla_{x_j} M_1$ the larger $n$ —the number of particles—is. When the $n$-particle ordered sets $x^n = (x_1^n, \ldots, x_n^n)$ are chosen in such a way that $\delta x^n \to \mu$, the extension just mentioned will reveal itself as the Wasserstein gradient in the first-order Taylor approximation derived in the preceding paragraphs; recall [25].

For fixed $n \in \mathbb{Z}^+$, let

$$\mathcal{B} := [0, T] \times [0, T] \times \mathbb{T}^d \times \{ (y_j, \mu^y) \mid y = (y_1, \ldots, y_n) \in (\mathbb{T}^d)^n, j \in \{1, \ldots, n\} \}.$$

A typical element of $\mathcal{B}$ is thus $(t, s, q, (y_j, \mu^y))$ where $y$ is any $n$-particle ordered set $(y_1, \ldots, y_n) \in (\mathbb{T}^d)^n$ and $y_j$ is any of its component particles. If $m \in \mathbb{Z}^+$ and $f : \mathcal{B} \to \mathbb{R}^m$ is a continuous function, let

$$\|f\|_{\mathcal{B}} := \sup\{|f(t, s, q, (x_j, \mu^x))| \mid t, s \in [0, T], q \in \mathbb{T}^d, x \in (\mathbb{T}^d)^n, j \in \{1, \ldots, n\}\}. $$

37
For any continuous function \( f = (f^{(1)}, \ldots, f^{(m)}) : \mathcal{B} \rightarrow \mathbb{R}^m \) such that

\[
|f(t, s, q, (y_j, \mu^y)) - f(t, s, q, (x_i, \mu^x))| \leq C(|x_i - y_j|_{\mathbb{T}^d} + \mathcal{W}(\mu^x, \mu^y) + \frac{1}{n}),
\]

(5.25)

where \( t, s \in [0, T], q \in \mathbb{T}^d, x, y \in (\mathbb{T}^d)^n, i, j \in \{1, \ldots, n\} \), define

\[
g^{(l)}(t, s, q, z, \mu) := \inf \left\{ f^{(l)}(t, s, q, (y_j, \mu^y)) + C(|z - y_j|_{\mathbb{T}^d} + \mathcal{W}(\mu, \mu^y)) \mid y \in (\mathbb{T}^d)^n, j \in \{1, \ldots, n\} \right\},
\]

\( l = 1, \ldots, m, \)

\[
g := (g^{(1)}, \ldots, g^{(m)}),
\]

at any fixed \( z \in \mathbb{T}^d, \mu \in \mathcal{P}(\mathbb{T}^d) \). The function \( g \) is thus an extension of \( f \) from \( \mathcal{B} \) to the full space \([0, T]^2 \times \mathbb{T}^d \times [\mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d)]\). The following is [25, Lemma 8.10].

**Proposition 5.8.** Suppose that (5.25) holds, and for any \( x \in (\mathbb{T}^d)^n, j \in \{1, \ldots, n\}, f(\cdot, \cdot, \cdot, (x_j, \mu^x)) \) is \( C \)-Lipschitz. Then

(i) \( g \) is \( \sqrt{3}C \)-Lipschitz,

(ii) \( \|g|_{\mathcal{B}} - f\|_{\mathcal{B}} \leq C/n \).

As in [25], we set, for \( s, t \in [0, T], q \in \mathbb{T}^d, x = (x_1, \ldots, x_n) \in (\mathbb{T}^d)^n, j = 1, \ldots, n, \)

\[
\zeta^n(t, s, q, (x_j, \mu^x)) = n \nabla_{x_j} M(t, s, q, x).
\]

(5.26)

The periodicity of \( M \) in \( q \) and \( x \) ensures that \( \zeta^n \) is well defined on \( \mathcal{B} \).

**Corollary 5.9.** (Extension of \( \zeta^n \)) For each \( n \in \mathbb{Z}^+ \), there is a function

\[
\chi^n : [0, T] \times [0, T] \times \mathbb{T}^d \times \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}^d \times \mathbb{R}^d
\]

such that \( \chi^n|_{\mathcal{B}} = \zeta^n \) and, with a larger value of \( C \) than before,

(i) \( \chi^n \) is \( C \)-Lipschitz,

(ii) \( \|\chi^n|_{\mathcal{B}} - \zeta^n\|_{\mathcal{B}} \leq \frac{C}{n} \).

**Proof.** We check that \( f = \zeta^n \) satisfies the conditions of Proposition 5.8. The Lipschitz property in \( t \) and \( q \) follows from (5.17), while, in \( s \), from (5.16). Hence, to obtain the Corollary, it is enough to prove that the condition (5.25) is satisfied by \( f = \zeta^n \). Fix then \( x, y \in (\mathbb{T}^d)^n, i, j \in \{1, \ldots, n\}, s, t \in [0, T], q \in \mathbb{T}^d \). Since the order in which we take the \( n \) particles \( x_1, \ldots, x_n \), which make up \( x \in (\mathbb{T}^d)^n \), does not change \( M(\cdot, \cdot, x) \), and \( \nabla_{x_j} M(\cdot, \cdot, x) \) is periodic in \( x \), it can be assumed that:

\[
\sum_{k \neq i, j} |x_k - y_k|^2 \leq \mathcal{W}^2(\mu^x, \mu^y), \quad |x_j - y_i| = |x_j - y_i|_{\mathbb{T}^d}, \quad |x_i - y_j| = |x_i - y_j|_{\mathbb{T}^d},
\]

\[
\nabla_{x_j} M(t, s, q, y) = \nabla_{x_i} M(t, s, q, y), \quad \nabla_{x_j} M(t, s, q, x) = \nabla_{x_i} M(t, s, q, x).
\]
where $\bar{y}$ denotes the result of shifting $y_j$ and $y_i$ to the first and second positions, respectively, in the $n$-uple $y$, and $\bar{x}$ denotes the result of shifting $x_i$ and $x_j$ to the first and second positions, respectively, in the $n$-uple $x$. Suppose, too, without loss of generality, that $i < j$. In view of these simplifications,

$$|\nabla_{x_j} M(t, s, q, y) - \nabla_{x_i} M(t, s, q, x)|$$

$$\leq |\nabla_{x_i} M(t, s, q, x)| y_j - x_i | + |\nabla_{x_j} M(t, s, q, x)| y_i - x_j | + \sum_{k=1}^{i-1} |\nabla_{x_{x_{k+2}}} M(t, s, q, x)| y_k - x_k|

+ \sum_{k=i+1}^{j-1} |\nabla_{x_{x_{k+1}}} M(t, s, q, x)| y_k - x_k|

+ \sum_{k=j+1}^{n} |\nabla_{x_{x_{k}}} M(t, s, q, x)| y_k - x_k|.

Therefore, by the bounds of Corollary 5.5,

$$|\nabla_{x_j} M(t, s, q, y) - \nabla_{x_i} M(t, s, q, x)| \leq \frac{C}{n} |y_j - x_i| + \frac{C}{n^2} |y_i - x_j| + \sum_{k \neq i,j} |y_k - x_k|

\leq \frac{C}{n} |y_j - x_i|_{T^d} + \frac{C}{n^2} |y_i - x_j|_{T^d} + \frac{C}{n^2} \sqrt{\sum_{k \neq i,j} |y_k - x_k|^2}

\leq \frac{C}{n} (|y_j - x_i|_{T^d} + \frac{\sqrt{d}}{2n} + \mathcal{W}(\mu^x, \mu^y)),

$$

where $\sqrt{d}/2$ is the diameter of $T^d$. Thus,

$$|n \nabla_{x_j} M(t, s, q, y) - n \nabla_{x_i} M(t, s, q, x)| \leq \sqrt{d}C(|y_j - x_i|_{T^d} + \mathcal{W}(\mu^x, \mu^y) + \frac{1}{n}), \quad (5.27)$$

which proves property (5.25) for $f = \zeta^n$, since $i$ and $j$ were arbitrary.

