ON A CLASSICAL SOLUTION TO THE MASTER EQUATION OF A FIRST ORDER MEAN FIELD GAME

A Thesis
Presented to
The Academic Faculty

by

Sergio Mayorga

in partial fulfillment
of the requirements for the degree
Doctor of Philosophy
in the School of Mathematics

Georgia Institute of Technology
August 2019

On a classical solution to the master equation of a first order mean field game

Approved by:

Dr. Wilfrid Gangbo, Advisor
Department of Mathematics
University of California, Los Angeles

Dr. Andrzej Święch, Advisor
School of Mathematics
Georgia Institute of Technology

Dr. Yao Yao
School of Mathematics
Georgia Institute of Technology

Dr. Rafael De La Llave
School of Mathematics
Georgia Institute of Technology

Dr. Predrag Cvitanovic
School of Physics
Georgia Institute of Technology

Date Approved: May 3rd, 2019

DEDICATION

To my parents.

ACKNOWLEDGEMENTS

I want to thank:

my advisors, Professors Wilfrid Gangbo and Andrzej Święch, for their unconditional, continuous and unbounded patience and support, expert advise and encouragement;

Professor Rafael De La Llave for his interest, encouragement and good advise;

Professor Ionel Popescu, for his wisdom and warm disposition;

the defense committee members Professors Yao Yao and Predrag Cvitanovic for devoting their time and offering input;

the Georgia Tech School of Mathematics for believing in me and allowing me to serve it in the course of my doctorate program;

my friends and colleagues, thanks to whom these five years have been less lonely;

Elizaveta, for everything.

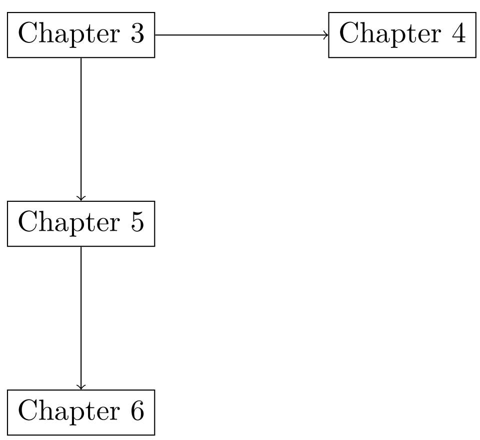
The research presented in this thesis was also partially supported by AFOSR MURI FA9550-18-1-0502.

Contents

SUMMARY

For a first order (deterministic) mean-field game with nonlocal couplings, a classical solution is constructed for the associated, so-called master equation, a partial differential equation in infinite-dimensional space with a nonlocal term, assuming the time horizon is sufficiently small and the coefficients are smooth enough, without convexity conditions on the individual Hamiltonian. The couplings, albeit smooth, are not assumed to derive from a potential, which makes the result currently the most general one for short time horizon. The approach to obtain the master equation is inspired by that of Gangbo and Święch [GŚ15] for the problem in which the Hamiltonian is quadratic and the couplings derive from a potential, but we use a non-variational method and require further results of the calculus on the Wasserstein space that has been advanced recently by Gangbo et. al. [GC17; Gan18].

Chapter dependence:



Chapter 1

INTRODUCTION

1.1 General background

One of the first economics models featuring an uncountable number of players was introduced by Robert Aumann in his paper [Aum64]. He argued that, under the conditions of *perfect competition*, a market economy should be modeled with a continuum of traders, rather than with a finite number of them. In broad terms, the rationale behind this is that, in the presence of a large number of players, the actions taken by a single player should be negligible, much in the same way that modifying the value of a function at a single point of its domain does not alter its Lebesgue integral. As Aumann pointed out, this idea had long been at the core of statistical and fluid mechanics, where, despite being clear to the physicist that there are finitely many particles in a gas or fluid, the aggregate effect of the massively high number of them makes the subject of study, for all practical purposes, no different from a continuum. This, in turn, forces the particle elements of the fluid to be treated as indistinguishable, a concept which, as we shall see, features in a fundamental way in the theory of mean field games (in shorthand: MFG). Furthermore, the negligibility of a single player's actions, mentioned above, must have the logical consequence that its choice be dictated not by the decision of any particular peer, but by the *whole* of them.

These ideas were picked up by Guilherme Carmona and others (see [Car04] and its references) in the early 2000s, and then, in a series of papers (see [LL07] and its references) by Pierre Lions and Jean-Michel Lasry, formulated in terms of partial differential equations (PDEs), thereby essentially founding MFG as a proper branch of mathematical research which at its core encompasses notions from deterministic and stochastic control theory, convex analysis, PDEs and, as this thesis tries to explain, optimal transport theory.

We will now move on to a more explicit description of some of the key concepts in MFG theory.

1.2 N-player games

Even though this thesis is not about game theory, we find it necessary to introduce some of the terminology that is ubiquitous in the MFG literature. This will also help pave the way towards a more concrete discussion of the problem that concerns us. The following discussion, however, is not fully rigorous, but is intended to motivate interest in the problem and provide some context. Every game involves a set of *players* or *agents*, who, according to the previous paragraph, can be represented by points in some d-dimensional manifold, which, in this work, is going to be the

d -dimensional torus $\mathbb{T}^d$. There is a set of *strategies*, or *controls*, available to the players and a *cost* function, assigning a number to each player that depends on his strategy and, in general, on the states or strategies of the other players. Each player, naturally, wants to minimize the cost associated with his strategy, and, given the choice between two strategies, he will act according to the one that leads to the least cost (this is what is known in game theory as the *rationality* of the agents). Let us fix some $s > 0$ and agree to denote the positions of N players at time $t = s$ on the torus by $q_1, \ldots, q_N$. Each player, say, the j-th player, is free to choose his own velocity $\beta_j(t)$, so long as $q_j(s) = q_j$, i.e.

$$\begin{cases} \dot{q}_j(t) &= \beta_j(t), \quad 0 \leqslant t < T, \\ q_j(s) &= q_j, \end{cases} \tag{1.2.1}$$

$j = 1, \ldots, N$ and T is some fixed positive number larger than s. The velocity is the player's strategy, or control.

Problem 1. To each choice of control, let us assign the cost:

$$C(\beta_j) = \int_0^s [L(q_j(t), \beta_j(t)) - F(q_j(t), \sigma_t)]dt + g(q_j(t), \sigma_0), \tag{1.2.2}$$

where σ_t, $0 \leqslant t \leqslant s$ is a certain parameter whose meaning we don't yet specify, and L, F, g are some given functions.

Find the strategy β_j that gives the minimum cost in (1.2.2), subject to (1.2.1) and the fixed parameter σ_t, $0 \leqslant t \leqslant s$.

Note that the cost functional here $C(\beta_j)$ *is the same for every* j, given L, F, g and σ. Let us set

$$U(s, q) := \inf\{C(\beta) \mid \dot{q}(t) = \beta(t), 0 \leqslant t < T, q(s) = q.\} \tag{1.2.3}$$

If g, F are bounded and continuous, and $L = L(q, v)$ is continuous and has superlinear growth in the v variable, i.e., $L(q, v)/|v| \to \infty$ as $|v| \to \infty$ for any q, then, 1) by elementary control theory (see, e.g., [FM93; Dac08]), the solution to the Hamilton-Jacobi-Bellman equation (HJB)[1]

$$\partial_t V(t, q) + H(q, \nabla_q V(t, q)) + F(q, \sigma_t) = 0 \qquad \text{in } (0, T) \times \mathbb{T}^d, \tag{1.2.4}$$

$$V(0, q) = g(0, \sigma_0) \quad \text{on } \mathbb{T}^d, \tag{1.2.5}$$

if smooth enough, provides a minimizer for (1.2.3) in the form $\beta(t) = \nabla_p H(q(t), \nabla_q V(t, q(t)))$, where $H = L^*$, the Legendre transform of L (see Chapter 2). To be clear, this means that the solution to the ordinary differential equation (ODE)

$$\dot{q}(t) = \nabla_p H(q(t), \nabla_q V(t, q(t))), \quad 0 \leqslant t \leqslant s, \quad q(s) = q, \tag{1.2.6}$$

determines the minimizer in (1.2.3). This is known as a *feedback control.* Moreover, 2) $V(s, q)$ is the minimum in (1.2.3), that is,

$$V(s, q) = U(s, q).$$

[1] Note that in our setting time is reversed with respect to the traditional control theory formulation that gives rise to a HJB with terminal value condition, instead of initial value.

In such a set-up, then, all the players follow the same rule, prescribed by the solution to (1.2.4-1.2.5). This is, then, essentially a one-player problem, in which the player "interacts" with the parameter σ. It should be emphasized that, for this very reason, the solution to this game (i.e., "solve (1.2.4-1.2.5)") is independent of N. Now let us modify Problem 1.

Problem 2. Consider, instead, the cost

$$\begin{aligned} J_j(\beta_j; \{\beta_k\}_{k\neq j}) &= g^0(q_j(0)) + \frac{1}{N-1}\sum_{k\neq j} g^1(q_j(0)-q_k(0)) \\ &\quad + \int_0^s [L(q_j(t),\beta_j(t)) - \frac{1}{N-1}\sum_{k\neq j}\phi(q_k(t)-q_j(t))]dt, \end{aligned} \tag{1.2.7}$$

where ϕ and g^1 are even functions, $j = 1, \dots, N$. Now the cost for the j-th player depends on the other players' strategies, or, equivalently, on the other players' states, because of (1.2.1). The notation $J_j(\beta_j; \{\beta_k\}_{k\neq j})$ is justified because the order of the β_k with $k \neq j$ is irrelevant. This means that for the j-th player, all the other players are *indistinguishable*. In this game, each player tries to minimize his cost, still subject to the condition (1.2.1). Expecting that all players can minimize their cost simultaneously is too much. More precisely: in general, there is no collection $(\hat\beta_1, \dots, \hat\beta_N)$ such that, for each $j \in \{1, \dots, N\}$, $(\hat\beta_j, \{\hat\beta_k\}_{k\neq j})$ is a minimum of J_j over all possible collections of strategies. One is hopeful for another type of settlement, that of *Nash equilibrium*: the collection $(\hat\beta_1, \dots, \hat\beta_N)$ constitutes a Nash equilibrium if

$$\text{for every } j \in \{1,\dots,N\}: \quad J_j(\hat\beta_j; \{\hat\beta_k\}_{k\neq j}) \leqslant J_j(\beta_j; \{\hat\beta_k\}_{k\neq j}) \quad \text{for all } \beta_j.$$

Thus, what characterizes a Nash equilibrium is that an agent cannot lower his cost by choosing another strategy while all the other players' strategies remain the same. Such an equilibrium is similar to a saddle point, rather than a global minimizer.

Find a Nash equilibrium for the game played with costs (1.2.7), subject to (1.2.1).

Looking back at Problem 1, we may ask ourselves whether Problem 2 has an associated differential equation or a system of equations, that would allow one to synthesize a Nash equilibrium. This is indeed the case, and has been known for a long time (see, e.g., [Fri75; BF83], where this is discussed in the stochastic setting): it is the *Nash system*

$$\begin{aligned} &\partial_t v^{N,j}(s,\boldsymbol{q}) + H(q_j, \nabla_{q_j} v^{N,j}(s,\boldsymbol{q})) \\ &\quad + \sum_{j\neq k} \nabla_p H(q_k, \nabla_{q_k} v^{N,k}(s,\boldsymbol{q}))\cdot\nabla_{q_k} v^{N,j}(s,\boldsymbol{q}) = -F^{N,j}(\boldsymbol{q}) \quad \text{in } (0,T)\times(\mathbb{T}^d)^N, \end{aligned} \tag{1.2.8}$$

$$v^{N,j}(0,\boldsymbol{q}) = g^{N,j}(\boldsymbol{q}) \quad \text{in } (\mathbb{T}^d)^N, \tag{1.2.9}$$

where $\boldsymbol{q} = (q_1, \dots, q_N)$, $F^{N,j}(\boldsymbol{q}) = \frac{1}{N-1}\sum_{k\neq j}\phi(q_k - q_j)$, and $g^{N,j}$ is analogously defined, $j = 1, \dots, N$.

Claim. Fix $(q_1, \dots, q_N)$ and suppose that a smooth solution $(v^{N,j})_{j=1}^N$ of (1.2.8-1.2.9) exists. For each $j \in \{1, \dots, N\}$, set $\hat\beta_j(t) = \dot{\hat q}_j(t)$, where $\hat q_j(t)$ solves

$$\dot{\hat q}_j(t) = \nabla_p H(\hat q_j(t), \nabla_{q_j} v^{j,N}(t, \hat{\boldsymbol{q}}(t))), \quad 0 \leqslant t \leqslant s, \quad \hat q_j(s) = q_j.$$

Then $(\hat\beta_1, \dots, \hat\beta_N)$ is a Nash equilibrium for the game (1.2.7).

Proof. Fix $j \in \{1, \dots, N\}$, let β_j be arbitrary and put $q_j(t) = q_j + \int_s^t \beta_j(\tau)d\tau$. Let $\hat{\boldsymbol{q}}_{-j}(t)$ the result of replacing the j-th component of the vector $\hat{\boldsymbol{q}}(t)$ by $q_j(t)$. We will drop the superscript N from v in the following computation. We have, at $0 < t \leqslant s$,

$$\frac{d}{dt}\big[v^j(t, \hat{\boldsymbol{q}}_{-j}(t)) - \int_0^t [L(q_j(\tau), \beta_j(\tau)) - \frac{1}{N-1}\sum_{k\neq j}\phi(\hat q_k(\tau) - q_j(\tau))]d\tau\big] =$$

$$= \partial_t v^j(t, \hat{\boldsymbol{q}}_{-j}(t)) + \nabla_{q_j} v^j(t, \hat{\boldsymbol{q}}_{-j}(t)) \cdot \beta_j(t) + \sum_{k\neq j} \nabla_{q_k} v^j(t, \hat{\boldsymbol{q}}_{-j}(t)) \cdot \hat\beta_k(t)$$

$$- [L(q_j(t), \beta_j(t)) - \frac{1}{N-1}\sum_{k\neq j}\phi(q_j(t) - \hat q_k(t))]$$

$$\leqslant \partial_t v^j(t, \hat{\boldsymbol{q}}_{-j}(t)) + H(q_j(t), \nabla_{q_j} v^j(t, \hat{\boldsymbol{q}}_{-j}(t))) + \sum_{k\neq j} \nabla_{q_k} v^j(t, \hat{\boldsymbol{q}}_{-j}(t)) \cdot \hat\beta_k(t) + \frac{1}{N-1}\sum_{k\neq j}\phi(q_j(t) - \hat q_k(t))]$$

$$= \partial_t v^j(t, \hat{\boldsymbol{q}}_{-j}(t)) + H(q_j(t), \nabla_{q_j} v^j(t, \hat{\boldsymbol{q}}_{-j}(t)))$$
$$+ \sum_{k\neq j} \nabla_{q_k} v^j(t, \hat{\boldsymbol{q}}_{-j}(t)) \cdot \nabla_p H(\hat q_k(t), \nabla_{q_k} v^j(t, \hat{\boldsymbol{q}}_{-j}(t))) + F^{N,j}(\hat{\boldsymbol{q}}_{-j}(t))$$
$$= 0,$$

where the inequality is due to the fact that $H = L^*$ and the last equality is due to v^j being a solution to the Nash system. Taking into account, furthermore, the initial condition (1.2.9), and $\hat{\boldsymbol{q}}_{-j}(s) = q$, integrating from $t = 0$ to $t = s$ we are led to:

$$v^{j,N}(s, \boldsymbol{q}) \leqslant J_j(\beta_j; \{\hat\beta_k\}_{k\neq j}),$$

and, if we repeat the computation with $\hat{\boldsymbol{q}}(t)$ instead of $\hat{\boldsymbol{q}}_{-j}(t)$, we obtain the equality to 0 (see the section on Legendre transform in the Preliminaries). Thus,

$$J_j(\hat\beta_j; \{\hat\beta_k\}_{k\neq j}) \leqslant J_j(\beta_j; \{\hat\beta_k\}_{k\neq j}),$$

and j was arbitrary, so this proves the claim. QED.

1.3 A continuum of players

However, conditions are not known under which the system (1.2.8-1.2.9) would have a solution, let alone a smooth one, despite the fact that the foregoing statement shows that this should be the correct system to find Nash equilibria of the game (1.2.2). Nevertheless, if N is very large, one should expect, and this is one of the key ideas in MFG theory, that a continuum version of Problem 2, in some sense the limit as $N \to \infty$, would provide a good approximation to the game with large N, and be solvable. Heuristically, we can proceed as follows. Let $q_j = q$, the position of the j-th player at time $t = s$, and $\beta_j = \beta$ his strategy. At this point we will switch to *Eulerian coordinates* for the velocity, meaning that the strategy β of the player who starts at q at time $t = s$ is equivalently described by $v_{(\ )}(q)$:

$$\beta(t) = v_t(q), \quad \text{i.e.} \quad \begin{cases} q(s) = q, \\ \dot q(t) = v_t(q(t)). \end{cases}$$

The fact that g^1 and ϕ are even functions allows us to rewrite (1.2.7) as

$$J(v_{(\,)}(q), \sigma^{(N)}) = g(q(0), \sigma_0^{(N)}) + \int_0^s [L(q(t), v_t(q(t)) - F(q(t), \sigma_t^{(N)})]dt,$$

where $\sigma_t^{(N)} = \frac{1}{N-1} \sum_{k \neq j} \delta_{q_k(t)}$, $0 \leqslant t \leqslant s$, and $\delta_{q_k(t)}$ is the Dirac point mass at $q_k(t)$. We can relate this now back to expression (1.2.2) and say that the domain of the functions g and F is $\mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d)$, where $\mathscr{P}(\mathbb{T}^d)$ is the space of Borel probability measures on $\mathbb{T}^d$. Formally, then, in the limit, his cost functional becomes

$$J(v_{(\,)}(q), \sigma) = g(q(0), \sigma_0) + \int_0^s [L(q(t), v_t(q(t))) - F(q(t), \sigma_t)]dt, \tag{1.3.1}$$

where σ is a path in $\mathscr{P}(\mathbb{T}^d)$ and describes the evolution of the mass distribution of *all* the agents (we normalize the total mass to be 1, so that it can be represented by a probability measure): indeed, if the measures σ_t have no atoms, the "distribution of all the other agents (except $q(t)$)" is the same as "the distribution of all the agents". This leads us to:

Problem 3. Suppose there is a continuum of players on $\mathbb{T}^d$, and the distribution of players at time $t = s$ is $\mu \in \mathscr{P}(\mathbb{T}^d)$. Each player seeks to minimize the cost (1.3.1) of his trajectory $q(\cdot)$, and this cost depends on the evolution of the distribution σ_t of the players. If each single player takes σ_t as given, then he can resort to the HJB (1.2.4-1.2.5) and move with the velocity given by (1.2.6). However, the parameter σ_t will now only make sense as the distribution of players if, upon taking the same course of action, the resulting players' trajectories will have precisely the distribution σ_t. This means that, if each player solves his minimization by moving with velocity given by (1.2.6), the evolution of the agents' distribution, $\tilde{\sigma}_t$, should be given by

$$\partial_t \tilde{\sigma}_t + \operatorname{div}(\nabla_p H(q, \nabla_q U(t, q))\sigma_t) = 0.$$

Therefore, we seek solutions to the (first order) **mean field game system**

$$\begin{cases} \partial_t U(t,q) + H(q, \nabla_q U(t,q)) + F(q, \sigma_t) = 0 \quad \text{in } (0,s) \times \mathbb{T}^d, & (1.3.2) \\ \partial_t \sigma_t + \operatorname{div}(\nabla_p H(q, \nabla_q U)\sigma_t) = 0 \quad \text{in } \mathcal{D}'((0,s) \times \mathbb{T}^d), & (1.3.3) \\ U(0, \cdot) = g(\cdot, \sigma_0), & (1.3.4) \\ \sigma_s = \mu. & (1.3.5) \end{cases}$$

Since it is not immediately clear what type of equilibrium is achieved by solving the MFG system, we will just say that Problem 3 is:

Solve the MFG system (1.3.2-1.3.5).

In Section 6.3, after our work constructing a solution to the master equation is done, we will explain the question that is answered by a solution of (1.3.2-1.3.5), in terms of the game (1.2.2).

To attempt at a clearer picture of what happens when $N \to \infty$, one should also heuristically pass to the limit as $N \to \infty$ in the Nash system (1.2.8-1.2.9). The reader may refer to Section 7 of [Car12] or the Introduction of [Car+19] for this. The idea is that for any fixed $q_j = q$, the functions

$v^{j,N}(t,q,\{q_k\}_{k\neq j})$ approach a function u defined on $[0,T)\times\mathbb{T}^d\times\mathscr{P}(\mathbb{T}^d)$, $u=u(t,q,\mu)$, which should satisfy the so called **master equation** (ME, for short) of *first order* mean field games, namely,

$$\begin{cases} \partial_s u(s,q,\mu) + \displaystyle\int_{\mathbb{T}^d} \nabla_\mu u(s,q,\mu)(x)\cdot\nabla_p H(x,\nabla_q u(s,x,\mu))\mu(dx) \\ \qquad + H(q,\nabla_q u(s,q,\mu)) + F(q,\mu) = 0 \qquad \text{in } (0,T)\times\mathbb{T}^d\times\mathscr{P}(\mathbb{T}^d), \\ \qquad\qquad u(0,q,\mu) = g(q,\mu) \qquad \text{on } \mathbb{T}^d\times\mathscr{P}(\mathbb{T}^d). \end{cases} \tag{1.3.6}$$

The meaning of $\nabla_\mu u$ is explained in the Preliminaries. We will show, in Section 6.3, that substituting an average of Dirac masses for μ in (1.3.6) yields the first order Nash system (1.2.8-1.2.9).

Before we say more about (1.3.6) and (1.3.2-1.3.5) in the next section, we need to say a few words about the second-order case, as this introduction cannot be complete without it.

Second order MFGs

Second order, or *stochastic*, mean field games, come about when we allow noise in the dynamics of the particles; for instance, instead of (1.2.1) we may have the SDE (stochastic differential equation)

$$\begin{cases} dX_t = b(t,X_t)dt + \sqrt{2\epsilon}\, dB_t, \\ X_0 = X, \end{cases}$$

where $(B_t)_{t\geqslant 0}$ is a standard d-dimensional Brownian motion on some probability space $(\Omega,\mathcal{F},\mathbb{P})$, $\epsilon>0$, and $X\in L^1(\Omega)$ is a random variable independent of the Brownian motion. In this case, the cost function (1.3.1) should be changed to

$$J(b_{(\,)}(X),\sigma) = \mathbb{E}\Big[g(X_T,\sigma_T) + \int_0^T [L(X_\tau,b_\tau(X_\tau)) - F(X_\tau,\sigma_\tau)]dt\Big],$$

The associated mean field game system is

$$\begin{cases} -\epsilon\Delta U - \partial_t U(t,q) + H(q,\nabla_q U(t,q)) + F(q,\sigma_t) = 0 \quad \text{in } (0,T)\times\mathbb{T}^d, \\ -\epsilon\Delta\sigma_t + \partial_t\sigma_t - \operatorname{div}(\nabla_p H(q,\nabla_q U)\sigma_t) = 0 \quad \text{in } \mathcal{D}'([0,T]\times\mathbb{T}^d), \\ \sigma_0 = \mu, \\ U(T,q) = g(q,\sigma_T). \end{cases} \tag{1.3.7}$$

and the second order master equation[2] is

$$\begin{cases} -\Delta_q u(t,q,\mu) - \partial_t u(t,q,\mu) + \displaystyle\int_{\mathbb{T}^d} \nabla_\mu u(t,q,\mu)(x)\cdot\nabla_p H(x,\nabla_q u(t,x,\mu))\mu(dx) \\ \quad - \displaystyle\int_{\mathbb{T}^d} \operatorname{div}_x \nabla_\mu u(t,q,\mu)(x)\mu(dx) + H(q,\nabla_q u(t,q,\mu)) + F(q,\mu) = 0 \quad \text{in } (0,T)\times\mathbb{T}^d\times\mathscr{P}(\mathbb{T}^d), \\ \qquad u(T,q,\mu) = g(q,\mu) \qquad \text{on } \mathbb{T}^d\times\mathscr{P}(\mathbb{T}^d). \end{cases}$$

[2]More precisely, this is called the master equation without common noise. We warn the reader that in [Car+19], the authors call this one a first order master equation because the highest derivative with respect to the measure argument is still of order 1.

1.4 Prior knowledge

Since its simultaneous introduction by Lasry and Lions [LL07] and, in the engineering community, by Huang, Caines and Malhamé [HCM07], the system (1.3.7), and similar second order systems, have been extensively investigated and conditions for existence and uniqueness are well understood. Existence of smooth (classical) solutions over arbitrary time horizons (i.e., no restrictions on T) is typically guaranteed by regularity of the couplings in both spatial and measure variables, convexity of $H(q,p)$ in the p variable, and a quadratic growth condition on $H(q,\cdot)$. Uniqueness is obtained under the condition of monotonicity of the couplings:

$$\int_{\mathbb{T}^d}(F(q,\mu)-F(q,\mu'))(\mu-\mu')(dq) \geqslant 0,\ \int_{\mathbb{T}^d}(g(q,\mu)-g(q,\mu'))(\mu-\mu')(dq) \geqslant 0 \quad \forall\ \mu,\mu' \in \mathscr{P}(\mathbb{T}^d).$$

One of the reasons that MFG systems have garnered so much interest in the mathematical community is that they presented a new challenge in the form of a coupled system of a forward HJB and a backward continuity equation (Fokker-Plack equation). First-order HJB equations on arbitrary time horizons generally call for the notion of weak solutions. Moreover, in most works on MFG systems, the measure of the terminal[3] condition ($\sigma_s = \mu$) is usually assumed to be absolutely continuous with respect to Lebesgue measure. Over the years since their inception in [LL07], everal significant modifications and refinements have been obtained, e.g., existence and uniqueness of weak solutions in the case of local couplings [CG15], first-order [CPT15; LS17; San18] and higher [GM18] Sobolev estimates of such solutions, with different growth conditions of the Hamiltonian H, whose convexity in the momentum variable is always required, and growth conditions on the couplings linked with those of H; all the approaches in these developments work for arbitrary time horizons. A full book on regularity theory for mean field games has been written [GPV16], while the recently published two volumes by R. Carmona and F. Delarue [CD18] constitute the most complete reference work on mean field game theory.

This is a good moment to break the only bad news to the reader, namely, that

the results of this thesis are valid not for arbitrary time T, but for a sufficiently small T, in a way that depends on the coefficients H, F, g.

(Alternatively, we could keep T arbitrary but impose smallness conditions on the coefficients.) In the course of our work towards the first order master equation (ME), we obtain classical solutions for small T of the system (1.3.2-1.3.5), with the conditions on the coefficients H, F, g stated in Section 2.2. From a theoretical point of view, the master equation presents a stronger challenge because the equation itself is written in an infinite dimensional space, which is "not flat", inviting techniques from non-linear analysis. The body of work published on MFG master equations is considerably thinner than that on MFG systems. Furthermore, many models are cast as MFG systems, but it is not clear what the corresponding master equations are, or how to solve them. Most available literature on mean field game MEs has dealt with the higher-order variants [BFY15; BFY17; CD18; CD14]. The recent paper by Pierre Cardaliaguet et al. [Car+19] includes proofs of existence and uniqueness of classical solutions for the master equations of second-order MFGs,

[3] We remind the reader that most works in MFGs are reversed in time with respect to ours, so they would say initial when we say terminal.

such as (1.3.2-1.3.5) and rigorous characterizations of master equations as limits of (second order) N-player Nash systems as $N \to \infty$.

1.5 Our contribution

As for the first-order master equation (1.3.6), the first major step was achieved by Wilfrid Gangbo and Andrzej Święch in [GŚ15], where short-time strong solutions to both the MFG system (1.3.2-1.3.5) and the ME are obtained for $H(q,p) = |p|^2/2$ and regularizing couplings (i.e., smooth enough with respect to both variables) that derive from a potential. The same result was later re-proved by Bessi [Bes16] using the language of random variables. In [May18], we keep the smoothness assumptions of the Hamiltonian and the couplings, but we do away with the convexity of H in the variable p (and, in particular, no growth condition is imposed). The main appeal of the Gangbo-Swiech approach is that it is intrinsic, working in the metric space of probability measures, which seems like a natural setting for mean-field theory, since, for example, an average of n Dirac masses is the same measure regardless of the order in which the point masses are taken, reflecting the interchangeability of players that is a core assumption in MFGs.

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 [GŚ15]. Given H, F, g as in Section 2.2.1, and given $0 < s < T$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, 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 Hamiltonian system

$$\partial_t \Sigma^1 = \nabla_p H(\Sigma^1, \Sigma^2), \qquad \partial_t \Sigma^2 = -\nabla_q H(\Sigma^1, \Sigma^2) - \nabla_q F(\Sigma^1, \Sigma^1_{\#}\mu) \tag{1.5.1}$$

with initial and terminal conditions

$$\Sigma^2(0,q) = \nabla_q g(\Sigma^1(0,q), \Sigma^1(0,\cdot)_{\#}\mu), \qquad \Sigma^1(s,q) = q,$$

providing us with a path in $\mathscr{P}(\mathbb{T}^d)$ 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[4]

$$\nabla_q U(t, \Sigma^1_t) = \Sigma^2_t. \tag{1.5.2}$$

On the other hand, we have that the velocity vector v_t driving the path σ_t satisfies

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

and the first equation in (1.5.1) implies

$$\nabla_p H(q, \nabla_q U(t, \Sigma^1_t)) = \partial_t \Sigma^1_t. \tag{1.5.3}$$

Comparing (1.5.3) and (1.5.2), we see that they are the same if $H(q,p) = \frac{1}{2}|p|^2$, which is the Hamiltonian in [GŚ15], and, indeed, in that case, $\partial_t \Sigma^1 = \Sigma^2$, with $\nabla_q U(t,\cdot)$ coinciding with the

[4] We follow the convention, common in this field, of using the subindex t to mean "at time t", and thus a shorthand for $(t, \dots)$ rather than the time derivate.

velocity v_t (*a posteriori* from (1.5.2)). Thus, the function Σ^2 is not present in [GŚ15], with $\partial_t\Sigma^1$ taking its place, while the relationship $\nabla_q U(t,\Sigma^1_t) = \partial_t \Sigma^1_t$ 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 (σ, v) is the unique minimizer[5] of

$$\mathcal{U}(s,\mu) = \inf_{(\sigma,v)} \Big\{ \int_0^s \int_{\mathbb{T}^d} \big(L(q,v_t(q)) - \mathcal{F}(\sigma_t)\sigma_t(dq)\big)dt + \mathcal{G}(\sigma_0) \;\Big|\; \sigma_s = \mu,\ \sigma \in AC^2(0,s;\mathscr{P}(\mathbb{T}^d))\Big\}, \tag{1.5.4}$$

where $\mathcal{F}, \mathcal{G} : \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ are functions whose Wasserstein gradients are F and g, the minimality of the norm of v_t then follows, leading to the symmetry of $\nabla_q v_t(q)$, which is used to establish the Hamilton-Jacobi equation in (1.3.2-1.3.5). In the case of a Hamiltonian with superlinear, polynomial growth (see condition 2.2.2), even though it is still true, as Chapter 4 will show, that the pair (σ, v) is the unique minimizer of (1.5.4), it is no longer clear how this approach can give us (1.5.2); if the growth condition is given up, then the minimality of (σ, v) is no longer known.

We encourage the reader to refer to the beginning of Chapter 4 for a more detailed explanation about why we had to take a different approach and how it resulted in a good generalization.

We turned, instead, to a more direct procedure (Lemma 5.0.2) that also helps to shed further light on how the equations (3.1.3) are the *characteristics* of (1.3.2-1.3.5). This optic allows us to present a sort of uniqueness counterpart (Theorem 5.0.5) to the existence result, namely, that if a solution $(\tilde{U}, \tilde{\sigma})$ is in $W^{2,3;\infty}((0,T)\times\mathbb{T}^d) \times AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$, then it must coincide, at least for a shorter time T, with the pair (U,σ) constructed from (Σ^1,Σ^2). With this approach we manage to circumvent the specific potential forms for F and g. In [May18], these assumptions are added for the theorem on the master equation, with the purpose of working out the differentiability of Σ in μ through the same discretization approach used in [GŚ15], a short description of which we have included in Section 1.5.1. It turns out, however, that it is not necessary to assume that the couplings are convolutions or that they derive from a potential. We instead provide some general, albeit strong, regularity assumptions on F and g that also lead to the regularity of Σ in the measure variable, and under which we still obtain the required first order Taylor estimates of the composite functions that enter the representation formula for u in (6.0.1). This is an additional improvement to [May18] and is first published in this thesis.

1.5.1 Summary of main results

We collect here the main statements of the thesis. Let H, F, g be as in Section 2.2.

Statement 1. *(Theorem 5.0.4) If T is sufficiently small, in a way that depends only on the coefficients (H,F,g), then, for every $0 < s < T$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, the MFG system (1.3.2-1.3.5) admits a classical solution (U,σ), in the sense of Section 2.3. Moreover, $(U,\sigma) \in W^{2,2;\infty}((0,T)\times\mathbb{T}^d) \times AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$.*

Statement 2. *(Theorem 5.0.5) If $(\tilde{U}, \tilde{\sigma}) \in W^{2,3;\infty}((0,T)\times\mathbb{T}^d) \times AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ is a classical solution to the MFG system (1.3.2-1.3.5), 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 5.0.4.*

[5] See also [Gho17] for a recent connection between value functionals such as (1.5.4) and Hopf-Lax formulae on the Wasserstein space.

Statement 3. *(Theorem 6.2.7) Let H, F, g be as in Sections 2.2.1 and 2.2.2. If T is small enough, in a way that depends only on the coefficients (H, F, g), then the master equation (1.3.6) admits a classical solution in the sense of Section 2.3.*

The following statement was proven before all the previous ones, and is not necessary for them, but at the beginning of Chapter 4 we explain why we have included it.

Statement 4. *(Theorem 4.1.4 and Lemma 4.2.3) Let H, F, g be as in Sections 2.2.1 and 2.2.3). Let $\mathcal{U}$ be defined as in (1.5.4), and $\sigma_t = (\Sigma^1_t)_{\#}\mu$, $v_t = \partial_t \Sigma^1 \circ (\Sigma^1_t)^{-1}$, where Σ is the solution to the Hamiltonian system (3.1.3). Then the pair (σ, v) is the unique minimizer for $\mathcal{U}(s, \mu)$, and*

$$\nabla_\mu \mathcal{U}(s, \mu) = \nabla_q U(s, q),$$

where U is the solution to (1.3.2,1.3.4).

The discretization approach Let $\Sigma^1[s, \mu]$, $\Sigma^2[s, \mu]$ be as in (1.5.1), i.e., the pair is a solution to (3.1.3) below. For an average of Dirac masses $\mu^x := \frac{1}{n}\sum_{j=1}^n \delta_{x_j}$, $x \in (\mathbb{T}^d)^n$, let $M(t, s, q, x) = \Sigma[s, \mu^x](t, q)$ (see Definition 6.1.1); *note that the number of coordinates in the domain of M depends on n, the number of particles.* Some calculations give us (Section 6.1.1) that

$$\|\nabla^2_{qx_j} M\|_\infty = \mathcal{O}(\frac{1}{n}), \quad \|\nabla^2_{x_j x_j} M\|_\infty = \mathcal{O}(\frac{1}{n}), \quad \text{while} \quad \|\nabla^2_{x_i x_j} M\|_\infty = \mathcal{O}(\frac{1}{n^2}) \text{ if } i \neq j.$$

This is precisely what is needed to obtain formula (6.1.21), namely,

$$n|\nabla_{x_j} M(t, s, q, y) - \nabla_{x_i} M(t, s, q, x)| \leqslant \sqrt{d} C(|y_j - x_i|_{\mathbb{T}^d} + \mathscr{W}(\mu^x, \mu^y) + \frac{1}{n}), \tag{6.1.21}$$

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 $\mathscr{P}(\mathbb{T}^d)$ (Corollary 6.1.9), becoming the Wasserstein gradient of Σ in μ (Lemma 6.1.10).

This thesis is organized as follows. In Chapter 2, we present the theoretical framework, notation, and technical facts that will be needed in the main body. We introduce the various assumptions on the coefficients of the equations for the results, and define the notions of classical solution for the MFG system and the ME. In Chapter 3, we solve the Hamiltonian system that can be interpreted, in a sense, to be the characteristics of the MFG system. Chapter 4 is not necessary to obtain either the solution to the MFG system of the master equation, but it is interesting in its own right, and its first paragraph clarifies how we managed to end up with a non-variational result for the first order mean field game. In Chapter 5 we used the solution to the Hamiltonian system to construct a solution to the MFG system. Finally, Chapter 6 is devoted to the regularity estimates and chain rules that lead to the final theorem, Theorem 6.2.7 on the existence of a classical solution to the master equation.

Chapter 2

PRELIMINARIES

In this chapter we will make a brief presentation of optimal transport, list a few facts (some with proof) that will be used along the way, and then present the conditions on the coefficients of the equations that will be held in some or all parts of the work. Finally, we will explain what we mean by classical solutions to the MFG system and to the master equation.

2.1 Theoretical framework

Since the theory of optimal transport is not commanded by every expert in the field of partial differential equations, or, more generally, by every student or professional with an interest in continuum differential games, we deem it convenient to set forth the bare minimum of the principles of optimal transport that are necessary for this thesis. Among the general references on this theory are [AGS08] and [San15].

2.1.1 Optimal transport basics

The set of Borel probability measures on the Euclidean space $\mathbb{R}^d$ is denoted by $\mathscr{P}(\mathbb{R}^d)$. For $p \geqslant 1$, we define

$$\mathscr{P}_p(\mathbb{R}^d) := \{\mu \in \mathscr{P}(\mathbb{R}^d) \mid \int_{\mathbb{R}^d} |x|^p \mu(dx) < \infty\},$$

the set of probability measures with finite p-moments. Given two measures $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$, which may describe, say, two different mass configurations in the space $\mathbb{R}^d$, the starting question of optimal transport theory is how to "transport" the mass μ to ν with "minimal cost": a *transport map* of μ to ν means a Borel map $S : \mathbb{R}^d \mapsto \mathbb{R}^d$ such that $S_{\#}\mu = \nu$, i.e.,

$$\nu(B) = \mu(S^{-1}(B)), \text{ for every Borel } B \subset \mathbb{R}^d,$$

equivalently,

$$\int_{\mathbb{R}^d} f(y)\nu(dy) = \int_{R^d} f \circ S(x)\mu(dx) \qquad \forall f \in L^1_\mu(\mathbb{R}^d; \mathbb{R}),$$

and the associated cost of the transport is

$$c_p^M(S)(\mu, \nu) := \int_{\mathbb{R}^d} |x - S(x)|^p \mu(dx);$$

the M stands for Monge, who was the first to formulate this problem. Even for the simple case $\mu = \delta_x$, $\nu = \frac{1}{2}\delta_{y_1} + \frac{1}{2}\delta_{y_2}$, where $x, y_1, y_2 \in \mathbb{R}^d$, there is no transport map of μ to ν. A relaxation of this problem is achieved by allowing the mass to "split". One seeks then a *transport plan* of μ to ν, which is a measure γ on the product $\mathbb{R}^d \times \mathbb{R}^d$, such that

$$\gamma(A \times \mathbb{R}^d) = \mu(A), \quad \gamma(\mathbb{R}^d \times B) = \nu(B),$$

for all Borel subsets $A, B \subset \mathbb{R}^d$, that minimizes the cost:

$$c_p^K(\gamma)(\mu, \nu) := \int_{R^d} |x - y|^p \gamma(dx, dy),$$

where the K now stands for Kantorovich, who came up with this idea. A minimizing γ for

$$W_p(\mu, \nu) := \inf\{c_p^K(\gamma)(\mu, \nu) \mid \gamma \in \Gamma(\mu, \nu)\},$$

where $\Gamma(\mu, \nu)$ is the set of all transport plans of μ to ν, always exists (see, e.g., [AGS08, Chapter 6]). Moreover, W_p, thus defined, constitutes a distance on $\mathscr{P}_p(\mathbb{R}^d)$, making $\mathscr{P}_p(\mathbb{R}^d)$ a metric space.

Since our MFG and master equations are written in the torus $\mathbb{T}^d$ and the space of probability measures on it, $\mathscr{P}(\mathbb{T}^d)$, we need to explain what we mean by this. However, 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 [GT14]. The set $\mathbb{T}^d$ is the set of equivalence classes on $\mathbb{R}^d$ with respect to the equivalence relation:

$$x \sim y \qquad \text{iff} \qquad \text{there exist integers } n_1, \dots, n_d \text{ such that } x^{(j)} - y^{(j)} = n_j, \ j = 1, \dots, d$$

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 μ, ν are Borel probability measures on $\mathbb{R}^d$, we already said what an optimal plan of μ to ν is, but the following notation is most common: $\pi^1_{\#}\gamma = \mu$ and $\pi^2_{\#}\gamma = \nu$, where $\pi^1, \pi^2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ are the first and second coordinate projections, respectively, where the subindex $_{\#}$ again stands for the pushforward operation. For $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$, $p > 1$, we define

$$\mathscr{W}_p(\mu, \nu) = \left(\inf_{\gamma \in \Gamma(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p_{\mathbb{T}^d} \gamma(dx, dy) \right)^{1/p}, \tag{2.1.1}$$

and let $\Gamma_0(\mu, \nu)$ denote the set of optimal transport plans γ between μ and ν, i.e. those for which the infimum in (2.1.1) is attained (this set, of course, depends on p, but this won't cause any problems for us). We will denote

$$\mathscr{W} := \mathscr{W}_2.$$

With the equivalence relation

$$\mu \sim \nu \qquad \text{iff} \qquad \int_{\mathbb{R}^d} \phi d\mu = \int_{\mathbb{R}^d} \phi d\nu \ \text{ for all } \ \phi \in C(\mathbb{T}^d)$$

on $\mathscr{P}_p(\mathbb{R}^d)$, where $C(\mathbb{T}^d)$ are all real-valued continuous functions ϕ on $\mathbb{R}^d$ such that $\phi(x) = \phi(x')$ whenever $x \sim x'$, it is true that $\mathscr{W}(\mu,\nu) = \mathscr{W}(\mu',\nu')$ whenever $\mu \sim \mu'$ and $\nu \sim \nu'$. In this way, $\mathscr{W}_p$ in formula (2.1.1) is defined on the set of equivalence classes, which we henceforth denote by $\mathscr{P}(\mathbb{T}^d)$. Moreover, $\mathscr{W}_p$ is a metric on $\mathscr{P}(\mathbb{T}^d)$, with respect to which $\mathscr{P}(\mathbb{T}^d)$ is compact. There is no ambiguity in the absence of a subindex in $\mathscr{P}(\mathbb{T}^d)$ because it can be shown that, unlike in the space of probability measures on $\mathbb{R}^d$, all the metrics $\mathscr{W}_p$, $p \geqslant 1$, on the space of probability measures on $\mathbb{T}^d$ are equivalent. It should be further said that the only portion of this thesis where we will allow $p \neq 2$ is in Chapter 4, which is independent of the rest.