**Lemma 5.10.** For every $s \in [0, T]$, the $T^d \times \mathbb{R}^d$-valued map $\Sigma[s, \cdot](\cdot, \cdot)$ is differentiable on $\mathcal{P}(T^d) \times [0, T] \times T^d$, that is: for every $s, t, t' \in [0, T], q, q' \in T^d, \mu, \nu \in \mathcal{P}(T^d), \gamma \in \Gamma_0(\mu, \nu)$,

$$|\Sigma[s, \nu](t', q') - \Sigma[s, \mu](t, q) - \partial_t \Sigma[s, \mu](t, q)(t' - t) - \nabla_q \Sigma[s, \mu](t, q) \cdot (q' - q)$$

$$- \int_{T^d \times T^d} \nabla_{\mu} \Sigma[s, \mu](t, q, x) \cdot (y - x) \gamma(dx, dy)|$$

$$\leq C(|t' - t|^2 + |q' - q|^2 + \mathcal{W}^2(\mu, \nu)), \quad (5.28)$$

where the mapping

$$\nabla_{\mu} \Sigma : [0, T] \times [0, T] \times T^d \times \mathcal{P}(T^d) \rightarrow \mathbb{R}^{d^2} \times \mathbb{R}^{d^2}$$

$$(t, s, q, \mu) \rightarrow \nabla_{\mu} \Sigma[s, \mu](t, q, x)$$

is Lipschitz.

**Proof.** Let $\mu, \nu \in \mathcal{P}(T^d)$, and let $\gamma \in \Gamma_0(\mu, \nu)$. Appealing to the last fact in the list of Section 2 there is a sequence $\{\gamma(n)\}_{n=1}^\infty$, converging narrowly to $\gamma$ in $\mathcal{P}(T^d \times T^d)$, such that

$$\gamma(n) = \frac{1}{n} \sum_{j=1}^{n} \delta(x_j(n), y_j(n)),$$
and for each \( j \in \{1, \ldots, n\} \), \((x_j(n), y_j(n))\) belongs to the support of \( \gamma \). Due to this latter fact (see, for instance, \cite[Theorem 6.1.4]{ref}), for each \( n \in \mathbb{Z}^+ \), the sequence \( \{(x_j(n), y_j(n))\}_{j=1}^n \) is \( |\cdot|_{\tau^d}\)-monotone, and, therefore,

\[
\gamma(n) \in \Gamma_0(\mu^{x(n)}, \mu^{y(n)}), \quad n \in \mathbb{Z}^+.
\]

It is also true that

\[
\lim_{n \to \infty} \mathcal{W}(\mu, \mu^{x(n)}) = 0, \quad \lim_{n \to \infty} \mathcal{W}(\nu, \mu^{y(n)}) = 0.
\]

Let \( \zeta^n \) be defined as in (5.23), that is, for each \( x(n) \) of our sequence,

\[
\zeta^n(t, s, q, (x_j(n), \mu^{x(n)})) = n \nabla_{x_j(n)} M(t, s, q, x(n)) = n \nabla_{x_j(n)} \Sigma[s, \mu^{x(n)}](t, q),
\]

\( j \in \{1, \ldots, n\} \). Recall the second-order estimate (5.18), now with \( x = x(n) \) and \( x' = y(n) \):

\[
|\Sigma[s, \mu^{y(n)}](t, q) - \Sigma[s, \mu^{x(n)}](t', q') - \partial_t |\Sigma[s, \mu^{x(n)}](t, q)(t' - t) - \nabla_q \Sigma[s, \mu^{x(n)}](t, q) \cdot (q' - q) \\
- \sum_{j=1}^{n} \nabla_{x_j(n)} \Sigma[s, \mu^{x(n)}](t, q) \cdot (y_j(n) - x_j(n))| \\
\leq C(|t' - t|^2 + |q - q'|^2 + \mathcal{W}^2(\mu^{x(n)}, \mu^{y(n)})).
\]

Since

\[
\frac{1}{n} \sum_{j=1}^{n} \zeta^n(t, s, q, (x_j(n), \mu^{x(n)})) \cdot (y_j(n) - x_j(n)) = \int_{\mathbb{T}^d \times \mathbb{T}^d} \zeta^n(t, s, q, (x, \mu^{x(n)})) \cdot (y - x) \gamma(n)(dx, dy),
\]

the latter inequality is the same as

\[
|\Sigma[s, \mu^{y(n)}](t', q') - \Sigma[s, \mu^{x(n)}](t, q) - \partial_t |\Sigma[s, \mu^{x(n)}](t, q)(t' - t) - \nabla_q \Sigma[s, \mu^{x(n)}](t, q) \cdot (q' - q) \\
- \int_{\mathbb{T}^d \times \mathbb{T}^d} \zeta^n(t, s, q, (x, \mu^{x(n)})) \cdot (y - x) \gamma(n)(dx, dy)| \\
\leq C(|t' - t|^2 + |q - q'|^2 + \mathcal{W}^2(\mu^{x(n)}, \mu^{y(n)})).
\]

Denote by \( \chi^n \) the extension of \( \zeta^n \) furnished by Corollary 5.9. The same inequality holds if we substitute \( \chi^n \) for \( \zeta^n \) in the previous inequality, because these functions coincide on the set \( \mathcal{B} \), which includes the support of \( \gamma(n) \). But, since we will pass to the limit, we rather write

\[
|\Sigma[s, \mu^{y(n)}](t, q) - \Sigma[s, \mu^{x(n)}](t', q') - \partial_t |\Sigma[s, \mu^{x(n)}](t, q)(t' - t) - \nabla_q \Sigma[s, \mu^{x(n)}](t, q) \cdot (q' - q) \\
- \int_{\mathbb{T}^d \times \mathbb{T}^d} \chi^n(t, s, q, (x, \mu^{x(n)})) \cdot (y - x) \gamma(n)(dx, dy)| \\
\leq C(|t' - t|^2 + |q - q'|^2 + \mathcal{W}^2(\mu^{x(n)}, \mu^{y(n)})) \\
+ \int_{\mathbb{T}^d \times \mathbb{T}^d} |\zeta^n(t, s, q, (x, \mu^{x(n)})) - \chi^n(t, s, q, (x, \mu^{x(n)}))| \cdot (y - x) \gamma(n)(dx, dy)| \\
\leq C(|t' - t|^2 + |q - q'|^2 + \mathcal{W}^2(\mu^{x(n)}, \mu^{y(n)})) + \frac{C}{n} \mathcal{W}(\mu^{x(n)}, \mu^{y(n)}),
\]

(5.29)
The continuity of Proof.

For every been known since Lemma 4.16 and formula (5.22) respectively. Let us put

Definition 5.11. For put

Next, we are going to prove the analogue of Lemma 5.10 for \( X = (\Sigma^1)^{-1} \), which amounts to smoothness of \( \mathcal{R} \) in the \( \mu \) variable.

**Lemma 5.12.** For every \( s \in [0, T] \), the \( \mathbb{R}^d \)-valued map \( X[s, \cdot](\cdot, \cdot) \) is differentiable on \( \mathcal{P}(\mathbb{T}^d) \times [0, T] \times \mathbb{T}^d \), i.e., there is a constant \( C > 0 \) such that for every \( s, t, t' \in [0, T] \), \( q, q' \in \mathbb{T}^d \), \( \mu, \nu \in \mathcal{P}(\mathbb{T}^d) \), \( \gamma \in \Gamma_0(\mu, \nu) \),

\[
|X[s, \nu](t', q') - X[s, \mu](t, q) - \tilde{c}_tX[s, \mu](t, q)(t' - t) - \nabla_qX[s, \mu](t, q)(q' - q)| - \int_{\mathbb{T}^d \times \mathbb{T}^d} \nabla_{\mu}X[s, \mu](t, q, x) \cdot (y - x) \gamma(dx, dy) \leq C(|t' - t|^2 + |q - q'|^2 + \mathcal{W}^2(\mu, \nu)),
\]

(5.30)

where \( \tilde{c}_tX \), \( \nabla_qX \), and the mapping \( \nabla_\mu X \) of Definition 5.11 are continuous.

Proof. The continuity of \( \nabla_\mu X \) is immediate from its definition, and the continuity of \( \nabla_q X \) and \( \tilde{c}_t X \) has been known since Lemma 4.10 and formula (5.22), respectively. Let us put

\[
\bar{q} = X[s, \mu](t, q), \quad \bar{q}' = X[s, \nu](t', q').
\]

(5.31)

We write out the expression on the left hand side of (5.30) and factor out \( \nabla_qX[s, \mu](t, q) \), while also using the fact that \( \nabla_qX[s, \mu](t, q) \) and \( \nabla_q\Sigma^1[s, \mu](t, \bar{q}) \) are inverses of one another:

\[
\left| \nabla_qX[s, \mu](t, q) \nabla_\Sigma \Sigma^1[s, \mu](t, \bar{q})(\bar{q}' - \bar{q}) + \tilde{c}_t\Sigma^1[s, \mu](t, \bar{q})(t' - t) - (\Sigma^1[s, \nu](t', \bar{q}') - \Sigma^1[s, \mu](t, \bar{q})) + \int_{\mathbb{T}^d \times \mathbb{T}^d} \nabla_{\mu}\Sigma^1[s, \mu](t, \bar{q}, x) \cdot (y - x) \gamma(dx, dy) \right|
\]

\[
= \left| - \nabla_qX[s, \mu](t, q) \left( \Sigma^1[s, \nu](t', q') - \Sigma^1[s, \mu](t, q) - \tilde{c}_t\Sigma^1[s, \mu](t, \bar{q})(t' - t) - \nabla_q\Sigma^1[s, \mu](t, \bar{q})(q' - q) - \int_{\mathbb{T}^d \times \mathbb{T}^d} \nabla_{\mu}\Sigma^1[s, \mu](t, \bar{q}, x) \cdot (y - x) \gamma(dx, dy) \right) \right|
\]

\[
\leq 4(1 + \sqrt{d})^{d-1}C(|t' - t|^2 + |q' - q|^2 + \mathcal{W}^2(\mu, \nu))
\]

\[
= 4(1 + \sqrt{d})^{d-1}C(|t' - t|^2 + |X[s, \nu](t', q') - X[s, \mu](t, q)|^2 + \mathcal{W}^2(\mu, \nu)).
\]
But, by Corollary 5.7, the term \(|X[s, ν](t', q') - X[s, μ](t, q)|\) is bounded by \(C(|t' - t| + |q' - q| + \mathcal{W}(μ, ν))\) for some \(C > 0\). Inserting this bound into the last expression, after expanding and raising the value of \(C\), one obtains (5.30). \(\Box\)

**5.3.1 Regularity of \(V\)**

Let us look back at the definition of \(V\), given by (1.23). Set now

\[
\nabla_μ V[s, μ](t, q, x) := \nabla_μ Σ^2[s, μ](t, X[s, μ](t, q), x) + \nabla_q Σ^2[s, μ](t, X[s, μ](t, q)) \nabla_μ X[s, μ](t, q, x),
\]

for \(s, t \in [0, T]\), \(q, x \in \mathbb{T}^d\), \(μ ∈ \mathcal{P}(\mathbb{T}^d)\).

**Lemma 5.13.** For every \(s ∈ [0, T]\), the \(\mathbb{R}^d\)-valued map \(V[s, \cdot](\cdot, \cdot)\) is differentiable on \(\mathcal{P}(\mathbb{T}^d) × [0, T] × \mathbb{T}^d\), i.e., there is a constant \(C > 0\) such that for every \(s, t, t' \in [0, T]\), \(q, q' \in \mathbb{T}^d\), \(μ, ν ∈ \mathcal{P}(\mathbb{T}^d)\), \(γ ∈ \Gamma_0(μ, ν)\),

\[
|V[s, ν](t', q') - V[s, μ](t, q) - \partial_t V[s, μ](t, q)(t' - t) - \nabla_q V[s, μ](t, q) \cdot (q' - q) - \int_{\mathbb{T}^d × \mathbb{T}^d} \nabla_μ V[s, μ](t, q, x) \cdot (y - x)γ(dx, dy)|
\]

\[
≤ C(|t' - t| + |q - q'|^2 + \mathcal{W}^2(μ, ν)),
\]

where the mapping \(\nabla_μ V\), defined by (5.32), and \(\partial_t V, \nabla_q V\), are continuous.

**Proof.** We know that

\[
\nabla_q V_t[s, μ] = \nabla_q Σ^2_t[s, μ] \nabla_q X_t[s, μ], \quad \partial_t V_t[s, μ] = \partial_t Σ^2_t[s, μ] + \nabla_q Σ^2_t[s, μ] \partial_t X_t[s, μ].
\]

Therefore, the continuity of the functions stated in the lemma follows from that of \(\nabla_q Σ, \nabla_q X, \partial_t Σ, \partial_t X\), and the continuity, proved above, of \(\nabla_μ Σ\) and \(\nabla_μ X\). Keeping the notation (5.31), we first write down the expression to estimate, i.e. the left hand side of (5.33), and factor out \(\nabla_q Σ^2_t[s, μ](t, \tilde{q})\):

\[
|Σ^2[s, ν](t', \tilde{q}) - Σ^2[s, μ](t, \tilde{q}) - \nabla_q Σ^2[s, μ](t, \tilde{q}) \nabla_q X[s, μ](t, \tilde{q})(q' - q) - \partial_t Σ^2_t[s, μ](t, \tilde{q}) \partial_t X[s, μ](t, \tilde{q})(t' - t) - \int_{\mathbb{T}^d × \mathbb{T}^d} \nabla_μ Σ^2_t[s, μ](t, \tilde{q}) \nabla_q X[s, μ](t, \tilde{q}, x) \cdot (y - x)γ(dx, dy)|
\]

\[
= |Σ^2[s, ν](t', \tilde{q}') - Σ^2[s, μ](t, \tilde{q}) - \nabla_q Σ^2[s, μ](t, \tilde{q}) \nabla_q X[s, μ](t, \tilde{q})(q' - q) + \partial_t X[s, μ](t, \tilde{q})(t' - t) + \int_{\mathbb{T}^d × \mathbb{T}^d} \nabla_μ X[s, μ](t, \tilde{q}) \cdot (y - x)γ(dx, dy) |
\]

\[
- \partial_t Σ^2_t[s, μ](t, \tilde{q})(t' - t) - \int_{\mathbb{T}^d × \mathbb{T}^d} \nabla_μ Σ^2_t[s, μ](t, \tilde{q}, x) \cdot (y - x)γ(dx, dy)|.
\]

Inside the large round brackets we add and substract \(\tilde{q}' - \tilde{q} = X_t'[s, ν](q') - X_t[s, μ](q)\), and apply (5.30), to obtain that the latter expression is no greater than

\[
|Σ^2_t[s, ν](q') - Σ^2_t[s, μ](q) - \nabla_q Σ^2_t[s, μ](q)(q - q) - \partial_t Σ^2_t[s, μ](q)(t' - t) - \int_{\mathbb{T}^d × \mathbb{T}^d} \nabla_μ Σ^2_t[s, μ](q, x) \cdot (y - x)γ(dx, dy)| + |\nabla_q Σ^2_t[s, μ](q)|C(|q' - q|^2 + |t' - t|^2 + \mathcal{W}^2(μ, ν)),
\]

42
which, in turn, by \([5.30]\), is bounded above by
\[
C(|X[s, \nu](t', q') - X[s, \mu](t, q)|^2 + |t' - t|^2 + |q' - q|^2 + \mathcal{W}^2(\mu, \nu))
+ \theta A_2 C(|q' - q|^2 + |t' - t|^2 + \mathcal{W}^2(\mu, \nu)),
\]
and, after using Corollary 5.7 again, simplifying and increasing the value of \(C\), inequality \([5.33]\) is obtained.