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 $\mathcal{F} : \mathscr{P}(\mathbb{T}^d) \to S$ is a function $\mathcal{F} : \mathscr{P}(\mathbb{R}^d) \to S$ that takes constant values on the equivalence classes of $\mathscr{P}(\mathbb{T}^d)$. 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 $\mu^x \in \mathscr{P}_2(\mathbb{R}^d)$ denotes the measure $\mu^x = \frac{1}{n}\sum_{j=1}^n \delta_{x_j}$. Such measures are called averages of Dirac masses.

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

$$W_2(f_{\#}\mu, g_{\#}\mu) \leqslant \|f - g\|_{L^2(\mathbb{R}^d,\mu)}. \tag{2.1.2}$$

Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. Then $L^2(\mathbb{T}^d,\mu)$ denotes the completion of $C(\mathbb{T}^d)$ with respect to the $L^2(\mathbb{R}^d,\mu)$ norm: $L^2(\mathbb{T}^d,\mu) = \overline{C(\mathbb{T}^d)}^{L^2(\mathbb{R}^d,\mu)}$. At the same time, we define the *tangent space to* $\mathscr{P}(\mathbb{T}^d)$ *at* μ, $\mathscr{T}_\mu\mathscr{P}(\mathbb{T}^d)$, 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 $\mathbb{R}^d$:

$$\mathscr{T}_\mu\mathscr{P}(\mathbb{T}^d) := \overline{\nabla C^\infty(\mathbb{T}^d;\mathbb{R})}^{L^2(\mathbb{R}^d,\mu)}.$$

Naturally, since $L^2(\mathbb{T}^d,\mu)$ is a Hilbert space, if we have $\xi, \eta \in L^2(\mathbb{T}^d,\mu)$, and $\eta \in \mathscr{T}_\mu\mathscr{P}(\mathbb{T}^d)$, then

$$\int_{\mathbb{T}^d} \xi(x)\cdot\eta(x)\mu(dx) = \int_{\mathbb{T}^d} \bar{\xi}(x)\cdot\eta(x)\mu(dx), \tag{2.1.3}$$

where $\bar{\xi}$ is the projection of ξ onto $\mathscr{T}_\mu\mathscr{P}(\mathbb{T}^d)$.

Wasserstein distance between average of Dirac masses. If $\mu,\nu \in \mathscr{P}_2(\mathbb{R}^d)$ are such that $\mu = \mu^x = \frac{1}{n}\sum_{j=1}^n \delta_{x_j}$ and $\nu = \mu^y = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$, 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

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

This fact will be used quite often in this work. For instance, if we consider $\gamma \in \Gamma_0(\mu^x,\mu^y)$, where μ^x and μ^y are as above, and we want to bound a certain expression by another that involves $\mathscr{W}(\mu^x,\mu^y)$, thanks to this fact we can always assume that the coordinates of $x = (x_1,\ldots,x_n)$ and $y = (y_1,\ldots,y_n)$ are already sorted in such a way that

$$\gamma = \frac{1}{n}\sum_{j=1}^n \delta_{(x_j,y_j)},$$

because, certainly, the measures μ^x and μ^y do not change regardless of the order in which these coordinates are taken. The following fact will also be crucial.

Density of average of Dirac masses in $\mathscr{P}(\mathbb{T}^d)$. Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. As it is well known (see, for instance, [Bog07, Ex. 8.1.6]), the set of average of Dirac masses is dense in $\mathscr{P}_2(\mathbb{R}^d)$ with respect to narrow convergence. In $\mathscr{P}(\mathbb{T}^d)$, this convergence coincides with convergence in $\mathscr{W}$. Thus, there exists a sequence $\{\mu(n)\}_1^\infty \subset \mathscr{P}_2(\mathbb{R}^d)$, with $\mu(n) = \frac{1}{n}\sum_{j=1}^n \delta_{x_j(n)}$, an average of Dirac masses, such that $\mathscr{W}(\mu, \mu(n)) \to 0$ as $n \to \infty$. Moreover, this sequence can be chosen so that each $x_j(n) \in \text{supp}(\mu)$, where $\text{supp}(\mu)$ is the support of the measure μ.

2.1.2 Absolutely continuous paths in the Wasserstein space

We denote by $AC^2(0,T;\mathscr{P}_p(\mathbb{T}^d))$ the set of all paths μ_t in $\mathscr{P}_p(\mathbb{T}^d)$ for which there exists $m \in L^2(0,T)$ such that

$$\mathscr{W}_p(\mu_{t_1}, \mu_{t_2}) \leqslant \int_{t_1}^{t_2} m(\tau)d\tau$$

whenever $0 < t_1 \leqslant t_2 < T$. We say that a time-dependent velocity vector field $v_t : \mathbb{T}^d \to \mathbb{R}^d$ is a velocity vector field for the absolutely continuous path μ_t if $v_t \in L^p(\mathbb{T}^d, \mu)$,

$$\int_0^T \int_{\mathbb{T}^d} |v_t(q)|\mu_t(dq)dt < \infty$$

and the *continuity equation* is true:

$$\partial_t \mu_t + \text{div}(v_t\mu_t) = 0 \quad \text{in } \mathcal{D}'((0,T) \times \mathbb{T}^d),$$

which means that for every $\varphi \in C_c^\infty((0,T) \times \mathbb{T}^d)$:

$$\int_0^T \int_{\mathbb{R}^d} [\partial_t\varphi(t,q) + \nabla\varphi(t,q) \cdot v_t(q))]\mu_t(dq)dt = 0.$$

A path μ_t in $\mathscr{P}(\mathbb{T}^d)$, $0 \leqslant t \leqslant 1$, is said to be a constant-speed geodesic if

$$\mathscr{W}(\mu_{t_1}, \mu_{t_2}) = |t_2 - t_1|\mathscr{W}(\mu_0, \mu_1), \quad t_1, t_2 \in [0,1].$$

Given a path μ_t in $\mathscr{P}(\mathbb{T}^d)$, a velocity v_t for μ_t is in $L^2(\mathbb{T}^d;\mu)$ but it may or may not be in $\mathscr{T}_{\mu_t}\mathscr{P}(\mathbb{T}^d)$. However, if μ_t is an $AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ path, a velocity field of minimal $L^2(\mathbb{T}^d,\mu)$-norm always exists, and it belongs to $\mathscr{T}_{\mu_t}\mathscr{P}(\mathbb{T}^d)$. This is the content of [AGS08, Theorem 8.3.1].

Remark 2.1.1. Let $\mu, \nu \in \mathscr{P}(\mathbb{T}^d)$, $\gamma \in \Gamma_0(\mu,\nu)$. For each $0 \leqslant \tau \leqslant 1$, let

$$\mu^\tau := [(1-\tau)\pi^1 + \tau\pi^2]_\#\gamma.$$

Let w^τ, $0 \leqslant \tau \leqslant 1$, be the velocity vector field of minimal norm for μ^τ.

(i) For $0 \leqslant \tau \leqslant 1$, $\|w^\tau\|_{L^2(\mu^\tau)} = \mathscr{W}(\mu,\nu)$.

(ii) For every $f \in C(\mathbb{T}^d;\mathbb{R}^d)$, every $0 \leqslant \tau \leqslant 1$,

$$\int\limits_{\mathbb{T}^d\times\mathbb{T}^d} f((1-\tau)x+\tau y) \cdot w^\tau((1-\tau)x+\tau y)\gamma(dx,dy) = \int\limits_{\mathbb{T}^d\times\mathbb{T}^d} f((1-\tau)x+\tau y)\cdot(y-x)\gamma(dx,dy).$$

(iii) Furthermore, fix $\tau \in (0,1)$, and let $\gamma^\tau \in \Gamma_0(\mu, \mu^\tau)$. Then, for every $f_1, f_2 \in C(\mathbb{T}^d; \mathbb{R}^d)$,

$$\int_{\mathbb{T}^d \times \mathbb{T}^d} [f_2(y) \cdot w^\tau(y) - f_1(x) \cdot w^0(x)] \gamma^\tau(dx, dy) = \int_{\mathbb{T}^d \times \mathbb{T}^d} [f_2(y) - f_1(x)] \cdot \frac{y - x}{\tau} \gamma^\tau(dx, dy).$$

Proof. The third fact is justified as follows. The optimal plans along geodesics are characterized by [AGS08, Chapter 7.2]:

$$\int_{\mathbb{T}^d \times \mathbb{T}^d} B(x, y) \gamma^\tau(dx, dy) = \int_{\mathbb{T}^d \times \mathbb{T}^d} B(x, (1-\tau)x + \tau z) \gamma(dx, dz), \tag{2.1.4}$$

$0 < \tau < 1, B \in C(\mathbb{T}^d \times \mathbb{T}^d)$. Therefore, the left-hand side of (iii) is equal to

$$\begin{aligned} &\int_{\mathbb{T}^d \times \mathbb{T}^d} [f_2((1-\tau)x + \tau z) \cdot w^\tau((1-\tau)x + \tau z) - f_1(x) \cdot w^0(x)] \gamma(dx, dz) \\ \overset{\text{(ii)}}{=} &\int_{\mathbb{T}^d \times \mathbb{T}^d} [f_2((1-\tau)x + \tau z) - f_1(x)] \cdot (z - x) \gamma(dx, dz) \\ = &\int_{\mathbb{T}^d \times \mathbb{T}^d} [f_2((1-\tau)x + \tau z) - f_1(x)] \cdot \frac{1}{\tau}((1-\tau)x + \tau z - x) \gamma(dx, dz) \\ \overset{(2.1.4)}{=} &\int_{\mathbb{T}^d \times \mathbb{T}^d} [f_2(y) - f_1(x)] \cdot \frac{y - x}{\tau} \gamma^\tau(dx, dy), \end{aligned}$$

which is the right-hand side of (iii). □

2.1.3 Differentiability in the Wasserstein space

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

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

We have chosen to present this section with a notation similar to the one found in the paper [GT18], which unifies the different notions of differentiability on $\mathscr{P}_2(\mathbb{R}^d)$ used in the literature. If $r > 0$, set

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

and

$$e^0[\xi, r] = \sup_{\gamma \in \Gamma_0(\mu, \nu)} \sup_{\nu \in \mathscr{P}_2(\mathbb{R}^d)} \left\{ \frac{|e(\nu, \xi, \gamma)|}{\|\pi^1 - \pi^2\|_\gamma} \mid \|\pi^1 - \pi^2\|_\gamma \leqslant r \right\}.$$

Definition 2.1.2. *With the preceding notation, we say that $\mathcal{W}$ is differentiable at μ if*

$$\lim_{r \to 0^+} e^0[\xi, r] = 0. \tag{2.1.5}$$

The set of all $\xi \in L^2(\mathbb{T}^d, \mu)$ for which (2.1.5) holds is denoted $\partial \mathcal{W}(\mu)$.

Lemma 2.1.3. *If $\xi \in \partial \mathcal{W}(\mu)$, then so is its projection $\bar{\xi}$ onto $\bar{\xi}$ onto $\mathscr{T}_\mu \mathscr{P}(\mathbb{T}^d)$, which is then the unique element of minimal $L^2(\mathbb{T}^d, \mu)$-norm in $\partial \mathcal{W}(\mu)$ and is denoted by*

$$\nabla_\mu \mathcal{W}(\mu).$$

We will call it the Wasserstein gradient of $\mathcal{W}$ at μ.

Its proof can be found in [GT14].

Remark 2.1.4. The following is an alternative characterization of a vector field $\xi \in L^2(\mathbb{T}^d, \mu)$ that satisfies (2.1.5):

$$\mathcal{W}(\nu) - \mathcal{W}(\mu) - \sup_{\gamma \in \Gamma_0(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \xi(x) \cdot (y-x) \gamma(dx, dy) = o(\mathscr{W}(\mu, \nu)). \tag{2.1.6}$$

Likewise, ξ satisfies

$$\lim_{r \to 0^+} e[\xi, r] = 0$$

if and only if

$$\mathcal{W}(\nu) - \mathcal{W}(\mu) - \sup_{\gamma \in \Gamma(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \xi(x) \cdot (y-x) \gamma(dx, dy) = o(\mathscr{W}(\mu, \nu)). \qquad /\!/$$

The following lemma will be used in the final section.

Lemma 2.1.5. *With the foregoing notation, if $\mathcal{W} : \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ is differentiable at μ, then*

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

By Remark 2.1.4, this is the same as

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

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

$$\Big| \int_{\mathbb{R}^d \times \mathbb{R}^d} \nabla \varphi(x) \cdot (y-x)(\gamma - \bar{\gamma})(dx, dy) \Big| \leqslant \frac{\|\pi^1 - \pi^2\|_\gamma^2 + \|\pi^1 - \pi^2\|_{\bar{\gamma}}^2}{2} \|\nabla^2 \varphi\|_\infty. \tag{2.1.7}$$

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 γ and then with respect to γ', remembering that γ and γ' have the same marginals μ and ν, and substracting one of the resulting expressions from the other, yields (2.1.7). Fix now $\mu \in \mathscr{P}(\mathbb{T}^d)$. Let $\nu \in \mathscr{P}(\mathbb{T}^d)$ and $\gamma \in \Gamma_0(\mu,\nu)$, $\bar{\gamma} \in \Gamma(\mu,\nu)$. 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 $\bar{\gamma}$ in place of γ. Substracting one from the other and taking absolute value gives, using Hölder's inequality,

$$\begin{aligned}|e(\nu, \nabla_\mu \mathcal{W}(\mu), \gamma) - e(\nu, \nabla_\mu \mathcal{W}(\mu), \bar{\gamma})| &\leqslant |e(\nu, \nabla\varphi, \gamma) - e(\nu, \nabla\varphi, \bar{\gamma})| \\ &\quad + \|\nabla_\mu \mathcal{W}(\mu) - \nabla\varphi\|_{L^2(\mu)} (\|\pi^2 - \pi^1\|_\gamma + \|\pi^2 - \pi^1\|_{\bar{\gamma}}).\end{aligned}$$

Now, $\|\pi^2 - \pi^1\|_\gamma \leqslant \|\pi^2 - \pi^1\|_{\bar{\gamma}}$, and, using (2.1.7),

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

Dividing by $\|\pi^2 - \pi^1\|_{\bar{\gamma}}$ and once again because $\|\pi^2 - \pi^1\|_\gamma \leqslant \|\pi^2 - \pi^1\|_{\bar{\gamma}}$, we obtain

$$\Big|\frac{e(\nu, \nabla_\mu \mathcal{W}(\mu), \bar{\gamma})}{\|\pi^1 - \pi^2\|_{\bar{\gamma}}}\Big| \leqslant \Big|\frac{e(\nu, \nabla_\mu \mathcal{W}(\mu), \gamma)}{\|\pi^1 - \pi^2\|_\gamma}\Big| + \big(\|\pi^1 - \pi^2\|_{\bar{\gamma}} \|\nabla^2\varphi\|_\infty + 2\|\nabla_\mu \mathcal{W}(\mu) - \nabla\varphi\|_{L^2(\mu)}\big).$$

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

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

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

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

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 φ, the conclusion follows. □

Twice differentiability

In [GC17], the notion of Hessian of a function on the Wasserstein space is defined. We will follow the same framework. Let ρ, ϵ be moduli of continuity, with ρ concave. We will say that a function

$$V : \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$$

is *twice differentiable* at (q,μ) if the following hold:

- the mapping $x \mapsto \nabla_\mu V(q,\nu)(x)$ exists and is differentiable for every ν in a neighbourhood of μ, with its derivative denoted by $\nabla^2_{x\mu} V(q,\nu)(x)$;

- the gradient $\nabla_q \nabla_\mu V(q,\mu)(x) =: \nabla^2_{q\mu} V(q,\mu)(x)$ exists;
- there exist a Borel, bounded matrix-valued function $A_{\mu\mu} : (\mathbb{T}^d)^3 \to \mathbb{R}^{d\times d}$ such that

$$\begin{aligned}&\sup_{\gamma\in\Gamma_0(\mu,\nu)} |\nabla_\mu V(\bar q,\nu)(y) - \nabla_\mu V(q,\mu)(x) - \nabla^2_{q\mu} V(q,\mu)(x)(\bar q - q) - P_\gamma[\mu](q,x,y)| \\ &\leqslant o(|\bar q - q|) + \big(\mathscr{W}(\mu,\nu) + |x-y|\big)\big(\rho(\mathscr{W}(\mu,\nu)) + \epsilon(|x-y|)\big),\end{aligned}$$

  where

$$P_\gamma[\mu](q,x,y) = \nabla^2_{x\mu} V(q,\mu)(x)(y-x) + \int_{\mathbb{T}^d\times\mathbb{T}^d} A_{\mu\mu}(q,x,a)(b-a)\gamma(da,db).$$

Without loss of generality, we may suppose that $A_{\mu\mu}(q,x,\cdot) \in \mathscr{T}_\mu \mathscr{P}(\mathbb{T}^d)$ for all q,x. We put

$$\nabla^2_{\mu\mu} V(q,\mu)(\cdot,\cdot) := A_{\mu\mu}(q,\cdot,\cdot), \qquad q \in \mathbb{T}^d.$$

Chain rules

Let I be an open interval, μ_t, $t \in I$ an absolutely continuous path in $\mathscr{P}(\mathbb{T}^d)$ defined on I, and v_t the velocity of minimal norm of μ_t. If the function $\mathcal{W} : \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ is differentiable at μ_{t_0}, where $t_0 \in I$, then

$$\lim_{h\to 0} \frac{\mathcal{W}(\mu_{t_0+h}) - \mathcal{W}(\mu_{t_0})}{h} = \int_{\mathbb{T}^d} \nabla_\mu \mathcal{W}(\mu_{t_0})(x)\cdot v_{t_0}(x)\mu_{t_0}(dx).$$

A simple proof can be found in [Gan18]. Due to (2.1.3), the conditions can be weakened: the formula holds if v_t is not of minimal norm, because $\nabla_\mu \mathcal{W}(\mu_{t_0}) \in \mathscr{T}_\mu \mathscr{P}(\mathbb{T}^d)$. Or, if v_t is of minimal norm, but ξ is merely in $\partial\mathcal{W}(\mu_{t_0})$, the formula also holds (with ξ in place of $\nabla_\mu \mathcal{W}$).

We are also going to need the following.

Proposition 2.1.6. *Let $V : \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ be twice differentiable, in the sense explained above. Let $h \mapsto q^h$, $h \mapsto x^h$, be differentiable paths in $\mathbb{T}^d$ defined on an interval I, and $\mu_h \in AC^2(I;\mathscr{P}(\mathbb{T}^d))$, with v_h a continuous in h velocity vector field for μ_h.*

(i) There exists a set $J \subset I$, of equal measure to that of I, such that, if $h_0 \in I$, then the function $h \mapsto V(q^h,\mu_h)(x^h)$ is differentiable at h_0 and

$$\begin{aligned}&\frac{d}{dh}[\nabla_\mu V(q^h,\mu_h)(x^h)]\big|_{h=h_0} = \\ &\qquad = \nabla^2_{qh} V(q^{h_0},\mu_{h_0})(x^{h_0})(q^h)'|_{h=h_0} + \nabla^2_{xh} V(q^{h_0}),\mu_{h_0})(x^{h_0})(x^h)'|_{h=h_0} \\ &\qquad\quad + \int_{\mathbb{T}^d} \nabla^2_{\mu\mu} V(q^{h_0},\mu_{h_0})(x^{h_0},r)v_{h_0}(r)\mu_{h_0}(dr).\end{aligned}$$

(ii) If $\nabla^2_{q\mu}V$, $\nabla^2_{x\mu}V$, $\nabla^2_{\mu\mu}V$ are continuous, and the paths $h \mapsto x^h$, $h \mapsto q^h$ are in $C^1(I)$, then

$$\nabla_\mu V(q^b,\mu_b)(x^b) - \nabla_\mu V(q^a,\mu_a)(x^a) = \int_a^b \frac{d}{dh}\nabla_\mu V(q^h,\mu_h)(x^h)dh$$

for any $a,b \in I$.

Proof. *(i)* Let us invoke Proposition 8.4.6 of [AGS08], to say that there exists a subset $J \in I$ whose measure equals that of I, such that, for every $V \in C(\mathbb{T}^d \times \mathbb{T}^d)$, $h_0 \in J$, we have

$$\lim_{h \to 0} \int_{\mathbb{T}^d \times \mathbb{T}^d} V\big(x, \frac{y-x}{h}\big)\gamma_h(dx, dy) = \int_{\mathbb{T}^d} V(x, \bar{v}_{h_0}(x))\mu_{h_0}(dx), \tag{2.1.8}$$

where $\bar{v}_{h_0}$ is the velocity vector field of minimal norm for μ_h at h_0, and $\{\gamma_h\}_{|h|>0}$ are optimal plans between μ_{h_0} and μ_{h_0+h}. Let then $h_0 \in J$. For h such that $h_0 + h \in I$, let $\gamma_h \in \Gamma_0(\mu_{h_0}, \mu_{h_0+h})$. By the twice differentiability of V, we have

$$\begin{aligned} &\big|\nabla_\mu V(q^{h_0+h}, \mu_{h_0+h})(x^{h_0+h}) - \nabla_\mu V(q^{h_0}, \mu_{h_0})(x^{h_0}) \\ &\qquad - \nabla^2_{q\mu} V(q^{h_0}, \mu_{h_0})(x^{h_0})(q^h - q^{h_0}) - P_\gamma[\mu_{h_0}](q^{h_0}, x^{h_0}, x^{h_0+h})\big| \\ &\leqslant o(|q^{h_0+h} - q^{h_0}|) \\ &\qquad + \big(\mathscr{W}(\mu_{h_0}, \mu_{h_0+h}) + |x^{h_0+h} - x^{h_0}|\big)\big(\rho(\mathscr{W}(\mu_{h_0}, \mu_{h_0+h})) + \epsilon(|x^{h_0+h} - x^{h_0}|)\big). \end{aligned} \tag{2.1.9}$$

Let $\bar{v}_{h_0}$ be the projection of v_{h_0} onto $\mathscr{T}_{\mu_{h_0}}\mathscr{P}(\mathbb{T}^d)$. Since

$$\nabla^2_{\mu\mu} V(q^{h_0}, \mu_{h_0})(x^{h_0}, \cdot) \in \mathscr{T}_{\mu_{h_0}}\mathscr{P}(\mathbb{T}^d),$$

(2.1.3) and (2.1.8) give us

$$\lim_{h \to 0} \int_{\mathbb{T}^d \times \mathbb{T}^d} \nabla^2_{\mu\mu} V(q^{h_0}, \mu_{h_0})(x^{h_0}, r)\frac{b-r}{h}\gamma_h(dr, dy) = \int_{\mathbb{T}^d} \nabla^2_{\mu\mu} V(q^{h_0}, \mu_{h_0})(x^{h_0}, r)v_{h_0}(r)\mu_{h_0}(dr).$$

Therefore, dividing both sides of inequality (2.1.9) by h, and passing to the limit as $h \to 0$, we obtain the desired formula.

(ii) Under those conditions, the formula for $\frac{d}{dh}\nabla_\mu V(q^h, \mu_h)(x^h)$ is continuous in h, and the claim follows. □

2.2 Assumptions on the coefficients of the equations

We call the triple (H, F, g) *the coefficients* of the mean-field game equations. The first list of conditions, 2.2.1 below, are for the entire thesis. They are enough for the two theorems of Chapter 5 —the chapter on the MFG system. The second list, 2.2.2 below, is more restrictive on F and g, implying the conditions in the first. They are sufficient for the theorem of Chapter 6 —the theorem on the solution to the master equation. The third list, 2.2.3, is meant only for Chapter 4.

2.2.1 Permanent conditions

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 $(\cdot, \cdot)$. Similarly for $\nabla_p H(\cdot, \cdot)$, and higher-order derivatives.

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

$$|\nabla_q F(q,\mu)|,\ |\nabla^2_{qq} F(q,\mu)|,\ |\nabla^3_{qqq} F(q,\mu)| \leqslant \kappa, \quad q \in \mathbb{T}^d, \mu \in \mathscr{P}(\mathbb{T}^d).$$

   Suppose, moreover, that $\nabla_q F$ is κ-Lipschitz on $\mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d)$, meaning that

$$|\nabla_q F(q_1,\mu_1) - \nabla_q F(q_2,\mu_2)| \leqslant \kappa \sqrt{|q_1 - q_2|^2 + \mathscr{W}^2(\mu_1,\mu_2)}, \quad q_1, q_2 \in \mathbb{T}^d,\ \mu_1, \mu_2 \in \mathscr{P}(\mathbb{T}^d).$$

3. Furthermore: we require that the vector field $\nabla_q F(q,\mu)$ is differentiable with respect to every μ, at every q, and

$$\nabla_\mu \nabla_q F(q,\mu)(x) =: \nabla^2_{\mu q} F(q,\mu)(x)$$

   is continuous in (q,μ,x) (hence, uniformly bounded).

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

An identical statement is assumed for g.

2.2.2 Conditions for the master equation

In addition to the previous set of conditions, here we suppose that the functions

$$(q,\mu) \mapsto F(q,\mu), \quad (q,\mu) \mapsto g(q,\mu) \quad \text{are twice differentiable}$$

in the sense explained below in Section 2.1.3, and that

$$\nabla_\mu F,\ \nabla^2_{q\mu} F,\ \nabla^2_{\mu\mu} F,\ \nabla^2_{x\mu} F \quad \text{are continuous in all its variables}$$

(and, therefore, uniformly bounded). We suppose an identical statement holds for g.

2.2.3 Conditions for Chapter 4

We will assume there that

$$\nabla^2_{pp} H(q,p) > 0, \tag{2.2.1}$$

i.e., $\nabla^2_{pp} H(q,p)$ is strictly positive-definite, for all $q \in \mathbb{T}^d$, $p \in \mathbb{R}^d$, and that there exist constants $c_0 > 0$, $c_1 > 0$, $r > 1$, such that

$$c_0^{-1}|p|^r - c_1 \leqslant H(q,p) \leqslant c_0(|p|^r + 1). \tag{2.2.2}$$

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

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

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

$$F(q,\mu) = \phi * \mu(q), \quad \mathcal{F}(\mu) = \int_{\mathbb{T}^d} \frac{1}{2} \phi * \mu(y)\mu(dy),$$

so that

$$\nabla_\mu \mathcal{F}(\mu)(\cdot) = \nabla_q F(\cdot, \mu) = \nabla\phi * \mu(\cdot);$$

and

$$g(q, \mu) = U^0(q) + U^1 * \mu(q), \quad \mathcal{G}(\mu) = \int_{\mathbb{T}^d} (U^0 + \frac{1}{2} U^1 * \mu)(y)\mu(dy),$$

so that

$$\nabla_\mu \mathcal{G}(q) = \nabla_q g(q, \mu) = \nabla U^0(q) + \nabla U^1 * \mu(q).$$

It is not difficult to verify that these instances of F and g are a particular case of the conditions of Section 2.2.2.

2.3 Definitions of classical (strong) solutions

Let $T > 0$, and $F, g : \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ be continuous; let $H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ be continuous and differentiable in p. The following fixes our definitions of classical solution to the MFG system (1.3.2-1.3.5) and the master equation (1.3.6).

MFG system Let $0 < s < T$, $\mu \in \mathscr{P}(\mathbb{T}^d)$. We say that the pair of functions $U : (0, T) \times \mathbb{T}^d \to \mathbb{R}$, $\sigma : (0, T) \to \mathscr{P}(\mathbb{T}^d)$ is a *classical solution to the first-order MFG system* (1.3.2-1.3.5) *on* $\mathbb{T}^d$ *with coefficients* (H, F, g) *and parameters* s, μ if the following hold:

- $U \in C^1((0, T) \times \mathbb{T}^d)$;
- the path $\sigma \in AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$ and (1.3.3) is true in the sense of distributions, i.e., for every $\varphi \in C_c^\infty((0, T) \times \mathbb{T}^d)$:

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

- equation (1.3.2) is satisfied pointwise, along with the condition (1.3.4) at time $t = 0$ for U and the condition (1.3.5) at time $t = s$ for σ.

We will often refer to the function U in (1.3.2) as the *value function*.

Master equation We say that the function $u : (0, T) \times \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ is a *classical solution the master equation of first-order MFGs* (1.3.6) *with coefficients* (H, F, g) if:

- u is differentiable in s, with $\partial_s u(\cdot, \cdot, \mu)$ continuous at every $\mu \in \mathscr{P}(\mathbb{T}^d)$;
- u is differentiable in q, with $\nabla_q u$ continuous in all three variables;
- u is differentiable in μ (see the following section), and u satisfies (1.3.6) pointwise.

We will refer to the function u in (1.3.6) as the *full value function*.

Chapter 3

THE HAMILTONIAN SYSTEM OF CHARACTERISTICS

In this chapter we obtain a unique solution to the system

$$\begin{aligned}
\partial_t Q(t,q) &= \nabla_p H(Q(t,q),P(t,q)),\\
\partial_t P(t,q) &= -\nabla_q H(Q(t,q),P(t,q)) - F(Q(t,q),Q(t,\cdot)_{\#}\mu),\\
Q(s,q) &= q,\\
P(0,q) &= g(Q(0,q),Q(0,\cdot)_{\#}\mu),
\end{aligned}$$

where $s < T$, and $\mu \in \mathscr{P}(\mathbb{T}^d)$, whose solution we will denote by $\Sigma = (\Sigma^1, \Sigma^2)$, Σ^1 standing for Q, Σ^2 for P. The solution is unique, provided T is small enough. This system of ODEs is the corresponding Hamiltonian system of the Euler-Lagrange equations for the action

$$\int_0^s [L(Q_t(q),\dot{Q}_t(q)) - F(Q_t(q),(Q_t)_{\#}\mu)]dt + g(Q_0,(Q_0)_{\#}\mu),$$

where $L = H^*$, which relates to (1.3.1) in the introduction. In Section 3.2 we prove some regularity properties of Σ that will be then used in Chapter 5 to obtain the solution to the MFG system.

3.1 Fixed-point argument

For $T > 0$, we will denote by $\mathcal{M}$ the space of continuous functions

$$Z = (Q,P) : [0,T] \times \mathbb{T}^d \longrightarrow \mathbb{T}^d \times \mathbb{R}^d,$$

endowed with the uniform norm,

$$\|Z\|_\infty = \max_{0 \leqslant t \leqslant T} |Z(t,q)| = \max\{(|Q(t,q)|^2 + |P(t,q)|^2)^{1/2} \mid t \in [0,T], q \in \mathbb{T}^d\}.$$

That is,

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

Similarly, let

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

Definition 3.1.1. *Let $\theta \in \mathbb{R}^+$, and fix $\mu \in \mathscr{P}(\mathbb{T}^d)$, $s \in [0,T]$.*

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

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

$$(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})(t,q) = q + \int_s^t \nabla_p H(\bar{Q}_\tau(q), \theta \bar{P}_\tau(q)) d\tau, \tag{3.1.1}$$

$$(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})(t,q) = \frac{1}{\theta} \nabla_q g(\bar{Q}_0(q), \bar{Q}_{0\#}\mu) - \frac{1}{\theta} \int_0^t \nabla_q H(\bar{Q}_\tau(q), \theta \bar{P}_\tau(q)) + \nabla_q F(\bar{Q}_\tau(q), \bar{Q}_{\tau\#}\mu) d\tau, \tag{3.1.2}$$

$0 \leqslant t \leqslant T$, $q \in \mathbb{R}^d$. *In equalities (3.1.1) and (3.1.2),* $\bar{Q}_\tau(q) := \bar{Q}(\tau, q)$, $\bar{P}_\tau(q) := \bar{P}(\tau, q)$, $\tau \in [0,T], q \in \mathbb{R}^d$.

2. *(Coefficient bounds I) For* $B > 0$, *let*

$$\bar{l}(B) := \max_{\substack{q \in \mathbb{R}^d, |p| \leqslant B, \\ \mu \in \mathscr{P}(\mathbb{T}^d)}} \left\{ \sqrt{2}\ |\nabla H(q, \theta p)| + |\nabla_q F(q,\mu)| \right\},$$

$$\bar{h}(B) := \max_{\substack{q \in \mathbb{R}^d, |p| \leqslant B, \\ \mu \in \mathscr{P}(\mathbb{T}^d)}} \left\{ \sqrt{2} |\nabla^2 H(q, \theta p)| + \sqrt{2} |\nabla^3 H(q, \theta p)| + |\nabla^2_{qq} F(q,\mu)| + |\nabla^3_{qqq} F(q,\mu)| \right\},$$

$$c := \max\{d, \kappa\}.$$

Thus, for a fixed B, *the numbers* $\bar{l}(B)$, $\bar{h}(B), c$ *depend only on the coefficients (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 $(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})(t,q) \sim (\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})(t,q')$, so $(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})(t,\cdot)$ is indeed a mapping into $\mathbb{T}^d$, in the sense explained in the Preliminaries.

(2) Both the fixed-point operator $\bar{\mathfrak{m}}^{s,\mu}$ and the coefficient bounds depend on the value of θ.

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

Suppose that the operator $\bar{\mathfrak{m}}^{s,\mu}$ has a fixed point $(\bar{Q}, \bar{P})$, so on the left-hand side of (3.1.1) and (3.1.2) we would see $\bar{Q}(t,q)$ and $\bar{P}(t,q)$ respectively. Set $Q := \bar{Q}$ and $P := \theta \bar{P}$. Then $Z := (Q, P)$ satisfies

$$\begin{cases} \partial_t Q(t,q) = \nabla_p H(Q(t,q), P(t,q)) \qquad \text{in } [0,s] \times \mathbb{T}^d , \\ \partial_t P(t,q) = -\nabla_q H(Q(t,q), P(t,q)) - \nabla_q F(Q(t,q), Q(t,\cdot)_\# \mu) \quad \text{in } [0,s] \times \mathbb{T}^d , \\ \ \ Q(s,q) = q \qquad \text{on } \mathbb{T}^d , \\ \ \ P(0,q) = \nabla_q g(Q(0,q), Q(0,\cdot)_\# \mu) \qquad \text{on } \mathbb{T}^d. \end{cases} \tag{3.1.3}$$

We will refer to the system (3.1.3) as the *Hamiltonian ODEs with parameters s and* μ.

Definition 3.1.2. *If* $T > 0$, $A_1, A_2, B, E, E_1, E_2 > 0$, *define*

$$\mathcal{M}_0(A_1, A_2, B, E, E_1, E_2, T) \subset \mathcal{M}$$

to be the subset of those $\bar{Z}(\cdot,\cdot) = (\bar{Q}(\cdot,\cdot), \bar{P}(\cdot,\cdot))$ such that:

(i) $\bar{Z}(\cdot,\cdot)$ belongs to $W^{2,2;\infty}([0,T]\times\mathbb{T}^d;\mathbb{T}^d\times\mathbb{R}^d)$;

(ii) the following bounds hold:

$$\begin{cases} \|\partial_t\bar{Q}\|_{\mathcal{M}^1} \leqslant A_1, \ \|\nabla_q\bar{Q}\|_{C([0,T]\times\mathbb{T}^d;\mathbb{T}^d\times\mathbb{T}^d)} \leqslant A_1, \ \|\nabla^2_{qq}\bar{Q}\|_{C([0,T]\times\mathbb{T}^d;\mathbb{T}^{2d}\times\mathbb{T}^d)} \leqslant A_1; \\ \|\partial_t\bar{P}\|_{\mathcal{M}^2} \leqslant A_2, \quad \|\nabla_q\bar{P}\|_{C([0,T]\times\mathbb{T}^d;\mathbb{R}^d\times\mathbb{R}^d)} \leqslant A_2, \ \|\nabla^2_{qq}\bar{P}\|_{C([0,T]\times\mathbb{T}^d;\mathbb{R}^{2d}\times\mathbb{R}^d)} \leqslant A_2 \\ \|\bar{P}\|_{\mathcal{M}^2} \leqslant B; \\ \|\nabla_q\bar{Q}_0\|_{C(\mathbb{T}^d;\mathbb{T}^d\times\mathbb{T}^d)}, \|\nabla^2_{qq}\bar{Q}_0\|_{C(\mathbb{T}^d;\mathbb{T}^{2d}\times\mathbb{T}^d)} \leqslant E; \end{cases} \tag{3.1.4}$$

(iii) $\|\partial^2_{tt}\bar{Q}\|_{\mathcal{M}} \leqslant E_1, \|\partial^2_{tt}\bar{P}\|_{\mathcal{M}} \leqslant E_2$.

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, $\mathcal{M}_0(A_1,A_2,B,E,E_1,E_2,T)$ is indeed a subset of $\mathcal{M}$. The following is a standard fact, but we will sketch its proof.