### 5.3.2 Regularity of \(H(q, \nu)\)

We finish this section by following the previous results with the regularity of what turned out to be the second function that appears in the MFG equation \((1.2)\), due to \((4.28)\). Set
\[
\hat{\nabla}_\mu H(q, \nu[s, \mu](t, q), \nu[s, \mu](t, q, x)),
\]
for \(s, t \in [0, T]\), \(q, x \in \mathbb{T}^d\), \(\mu \in \mathcal{P}(\mathbb{T}^d)\).

**Lemma 5.14.** For every \(s \in [0, T]\), the \(\mathbb{R}\)-valued map \(H(\cdot, \nu[s, \cdot](\cdot, \cdot))\) is differentiable on \(\mathcal{P}(\mathbb{T}^d) \times [0, T] \times \mathbb{T}^d\), and there is a constant \(C > 0\) such that for every \(s, t, t' \in [0, T]\), \(q, q' \in \mathbb{T}^d\), \(\mu, \nu \in \mathcal{P}(\mathbb{T}^d)\), \(\gamma \in \Gamma_0(\mu, \nu)\),
\[
|H(q', \nu[s, \nu](t', q')) - H(q, \nu[s, \mu](t, q))) - (\partial_t)(H(q, \nu[s, \mu](t, q)))(t' - t)
- (\nabla_q)(H(q, \nu[s, \mu](t, q)))(q' - q) - \int_{\mathbb{T}^d \times \mathbb{T}^d} \hat{\nabla}_\mu H(q, \nu[s, \mu](t, q))(x) \cdot (y - x)\gamma(dx, dy)|
\leq C(|t' - t|^2 + |q' - q'|^2 + \mathcal{W}^2(\mu, \nu)),
\]
where the mapping \(\hat{\nabla}_\mu H\), defined by \([5.34]\), is continuous.

**Proof.** Let us abbreviate \(\nu = \nu[s, \mu](q)\) and \(\nu' = \nu[s, \nu](q')\). Since
\[
(\partial_t)(H(q, \nu)) = \nabla_p H(q, \nu)\partial_t \nu, \quad (\nabla_q)(H(q, \nu)) = \nabla_q H(q, \nu) + \nabla_p H(q, \nu)\nabla_q \nu,
\]
the left hand side of \([5.33]\) is, after factoring out \(\nabla_p H(q, \nu)\),
\[
|H(q', \nu') - H(q, \nu) - \nabla_p H(q, \nu)(\nu' - \nu) + \nabla_q H(q, \nu)(q' - q) + \nabla_p H(q, \nu)\nabla_q \nu|(t' - t)
- \nabla_q H(q, \nu)(q' - q) - \nabla_p H(q, \nu')(\nu' - \nu')|
\leq |H(q', \nu') - H(q, \nu) - \nabla_q H(q, \nu)(q' - q) - \nabla_p H(q, \nu)(\nu' - \nu)|
+ |\nabla_p H(q, \nu)|C(|t' - t|^2 + |q' - q|^2 + \mathcal{W}^2(\mu, \nu)).
\]
Remember now that \(|\Sigma^2| \leq \theta B\) (see Corollary \([4.7]\),\(\text{ii}\)) at any \(t, q, s, \mu\); recall Definition \([4.10]\). Therefore, the right-hand side of this inequality is bounded by
\[
h(\theta B)(|q' - q|^2 + |\nu' - \nu|^2) + l(\theta B)C(|t' - t|^2 + |q' - q|^2 + \mathcal{W}^2(\mu, \nu)).
\]
To deal with the term $|\mathcal{V}' - \mathcal{V}|^2$, note that Corollary 5.7 is also valid for $\Sigma$ in place of $X$, following a similar argument. With the notation (5.31),

$$|\mathcal{V}' - \mathcal{V}| = |\Sigma^2[s, \nu](t', \tilde{q}') - \Sigma^2[s, \mu](t, \tilde{q})| \leq C(|t' - t| + |\tilde{q}' - \tilde{q}| + H(\mu, \nu)).$$

Applying Corollary 5.7 and raising the value of $C$, we get

$$|\mathcal{V}' - \mathcal{V}| \leq C(|t' - t| + |q' - q| + H(\mu, \nu)).$$

Substituting this into the bounding expression, and simplifying, one arrives at (5.35), for some larger value of $C$. The continuity of $\bar{\nabla}_\mu H$ in all its variables is clear from the definition.

6 Solution to the master equation

Let us recall the definition of the function $u$ from Section 4.3:

$$u(s, q, \mu) = g(q, \sigma_0) - \int_0^s \left[ H(q, \mathcal{V}[s, \mu](\tau, q)) + F(q, \sigma_\tau) \right] d\tau,$$

where $\sigma_\tau = \Sigma^1[s, \mu](\tau, \cdot)_\# \mu$, $0 \leq \tau \leq s$. By the particular forms of the functions $F$ and $g$ assumed at the beginning of Section 5,

$$u(s, q, \mu) = U^0(q) + \int_{T^d} U^1(q - \Sigma^1_0[s, \mu](y))\mu(dy),$$

$$- \int_0^s \left[ H(q, \mathcal{V}[s, \mu](\tau, q)) + \int_{T^d} \phi(q - \Sigma^1_\tau[s, \mu](y))\mu(dy) \right] d\tau. \tag{6.1}$$

6.1 Gradients of $F(q, \cdot)$ and $g(q, \cdot)$

The differentiability in $\mu$ of the expression that substitutes $F(q, \sigma_t)$ in (6.1), as well as that which substitutes $g(q, \mu)$, are dealt with in [25], and we rewrite the proof here without any changes except for a slight change in notation. We just draw the readers’ attention to the fact that, in this case, the formula for what should be the derivative in $\mu$ of those expressions is not easily guessed by formally applying the chain rule, since $\mu$ appears not only in the integrand, but as the integrating measure.