Proposition 3.1.3. *For any $A_1,A_2,B,E,E_1,E_2,T>0$, $\mathcal{M}_0(A_1,A_2,B,E,E_1,E_2,T)$ is closed in $\mathcal{M}$.*

Proof. For simplicity, let us just show that

$$\{Q\in W^{1;\infty}(\mathbb{T}^d;\mathbb{T}^d) \mid \|\nabla_q Q\|_{L^\infty(\mathbb{T}^d;\mathbb{R}^{d^2})} \leqslant A\}$$

is closed in $C(\mathbb{T}^d;\mathbb{T}^d)$. Suppose that $\{Q_n\}_1^\infty$ is a sequence in $W^{1;\infty}(\mathbb{T}^d;\mathbb{T}^d)$, and that $Q_n\to Q$ in $C(\mathbb{T}^d;\mathbb{T}^d)$, that is, Q_n converges to $Q\in C(\mathbb{T}^d;\mathbb{T}^d)$ uniformly, and $\|\nabla_q Q_n\|_{L^\infty(\mathbb{T}^d;\mathbb{R}^{d^2})}\leqslant A$, for every n. Then the sequence $\{\nabla_q Q_n\}_1^\infty$ is also bounded as a sequence of functionals in $(L^1(\mathbb{T}^d;\mathbb{R}^{d^2}))^*$. By Alaoglu's theorem, there exists $Q'\in L^\infty(\mathbb{T}^d;\mathbb{R}^{d^2})$ and a subsequence $\{\nabla_q Q_{n_k}\}_{k=1}^\infty$ such that

$$\nabla_q Q_{n_k} \underset{k\to\infty}{\longrightarrow} Q' \quad \text{weak} * \ \text{in } (L^1(\mathbb{T}^d;\mathbb{R}^{d^2}))^*,$$

and $\|Q'\|_{L^\infty(\mathbb{T}^d;\mathbb{R}^{d^2})}\leqslant A$. Then, if $\varphi\in C_c^\infty(\mathbb{T}^d;\mathbb{R}^{d^2})$,

$$\int_{\mathbb{T}^d}\nabla_q Q_{n_k}(q)\cdot\varphi(q)dq \underset{k\to\infty}{\longrightarrow} \int_{\mathbb{T}^d} Q'(q)\cdot\varphi(q)dq. \tag{3.1.5}$$

On the other hand, from uniform convergence, we have that if $\varphi\in C_c^\infty(\mathbb{T}^d;\mathbb{R}^{d^2})$, then

$$\int_{\mathbb{T}^d}\nabla_q\varphi(q)\cdot Q_n(q)dq \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{T}^d}\nabla_q\varphi(q)\cdot Q(q)dq. \tag{3.1.6}$$

Since each $Q_{n_k}\in W^{1;\infty}(\mathbb{T}^d;\mathbb{T}^d)$, (3.1.5) gives us

$$\int_{\mathbb{T}^d}\nabla_q\varphi(q)\cdot Q_{n_k}(q)dq \underset{k\to\infty}{\longrightarrow} \int_{\mathbb{T}^d} Q'(q)\cdot\varphi(q)dq. \tag{3.1.7}$$

Lines (3.1.7) and (3.1.6), since they hold for arbitrary $\varphi\in C_c^\infty(\mathbb{T}^d;\mathbb{R}^{d^2})$, imply that $Q\in W^{1;\infty}(\mathbb{T}^d;\mathbb{T}^d)$ and $\nabla_q Q = Q'$. The proof for $W^{1,2;\infty}$ is obtained by taking a sub-subsequence $\{\partial_t\nabla^2_{qq}Q_{n_{k_l}}\}_{l=1}^\infty$ and proceeding in similar fashion. ☐

Lemma 3.1.4. *Let $\theta > 0$ of Definition 3.1.1 be arbitrary. There exist $A_1, A_2, B, E, E_1, E_2 > 0$, and $T > 0$, such that $\bar{\mathfrak{m}}^{s,\mu}$ maps $\mathcal{M}_0(A_1, A_2, B, E, E_1, E_2, T)$ into itself, for any $s \in (0, T)$ and $\mu \in \mathscr{P}(\mathbb{T}^d)$. The numbers $A_1, A_2, B, E, E_1, E_2, T$ depend only on the coefficients and θ.*

Proof. Observe that[1]

$$|(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})(t,q)| \leqslant \frac{1}{\theta}|\nabla_q g(\bar{Q}_0(q), \bar{Q}_{0\#}\mu)| + t\frac{1}{\theta}\sup_{0\leqslant\tau\leqslant t}[|\nabla_q H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q))| + |\nabla_q F(\bar{Q}_\tau(q), \bar{Q}_{\tau\#}\mu)|],$$

$$\begin{aligned}|\partial_t(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})(t,q))| &\leqslant |\nabla_p H(\bar{Q}_t(q), \theta\bar{P}_t(q))| \ , \\ |\partial_t(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})(t,q)| &\leqslant \frac{1}{\theta}|\nabla_q H(\bar{Q}_t(q), \theta\bar{P}_t(q))| + \frac{1}{\theta}|\nabla_q F(\bar{Q}_t(q), \bar{Q}_{t\#}\mu)| \ ;\end{aligned}$$

$$|\nabla_q(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})(t,q)| \leqslant \sqrt{d} + \int_s^t \big|\nabla^2_{qp}H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q))\nabla_q\bar{Q}_\tau(q) + \nabla^2_{pp}H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q))\theta\nabla_q\bar{P}_\tau(q)\big|d\tau \ ;$$

$$\begin{aligned}&|\nabla_q(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})(t,q)| \\ &\leqslant \frac{1}{\theta}|\nabla^2_{qq}g(\bar{Q}_0(q), \bar{Q}_{0\#}\mu)\nabla_q\bar{Q}_0(q)| + \frac{1}{\theta}\int_0^t |\nabla^2_{qq}F(\bar{Q}_\tau(q), \bar{Q}_{\tau\#}\mu)\nabla_q\bar{Q}_\tau(q)|d\tau \\ &\quad + \frac{1}{\theta}\int_0^t \big|\nabla^2_{qq}H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q))\nabla_q\bar{Q}_\tau(q) + \nabla^2_{pq}H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q))\theta\nabla_q\bar{P}_\tau(q)\big|d\tau\end{aligned}$$

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

$$\begin{aligned}&|\nabla^2_{qq}(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})(t,q)| \\ &\qquad \leqslant \int_s^t \big|(\nabla^3_{qqp}H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q))\nabla_q\bar{Q}_\tau(q) + \nabla^3_{pqp}H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q))\theta\nabla_q\bar{P}_\tau(q))\nabla_q\bar{Q}_\tau(q) \\ &\qquad\quad + \nabla^2_{qp}H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q))\nabla^2_{qq}\bar{Q}_\tau(q) \\ &\qquad\quad + (\nabla^3_{qpp}H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q))\nabla_q\bar{Q}_\tau(q) + \nabla^3_{ppp}H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q))\theta\nabla_q\bar{P}_\tau(q))\theta\nabla_q\bar{P}_\tau(q) \\ &\qquad\quad + \nabla^2_{pp}H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q))\theta\nabla^2_{qq}\bar{P}_\tau(q)\big|d\tau\end{aligned}$$

[1]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 $\bar{\mathfrak{m}}$, and, for the second component,

$$
\begin{aligned}
&|\nabla^2_{qq}(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})(t,q)| \\
&\qquad \leqslant \frac{1}{\theta}\big|(\nabla^3_{qqq}g(\bar{Q}_0(q),\bar{Q}_{0\#}\mu)\nabla_q\bar{Q}_0(q))\nabla_q\bar{Q}_0(q)+\nabla^2_{qq}g(\bar{Q}_0(q),\bar{Q}_{0\#}\mu)\nabla^2_{qq}\bar{Q}_0(q)\big| \\
&\qquad + \frac{1}{\theta}\int_0^t \big|(\nabla^3_{qqq}H(\bar{Q}_\tau(q),\theta\bar{P}_\tau(q))\nabla_q\bar{Q}_\tau(q)+\nabla^3_{pqq}H(\bar{Q}_\tau(q),\theta\bar{P}_\tau(q))\theta\nabla_q\bar{P}_\tau(q))\nabla_q\bar{Q}_\tau(q) \\
&\qquad\qquad + \nabla^2_{qq}H(\bar{Q}_\tau(q),\theta\bar{P}_\tau(q))\nabla^2_{qq}\bar{Q}_\tau(q) \\
&\qquad\qquad + (\nabla^3_{qpq}H(\bar{Q}_\tau(q),\theta\bar{P}_\tau(q))\nabla_q\bar{Q}_\tau(q)+\nabla^3_{ppq}H(\bar{Q}_\tau(q),\theta\bar{P}_\tau(q))\theta\nabla_q\bar{P}_\tau(q))\theta\nabla_q\bar{P}_\tau(q) \\
&\qquad\qquad + \nabla^2_{pq}H(\bar{Q}_\tau(q),\theta\bar{P}_\tau(q))\theta\nabla^2_{qq}\bar{P}_\tau(q)\big|d\tau \\
&\qquad + \frac{1}{\theta}\int_0^t \big|(\nabla^3_{qqq}F(\bar{Q}_\tau(q),\bar{Q}_{\tau\#}\mu)\nabla_q\bar{Q}_\tau(q))\nabla_q\bar{Q}_\tau(q)+\nabla^2_{qq}F(\bar{Q}_\tau(q),\bar{Q}_{\tau\#}\mu)\nabla^2_{qq}\bar{Q}_\tau(q)\big|d\tau.
\end{aligned}
$$

We deal with A_1, A_2, B, E, T first. 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 (3.1.4). From the latter inequalities we see that:

$$
\begin{aligned}
&\text{(a)} \quad c/\theta + T\bar{h}(B)/\theta \leqslant B \qquad \text{implies} \quad |(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})| \leqslant B; \\
&\text{(b1)} \quad \bar{l}(B) \leqslant A_1 \qquad \text{implies} \quad |\partial_t(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})| \leqslant A_1; \\
&\text{(b2)} \quad \bar{l}(B)/\theta \leqslant A_2 \qquad \text{implies} \quad |\partial_t(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})| \leqslant A_2;
\end{aligned}
$$

$$
\begin{aligned}
&\text{(c)} \quad c + T\bar{h}(B)(A_1+\theta A_2) \leqslant A_1 \qquad \text{implies} \quad |\nabla_q(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})| \leqslant A_1; \\
&\text{(d)} \quad cE/\theta + \frac{T}{\theta}\bar{h}(B)(A_1+\theta A_2) \leqslant A_2 \qquad \text{implies} \quad |\nabla_q(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})| \leqslant A_2; \\
&\text{(e)} \quad c + T\bar{h}(B)(A_1+\theta A_2) \leqslant E \qquad \text{implies} \quad |\nabla_q(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})|\big|_{t=0} \leqslant E;
\end{aligned}
$$

$$
\begin{aligned}
&\text{(f1)} \quad T\bar{h}(B)(A_1+\theta A_2)(A_1+\theta A_2+1) \leqslant A_1 \qquad \text{implies} \quad |\nabla^2_{qq}(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})| \leqslant A_1; \\
&\text{(f2)} \quad T\bar{h}(B)(A_1+\theta A_2)(A_1+\theta A_2+1) \leqslant E \qquad \text{implies} \quad |\nabla^2_{qq}(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})|\big|_{t=0} \leqslant E; \\
&\text{(g)} \quad \frac{1}{\theta}cE(E+1) + \frac{T}{\theta}\bar{h}(B)(A_1+\theta A_2)(A_1+\theta A_2+1) + \frac{T}{\theta}\bar{h}(B)A_1(A_1+1) \leqslant A_2 \\
&\qquad\qquad \text{implies} \quad |\nabla^2_{qq}(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})| \leqslant A_2.
\end{aligned}
$$

We need to set A_1, A_2, B, E, T so that the above inequalities hold simultaneously. First choose $B > c/\theta$. The number B now depends only on the coefficients and θ (through c), and thus $\bar{l}(B), \bar{h}(B)$ depend only on the coefficients and θ, through B. Let T be small enough that $c/\theta + (T/\theta)\bar{l}(B) \leqslant B$ (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

$$A_1 > \max\{\bar{l}(B), E, c\}, \tag{3.1.8}$$

$$A_2 > \max\{\frac{\bar{l}(B)}{\theta}, \frac{1}{\theta}cE(E+1)\}. \tag{3.1.9}$$

This gives (b1), (b2). Making T possibly smaller by letting

$$T < R := \min\Big\{\frac{\theta B - c}{\bar{l}(B)}, \frac{E-c}{\bar{h}(B)(A_1+\theta A_2)}, \frac{A_2 - cE(E+1)/\theta}{(1/\theta)\bar{h}(B)[(A_1+\theta A_2)(A_1+\theta A_2+1)+A_1(A_1+1)]}, \frac{E}{\bar{h}(B)(A_1+\theta A_2)(A_1+\theta A_2+1)}\Big\}, \tag{3.1.10}$$

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. Finally, let us answer the existence of the constants E_1, E_2. We compute $\partial^2_{tt}(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})(t,q)$ to be

$$\partial^2_{tt}(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})(t,q) = \nabla^2_{pq}H(\bar{Q}_t(q), \theta\bar{P}_t(q))\partial_t\bar{Q}_t(q) + \theta\nabla^2_{pp}H(\bar{Q}_t(q), \theta\bar{P}_t(q))\partial_t\bar{P}_t(q)$$

and, thanks to the conditions in Section 2.2.1, by arguing as in Proposition 2.1.6 to address the time derivative of $\nabla_q F(\bar{Q}_t(q), (\bar{Q}_t)_{\#}\mu)$, we get for $\partial^2_{tt}(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})(t,q)$:

$$\begin{aligned}\partial^2_{tt}(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})(t,q) = &-\frac{1}{\theta}\nabla^2_{qq}H(\bar{Q}_t(q), \theta\bar{P}_t(q))\partial_t\bar{Q}_t(q) + \nabla^2_{pq}H(\bar{Q}_t(q), \theta\bar{P}_t(q))\partial_t\bar{P}_t(q)\\ &-\frac{1}{\theta}\nabla^2_{qq}F(\bar{Q}_t(q), \bar{Q}_{t\#}\mu)\partial_t\bar{Q}_t(q) - \frac{1}{\theta}\int_{\mathbb{T}^d}\nabla^2_{\mu q}F(\bar{Q}_t(q), \bar{Q}_{t\#}\mu)(\bar{Q}_t(x))\partial_t\bar{Q}_t(x)\mu(dx).\end{aligned}$$

Therefore, it is enough to choose E_1, E_2 large enough such that

$$\begin{aligned}\bar{h}(B)A_1 + \theta\bar{h}(B)A_2 &\leqslant E_1/\sqrt{2},\\ \frac{1}{\theta}\bar{h}(B)A_1 + \bar{h}(B)A_2 + \frac{1}{\theta}cA_1 + \frac{1}{\theta}\|\nabla^2_{\mu q}F\|_\infty A_1 &\leqslant E_2/\sqrt{2}.\end{aligned}$$

□

Proposition 3.1.5. *(Contraction property) Let $\theta > 2\kappa$. Then there exist positive numbers A_1,A_2, B, E, E_1, T such that for any $s \in [0,T]$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, the operator $\bar{\mathfrak{m}}^{s,\mu}$ maps $\mathcal{M}_0(A,B,E,E_1,T)$ into itself and is a contraction.*

Proof. We run the previous lemma to obtain the numbers A_1, A_2, B, E, E_1, and T, and decrease T, if necessary, so that

$$T < \min\Big\{R, \frac{1-\frac{2\kappa}{\theta}}{\bar{h}(B)\sqrt{2}(1+1/\theta)+2\kappa/\theta}\Big\}, \tag{3.1.11}$$

where R is the number defined in (3.1.10). Let $\bar{Z} = (\bar{Q}, \bar{P}), \bar{Z}' = (\bar{Q}', \bar{P}') \in \mathcal{M}_0$. Let $s \in [0,T]$, $\mu \in \mathscr{P}(\mathbb{T}^d)$ be arbitrary. We have, for the first component of $\bar{\mathfrak{m}}^{s,\mu}$, that

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

Since H is C^2, we can write

$$|\nabla_p H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q)) - \nabla_p H(\bar{Q}'_\tau(q), \theta\bar{P}'_\tau(q))| \leqslant M^1_{\tau,q}|\bar{Z}(\tau,q) - \bar{Z}'(\tau,q)|,$$

where

$$M^1_{\tau,q} = \max_{0\leqslant\lambda\leqslant 1} |\nabla(\nabla_p H)[(1-\lambda)\bar{Q}_\tau(q) + \lambda\bar{Q}'_\tau(q), (1-\lambda)\theta\bar{P}_\tau(q) + \lambda\theta\bar{P}'_\tau(q)]|.$$

For the second component of $\bar{\mathfrak{m}}$ we apply (2.1.2) to get

$$\begin{aligned}
&|(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})(t,q) - (\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z}')(t,q)| \\
&\leqslant \frac{\kappa}{\theta}\sqrt{|\bar{Q}_0(q) - \bar{Q}'_0(q)|^2 + \|\bar{Q}_0 - \bar{Q}'_0\|^2_{L^2(\mu)}} + \frac{1}{\theta} t \max_{0\leqslant\tau\leqslant t}\big[|\nabla_q H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q)) - \nabla_q H(\bar{Q}'_\tau(q), \theta\bar{P}'_\tau(q))| \\
&\quad + \kappa\sqrt{|\bar{Q}_\tau(q) - \bar{Q}'_\tau(q)|^2 + \|\bar{Q}_\tau - \bar{Q}'_\tau\|^2_{L^2(\mu)}}\ \big] \\
&\leqslant \sqrt{2}\frac{\kappa}{\theta}(1+t)\|\bar{Z} - \bar{Z}'\|_\infty + \frac{t}{\theta}\max_{0\leqslant\tau\leqslant t}|\nabla_q H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q)) - \nabla_q H(\bar{Q}'_\tau(q), \theta\bar{P}'_\tau(q))|,
\end{aligned}$$

with

$$|\nabla_q H(\bar{Q}_\tau(q), \theta\bar{P}_\tau(q)) - \nabla_q H(\bar{Q}'_\tau(q), \theta\bar{P}'_\tau(q))| \leqslant M^2_{\tau,q}|\bar{Z}(\tau,q) - \bar{Z}'(\tau,q)|,$$

where

$$M^2_{\tau,q} = \max_{0\leqslant\lambda\leqslant 1} |\nabla(\nabla_q H)[(1-\lambda)\bar{Q}_\tau(q) + \lambda\bar{Q}'_\tau(q), (1-\lambda)\theta\bar{P}_\tau(q) + \lambda\theta\bar{P}'_\tau(q)]|.$$

But, since $\mathcal{M}_0$ is a convex subset of $\mathcal{M}$, it is true that $M^1_{\tau,q}, M^2_{\tau,q} \leqslant \bar{h}(B)$, $(\tau,q) \in [0,s]\times\mathbb{T}^d$. It follows that

$$|(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z})(t,q) - (\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z}')(t,q)| \leqslant |s-t|\bar{h}(B)\|\bar{Z} - \bar{Z}'\|_\infty\ , \quad 0\leqslant t\leqslant s,\ q\in\mathbb{T}^d,$$

and

$$|(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z})(t,q) - (\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z}')(t,q)| \leqslant \sqrt{2}\frac{1}{\theta}\kappa(1+t)\|\bar{Z} - \bar{Z}'\|_\infty + \frac{1}{\theta}t\bar{h}(B)\|\bar{Z} - \bar{Z}'\|_\infty.$$

Consequently, since $0\leqslant t,s\leqslant T$, we obtain

$$\begin{aligned}
\|\bar{\mathfrak{m}}^{s,\mu}(\bar{Z}) - \bar{\mathfrak{m}}^{s,\mu}(\bar{Z}')\|_\infty &\leqslant \sqrt{2}T\bar{h}(B) + \sqrt{2}(\sqrt{2}\frac{1}{\theta}\kappa(1+T) + \frac{1}{\theta}T\bar{h}(B))\|\bar{Z} - \bar{Z}'\|_\infty \\
&= \Big[\frac{2\kappa}{\theta} + T\big(\bar{h}(B)\sqrt{2}(1+\frac{1}{\theta}) + \frac{2\kappa}{\theta}\big)\Big]\|\bar{Z} - \bar{Z}'\|_\infty. \qquad (3.1.12)
\end{aligned}$$

Due to (3.1.11), the expression inside the square brackets in (3.1.12) is less than 1. □

It follows now that the operator (3.1.1-3.1.2) has a unique fixed point in $\mathcal{M}_0(A_1, A_2, B, E, E_1, T)$, where A_1, A_2, B, E, E_1, T are as above.

Definition 3.1.6. *Fix $\mu \in \mathscr{P}(\mathbb{T}^d)$, $s \in [0,T]$. Define the operator $\mathfrak{m}^{s,\mu} : \mathcal{M} \to \mathcal{M}$, $\mathfrak{m} = ((\mathfrak{m}^{s,\mu})^1, (\mathfrak{m}^{s,\mu})^2)$ as follows:*

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

$$(\mathfrak{m}^{s,\mu})^1(Z)(t,q) = q + \int_s^t \nabla_p H(Q_\tau(q), P_\tau(q))d\tau, \qquad (3.1.13)$$

$$(\mathfrak{m}^{s,\mu})^2(Z)(t,q) = \nabla_q g(Q_0(q), Q_{0\#}\mu) - \int_0^t \nabla_q H(Q_\tau(q), P_\tau(q)) + \nabla_q F(Q_\tau(q), Q_{\tau\#}\mu)d\tau, \qquad (3.1.14)$$

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

Corollary 3.1.7. *(i) For any $T > 0$, $s \in [0,T], \mu \in \mathscr{P}(\mathbb{T}^d)$, $\theta > 0$, the operator $\bar{\mathfrak{m}}$ has a unique fixed point in $\mathcal{M}_0(A_1, A_2, B, E, E_1, E_2, T)$ if, and only if, $\mathfrak{m}$ has a unique fixed point in $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$.*

(ii) With $\theta > 2\kappa$, fix $s \in [0,T]$, $\mu \in \mathscr{P}(\mathbb{T}^d)$. Let $A_1, A_2, B, E, E_1, E_2, T$ be as obtained in Lemma 3.1.4. Then $\mathfrak{m}^{s,\mu}$ maps $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$ into itself and the system (3.1.3) has a unique solution in $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$.

Proof. (i) Suppose $\bar{\mathfrak{m}}^{s,\mu}$ has a unique fixed point $\bar{\Sigma}[s,\mu] = (\bar{\Sigma}^1[s,\mu], \bar{\Sigma}^2[s,\mu])$ that satisfies the bounds of Definition 3.1.2 with $\bar{Q} = \bar{\Sigma}^1$ and $\bar{P} = \bar{\Sigma}^2$. Define

$$\Sigma^2[s,\mu] := \bar{\Sigma}^1[s,\mu], \quad \Sigma^2[s,\mu] := \theta\bar{\Sigma}^2[s,\mu]. \tag{3.1.15}$$

Then it is straighforward to check that $\Sigma[s,\mu] = (\Sigma^1[s,\mu], \Sigma^2[s,\mu])$ is the unique fixed point of the operator $\mathfrak{m}^{s,\mu}$ such that the inequalities of Definition 3.1.2 are true for Σ^1, Σ^2 with the new constants θA_2, θB, θE_2 in place of A_2 and B, E_2 respectively, that is, such that $Q = \Sigma^1$, $P = \Sigma^2$ satisfy

$$\begin{cases} \|\partial_t Q\|_{\mathcal{M}^1} \leqslant A_1, \ \|\nabla_q Q\|_{C([0,T]\times\mathbb{T}^d;\mathbb{T}^d\times\mathbb{T}^d)} \leqslant A_1, \ \|\nabla^2_{qq} Q\|_{C([0,T]\times\mathbb{T}^d;\mathbb{T}^{2d}\times\mathbb{T}^d)} \leqslant A_1; \\ \|\partial_t P\|_{\mathcal{M}^2} \leqslant \theta A_2, \ \ \|\nabla_q P\|_{C([0,T]\times\mathbb{T}^d;\mathbb{R}^d\times\mathbb{R}^d)} \leqslant \theta A_2, \ \|\nabla^2_{qq} P\|_{C([0,T]\times\mathbb{T}^d;\mathbb{R}^{2d}\times\mathbb{R}^d)} \leqslant \theta A_2; \\ \|P\|_{\mathcal{M}^2} \leqslant \theta B; \\ \|\nabla_q Q_0\|_{C(\mathbb{T}^d;\mathbb{T}^d\times\mathbb{T}^d)}, \|\nabla^2_{qq} Q_0\|_{C(\mathbb{T}^d;\mathbb{T}^{2d}\times\mathbb{T}^d)} \leqslant E, \\ \|\partial^2_{tt} Q\|_{\mathcal{M}} \leqslant E_1, \|\partial^2_{tt} P\|_{\mathcal{M}} \leqslant \theta E_2. \end{cases} \tag{3.1.16}$$

The sufficiency part of the statement is equally easily verified.

(ii) We take $Z = (Q,P) \in \mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$, and evaluate $\bar{\mathfrak{m}}^{s,\mu}$ at $\bar{Z} = (\bar{Q}, \bar{P})$ where $\bar{Q} = Q$ and $\bar{P} = P/\theta$, $\bar{Z} \in \mathcal{M}_0(A_1, A_2, B, E, E_1, E_2, T)$. Then, by Lemma 3.1.4, $\bar{\mathfrak{m}}^{s,\mu}(\bar{Z}) \in \mathcal{M}_0(A_1, A_2, B, E, E_1, E_2, T)$. But $(\bar{\mathfrak{m}}^{s,\mu})^1(\bar{Z}) = (\mathfrak{m}^{s,\mu})^1(Z)$ and $(\bar{\mathfrak{m}}^{s,\mu})^2(\bar{Z}) = \frac{1}{\theta}(\mathfrak{m}^{s,\mu})^1(Z)$, thus, $\mathfrak{m}^{s,\mu}(Z) \in \mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$. Furthermore, Proposition 3.1.5 provides a unique fixed point $\bar{\Sigma}[s,\mu]$ of $\bar{\mathfrak{m}}$ in $\mathcal{M}_0(A_1, A_2, B, E, E_1, E_2, T)$. Defining $\Sigma[s,\mu]$ as in (3.1.15), then, by (i), $\Sigma[s,\mu]$ is the unique fixed point of the operator $\mathfrak{m}^{s,\mu}$ on $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$, so it is the unique solution to (3.1.3) in $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$.

□

We stress that, as long as T is as in (3.1.11), we have a solution $\Sigma[s,\mu]$ to (3.1.3) for any $\mu \in \mathscr{P}(\mathbb{T}^d)$, and $0 \leqslant s \leqslant T$, and moreover, its Q and P components $\Sigma^1[s,\mu]$ and $\Sigma^2[s,\mu]$ satisfy the bounds of Definition 3.1.2 with θA_2, and θB, θE_2 in place of A, B, θE_2 respectively, *independently of* μ, with solutions being continuous and differentiable in t and q.

These solutions will always be denoted by $\Sigma[s,\mu] = (\Sigma^1[s,\mu], \Sigma^2[s,\mu])$.

Remark 3.1.8. 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 (3.1.11) above that would make their purpose unclear at first reading. We may sometimes just say "T is small", having this remark in mind. //

3.2 First regularity properties of the solution

Lemma 3.2.1. *For any fixed $Z = (Q, P) \in \mathcal{M}$, $t \in [0, T]$, $q \in \mathbb{R}^d$, and $\mu \in \mathscr{P}(\mathbb{T}^d)$, the function*

$$s \mapsto \mathfrak{m}^{s,\mu}(Z)(t, q)$$

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

$$\mu \mapsto \mathfrak{m}^{s,\mu}(Z)(t, q)$$

is continuous.

Proof. Continuity of $s \mapsto \mathfrak{m}^{s,\mu}(Z)(t, q)$ for fixed t, q, μ is immediate from Definition 3.1.6 (formulas (3.1.13) and (3.1.14)). Looking at the same definition for the continuity with respect to μ, note that for each τ, $0 \leqslant \tau \leqslant T$, the function $q \mapsto Q_\tau(q)$ is Lipschitz. This implies (see, e.g., [GŚ15, Remark 3.3]) that the function $\mu \mapsto Q_{\tau\#}\mu$ is Lipschitz from $\mathscr{P}(\mathbb{T}^d)$ 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 τ, $0 \leqslant \tau \leqslant T$. These facts, combined with the Lipschitz continuity of $\nabla_q F$ and $\nabla_q g$, show that

$$\mu_n \to \mu \text{ in } \mathscr{P}(\mathbb{T}^d) \quad \Longrightarrow \quad \mathfrak{m}^{s,\mu_n}(Z)(t, q) \to \mathfrak{m}^{s,\mu}(Z)(t, q)$$

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 μ. This is addressed in Lemmas 3.2.4 and 3.2.7 below. Before that, let us name the coefficient bounds that will appear in the calculations.

Definition 3.2.2. *(Coefficient bounds II) For $B > 0$, let*

$$\begin{aligned} l(B) &:= \max_{\substack{q\in\mathbb{R}^d, |p|\leqslant B,\\ \mu\in\mathscr{P}(\mathbb{T}^d)}} \sqrt{2}\ |\nabla H(q, p)| + |\nabla_q F(q, \mu)|, \\ h(B) &:= \max_{\substack{q\in\mathbb{R}^d, |p|\leqslant B,\\ \mu\in\mathscr{P}(\mathbb{T}^d)}} \sqrt{2}|\nabla^2 H(q, p)| + \sqrt{2}|\nabla^3 H(q, p)| + |\nabla^2_{qq} F(q, \mu)| + |\nabla^3_{qqq} F(q, \mu)|. \end{aligned}$$

Unlike the coefficient bounds $\bar{l}(B)$ and $\bar{h}(B)$, here $l(B)$ and $h(B)$ are independent of the number θ. However, if, say, $(Q, P) \in \mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$, then $|\nabla_p H(Q(t, q), P(t, q))| \leqslant l(\theta B)$.

Definition 3.2.3. *For any $D = (D_1, D_2)$, $D_1, D_2 > 0$, define:*

(i)

$$\mathcal{M}^*_{0,D}(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T) \subset W^{1,2,2;\infty}([0, T] \times [0, T] \times \mathbb{T}^d; \mathbb{T}^d \times \mathbb{R}^d)$$

as the subset of those $Z(\cdot; \cdot, \cdot)$ such that, for each $s \in [0, T]$,

$$Z(s; \cdot, \cdot) \in \mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T),$$

and

$$\|\partial_s Q(s; \cdot, \cdot)\|_\infty \leqslant D_1, \quad \|\partial_s P(s; \cdot, \cdot)\|_\infty \leqslant D_2 \tag{3.2.1}$$

wherever $\partial_s Q$, $\partial_s P$ *are defined;*

(ii)

$$\mathcal{Q}^*_{0,D_1}(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T) \subset W^{1,2,2;\infty}([0,T] \times [0,T] \times \mathbb{T}^d; \mathbb{T}^d)$$

as the subset of those $Q(\cdot;\cdot,\cdot)$ *such that, for each* $s \in [0,T]$,

$$(Q(s;\cdot,\cdot), 0) \in \mathcal{M}^*_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T),$$

and

$$\|\partial_s Q(s;\cdot,\cdot)\|_\infty \leqslant D_1$$

wherever $\partial_s Q$ *is defined.*

By an argument similar to that of Proposition 3.1.3, the sets $\mathcal{M}^*_0$ and $\mathcal{Q}^*_0$ just defined are closed subsets for the uniform convergence of $C([0,T] \times [0,T] \times \mathbb{T}^d; \mathbb{T}^d \times \mathbb{R}^d)$ and $C([0,T] \times [0,T] \times \mathbb{T}^d; \mathbb{T}^d)$ respectively.

Lemma 3.2.4. *For fixed* $\mu \in \mathscr{P}(\mathbb{T}^d)$*, using the same notation of Definition 3.2.3, and* $\theta > 2\kappa$*:*

(i) There exists a pair of positive constants $D = (D_1, D_2)$ *such that, if* $Z \in \mathcal{M}^*_{0,D}(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$*, then the function*

$$(s,t,q) \mapsto \mathfrak{m}^{s,\mu}(Z(s;\cdot,\cdot))(t,q)$$

belongs to $\mathcal{M}^*_{0,D}(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$ *for any* $\mu \in \mathscr{P}(\mathbb{T}^d)$.

(ii) The mapping

$$s \mapsto \Sigma[s,\mu](t,q)$$

is differentiable in s *a.e. on the interval* $0 < s < T$*, for every* $\mu \in \mathscr{P}(\mathbb{T}^d)$*,* $t \in [0,T]$*, and* $q \in \mathbb{T}^d$*, and it satisfies*

$$\|\partial_s \Sigma^1[s,\mu](\cdot,\cdot)\|_{C([0,T]\times\mathbb{T}^d;\mathbb{R}^d)} \leqslant D_1, \quad \|\partial_s \Sigma^2[s,\mu](\cdot,\cdot)\|_{C([0,T]\times\mathbb{T}^d;\mathbb{R}^d)} \leqslant D_2$$

for a.e. $s \in [0,T], \mu \in \mathscr{P}(\mathbb{T}^d)$.

Proof. Given a function $Z \in \mathcal{M}^*_{0,D}(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$, define $\mathfrak{m}^\mu(Z) \in C([0,T] \times [0,T] \times \mathbb{T}^d; \mathbb{T}^d \times \mathbb{R}^d)$ to be the first function displayed in the statement.

(i) Let us show that $\mathfrak{m}^\mu(Z) \in \mathcal{M}^*_{0,D}(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$ for an appropriate D. Indeed, that $\mathfrak{m}^\mu(Z)$ is continuous is evident. For a.e. $s \in (0,T)$,

$$\begin{aligned}\partial_s Q'(s;t,q) = -\nabla_p H(Q(s;s,q), P(s;s,q)) + \int_s^t &[\nabla^2_{qp} H(Q(s;\tau,q), P(s;\tau,q))\partial_s Q(s;\tau,q) \\ &+ \nabla^2_{pp} H(Q(s;\tau,q), P(s;\tau,q))\partial_s P(s;\tau,q)]d\tau,\end{aligned}$$

so

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

Put $Q(s;\tau,\cdot)_{\#}\mu =: \sigma^s_\tau$, $0 \leqslant \tau \leqslant T$. Then, for a.e. $s \in (0,T)$,

$$\begin{aligned}\partial_s P'(s;t,q) = {} & (\partial_s)(g(Q(s;0,q),\sigma^s_0)) \\ & - \int_0^t [\nabla^2_{qq} H(Q(s;\tau,q),P(s;\tau,q))\partial_s Q(s;\tau,q) \\ & \qquad + \nabla^2_{pq} H(Q(s;\tau,q),P(s;\tau,q))\partial_s P(s;\tau,q) + (\partial_s)(g(Q(s;\tau,q),\sigma^s_\tau))]d\tau.\end{aligned}$$

By the joint Lipschitz constants of g and F being bounded by κ, we obtain

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

Thus, if $D_1 > l(\theta B)$, and $D_2 > 2\kappa D_1$, we refer to Remark 3.1.8 and assume that

$$T < \min\Big\{\frac{D_1 - l(\theta B)}{h(\theta B)(D_1 + D_2)}, \frac{D_2 - 2\kappa D_1}{2\kappa D_1 + h(\theta B)(D_1 + D_2)}\Big\}$$

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

The fact that $\mathfrak{m}^\mu(Z)(s;\cdot,\cdot) \in \mathcal{M}_0(A_1,\theta A_2,\theta B,E,E_1,\theta E_2,T)$ follows from Corollary 3.1.7.

(ii) Let $Z^0 \in \mathcal{M}_{0,D}(A_1,\theta A_2,\theta B,E,E_1,\theta E_2,T)$ be arbitrary, with D from part (i). Define inductively $Z^k = \mathfrak{m}^\mu(Z^{k-1})$, $k = 1,\ldots$ Then $\{Z^k\}_{k=0}^\infty$ is a sequence in $\mathcal{M}^*_{0,D}(\cdots)$, and for each $s \in [0,T]$, $Z^k(s;\cdot,\cdot) = \mathfrak{m}^{s,\mu}(Z^{k-1}(s;\cdot,\cdot))$, so for each fixed $s \in [0,T]$,

$$Z^k(s;\cdot,\cdot) \longrightarrow \Sigma[s,\mu](\cdot,\cdot) \quad \text{uniformly in } \mathcal{M}_0(A_1,\theta A_2,\theta B,E,E_1,\theta E_2,T),$$

by the fixed point theorem. We thus have pointwise convergence of $Z^k(\cdot;\cdot,\cdot)$ to $Z[\cdot,\mu](\cdot,\cdot)$. 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 $\mathcal{M}^*_{0,D}(\cdots)$ with respect to uniform convergence now ensures that $\Sigma[\cdot,\mu](\cdot,\cdot)$ belongs to this subspace, so it is differentiable with respect to s for a.e. $s \in [0,T]$ and satisfies (3.2.1). ☐

In the following, the constants D_1, D_2 will always be as in Lemma 3.2.4. The next remark will not be used before Section 6.1.

Remark 3.2.5. If $Z = (Q,P)$ and $\bar{Z} = (\bar{Q},\bar{P})$ are related as in the proof of Corollary 3.1.7(ii), that is, $Q = \bar{Q}$, $P = \theta\bar{P}$, then $Z \in \mathcal{M}^*_{0,D}(A_1,\theta A_2,\theta B,E,E_1,\theta E_2,T)$ if, and only if, $\bar{Z} \in \mathcal{M}^*_{0,\bar{D}}(A_1,A_2,B,E,E_1,E_2,T)$, where

$$\bar{D} = (\bar{D}_1,\bar{D}_2), \quad \bar{D}_1 := D_1,\ \bar{D}_2 := D_2/\theta. \qquad /\!/$$

Definition 3.2.6. *(Master map) The mapping*

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

given by

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

will be called the master map*.*

Lemma 3.2.7. *Let $\theta > 2\kappa$. The master map $\mathfrak{M}$ is continuous, and for any fixed $\mu \in \mathscr{P}(\mathbb{T}^d)$, $\mathfrak{M}(\cdot,\cdot,\cdot,\mu)$ is a C^1 diffeomorphism.*

Proof. Let $\{\mu_k\}_{k=1}^{\infty}$ be a sequence in $\mathscr{P}(\mathbb{T}^d)$ converging to $\mu \in \mathscr{P}(\mathbb{T}^d)$. Consider the sequence $\{\Sigma^1[\cdot,\mu_k](\cdot,\cdot)\}_1^{\infty}$ and an arbitrary subsequence $\{\Sigma^1[\cdot,\mu_{k_j}](\cdot,\cdot)\}_{j=1}^{\infty}$. Being in $\mathcal{Q}^*_{0,D}(A_1,\theta A_2,\theta B,E,E_1,\theta E_2,T)$, the latter is equicontinuous and uniformly bounded, so there is a sub-subsequence, which we still index with j, converging to S for some $S \in \mathcal{Q}^*_{0,D}(\cdots)$ as $j \to \infty$. For each $s \in [0,T]$, on one hand, $\Sigma^1[s,\mu^{k_j}](\cdot,\cdot) \underset{j\to\infty}{\longrightarrow} S(s;\cdot,\cdot)$. On the other, since, by Corollary 3.1.7 and Lemma 3.2.1, the mapping $(s,Z,\mu) \mapsto \mathfrak{m}^{s,\mu}(Z)$ is a continuous mapping of $[0,T] \times \mathcal{M}_0(A_1,\theta A_2,\theta B,E,E_1,\theta E_2,T) \times \mathscr{P}(\mathbb{T}^d)$ into $\mathcal{M}_0(A_1,\theta A_2,\theta B,E,E_1,\theta E_2,T)$ (because the Lipschitz constant of $\mathfrak{m}^{s,\mu}$ is independent of s and μ), we have

$$\Sigma^1[s,\mu^{k_j}](\cdot,\cdot) = (\mathfrak{m}^1)^{s,\mu^{k_j}}(\Sigma^1[s,\mu^{k_j}](\cdot,\cdot)) \underset{j\to\infty}{\longrightarrow} (\mathfrak{m}^1)^{s,\mu}(S(s;\cdot,\cdot)).$$

Therefore, for each $s \in [0,T]$, $S(s;\cdot,\cdot) = (\mathfrak{m}^1)^{s,\mu}(S(s;\cdot,\cdot))$, that is, $S(s;\cdot,\cdot)$ is a fixed point of $(\mathfrak{m}^1)^{s,\mu}$, so, by uniqueness, $S(s;\cdot,\cdot) = \Sigma^1[s,\mu](\cdot,\cdot)$. Thus, every subsequence of $\{\Sigma^1[\cdot,\mu_k](\cdot,\cdot)\}_1^{\infty}$ has a subsequence that converges to $\Sigma^1[\cdot,\mu](\cdot,\cdot) \in \mathcal{Q}^*_{0,D}(A_1,\theta A_2,\theta B,E,E_1,\theta E_2,T)$. Hence,

$$\Sigma^1[\cdot,\mu_k](\cdot,\cdot) \longrightarrow \Sigma^1[\cdot,\mu](\cdot,\cdot) \qquad \text{uniformly,}$$

and this implies that $\mathfrak{M}$ is continuous.

The second assertion of the lemma will be an immediate consequence of Lemma 3.2.8. ☐

Lemma 3.2.8. *Let $\theta > 2\kappa$, $0 \leqslant s \leqslant T$, $\mu \in \mathscr{P}(\mathbb{T}^d)$. The mapping $q \mapsto \Sigma^1[s,\mu](t,q)$ is a C^1 diffeomorphism, for $0 \leqslant t \leqslant T$, with*

$$\frac{1}{2} < \det \nabla_q \Sigma^1[s,\mu](t,q), \qquad |(\nabla_q \Sigma^1[s,\mu](t,q))^{-1}| < 4(1+\sqrt{d})^{d-1}, \tag{3.2.2}$$

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

$$|\nabla_q\Theta(t,q)| \leqslant |s-t|(A_1+\theta A_2)h(\theta B) \tag{3.2.3}$$

because $\Sigma \in \mathcal{M}_0(A_1,\theta A_2,\theta B,E,E_1,\theta E_2,T)$. By Remark 3.1.8, 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 $0 \leqslant t \leqslant T$.

To prove that $q \mapsto \Sigma^1(t,q)$ is onto, note that $\sup_{q\in\mathbb{R}^d} |\Sigma^1(t,q) - q| \leqslant Tl(\theta B) < 2Tl(\theta B)$, for $0 \leqslant t \leqslant T$. 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 $0 \leqslant t \leqslant T$. Therefore

$$f(t) := \deg(\Sigma^1_t, B_R(0), y), \quad 0 \leqslant t \leqslant T,$$

the topological degree of Σ^1_t is well defined at $y \in B_{R-2Tl(\theta B)}(0)$. This counts the number of "signed" solutions (see, e.g., [FG95]) x in $B_R(0)$ of the equation $\Sigma^1_t(x) = y$. 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 Σ^1_t includes $B_{R-2Tl(\theta B)}(0)$. Since $R > 1$ is arbitrary, we conclude that the range of Σ^1_t is $\mathbb{R}^d$.

We will denote the inverse of Σ^1 by X, *so*

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

for $0 < s \leqslant T$, $0 \leqslant t \leqslant T$, $q \in \mathbb{T}^d$, $\mu \in \mathscr{P}(\mathbb{T}^d)$. Next, note that for $0 \leqslant t \leqslant T$, $q \in \mathbb{T}^d$, $|\nabla_q \Sigma^1_t(q) - I_d| \leqslant T(A_1 + \theta A_2)h(\theta B) < 1$, 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 $0 \leqslant t \leqslant T$, $q \in \mathbb{T}^d$. By the inverse function theorem, X_t is differentiable. Moreover, since

$$\nabla_q X_t(q) = [\nabla_q \Sigma^1_t(X_t(q))]^{-1}, \tag{3.2.4}$$

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

To show (3.2.2), we may use the fact that[2] the determinant function $\det : \mathbb{R}^{d^2} \to \mathbb{R}$ has derivative $\nabla \det$ satisfying

$$|\nabla \det(\xi)| \leqslant 2|\xi|^{d-1}, \quad \xi \in \mathbb{R}^{d^2}$$

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 $\det(I_d + \frac{s}{T}\nabla_q \Theta) - \det I_d = \nabla \det(I_d + \tau \frac{s}{T} \nabla_q \Theta) \cdot \frac{s}{T} \nabla_q \Theta$, where $\nabla\Theta$ abbreviates $\nabla\Theta(t,q)$ at an arbitrary $(t,q) \in [0,T] \times \mathbb{T}^d$. Hence, by the aforementioned fact,

$$\begin{aligned} |\det(I_d + \frac{s}{T}\nabla_q\Theta) - \det I_d| &\leqslant \frac{s}{T} |\nabla \det(I_d + \tau \frac{s}{T} \nabla_q \Theta)| |\nabla_q \Theta| \leqslant \frac{s}{T} 2 |I_d + \tau \frac{s}{T} \nabla_q \Theta|^{d-1} |\nabla_q \Theta| \\ &\leqslant 2(\sqrt{d} + |\nabla_q \Theta|)^{d-1} |\nabla_q \Theta| \leqslant 2(1+\sqrt{d})^{d-1} T(A_1 + A_2) h(\theta B) \\ &< \frac{1}{2}, \end{aligned}$$

by (3.2.3) 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 (3.2.2). Using the inverse matrix formula and the inequality $|\nabla \det(\xi)| \leqslant 2|\xi|^{d-1}$ once more, we have

$$\begin{aligned} |\nabla_q X(t,q)| = |(I_d + \nabla_q \Theta)^{-1}| &= \left| (\det(I_d + \nabla_q\Theta))^{-1} [\nabla \det(I_d + \nabla_q \Theta)]^t \right| \\ &\leqslant \frac{2}{\det(I_d + \nabla_q \Theta)} |I_d + \nabla_q \Theta|^{d-1} \leqslant \frac{2}{\det(I_d + \nabla_q\Theta)} (1+\sqrt{d})^{d-1} \\ &< 4(1+\sqrt{d})^{d-1}, \end{aligned}$$

[2] See, for instance, [GŚ15, Remark 3.11].

and since this holds for any $t \in [0,T]$ and $q \in \mathbb{T}^d$, we have obtained the second inequality in (3.2.2).