Set

$$\sigma'_\nu = \Sigma^1_\nu[s, \nu]_\# \nu, \quad \sigma_t = \Sigma^1_t[s, \mu]_\# \mu,$$

and define the functions

$$(\hat{\sigma}_t)(F(q, \Sigma^1_t[s, \mu]_\# \mu)) := - \int_{T^d} \nabla \phi(q - \Sigma^1_t[s, \mu](y)) \hat{\sigma}_t \Sigma^1_t[s, \mu](y)\mu(dy);$$

$$(\bar{\nabla}_\mu)(F(q, \Sigma^1_t[s, \mu]_\# \mu))(y) := - \nabla \phi(q - \Sigma^1_t[s, \mu](y)) \nabla \Sigma^1_t[s, \mu](y)$$

$$- \int_{T^d} \nabla \phi(q - \Sigma^1_t[s, \mu](u)) \bar{\nabla}_\mu \Sigma^1_t[s, \mu](u, y)\mu(du),$$

for $s, t \in [0, T]$, $q, y \in T^d$, $\mu \in \mathcal{P}(T^d)$.
Lemma 6.1. The function \((t, s, q, \mu) \mapsto F(q, \Sigma_t^1[s, \mu]_{\#})\) is continuously differentiable in the sense that, with a larger value of \(C\),

\[
|F(q', \sigma') - F(q, \sigma) - (\hat{c}_t)(F(q, \sigma_t))(t' - t) - \nabla_q F(q, \sigma_t) \cdot (q' - q) \\
- \int_{T^d \times T^d} (\nabla_\mu)(F(q, \sigma_t))(y) \cdot (x - y) \gamma(dy, dx)| \\
\leq C(|t' - t|^2 + |q' - q|^2 + \mathcal{W}^2(\mu, \nu)),
\]

for any \(q, q' \in T^d\), \(\gamma \in \Gamma_0(\mu, \nu), s, t, t' \in [0, T]\), where the functions \((\hat{c}_t)(F(q, \Sigma_t^1[s, \mu]_{\#})),
(\nabla_\mu)(F(q, \Sigma_t^1[s, \mu]_{\#}))(y)\), just defined, are continuous in all its variables.

Proof. The continuity of said functions is straightforward from the continuity of its parts. To obtain the estimate (6.2), the proof begins by noting that

\[
|F(q', \sigma') - F(q, \sigma_t)| = \left| \int_{T^d \times T^d} [\phi(q' - \Sigma_t^1[s, \sigma](x)) - \phi(q - \Sigma_t^1[s, \sigma](y))] \gamma(dy, dx) \right|
\]

for any \(\gamma \in \Gamma(\mu, \nu)\), in particular for \(\gamma \in \Gamma_0(\mu, \nu)\). Let then \(\gamma \in \Gamma_0(\mu, \nu)\), and apply a first-order Taylor estimate of \(\phi\), remembering that \(\kappa\) bounds \(\|\nabla^2 \phi\|_{L^1}\):

\[
|\int_{T^d \times T^d} [\phi(q' - \Sigma_t^1[s, \sigma](x)) - \phi(q - \Sigma_t^1[s, \sigma](y)) - \nabla \phi(q - \Sigma_t^1[s, \sigma](y)) \cdot (q - q') \\
- \nabla \phi(q - \Sigma_t^1[s, \sigma](y)) \cdot (\Sigma_t^1[s, \sigma](x) - \Sigma_t^1[s, \sigma](y))] \gamma(dy, dx)| \\
\leq \frac{\kappa}{2} |q' - q - \Sigma_t^1[s, \sigma](x) + \Sigma_t^1[s, \sigma](y)|^2 \gamma(dy, dx);
\]

then, because of the Lipschitz continuity of \(\Sigma^1\) and the definition of Wasserstein distance,

\[
\text{RHS of (6.3)} \leq \frac{\kappa}{2} C \int_{T^d \times T^d} (|q' - q|^2 + |y - x|^2 + |t' - t|^2 + \mathcal{W}^2(\mu, \nu)) \gamma(dy, dx) \\
\leq C'(|q' - q|^2 + |t' - t|^2 + 2\mathcal{W}^2(\mu, \nu))
\]

for some constant \(C'\). Since

\[
\int_{T^d \times T^d} \nabla \phi(q - \Sigma_t^1[s, \sigma](y)) \cdot (q - q') \gamma(dy, dx) = \nabla_q F(q, \sigma_t) \cdot (q' - q),
\]

we can rewrite the resulting inequality as

\[
|F(q', \sigma') - F(q, \sigma_t) - \nabla_q F(q, \sigma_t) \cdot (q' - q) - B| \leq C'(|q' - q|^2 + |t' - t|^2 + 2\mathcal{W}^2(\mu, \nu)),
\]

where

\[
B := \int_{T^d \times T^d} \nabla \phi(q - \Sigma_t^1[s, \sigma](y)) \cdot (\Sigma_t^1[s, \sigma](x) - \Sigma_t^1[s, \sigma](y)) \gamma(dy, dx).
\]

We see that the part of the desired estimate (6.2) corresponding to the increment in \(q\) is taken care of. The key of the proof is that the remaining part is also already contained in (6.3), if one works out the expression for \(B\). Indeed, let us put

\[
B = B_1 + B_2,
\]
where
\[
B_1 := \int_{T^d \times T^d} \nabla \phi(q - \Sigma^1_t[s, \mu](y)) \cdot \left( \Sigma^1_t[s, \mu](y) - \Sigma^1_t[s, \mu](y) - \partial_t \Sigma^1_t[s, \mu](y)(t' - t) \right. \\
\left. - \nabla_q \Sigma^1_t[s, \mu](y)(x - y) - \int_{T^d \times T^d} \nabla_\mu \Sigma^1_t[s, \mu](y, r) \cdot (v - r) \gamma(dr, dv) \right) \gamma(dy, dx)
\]
and
\[
B_2 := \int_{T^d \times T^d} \nabla \phi(q - \Sigma^1_t[s, \mu](y)) \cdot \left( \partial_t \Sigma^1_t[s, \mu](y)(t' - t) + \nabla_q \Sigma^1_t[s, \mu](y)(y - x) \\
+ \int_{T^d \times T^d} \nabla_\mu \Sigma^1_t[s, \mu](y, r) \cdot (v - r) \gamma(dr, dv) \right) \gamma(dy, dx).
\]
Using (5.28),
\[
|B_1| \leq \int_{T^d \times T^d} |\nabla \phi(q - \Sigma^1_t[s, \mu](y))|C(|t' - t|^2 + |x - y|^2 + W^2(\mu, \nu)) \gamma(dy, dx) \\
\leq \kappa C(|t' - t|^2 + 2W^2(\mu, \nu)).
\]
Again by the fact that \( \gamma \) is a transport plan between \( \mu \) and \( \nu \), we observe that the term in \( B_2 \) involving the time derivative of \( \Sigma^1_t \) is \( (\partial_t)(F(q, \sigma_t))(t' - t) \), according to our definition above. The term in \( B_2 \) involving \( \nabla_\mu \Sigma^1 \) is
\[
\int_{T^d \times T^d} \nabla \phi(q - \Sigma^1_t[s, \mu](y)) \cdot \int_{T^d \times T^d} \nabla_\mu \Sigma^1_t[s, \mu](y, r) \cdot (v - r) \gamma(dr, dv) \gamma(dy, dx) \\
= \int_{T^d} \nabla \phi(q - \Sigma^1_t[s, \mu](y)) \cdot \int_{T^d \times T^d} \nabla_\mu \Sigma^1_t[s, \mu](y, r) \cdot (v - r) \gamma(dr, dv) \mu(dy) \\
= \int_{T^d \times T^d} \int_{T^d} \nabla \phi(q - \Sigma^1_t[s, \mu](y)) \nabla_\mu \Sigma^1_t[s, \mu](y, r) \cdot (v - r) \mu(dy) \gamma(dr, dv),
\]
which, after renaming variables, and adding back the term with \( \nabla_q \Sigma^1_t[s, \mu](y)(y - x) \), is seen to be the same as the integral term in (6.2). Thus, the expression in the left-hand side of (6.2) is
\[
|F(q', \sigma_{t'}) - F(q, \sigma_t) - \nabla_q F(q, \sigma_t) \cdot (q' - q) - B_2| \\
\leq |B_1| + |F(q', \sigma_{t'}) - F(q, \sigma_t) - \nabla_q F(q, \sigma_t) \cdot (q' - q) - B|.
\]
The bound on \( B_1 \) and (6.4) now yield (6.2). □