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 (3.2.4), the second inequality in (3.2.2) implies

$$\|\nabla_q X[s,\mu](t,\cdot)\|_{C(\mathbb{T}^d;\mathbb{T}^d\times\mathbb{T}^d)} < 4(1+\sqrt{d})^{d-1}. \tag{3.2.5}$$

□

Definition 3.2.9. *Let $\theta > 2\kappa$. Given $\mu \in \mathscr{P}(\mathbb{T}^d)$, $s \in [0,T]$, and Σ the unique fixed point of $\mathfrak{m}^{s,\mu}$ in $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$, set*

$$\sigma_t = \sigma(t) := \Sigma^1[s,\mu](t,\cdot)_{\#}\mu, \quad v[s,\mu](t,q) := \partial_t \Sigma_t^1[s,\mu] \circ X_t[s,\mu](q), \tag{3.2.6}$$

for $0 \leqslant t \leqslant T$.

It should be kept in mind that the path σ_t depends on s and μ. Also, the arguments s,μ may often be omitted in the notation for v, as has been done for Σ.

Proposition 3.2.10. *The path σ belongs to $AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ and v is a velocity associated to σ, that is, $\partial_t\sigma + div(\sigma v) = 0$ in the distribution sense, with $v_t = v(t,\cdot) \in L^2(\mathbb{T}^d, \sigma_t)$, $0 < t < T$.*

Proof. Using (2.1.2),

$$\begin{aligned}
&\mathscr{W}^2(\sigma_t,\sigma_{t+h}) \leqslant \|\Sigma_t^1 - \Sigma_{t+h}^1\|_{L^2(\mu)}^2 = \int_{\mathbb{R}^d} |\Sigma_t^1(q) - \Sigma_{t+h}^1(q)|^2 \mu(dq) \\
&= \int_{\mathbb{R}^d} \left| \int_t^{t+h} \nabla_p H(\Sigma^1(\tau,q),\Sigma^2(\tau,q)) d\tau \right|^2 \mu(dq) \leqslant \int_{\mathbb{R}^d} \left(\int_t^{t+h} |\nabla_p H(\Sigma^1(\tau,q),\Sigma^2(\tau,q))| d\tau \right)^2 \mu(dq) \\
&\leqslant l(\theta B)^2 h^2 \leqslant \left(\int_t^{t+h} l(\theta B) d\tau \right)^2.
\end{aligned}$$

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

$$\begin{aligned}
\int_{\mathbb{R}^d} \nabla\varphi(q)\cdot v_t(q)\sigma_t(dq) &= \int_{\mathbb{R}^d} \nabla\varphi(q)\cdot[\partial_t\Sigma_t^1 \circ X_t(q)](\Sigma_t^1)_{\#}\mu(dq) = \int_{\mathbb{R}^d} \nabla\varphi(\Sigma_t^1(q))\cdot\partial_t\Sigma_t^1(q)\mu(dq) \\
&= \frac{d}{dt}\int_{\mathbb{R}^d} \varphi(\Sigma_t^1(q))\mu(dq) = \frac{d}{dt}\int_{\mathbb{R}^d}\varphi(q)\sigma_t(dq),
\end{aligned}$$

which shows that v is a velocity vector field for σ. 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. Finally, the boundedness of v_t ensures that $\int_0^T \int_{\mathbb{T}^d} |v_t(q)|\sigma_t(dq)dt < \infty$. □

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.

Remark 3.2.11. In the same way one proves that, for any $p > 1$, $\sigma \in AC^2(0,s;\mathscr{P}_p(\mathbb{T}^d))$, with $v_t \in L^p(\mathbb{T}^d,\sigma_t)$, $0 < t < T$. //

Proposition 3.2.12. *Let $\mu_n \to \mu$ in $\mathscr{P}(\mathbb{T}^d)$. Then*

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

Proof. Due to the uniform bound on $\nabla_q X$ (3.2.5), 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 3.2.7 yields the conclusion. □

Chapter 4

THE HAMILTON-JACOBI EQUATION ON $\mathscr{P}(\mathbb{T}^d)$

As was announced in the Preliminaries, for this chapter only we will, in addition to the conditions laid out in Section 2.2.1, uphold the conditions of Section 2.2.3.

We displayed equation (1.5.4) in the Introduction, saying that it played a crucial role in the work [GŚ15]. Let us remember that the case there was $H(q,p) = \frac{1}{2}|p|^2$. The authors showed that (σ, v) is the unique minimizer for the function $\mathcal{U}(s,\mu)$. Since $L(q,v) = \frac{1}{2}|v|^2$, the Lagrangian is increasing with respect to the modulus of v. It is then shown that this implies that v is the velocity field of minimal norm for σ. As such, $v_t \in \mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d)$ and it further follows that v_t is the gradient of a function, with the consequence that ∇v_t is a symmetric matrix. Defining u as in (6.0.1), they prove the pathwise total derivative formula (6.0.6), using the fact that ∇v_t is symmetric. The solution to the MFG system then follows, by *defining* $U(t,q) := u(t,q,\sigma_t)$.

In the course of this project, we started with a convex Hamiltonian in the p variable, satisfying (2.2.1) and (2.2.2), and we tried to follow [GŚ15] as closely as possible. We obtained the fact that (σ, v) is a minimizer for U (Theorem 4.1.4), and also that the Wasserstein gradient of $\mathcal{U}$ coincides with $\nabla_q U$ (Lemma 4.2.3). However, this is as far as one could go, because for a general Hamiltonian, it will not be true that $v_t(q)$ is a gradient; moreover, it is not $v_t(q)$ what we now need to be a gradient, but $\mathcal{V}_t(q)$, which does not appear in the minimization problem. This forced us to work directly with the MFG system and extract the desired property of $\mathcal{V}(t,q)$, as we will see in the next chapter.

Define

$$L(q,v) = \sup_{p\in\mathbb{R}^d} p\cdot v - H(q,p), \quad q\in\mathbb{T}^d, p\in\mathbb{R}^d.$$

Proposition 4.0.1. *The following bounds follow:*

- *There exists a constant $c_2 > 0$ such that*

$$|\nabla_p H(q,p)| \leqslant c_2(|p|^{r-1} + 1), \quad q\in\mathbb{T}^d, p\in\mathbb{R}^d. \tag{4.0.1}$$

- *There exist positive constants c_3, c_4, c_5 such that*

$$c_3|v|^{r'} - c_4 \leqslant L(q,v) \leqslant c_5(|v|^{r'} + 1), \tag{4.0.2}$$

where

$$r' := \frac{r}{r-1}.$$

Proof. We use standard results found in the literature.

- Fix $q \in \mathbb{T}^d$, $p \in \mathbb{R}^d$, and $j \in \{1, \ldots, d\}$. Defining $f(t) = H(q, p + te_j)$, where e_j is the j-th canonical basis vector of $\mathbb{R}^d$, we use [Dac08, p. 51], which says there is a constant $c' \geqslant 0$ such that

$$|H(q, p + te_j) - H(q, p)| \leqslant c'(1 + |p|^{r-1} + |p + te_j|^{r-1})|t|,$$

from which

$$\partial_{p_j} H(q, p)| \leqslant 2c'(1 + |p|^{r-1}), \qquad q \in \mathbb{T}^d, p \in \mathbb{R}^d.$$

Then (4.0.1) follows.

- From the lower bound on H, we have

$$L(q, v) = \sup_{p \in \mathbb{R}^d} p \cdot v - H(q, p) \leqslant \sup_{p \in \mathbb{R}^d} p \cdot v + c_1 - c_0^{-1}|p|^r.$$

It is readily verified that the mapping $p \mapsto p \cdot v + c_1 - c_0^{-1}|p|^r$ attains its maximum value at $p = (|v|^{2-r} c_0 r^{-1})^{1/(r-1)} v$. Then

$$\begin{aligned} L(q, v) &\leqslant (|v|^{2-r} c_0 r^{-1})^{\frac{1}{r-1}} v \cdot v + c_1 - c_0^{-1}(|v|^{2-r} c_0 r^{-1})^{\frac{r}{r-1}} |v|^r = \\ &= c_1 + c_0^{\frac{1}{r-1}} (r^{-\frac{1}{r-1}} - r^{-\frac{r}{r-1}})|v|^{\frac{r}{r-1}}. \end{aligned}$$

Setting $c_5 := \max\{c_1, c_0^{\frac{1}{r-1}} (r^{-\frac{1}{r-1}} - r^{-\frac{r}{r-1}})\}$, we have obtained the second inequality in (4.0.2). For the first one, note that, for any fixed $v \in \mathbb{R}^d$ and every $p \in \mathbb{R}^d$, $-L(q, v) \leqslant \inf_{p \in \mathbb{R}^d} H(q, p) - p \cdot v$, so

$$-L(q, v) \leqslant \inf_{p \in \mathbb{R}^d} c_0(|p|^r + 1) - p \cdot v.$$

The mapping $p \mapsto c_0(|p|^r + 1) - p \cdot v$ attains its minimum value at $p = (|v|^{2-r} r^{-1} c_0^{-1})^{1/(r-1)} v$, so

$$\begin{aligned} -L(q, v) &\leqslant c_0(|(|v|^{2-r} r^{-1} c_0^{-1})^{\frac{1}{r-1}} + 1) - (|v|^{2-r} r^{-1} c_0^{-1})^{\frac{1}{r-1}} v \cdot v \\ &= c_0 + c_0^{-\frac{1}{r-1}} (r^{-\frac{r}{r-1}} - r^{-\frac{1}{r-1}})|v|^{\frac{r}{r-1}}, \end{aligned}$$

from which the first inequality in (4.0.2) follows, with $c_4 = c_0$ and $c_3 = c_0^{-\frac{1}{r-1}} (-r^{-\frac{r}{r-1}} + r^{-\frac{1}{r-1}})$.

□

4.1 Minimization problem for an action on $AC^2(0, s; \mathscr{P}(\mathbb{T}^d))$

Since $L(q, \cdot)$ is the Legendre transform of $H(q, \cdot)$, $q \in \mathbb{T}^d$, we know[1] that $L \in C^3(\mathbb{T}^d \times \mathbb{R}^d)$ and that there exists $\gamma^* : [0, \infty) \to \mathbb{R}$ such that: $\lim_{t \to \infty} \gamma^*(t)/t = \infty$ and

$$L(q, b) \geqslant \gamma^*(|b|)$$

[1] The essentials of the proof of this fact are in [Dac08, p. 138].

for all $q \in \mathbb{T}^d$, $b \in \mathbb{R}^d$, because L satisfies (4.0.2). Define

$$\mathcal{L}(\bar{\mu}, w) := \int_{\mathbb{T}^d} L(x, w(x))\bar{\mu}(dx) - \mathcal{F}(\bar{\mu})$$

for $\bar{\mu} \in \mathscr{P}(\mathbb{T}^d)$ and $w : \mathbb{T}^d \to \mathbb{R}^d$ Borel such that $L(\cdot, w(\cdot)) \in L^1(\mu)$. Given $\mu \in \mathscr{P}(\mathbb{T}^d)$, recall the definition (3.2.6).

Remark 4.1.1. Proposition (3.2.10) showed that $\sigma \in AC^2(0, s; \mathscr{P}(\mathbb{T}^d))$ for every $s \in (0, T)$. Since $\mathscr{P}(\mathbb{T}^d)$ and $\mathscr{P}_{r'}(\mathbb{T}^d)$ are metrically equivalent, we also have that σ belongs to $AC^2(0, s; \mathscr{P}_{r'}(\mathbb{T}^d))$ and with velocity v. //

For any $0 \leqslant s \leqslant T$, define the augmented action

$$\mathcal{A}(s; \bar{\sigma}, \bar{v}) := \int_0^s \mathcal{L}(\bar{\sigma}, \bar{v})dt + \mathcal{G}(\bar{\sigma}_0) = \int_0^s \Big(\int_{\mathbb{T}^d} L(x, \bar{v}_t(x))\bar{\sigma}_t(dx) - \mathcal{F}(\bar{\sigma}_t)\Big)dt + \mathcal{G}(\bar{\sigma}_0),$$

where $\bar{v}$ is a velocity field for $\bar{\sigma}$. We are interested in the problem

$$\inf_{(\bar{\sigma}, \bar{v})} \big\{\mathcal{A}(s; \bar{\sigma}, \bar{v}) \ \big| \ \bar{\sigma} \in AC^2(0, s; \mathscr{P}_{r'}(\mathbb{T}^d)),\ \bar{\sigma}_s = \mu\big\}, \tag{4.1.1}$$

in which $\bar{v}$ is a velocity field for $\bar{\sigma}$. Here $r' = r/(r-1)$, as defined in Proposition 4.0.1. Note that, even though the distances $\mathscr{W}_r$, $r > 1$, are all equivalent on $\mathscr{P}(\mathbb{T}^d)$, we expect the measurement of velocities to depend on the particular r, because the degree of integrability of v_t is exactly the same as the power chosen for the distance on $\mathscr{P}(\mathbb{T}^d)$ (r', in the statement of the theorem). We are going to prove in Theorem 4.1.4 below that (σ_t, v_t) given as in (3.2.6) (see also Remark 3.2.11), is a minimizer of (4.1.1).

4.1.1 The n particles case

We first solve problem (4.1.1) in the subset of $AC^2(0, s; \mathscr{P}_{r'}(\mathbb{T}^d))$ consisting of paths of Dirac masses.

Lemma 4.1.2. *Let $x_1, \dots, x_n \in \mathbb{T}^d$. Assume*

$$\mu = \frac{1}{n}\sum_{j=1}^n \delta_{x_i}.$$

Then the path $\sigma_t = (\Sigma_t^1)_{\#}\mu$, which belongs to the set of $\mathcal{C}_n$ consisting of $\bar{\sigma} \in AC^2(0, s; \mathscr{P}_{r'}(\mathbb{T}^d))$ of the form

$$\bar{\sigma}_t = \frac{1}{n}\sum_{j=1}^n \delta_{y_j(t)},$$

$y_j(\cdot) \in AC^2(0, s; \mathbb{T}^d)$, $j = 1, \dots, n$, with its corresponding velocity $v_t[s, \mu]$, solves (4.1.1) **in** *$\mathcal{C}_n$, that is,*

$$\mathcal{A}(s; \sigma, v) = \inf_{(\bar{\sigma}, \bar{v})} \big\{\mathcal{A}(s; \bar{\sigma}, \bar{v}) \ \big| \ \bar{\sigma} \in \mathcal{C}_n \subset AC^2(0, s; \mathscr{P}_{r'}(\mathbb{T}^d)),\ \bar{\sigma}_s = \mu\big\}. \tag{4.1.2}$$

Proof. The fact that $\sigma \in \mathcal{C}_n$ is straightforward. When $\bar{\sigma} \in \mathcal{C}_n$, that is, if $\bar{\sigma} = \frac{1}{n}\sum_{j=1}^{n} \delta_{y_j(t)}$, with $y_j(\cdot) \in AC^2(0, s; \mathbb{T}^d)$, $j = 1, \dots, n$ then for a.e. $t \in [0, s]$ the velocity field is unique and equal to

$$v_t(q) = \dot{\bar{\sigma}}_t(q) := \sum_{j=1}^{n} \dot{y}_j(t)\chi_{y_j(t)}(q), \qquad q \in \mathbb{T}^d,$$

where $\chi : \mathbb{T}^d \to \{0, 1\}$ is the characteristic function. Thus, in the case $\bar{\sigma} \in \mathcal{C}_n$, the action takes the particular form

$$\mathcal{A}(s; \bar{\sigma}, \dot{\bar{\sigma}}) = \int_0^s \big[\frac{1}{n}\sum_{j=1}^{n} L(y_j(t), \dot{y}_j(t)) - \mathcal{F}\big(\frac{1}{n}\sum_{j=1}^{n} \delta_{y_j(t)}\big)\big]dt + \mathcal{G}\big(\frac{1}{n}\sum_{j=1}^{n} \delta_{y_j(0)}\big).$$

<u>*Step 1.*</u> Existence of a minimizer for

$$\inf \{ \ \mathcal{A}(s; \bar{\sigma}, \dot{\bar{\sigma}}) \bigm| \bar{\sigma}_t = \frac{1}{n}\sum_{j=1}^{n} \delta_{y_j(t)}, \ y(\cdot) \in AC^2(0, s; (\mathbb{T}^d)^n), \ y(s) = x \ \},$$

where $x = (x_1, \dots, x_n) \in (\mathbb{T}^d)^n$ is given, is obtained by standard methods of the calculus of variations. However, as we will see below in Step 3, we do not need to assume existence of a minimizer, because we will find one explicitly. Let $\bar{\sigma}(t) = \frac{1}{n}\sum_{j=1}^{n} \delta_{y_j(t)}$, $0 \leqslant t \leqslant s$ be such a minimizer. Let $h(\cdot) \in AC^2(0, s; (\mathbb{T}^d)^n)$ such that $h(0) = h(s) = 0$. Then, using the chain rule to compute the first variation of $\mathcal{A}$ at $y(\cdot)$,

$$\begin{aligned} 0 &= \frac{d}{d\tau}\mathcal{A}(s; \frac{1}{n}\sum_{j=1}^{n} \delta_{y_j(\cdot)+\tau h_j(\cdot)}, \sum_{j=1}^{n} \dot{y}_j(\cdot)\chi_{y_j(\cdot)}))\big|_{\tau=0} \\ &= \int_0^s \big(\frac{1}{n}\sum_{j=1}^{n} \nabla_q L(y_j(t), \dot{y}_j(t)) - \frac{d}{dt}\nabla_v L(y_j(t), \dot{y}_j(t)) - \frac{1}{n}\sum_{j=1}^{n} \nabla_q F(y_j(t), \bar{\sigma}(t))\big) \cdot h_j(t)dt \ ; \end{aligned}$$

and since $h(\cdot) = (h_1(\cdot), \dots, h_n(\cdot))$ is arbitrary (with $h(0) = h(s) = 0$), we have

$$\nabla_q L(y_j(t), \dot{y}_j(t)) - \nabla_q F(y_j(t) = \frac{d}{dt}\nabla_v L(y_j(t), \dot{y}_j(t)), \bar{\sigma}(t)), \qquad 0 \leqslant t \leqslant s, \quad j = 1, \dots n.$$

If we require of $h(\cdot) \in AC^2(0, s; (\mathbb{T}^d)^n)$ to merely satisfy $h(s) = 0$, using the previous equality in the integration by parts, the fact that the $y_j(0)$ can be any vectors gives us

$$\nabla_v L(y_j(0), \dot{y}_j(0)) = \nabla_q g(y_j(0), \bar{\sigma}(0)), \quad j = 1, \dots n.$$

Thus, the minimizer $\bar{\sigma}(t) = \frac{1}{n}\sum_{j=1}^{n} \delta_{y_j(t)}$ satisfies the Euler-Lagrange equations

$$\begin{cases} \nabla_q L(y_j(t), \dot{y}_j(t)) = \dfrac{d}{dt}\nabla_v L(y_j(t), \dot{y}_j(t)) + \nabla_q F(y_j(t), \bar{\sigma}(t)), & 0 \leqslant t \leqslant s, \quad j = 1, \dots, n \ ; \\ \nabla_v L(y_j(0), \dot{y}_j(0)) = \nabla_q g(y_j(0), \bar{\sigma}(0)), & j = 1, \dots, n \ ; \\ y_j(s) = x_j, & j = 1, \dots, n. \end{cases} \tag{4.1.3}$$

Step 2. The variables p and v that appear in $H(y,p)$ and $L(y,v)$ are related by

$$p = \nabla_v L(y, v(p)), \qquad v = \nabla_p H(y, p(v)).$$

If $t \mapsto (y(t), v(t)) \in (\mathbb{T}^d)^n \times (\mathbb{R}^d)^n$ is a trajectory in y-v space that satisfies $v(t) = \dot{y}(t)$, then (4.1.3) is satisfied if and only if

$$\begin{cases} \dot{y_j} = \nabla_p H(y_j(t), p_j(t)) \ , \quad j = 1, \ldots, n; \\ \dot{p_j} = \ - \nabla_q H(y_j(t), p_j(t)) - \nabla_q F(y_j(t), \bar{\sigma}(t)) \ , \quad j = 1, \ldots, n; \\ y_j(s) \ = x_j \ , \quad j = 1, \ldots, n; \\ p_j(0) \ = \nabla_q g(y_j(0), \bar{\sigma}(0)) \ , \quad j = 1, \ldots, n \end{cases} \tag{4.1.4}$$

is satisfied. Define the path $x(\cdot)$ by

$$x_j(t) = \Sigma^1(t, x_j), \quad j = 1, \ldots, n, \tag{4.1.5}$$

and set

$$p_j(t) := \Sigma^2(t, x_j), \quad j = 1, \ldots, n.$$

Then, since Σ is a solution to (3.1.3), we have that $x_1(\cdot), \ldots, x_n(\cdot), p_1(\cdot), \ldots, p_n(\cdot)$ is a solution of (4.1.4), with $y(\cdot) = x(\cdot)$ and $\bar{\sigma}(t) = (\Sigma^1_t)_{\#}\mu$, $0 \leqslant t \leqslant s$, $\mu = \sum_{j=1}^n \delta_{x_j}$.

Conclusion of Steps 1 & 2: $x(\cdot)$, defined by (4.1.5), satisfies the Euler-Lagrange equations, (4.1.3); that is, $x(\cdot)$ provides a critical point for the action $\mathcal{A}$ restricted to $\mathcal{C}_n$.

Step 3. Now we are going to show that the critical point provided by $x(\cdot)$, that is, $\sigma(\cdot) = \frac{1}{n}\sum_{j=1}^n \delta_{x_j(\cdot)}$, is actually the minimizer of $\mathcal{A}$ over $\mathcal{C}_n$. We will work it out with $n = 1$ first, and then generalize to $n \in \mathbb{Z}^+$.

Case $n = 1$**.** In this case, if $\bar{\sigma}(t) = \delta_{y(t)}$, $0 \leqslant t \leqslant s$, then

$$\mathcal{A}(s; \bar{\sigma}, \dot{\bar{\sigma}}) = \int_0^s L(y(t), \dot{y}(t))dt + \mathcal{G}(\delta_{y(0)}).$$

Let $h(\cdot) \in AC^2(0, s; \mathbb{T}^d)$ with $h(s) = 0$, so that $t \mapsto \delta_{x(t)+\tau h(t)}$ is an admissible path, for any $\tau \in \mathbb{R}$. Set

$$g(\tau) := \mathcal{A}(s; \delta_{x(\cdot)+\tau h(\cdot)}, \dot{x}(\cdot) + \tau \dot{h}(\cdot)), \quad \tau \in \mathbb{R}.$$

We have

$$\begin{aligned} \frac{dg(\tau)}{d\tau} = & \int_0^s \left[\nabla_q L(x(t) + \tau h(t), \dot{x}(t) + \tau \dot{h}(t)) \cdot h(t) + \nabla_v L(x(t) + \tau h(t), \dot{x}(t) + \tau \dot{h}(t)) \cdot \dot{h}(t)\right] dt \\ & + \nabla U^0(y(0) + \tau h(0)) \cdot h(0), \end{aligned}$$

and this equals zero $at\tau = 0$ because $x(\cdot)$ satisfies the Euler-Lagrange equations. The second derivative at $\tau = 0$ is

$$\begin{aligned} B(h(\cdot)) := & \frac{d^2 g(0)}{d\tau^2} \\ = & \int_0^s \left[(\nabla^2_{qq} L(x(t), \dot{x}(t)) \cdot h(t)) \cdot h(t) + 2(\nabla^2_{vq} L(x(t), \dot{x}(t)) \cdot h(t)) \cdot \dot{h}(t) \right. \\ & \left. + (\nabla^2_{vv} L(x(t), \dot{x}(t)) \cdot \dot{h}(t)) \cdot \dot{h}(t)\right] dt \\ & + [\nabla^2 U^0(y(0)) h(0)] \cdot h(0). \end{aligned}$$

We will use *Poincaré's inequality*: there is an absolute constant C^1 such that for any $u \in W_0^{1,2}(0,1;\mathbb{T}^d)$,

$$\int_0^1 |u(t)|^2 dt \leqslant C^1 \int_0^1 |\dot{u}(t)|^2 dt.$$

If $v \in W_0^{1,2}(0,b;\mathbb{T}^d)$, $b > 0$, letting $\tilde{v}(t) = v(bt)$, so that $\tilde{v} \in W_0^{1,2}(0,1;\mathbb{T}^d)$, Poincaré's inequality and a change of variable yield

$$\int_0^b |v(t)|^2 dt \leqslant C^1 b^2 \int_0^b |\dot{v}(t)|^2 dt. \tag{4.1.6}$$

Let

$$\tilde{h}(t) := \begin{cases} h(s-t), & 0 \leqslant t \leqslant s, \\ h(t-s), & s \leqslant t \leqslant 2s. \end{cases}$$

Then $\tilde{h}(0) = \tilde{h}(2s) = 0$, so $\tilde{h}(\cdot) \in W_0^{1,2}(0,2s;\mathbb{T}^d)$. By (4.1.6),

$$\int_0^{2s} |\tilde{h}|^2 dt \leqslant (2s)^2 C^1 \int_0^{2s} |\dot{\tilde{h}}|^2 dt.$$

Hence, since

$$\int_s^{2s} |h(t-s)|^2 dt = \int_0^s |h(t)|^2 dt = \int_0^s |h(s-t)|^2 dt,$$

we get

$$\int_0^s |h(t)|^2 dt \leqslant 2s^2 C^1 \int_0^s |\dot{h}(t)|^2 dt. \tag{4.1.7}$$

Keeping in mind that

$$\sup\{|\nabla^2_{qq} H(q,p) \mid q \in \mathbb{T}^d, |p| \leqslant l(B)\} < \infty,$$

and $\dot{x}(t) = \nabla_p H(x(t),p(t))$, with $|p(t)| \leqslant B$, $0 \leqslant t \leqslant s$, the fact that $L \in C^2(\mathbb{T}^d,\mathbb{R}^d)$ gives us that there is a constant $\beta_1 \neq 0$ such that

$$\int_0^s (\nabla^2_{qq} L(x(t),\dot{x}(t)) \cdot h(t)) \cdot h(t) \geqslant -\beta_1^2 \int_0^s |h(t)|^2 dt. \tag{4.1.8}$$

Since

$$v(x,p) = \nabla_p H(x,p(x,v)),$$

we have

$$\nabla_v p(x,v) = (\nabla^2_{pp} H(x,p))^{-1}; \tag{4.1.9}$$

hence

$$\nabla^2_{vq} L(x,v) = -\nabla_v \nabla_q H(x,p(x,v)) = -\nabla^2_{pq} H(x,p) \cdot [\nabla^2_{pp} H(x,p)]^{-1},$$

so $\nabla^2_{vq} L(x(t),\dot{x}(t)) = -\nabla^2_{pq} H(x(t),p(t)) \cdot [\nabla^2_{pp} H(x(t),p(t))]^{-1}$. The mapping $p \mapsto \nabla^2_{pp} H(x,p)$ is continuous, so the compactness of $[0,s]$ implies there is a constant $\beta_2 \neq 0$, depending only on the coefficients, such that $|\nabla^2_{qv} L(x(t),\dot{x}(t))| \leqslant \beta_2^2$, for $0 \leqslant t \leqslant s$. Using Cauchy's inequality,

$$\int_0^s 2\nabla^2_{vq}(L(x(t),\dot{x}(t)) \cdot h(t)) \cdot \dot{h}(t)\ dt \geqslant -\beta_2^2 \big(\varepsilon \int_0^s |\dot{h}(t)|^2 dt + \frac{1}{4\varepsilon} \int_0^s |h(t)|^2 dt\big), \tag{4.1.10}$$

with arbitrary $\varepsilon > 0$. As for g, since $h(s) = 0$ we get, from the fundamental theorem of calculus and the Cauchy-Schwarz inequality that

$$|h(0)| \leqslant \int_0^s |\dot{h}(t)| dt \leqslant \sqrt{s}\Big(\int_0^s |\dot{h}(t)|^2 dt\Big)^{1/2}.$$

Therefore, since $\|\nabla^2 U^0\| \leqslant \kappa$, and (4.1.8) and (4.1.10) hold, we get

$$\begin{aligned} B(h(\cdot)) \ &\geqslant -\beta_1^2 \|h(\cdot)\|^2_{L^2(0,s)} - \beta_2^2 (\varepsilon \|\dot{h}(\cdot)\|^2_{L^2(0,s)} + \frac{1}{4\varepsilon}\|h(\cdot)\|^2_{L^2(0,s)}) \\ &+ \int_0^s (\nabla^2_{vv} L(x(t), \dot{x}(t)) \cdot \dot{h}(t)) \cdot \dot{h}(t)\ dt - \kappa s \|\dot{h}(\cdot)\|^2_{L^2(0,s)}. \end{aligned}$$

To deal with $\nabla^2_{vv} L$, we use the fact that $\nabla^2_{pp} H(x,p)$ and $\nabla^2_{vv} L(x,v)$ are inverses of each other (indeed, from $p(x,v) = \nabla_v L(x,v)$, $\nabla_p v(x,p) = (\nabla^2_{vv} L(x,v))^{-1}$ is obtained; compare with (4.1.9)). Let

$$C_0 := \max_{x \in \mathbb{T}^d, |p| \leqslant B} \{\nabla^2_{pp} H(x,p) \cdot w \cdot w \mid w \in \mathbb{R}^d, |w| = 1\}.$$

Setting $D_0 := 1/C_0$, we obtain[2]

$$\nabla^2_{vv} L(x(t), \dot{x}(t)) \cdot w \cdot w \geqslant D_0 |w|^2, \quad w \in \mathbb{R}^d,\ 0 \leqslant t \leqslant s,$$

The constant D_0 depends only on the coefficients. Employing now (4.1.7) and the latter inequality in the estimate of $B(h)$ we obtain

$$B(h) \geqslant [(-\beta_1^2 - \frac{\beta_2^2}{4\varepsilon})(2s)^2 C^1 - \beta_2^2 \varepsilon - \kappa s + D_0] \|\dot{h}\|^2_{L^2(0,s)}$$

The number ε can be chosen so that $D_0 - \beta_2^2 \varepsilon > 0$. But then, since $s \leqslant T$ is small, the full coefficient of $\|\dot{h}\|_{L^2(0,s)}$ in the inequality is positive, say, β^2. Thus, we have

$$g(\tau) = \mathcal{A}(\delta_{x+\tau h}) = \mathcal{A}(\delta_x) + \frac{1}{2} B(h) \tau^2 \geqslant \mathcal{A}(\delta_x) + \frac{1}{2}\tau^2 \beta^2 \|\dot{h}\|^2_{L^2(0,s)}.$$

Case of general $n \in \mathbb{Z}^+$. This is the case of n particles, which means, $\sigma(t) = \frac{1}{n}\sum_{j=1}^n \delta_{y_j(t)}$ and $v(t) = \sum_1^n \dot{y}_j(t) \chi_{y_j(t)}$, and the expression for the Lagrangian is

$$\mathcal{L}(\sigma(t), v(t)) = \frac{1}{n}\sum_{j=1}^n L(y_j(t), \dot{y}_j(t)) + \frac{1}{2n^2}\sum_{j,k=1}^n \phi(y_j(t) - y_k(t)).$$

The corresponding action $\mathcal{A}$ on $AC^2(0, s; \mathbb{T}^{d\times n})$ is

$$\begin{aligned} \mathcal{A}(s; \sigma, \dot{\sigma}) = &\int_0^s \Big[\frac{1}{n}\sum_{j=1}^n L(y_j, \dot{y}_j) + \frac{1}{2n^2}\sum_{j,k=1}^n \phi(y_j(t) - y_k(t))\Big] dt \\ &+ \frac{1}{n}\sum_{j=1}^n U^0(y_j(0)) + \frac{1}{2n^2}\sum_{k,j=1}^n U^1(y_j(0) - y_k(0)). \end{aligned}$$

[2]Because

$$1 = |w|^2 = w^T \nabla^2_{vv} L(x,v) \nabla^2_{pp} H(x,p) w = (w^T \nabla^2_{vv} L(x,v) w)(w^T \nabla^2_{pp} H(x,p) w)$$

where w^T denotes the transpose of the unit vector w.

Now we let $x(t) = (x_1(t), \ldots, x_n(t))$, $0 \leqslant t \leqslant s$ be again as in (4.1.5), thus, $\frac{1}{n}\sum_{j=1}^n \delta_{x_j(\cdot)}$ is a critical point of $\mathcal{A}$ restricted to $\mathcal{C}_n$, subject to $\sigma(s) = \sum_{j=1}^n \delta_{x_j}$ (i.e., $x(s) = x = (x_1, \ldots, x_n)$). Let $h \in AC^2(0, s; \mathbb{T}^{d\times n})$ with $h(s) = 0$. Set

$$g_n(\tau) := \mathcal{A}\big(s; \frac{1}{n}\sum_{j=1}^n \delta_{x_j(\cdot)+\tau h_j(\cdot)}, \sum_{j=1}^n \dot{x_j}(\cdot)\chi_{x_j(\cdot)}\big), \quad \tau \in \mathbb{R}.$$

Define

$$\begin{aligned}
B_n(h) &:= \frac{d^2 g_n(0)}{d\tau^2} \\
&= \int_0^s \big[\frac{1}{n}\sum_{j=1}^n (\nabla^2_{qq}L(x_j,\dot{x}_j)\cdot h_j)\cdot h_j + \frac{2}{n}\sum_{j=1}^n (\nabla^2_{vq}L(x_j,\dot{x}_j)\cdot h_j)\cdot \dot{h}_j + \frac{1}{n}\sum_{j=1}^n (\nabla^2_{v}L(x_j,\dot{x}_j)\cdot \dot{h}_j)\cdot \dot{h}_j \\
&\qquad \frac{1}{2n^2}\sum_{j,k=1}^n (\nabla^2\phi(x_j-x_k)\cdot(h_j-h_k))\cdot(h_j-h_k)\big]dt \\
&\quad + \frac{1}{n}\sum_{j=1}^n (\nabla^2 U^0(x_j(0))\cdot h_j(0))\cdot h_j(0) \\
&\quad + \frac{1}{2n^2}\sum_{j,k=1}^n (\nabla^2 U^1(x_j(0)-x_k(0))\cdot(h_j(0)-h_k(0))\cdot(h_j(0)-h_k(0)).
\end{aligned}$$

Using the single-particle estimate obtained above for each $j = 1, \ldots, n$ we get

$$\begin{aligned}
B_n(h) &\geqslant \frac{1}{n}\sum_{j=1}^n \beta^2\|\dot{h}_j\|^2_{L^2(0,s)} - \frac{1}{2n^2}\sum_{j,k=1}^n \kappa\|h_j-h_k\|^2_{L^2(0,s)} \\
&\quad - \frac{1}{n}\sum_{j=1}^n \kappa s\|\dot{h}_j\|^2_{L^2(0,s)} - \frac{1}{2n^2}\sum_{j,k=1}^n \kappa s\|\dot{h}_j-\dot{h}_k\|^2_{L^2(0,s)} \\
&\geqslant \frac{1}{n}\sum_{j=1}^n \beta^2\|\dot{h}_j\|^2_{L^2(0,s)} - \frac{2C^1 s^2\kappa}{n^2}\sum_{j,k=1}^n \|\dot{h}_j-\dot{h}_k\|^2_{L^2(0,s)} \\
&\quad - \frac{\kappa s}{n}\sum_{j=1}^n \|\dot{h}_j\|_{L^2(0,s)} - \frac{\kappa s}{2n^2}\sum_{k,j=1}^n \|\dot{h}_j-\dot{h}_k\|^2_{L^2(0,s)} \\
&\geqslant \frac{1}{n}\sum_{j=1}^n \beta^2\|\dot{h}_j\|^2_{L^2(0,s)} - \left(\frac{4C^1 s^2 n\kappa}{n^2} + \frac{\kappa s}{n} + \frac{2n\kappa s}{n^2}\right)\sum_{j=1}^n \|\dot{h}_j\|^2_{L^2(0,s)} \\
&= \frac{1}{n}[\beta^2 - \kappa s(4C^1 s + 3)]\sum_{j=1}^n \|\dot{h}_j\|^2_{L^2(0,s)},
\end{aligned}$$

where we have used $\sum_{k,j}\|\dot{h}_j - \dot{h}_k\|^2 = \sum_{k\neq j}\|\dot{h}_j - \dot{h}_k\|^2 \leqslant \sum_{j\neq k} 2(\|\dot{h}_j\|^2 + \|\dot{h}_k\|^2) = 2n\sum_j \|\dot{h}_j\|^2$. Since s is small, we have that $\beta^2 - \kappa s(4C^1 s + 3)$ is a positive number, *still denoted* by β^2. Since $(d/d\tau)g_n(0) = 0$ because $\frac{1}{n}\sum_{j=1}^n \delta_{x_j(\cdot)}$ is a critical point, we have proved that

$$g_n(\tau) = \mathcal{A}(\frac{1}{n}\sum_{j=1}^n \delta_{x_j(\cdot)+\tau h_j(\cdot)}) \geqslant \mathcal{A}(\frac{1}{n}\sum_{j=1}^n \delta_{x_j(\cdot)}) + \frac{1}{2}\tau^2\beta^2\frac{1}{n}\sum_{j=1}^n \|\dot{h}_j\|^2_{L^2(0,s)}.$$

Putting $y(\cdot) := x(\cdot) + h(\cdot)$ and denoting by σ^y, σ^x the respective paths in $\mathscr{P}(\mathbb{T}^d)$ defined by the paths $y(\cdot)$ and $x(\cdot)$ in $AC^2(0, s; (\mathbb{T}^d)^n)$, employing (4.1.7) once more, we get, with $\tau = 1$,

$$\mathcal{A}(s; \sigma^y, \dot{\sigma}^y) - \mathcal{A}(s; \sigma^x, \dot{\sigma}^x) \geqslant \frac{\beta^2}{4ns^2C^1} \sum_{j=1}^{n} \|y_j(\cdot) - x_j(\cdot)\|^2_{L^2(0,s)},$$

holding for every $y \in AC^2(0, s; (\mathbb{T}^d)^n)$. This concludes the proof of (4.1.2), because the latter inequality implies

$$\mathcal{A}(s; \sigma^y, \dot{\sigma}^y) - \mathcal{A}(s; \sigma^x, \dot{\sigma}^x) \geqslant \frac{\beta^2}{4s^2C^1} \int_0^s \mathscr{W}^2(\sigma^x(t), \sigma^y(t))dt. \tag{4.1.11}$$

□

4.1.2 The general case

Now we wish to start with an arbitrary Borel probability measure μ on $\mathbb{T}^d$, and we are going to define $\sigma \in AC^2(0, s; \mathscr{P}_{r'}(\mathbb{T}^d))$ and its corresponding velocity field v as in (3.2.6). Using an approximation argument, we are going to show that (σ, v) solves (4.1.1). To do this, we will need the following, general lemma.

Lemma 4.1.3. *Let $\mu \in \mathscr{P}(\mathbb{T}^d)$ and let $\sigma \in AC^2(0, s; \mathscr{P}_{r'}(\mathbb{T}^d))$ be a path of velocity w such that $\sigma_s = \mu$. Then there exists a sequence $\{\sigma^m\}_{m=1}^{\infty}$, $\sigma^m \in AC^2(0, s; \mathcal{C}_m)$ (see Lemma 4.1.2 for $\mathcal{C}_m$), with corresponding velocity vector fields w^m, $m \in \mathbb{N}$, and a sequence of real numbers $\{r_m\}_1^{\infty} \subset (0, 1)$, $r_m \searrow 0$, such that*

$$\sup_{0 \leqslant t \leqslant s} \mathscr{W}_{r'}(\sigma_t, \sigma_t^m) \leqslant r_m \ ,$$

$$\int_0^s \int_{\mathbb{T}^d} L(q, w_t^m(q))\sigma_t^m(dq)dt \leqslant \int_0^s \int_{\mathbb{T}^d} L(q, w_t(q))\sigma_t(dq)dt + r_m.$$

Here

$$\sigma_t^m = \frac{1}{m} \sum_{j=1}^{m} \delta_{y_j^m(t)}, \qquad w_t^m = \sum_{j=1}^{m} \dot{y}_j^m(t)\chi_{y_j^m(t)},$$

so

$$\int_0^s \int_{\mathbb{T}^d} L(q, w_t^m(q))\sigma_t^m(dq)dt = \int_0^s \frac{1}{m} \sum_{j=1}^{m} L(y_j^m(t), \dot{y}_j^m(t))dt.$$

Proof. We are first going to approximate σ by measures with smooth density (and w by a smooth field). To this end, define the periodic mollifying kernel, following [GT14], as

$$\mathcal{E}^{\varepsilon}(x) = \sum_{k \in \mathbb{Z}^d} \eta^{\varepsilon}(x + k), \quad x \in \mathbb{R}^d, \ \varepsilon > 0,$$

where

$$\eta^{\varepsilon}(x) = \frac{1}{\varepsilon^d}\eta(x/\varepsilon),$$

with

$$\eta(x) = \frac{1}{(4\pi)^{d/2}} \exp(-|x|^2/4).$$

Set

$$\rho_t^\varepsilon := \sigma_t * \mathcal{E}^\varepsilon, \quad w_t^\varepsilon := \frac{(w_t \sigma_t) * \mathcal{E}^\varepsilon}{\rho_t^\varepsilon}, \quad \sigma_t^\varepsilon := \rho_t^\varepsilon|_{[0,1)^d}. \tag{4.1.12}$$

We won't distinguish between measures that are absolutely continuous with respect to Lebesgue measure, and their densities. The measures and functions defined in (4.1.12) are periodic, in the sense explained in [GT14, Sec. 2].