The next lemma addresses the differentiability of \( g \) in the \( \mu \) variable, and is proved in the same way, so we will omit the proof. Define
\[
(\nabla_\mu)(g(q, \Sigma^1_0[s, \mu])\mu)(y) := -\nabla U^1(q - \Sigma^1_0[s, \mu](y)) \nabla \Sigma^1_0[s, \mu](y) \\
- \int_{T^d} \nabla U^1(q - \Sigma^1_0[s, \mu](u)) \nabla \Sigma^1_0[s, \mu](u, y) \mu(du).
\]
Lemma 6.2. The function \((s, q, \mu) \mapsto g(q, \sigma_0^1[s, \mu]_{\#}\mu)\) is continuously differentiable, meaning that, with a constant \(C\) having a larger value than before,

\[
|u(q', \sigma_0') - F(q, \sigma_0) - \nabla_q F(q, \sigma_0) \cdot (q' - q) - \int_{T^d \times T^d} (\nabla_\mu)(g(q, \sigma_0))(y) \cdot (x - y) \gamma(dy, dx)| \leq C(|q' - q|^2 + \mathcal{W}^2(\mu, \nu)),
\]

for any \(q, q' \in T^d, \gamma \in \Gamma_0(\mu, \nu), s, t, t' \in [0, T]\), with the function \((\nabla_\mu)(g(q, \Sigma_1^1[s, \mu]_{\#}\mu))(y)\) being continuous in all its variables.

6.2 Gradient of \(u(s, q, \cdot)\) and chain rule

We collect now the results on differentiability in \(\mu\) of the functions \(g, F, H(q, \nu)\), that go into the definition of \(u\), with the following definition and corollary. Define the \(\mathbb{R}^d\)-valued function

\[
\Upsilon[s, \mu](q, y) := (\nabla_\mu)(g(q, \sigma_0))(y) + \int_0^s \{\nabla_\mu H(q, \nu_t[s, \mu](q, y)) + (\nabla_\mu)(F(q, \sigma_t))(y)\} dt,
\]

where \(s \in [0, T], q \in T^d, y \in \mathbb{R}^d, \mu \in \mathcal{P}(T^d), \sigma_t = \Sigma_1^1[s, \mu]_{\#}\mu\).

Corollary 6.3. The function \(\Upsilon\), just defined, is continuous on \([0, T] \times T^d \times T^d \times \mathcal{P}(T^d)\), and \(u(s, q, \cdot)\), defined by \((4.1)\), is differentiable on \(\mathcal{P}(T^d)\), in the sense that there exists a constant \(C\) such that

\[
|u(s, q, \nu) - u(s, q, \mu) - \int_{T^d \times T^d} \Upsilon[s, \mu](q, y) \cdot (x - y) \gamma(dy, dx)| \leq C\mathcal{W}^2(\mu, \nu)
\]

for every \(\mu, \nu \in \mathcal{P}(T^d), \gamma \in \Gamma_0(\mu, \nu), s \in [0, T], q \in T^d\).

Proof. The continuity of \(\Upsilon\) is a consequence of the continuity of its parts, and combining Lemma with Lemmas 6.1 and 6.2 produces the stated estimate. \(\Box\)

We refer back to Section 2.2 for the definition of \(\mathcal{F}_\mu \mathcal{P}(T^d)\). Since we do not know whether \(\Upsilon[s, \mu](q, \cdot)\) belongs to the \(L^2(\mu)\) closure of \(\{\nabla \varphi \mid \varphi \in C_c^\infty(T^d)\}\), we make the following definition.

Definition 6.4. Let \(u = u(s, q, \mu)\) be as in \((6.7)\), for \(s \in [0, T], q \in T^d, \mu \in \mathcal{P}(T^d), \) and \(\Upsilon\) be as in \((6.6)\). At every \(s, q, \mu\), by

\[
\nabla_\mu u(s, q, \mu)
\]

we will mean the projection of \(\Upsilon[s, \mu](q, \cdot)\) onto \(\mathcal{F}_\mu \mathcal{P}(T^d)\).

We need to note that the velocity vector fields \(v[s, \mu](t, \cdot)\) are not necessarily elements of \(\mathcal{F}_\sigma \mathcal{P}(T^d)\), even though this is true in the case \(H(q, p) = \frac{1}{2}|p|^2\) (see (25) Theorem 5.1]). This leads to the following definition.

Definition 6.5. At every \(s, t \in [0, T], \mu \in \mathcal{P}(T^d), \) by

\[
\vec{v}[s, \mu](t, \cdot)
\]

we will mean the projection of \(v[s, \mu](t, \cdot)\) onto \(\mathcal{F}_\sigma \mathcal{P}(T^d), \) where \(\sigma_t = \Sigma_1^1[s, \mu]_{\#}\mu\).
Note that, if \( w \in L^2(\mathbb{T}^d, \mu) \) is arbitrary and \( \bar{w} \) is its projection onto \( \mathcal{T}_\mu \mathcal{P}(\mathbb{T}^d) \), then
\[
\int_{\mathbb{T}^d} Y[s, \mu](q, x) \cdot \bar{w}(x) \mu(dx) = \int_{\mathbb{T}^d} \nabla \mu u(s, q, \mu)(x) \cdot w(x) \mu(dx).
\] (6.7)

We are now ready to prove the third main statement of the paper:

**Theorem 6.6.** Let \( H, F, g \) be as in Section 2.1 with \( F \) and \( g \) having the specific convolution forms \( \text{[5.1]}, \text{[5.2]} \). Let \( u = u(s, q, \mu) \) be defined as in \( \text{[4.41]} \). Then:

(i) For any \( s \in [0, T] \), \( \mu \in \mathcal{P}(\mathbb{T}^d) \), there exists \( \sigma \in AC^2(0, T; \mathcal{P}(\mathbb{T}^d)) \) such that \( \sigma_s = \mu \) and the continuity equation
\[
\partial_t \sigma_t + \nabla \cdot (\sigma_t \nabla_H(q, \nabla_u(t, q, \sigma_t)) = 0 \quad \text{in } \mathcal{D}'((0, T) \times \mathcal{P}(\mathbb{T}^d))
\]
holds;

(ii) The function \( u \) is a classical solution to the master equation \( \text{[1.1]} \) in the sense explained in Section 2.1.1.

The full value function \( u = u(s, q, \mu) \) is the value of the solution \( U \) of the MFG system’s Hamilton-Jacobi equation \( \text{[1.2]} \) at the time \( t = s \) at which the terminal condition \( \sigma_{t=s} = \mu \) is prescribed for the continuity equation \( \text{[1.3]} \).

**Proof.** (i) Let \( s \in [0, T] \), \( \mu \in \mathcal{P}(\mathbb{T}^d) \). Set \( \sigma_t := \Sigma^1_t[s, \mu] \neq \mu \). Then the statement follows from Proposition 4.18, Corollary 4.23, formula \( \text{[4.29]} \) and Lemma 4.27.