Step 1. (a) We claim that

$$\mathscr{W}_{r'}(\sigma_t^\varepsilon, \sigma_t) \to 0 \quad \text{uniformly on } 0 \leqslant t \leqslant s. \tag{4.1.13}$$

Since all Wasserstein distances on $\mathscr{P}(\mathbb{T}^d)$ are equivalent, it suffices to show (4.1.13) for $r' = 2$. According to [GT14], if $\mu, \nu \in \mathscr{P}(\mathbb{T}^d)$, then

$$\mathscr{W}^2(\mu, \nu) = \sup \Big\{ \int_{\mathbb{R}^d} \zeta d\mu + \int_{\mathbb{R}^d} \theta d\nu \;\Big|\; \zeta, \theta \in C(\mathbb{T}^d), \zeta(x) + \theta(y) \leqslant |x-y|_{\mathbb{T}^d}^2, (x,y) \in \mathbb{R}^d \times \mathbb{R}^d \Big\}.$$

A calculation shows that if $\varphi \in C(\mathbb{T}^d)$, then

$$\mathcal{E}^\varepsilon * (\varphi \chi_{[0,1)^d}) = \eta^\varepsilon * \varphi. \tag{4.1.14}$$

Thus, one verifies that σ_t^ε has mass 1, by setting $\varphi \equiv 1$ in (4.1.14). Formula (4.1.14) also implies, in particular, that

$$\int_{[0,1)^d} \mathcal{E}^\varepsilon(x-y) dy = 1$$

for any $x \in \mathbb{R}^d$ and $\varepsilon > 0$. Let $\zeta, \theta \in C(\mathbb{T}^d)$. We have

$$\begin{aligned}
\int_{\mathbb{R}^d} \zeta(x) \sigma_t^\varepsilon(dx) &= \int_{\mathbb{R}^d} \zeta(x) (\rho_t^\varepsilon|_{[0,1)^d})(dx) = \int_{\mathbb{R}^d} \zeta(x) \chi_{[0,1)^d}(x) \int_{\mathbb{R}^d} \mathcal{E}^\varepsilon(x-y)] \sigma_t(dy) dx \\
&= \int_{\mathbb{R}^d} \zeta(x) \chi_{[0,1)^d}(x) (\sigma_t * \mathcal{E}^\varepsilon)(dx) = \int_{\mathbb{R}^d} (\mathcal{E}^\varepsilon * \zeta \chi_{[0,1)^d})(x) \sigma_t(dx) \\
&= \int_{\mathbb{R}^d} (\eta^\varepsilon * \zeta)(x) \sigma_t(dx),
\end{aligned}$$

by (4.1.14). Then

$$\begin{aligned}
\int_{\mathbb{R}^d} \zeta(x) \sigma_t^\varepsilon(dx) + \int_{\mathbb{R}^d} \theta(y) \sigma_t(dy) &= \int_{\mathbb{R}^d} (\eta^\varepsilon * \zeta)(x) \sigma_t(dx) + \int_{\mathbb{R}^d} \theta(y) \sigma_t(dy) \\
&= \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} \eta^\varepsilon(y) [\zeta(x-y) + \theta(x)] dy \Big) \sigma_t(dx).
\end{aligned}$$

Suppose now that ζ and θ are such that $\zeta(x) + \theta(y) \leqslant |x-y|_{\mathbb{T}^d}^2$. This implies $\zeta(x) + \theta(y) \leqslant |x-y|^2$. Hence,

$$\mathscr{W}(\sigma_t^\varepsilon, \sigma_t) \leqslant \Big(\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \eta^\varepsilon(y) |y|^2 dy \sigma_t(dx) \Big)^{1/2} = \varepsilon \Big(\int_{\mathbb{R}^d} \eta(y) |y|^2 dy \Big)^{1/2},$$

which proves (4.1.13).

(b) Next, we prove that

$$\limsup_{\varepsilon\to 0}\int_0^s\int_{\mathbb{T}^d} L(q,w_t^\varepsilon)\sigma_t^\varepsilon(dq)dt \leqslant \int_0^s\int_{\mathbb{T}^d} L(q,w_t(q))\sigma_t(dq)dt. \tag{4.1.15}$$

Write

$$w_t^\varepsilon(q) = \frac{\int_{\mathbb{R}^d}\mathcal{E}^\varepsilon(q-y)w_t(y)\sigma_t(dy)}{\int_{\mathbb{R}^d}\mathcal{E}^\varepsilon(q-y)\sigma_t(dy)} = \int_{\mathbb{R}^d} w_t(y)j^\varepsilon(q,y)\sigma_t(dy)$$

where

$$j^\varepsilon(q,y) := \frac{\mathcal{E}^\varepsilon(q-y)}{\int_{\mathbb{R}^d}\mathcal{E}^\varepsilon(q-y)\sigma_t(dy)}.$$

Since $\int_{\mathbb{R}^d} j^\varepsilon(q,y)\sigma_t(dy) = 1$, $j^\varepsilon(q,\cdot)\sigma_t$ is a probability measure on $\mathbb{R}^d$ for every $q\in\mathbb{T}^d$, moreover, $j^\varepsilon(q,\cdot)$ is periodic, so $j^\varepsilon(q,\cdot)\sigma_t\in\mathscr{P}(\mathbb{T}^d)$ for every $q\in\mathbb{T}^d$. Fix $q\in\mathbb{T}^d$. Since $L(q,\cdot)$ is convex, we have

$$L(q,w_t^\varepsilon(q)) = L(q,\int_{\mathbb{R}^d} w_t(y)j^\varepsilon(q,y)\sigma_t(dy)) \leqslant \int_{\mathbb{R}^d} L(q,w_t(y))j^\varepsilon(q,y)\sigma_t(dy) \tag{4.1.16}$$

by Jensen's inequality. For every q, the expression $\int_{\mathbb{R}^d} L(q,w_t(y))j^\varepsilon(q,y)\sigma_t(dt)$ is finite because $|L(q,w_t(y))| \leqslant c_5(1+|w_t(y)|^{r'})$, j^ε is bounded and $w_t\in L^{r'}(\sigma_t)$. Observe that the function $q\mapsto \int_{\mathbb{R}^d} L(q,w_t(y))j^\varepsilon(q,y)\sigma_t(dy)$ belongs to $L^1(\sigma_t^\varepsilon)$:

$$\begin{aligned}
&\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|L(q,w_t(y))|j^\varepsilon(q,y)\sigma_t(dy)\sigma_t^\varepsilon(dq) \leqslant \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} c_5(1+|w_t(y)|^{r'})j^\varepsilon(q,y)\sigma_t(dy)\sigma_t^\varepsilon(dq)\\
&= c_5 + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|w_t(y)|^{r'}j^\varepsilon(q,y)\sigma_t(dy)\sigma_t^\varepsilon(dq) = c_5 + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|w_t(y)|^{r'}j^\varepsilon(q,y)\sigma_t^\varepsilon(dq)\sigma_t(dy)\\
&= c_5 + \int_{\mathbb{R}^d}|w_t(y)|^{r'}\int_{[0,1)^d}\frac{\mathcal{E}^\varepsilon(q-y)}{\int_{\mathbb{R}^d}\mathcal{E}^\varepsilon(q-y)\sigma_t(dy)}\rho_t^\varepsilon(dq)\sigma_t(dy)\\
&= c_5 + \int_{\mathbb{R}^d}|w_t(y)|^{r'}\Big[\int_{[0,1)^d}\mathcal{E}^\varepsilon(q-y)dq\Big]\sigma_t(dy)\\
&\leqslant c_5 + C\|w_t\|^{r'}_{L^{r'}(\sigma_t)}.
\end{aligned}$$

Hence

$$\begin{aligned}
\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} L(q,w_t(y))j^\varepsilon(q,y)\sigma_t(dy)\sigma_t^\varepsilon(dq) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} L(q,w_t(y))j^\varepsilon(q,y)\sigma_t^\varepsilon(dq)\sigma_t(dy)\\
&= \int_{\mathbb{R}^d}\int_{[0,1)^d} L(q,w_t(y))\mathcal{E}^\varepsilon(q-y)dq\sigma_t(dy).
\end{aligned}$$

Thus, integrating (4.1.16) on both sides, we get

$$\int_{\mathbb{R}^d} L(q,w_t^\varepsilon(q))\sigma_t^\varepsilon(dq) \leqslant \int_{\mathbb{R}^d}\int_{[0,1)^d} L(q,w_t(y))\mathcal{E}^\varepsilon(q-y)dq\sigma_t(dy). \tag{4.1.17}$$

We claim that

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^d}\int_{[0,1)^d} L(q,w_t(y))\mathcal{E}^\varepsilon(q-y)dq\sigma_t(dy) = \int_{\mathbb{R}^d} L(y,w_t(y))\sigma_t(dy),$$

being clear that from this claim follows, because of (4.1.17), that

$$\limsup_{\varepsilon\to 0}\int_{\mathbb{T}^d} L(q,w_t^\varepsilon)\sigma_t^\varepsilon(dq) \leqslant \int_{\mathbb{T}^d} L(q,w_t(q))\sigma_t(dq),$$

which, in turn, by Fatou's lemma, yields (4.1.15).

Proof of the claim: We have

$$\begin{aligned}
&\Big|\int_{\mathbb{R}^d}\big(L(y,w_t(t)) - \int_{[0,1)^d} L(q,w_t(y))\mathcal{E}^\varepsilon(q-y)dq\big)\sigma_t(dy)\Big| \\
\leqslant\ & \int_{\mathbb{R}^d}\int_{[0,1)^d} |L(y,w_t(y)) - L(q,w_t(y))|\mathcal{E}^\varepsilon(q-y)dq\sigma_t(dy) \\
=\ & \int_{\mathbb{R}^d}\int_{[0,1)^d} |l(q,y)|\mathcal{E}^\varepsilon(q-y)dq\sigma_t(dy),
\end{aligned}$$

where $l(q,y) := L(q,w_t(y)) - L(y,w_t(y))$, $q \in \mathbb{T}^d$, σ_t-a.e. $y \in \mathbb{R}^d$. For σ_t-a.e. $y \in \mathbb{R}^d$, $l(\cdot,y)$ is continuous and $\lim_{q\to y} l(q,y) = 0$. Now, fix $y \in \mathbb{R}^d$. Then

$$\int_{[0,1)^d} |l(q,y)|\mathcal{E}^\varepsilon(q-y)dq = \mathcal{E}^\varepsilon * |l(\cdot,y)|\chi_{[0,1)^d} = (\eta^\varepsilon * |l(\cdot,y)|)(y) = \int_{\mathbb{R}^d} |l(q,y)|\frac{1}{(4\pi)^{d/2}\varepsilon^d}e^{-\frac{|q-y|^2}{4\varepsilon^2}}dq,$$

where formula (4.1.14) has been used. Let $\epsilon > 0$. And let $\delta > 0$ be such that $|l(q,y)| < \epsilon$ if $|q-y| < \delta$. Then

$$\begin{aligned}
\int_{[0,1)^d} |l(q,y)|\mathcal{E}^\varepsilon(q-y)dq =\ & \int_{B(y;\delta)} |l(q,y)|\frac{1}{(4\pi)^{d/2}\varepsilon^d}e^{-\frac{|q-y|^2}{4\varepsilon^2}}dq \\
&+ \int_{B(y;\delta)^c} |l(q,y)|\frac{1}{(4\pi)^{d/2}\varepsilon^d}e^{-\frac{|q-y|^2}{4\varepsilon^2}}dq \\
\leqslant\ & \epsilon + \int_{B(y;\delta)^c}|l(q,y)|\frac{1}{(4\pi)^{d/2}\varepsilon^d}e^{-\frac{|q-y|^2}{4\varepsilon^2}}dq = \epsilon + \int_{B(0;\delta/\varepsilon)^c}|l(q+y,y)|\frac{1}{(4\pi)^{d/2}}e^{-\frac{|q|^2}{4}}dq \\
\leqslant\ & \epsilon + \|l(\cdot,y)\|_{L^\infty(\mathbb{T}^d)}\int_{B(0;\delta/\varepsilon)}\frac{1}{(4\pi)^{d/2}}e^{-\frac{|q|^2}{4}}dq \underset{\varepsilon\to 0}{\longrightarrow} \epsilon.
\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, we obtain

$$\lim_{\varepsilon\to 0}\int_{[0,1)^d} |l(q,y)|\mathcal{E}^\varepsilon(q-y)dq = 0.$$

Also,

$$\int_{[0,1)^d} |l(q,y)|\mathcal{E}^\varepsilon(q-y)dq \leqslant \int_{[0,1)^d} 2c_5(1+|w_t(y)|^{r'})\mathcal{E}^\varepsilon(q-y)dq \leqslant 2c_5(1+|w_t(y)|^{r'}) \in L^1(\sigma_t).$$

Hence

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^d}\int_{[0,1)^d} |l(q,y)|\mathcal{E}^\varepsilon(q-y)dq\sigma_t(dy) = 0.$$

This proves the claim, and, in turn, the veracity of (4.1.15).

Step 2. The constructed σ_t^ε solve the continuity equation

$$\partial_t \sigma_t^\varepsilon + \operatorname{div}(w_t^\varepsilon \sigma_t^\varepsilon) = 0 \qquad \text{in } \ (0,s) \times \mathbb{T}^d \ ,$$

because, by construction,

$$\operatorname{div}(w_t^\varepsilon \sigma_t^\varepsilon) = \operatorname{div}((v_t\sigma_t) * \mathcal{E}^\varepsilon \chi_{[0,1)^d}) = (\operatorname{div}(w_t\sigma_t)) * \mathcal{E}^\varepsilon \chi_{[0,1)^d}.$$

Thus,

$$\partial_t \sigma_t(q) = -\sigma_t^\varepsilon(q) \ \operatorname{div} \ w_t^\varepsilon(q), \qquad q \in \mathbb{T}^d, \ 0 < t < s.$$

Step 3. Because each $(\sigma_t^\varepsilon, w_t^\varepsilon)$ is smooth, for arbitrary $\delta > 0$ we can find $\sigma_t^{\varepsilon,n} \in AC^2(0,s;\mathcal{C}_n)$, $\sigma_t^{\varepsilon,n} = \frac{1}{n}\sum_{j=1}^n \delta_{y_j^{\varepsilon,n}(t)}$, such that

$$\sup_{0 \leqslant t \leqslant s} \mathscr{W}(\sigma_t^\varepsilon, \sigma_t^{\varepsilon,n}) \leqslant \delta,$$

$$\int_0^s \frac{1}{n}\sum_{j=1}^n L(y_j^{\varepsilon,n}(t), \dot{y}_j^{\varepsilon,n}(t))dt = \int_0^s \int_{\mathbb{T}^d} L(q, w_t^{\varepsilon,n}(q))\sigma_t^{\varepsilon,n}(dq)dt \leqslant \int_0^s \int_{\mathbb{T}^d} L(q, w_t^\varepsilon(q))\sigma_t^\varepsilon(dq)dt + \delta.$$

Finally, it is clear that combining the latter inequalities, together with (4.1.15) and (4.1.13), give the statement. □

Theorem 4.1.4. *Let* $0 < s < T$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, $\sigma \in AC^2(0,s;\mathscr{P}_{r'}(\mathbb{T}^d))$ *and its corresponding velocity field* v *be as in (3.2.6). Then* (σ, v) *is the unique minimizer for (4.1.1) up to a* $\mathscr{L}^1$*-zero measure subset of* $(0,s)$, *i.e., if*

$$\mathcal{A}(s;\sigma', v') = \inf_{(\bar\sigma,\bar v)} \left\{ \mathcal{A}(s;\bar\sigma,\bar v) \ \middle| \ \bar\sigma_s = \mu \right\}$$

then there is $I \subset (0,s)$ *such that* $\mathscr{L}^1(I) = 0$ *and* $\sigma'_t = \sigma_t$ *for every* $t \in (0,s)\backslash I$.

Proof. Let $(\bar\sigma, \bar v)$ be an arbitrary admissible pair for the minimization problem. Thus, in particular, $\bar\sigma(s) = \sigma(s) = \mu$. Let $\{\bar\sigma^m\}_{m=1}^\infty$ be the corresponding sequence of approximating discrete paths provided by Lemma 4.1.3, with the following notation:

$$\bar\sigma^m(t) = \frac{1}{m}\sum_{j=1}^m \delta_{\bar y_j^m(t)}, \quad 0 \leqslant t \leqslant s \ ;$$

$$\sup_{0 \leqslant t \leqslant s} \mathscr{W}_{r'}(\bar\sigma_t, \bar\sigma_t^m) \leqslant \bar r_m \ ,$$

$$\int_0^s \frac{1}{m}\sum_{j=1}^m L(\bar y_j^m(t), \dot{\bar y}_j^m(t))dt \leqslant \int_0^s \int_{\mathbb{T}^d} L(q, \bar v_t(q))\bar\sigma_t(dq)dt + \bar r_m,$$

with $\bar r_m \searrow 0$. It follows from this that, since $\mathcal{F}, \mathcal{G}$ are κ-Lipschitz,

$$\mathcal{A}(s;\bar\sigma^m, \dot{\bar\sigma}^m) \leqslant \mathcal{A}(s;\bar\sigma,\bar v) + \bar r^m, \tag{4.1.18}$$

for $m = 1, \dots$ Define

$$\sigma^m(t) = \Sigma_t^1[s, \bar{\sigma}^m(s)]_\# \bar{\sigma}^m(s), \quad 0 \leqslant t \leqslant s,$$

for $m = 1, \dots$ Recall that $\mathcal{C}_m \subset AC^2(0, s; \mathscr{P}_{r'}(\mathbb{T}^d))$ denotes the subset of paths ν_t of the form $\nu_t = \frac{1}{m} \sum_{j=1}^m \delta_{y_j(t)}$ where $y_j(\cdot) \in AC^2(0, s; \mathbb{T}^d)$, $j = 1, \dots, m$.

1. Note that

$$\bar{\sigma}^m, \sigma^m \in \mathcal{C}_m ;$$
$$\sigma_s^m = \bar{\sigma}_s^m,$$

and σ^m is the path obtained by following the flow lines of Σ^1 from $\bar{\sigma}_s^m$, the configuration of $\bar{\sigma}^m$ at time $t = s$. The corresponding velocity vector field for σ^m is

$$v_t^m = \partial_t \Sigma_t^1[s, \bar{\sigma}^m(s)] \circ X_t[s, \bar{\sigma}^m(s)],$$

for $m = 1, \dots$ and $0 \leqslant t \leqslant T$. Since $\bar{\sigma}_s^m \to \mu = \sigma_s^m$, we know, by Proposition 3.2.12, that $v_t^m \to v_t$ in $C[\mathbb{T}^d; \mathbb{T}^d]$ uniformly with respect to t.

2. We have

$$\begin{aligned}
&\int_0^s \mathcal{L}(\sigma_t^m, \dot{\sigma}_t^m) dt - \int_0^s \mathcal{L}(\sigma_t, v_t) dt \\
&= \int_0^s \Big[\int_{\mathbb{T}^d} L(q, v_t^m(q)) \sigma_t^m(dq) - \mathcal{F}(\sigma_t^m) \Big] dt - \int_0^s \Big[\int_{\mathbb{T}^d} L(q, v_t(q)) \sigma_t(dq) - \mathcal{F}(\sigma_t) \Big] dt \\
&= \int_0^s \Big[\int_{\mathbb{T}^d} L(q, v_t^m \circ \Sigma_t^1[s, \bar{\sigma}_s^m](q)) \bar{\sigma}_s^m(dq) - \int_{\mathbb{T}^d} L(q, v_t \circ \Sigma_t^1[s, \mu](q)) \mu(dq) - (\mathcal{F}(\bar{\sigma}_t^m) - \mathcal{F}(\sigma_t)) \Big] dt.
\end{aligned}$$

From the uniform convergence properties shown earlier, it follows that the latter expression approaches 0 as $m \to \infty$. Therefore,

$$\mathcal{A}(s; \sigma^m, v^m) \to \mathcal{A}(s; \sigma, v) \quad \text{as } m \to \infty.$$

3. We now invoke inequality (4.1.11), which was obtained for discrete paths having the same value at time $t = s$, one of which was the solution to the minimization problem in $\mathcal{C}_m$, with $n = m$ there (see part 2 of the proof of Lemma 4.1.2, which leads to this inequality). Note that, by the equivalence of metrics $\mathscr{W}$ and $\mathscr{W}_{r'}$, we can substitute the former for the latter in (4.1.11). Applying it to σ^m, $\bar{\sigma}^m$:

$$\mathcal{A}(s; \bar{\sigma}^m, \dot{\bar{\sigma}}^m) - \mathcal{A}(s; \sigma^m, v^m) \geqslant \frac{\beta^2}{4s^2 C^1} \int_0^s \mathscr{W}_{r'}^2(\sigma_t^m, \bar{\sigma}_t^m) dt.$$

Then, using (4.1.18), we have

$$\mathcal{A}(s; \bar{\sigma}, \bar{v}) + \bar{r}^m - \mathcal{A}(s; \sigma^m, v^m) \geqslant \frac{\beta^2}{4s^2 C^1} \int_0^s \mathscr{W}_{r'}^2(\sigma_t^m, \bar{\sigma}_t^m) dt.$$

Letting $m \to \infty$, it follows that

$$\mathcal{A}(s; \bar{\sigma}, \bar{v}) - \mathcal{A}(s; \sigma, v) \geqslant \frac{\beta^2}{4s^2 C^1} \int_0^s \mathscr{W}_{r'}^2(\sigma_t, \bar{\sigma}_t) dt.$$

This concludes the proof. □

For $\mu \in \mathscr{P}(\mathbb{T}^d)$, consider now the problem

$$\inf\Big\{\int_0^s \int_{\mathbb{T}^d} L(Q_t,\dot{Q}_t)\mu(dq) - \mathcal{F}((Q_t)_{\#}\mu)dt + \mathcal{G}((Q_0)_{\#}\mu) \mid Q_s = id,\ Q \in AC^2(0,s;C(\mathbb{T}^d;\mathbb{T}^d))\Big\}, \tag{4.1.19}$$

where id is the identity map on $\mathbb{T}^d$. The problem (4.1.19) is embedded in (4.1.1), in the sense that if we put $\bar{\sigma}_t := (Q_t)_{\#}\mu$, then $\bar{\sigma}_t$ is an admissible path in (4.1.1). In other words, (4.1.1) is a relaxation of (4.1.19). Thus, we have a unique minimizer for (4.1.19), the one provided by the solution to (3.1.3). Every minimizer of (4.1.19) must satisfy the Euler-Lagrange equations

$$\begin{cases} \nabla_q L(Q_t,\dot{Q}_t) = \dfrac{d}{dt}\nabla_v L(Q_t,\dot{Q}_t) + \nabla_q F(Q_t,(Q_t)_{\#}\mu) \quad \text{in } [0,s]\times\mathbb{T}^d\ ; \\ \nabla_v L(Q_0,\dot{Q}_0) = \nabla_q g(Q_0,(Q_0)_{\#}\mu) \quad \text{on } \mathbb{T}^d \\ \qquad\qquad Q_s = id. \end{cases}$$

Through $H = L^*$ and $p := \nabla_v L(x,v)$, these equations are equivalent to (3.1.3). Therefore, since the minimizer of (4.1.19) is unique and it is Σ_t^1, we have deduced:

Corollary 4.1.5. *(An alternative proof of uniqueness of the fixed point) The solution to (3.1.3) obtained in Corollary 3.1.7 is unique. The fixed point of the operator* $\mathfrak{m}^{s,\mu}$ *defined therein is unique.*

For $s \in (0,T), \mu \in \mathscr{P}(\mathbb{T}^d)$, define

$$\begin{aligned} \mathcal{V}[s,\mu](t,q) &:= \Sigma_t^2[s,\mu]\circ X_t[s,\mu](q) \\ &= \Sigma_t^2[s,\mu]\circ(\Sigma_t^1[s,\mu])^{-1}(q), \quad s,t\in[0,T], \mu\in\mathscr{P}(\mathbb{T}^d), q\in\mathbb{T}^d. \end{aligned} \tag{4.1.20}$$

The next list of properties of the mappings $(t,s,\mu)\mapsto\Sigma_t[s,\mu](\cdot)$ is vital both for this chapter and the next.

Proposition 4.1.6. *Let* $0 \leqslant s, t_0 \leqslant T$, $\mu \in \mathscr{P}(\mathbb{T}^d)$. *Set*

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

Then,

(i) For every $0 \leqslant t \leqslant T$*:*

$$\Sigma_t[t_0,\sigma_{t_0}]\circ\Sigma_{t_0}^1[s,\mu] = \Sigma_t[s,\mu]. \tag{4.1.21}$$

(ii) For every $0 \leqslant t \leqslant T$*:*

$$\Sigma_t^1[t_0,\Sigma_{t_0}^1[t,\mu]_{\#}\mu] \quad \text{and} \quad \Sigma_{t_0}^1[t,\mu] \quad \text{are inverses of each other}, \tag{4.1.22}$$

$$v_t[s,\mu] = v_t[t_0,\sigma_{t_0}], \tag{4.1.23}$$

$$\Sigma_t^2[t_0,\sigma_{t_0}]\circ\Sigma_{t_0}^1[s,\mu] = \Sigma_t^2[s,\mu], \tag{4.1.24}$$

$$\partial_s\Sigma_t^1[s,\mu] = -\nabla_q\Sigma_t^1[s,\mu]v_s[t,\sigma_t]. \tag{4.1.25}$$

(iii) If $0 \leqslant \tau, t \leqslant T$, *then*

$$\Sigma_\tau^2[t,\sigma_t]\circ(\Sigma_\tau^1[t,\sigma_t])^{-1} = \Sigma_\tau^2[s,\mu]\circ\Sigma_\tau^1[s,\mu]^{-1}. \tag{4.1.26}$$

Proof. (i) Let

$$\begin{aligned} Q(t,q) &= (\Sigma^1_t[s,\mu] \circ \Sigma^1_{t_0}[s,\mu]^{-1})(q), \\ P(t,q) &= (\Sigma^2_t[s,\mu] \circ \Sigma^1_{t_0}[s,\mu]^{-1})(q), \end{aligned}$$

for $0 \leqslant t \leqslant T$, $q \in \mathbb{T}^d$. 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 (3.1.3) with

$$s = t_0, \quad \mu = \sigma_{t_0}.$$

Since solutions to (3.1.3) 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 (4.1.21).

(ii) Fact (4.1.22) follows readily from (i), by setting $t = s$. For (4.1.23), see [GŚ15, p. 6593]. Formula (4.1.24) is just the second component of (4.1.21).

By (4.1.22), with $t = s$, $t_0 = t$, we have

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

By Lemma 3.2.4, 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

$$\begin{aligned} 0 &= \partial_s \Sigma^1[s,\mu](t,q) + \nabla_q \Sigma^1[s,\mu](t,q) \partial_s \Sigma^1_s[t,\sigma_t](\Sigma^1_s[t,\sigma_t]^{-1}) \\ &= \partial_s \Sigma^1[s,\mu](t,q) + \nabla_q \Sigma^1[s,\mu](t,q) v_s[t,\sigma_t], \end{aligned}$$

which gives (4.1.25).

(iii) For (4.1.26), simply use (4.1.24) with τ in place of t and t in place of t_0:

$$\begin{aligned} \Sigma^2_\tau[t,\sigma_t] \circ (\Sigma^1_\tau[t,\sigma_t])^{-1} &= \Sigma^2_\tau[s,\mu] \circ \Sigma_t[s,\mu]^{-1} \circ \Sigma^1_\tau[t,\sigma_t]^{-1} \\ &= \Sigma^2_\tau \circ (\Sigma^1_\tau[t,\sigma_t] \circ \Sigma_t[s,\mu])^{-1} = \Sigma^2_\tau[s,\mu] \circ \Sigma^1_\tau[s,\mu]^{-1}. \end{aligned}$$

□

Corollary 4.1.7. *Let $0 \leqslant s \leqslant T$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, and, as before, set*

$$\sigma_t = \Sigma^1_t[s,\mu]_\#\mu, \quad v_t = \partial_t \Sigma^1_t[s,\mu] \circ X_t[s,\mu].$$

Then, if $0 < r \leqslant T$ and $\bar\sigma \in AC^2(0,r;\mathscr{P}(\mathbb{T}^d))$, with velocity $\bar v$, satisfies

$$\bar\sigma_r = \sigma_r,$$

then

$$\mathcal{A}(r;\bar\sigma,\bar v) > \mathcal{A}(r;\sigma,v),$$

unless $\bar\sigma = \sigma$, $\mathscr{L}^1$-a.e. on $(0,r)$.

Proof. This is a consequence of Proposition 4.1.6. Let

$$\sigma_t^* := \Sigma_t^1[r, \bar{\sigma}_r]_\# \bar{\sigma}_r, \quad v_t^* := v_t[r, \bar{\sigma}_r], \qquad 0 \leqslant t \leqslant r.$$

By Theorem 4.1.4,

$$\mathcal{A}(r; \bar{\sigma}, \bar{v}) > \mathcal{A}(r; \sigma^*, v^*) \tag{4.1.27}$$

unless $\bar{\sigma}_t = \sigma_t^*$ for $\mathscr{L}^1$-a.e. t in $(0, r)$. However,

$$\begin{aligned} \sigma_t^* &= \Sigma_t^1[r, \bar{\sigma}_r]_\# \bar{\sigma}_r = \Sigma_t^1[r, \bar{\sigma}_r]_\# (\Sigma_r^1[s, \mu]_\# \mu) = \Sigma_t^1[r, \sigma_r]_\# (\Sigma_r^1[s, \mu]_\# \mu) \\ &= \Sigma_t^1[s, \mu]_\# \mu = \sigma_t \ , \end{aligned}$$

$0 \leqslant t \leqslant r$, where the second line follows from the first by Proposition 4.1.6, another consequence of which is

$$v_t^* = v_t[r, \bar{\sigma}_r] = v_t[s, \mu] = v_t.$$

Therefore $(\sigma, v) = (\sigma^*, v^*)$ $\mathscr{L}^1$-a.e. on $(0, r)$, and returning to (4.1.27), the assertion is proven. ☐

4.2 The aggregate value function $\mathcal{U}$ and its Wasserstein gradient

For $s \in [0, T]$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, we define the *value function* $\mathcal{U}$:

$$\mathcal{U}(s, \mu) = \inf \Big\{ \int_0^s \mathcal{L}(\bar{\sigma}, \bar{v}) dt + \mathcal{G}(\bar{\sigma}_0) \mid \bar{\sigma}_s = \mu \Big\}, \tag{4.2.1}$$

where the infimum is taken over the set of all pairs $(\bar{\sigma}, \bar{v})$ such that $\bar{\sigma} \in AC^2(0, s; \mathscr{P}_{r'}(\mathbb{T}^d))$ and $\bar{v}$ is a velocity for $\bar{\sigma}$. By the way $\mathcal{L}$ was defined at the beginning of Section 4.1, $\mathcal{U}(s, \mu)$ can be interpreted as the best *average* cost of displacement for the whole ensemble of players that begins at $t = s$ with distribution μ and is penalized according to the running cost $\mathcal{L}$ and the initial cost $\mathcal{G}(\sigma_0)$ at $t = s$.

From its definition, it follows that $\mathcal{U}$ satisfies the *dynamic programming principle*:

$$\mathcal{U}(s, \mu) = \inf_{(\bar{\sigma}, \bar{v})} \Big\{ \int_r^s \mathcal{L}(\bar{\sigma}, \bar{v}) dt + \mathcal{U}(r, \bar{\sigma}_r) \ \Big| \ \bar{\sigma}_s = \mu \Big\}. \tag{4.2.2}$$

Proposition 4.2.1. *Fix $s \in [0, T]$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, and set*

$$\sigma_t = \Sigma_t^1[s, \mu]_\# \mu, \quad v_t = \partial_t \Sigma_t^1[s, \mu] \circ X_t[s, \mu], \qquad 0 \leqslant t \leqslant T.$$

Then, for any $r \in [0, T]$, we have

$$\mathcal{U}(r, \sigma_r) = \mathcal{A}(r; \sigma, v),$$

in particular, $\mathcal{U}(s, \mu) = \mathcal{A}(s; \sigma, v)$.

Proof. This follows directly from Corollary 4.1.7. ☐

For arbitrary vectors $q = (q_1, \dots, q_n) \in (\mathbb{T}^d)^n$, $v = (v_1, \dots, v_n) \in (\mathbb{R}^d)^n$, let $\mu^q \in \mathscr{P}(\mathbb{T}^d)$ be $\mu^q = \frac{1}{n}\sum_{j=1}^n \delta_{q_j}$, and $v^q \in L^{r'}(\mathbb{T}^d)$ be $v^q(x) = \sum_{j=1}^n v_j \chi_{\{q_j\}}(x)$, $x \in \mathbb{T}^d$. Define

$$L_n(q,v) := \mathcal{L}(\mu^q, v^q) = \int_{\mathbb{T}^d} L(x, v^q(x))\mu^q(dx) - \mathcal{F}(\mu^q) = \frac{1}{n}\sum_{j=1}^n L(q_j, v_j) - \frac{1}{2n^2}\sum_{j,k=1}^n \phi(q_k - q_l);$$

$$G_n(q) := \mathcal{G}(\mu^q) = \frac{1}{n}\sum_{j=1}^n U^0(q_j) + \frac{1}{2n^2}\sum_{j,k=1}^n U^1(q_j - q_k);$$

$$\mathcal{U}^n(t,q) := \inf\left\{ \int_0^s L_n(x(t), \dot{x}(t))dt + G_n(x(0)) \;\middle|\; x(s) = q,\ x(\cdot) \in AC^2(0,s;(\mathbb{T}^d)^n)\right\} .$$

Because of Theorem 4.1.4,

$$\mathcal{U}^n(t,q) = \mathcal{U}(t,\mu^q).$$

We want to obtain semiconcavity inequalities for $\mathcal{U}^n$ in the second variable, using the properties of the coefficients $L, \mathcal{F}, \mathcal{G}$, as a way of discovering what the derivative in measure of the value function $\mathcal{U}$ should be.

We need the following calculations. Since we are assuming that $\mathcal{F}$, $\mathcal{G}$ satisfy the conditions of Section 2.2.3, we have:

$$\begin{aligned}\nabla_{q_j} L_n(q,v) &= \frac{1}{n}\nabla_q L(q_j, v_j) - \frac{1}{n^2}\sum_{k=1}^n \nabla\phi(q_j - q_k) = \frac{1}{n}(\nabla_q L(q_j, v_j) - \nabla\phi * \mu(q_j)) \\ &= \frac{1}{n}(\nabla_q L(q_j, v_j) - \nabla_q F(q_j, \mu^q)) \ ; \end{aligned} \tag{4.2.3}$$

$$\nabla_{q_j} G_n(q) = \frac{1}{n}\nabla U^0(q_j) + \frac{1}{n^2}\sum_{k=1}^n \nabla U^1(q_j - q_k) = \frac{1}{n}\nabla_q g(q_j, \mu^q). \tag{4.2.4}$$

Furthermore,

$$\nabla^2_{q_k q_j} L_n(q,v) = 0 \ \text{ if } j \neq k, \quad \nabla^2_{q_j q_j} L_n(q,v) = \frac{1}{n}(\nabla^2_{qq} L(q_j, v_j) - \nabla^2_{qq} F(q_j, \mu^q)), \tag{4.2.5}$$

$$\nabla^2_{q_k v_j} L_n(q,v) = 0 \ \text{ if } k \neq j, \quad \nabla^2_{q_j v_j} L_n(q,v) = \frac{1}{n}\nabla^2_{qv} L(q_j, v_j), \tag{4.2.6}$$

$$\nabla^2_{v_k v_j} L_n(q,v) = 0 \ \text{ if } k \neq j, \quad \nabla^2_{v_j v_j} L_n(q,v) = \frac{1}{n}\nabla^2_{vv} L(q_j, v_j) \ ; \tag{4.2.7}$$

$$\nabla^2_{q_k q_j} G_n(q) = 0 \ \text{ if } j \neq k, \quad \nabla^2_{q_j q_j} G_n(q) = \frac{1}{n}\nabla^2_{qq} g(q_j, \mu^q).$$

Lemma 4.2.2. *Let $s \in (0,T)$, let $q = (q_1, \dots, q_n) \in (\mathbb{T}^d)^n, q^* = (q_1^*, \dots, q_n^*) \in (\mathbb{T}^d)^n$. Then*

$$\mathcal{U}(s, \mu^{q^*}) \leqslant \mathcal{U}(s, \mu^q) + \int_{\mathbb{T}^d \times \mathbb{T}^d} \Sigma^2_s[s, \mu^q](x)(b - x)\gamma(dx, db) + [\kappa(1+s) + \ell]\mathscr{W}^2(\mu^q, \mu^{q^*}) \ , \tag{4.2.8}$$

$$\mathcal{U}(s,\mu^{q^*}) \geqslant \mathcal{U}(s,\mu^q)+\int_{\mathbb{T}^d\times\mathbb{T}^d}\Sigma^2_s[s,\mu^q](x)(b-x)\gamma(dx,db)-\Big[\frac{1}{3}(T-s)(\kappa+\ell)+\frac{1}{2}\ell+\frac{\ell}{T-s}\Big]\mathscr{W}^2(\mu^q,\mu^{q^*}). \tag{4.2.9}$$

Here

$$\ell = \sup\{|\nabla^2 L(x,v)| \mid x \in \mathbb{T}^d, |v| \leqslant A + \frac{\sqrt{d}}{T-s}\}.$$

Proof. Define the path $q(\cdot)$ by

$$q_j(t) = \Sigma^1_t[s,\mu^q](q_j), \quad 0 \leqslant t \leqslant T,\ j = 1,\dots,n.$$

Evidently, $q(s) = q$. By Proposition 4.2.1,

$$\mathcal{U}^n(s,q) = \mathcal{U}(s,\mu^q) = \int_0^s L_n(q(t),\dot q(t))dt + G_n(q(0)). \tag{4.2.10}$$

Remembering that Σ is a solution to the hamiltonian ODEs (3.1.3), and the relationship between H and L (see (4.1.3)), we have, following (4.2.3) and (4.2.4):

$$\begin{aligned}\nabla_{q_j} G_n(q(0)) &= \frac{1}{n}\nabla_q g(q_j(0),\mu^{q(0)}) = \frac{1}{n}\nabla_q g(\Sigma^1[s,\mu^q](0,q_j),\Sigma^1[s,\mu^q](0,\cdot)_{\#}\mu^q)\\ &= \frac{1}{n}\Sigma^2[s,\mu^q](0,q_j)\ ;\end{aligned}$$

$$\begin{aligned}\nabla_{q_j} L_n(q(t),\dot q(t)) &= \frac{1}{n}(\nabla_q L(q_j(t),\dot q_j(t)) - \nabla_q F(q_j(t),\mu^q))\\ &= \frac{1}{n}\partial_t\Sigma^2[s,\mu^q](t,q_j)\ .\end{aligned}$$

A first-order Taylor expansion gives, for any $y(\cdot) \in AC^2(0,s;(\mathbb{T}^d)^n)$,

$$\begin{aligned}L_n(y(t),\dot y(t)) - L_n(q(t),\dot q(t)) =& \sum_{j=1}^n \nabla_{q_j} L_n(q(t),\dot q(t))\cdot(y_j(t)-q_j(t))\\ &+ \sum_{j=1}^n \nabla_{v_j} L_n(q(t),\dot q(t))\cdot(\dot y_j(t)-\dot q_j(t))\\ &+ \varepsilon_n(t)(|y(t)-q(t)|^2 + |y(t)-q(t)||\dot y(t)-\dot q(t)| + |\dot y(t)-\dot q(t)|^2),\end{aligned}$$

where

$$|\varepsilon_n(t)| \leqslant \max\{|\nabla^2 L_n(x,v)| \bigm| (x,v) \in [(q(t),\dot q(t)),(y(t),\dot y(t))]\}$$

and $[(q(t),\dot q(t)),(y(t),\dot y(t))]$ is the straight segment joining the points $(q(t),\dot q(t))$ and $(y(t),\dot y(t))$. Using (4.2.5), (4.2.6), (4.2.7), then,

$$\begin{aligned}L_n(y(t),\dot y(t)) - L_n(q(t),\dot q(t)) =& \sum_{j=1}^n \frac{1}{n}\partial_t\Sigma^2[s,\mu^q](t,q_j)\cdot(y_j(t)-q_j(t))\\ &+ \frac{1}{n}\varepsilon^1(t)|y(t)-q(t)|^2 + \frac{1}{n}\varepsilon^2(t)|y(t)-q(t)||\dot y(t)-\dot q(t)|\\ &+ \frac{1}{n}\varepsilon^3(t)|\dot y(t)-\dot q(t)|^2,\end{aligned}$$

where

$$|\varepsilon^1(t)| \leqslant \max\{|\nabla^2_{qq}L(x,v)| \bigm| (x,v) \in [(q(t),\dot q(t)),(y(t),\dot y(t))]\} + \max\{|\nabla^2 F_{qq}(x,\mu^x)| \bigm| x \in [q(t),(y(t)]\},$$

$$\begin{aligned}|\varepsilon^2(t)| \leqslant\ &\max\{|\nabla^2_{qv}L(x,v)| \bigm| (x,v) \in [(q(t),\dot q(t)),(y(t),\dot y(t))]\},\\ |\varepsilon^3(t)| \leqslant\ &\max\{|\nabla^2_{vv}L(x,v)| \bigm| (x,v) \in [(q(t),\dot q(t)),(y(t),\dot y(t))]\}.\end{aligned}$$