(ii) The regularity of \( u \) in \( q \) is the same as the regularity of \( U \) in \( q \), which was discussed in Lemma 4.22. Fix \( 0 < s < T, q \in \mathbb{T}^d, \mu \in \mathcal{P}(\mathbb{T}^d) \). As usual, \( \sigma_t = \Sigma^1_t[s, \mu] \neq \mu \), and \( v_t = \partial_t \Sigma^1_t[s, \mu] \circ X_t[s, \mu], 0 \leq t \leq T \). Set
\[
\hat{\sigma}_t := (id + (t-s)\bar{v}_s) \neq \mu, \quad \hat{\sigma}_t := (id + (t-s)\bar{v}_s) \neq \mu,
\]
where \( \bar{v}_s \) is the projection of \( \nu_s \) to \( \mathcal{T}_\mu \mathcal{P}(\mathbb{T}^d) \). Through \( \sigma_{s+h} \times \hat{\sigma}_{s+h} \), we estimate \( \mathcal{W}(\sigma_{s+h}, \hat{\sigma}_{s+h}) \):
\[
\mathcal{W}^2(\sigma_{s+h}, \hat{\sigma}_{s+h}) \leq \int_{T^d \times T^d} |x-y|^2 (\sigma_{s+h} \times \hat{\sigma}_{s+h})(dx, dy)
\]
\[
= \int_{T^d} |\Sigma^1_{s+h}[s, \mu](y) - \Sigma^1_s[s, \mu](y) - hv_s[s, \mu](y)|^2 \mu(dy).
\]
Note that \( \nu_s[s, \mu](q) = \partial_t \Sigma^1_t[s, \mu](q) \big|_{t=s} \), since \( X_s[s, \mu] = id \). Therefore,
\[
\mathcal{W}(\sigma_{s+h}, \hat{\sigma}_{s+h}) \leq |h|^2 \| \partial_t \Sigma^1 \|^2_{\infty}.
\] (6.8)

Let
\[
\gamma_h := (id \times (id + hv_s)) \neq \mu \in \Gamma(\mu, \hat{\sigma}_{s+h}).
\]
Since, by definition, \( \nabla \mu u(s + h, q, \mu) \in \mathcal{T}_\mu \mathcal{P}(\mathbb{T}^d) \), we apply Lemma 2.4 to write
\[
|u(s + h, q, \hat{\sigma}_{s+h}) - u(s + h, q, \mu) - \int_{T^d \times T^d} \nabla \mu u(s + h, q, \mu)(x) \cdot (y - x) \gamma_h(dx, dy)| = o(\|\pi - \pi^1\|_{\gamma_h}),
\]
48
which is the same as

\[ |u(s + h, q, \hat{\sigma}_{s+h}) - u(s + h, q, \mu)| = o(|h|), \quad (6.9) \]

because \( o(\|\pi^2 - \pi^1\|^2_{\gamma_{0}}) = o(|h|) \) as can be easily checked. Recall now (6.7). Formula (6.6) shows that \( \Psi[\cdot, \mu](q, y) \) is continuous, so there is a modulus of continuity \( \omega \) such that

\[
\int_{T^d} \nabla_{\mu} u(s + h, q, \mu)(x) \cdot v_s(x) \mu(dx) = \int_{T^d} \Psi[s + h, \mu](q, x) \cdot \bar{v}_s(x) \mu(dx) \\
= \int_{T^d} \Psi[s, \mu](q, x) \cdot \bar{v}_s(x) \mu(dx) + \omega(|h|) \\
= \int_{T^d} \nabla_{\mu} u(s, q, \mu)(x) \cdot v_s(x) \mu(dx) + \omega(|h|).
\]

Therefore

\[
|u(s + h, q, \hat{\sigma}_{s+h}) - u(s + h, q, \mu)| = o(|h|) + |h|\omega(|h|). \quad (6.10)
\]

Corollary 6.3 shows that \( u(s, q, \cdot) \) is \( \kappa_1 \)-Lipschitz for some constant \( \kappa_1 \), because \([0, T] \times T^d \) is compact. Using the bound (6.8), we then have

\[
|u(s + h, q, \hat{\sigma}_{s+h}) - u(s + h, q, \sigma_{s+h})| \leq \kappa_1 h^2 \|\bar{\sigma}_{\mu}^{2} \Sigma_{1}^{2}\|_{\infty}^{2}. \quad (6.11)
\]

Invoking (4.47), we write

\[
|u(s + h, q, \sigma_{s+h}) - u(s, q, \mu) + h[H(q, \nabla_q u(s, q, \sigma_s)) + F(q, \sigma_s)]| = o(|h|). \quad (6.12)
\]

Finally, (6.10), (6.11) and (6.12) are needed to obtain:

\[
|u(s + h, q, \mu) - u(s, q, \mu) + h\int_{T^d} \nabla_{\mu} u(s, q, \mu)(x) \cdot v_s(x) \mu(dx) + h[H(q, \nabla_q u(s, q, \sigma_s)) + F(q, \sigma_s)]| \\
= |u(s + h, q, \mu) - u(s + h, q, \hat{\sigma}_{s+h}) + h\int_{T^d} \nabla_{\mu} u(s, q, \mu)(x) \cdot v_s(x) \mu(dx) \\
+ u(s + h, q, \hat{\sigma}_{s+h}) - u(s + h, q, \sigma_{s+h}) \\
+ u(s + h, q, \sigma_{s+h}) - u(s, q, \mu) + h[H(q, \nabla_q u(s, q, \sigma_s)) + F(q, \sigma_s)]| \\
= o(|h|) + |h|\omega(|h|) + \kappa_1 h^2 \|\bar{\sigma}_{\mu}^{2} \Sigma_{1}^{2}\|_{\infty}^{2} + o(|h|) = o(|h|). \quad (6.13)
\]

We divide by \( h \), remember that \( v_s(x) = \nabla_p H(q, \nabla_q u(s, q, \mu)) \), \( \mu = \sigma_s \) and let \( h \to 0 \) to obtain

\[-\partial_s u(s, q, \mu) = \int_{T^d} \nabla_{\mu} u(s, q, \mu)(x) \cdot \nabla_p H(q, \nabla_q u(s, q, \mu)) \mu(dx) + H(q, \nabla_q u(s, q, \mu)) + F(q, \mu).\]

Let us check the continuity of \( s \mapsto \partial_s u(s, q, \mu) \). Due to (4.46), \( \nabla_q u(s, q, \mu) = \Psi[s, \mu](s, q) = \Sigma^2[s, \mu](s, q) \), which is continuous in \( s \), and the continuity of \( H \) and \( F \) takes care of the non-integral term in the formula.
for $\partial_s u$. For the integral term, we use once again (6.7). Let $s' \in (0, T)$. Then

$$\left| \int_{T^d} \mathcal{Y}[s, \mu](q, x) \cdot \bar{v}_s(x) \mu(dx) - \int_{T^d} \mathcal{Y}[s', \mu](q, x) \cdot \bar{v}_{s'}(x) \mu(dx) \right|
$$

$$\leq \left| \int_{T^d} \mathcal{Y}[s, \mu](q, x) \cdot \bar{v}_s(x) \mu(dx) \right| - \left| \int_{T^d} \mathcal{Y}[s', \mu](q, x) \cdot \bar{v}_{s'}(x) \mu(dx) \right|
$$

$$+ \left| \int_{T^d} \mathcal{Y}[s, \mu](q, x) \cdot \bar{v}_s(x) \mu(dx) - \int_{T^d} \mathcal{Y}[s, \mu](q, x) \cdot \bar{v}_s(x) \mu(dx) \right|
$$

$$\leq \left\| \mathcal{Y}[s, \mu](q, \cdot) \right\|_{L^2(\mu)} \left\| \bar{v}_s - \bar{v}_{s'} \right\|_{L^2(\mu)} + \left\| \mathcal{Y}[s, \mu](q, \cdot) - \mathcal{Y}[s', \mu](q, \cdot) \right\|_{L^2(\mu)} \left\| \bar{v}_{s'} \right\|_{L^2(\mu)}.$$

By the fact that $\bar{v}_s, \bar{v}_{s'}$ are the projections of $v_s, v_{s'}$ on a subspace of $L^2(\mu)$, we know that $\left\| \bar{v}_s - \bar{v}_{s'} \right\|_{L^2(\mu)} \leq \left\| v_s - v_{s'} \right\|_{L^2(\mu)}$. Letting $s' \to s$ we conclude the continuity. The continuity of $\partial_s u(s, \cdot, \mu)$ is treated in the same fashion, since $v[s, \mu](s, \cdot)$ is continuous. This completes the proof.