Let

$$y(t) := q(t) + q^* - q,$$

so $y(\cdot)$ is a translation of $q(\cdot)$. Then $\dot y(t) = \dot q(t)$ and $y(t) - q(t) = q^* - q$ for all t. Since $y(s) = q^*$, we know that $\mathcal{U}^n(s,q^*) \leqslant \int_0^s L_n(y(t),\dot y(t))dt + G_n(y(0))$. By (4.2.10), we have

$$\mathcal{U}^n(s,q^*) \leqslant \mathcal{U}^n(s,q) + \int_0^s [L_n(y(t),\dot y(t)) - L_n(q(t),\dot q(t))]dt + G_n(y(0)) - G_n(q(0)).$$

Using the Taylor expansion above for this particular $y(\cdot)$, and remembering again that $F(\cdot,\mu) = \phi * \mu(\cdot)$, $g(\cdot,\mu) = U^0(\cdot) + U^1 * \mu(\cdot)$, with the C^3-norms of ϕ, U^0, U^1 bounded by the constant κ, we estimate that

$$\begin{aligned}\mathcal{U}^n(s,q^*) \leqslant\ &\mathcal{U}^n(s,q) + \int_0^s [\sum_{j=1}^n \nabla_{q_j} L_n(q(t),\dot q(t)) \cdot (q_j^* - q_j) + \frac{1}{n}(\ell+\kappa)|q^*-q|^2]\ dt\\ &+ \sum_{j=1}^n \nabla_{q_j} g(q(0),\mu^{q(0)}) \cdot (q_j^* - q_j) + \frac{1}{n}\kappa|q^*-q|^2\\ \leqslant\ &\mathcal{U}^n(s,q) + \int_0^s [\frac{1}{n}\sum_{j=1}^n \partial_t \Sigma^2[s,\mu^q](t,q_j) \cdot (q_j^* - q_j) + \frac{1}{n}(\ell+\kappa)|q^*-q|^2]\ dt\\ &+ \frac{1}{n}\sum_{j=1}^n \nabla_{q_j} g(q(0),\mu^{q(0)}) \cdot (q_j^* - q_j) + \frac{1}{n}\kappa|q^*-q|^2\\ =\ &\mathcal{U}^n(s,q) + \frac{1}{n}\sum_{j=1}^n \Sigma^2[s,\mu^q](s,q_j) \cdot (q_j^* - q_j) + \frac{s}{n}(\kappa+\ell)|q^*-q|^2 + \frac{1}{n}\kappa|q^*-q|^2.\end{aligned}$$

Since the inequality that we seek to prove is independent of the order of the coordinates of q and q^*, we can assume that the points $q_1^*, \dots, q_n^*$ are already ordered in such a way that

$$\frac{1}{n}\sum_{j=1}^n |q_j - q_j^*|^2 = \mathscr{W}^2(\mu^q,\mu^{q^*}),$$

so that

$$\frac{1}{n}\sum_{j=1}^n \delta_{(q_i,q_i^*)} \in \Gamma_0(\mu^q,\mu^{q^*}).$$

Then the latter inequality becomes (4.2.8). For the second inequality, we begin with

$$\begin{aligned}
\mathcal{U}^n(T, q(T)) &= \int_0^T L_n(q(t), \dot{q}(t))dt + G_n(q(0)) \\
&= \int_0^s L_n(q(t), \dot{q}(t))dt + G_n(q(0)) + \int_s^T L_n(q(t), \dot{q}(t))dt + \mathcal{U}^n(s, q) - \mathcal{U}^n(s, q) \\
&= \int_s^T L_n(q(t), \dot{q}(t))dt + \mathcal{U}^n(s, q),
\end{aligned}$$

where Proposition 4.2.1 has been applied. This time we use the auxiliary path

$$y(t) = q(t) + \frac{T-t}{T-s}(q^* - q),$$

which satifisfies $y(s) = q^*$ and $y(T) = q(T)$, so, using the dynamic programming principle, (4.2.2), we know that

$$\begin{aligned}
\mathcal{U}^n(T, q(T)) &= \inf \left\{ \int_s^T L_n(w(t), \dot{w}(t))dt + \mathcal{U}^n(s, w(s)) \;\middle|\; w(T) = q(T) \right\} \\
&\leqslant \int_s^T L_n(y(t), \dot{y}(t))dt + \mathcal{U}^n(s, y(s)).
\end{aligned}$$

Comparing this with the equality for $\mathcal{U}^n(T, q(T))$ above, we have

$$\mathcal{U}^n(s, q^*) \geqslant \mathcal{U}^n(s, q) + \int_s^T [L_n(q(t), \dot{q}(t)) - L_n(y(t), \dot{y}(t)]dt.$$

Note that

$$\dot{y}_j(t) = \dot{q}_j(t) - \frac{q_j^* - q_j}{T-s}.$$

Therefore, using Taylor's formula and the bounds,

$$\begin{aligned}
\mathcal{U}^n(s, q^*) \geqslant \mathcal{U}^n(s, q) &+ \int_s^T \sum_{j=1}^n \frac{1}{n} \partial_t \Sigma^2[s, \mu^q](t, q_j) \cdot (y_j(t) - q_j(t)) + \sum_{j=1}^n \frac{1}{n} \Sigma^2[s, \mu^q](t, q_j) \cdot (\dot{y}_j(t) - \dot{q}_j(t)) \; dt \\
&- \int_s^T \frac{1}{n} \sum_{j=1}^n \left(\frac{T-t}{T-s} \right)^2 (\ell + \kappa) |q_j^* - q_j|^2 \; dt - \int_s^T \frac{1}{n} \sum_{j=1}^n \frac{T-t}{(T-s)^2} \ell |q_j^* - q_j|^2 \; dt \\
&- \int_s^T \frac{1}{n} \sum_{j=1}^n \frac{\ell}{(T-s)^2} |q_j^* - q_j|^2 \; dt
\end{aligned}$$

$$= \mathcal{U}^n(s,q) - \frac{1}{n}\sum_{j=1}^n \Sigma^2[s,\mu^q](s,q_j)\cdot(y_j(s)-q_j(s))$$
$$- \frac{1}{3}(T-s)(\ell+\kappa)\frac{1}{n}\sum_{j=1}^n |q_j^* - q_j|^2 - \frac{1}{2}\ell\frac{1}{n}\sum_{j=1}^n |q_j^* - q_j|^2 - \frac{\ell}{T-s}\frac{1}{n}\sum_{j=1}^n |q_j^* - q_j|^2$$
$$= \mathcal{U}^n(s,q) - \frac{1}{n}\sum_{j=1}^n \Sigma^2[s,\mu^q](s,q_j)\cdot(y_j(s)-q_j(s))$$
$$- \big[\frac{1}{3}(T-s)(\ell+\kappa) + \frac{1}{2}\ell + \frac{\ell}{T-s}\big]\frac{1}{n}\sum_{j=1}^n |q_j^* - q_j|^2,$$

and arguing in the same way as for the first inequality, we obtain (4.2.9). □

Now we can answer the question about the derivative in measure of the value function $\mathcal{U}$. Recall that

$$\begin{aligned}\mathcal{V}[s,\mu](t,q) &:= \Sigma_t^2[s,\mu]\circ X_t[s,\mu](q) \\ &= \Sigma_t^2[s,\mu]\circ(\Sigma_t^1[s,\mu])^{-1}(q), \quad s,t\in[0,T], \mu\in\mathscr{P}(\mathbb{T}^d), q\in\mathbb{T}^d.\end{aligned} \tag{4.1.20}$$

Lemma 4.2.3. *Let μ, $\mu^*\in\mathscr{P}(\mathbb{T}^d)$, $\gamma\in\Gamma_0(\mu,\mu^*)$. If $0\leqslant s\leqslant T$, then*

$$\begin{aligned}\mathcal{U}(s,\mu^*) \leqslant \mathcal{U}(s,\mu) &+ \int_{\mathbb{T}^d\times\mathbb{T}^d}\mathcal{V}[s,\mu](s,q)\cdot(b-q)\gamma(dq,db) \\ &+ (\kappa(1+s)+\ell)\mathscr{W}^2(\mu,\mu^*)\ ;\end{aligned} \tag{4.2.11}$$

$$\begin{aligned}\mathcal{U}(s,\mu^*) \geqslant \mathcal{U}(s,\mu) &+ \int_{\mathbb{T}^d\times\mathbb{T}^d}\mathcal{V}[s,\mu](s,q)\cdot(b-q)\gamma(dq,db) \\ &- ((T-s)(\kappa+\ell)/3 + \ell/2 + \ell/(T-s))\mathscr{W}^2(\mu,\mu^*)\ ;\end{aligned} \tag{4.2.12}$$

and for any $s\in(0,T)$, $t\in[0,T)$, there holds

$$\nabla_\mu\mathcal{U}(t,\sigma_t) = \mathcal{V}_t[s,\mu]. \tag{4.2.13}$$

Proof. It is easily verified that the function $t\mapsto\mathcal{L}(\sigma(t),v(t))$ is continuous. Recall that $\sigma(t) = \Sigma_t^1[s,\mu]_\#\mu$, $v(t) = \partial_t\Sigma_t^1[s,\mu]\circ X_t[s,\mu]$. Since $\mathcal{U}(t,\mu) = \mathcal{A}(s;\sigma,v)$, it follows that $\mathcal{U}(\cdot,\mu)$ is continuous, for any $\mu\in\mathscr{P}(\mathbb{T}^d)$. Thus, it will be enough to prove (4.2.11), (4.2.12), for $s\in(0,T)$. In fact, since

$$\begin{aligned}\mathcal{A}(s;\sigma,v) = &\int_0^s\int_{\mathbb{T}^d}[L(\Sigma_t^1[s,\mu](q),\partial_t\Sigma_t^2[s,\mu](q))\mu(dq) - \frac{1}{2}\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\phi(q-\Sigma_t^1[s,\mu](y))\mu(dy)dq\ ]\ dt \\ &+ \int_{\mathbb{T}^d}[U_0(\Sigma_t^1[s,\mu](q)) + \frac{1}{2}\int_{\mathbb{T}^d}U^1(q-\Sigma_t^1[s,\mu](y))\mu(dy)]\mu(dq),\end{aligned}$$

the convergence results of Chapter 3 give the continuity of $\mathcal{U}$ with respect to μ as well. The fact that any measure $\mu\in\mathscr{P}(\mathbb{T}^d)$ can be approximated by averages of point masses, implies, then, because of

Lemma 4.2.2, that for given $\mu, \mu^* \in \mathscr{P}(\mathbb{T}^d)$, there is some $\gamma \in \Gamma_0(\mu, \mu^*)$ such that

$$\begin{aligned}\mathcal{U}(s,\mu^*) &\leqslant \mathcal{U}(s,\mu) + \int_{\mathbb{T}^d\times\mathbb{T}^d} \Sigma^2[s,\mu](s,q)\cdot(b-q)\gamma(dq,db)\\ &\quad + (\kappa(1+s)+\ell)\mathscr{W}^2(\mu,\mu^*)\ ; \end{aligned}\tag{4.2.14}$$

$$\begin{aligned}\mathcal{U}(s,\mu^*) &\geqslant \mathcal{U}(s,\mu) + \int_{\mathbb{T}^d\times\mathbb{T}^d} \Sigma^2[s,\mu](s,q)\cdot(b-q)\gamma(dq,db)\\ &\quad - ((T-s)(\kappa+\ell)/3+\ell/2+\ell/(T-s))\mathscr{W}^2(\mu,\mu^*)\ ; \end{aligned}\tag{4.2.15}$$

hold. Let now $\mu, \mu^* \in \mathscr{P}(\mathbb{T}^d)$ be fixed, as in the statement, and let $\gamma \in \Gamma_0(\mu, \mu^*)$ such that (4.2.14) holds. For each $\lambda \in (0,1)$, define $\mu_\lambda \in \mathscr{P}(\mathbb{T}^d)$ through

$$\int_{\mathbb{T}^d} \phi(q)\mu_\lambda(dq) = \int_{\mathbb{T}^d\times\mathbb{T}^d} \phi((1-\lambda)q+\lambda b)\gamma(dq,db), \quad \phi \in C_b(\mathbb{T}^d).$$

According to ([AGS08, Lemmas 7.2.2, 7.2.1]), for each $\lambda \in (0,1)$, there is a unique optimal transport plan γ_λ between $\mu_0 = \mu$ and μ_λ, so $\Gamma_0(\mu,\mu_\lambda) = \{\gamma_\lambda\}$. Thus, (4.2.14) holds for $\mu = \mu$ and $\mu^* = \mu_\lambda$, $0 < \lambda < 1$, for in this case existence in the previous lemma covers all (only one) optimal plans. Letting $\lambda \to 1$, and using the continuity of $\mathcal{U}$, (4.2.14) is established, and (4.2.15) is treated analogously. If we now refer back to our adopted definition of Wasserstein gradient, we see that (4.2.14) and (4.2.15) mean that

$$\nabla_\mu \mathcal{U}(s,\mu) = \Sigma_s^2[s,\mu].$$

Choosing $\mu = \sigma_t$ and $s = t$, the latter reads $\nabla_\mu \mathcal{U}(t,\sigma_t) = \Sigma_t^2[t,\sigma_t]$. Finally, using (4.1.24), we obtain (4.2.13), indeed:

$$\Sigma_t^2[t,\sigma_t]\circ\Sigma_t^1[s,\mu] = \Sigma_t^2[s,\mu] \implies \Sigma_t^2[t,\sigma_t] = \Sigma_t^2[s,\mu]\circ(\Sigma_t^1[s,\mu])^{-1} = \mathcal{V}_t[s,\mu].$$

□

4.3 The HJB equation for $\mathcal{U}$

This short section justifies the title of the chapter. Recall the definition of $\mathcal{L}$ at the beginning of Section 4.1. Define $\mathcal{H}$ as

$$\mathcal{H}(\mu,\zeta) := \sup_{\xi\in L^{r'}(\mu)} \Big\{ \int_{\mathbb{T}^d} \xi(q)\cdot\zeta(q)\mu(dq) - \mathcal{L}(\mu,\xi)\Big\},$$

for $\mu \in \mathscr{P}(\mathbb{T}^d)$, $\zeta \in L^r(\mu)$.

Proposition 4.3.1. *Let $\mu \in \mathscr{P}(\mathbb{T}^d)$, $\zeta \in L^r(\mu)$. Then*

$$\mathcal{H}(\mu,\zeta) = \int_{\mathbb{T}^d} H(q,\zeta(q))\mu(dq) + \mathcal{F}(\mu).\tag{4.3.1}$$

Proof. Since $H(q,\zeta(q)) = \sup_{v\in\mathbb{R}^d} v\cdot\zeta(q) - L(q,v)$, for every $q \in \mathbb{T}^d$, it is apparent that in (4.3.1), the inequality $\leqslant$ holds. To prove the equality, fix $\zeta \in L^r(\mu)$. For any fixed $q \in \mathbb{T}^d$,

$$H(q,\zeta(q)) = \sup_{v\in\mathbb{R}^d} v\cdot\zeta(q) - L(q,v) = \xi(q)\cdot\zeta(q) - L(q,\xi(q)),$$

where $\xi(q)$ is defined by $\zeta(q) = \nabla_v L(q,\xi(q))$, that is,

$$\xi(q) := [\nabla_v L(q,\cdot)]^{-1}(\zeta(q)). \tag{4.3.2}$$

(The inverse in (4.3.2) exists because, if $\nabla_v L(q,v_1) = \nabla_v L(q,v_2)$, let $f(\theta) := L(q,(1-\theta)v_1+\theta v_2)$. Then $f'(\theta) = \nabla_v L(q,(1-\theta)v_1+\theta v_2)\cdot(v_2-v_1)$, and $f'(0) = f'(1)$. The condition $\nabla^2_{pp}H > 0$ implies $\nabla^2_{vv}L > 0$, so $f'' > 0$. Therefore one must have $v_1 = v_2$.) Recall that p and v are related to one another by the formula $H(q,p)+L(q,v) = p\cdot v$, and this gives $v = \nabla_p H(q,p)$, $p = p(q,v) = \nabla_v L(q,v)$, so $v = \nabla_p H(q,\nabla_v L(q,v)), q \in \mathbb{T}^d, v \in \mathbb{R}^d$. Thus, for any fixed q, the mapping $\nabla_p H(q,\nabla_v L(q,\cdot))$ is the identity. Hence (4.3.2) is followed by

$$\xi(q) = \nabla_p H(q,\zeta(q)).$$

But $\nabla_p H$ satisfies (4.0.1), so

$$|\nabla_p H(q,\zeta(q))|^{r'} \leqslant [c_2(|\zeta(q)|^{r-1}+1)]^{r'} \leqslant 2^{r'-1}|\zeta(q)|^{r'(r-1)}+2^{r'-1}c_2 = 2^{r'-1}|\zeta(q)|^{r}+2^{r'-1}c_2, \quad q\in\mathbb{T}^d.$$

This means that ξ is actually in $L^{r'}(\mu)$, and this completes the proof. □

Now suppose that $r = 2$. In a recent work [GT18] by Gangbo and Tudorascu, it is proven that the value function $\mathcal{U}$, defined in (4.2.1), is the unique solution *in the viscosity sense* to the Hamilton-Jacobi equation

$$\begin{cases} \partial_s\mathcal{U}(s,\mu) + \mathcal{H}(\mu,\nabla_\mu\mathcal{U}(s,\mu)) = 0, \\ \mathcal{U}(0,\mu) = \mathcal{G}(\mu). \end{cases}$$

Chapter 5

THE FIRST ORDER MEAN FIELD GAME SYSTEM

In this short but important chapter, we construct a short-time classical solution to the system (1.3.2-1.3.5), based on the solutions Σ to the Hamiltonian equations (3.1.3).

In what follows, we assume $\theta > 2\kappa$, as in Corollary 3.1.7(ii). By Σ, as always, we denote the solution to (3.1.3), obtained as the fixed point of the operator $\mathfrak{m}^{s,\mu}$ of Definition 3.1.6. 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 5.0.1. *For every $s \in [0,T]$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, the function $X[s,\mu](\cdot,\cdot)$ is in the Sobolev space $W^{2,2;\infty}((0,T) \times \mathbb{T}^d; \mathbb{T}^d)$.*

Let us recall the definition of $\mathcal{V}$:

$$\mathcal{V}[s,\mu](t,q) := \Sigma_t^2[s,\mu] \circ X_t[s,\mu](q) = \Sigma_t^2[s,\mu] \circ (\Sigma_t^1[s,\mu])^{-1}(q), \tag{4.1.20}$$

for $s,t \in [0,T], \mu \in \mathscr{P}(\mathbb{T}^d), q \in \mathbb{T}^d$. Alternatively, we may write $\mathcal{V}_t[s,\mu](q)$.

At this point it is worth saying a couple of words about the method of proof for the MFG system. With (σ, v) defined as in (3.2.6), (σ, v) satisfies the continuity equation. Suppose, for a moment, that we have solved the system (1.3.2-1.3.5) and we wish to calculate the total derivative with respect to t of U *along the trajectory* of a particle given by Σ^1. We have

$$(\partial_t)(U(t,\Sigma_t^1(q))) = \partial_t U(t,\Sigma_t^1(q)) + \nabla U(t,\Sigma_t^1(q)) \cdot \partial_t \Sigma_t^1(q), \tag{5.0.1}$$

and by (1.3.2),

$$(\partial_t)(U(t,\Sigma_t^1(q))) = -H(\Sigma_t^1(q), \nabla_q U(t,\Sigma_t^1(q))) - F(\Sigma_t^1(q),\sigma_t) + \nabla U(t,\Sigma_t^1(q)) \cdot \partial_t \Sigma_t^1(q).$$

Changing from $\Sigma_t^1(q)$ to q, and recalling that $\partial_t \Sigma_t^1 = \nabla_p H(\Sigma_t^1, \Sigma_t^2)$,

$$(\partial_t)(U(t,q)) = -H(q, \nabla_q U(t,q)) - F(q,\sigma_t) + \nabla U(t,q) \cdot \nabla_p H(q, \mathcal{V}_t(q)). \tag{5.0.2}$$

If we revisit the old, classical method of characteristics to solve first-order PDEs and its application to the simplest HJB equation (see, e.g., [Eva10]), we are strongly persuaded to believe that, in our case as well, the "momentum" variable (**p** in [Eva10]) in the system of characteristics should coincide with the gradient of the solution to the PDE, that is:

$$\Sigma_t^2(X_t(q)) = \nabla_q U(t,q), \tag{5.0.3}$$

i.e. formula (5.0.8). If we substitute this into (5.0.2) we have:

$$(\partial_t)(U(t,q)) = -H(q,\mathcal{V}_t(q))) - F(q,\sigma_t) + \mathcal{V}_t(q)\cdot\nabla_p H(q,\mathcal{V}_t(q)). \tag{5.0.4}$$

Now, note that (5.0.4) is not written in terms of U. If we take (5.0.4) as our starting point, then, combining with (5.0.1) and (5.0.3) again, we obtain none other than the HJB (1.3.2). Equation (5.0.4) is straightforward to solve.

We see then that, to solve (1.3.2), we should aim to obtain (5.0.3). That step will also follow the idea for the analogous fact in the general method of characteristics. We hope that the foregoing discussion helps clarify the proof below. Let us also mention that (5.0.8) shows that $\nabla_q\mathcal{V}_t(q)$ is symmetric, a fact that is an essential ingredient to prove formula (6.0.6) at the beginning of Chapter 6.

Lemma 5.0.2. *Let T be small according to Remark 3.1.8, $s\in[0,T]$, $\mu\in\mathscr{P}(\mathbb{T}^d)$, $\sigma_t=\Sigma^1[s,\mu](t,\cdot)_{\#}\mu$, and for each $q\in\mathbb{T}^d$, let*

$$U(t,q) = z(t, X[s,\mu](t,q)),\quad t\in[0,T],$$

where $z(\cdot,q)$ satisfies

$$\begin{aligned}\partial_t z(t,q) &= \Sigma^2[s,\mu](t,q)\cdot\nabla_p H(\Sigma^1[s,\mu](t,q),\Sigma^2[s,\mu](t,q)) \\ &\qquad - H(\Sigma^1[s,\mu](t,q),\Sigma^2[s,\mu](t,q)) - F(\Sigma^1[s,\mu](t,q),\sigma_t) \quad \text{in } (0,T), \\ z(0,q) &= g(\Sigma^1[s,\mu](0,q),\sigma_0).\end{aligned}\tag{5.0.5}$$

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

$$\partial_t U(t,x) + H(x,\nabla_q U(t,x)) + F(x,\sigma_t) = 0 \quad \text{in } (0,T)\times\mathbb{T}^d, \tag{1.3.2}$$

$$U(0,\cdot) = g(\cdot,\sigma_0). \tag{1.3.4}$$

Proof. Since s and μ are fixed, we abbreviate $\Sigma[s,\mu](t,q)=\Sigma_t(q)$. Observe that the right-hand side of (5.0.5) 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 5.0.1. Moreover, since $\partial_t z(t,q)$ is C^1 in q, then (e.g. see [Rud76, Thm. 9.41]) $\nabla^2_{tq}z$ exists and is equal to $\nabla^2_{qt}z$. Thus, the calculations below are legitimate. We have $U(t,\Sigma^1_t(q)) = z(t,q)$, so $\partial_t z(t,q) = \partial_t(U(t,\Sigma^1_t(q))$ and

$$\begin{aligned}\partial_t z(t,q) &= \partial_t(U(t,\Sigma^1_t(q))) = \partial_t U(t,\Sigma^1_t(q)) + \nabla_q U(t,\Sigma^1_t(q))\cdot\partial_t\Sigma^1_t(q) \\ &= \partial_t U(t,\Sigma^1_t(q)) + \nabla_q U(t,\Sigma^1_t(q))\cdot\nabla_p H(\Sigma^1_t(q),\Sigma^2_t(q))\end{aligned}\tag{5.0.6}$$

Now, *if*

$$\nabla_q U(t,\Sigma^1_t(q)) = \Sigma^2_t(q),\quad t\in(0,T), q\in\mathbb{T}^d, \tag{5.0.7}$$

then, comparing (5.0.6) and (5.0.5), 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 (1.3.2), (1.3.4). We set out to prove (5.0.7) now. Let

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

$i = 1, \ldots, d$, where $x = \Sigma_t^1(q)$. We know that $r_i(0) = 0$, from the initial conditions, and

$$\dot{r}_i(q) = \frac{\partial^2 z(t,q)}{\partial t \partial q^{(i)}} - \sum_{j=1}^{d} \partial_t (\Sigma_t^2)^{(j)}(q) \frac{\partial}{\partial q^{(i)}} \frac{\partial}{\partial t} x^{(j)} =: a - b.$$

Using the first line, $\partial_t \Sigma_t^1(q) = \nabla_p H(\Sigma_t^1(q), \Sigma_t^2(q))$, in (3.1.3), we have

$$\begin{aligned}
a &= \partial_{q^{(i)}} \partial_t z(t,q) \\
&= \sum_{k=1}^{d} \Big[\sum_{k=1}^{d} \partial_{q^{(i)}} (\Sigma_t^2)^{(k)}(q) \partial_t (\Sigma^1)_t^{(k)}(q) + (\Sigma_t^2)^{(k)}(q) \partial^2_{t,q^{(i)}} (\Sigma_t^1)^{(k)}(q) \Big] \\
&\quad - \sum_{l=1}^{d} \big[\partial_{q^{(l)}} H(\Sigma_t^1(q), \Sigma_t^2(q)) \partial_{q^{(i)}} (\Sigma_t^1)^{(l)}(q) + \partial_{p^{(l)}} H(\Sigma_t^1(q), \Sigma_t^2(q)) \partial_{q^{(i)}} (\Sigma_t^2)^{(l)}(q) \big] \\
&\quad - \sum_{l=1}^{d} \partial_{q^{(l)}} F(\Sigma_t^1(q), \sigma_t) \partial_{q^{(i)}} (\Sigma_t^1)^{(l)}(q),
\end{aligned}$$

and, by the second line in (3.1.3), namely, $\partial_t \Sigma_t^2(q) = -\nabla_q H(\Sigma_t^1(q), \Sigma_t^2(q)) - \nabla_q F(\Sigma_t^1(q), \sigma_t)$, this simplifies to

$$a = \sum_{k=1}^{d} \big[(\Sigma_t^2)^{(k)}(q) \partial^2_{tq^{(i)}} (\Sigma_t^1)^{(k)}(q) + \partial_t (\Sigma_t^2)^{(k)}(q) \partial_{q^{(i)}} (\Sigma_t^1)^{(k)}(q) \big].$$

As for b,

$$\begin{aligned}
b &= \partial_t \Big(\sum_{j=1}^{d} (\Sigma_t^2)^{(j)}(q) \partial_{q^{(i)}} (\Sigma_t^1)^{(j)}(q) \Big) \\
&= \sum_{j=1}^{d} \big[\partial_t (\Sigma_t^2)^{(j)}(q) \partial_{q^{(i)}} (\Sigma_t^1)^{(j)}(q) + (\Sigma_t^2)^{(j)}(q) \partial^2_{tq^{(i)}} (\Sigma_t^1)^{(j)}(q) \big],
\end{aligned}$$

so $a = b$. Therefore $\dot{r}_i(t) \equiv 0$, and $r_i(t) \equiv 0$ on $(0, T]$, by the uniqueness of (5.0.5) $[0, T]$. Now we differentiate U, keeping in mind that $U(t,x) = z(t,q)$; using the fact that $r_i(t) = 0$, $0 \leqslant t \leqslant T$, we have

$$\partial_{x^{(i)}} U(t,x) = \sum_{j=1}^{d} \frac{\partial z(t,q)}{\partial q^{(j)}} \frac{\partial q^{(j)}}{\partial x^{(i)}} = \sum_{j=1}^{d} \sum_{k=1}^{d} (\Sigma_t^2)^{(k)}(q) \frac{\partial x^{(k)}}{\partial q^{(i)}} \frac{\partial q^{(j)}}{\partial x^{(i)}} = \sum_{k=1}^{d} (\Sigma_t^2)^{(k)}(q) \frac{\partial x^{(k)}}{\partial x^{(i)}} = (\Sigma_t^2)^{(i)}(q),$$

for $i = 1, \ldots, d$ and $0 \leqslant t \leqslant T$. This proves (5.0.7), completing the proof of the lemma. □

Corollary 5.0.3. *By (5.0.7) in the proof of the preceding lemma, we have*

$$\nabla_q U(t,q) = \mathcal{V}[s,\mu](t,q), \qquad t \in (0,T), q \in \mathbb{T}^d. \tag{5.0.8}$$

The dependence of U on the parameters s and μ, 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

$$v_t(q) = \partial_t \Sigma_t^1((\Sigma_t^1)^{-1}(q)) = \nabla_p H(q, \mathcal{V}(t,q)) = \nabla_p H(q, \nabla_q U(t,q)). \tag{5.0.9}$$

Hence, combining with Proposition 3.2.10, we have the following.

Theorem 5.0.4. *(Existence of solution to the MFG system) Let $\mu \in \mathscr{P}(\mathbb{T}^d)$, T be in accordance with Remark 3.1.8 and Proposition 3.1.5, $0 < s < T$, and let σ_t, v_t be as in (3.2.6), where $(\Sigma^1[s,\mu], \Sigma^2[s,\mu])$ is the unique solution to (3.1.3) with parameters s, μ. Then the pair (U,σ), where U is as in Lemma 5.0.2, is a classical solution to the mean-field game system (1.3.2-1.3.5) in the sense explained in Section 2.3.*

Note that, by Proposition 5.0.1, the function U in the pair (U,σ) constructed above is in $W^{2,2;\infty}((0,T) \times \mathbb{T}^d) \times AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$. The following is, in a sense, a consistency, or restricted uniqueness, complement to the latter theorem.

Theorem 5.0.5. *(The case of $W^{2,3;\infty}((0,T)\times\mathbb{T}^d)\times AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ solutions to the MFG system) Let $(\tilde{U},\tilde{\sigma})$ be a classical solution to the MFG system (1.3.2-1.3.5), in the sense explained in Section 2.3, such that $U \in W^{2,3;\infty}((0,T) \times \mathbb{T}^d)$. Then $(\tilde{U},\tilde{\sigma}) = (U,\sigma)$, where (U,σ) is the pair constructed for Theorem 5.0.4.*

Proof. Let $(\tilde{U},\tilde{\sigma})$ be a solution to the system (1.3.2-1.3.5) with parameters $s \in (0,T)$ and $\mu \in \mathscr{P}(\mathbb{T}^d)$, according to the definition of Section 2.3, 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 (3.1.3). Set

$$\tilde{v}_t(q) = \tilde{v}(t,q) := \nabla_p H(q, \nabla_q \tilde{U}(t,q)),$$

$0 \leqslant t \leqslant T$, $q \in \mathbb{T}^d$. Since $\tilde{U} \in W^{2,3;\infty}((0,T)\times\mathbb{T}^d)$ and $H \in C^3$, we have that $\mathrm{Lip}(\tilde{v}_t, K) + \sup_{q\in K} |\tilde{v}_t(q)|$ is bounded on $[0,T]$, where K is any compact subset of $\mathbb{R}^d$ and $\mathrm{Lip}(\tilde{v}_t, K)$ is the Lipschitz constant of $\tilde{v}_t|_K$. By elementary ODE theory (see, e.g., [AGS08, Lemma 8.1.4]), if $q \in \mathbb{R}^d$, the ODE

$$\tilde{\Sigma}^1(s,q) = q, \quad \frac{\partial}{\partial t}\tilde{\Sigma}^1(t,q) = \tilde{v}_t(\tilde{\Sigma}^1(t,q)) \tag{5.0.10}$$

has a unique maximal solution in a neighborhood $I(q,s) \subset (0,T)$ of s, but, since $\tilde{\Sigma}^1(t,\cdot)$ is periodic, and therefore bounded for every $t \in (0,T)$, then $I(q,s) = (0,T)$. Clearly, the path $t \mapsto \tilde{\Sigma}^1(t,\cdot)_{\#}\mu$ solves the continuity equation with velocity $\tilde{v}_t$. We can apply Proposition 8.1.7 of [AGS08] to conclude that

$$\tilde{\sigma}_t = \tilde{\Sigma}^1(t,\cdot)_{\#}\mu, \tag{5.0.11}$$

$0 \leqslant t \leqslant T$. Let

$$\tilde{\mathcal{V}}(t,q) := \nabla_q \tilde{U}(t,q) \ , \qquad \tilde{\Sigma}^2(t,q) := \tilde{\mathcal{V}}(t, \tilde{\Sigma}^1(t,q)), \tag{5.0.12}$$

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

$$\partial_t \tilde{\Sigma}^1(t,q) = \tilde{v}_t(\tilde{\Sigma}^1(t,q)) = \nabla_p H(\tilde{\Sigma}^1(t,q), \tilde{\Sigma}^2(t,q)), \tag{5.0.13}$$

which is the first equation in (3.1.3). To obtain the second one, observe that

$$\partial_t \tilde{\Sigma}^2(t,q) = \partial_t \tilde{\mathcal{V}}(t, \tilde{\Sigma}^1(q,t)) + \nabla_q \tilde{\mathcal{V}}(t, \tilde{\Sigma}^1(t,q)) \partial_t \tilde{\Sigma}^1(t,q) = \nabla^2_{tq} \tilde{U}(t, \tilde{\Sigma}^1(t,q)) + \nabla^2_{qq} \tilde{U}(t, \tilde{\Sigma}^1(t,q)) \partial_t \tilde{\Sigma}^1(t,q),$$

while, differentiating the Hamilton-Jacobi equation with respect to q gives

$$\nabla^2_{qt} \tilde{U}(t,q) + \nabla_q H(q, \nabla_q \tilde{U}(t,q)) + \nabla_p H(q, \nabla_q \tilde{U}(t,q)) \nabla^2_{qq} \tilde{U}(t,q) + \nabla_q F(q, \tilde{\sigma}_t) = 0,$$

which, evaluating at $\tilde{\Sigma}^1(t,q)$ in place of q, and using (5.0.12), (5.0.13), serves to simplify the former equality to

$$\partial_t \tilde{\Sigma}^2(t,q) = -\nabla_q H(\tilde{\Sigma}^1(t,q), \tilde{\Sigma}^2(t,q)) - \nabla_q F(\tilde{\Sigma}^1(t,q), \tilde{\sigma}_t),$$

which is the second equation in (3.1.3). The condition $\tilde{\Sigma}^2(0,q) = \nabla_q g(\tilde{\Sigma}^1(0,q), \tilde{\Sigma}^1(0,\cdot)_{\#}\mu)$ follows readily from (5.0.12) and (1.3.4).

2. We prove that, for a possibly smaller T, the solutions $(\tilde{\Sigma}^1, \tilde{\Sigma}^2)$ to (3.1.3) from the previous paragraph belongs to $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$, i.e., they satisfy the bounds (3.1.16). If T is small enough, since $|\tilde{\Sigma}^2(0,q)| = |\nabla_q g(\tilde{\Sigma}^1_0(q), \tilde{\sigma}_0)| \leqslant \kappa$, continuity implies $|\tilde{\Sigma}^2(t,q)| \leqslant \kappa + \varepsilon$, $0 \leqslant t \leqslant T$, $q \in \mathbb{T}^d$ for an $\varepsilon > 0$ such that

$$|\tilde{\Sigma}^2(t,q)| \leqslant \theta \frac{\kappa}{\theta} + \varepsilon \leqslant \theta \frac{1}{\theta} \max\{d, \kappa\} + \varepsilon = \theta c/\theta + \varepsilon \leqslant \theta B,$$

because B in the proof of Lemma 3.1.4 was chosen as $B > c/\theta$ (in these lines we are referring back to the proof of Lemma 3.1.4, in particular (3.1.8), (3.1.9) and the paragraph preceding those inequalities). This is the third line in (3.1.16). Since $\|\tilde{\Sigma}^2\|_{\mathcal{M}^2} \leqslant \theta B$, we have $|\partial_t \tilde{\Sigma}^1(t,q)| = |\nabla_p H(\tilde{\Sigma}^1(t,q), \tilde{\Sigma}^2(t,q))| \leqslant \bar{l}(B)$ (see Definition 3.1.1, "Coefficient bounds I") for small enough T and all q, and since A_1 in Lemma 3.1.4 was chosen to be larger than $\bar{l}(B)$, we obtain the bound for $\|\partial_t \tilde{\Sigma}^1\|$ in (3.1.16). The one for $\|\partial_t \tilde{\Sigma}^2\|_{\mathcal{M}^2}$ goes in a similar way, because A_2 in Lemma 3.1.4 was chosen larger than $\bar{l}(B)/\theta$. The bounds for the second-order time derivatives are dealt with in a similar way, keeping in mind the way E_1 and E_2 were chosen in the proof of Lemma 3.1.4.

From (5.0.10), $\tilde{\Sigma}^1(t,q) = q + \int_s^t v_\tau(\tilde{\Sigma}^1(\tau,q))d\tau$, 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 $\sqrt{d}$, making it less than A_1, because of (3.1.8). The bound for $\|\nabla^2_{qq} \tilde{\Sigma}^1\|$, due to $\nabla^2_{qq}(q) \equiv 0$, can actually be made arbitrarily small by choosing T small enough. To address $\nabla_q \tilde{\Sigma}^2$ 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 $\nabla_q \tilde{\Sigma}^2(t,q) = \nabla^2_{qq} \tilde{U}(t, \tilde{\Sigma}^1(t,q)) \nabla_q \tilde{\Sigma}^1(t,q)$, the norm of $\nabla_q \tilde{\Sigma}^2$ is the product of a number slightly larger than κ and one slightly larger than $\sqrt{d}$, for small times T. But the constant E in the proof of Lemma 3.1.4 is larger than $c = \max\{d, \kappa\}$, and $A_2 > \frac{1}{\theta} c E(E+1)$. This ensures that $\|\nabla_q \tilde{\Sigma}^2\| \leqslant \theta A_2$. Next, given that

$$\nabla^2_{qq} \tilde{\Sigma}^2 = \nabla^3_{qqq} \tilde{U} \nabla_q \tilde{\Sigma}^1 \nabla_q \tilde{\Sigma}^1 + \nabla^2_{qq} \tilde{U} \nabla^2_{qq} \tilde{\Sigma}^1$$

(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 θA_2, since the norm of $\nabla^3_{qqq}\tilde{U}\nabla_q\tilde{\Sigma}^1\nabla_q\tilde{\Sigma}^1$ is the product of a number slightly larger than κ and one slightly than d. Finally, $\|\nabla_q\tilde{\Sigma}^1_0\|, \|\nabla^2_{qq}\tilde{\Sigma}^1_0\| \leqslant E$ follow from E having been picked larger than d (and, therefore, than $\sqrt{d}$), and taking T smaller if necessary.

Thus, the mapping $(\tilde{\Sigma}^1, \tilde{\Sigma}^2)$ constructed in (5.0.10) and (5.0.12) from $(\tilde{U}, \tilde{\sigma})$ coincides with the unique solution $(\Sigma^1[s,\mu], \Sigma^2[s,\mu])$ of (3.1.3) in $\mathcal{M}_0(A_1, \theta A_2, \theta B, E, E_1, \theta E_2, T)$ during a possibly shorter interval $[0,T]$. Consequently, by (5.0.11), we further have that $\tilde{\sigma} = \sigma$. Also, now that $\tilde{\mathcal{V}}(t,q) = \mathcal{V}(t,q)$, $0 \leqslant t \leqslant T$, we get

$$\begin{aligned}\tilde{U}(t,q) &= g(x, \tilde{\sigma}_0) - \int_0^t [H(q, \tilde{\mathcal{V}}(\tau, q) - F(x, \tilde{\sigma}_\tau)]d\tau \\ &= g(x, \sigma_0) - \int_0^t [H(x, \mathcal{V}(\tau, q)) - F(q, \sigma_\tau)]d\tau = U(t,q),\end{aligned}$$

for any $t \in [0,T]$. Thus, $(\tilde{U}, \tilde{\sigma}) = (U, \sigma)$ on the possibly smaller interval $[0,T]$.

□

Finally, let us stress the following. If (U, σ) is a solution to the MFG system (1.3.2-1.3.5), then, with the knowledge of formula (5.0.8), i.e.,

$$\nabla_q U(t,q) = \mathcal{V}[s,\mu](t,q),$$

we can integrate the HJB equation of the MFG system, to obtain

$$\begin{aligned}U(t,q) &= g(q, \sigma_0) - \int_0^t [H(q, \nabla U(\tau, q)) + F(q, \sigma_\tau)]d\tau \\ &= g(q, \Sigma^1_\tau[s,\mu]_\# \mu) - \int_0^t [H(q, \mathcal{V}[s,\mu](\tau, q) + F(q, \Sigma^1_\tau[s,\mu]_\# \mu)]d\tau. \qquad (5.0.14)\end{aligned}$$

Chapter 6

THE FIRST ORDER MASTER EQUATION

In this chapter we undertake the study of the dependence of our solution U to (1.3.2) on the parameter μ. As mentioned earlier, the theorem on the master equation does not the depend on the main results of Chapter 4, except for Proposition 4.1.6, which we include here again for convenience:

Let $0 \leqslant s, t_0 \leqslant T$, $\mu \in \mathscr{P}(\mathbb{T}^d)$. *Set*

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

Then,

(i) For every $0 \leqslant t \leqslant T$*:*

$$\Sigma_t[t_0,\sigma_{t_0}] \circ \Sigma^1_{t_0}[s,\mu] = \Sigma_t[s,\mu]. \tag{4.1.21}$$

(ii) For every $0 \leqslant t \leqslant T$*:*

$$\begin{aligned}
\Sigma^1_t[t_0,\Sigma^1_{t_0}[t,\mu]_{\#}\mu] \quad &\textit{and} \quad \Sigma^1_{t_0}[t,\mu] \quad \textit{are inverses of each other ,} && (4.1.22)\\
v_t[s,\mu] \; &= v_t[t_0,\sigma_{t_0}]\ , && (4.1.23)\\
\Sigma^2_t[t_0,\sigma_{t_0}] \circ \Sigma^1_{t_0}[s,\mu] \; &= \Sigma^2_t[s,\mu]\ , && (4.1.24)\\
\partial_s\Sigma^1_t[s,\mu] \; &= -\nabla_q\Sigma^1_t[s,\mu]v_s[t,\sigma_t]. && (4.1.25)
\end{aligned}$$

(iii) If $0 \leqslant \tau, t \leqslant T$, *then*

$$\Sigma^2_\tau[t,\sigma_t] \circ (\Sigma^1_\tau[t,\sigma_t])^{-1} = \Sigma^2_\tau[s,\mu] \circ \Sigma^1_\tau[s,\mu]^{-1}. \tag{4.1.26}$$

The following is the definition of the function that will turn out to be the solution to the master equation. It is given by the representation formula (5.0.14) at $t = s$, including the parameter μ explicitly. That is: given $s \in [0,T]$, $q \in \mathbb{T}^d$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, define

$$u(s,q,\mu) = g(q,\Sigma^1[s,\mu](0,\cdot)_{\#}\mu) - \int_0^s \big[H(q,\mathcal{V}[s,\mu](\tau,q)) + F(q,\Sigma^1[s,\mu](\tau,\cdot)_{\#}\mu\big]d\tau, \tag{6.0.1}$$

and, as before, $\sigma_t = \Sigma^1_t[s,\mu]_{\#}\mu$, $0 \leqslant t \leqslant T$. Note that (4.1.26) reads now as

$$\mathcal{V}_\tau[t,\sigma_t] = \mathcal{V}_\tau[s,\mu].$$

Hence, coupled with the fact that, by (4.1.21), $\Sigma^1_\tau[t,\sigma_t]\circ\Sigma^1_t[s,\mu]=\Sigma^1_\tau[s,\mu]$, $0\leqslant\tau\leqslant T$, we have

$$\begin{aligned} u(t,q,\sigma_t) &= g(q,\Sigma^1_0[t,\sigma_t]_\#\sigma_t)-\int_0^t\big[H(q,\mathcal{V}[t,\sigma_t](\tau,q))+F(q,\Sigma^1[t,\sigma_t](\tau,\cdot)_\#\mu)\big]d\tau\\ &= g(q,\sigma_0)-\int_0^t\big[H(q,\mathcal{V}[s,\mu](\tau,q))+F(q,\Sigma^1_\tau[s,\mu]_\#\mu)\big]d\tau. \end{aligned} \tag{6.0.2}$$

Since

$$\mathcal{V}[s,\mu](t,\Sigma^1[s,\mu](t,q))=\Sigma^2[s,\mu](t,q),$$

and Σ^2 satisfies the second of the Hamiltonian ODEs (3.1.3), it follows, by taking the total time derivative of $\mathcal{V}[s,\mu](t,\Sigma^1[s,\mu](t,q))$, 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

$$\partial_t\mathcal{V}(t,q)+\nabla_q\mathcal{V}(t,q)\nabla_pH(q,\mathcal{V}(t,q)) = -\nabla_qH(q,\mathcal{V}(t,q))-\nabla_qF(q,\Sigma^1_{t\,\#}\mu), \tag{6.0.3}$$

$$\mathcal{V}(0,q)=\nabla_qg(q,\sigma_0). \tag{6.0.4}$$

If we differentiate $u(t,q,\sigma_t)$ with respect to q in (6.0.2), and use (6.0.3) and (6.0.4), we get

$$\begin{aligned} \nabla_qu(t,q,\sigma_t)=\nabla_qg(q,\sigma_0)+\int_0^t&\big[\partial_t\mathcal{V}[s,\mu](\tau,q)+\nabla_q\mathcal{V}[s,\mu](\tau,q)\nabla_pH(q,\mathcal{V}[s,\mu](\tau,q))\\ &-\nabla_pH(q,\mathcal{V}[s,\mu](\tau,q)\nabla_q\mathcal{V}[s,\mu](\tau,q)\big]\,d\tau. \end{aligned}$$

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

$$\nabla_qu(t,q,\sigma_t)=\mathcal{V}[s,\mu](t,q),\quad 0\leqslant t\leqslant T,\ q\in\mathbb{T}^d. \tag{6.0.5}$$

Differentiating now with respect to t in (6.0.2), and substituting (6.0.5) into it, we conclude that

$$\partial_t(u(t,q,\sigma_t))+H(q,\nabla_qu(t,q,\sigma_t))+F(q,\sigma_t)=0. \tag{6.0.6}$$

Thus, we have shown:

Lemma 6.0.1. *For $s\in[0,T]$, $\mu\in\mathscr{P}(\mathbb{T}^d)$, we have:*

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

$$u(0,\cdot,\mu)=g(\cdot,\mu)\ ,\quad \nabla_qu(t,q,\sigma_t)=\mathcal{V}[s,\mu](t,q).$$

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

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

This lemma and formula (6.0.2) clarify something important about the relationship between U and u: if (U,σ) is the solution to the MFG system (1.3.2-1.3.5), then the mapping $t\mapsto U(t,q)$ is the same as the mapping $t\mapsto u(t,q,\sigma_t)$.

As explained in the Preliminaries and the Introduction, besides the conditions of Section 2.2.1, we now activate the conditions of Section 2.2.2, for ther remainder of the thesis.

6.1 Regularity of $\Sigma[s, \cdot](t, q)$

Since the solution U to the MFG system is given by an explicit representation formula, namely, formula (5.0.14), which is written in terms of $\Sigma[s, \mu]$, if we want to understand the dependence of u on μ and s, we must investigate the dependece of $\Sigma[s, \mu]$ on μ and s.

6.1.1 The discretized map M

For the remainder of the paper, let

$$\theta > \max\{1, 5\sqrt{2}\kappa\}, \tag{6.1.1}$$

and A_1, A_2, B, E, T, D be as in Proposition 3.1.5 and Corollary 3.1.7, with T being subject to Remark 3.1.8. The functions $\Sigma[s, \mu]$, $s \in [0, T], \mu \in \mathscr{P}(\mathbb{T}^d)$, as always, denote the fixed points of the operators $\mathfrak{m}^{s,\mu}$, while $\bar{\Sigma}[s, \mu] = (\Sigma^1[s, \mu], \frac{1}{\theta}\Sigma^2[s, \mu]) =: (\bar{\Sigma}^1[s, \mu], \bar{\Sigma}^2[s, \mu])$ are the fixed points of the operators $\bar{\mathfrak{m}}^{s,\mu}$; recall (3.1.15) from the proof of Corollary 3.1.7. The master map was defined in Definition 3.2.6 as $\mathfrak{M}(t, s, q, \mu) = (t, s, \Sigma^1[s, \mu](t, q), \mu)$. Let $M = (M_1, M_2)$ be the map $\Sigma = (\Sigma^1, \Sigma^2)$ restricted to average of Dirac masses. Namely:

Definition 6.1.1. *For any $s, t \in [0, T]$, $q \in \mathbb{T}^d$, $x \in (\mathbb{T}^d)^n$, let*

$$\begin{aligned} M = (M_1, M_2) : [0, T] \times [0, T] \times \mathbb{T}^d \times (\mathbb{T}^d)^n &\longrightarrow \mathbb{T}^d \times \mathbb{R}^d \\ (t, s, q, x) &\longmapsto \Sigma[s, \mu^x](t, q) \\ &\qquad = (\Sigma^1[s, \mu^x](t, q), \Sigma^2[s, \mu^x](t, q)). \end{aligned}$$

Note. The domain of the mapping M depends on $n \in Z^+$. //

As declared earlier, this section deals with the regularity of Σ in the measure variable μ. See Section 1.5.1 or the first paragraph of Section 6.1.3 for the general plan. Several preliminary smoothness estimates are set down first, before the main result is presented in Section 6.1.3. 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 Chapter 3.

To understand the following definition, the reader is encouraged to verify the following calculation. For $q \in \mathbb{T}^d$, $x \in (\mathbb{T}^d)^n$, let $g_n(q, x) := g(q, \mu^x)$. Then

$$g(\Sigma_t^1[s, \mu^x](q), \Sigma_t^1[s, \mu^x]_{\#}\mu^x] = g_n(M_1(t, s, q, x), M_1(t, s, x_1, x), \ldots, M_1(t, s, x_n, x)).$$

Definition 6.1.2. *(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*

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

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,

$$\bar{M}_1^{k+1}(t, s, q, x) = q + \int_s^t \nabla_p H(\bar{M}_1^k(\tau, s, q, x), \theta\bar{M}_2^k(\tau, s, q, x))d\tau$$

and

$$\bar{M}_2^{k+1}(t,s,q,x) = \frac{1}{\theta}\nabla_q g_n(\bar{M}_1^k(0,s,q,x),\bar{M}_1^k(0,s,x_1,x),\cdots,\bar{M}_1^k(0,s,x_n,x))$$
$$-\frac{1}{\theta}\int_0^t [\nabla_p H(\bar{M}_1^k(\tau,s,q,x),\theta\bar{M}_2^k(\tau,s,q,x))$$
$$+\nabla_q F_n(\bar{M}_1^k(t,s,q,x),\bar{M}_1^k(t,s,q,x_1),\cdots,\bar{M}_1^k(t,s,q,x_n))]d\tau,$$

where

$$F_n(q,x) := F(q,\mu^x), \quad g_n(q,x) := g(q,\mu^x).$$

(ii) Let $M^k, k = 0,1,\dots$ be the 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

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

Remark 6.1.3. It follows that, for every $k = 0,1,\dots$

$$M_1^k(t,s,q,x) = \bar{M}_1^k(t,s,q,x); \quad M_2^k(t,s,q,x) = \theta\bar{M}_2^k(t,s,q,x),$$

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

The objective is to obtain estimates on the derivatives of M^k with respect to x. Using the definition of Wasserstein gradient directly, one verifies the formulas

$$\nabla^2_{x_iq}F_n(q,x) = \frac{1}{n}\nabla_\mu\nabla_q F(q,\mu^x)(x_i), \qquad \nabla^2_{x_iq}g_n(q,x) = \frac{1}{n}\nabla_\mu\nabla_q g(q,\mu^x)(x_i). \tag{6.1.2}$$

Keep in mind that this is, actually, a matrix equality: the entries on the left hand side are $\partial_{x_i^{(j)}}\partial_{q^{(l)}}F_n(q,x)$ and those on the right are $\frac{1}{n}\nabla_{\mu_j}\partial_{q^{(l)}}F(q,\mu^x)(x_i)$, for $j,l = 1,\dots,d$, where $\nabla_{\mu_j}F(q,\mu)$ denotes the j-th component of the Wasserstein gradient of $\partial_{q^{(l)}}F(q,\mu^x)$ at x_i. Furthermore, since $\nabla_q F$ and $\nabla_q g$ are twice differentiable in the measure variable, we further know that

$$\nabla^2_{x_jx_i}\nabla_q F_n(q,x) = \frac{1}{n^2}\nabla^2_{\mu\mu}\nabla_q F(q,\mu^x)(x_i,x_j), \qquad \nabla^2_{x_jx_i}\nabla_q g_n(q,x) = \frac{1}{n^2}\nabla^2_{\mu\mu}\nabla_q g(q,\mu^x)(x_i,x_j). \tag{6.1.3}$$

We begin with the x_j-derivative of the $(k+1)$-th iteration ($j = 1,\dots,n$), and in this and subsequent calculations, we will include in the arguments of the functions only the variables that are relevant to them.

$$\nabla_{x_j}\bar{M}_1^{k+1}(t,s,q,x) = \int_s^t [\nabla^2_{qp}H(\bar{M}_1^k,\theta\bar{M}_2^k)\nabla_{x_j}\bar{M}_1^k + \theta\nabla^2_{pp}H(\bar{M}_1^k,\theta\bar{M}_2^k)\nabla_{x_j}\bar{M}_2^k]d\tau, \tag{6.1.4}$$

$$
\begin{aligned}
\nabla_{x_j}\bar{M}_2^{k+1}(t,s,q,x) = {} & \frac{1}{\theta}\nabla^2_{qq}g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_j}\bar{M}_1^k(q,x)\\
& + \frac{1}{\theta}\nabla^2_{x_jq}g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))[\nabla_q\bar{M}_1^k(x_j,x)+\nabla_{x_j}\bar{M}_1^k(x_j,x)]\\
& + \frac{1}{\theta}\sum_{i\neq j}\nabla^2_{x_iq}g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_j}\bar{M}_1^k(x_i,x)\\
& - \frac{1}{\theta}\int_0^t[\nabla^2_{qq}H(\bar{M}_1^k,\theta\bar{M}_2^k)\nabla_{x_j}\bar{M}_1^k+\theta\nabla^2_{pq}H(\bar{M}_1^k,\theta\bar{M}_2^k)\nabla_{x_j}\bar{M}_2^k]d\tau\\
& - \frac{1}{\theta}\int_0^t[\nabla^2_{qq}F_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_j}\bar{M}_1^k(q,x)\\
& \quad + \nabla^2_{x_jq}F_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))[\nabla_q\bar{M}_1^k(x_j,x)+\nabla_{x_j}\bar{M}_1^k(x_j,x)]\\
& \quad + \sum_{i\neq k}\nabla^2_{x_iq}F_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_j}\bar{M}_1^k(x_i,x)]d\tau.
\end{aligned}
\tag{6.1.5}
$$

For the next lemma, we remind the reader of Remark 3.2.5.

Lemma 6.1.4. *Fix* $n \in \mathbb{Z}^+$. *Using the terminology of Definition 6.1.2, the following hold:*

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

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

$$
\|\nabla_{x_j}M^k\|_\infty \leqslant \frac{C}{n}. \tag{6.1.6}
$$

Proof. (i) Clearly, $\bar{M}^0(\cdot,\cdot,\cdot,x) \in \mathcal{M}^*_{0,\bar{D}}(A_1, A_2, B, E, T)$, and by Lemma 3.2.4, each $\bar{M}^k(\cdot,\cdot,\cdot,x) \in \mathcal{M}^*_{0,\bar{D}}(A_1, A_2, B, E, T)$. By Remark 3.2.5, each $M^k(\cdot,\cdot,\cdot,x) \in \mathcal{M}^*_{0,D}(A_1, \theta A_2, \theta B, E, T)$.

(ii) Recall formulas (6.1.2), (6.1.3)—since $\nabla_q F$, $\nabla_q g$ and their first and second order Wasserstein gradients are uniformly bounded (conditions imposed in Section 2.2.2), we get

$$
\begin{aligned}
\|\nabla_{x_j}\bar{M}^{k+1}\|_\infty \leqslant & \big(\frac{\sqrt{2}\kappa}{\theta}E + \frac{\sqrt{2}\kappa}{\theta}TA_1\big)\frac{1}{n}\\
& + \|\nabla_{x_j}\bar{M}\|_\infty\big(\frac{\sqrt{2}\kappa}{\theta} + \frac{\sqrt{2}\kappa}{\theta}\frac{1}{n} + \frac{\sqrt{2}\kappa}{\theta}\frac{n-1}{n} +\\
& \qquad T\big[2\sqrt{2}(1+\theta)\bar{h}(B) + \frac{\sqrt{2}\kappa}{\theta} + \frac{\sqrt{2}\kappa}{\theta}\frac{1}{n} + \frac{\sqrt{2}\kappa}{\theta}\frac{n-1}{n}\big]\big)\\
\leqslant & \frac{\sqrt{2}\kappa}{\theta}(E + TA_1)\frac{1}{n} + \|\nabla_{x_j}\bar{M}^k\|_\infty\big(3\frac{\sqrt{2}\kappa}{\theta} + T(2\sqrt{2}(1+\theta)\bar{h}(B) + 3\frac{\sqrt{2}\kappa}{\theta})\big).
\end{aligned}
$$

By (6.1.1), we can invoke Remark 3.1.8 to obtain that the latter is an expression of the form

$$
\|\nabla_{x_j}\bar{M}^{k+1}\|_\infty \leqslant \frac{a}{n} + b\|\nabla_{x_j}\bar{M}^k\|_\infty
$$

with positive constants a, b in which $b < 1$. This holds for every $k = 0, 1, \dots$ Applying this inequality recursively (see [GŚ15, Remark 8.1]), we find a constant $C > 0$ such that $\|\nabla_{x_j} \bar{M}^k\|_\infty \leqslant C/n$, for every $k \in \mathbb{Z}^+$. Thus,

$$(\|\nabla_{x_j} M_1^k\|_\infty^2 + \|\frac{1}{\theta}\nabla_{x_j} M_2^k\|_\infty^2)^{1/2} = (\|\nabla_{x_j} \bar{M}_1^k\|_\infty^2 + \|\nabla_{x_j} \bar{M}_2^k\|_\infty^2)^{1/2} \leqslant \frac{C}{n}. \tag{6.1.7}$$

Multiplying by θ on both sides we get, since $\theta > 1$, that $(\|\nabla_{x_j} M_1^k\|^2 + \|\nabla_{x_j} M_2^k\|^2)^{1/2} \leqslant \frac{\theta C}{n}$, which is (6.1.6) for a larger constant C. □

Regularity of M in x, q and s

Corollary 6.1.5. *The sequence $\{M^k\}_1^\infty$ of Definition 6.1.2(ii) converges uniformly to the function M of Definition 6.1.1, with $M(\cdot,\cdot,\cdot,x) \in \mathcal{M}^*_{0,D}(A_1, \theta A_2, \theta B, E, T)$ 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*

$$\|\nabla^2_{qx_j} M\|_\infty \leqslant \frac{C}{n}, \quad j = 1, \dots, n\ , \tag{6.1.8}$$

$$\|\nabla^2_{x_i x_j} M\|_\infty \leqslant \frac{C}{n^2}, \quad i \neq j,\ i, j \in \{1, \dots, n\}\ , \tag{6.1.9}$$

$$\|\nabla^2_{x_j x_j} M\|_\infty \leqslant \frac{C}{n}, \quad j = 1, \dots, n\ , \tag{6.1.10}$$

$$\|\nabla^2_{s x_j} M\|_\infty \leqslant \frac{C}{n}, \quad j = 1, \dots, n. \tag{6.1.11}$$

Note. Since on $W^{2;\infty}(\mathbb{T}^d \times \mathbb{T}^d)$ the mixed partial derivatives $\nabla^2_{x_j x_i}$ and $\nabla^2_{x_i x_j}$ are equal, estimate (6.1.8) holds for $\nabla^2_{x_j q} M$ too.

Proof. The proof of Lemma 3.2.4 shows that for every $x \in (\mathbb{T}^d)^n$, $M^k(\cdot,\cdot,\cdot,x)$ converges uniformly to $\Sigma[\cdot,\mu^x](\cdot,\cdot)$. Formula (6.1.6) 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 (6.1.6) of Lemma 6.1.4; to be more precise:

$$\|\nabla_{x_j} M\|_\infty \leqslant \frac{C}{n}. \tag{6.1.12}$$

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 κ, so we again do it for $\bar{M}^k$ first. Differentiating (6.1.4) and (6.1.5) with respect to q, we get

$$\begin{aligned}\nabla^2_{qx_j} \bar{M}_1^{k+1}(t,s,q,x) = \int_s^t &[(\nabla^3_{qqp} H(\bar{M}_1^k, \theta\bar{M}_2^k)\nabla_q \bar{M}_1^k + \theta\nabla^3_{pqp} H(\bar{M}_1^k, \theta\bar{M}_2^k)\nabla_q \bar{M}_2^k)\nabla_{x_j}\bar{M}_1^k \\ &+ \nabla^2_{qp} H(\bar{M}_1^k, \theta\bar{M}_2^k)\nabla^2_{qx_j}\bar{M}_1^k \\ &+ \theta(\nabla^3_{qpp} H(\bar{M}_1^k, \theta\bar{M}_2^k)\nabla_q \bar{M}_1^k + \theta\nabla^3_{ppp} H(\bar{M}_1^k, \theta\bar{M}_2^k)\nabla_q \bar{M}_2^k)\nabla_{x_j}\bar{M}_2^k \\ &+ \theta\nabla^2_{pp} H(\bar{M}_1^k, \bar{M}_2^k)\nabla^2_{qx_j}\bar{M}_2^k]d\tau,\end{aligned}$$

and[1]

$$
\begin{aligned}
&\qquad\quad \nabla^2_{qx_j}\bar{M}_2^{k+1}(t,s,q,x) = \\
&= \frac{1}{\theta}\nabla^3_{qqq}g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_q\bar{M}_1^k(q,x)\nabla_{x_j}\bar{M}_1^k(q,x)\\
&+\frac{1}{\theta}\nabla^2_{qq}g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla^2_{x_jq}\bar{M}_1^k(q,x)\\
&+\frac{1}{\theta}\sum_{l\neq k}\big[\nabla^3_{qx_lq}g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_q\bar{M}_1^k(q,x)\nabla_{x_j}\bar{M}_1^k(x_l,x)\\
&\qquad\quad+\nabla^2_{x_lq}g_n(M_1(q,x),M_1(x_1,x),\dots,M_1(x_n,x))\nabla^2_{qx_j}M_1(x_l,x)\big]\\
&+\frac{1}{\theta}\nabla^3_{qx_jq}g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_q\bar{M}_1^k(q,x)(\nabla_q\bar{M}_1^k(x_j,x)+\nabla_{x_j}\bar{M}_1^k(x_j,x))\\
&+\frac{1}{\theta}\nabla^2_{x_jq}g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))[\nabla^2_{qq}\bar{M}_1^k(x_j,x)+\nabla^2_{qx_j}\bar{M}_1^k(x_j,x)]\\
&-\frac{1}{\theta}\int_0^t\cdots\ d\tau.
\end{aligned}
$$

Using (6.1.6) and the bounds on the coefficients, we get:

$$
\begin{aligned}
\|\nabla^2_{qx_j}\bar{M}^{k+1}\|_\infty \leqslant\ & 2\sqrt{2}T(1+\theta)\bar{h}(B)(A_1+\theta A_2)\frac{C}{n}+\sqrt{2}\frac{\kappa}{\theta}\frac{1}{n}(E(1+C)+\frac{n-1}{n}E\frac{C}{n}+E(E+\frac{C}{n}))\\
&+\sqrt{2}\frac{\kappa}{\theta}T\frac{1}{n}(A_1(1+C)+\frac{n-1}{n}A_1\frac{C}{n}+A_1(A_1+\frac{C}{n}))\\
&+\|\nabla^2_{qx_j}\bar{M}^k\|_\infty\big(\sqrt{2}T\frac{1+\theta}{\theta}\bar{h}(B)+\frac{\sqrt{2}\kappa}{\theta}(\frac{1}{n}+1+\frac{n-1}{n})+T\frac{\sqrt{2}\kappa}{\theta n}\big),
\end{aligned}
$$

which is an inequality of the form

$$
\|\nabla^2_{qx_j}\bar{M}^{k+1}\|_\infty \leqslant \frac{a}{n}+b\|\nabla^2_{qx_j}\bar{M}^k\|_\infty
$$

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 (6.1.7), we obtain (6.1.8) in the limit as $k\to\infty$.

Case $i\neq j$:

$$
\begin{aligned}
\nabla^2_{x_jx_i}\bar{M}_1^{k+1}(t,s,q,x) = \int_s^t\ &[(\nabla^3_{qqp}H(\bar{M}_1^k,\theta\bar{M}_2^k)\nabla_{x_j}\bar{M}_1^k+\theta\nabla^3_{pqp}H(\bar{M}_1^k,\bar{M}_2^k)\nabla_{x_j}\bar{M}_2^k)\nabla_{x_i}\bar{M}_1^k\\
&+\nabla^2_{qp}H(\bar{M}_1^k,\theta\bar{M}_2^k)\nabla^2_{x_jx_i}\bar{M}_1^k\\
&+\theta(\nabla^3_{qpp}H(\bar{M}_1^k,\theta\bar{M}_2^k)\nabla_{x_j}\bar{M}_1^k+\theta\nabla^3_{ppp}H(\bar{M}_1^k,\theta\bar{M}_2^k)\nabla_{x_j}\bar{M}_2^k)\nabla_{x_i}\bar{M}_2^k\\
&+\nabla^2_{pp}H(\bar{M}_1^k,\theta\bar{M}_2^k)\nabla^2_{x_jx_i}\bar{M}_2^k]d\tau,
\end{aligned}
$$

[1]For the second order gradients of $\bar{K}_2$, we will only write the part corresponding to g_n, knowing that the one corresponding to F_n has the exactly the same structure, and the expression coming from H is the same as the one just displayed, except that the last subindex in $\nabla^3_{\square\square\square}$ is q.

$$\nabla^2_{x_j x_i}\bar{M}_2^{k+1}(t,s,q,x) =$$

$$\begin{aligned}
&= \frac{1}{\theta}\big[\nabla^3_{qqq} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_j}\bar{M}_1^k(q,x)\\
&\quad + \nabla^3_{x_j qq} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))[\nabla_q\bar{M}_1^k(x_j,x)+\nabla_{x_j}\bar{M}_1^k(x_j,x)]\\
&\quad + \sum_{m\neq j}\nabla^3_{x_m qq} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_j}\bar{M}_1^k(x_m,x)\big]\nabla_{x_i}\bar{M}_1^k(q,x)\\
&\quad + \frac{1}{\theta}\nabla^2_{qq} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla^2_{x_j x_i}\bar{M}_1^k(q,x)\\
&+\frac{1}{\theta}\sum_{l\notin\{j,i\}}\Big[\big[\nabla^3_{q x_l q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_j}\bar{M}_1^k(q,x)\\
&\qquad + \nabla^3_{x_j x_l q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))[\nabla_q\bar{M}_1^k(x_j,x)+\nabla_{x_j}\bar{M}_1^k(x_j,x)]\\
&\qquad + \sum_{m\neq j}\nabla^3_{x_m x_l q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_j}\bar{M}_1^k(x_m,x)\big]\nabla_{x_i}\bar{M}_1^k(q,x)\\
&\qquad + \nabla^2_{x_l q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla^2_{x_j x_i}\bar{M}_1^k(x_l,x)\Big]\\
&\quad + \frac{1}{\theta}\big[\nabla^3_{q x_j q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_j}\bar{M}_1^k(q,x)\\
&\qquad + \nabla^3_{x_j x_j q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))[\nabla_q\bar{M}_1^k(x_j,x)+\nabla_{x_j}\bar{M}_1^k(x_j,x)]\\
&\qquad + \sum_{m\neq i}\nabla^3_{x_m x_j q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_j}\bar{M}_1^k(x_m,x)\big]\nabla_{x_i}\bar{M}_1^k(x_j,x)\\
&\quad + \frac{1}{\theta}\nabla^2_{x_j q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))(\nabla^2_{q x_i}\bar{M}_1^k(x_j,x)+\nabla^2_{x_j x_i}\bar{M}_1^k(x_j,x))\\
&\quad + \frac{1}{\theta}\big[\nabla^3_{q x_i q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_j}\bar{M}_1^k(q,x)\\
&\qquad + \nabla^3_{x_j x_i q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))[\nabla_q\bar{M}_1^k(x_i,x)+\nabla_{x_i}\bar{M}_1^k(x_i,x)]\\
&\qquad + \sum_{m\neq i}\nabla^3_{x_m x_i q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))\nabla_{x_i}\bar{M}_1^k(x_m,x)\big](\nabla_q\bar{M}_1^k(x_i,x)+\nabla_{x_i}\bar{M}_1^k(x_i,x))\\
&\quad + \frac{1}{\theta}\nabla^2_{x_i q} g_n(\bar{M}_1^k(q,x),\bar{M}_1^k(x_1,x),\dots,\bar{M}_1^k(x_n,x))[\nabla^2_{x_j q}\bar{M}_1^k(x_i,x)+\nabla^2_{x_j x_i}\bar{M}_1^k(x_i,x)]\\
&\quad - \frac{1}{\theta}\int_0^t \cdots\ d\tau.
\end{aligned}$$

Knowing the bound (6.1.8), we can now estimate:

$$\|\nabla^2_{x_j x_i}\bar{M}_1^{k+1}\|_\infty \leqslant T(1+\theta)[\big(\bar{h}(B)\frac{C}{n}+\theta\bar{h}(B)\frac{C}{n}\big)\frac{C}{n}+\bar{h}(B)\|\nabla^2_{x_j x_i}\bar{M}^k\|_\infty],$$

$$
\begin{aligned}
\|\nabla^2_{x_jx_i}\bar{M}_2^{k+1}\|_\infty \leqslant \big[&\frac{\kappa}{\theta n}C + \frac{\kappa}{\theta n}(E+\frac{C}{n}) + \frac{\kappa}{\theta n}(n-1)\frac{C}{n}\big]\frac{C}{n} + \frac{\kappa}{\theta}\|\nabla^2_{x_jx_i}\bar{M}^k\|_\infty \\
&+ (n-1)\big[\big(\frac{\kappa}{\theta n}\frac{C}{n} + \frac{\kappa}{\theta n^2}(E+\frac{C}{n}) + (n-1)\frac{\kappa}{\theta n^2}\frac{C}{n}\big) + \frac{\kappa}{\theta n}\|\nabla^2_{x_jx_i}\bar{M}^k\|_\infty\big] \\
&+ \big[\frac{\kappa}{\theta n}\frac{C}{n} + \frac{\kappa}{\theta n^2}(E+\frac{C}{n}) + (n-1)\frac{\kappa}{\theta n^2}\frac{C}{n}\big]\frac{C}{n} + \frac{\kappa}{\theta n}\big(\frac{C}{n} + \|\nabla^2_{x_jx_i}\bar{M}^k\|_\infty\big) \\
&+ \ \big[\frac{\kappa}{\theta n}\frac{C}{n} + \frac{\kappa}{\theta n^2}(E+\frac{C}{n}) + (n-1)\frac{\kappa}{\theta n^2}\frac{C}{n}\big]\big(E+\frac{C}{n}\big) + \frac{\kappa}{\theta n}\big(\frac{C}{n} + \|\nabla^2_{x_jx_i}\bar{M}^k\|_\infty\big) \\
&+ T(1+\theta)[\big(\bar{h}(B)\frac{C}{n} + \theta\bar{h}(B)\frac{C}{n}\big)\frac{C}{n} + \bar{h}(B)\|\nabla^2_{x_jx_i}\bar{M}^k\|_\infty] \\
&+ T\Bigg[\big[\frac{\kappa}{\theta n}C + \frac{\kappa}{\theta n}(A_1+\frac{C}{n}) + \frac{\kappa}{\theta n}(n-1)\frac{C}{n}\big]\frac{C}{n} + \frac{\kappa}{\theta}\|\nabla^2_{x_jx_i}\bar{M}^k\|_\infty \\
&\quad + (n-1)\big[\big(\frac{\kappa}{\theta n}\frac{C}{n} + \frac{\kappa}{\theta n^2}(A_1+\frac{C}{n}) + (n-1)\frac{\kappa}{\theta n^2}\frac{C}{n}\big) + \frac{\kappa}{\theta n}\|\nabla^2_{x_jx_i}\bar{M}^k\|_\infty\big] \\
&\quad + \big[\frac{\kappa}{\theta n}\frac{C}{n} + \frac{\kappa}{\theta n^2}(A_1+\frac{C}{n}) + (n-1)\frac{\kappa}{\theta n^2}\frac{C}{n}\big]\frac{C}{n} + \frac{\kappa}{\theta n}\big(\frac{C}{n} + \|\nabla^2_{x_jx_i}\bar{M}^k\|_\infty\big) \\
&\quad + \ \big[\frac{\kappa}{\theta n}\frac{C}{n} + \frac{\kappa}{\theta n^2}(A_1+\frac{C}{n}) + (n-1)\frac{\kappa}{\theta n^2}\frac{C}{n}\big]\big(A_1+\frac{C}{n}\big) + \frac{\kappa}{\theta n}\big(\frac{C}{n} + \|\nabla^2_{x_jx_i}\bar{M}^k\|_\infty\big)\Bigg].
\end{aligned}
$$

We see that $1/n^2$ factors out from the terms that do not involve $\|\nabla^2_{x_jx_i}\bar{M}^k\|_\infty$, while the term multiplying $\|\nabla^2_{x_jx_i}\bar{M}^k\|_\infty$ is

$$
\frac{\kappa}{\theta} + \frac{(n-2)\kappa}{\theta n} + \frac{\kappa}{\theta n} + \frac{\kappa}{\theta n} + T(\cdots).
$$

Thus, because of the lower bound imposed on θ, we get that

$$
\|\nabla^2_{x_jx_i}\bar{M}^{k+1}\|_\infty \leqslant \frac{a}{n^2} + b\|\nabla^2_{x_jx_i}\bar{M}^{k+1}\|_\infty,
$$

a, b, with $b<1$. The conclusion (6.1.9) follows.

Case $i=j$: since the formula for $\nabla^2_{x_jx_j}\bar{M}_1^{k+1}$ is the same as that for $\nabla^2_{x_jx_i}\bar{M}_1^{k+1}$, except with i

substituted by j, we only write out $\nabla^2_{x_jx_j}\bar{M}_2^{k+1}$:

$$
\begin{aligned}
&\nabla^2_{x_jx_j}\bar{M}_2^{k+1}(t,s,q,x) = \\
&= \frac{1}{\theta}\big[\nabla^3_{qqq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))\nabla_{x_j}\bar{M}_1(q,x) \\
&\quad + \nabla^3_{x_jqq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))[\nabla_q\bar{M}_1(x_j,x)+\nabla_{x_j}\bar{M}_1(x_j,x)] \\
&\quad + \frac{1}{\theta}\sum_{i\neq j}\nabla^3_{x_iqq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))\nabla_{x_j}\bar{M}_1(x_i,x)\big]\nabla_{x_j}\bar{M}_1(q,x) \\
&\quad + \frac{1}{\theta}\nabla^2_{qq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))\nabla^2_{x_jx_j}\bar{M}_1(q,x) \\
&+\frac{1}{\theta}\sum_{l\neq j}\Bigg[\big[\nabla^3_{qx_lq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))\nabla_{x_j}\bar{M}_1(q,x) \\
&\qquad + \nabla^3_{x_jx_lq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))[\nabla_q\bar{M}_1(x_j,x)+\nabla_{x_j}\bar{M}_1(x_j,x)] \\
&\qquad + \sum_{i\neq j}\nabla^3_{x_ix_lq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))\nabla_{x_j}\bar{M}_1(x_i,x)\big]\nabla_{x_j}\bar{M}_1(q,x) \\
&\qquad + \nabla^2_{x_lq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))\nabla^2_{x_jx_j}\bar{M}_1(x_l,x)\Bigg] \\
&\quad + \frac{1}{\theta}\big[\nabla^3_{qx_jq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))\nabla_{x_j}\bar{M}_1(q,x) \\
&\qquad + \nabla^3_{x_jx_jq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))[\nabla_q\bar{M}_1(x_j,x)+\nabla_{x_j}\bar{M}_1(x_j,x)] \\
&\qquad + \sum_{i\neq j}\nabla^3_{x_ix_jq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))\nabla_{x_j}\bar{M}_1(x_i,x)\big](\nabla_q\bar{M}_1(x_j,x)+\nabla_{x_j}\bar{M}_1(x_j,x)) \\
&\quad + \frac{1}{\theta}\nabla^2_{x_jq}g_n(\bar{M}_1(q,x),\bar{M}_1(x_1,x),\dots,\bar{M}_1(x_n,x))(\nabla^2_{qq}\bar{M}_1(x_j,x)+\nabla^2_{x_jq}\bar{M}(x_j,x) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \nabla^2_{qx_j}\bar{M}_1(x_j,x)+\nabla^2_{x_jx_j}\bar{M}_1(x_j,x)) \\
&\quad - \frac{1}{\theta}\int_0^t \cdots\ d\tau.
\end{aligned}
$$

Bounding yields

$$
\begin{aligned}
\|\nabla^2_{x_jx_j}\bar{M}_2^{k+1}\| \leqslant{}& \big[\frac{\kappa}{\theta}\frac{C}{n}+\frac{\kappa}{\theta n}(E+\frac{C}{n})+\frac{\kappa(n-1)}{\theta n}\frac{C}{n}\big]\frac{C}{n}+\frac{\kappa}{\theta}\|\nabla^2_{x_jx_j}\bar{M}^k\|_\infty \\
&+(n-1)\big[\big(\frac{\kappa}{\theta n}\frac{C}{n}+\frac{\kappa}{\theta n^2}(E+\frac{C}{n})+(n-1)\frac{\kappa}{\theta n^2}\big)\frac{C}{n}+\frac{\kappa}{\theta n}\|\nabla^2_{x_jx_j}\bar{M}^k\|_\infty\big] \\
&+\big(\frac{\kappa}{\theta n}\frac{C}{n}+\frac{\kappa}{\theta n^2}(E+\frac{C}{n})+(n-1)\frac{\kappa}{\theta n^2}\frac{C}{n}\big)(E+\frac{C}{n}) \\
&+\frac{\kappa}{\theta n}\big(E+\frac{C}{n}+\frac{C}{n}+\|\nabla^2_{x_jx_j}\bar{M}^k\|_\infty\big) \\
&+T(\cdots).
\end{aligned}
$$

This time, however, only $1/n$, not $1/n^2$, factors out from the terms not involving $\|\nabla^2_{x_jx_j}\bar{M}^k\|$, and the term multiplying $\|\nabla^2_{x_j x_j}\bar{M}^k\|$ is

$$\frac{\kappa}{\theta} + (n-1)\frac{\kappa}{\theta n} + \frac{\kappa}{\theta n} + T(\cdots).$$

Hence

$$\|\nabla^2_{x_j x_j}\bar{M}^{k+1}\|_\infty \leqslant \frac{a}{n} + b\|\nabla^2_{x_j x_j}\bar{M}^k\|_\infty$$

with $b < 1$, for all $k \in \mathbb{Z}^+$. Therefore (6.1.10) holds. Going through the same process once more, this time with the operator $\nabla^2_{s x_j}$, ends up in

$$\|\nabla^2_{s x_j}\bar{M}^{k+1}\|_\infty \leqslant a/n + b\|\nabla^2_{s x_j}\bar{M}^k\|_\infty,$$

with $b < 1$, from which the estimate (6.1.11) follows. □

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

$$\begin{aligned}
\partial_t M_1(t,s,q,x) &= \nabla_p H(M_1(t,s,q,x), M_2(t,s,q,x)),\\
\partial_t M_2(t,s,q,x) &= -\nabla_q H(M_1(t,s,q,x), M_2(t,s,q,x))\\
&\quad - \nabla_q F_n(M_1(t,s,q,x), M_1(t,s,x_1,x), \ldots, M_1(t,s,x_n,x)),
\end{aligned}$$

differentiating once more with respect to t and knowing that $M(\cdot,\cdot,\cdot,x) \in \mathcal{M}^*_{0,D}(A_1, \theta A_2, \theta B, E, T)$ for every $x \in (\mathbb{T}^d)^n$, we obtain

$$\|\partial^2_{tt}M\|_\infty,\ \|\nabla^2_{qt}M\|_\infty \leqslant 2\sqrt{2}(h(\theta B)(A_1+\theta A_2)+\kappa A_1), \quad \|\nabla^2_{x_j t}M\|_\infty \leqslant \sqrt{2}(4h(\theta B)\frac{C}{n}+\frac{\kappa}{n}(A_1+2C)), \tag{6.1.13}$$

for $j = 1, \ldots, n$. Equipped now with estimates on all the second order derivatives of $M(\cdot, s, \cdot, \cdot)$, we proceed to obtain the Taylor estimate in the following corollary.