**Remark 6.7.** We do not claim that the function $\nabla_\mu u(s, q, \cdot)$ is continuous on $P^d$, which is true in the case $\left\| H(q, u) \right\| \leq \frac{1}{2} |q|^2$. The reason is that we have had to define $\nabla_\mu u$ as the projection of a vector field (Definition 6.4) that, in general, is not in the tangent space $T_{\mu} \mathcal{P}(\mathbb{R}^d)$, whereas for the quadratic Hamiltonian, $\mathcal{Y}[s, \mu](q, \cdot)$ and $\nabla_\mu u(s, q, \mu)$ are the same.

### 7 Acknowledgements

The work presented here would not have been possible at all without the careful, patient and diligent advise of my advisors, Professors Wilfrid Gangbo and Andrzej Święch, who taught me the main ideas and methods involved in it, and provided me with frequent and valuable feedback.

### References

[1] D. M. Ambrose, Small strong solutions for time-dependent mean field games with local coupling, C. R. Acad. Sci. Paris, Ser. I 354 (2016) 589-594.

[2] D. M. Ambrose, Strong solutions for time-dependent mean field games with non-separable Hamiltonians, J. Math. Pures Appl. 113 (2018) 141-154.

[3] L. Ambrosio, N. Gigli, G. Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures, Birkhäuser, 2nd ed., 2008.

[4] Y. Averbukh, Minimax approach to first-order mean field games, arXiv:1312.6627v2 [math.AP], 2014.

[5] J-D. Benamou, G. Carlier, F. Santambrogio, Variational mean field games. In: Active Particles, Volume 1, Birkhäuser, 2017.

[6] A. Bensoussan, J. Frehse, P. Yam, Mean Field Games and Mean Field Type Control Theory, Springer, 2013.

[7] A. Bensoussan, J. Frehse, P. Yam, The master equation in mean field theory, J. Math. Pures Appl. 103 (2015) 1441-1474.

[8] A. Bensoussan, J. Frehse, P. Yam, On the interpretation of the master equation, Stochastic Processes and their Applications 127 (2017) 2093-2137.
[9] U. Bessi, Existence of solutions of the master equation in the smooth case, SIAM J. MATH. ANAL., Vol. 48, No. 1, pp. 204-228.

[10] V. Bogachev, Measure Theory, Vol. 2, Springer, 2007.

[11] P. Cannarsa, R. Capuani, Existence and uniqueness for mean field games with state constraints, arXiv:1711.01063v2 [math.OC], 2017.

[12] P. Cardaliaguet, Notes on Mean Field Games, 2012, https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf

[13] P. Cardaliaguet, P. J. Graber, Mean field games systems of first order, ESAIM: COCV Vol 21, 3, p. 690-722, 2015.

[14] P. Cardaliaguet, A. Porretta, D. Tonon, Sobolev regularity for the first order Hamilton-Jacobi equation, Calc. Var. (2015) 54: 3037.

[15] P. Cardaliaguet, Long time average of first order mean field games and weak KAM theory, P. Dyn Games Appl (2013) 3: 473.

[16] R. Carmona, F. Delarue, Probabilistic Theory of Mean Field Games with Applications I and II, Springer, 2018.

[17] R. Carmona, F. Delarue, The master equation for large population equilibriums. In: Crisan D., Hambly B., Zariphopoulou T. (eds) Stochastic Analysis and Applications 2014. Springer Proceedings in Mathematics & Statistics, vol 100, Springer.

[18] P. Cardaliaguet, F. Delarue, J. M. Lasry, P. L. Lions, The Master Equation and the Convergence Problem in Mean Field Games, Princeton University Press, forthcoming August 2019.

[19] J-F. Chassagneux, D. Crisan, F. Delarue, A probabilistic approach to classical solutions of the master equation for large population equilibria, arXiv:1411.3009v2 [math.PR], 2014.

[20] M. G. Crandall, P. L. Lions, Hamilton-Jacobi equations in infinite dimensions I. Uniqueness of viscosity solutions, J. Funct. Anal. 62 (1985), 379-396.

[21] M. G. Crandall, P. L. Lions, Hamilton-Jacobi equations in infinite dimensions III, J. Funct. Anal. 68 (1986), 214-247.

[22] I. Fonseca, W. Gangbo, Degree Theory in Analysis and its Applications, GMT, Oxford University Press, 1995.

[23] W. Gangbo, On some analytical aspects of Mean Field Games, Notes of Berkeley lectures, 2018, https://math.berkeley.edu/~wgangbo/math_275/MFG-1-18.pdf

[24] W. Gangbo, T. Nguyen, A. Tudorascu, Hamilton-Jacobi equations in the Wasserstein space, Methods Appl. Anal. 15 (2) (2008) 155-183.

[25] W. Gangbo, A. Święch, Existence of a solution to an equation arising from the theory of mean field games, Journal of Differential Equations, 259 (2015), 6573-6643.

[26] W. Gangbo, A. Tudorascu, Weak KAM theory on the Wasserstein torus with multidimensional underlying space, Communications on Pure and Applied Mathematics, Vol. LXVII (2014), 0408-0463.

[27] W. Gangbo, A. Tudorascu, On differentiability in the Wasserstein space and well-posedness for Hamilton-Jacobi Equations, to appear in J. Math. Pures Appl.
[28] N. Ghoussoub, Optimal ballistic transport and Hopf-Lax formulae on Wasserstein space, arXiv:1705.05951v2 [math.AP], 2017.

[29] D. Gomes, J. Saúde, Mean field games models—a brief survey, Dyn Games Appl (2014) 4:110-154.

[30] D. Gomes, E. Pimentel, V. Voskanyan, Regularity Theory for Mean-Field Game Systems, Springer, 2016.

[31] P. J. Graber, A. Mészáros, Sobolev regularity for first order mean field games, Ann. I. H. Poincaré – AN 35 (2018) 1557-1576.

[32] O. Guéant, J. M. Lasry, P. L. Lions, Mean field games and applications, in: Paris-Princeton Lectures on Mathematical Finance 2010, Lecture Notes in Mathematics, vol 2003, Springer.

[33] M. Huang, P.E. Caines, R. P. Malhamé, Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized -Nash equilibria, IEEE Trans. Automat. Control 52 (9) (2007), 1560-1571.

[34] D. Lacker, A general characterization of the mean field limit for stochastic differential games, Probab. Theory Relat. Fields (2016) 165: 581.

[35] H. Lavenant, F. Santambrogio, Optimal density evolution with congestion: $L^\infty$ bounds via flow interchange techniques and applications to variational mean field games, arXiv:1705.05658v1 [math.AP], 2017.

[36] J. M. Lasry, P. L. Lions, Mean field games, Jpn. J. Math. 2 (2007), 229-260.

[37] P. L. Lions, Cours au Collège de France. [http://www.college-de-france.fr](http://www.college-de-france.fr)

[38] S. Mayorga, Short time Solution to the Master Equation of a First Order Mean Field Game System (doctoral dissertation), School of Mathematics, Georgia Institute of Technology, to be available in the Spring of 2019.

[39] W. Rudin, Principles of Mathematical Analysis, Third Edition, Mc-Graw Hill 1976.xs.

[40] F. Santambrogio, Regularity via duality in calculus of variations and degenerate elliptic PDEs, Journal of Mathematical Analysis and Applications, J. Math. Anal. Appl. 457 (2) (2018) 1649-1674.