Corollary 6.1.6. *Let $M = M(t,s,q,x)$, $M' = M(t',s,q',x')$, with the notation of Corollary 6.1.2. There is a constant $C > 0$ such that*

$$\begin{aligned}
&|M' - M - \partial_t M(t'-t) - \nabla_q M \cdot (q'-q) - \nabla_x M \cdot (x'-x)|\\
&\leqslant C(|t'-t|^2 + |q-q'|^2 + |x-x'|^2).
\end{aligned}$$

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 $\mathbb{R}^{2d}$. //

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

$$\begin{aligned}
&|M_i'^{(j)} - M_i^{(j)} - \partial_t M_i^{(j)}(t'-t) - \nabla_q M_i^{(j)} \cdot (q'-q) - \nabla_x M_i^{(j)}(x'-x)| \\
&\quad \leqslant \|\partial^2_{tt} M_i^{(j)}\|_\infty |\Delta t|^2 + 2\|\nabla^2_{tq} M_i^{(j)}\|_\infty |\Delta t||\Delta q| + 2\sum_{l=1}^n \|\nabla^2_{tx_l} M_i^{(j)}\|_\infty |\Delta t||\Delta x_l| + \|\nabla^2_{qq} M_i^{(j)}\|_\infty |\Delta q|^2 \\
&\qquad + 2\sum_{l=1}^n \|\nabla^2_{qx_l} M_i^{(j)}\|_\infty |\Delta q||\Delta x_l| + \sum_{\substack{l,m=1\\ l\neq m}}^n \|\nabla^2_{x_l x_m} M_i^{(l)}\|_\infty |\Delta x_l||\Delta x_m| + \sum_{l=1}^n \|\nabla^2_{x_l x_l} M_i^{(j)}\|_\infty |\Delta x_l|^2.
\end{aligned}$$

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

$$\begin{aligned}
&|M' - M - \partial_t M(t'-t) - \nabla_q M \cdot (q'-q) - \nabla_x M \cdot (x'-x)| \\
&\leqslant \|\partial^2_{tt} M\|_\infty |\Delta t|^2 + 2\|\nabla^2_{tq} M\|_\infty |\Delta t||\Delta q| + 2\sum_{l=1}^n \|\nabla^2_{tx_l} M\|_\infty |\Delta t||\Delta x_l| + \|\nabla^2_{qq} M\|_\infty |\Delta q|^2 \\
&\quad + 2\sum_{l=1}^n \|\nabla^2_{qx_l} M\|_\infty |\Delta q||\Delta x_l| + \sum_{\substack{l,m=1\\ l\neq m}}^n \|\nabla^2_{x_l x_m} M_i^{(l)}\|_\infty |\Delta x_l||\Delta x_m| + \sum_{l=1}^n \|\nabla^2_{x_l x_l} M\|_\infty |\Delta x_l|^2 \\
&\leqslant 2\sqrt{2}(A_1 + A_2)(\kappa + h(\theta B))(|\Delta t|^2 + 2|\Delta t||\Delta q|) + 4\sqrt{2}\sqrt{2}n\frac{C}{n}(2h(\theta B) + \kappa)|\Delta x||\Delta t| \\
&\quad + \sqrt{2}(A_1 + A_2)|\Delta q|^2 + 2n\frac{C}{n}\sqrt{2}|\Delta x||\Delta q| + n(n-1)\frac{C}{n^2}\sqrt{2}\sqrt{2}|\Delta x|^2 + n\frac{C}{n}|\Delta x|^2.
\end{aligned}$$

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 $x' = (x'_1, \dots, x'_n)$ so that $|x - x'|^2 = \mathscr{W}^2(\mu^x, \mu^{x'})$. Thus, the inequality of Corollary 6.1.6 reads

$$\begin{aligned}
&|M' - M - \partial_t M(t'-t) - \nabla_q M \cdot (q'-q) - \nabla_x M \cdot (x'-x)| \\
&\quad \leqslant C(|t'-t|^2 + |q-q'|^2 + \mathscr{W}^2(\mu^x, \mu^{x'})).
\end{aligned} \tag{6.1.14}$$

6.1.2 The discretized map N $(= M^{-1})$

Let us define

$$\begin{aligned}
N : [0,T] \times [0,T] \times \mathbb{T}^d \times (\mathbb{T}^d)^n &\longrightarrow \mathbb{T}^d \\
(t,s,q,x) &\longmapsto X[s,\mu^x](t,q);
\end{aligned} \tag{6.1.15}$$

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

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

Let us define also

$$\begin{aligned}\mathfrak{M}^n : [0,T]\times[0,T]\times\mathbb{T}^d\times(\mathbb{T}^d)^n &\longrightarrow [0,T]\times[0,T]\times\mathbb{T}^d\times(\mathbb{T}^d)^n\\ (t,s,q,x) &\longmapsto (t,s,M_1(t,s,q,x),x),\end{aligned}$$

where the superindex n now makes it explicit that n is the number of particles. The map $\mathfrak{M}^n$ is a "discretized" version of the master map $\mathfrak{M}$ of Definition 3.2.6. By Lemma 3.2.8, $\mathfrak{M}^n$ is a diffeomorphism, with inverse

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

for $t,s\in[0,T], q\in\mathbb{T}^d, x\in(\mathbb{T}^d)^n$; $\mathfrak{N}^n$ is thus the "discretized" version of the inverse of the master map. By Corollary 6.1.5 and (6.1.13), for each fixed $s\in[0,T]$, $\mathfrak{M}^n(\cdot,s,\cdot,\cdot)\in W^{2,2,2;\infty}([0,T]\times\mathbb{T}^d\times(\mathbb{T}^d)^n;[0,T]\times\mathbb{T}^d\times(\mathbb{T}^d)^n)$. 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 (3.2.5):

$$\|\nabla_q X[s,\mu](t,\cdot)\|_\infty < 4(1+\sqrt{d})^{d-1}, \qquad s,t\in[0,T],\mu\in\mathscr{P}(\mathbb{T}^d).$$

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

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

from which

$$\partial_t X[s,\mu](t,q) = -\nabla_q X[s,\mu](t,q)v[s,\mu](t,q), \tag{6.1.16}$$

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

$$\|\partial_t X_t[s,\mu]\|_\infty \leqslant \|\nabla_q X_t[s,\mu]\|_\infty\|v_t[s,\mu]\|_\infty \leqslant 4A_1(1+\sqrt{d})^{d-1}. \tag{6.1.17}$$

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

$$M_1(t,s,N(t,s,q,x),x)\equiv q,$$

which holds by definition. Taking the derivative with respect to x_j, $j=1,\dots,n$, 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, $\|\nabla_{x_j}N\|_\infty \leqslant 4(1+\sqrt{d})^{d-1}\frac{C}{n}$, which, increasing the value of C, gives

$$\|\nabla_{x_j}N\|_\infty \leqslant \frac{C}{n}. \tag{6.1.18}$$

Corollary 6.1.7. *Let $t,t',s\in[0,T]$, $q',q\in\mathbb{T}^d$, $\mu,\nu\in\mathscr{P}(\mathbb{T}^d)$. Then there is a constant $C>0$ such that*

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

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 (6.1.15). By the bounds (6.1.17), (3.2.5), (6.1.18), and relabeling the sequence $x_1, \dots, x_n$ and shifting the points so that $\mathscr{W}^2(\mu^x, \mu^{x'}) = \sum_{j=1}^n |x_j - x_j'|^2$,

$$\begin{aligned}|N(t',s,q',x') - N(t,s,q,x)| &\leqslant \|\partial_t N\|_\infty |t'-t| + \|\nabla_q N\|_\infty |q'-q| + \sum_{j=1}^n \|\nabla_{x_j} N\|_\infty |x_j' - x_j| \\ &\leqslant 4A_1(1+\sqrt{d})^{d-1}|t'-t| + 4(1+\sqrt{d})^{d-1})|q'-q| + \sum_{j=1}^n \frac{C}{n}|x_j'-x_j|;\end{aligned}$$

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

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

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

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

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

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

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

Regularity of the master map is obtained in the next paragraphs from the foregoing properties of its discretized version.

6.1.3 Regularity properties of the master map

We follow [GŚ15] closely here. The idea, roughly speaking, begins with introducing a Lipschitz extension of the "discretized derivative" $\nabla_{x_j} M_1$, $j = 1, \dots, n$, that is defined at every measure μ, 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 $x^n = (x_1^n, \dots, x_n^n)$ 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 (2.1.6).

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

$$\begin{aligned}\mathcal{B} &:= [0,T] \times [0,T] \times \mathbb{T}^d \times \{(y_j, \mu^y) \mid y = (y_1, \dots, y_n) \in (\mathbb{T}^d)^n, j \in \{1, \dots, n\}\} \\ &\subset [0,T] \times [0,T] \times \mathbb{T}^d \times [(\mathbb{T}^d) \times \mathscr{P}(\mathbb{T}^d)].\end{aligned}$$

A typical element of $\mathcal{B}$ is thus $(t,s,q,(y_j,\mu^y))$ where y is any n-particle ordered set $(y_1, \dots, y_n) \in (\mathbb{T}^d)^n$ 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

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

For any continuous function $f = (f^{(1)}, \dots, f^{(m)}) : \mathcal{B} \to \mathbb{R}^m$ such that

$$|f(t,s,q,(y_j,\mu^y)) - f(t,s,q,(x_i,\mu^x))| \leqslant C(|x_i - y_j|_{\mathbb{T}^d} + \mathscr{W}(\mu^x,\mu^y) + \frac{1}{n}), \tag{6.1.19}$$

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

$$g^{(l)}(t,s,q,z,\mu) := \inf\left\{ f^{(l)}(t,s,q,(y_j,\mu^y)) + C(|z-y_j|_{\mathbb{T}^d} + \mathscr{W}(\mu,\mu^y)) \;\middle|\; y \in (\mathbb{T}^d)^n, j \in \{1,\dots,n\}\right\},$$
$$l = 1,\dots,m,$$
$$g := (g^{(1)},\dots,g^{(m)}),$$

at any fixed $z \in \mathbb{T}^d, \mu \in \mathscr{P}(\mathbb{T}^d)$. The function g is thus an extension of f from $\mathcal{B}$ to the full space $[0,T]^2 \times \mathbb{T}^d \times [\mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d)]$. The following is [GŚ15, Lemma 8.10].

Proposition 6.1.8. *Suppose that (6.1.19) holds, and for any $x \in (\mathbb{T}^d)^n$, $j \in \{1,\dots,n\}$, $f(\cdot,\cdot,\cdot,(x_j,\mu^x))$ is C-Lipschitz. Then*

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

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

As in [GŚ15], we set, for $s,t \in [0,T]$, $q \in \mathbb{T}^d$, $x = (x_1,\dots,x_n) \in (\mathbb{T}^d)^n$, $j = 1,\dots,n$,

$$\zeta^n(t,s,q,(x_j,\mu^x)) = n\nabla_{x_j} M(t,s,q,x). \tag{6.1.20}$$

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

Corollary 6.1.9. *(Extension of ζ^n) For each $n \in \mathbb{Z}^+$, there is a function*

$$\chi^n : [0,T] \times [0,T] \times \mathbb{T}^d \times \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}^{d^2} \times \mathbb{R}^{d^2}$$

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

(i) χ^n is C-Lipschitz,

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

Proof. We check that $f = \zeta^n$ satisfies the conditions of Proposition 6.1.8. The Lipschitz property in t and q follows from (6.1.13), while, in s, from (6.1.11). Hence, to obtain the Corollary, it is enough to prove that the condition (6.1.19) is satisfied by $f = \zeta^n$. Fix then $x,y \in (\mathbb{T}^d)^n$, $i,j \in \{1,\dots,n\}$, $s,t \in [0,T]$, $q \in \mathbb{T}^d$. Since the order in which we take the n particles $x_1,\dots,x_n$, which make up $x \in (\mathbb{T}^d)^n$, does not change $M(\cdots,x)$, and $\nabla_{x_j} M(\cdots,x)$ is periodic in x, it can be assumed that:

$$\sum_{k \neq i,j} |x_k - y_k|^2 \leqslant \mathscr{W}^2(\mu^x,\mu^y), \qquad |x_j - y_i| = |x_j - y_i|_{\mathbb{T}^d}, \quad |x_i - y_j| = |x_i - y_j|_{\mathbb{T}^d},$$
$$\nabla_{x_j} M(t,s,q,y) = \nabla_{x_1} M(t,s,q,\bar{y}), \quad \nabla_{x_i} M(t,s,q,x) = \nabla_{x_1} M(t,s,q,\bar{x}),$$

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,

$$\begin{aligned}
|\nabla_{x_j} M(t,s,q,y) - &\nabla_{x_i} M(t,s,q,x)| \\
&\leqslant \|\nabla^2_{x_1,x_1} M\|_\infty |y_j - x_i| + \|\nabla^2_{x_1,x_2} M\|_\infty |y_i - x_j| + \sum_{k=1}^{i-1} \|\nabla^2_{x_1,x_{k+2}} M\|_\infty |y_k - x_k| \\
&\quad + \sum_{k=i+1}^{j-1} \|\nabla^2_{x_1,x_{k+1}} M\|_\infty |y_k - x_k| + \sum_{k=j+1}^{n} \|\nabla^2_{x_1,x_k} M\|_\infty |y_k - x_k|.
\end{aligned}$$

Therefore, by the bounds of Corollary 6.1.5,

$$\begin{aligned}
|\nabla_{x_j} M(t,s,q,y) - \nabla_{x_i} M(t,s,q,x)| &\leqslant \frac{C}{n}|y_j - x_i| + \frac{C}{n^2}|y_i - x_j| + \frac{C}{n^2}\sum_{k\neq i,j} |y_k - x_k| \\
&\leqslant \frac{C}{n}|y_j - x_i|_{\mathbb{T}^d} + \frac{C}{n^2}|y_i - x_j|_{\mathbb{T}^d} + \frac{C\sqrt{n}}{n^2}\sqrt{\sum_{k\neq i,j} |y_k - x_k|^2} \\
&\leqslant \frac{C}{n}(|y_j - x_i|_{\mathbb{T}^d} + \frac{\sqrt{d}}{2n} + \mathscr{W}(\mu^x, \mu^y)),
\end{aligned}$$

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

$$|n\nabla_{x_j} M(t,s,q,y) - n\nabla_{x_i} M(t,s,q,x)| \leqslant \sqrt{d}C(|y_j - x_i|_{\mathbb{T}^d} + \mathscr{W}(\mu^x, \mu^y) + \frac{1}{n}), \tag{6.1.21}$$

which proves property (6.1.19) for $f = \zeta^n$, since i and j were arbitrary. □

Lemma 6.1.10. *For every $s \in [0,T]$, the $\mathbb{T}^d \times \mathbb{R}^d$-valued map $\Sigma[s,\cdot](\cdot,\cdot)$ is differentiable on $\mathscr{P}(\mathbb{T}^d) \times [0,T] \times \mathbb{T}^d$, that is: there exists a mapping*

$$\begin{aligned}
\bar{\nabla}_\mu \Sigma : [0,T] \times [0,T] \times \mathbb{T}^d \times \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d) &\longrightarrow \mathbb{R}^{d^2} \times \mathbb{R}^{d^2} \\
(t,s,q,x,\mu) &\longmapsto \bar{\nabla}_\mu \Sigma[s,\mu](t,q,x)
\end{aligned}$$

such that, for every $s,t,t' \in [0,T]$, $q,q' \in \mathbb{T}^d$, $\mu,\nu \in \mathscr{P}(\mathbb{T}^d)$, $\gamma \in \Gamma_0(\mu,\nu)$,

$$\begin{aligned}
&\big|\Sigma[s,\nu](t',q') - \Sigma[s,\mu](t,q) - \partial_t \Sigma[s,\mu](t,q)(t'-t) - \nabla_q \Sigma[s,\mu](t,q)\cdot(q'-q) \\
&\qquad - \int_{\mathbb{T}^d\times\mathbb{T}^d} \bar{\nabla}_\mu \Sigma[s,\mu](t,q,x)\cdot(y-x)\gamma(dx,dy)\big| \\
&\leqslant C(|t'-t|^2 + |q'-q|^2 + \mathscr{W}^2(\mu,\nu)).
\end{aligned} \tag{6.1.22}$$

Moreover, the mapping $\bar{\nabla}_\mu \Sigma$ is Lipschitz.

Proof. First, let us define $\bar{\nabla}\Sigma$. Let ζ^n be defined as in (6.1.20). Denote by χ^n the extension of ζ^n furnished by Corollary 6.1.9. By Corollary 6.1.9(i), each χ^n, $n = 1,\dots$ is C-Lipschitz on the bounded domain $[0,T]^2 \times (\mathbb{T}^d)^2 \times \mathscr{P}(\mathbb{T}^d)$. The functions χ^n are also pointwise uniformly bounded, because of (6.1.12). Thus, the sequence $\{\chi^n\}_{n=1}^{\infty}$ is equicontinuous and pointwise uniformly bounded, so a subsequence of it converges to a C-Lipschitz mapping, which we define as the mapping $\bar{\nabla}_\mu \Sigma$.

Let $\mu, \nu \in \mathscr{P}(\mathbb{T}^d)$, and let $\gamma \in \Gamma_0(\mu,\nu)$. Appealing again to Chapter 2, there is a sequence $\{\gamma(n)\}_{n=1}^{\infty}$, converging narrowly to γ in $\mathcal{P}(\mathbb{T}^d \times \mathbb{T}^d)$, such that

$$\gamma(n) = \frac{1}{n}\sum_{j=1}^{n} \delta_{(x_j(n),y_j(n))},$$

and for each $j \in \{1,\dots,n\}$, $(x_j(n), y_j(n))$ belongs to the support of γ. Due to this latter fact (see, for instance, [AGS08, Theorem 6.1.4]), for each $n \in \mathbb{Z}^+$, the sequence $\{(x_j(n), y_j(n))\}_{j=1}^{n}$ is $|\cdot|_{\mathbb{T}^d}$-monotone, and, therefore,

$$\gamma(n) \in \Gamma_0(\mu^{x(n)}, \mu^{y(n)}), \qquad n \in \mathbb{Z}^+.$$

It is also true that

$$\lim_{n\to\infty} \mathscr{W}(\mu, \mu^{x(n)}) = 0, \qquad \lim_{n\to\infty} \mathscr{W}(\nu, \mu^{y(n)}) = 0.$$

For each $x(n)$ of our sequence,

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

$j \in \{1,\dots,n\}$. Recall the second-order estimate (6.1.14), now with $x = x(n)$ and $x' = y(n)$:

$$\begin{aligned}
&\big|\Sigma[s,\mu^{y(n)}](t,q) - \Sigma[s,\mu^{x(n)}](t',q') - \partial_t\Sigma[s,\mu^{x(n)}](t,q)(t'-t) - \nabla_q\Sigma[s,\mu^{x(n)}](t,q)\cdot(q'-q)\\
&\qquad - \sum_{j=1}^{n} \nabla_{x_j(n)}\Sigma[s,\mu^{x(n)}](t,q)\cdot(y_j(n)-x_j(n))\big|\\
&\leqslant C(|t'-t|^2 + |q-q'|^2 + \mathscr{W}^2(\mu^{x(n)},\mu^{y(n)})).
\end{aligned}$$

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

$$\begin{aligned}
&\big|\Sigma[s,\mu^{y(n)}](t',q') - \Sigma[s,\mu^{x(n)}](t,q) - \partial_t\Sigma[s,\mu^{x(n)}](t,q)(t'-t) - \nabla_q\Sigma[s,\mu^{x(n)}](t,q)\cdot(q'-q)\\
&\qquad - \int_{\mathbb{T}^d\times\mathbb{T}^d} \zeta^n(t,s,q,(x,\mu^{x(n)}))\cdot(y-x)\gamma(n)(dx,dy)\big|\\
&\leqslant C(|t'-t|^2 + |q-q'|^2 + \mathscr{W}^2(\mu^{x(n)},\mu^{y(n)})).
\end{aligned}$$

The same inequality holds if we substitute χ^n for ζ^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

$$
\begin{aligned}
&\big|\Sigma[s,\mu^{y(n)}](t,q)-\Sigma[s,\mu^{x(n)}](t,q)-\partial_t\Sigma[s,\mu^{x(n)}](t,q)(t'-t)-\nabla_q\Sigma[s,\mu^{x(n)}](t,q)\cdot(q'-q)\\
&\quad-\int\limits_{\mathbb{T}^d\times\mathbb{T}^d}\chi^n(t,s,q,(x,\mu^{x(n)}))\cdot(y-x)\gamma(n)(dx,dy)\big|\\
&\leqslant C(|t'-t|^2+|q-q'|^2+\mathscr{W}^2(\mu^{x(n)},\mu^{y(n)}))\\
&\quad+\big|\int\limits_{\mathbb{T}^d\times\mathbb{T}^d}[\zeta^n(t,s,q,(x,\mu^{x(n)}))-\chi^n(t,s,q,(x,\mu^{x(n)}))]\cdot(y-x)\gamma(n)(dx,dy)\big|\\
&\leqslant C(|t'-t|^2+|q-q'|^2+\mathscr{W}^2(\mu^{x(n)},\mu^{y(n)}))+\frac{C}{n}\mathscr{W}(\mu^{x(n)},\mu^{y(n)}),
\end{aligned}
\tag{6.1.23}
$$

by Corollary 6.1.9(ii) and the fact $\gamma(n)\in\Gamma_0(\mu^{x(n)},\nu^{y(n)})$. Passing to the limit as $n\to\infty$ in (6.1.23), we prove (6.1.22), in particular, that Σ (and, therefore, the master map) is differentiable in the μ variable. ☐

We will denote by $\bar\nabla_\mu\Sigma^1$ the first component (the $\mathbb{T}^d$-valued part) of $\bar\nabla_\mu\Sigma$, and $\bar\nabla_\mu\Sigma^2$ the second component ($\mathbb{R}^d$-valued) of $\bar\nabla_\mu\Sigma$. Note the bar over the nabla in $\bar\nabla_\mu\Sigma$. We have put the bar so as to not mislead the reader into thinking that $\bar\nabla_\mu\Sigma$ is *the* Wasserstein gradient, which we would denote by $\nabla_\mu\Sigma$; recall that, for us, the Wasserstein gradient of a function $\mu\mapsto\mathcal{W}(\mu)$ is the element of minimal norm in the set $\partial\mathcal{W}(\mu)$. Lemma 6.1.10 shows that

$$\bar\nabla_\mu\Sigma\in\partial[\Sigma_t[s,\cdot](q)](\mu).$$

Next, we are going to prove the analogue of Lemma 6.1.10 for $X=(\Sigma^1)^{-1}$, which amounts to smoothness of $\mathfrak{N}$ in the μ variable.

Definition 6.1.11. *For $t,s\in[0,T]$, $\mu\in\mathscr{P}(\mathbb{T}^d)$, $q,x\in\mathbb{T}^d$, put*

$$\bar\nabla_\mu X[s,\mu](t,q,x):=-\nabla_q X[s,\mu](t,q)\bar\nabla_\mu\Sigma^1[s,\mu](t,X[s,\mu](t,q),x).$$

Before stating and proving the lemma, we recall formula (6.1.16):

$$\partial_t X[s,\mu](t,q)=-\nabla_q X[s,\mu](t,q)\partial_t\Sigma^1[s,\mu](t,\cdot)\circ X[s,\mu](t,q).$$

Lemma 6.1.12. *For every $s\in[0,T]$, the $\mathbb{R}^d$-valued map $X[s,\cdot](\cdot,\cdot)$ is differentiable on $\mathscr{P}(\mathbb{T}^d)\times[0,T]\times\mathbb{T}^d$, i.e., there is a constant $C>0$ such that for every $s,t,t'\in[0,T]$, $q,q'\in\mathbb{T}^d$, $\mu,\nu\in\mathscr{P}(\mathbb{T}^d)$, $\gamma\in\Gamma_0(\mu,\nu)$,*

$$
\begin{aligned}
&\big|X[s,\nu](t',q')-X[s,\mu](t,q)-\partial_t X[s,\mu](t,q)(t'-t)-\nabla_q X[s,\mu](t,q)\cdot(q'-q)\\
&\quad-\int\limits_{\mathbb{T}^d\times\mathbb{T}^d}\bar\nabla_\mu X[s,\mu](t,q,x)\cdot(y-x)\gamma(dx,dy)\big|\\
&\leqslant C(|t'-t|^2+|q-q'|^2+\mathscr{W}^2(\mu,\nu)),
\end{aligned}
\tag{6.1.24}
$$

where $\partial_t X$, $\nabla_q X$, and the mapping $\bar\nabla_\mu X$ of Definition 6.1.11, are continuous.

Proof. The continuity of $\bar{\nabla}_\mu X$ is immediate from its definition, and the continuity of $\nabla_q X$ and $\partial_t X$ has been known since Lemma 3.2.8 and formula (6.1.16) respectively. Let us put

$$\tilde{q} = X[s,\mu](t,q), \quad \tilde{q}' = X[s,\nu](t',q'). \tag{6.1.25}$$

We write out the expression on the left hand side of (6.1.24) and factor out $\nabla_q X[s,\mu](t,q)$, while also using the fact that $\nabla_q X[s,\mu](t,q)$ and $\nabla_q \Sigma^1[s,\mu](t,\tilde{q})$ are inverses of one another:

$$\begin{aligned}
&\Bigg| \nabla_q X[s,\mu](t,q)\Bigg[\nabla_q \Sigma^1[s,\mu](t,\tilde{q})(\tilde{q}'-\tilde{q}) + \partial_t \Sigma^1[s,\mu](t,\tilde{q})(t'-t) \\
&\quad - (\Sigma^1[s,\nu](t',\tilde{q}') - \Sigma^1[s,\mu](t,\tilde{q})) + \int_{\mathbb{T}^d\times\mathbb{T}^d} \bar{\nabla}_\mu \Sigma^1[s,\mu](t,\tilde{q},x)\cdot(y-x)\gamma(dx,dy)\Bigg]\Bigg| \\
= &\Bigg| -\nabla_q X[s,\mu](t,q)\Bigg[\Sigma^1[s,\nu](t',\tilde{q}') - \Sigma^1[s,\mu](t,\tilde{q}) - \partial_t\Sigma^1[s,\mu](t,\tilde{q})(t'-t) \\
&\quad - \nabla_q\Sigma^1[s,\mu](t,\tilde{q})(\tilde{q}'-\tilde{q}) - \int_{\mathbb{T}^d\times\mathbb{T}^d} \bar{\nabla}_\mu \Sigma^1[s,\mu](t,\tilde{q},x)\cdot(y-x)\gamma(dx,dy)\Bigg]\Bigg| \\
\leqslant &\ 4(1+\sqrt{d})^{d-1}C(|t'-t|^2 + |\tilde{q}'-\tilde{q}|^2 + \mathscr{W}^2(\mu,\nu)) \\
= &\ 4(1+\sqrt{d})^{d-1}C(|t'-t|^2 + |X[s,\nu](t',\tilde{q}') - X[s,\mu](t,q)|^2 + \mathscr{W}^2(\mu,\nu)).
\end{aligned}$$

But, by Corollary 6.1.7, the term $|X[s,\nu](t',q') - X[s,\mu](t,q)|$ is bounded by $C(|t'-t| + |q'-q| + \mathscr{W}(\mu,\nu))$ for some $C > 0$. Inserting this bound into the last expression, after expanding and raising the value of C, one obtains (6.1.24). □

Regularity of $\mathcal{V}$

Let us look back at the definition of $\mathcal{V}$, given by (4.1.20). Set now

$$\bar{\nabla}_\mu \mathcal{V}[s,\mu](t,q,x) := \bar{\nabla}_\mu \Sigma^2[s,\mu](t,X[s,\mu](t,q),x) + \nabla_q \Sigma^2[s,\mu](t,X[s,\mu](t,q))\bar{\nabla}_\mu X[s,\mu](t,q,x), \tag{6.1.26}$$

for $s,t \in [0,T]$, $q,x \in \mathbb{T}^d$, $\mu \in \mathscr{P}(\mathbb{T}^d)$.

Lemma 6.1.13. *For every $s \in [0,T]$, the $\mathbb{R}^d$-valued map $\mathcal{V}[s,\cdot](\cdot,\cdot)$ is differentiable on $\mathscr{P}(\mathbb{T}^d) \times [0,T] \times \mathbb{T}^d$, i.e., there is a constant $C > 0$ such that for every $s,t,t' \in [0,T]$, $q,q' \in \mathbb{T}^d$, $\mu,\nu \in \mathscr{P}(\mathbb{T}^d)$, $\gamma \in \Gamma_0(\mu,\nu)$,*

$$\begin{aligned}
&\big|\mathcal{V}[s,\nu](t',q') - \mathcal{V}[s,\mu](t,q) - \partial_t \mathcal{V}[s,\mu](t,q)(t'-t) - \nabla_q \mathcal{V}[s,\mu](t,q)\cdot(q'-q) \\
&\qquad - \int_{\mathbb{T}^d\times\mathbb{T}^d} \bar{\nabla}_\mu \mathcal{V}[s,\mu](t,q,x)\cdot(y-x)\gamma(dx,dy)\big| \\
&\leqslant C(|t'-t|^2 + |q-q'|^2 + \mathscr{W}^2(\mu,\nu)),
\end{aligned} \tag{6.1.27}$$

where the mapping $\bar{\nabla}_\mu\mathcal{V}$, defined by (6.1.26), and $\partial_t\mathcal{V}$, $\nabla_q\mathcal{V}$, are continuous.

Proof. We know that

$$\nabla_q \mathcal{V}_t[s,\mu] = \nabla_q\Sigma^2_t[s,\mu]\nabla_q X_t[s,\mu], \quad \partial_t\mathcal{V}_t[s,\mu] = \partial_t\Sigma^2_t[s,\mu] + \nabla_q\Sigma^2_t[s,\mu]\partial_t X_t[s,\mu].$$

Therefore, the continuity of the functions stated in the lemma follows from that of $\nabla_q\Sigma$, $\nabla_q X$, $\partial_t\Sigma$, $\partial_t X$, and the continuity, proved above, of $\bar\nabla_\mu\Sigma$ and $\bar\nabla_\mu X$. Keeping the notation (6.1.25), we first write down the expression to estimate, i.e. the left hand side of (6.1.27), and factor out $\nabla_q\Sigma^2[s,\mu](t,\tilde q)$:

$$\begin{aligned}
&\big|\Sigma^2[s,\nu](t',\tilde q') - \Sigma^2[s,\mu](t,\tilde q) - \nabla_q\Sigma^2[s,\mu](t,\tilde q)\nabla_q X[s,\mu](t,q)(q'-q)\\
&\quad - (\partial_t\Sigma^2[s,\mu](t,\tilde q) + \nabla_q\Sigma^2[s,\mu](t,\tilde q)\partial_t X[s,\mu](t,q))(t'-t)\\
&\quad - \int_{\mathbb{T}^d\times\mathbb{T}^d}\big[\bar\nabla_\mu\Sigma^2[s,\mu](t,\tilde q,x) + \nabla_q\Sigma^2[s,\mu](t,\tilde q)\bar\nabla_\mu X[s,\mu](t,q,x)\big]\cdot(y-x)\gamma(dx,dy)\big|\\
=\ &\big|\Sigma^2[s,\nu](t',\tilde q') - \Sigma^2[s,\mu](t,\tilde q) - \nabla_q\Sigma^2[s,\mu](t,\tilde q)\ \Big(\\
&\qquad \nabla_q X[s,\mu](t,q)(q'-q) + \partial_t X[s,\mu](t,q)(t'-t) + \int_{\mathbb{T}^d\times\mathbb{T}^d}\bar\nabla_\mu X[s,\mu](t,q)\cdot(y-x)\gamma(dx,dy)\ \Big)\\
&\quad - \partial_t\Sigma^2[s,\mu](t,\tilde q)(t'-t) - \int_{\mathbb{T}^d\times\mathbb{T}^d}\bar\nabla_\mu\Sigma^2[s,\mu](t,\tilde q,x)\cdot(y-x)\gamma(dx,dy)\big|.
\end{aligned}$$

Inside the large round brackets we add and substract $\tilde q' - \tilde q = X_{t'}[s,\nu](q') - X_t[s,\mu](q)$, and apply (6.1.24), to obtain that the latter expression is no greater than

$$\begin{aligned}
&\big|\Sigma^2_{t'}[s,\nu](\tilde q') - \Sigma^2_t[s,\mu](\tilde q) - \nabla_q\Sigma^2_t[s,\mu](\tilde q)(\tilde q - q) - \partial_t\Sigma^2_t[s,\mu](\tilde q)(t'-t)\\
&- \int_{\mathbb{T}^d\times\mathbb{T}^d}\bar\nabla_\mu\Sigma^2_t[s,\mu](\tilde q,x)\cdot(y-x)\gamma(dx,dy)\big| + |\nabla_q\Sigma^2_t[s,\mu](\tilde q)|C(|q'-q|^2 + |t'-t|^2 + \mathscr{W}^2(\mu,\nu)),
\end{aligned}$$

which, in turn, by (6.1.24), is bounded above by

$$\begin{aligned}
&C(|X[s,\nu](t',q') - X[s,\mu](t,q)|^2 + |t'-t|^2 + |q'-q|^2 + \mathscr{W}^2(\mu,\nu))\\
&+ \theta A_2 C(|q'-q|^2 + |t'-t|^2 + \mathscr{W}^2(\mu,\nu)),
\end{aligned}$$

and, after using Corollary 6.1.7 again, simplifying and increasing the value of C, inequality (6.1.27) is obtained. □

Regularity of $H(q,\mathcal{V})$

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 (1.3.2), due to (5.0.8). Set

$$\bar\nabla_\mu H(q,\mathcal{V}[s,\mu](t,q))(x) := \nabla_p H(q,\mathcal{V}[s,\mu](t,q))\bar\nabla_\mu\mathcal{V}[s,\mu](t,q,x), \tag{6.1.28}$$

for $s,t\in[0,T]$, $q,x\in\mathbb{T}^d$, $\mu\in\mathscr{P}(\mathbb{T}^d)$.

Lemma 6.1.14. *For every $s\in[0,T]$, the $\mathbb{R}$-valued map $H(\cdot,\mathcal{V}[s,\cdot](\cdot,\cdot))$ is differentiable on $\mathscr{P}(\mathbb{T}^d)\times[0,T]\times\mathbb{T}^d$, and there is a constant $C>0$ such that for every $s,t,t'\in[0,T]$, $q,q'\in\mathbb{T}^d$, $\mu,\nu\in\mathscr{P}(\mathbb{T}^d)$, $\gamma\in\Gamma_0(\mu,\nu)$,*

$$\begin{aligned}
&\big|H(q',\mathcal{V}[s,\nu](t',q')) - H(q,\mathcal{V}[s,\mu](t,q)) - (\partial_t)(H(q,\mathcal{V}[s,\mu](t,q))(t'-t)\\
&\quad - (\nabla_q)(H(q,\mathcal{V}[s,\mu](t,q)))\cdot(q'-q) - \int_{\mathbb{T}^d\times\mathbb{T}^d}\bar\nabla_\mu H(q,\mathcal{V}[s,\mu](t,q))(x)\cdot(y-x)\gamma(dx,dy)\big|\\
&\leqslant C(|t'-t|^2 + |q-q'|^2 + \mathscr{W}^2(\mu,\nu)),
\end{aligned}\tag{6.1.29}$$

where the mapping $\bar{\nabla}_\mu H$, defined by (6.1.28), is continuous.

Proof. Let us abbreviate $\mathcal{V} = \mathcal{V}_t[s,\mu](q)$ and $\mathcal{V}' = \mathcal{V}_{t'}[s,\nu](q')$. Since

$$(\partial_t)(H(q,\mathcal{V})) = \nabla_p H(q,\mathcal{V})\partial_t\mathcal{V}, \quad (\nabla_q)(H(q,\mathcal{V})) = \nabla_q H(q,\mathcal{V}) + \nabla_p H(q,\mathcal{V})\nabla_q\mathcal{V},$$

the left hand side of (6.1.29) is, after factoring out $\nabla_p H(q,\mathcal{V})$,

$$\begin{aligned}
&\big|H(q',\mathcal{V}') - H(q,\mathcal{V}) - \nabla_p H(q,\mathcal{V})\big[-(\mathcal{V}'-\mathcal{V}) + \nabla_q\mathcal{V}(q'-q) + \partial_t\mathcal{V}(t'-t) \\
&\qquad\qquad\qquad + \int_{\mathbb{T}^d\times\mathbb{T}^d} \bar{\nabla}_\mu \mathcal{V}_t[s,\mu](q)\cdot(y-x)\gamma(dx,dy)\big] \\
&\quad - \nabla_q H(q,\mathcal{V})(q'-q) - \nabla_p H(q,\mathcal{V}')(\mathcal{V}-\mathcal{V}')\big| \\
&\leqslant |H(q',\mathcal{V}') - H(q,\mathcal{V}) - \nabla_q H(q,\mathcal{V})(q'-q) - \nabla_p H(q,\mathcal{V})(\mathcal{V}'-\mathcal{V})| \\
&\quad + |\nabla_p H(q,\mathcal{V})|C(|t'-t|^2 + |q'-q|^2 + \mathscr{W}^2(\mu,\nu)).
\end{aligned}$$

Remember now that $|\Sigma^2| \leqslant \theta B$ (see Corollary 3.1.7(ii)) at any t,q,s,μ; recall Definition 3.2.2. Therefore, the right-hand side of this inequality is bounded by

$$h(\theta B)(|q'-q|^2 + |\mathcal{V}'-\mathcal{V}|^2) + l(\theta B)C(|t'-t|^2 + |q'-q|^2 + \mathscr{W}^2(\mu,\nu)).$$

To deal with the term $|\mathcal{V}'-\mathcal{V}|^2$, note that Corollary 6.1.7 is also valid for Σ in place of X, following a similar argument. With the notation (6.1.25),

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

Applying Corollary 6.1.7, and raising the value of C, we get

$$|\mathcal{V}'-\mathcal{V}| \leqslant C(|t'-t| + |q'-q| + \mathscr{W}(\mu,\nu)).$$

Substituting this into the bounding expression, and simplifying, one arrives at (6.1.29), for some larger value of C. The continuity of $\bar{\nabla}_\mu H$ in all its variables is clear from the definition. □

6.2 Solution to the master equation

Let us recall, once again, the definition of the function u :

$$u(s,q,\mu) = g(q,\Sigma^1[s,\mu](0,\cdot)_\#\mu) - \int_0^s \big[H(q,\mathcal{V}[s,\mu](\tau,q)) + F(q,\Sigma^1[s,\mu](\tau,\cdot)_\#\mu\big]d\tau. \tag{6.0.1}$$

6.2.1 Pathwise gradients of the couplings

For any $s,t \in [0,T)$, $q \in \mathbb{T}^d$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, define

$$\begin{aligned}
\mathcal{N}_t^F[s,\mu](q)(x) := &-\nabla_\mu F(q,\sigma_t)(\Sigma_t^1[s,\mu](x))\nabla_q\Sigma_t^1[s,\mu](x) \\
&+ \int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(\Sigma_t^1[s,\mu](r))\bar{\nabla}_\mu\Sigma_t^1[s,\mu](r)(x)\mu(dr),
\end{aligned}$$

and

$$\mathcal{N}_t^g[s,\mu](q)(x) := -\nabla_\mu g(q,\sigma_t)(\Sigma_t^1[s,\mu](x))\nabla_q\Sigma^1[s,\mu](t,x) + \int_{\mathbb{T}^d} \nabla_\mu g(q,\sigma_t)(\Sigma_t^1[s,\mu](r))\bar{\nabla}_\mu\Sigma_t^1[s,\mu](r)(x)\mu(dr).$$

Likewise,

$$(\partial_t)(F(q,\Sigma_t^1[s,\mu]_\#\mu)) := \nabla_\mu F(q,\sigma_t)(\Sigma_t^1[s,\mu](x))\cdot\partial_t\Sigma^1[s,\mu](t,x)\mu(dx),$$

with an analogous definition for $(\partial_t)(g(q,\Sigma_t^1[s,\mu]_\#\mu))$. In preparation for Lemma 6.2.3 below, we are going to adopt this notation: for $t,t'\in(0,T)$, $q,q'\in\mathbb{T}^d$, $\mu,\nu\in\mathscr{P}(\mathbb{T}^d)$, and $0\leqslant\tau\leqslant 1$, and $\gamma\in\Gamma(\mu,\nu)$,

$$\begin{cases} t^\tau = (1-\tau)t+\tau t', \quad q^\tau = (1-\tau)q+\tau q', \quad \mu^\tau = ((1-\tau)\pi^1+\tau\pi^2)_\#\gamma, \\ \sigma'_{t'} = \Sigma^1[s,\nu](t',q'), \quad \sigma_t = \Sigma^1[s,\mu](t,q), \\ \sigma^\tau = \Sigma_{t^\tau}^1[s,\mu^\tau]_\#\mu^\tau. \end{cases} \tag{6.2.1}$$

Recall that, by definition, $X=(\Sigma^1)^{-1}$. In this context, we are going to need the following:

Proposition 6.2.1. *Let $\mu,\nu\in\mathscr{P}(\mathbb{T}^d)$, $q,q'\in\mathbb{T}^d$, $t,t'\in(0,T)$, and let the notation (6.2.1) be in force. Let w^τ be the velocity vector field of the geodesic μ^τ.*

(i) The vector field

$$v^\tau(y) = (t'-t)v_{t^\tau}[s,\mu^\tau](y) + \int_{\mathbb{T}^d}\bar{\nabla}_\mu\Sigma_{t^\tau}^1[s,\mu^\tau](X_{t^\tau}[s,\mu^\tau](y))(x)w^\tau(x)\mu^\tau(dx) - \nabla_q\Sigma_{t^\tau}^1[s,\mu^\tau](X_{t^\tau}[s,\mu^\tau](y))w^\tau(X_{t^\tau}[s,\mu^\tau](y)), \quad y\in\mathbb{T}^d, \tag{6.2.2}$$

is a velocity vector field for the path σ^τ.

(ii)

$$\|v^\tau\|_{L^2(\sigma^\tau)} \leqslant C(|t'-t|+\mathscr{W}(\mu,\nu)), \quad 0\leqslant\tau\leqslant 1$$

for some constant C.

Proof. *(i)* Fix any $\gamma\in\Gamma_0(\mu,\nu)$ and for $0\leqslant\tau\leqslant 1$, put μ^τ as in (6.2.1). By definition, v^τ must satisfy

$$\frac{d}{d\tau}\int_{\mathbb{T}^d}\varphi(y)\sigma^\tau(dy) = \int_{\mathbb{T}^d}\nabla\varphi(y)\cdot v^\tau(y)\sigma^\tau(dy) \tag{6.2.3}$$

for every $\varphi\in C^\infty(\mathbb{T}^d)$ and for $\mathscr{L}^1$-a.e. $t\in(0,1)$. Computing the derivative on the left hand side,

$$\frac{d}{d\tau}\int_{\mathbb{T}^d}\varphi(\Sigma^1[s,\mu^\tau](t^\tau,y))\mu^\tau(dy) = = \int_{\mathbb{T}^d}\frac{d}{d\tau}\varphi(\Sigma^1[s,\mu^\tau](t^\tau,y))\mu^\tau(dy) - \int_{\mathbb{T}^d}(\nabla_y)[\varphi(\Sigma^1[s,\mu^\tau](t^\tau,y))]w^\tau(y)\mu^\tau(dy),$$

where w^τ is the velocity vector field of the geodesic μ^τ, and we have again used the definition of velocity, since $\varphi \circ \Sigma^1$ is C^∞ in y. Doing the differentiation with respect to τ and y, we get

$$\frac{d}{d\tau}\int_{\mathbb{T}^d} \varphi(y)\sigma^\tau(dy) =$$
$$= \int_{\mathbb{T}^d} \big[\nabla\varphi(\Sigma^1[s,\mu^\tau](t^\tau,y)\cdot\Big(\partial_t\Sigma^1[s,\mu^\tau](t^\tau,y)(t'-t) + \int_{\mathbb{T}^d}\bar\nabla_\mu\Sigma^1[s,\mu^\tau](t^\tau,y)(x)w^\tau(x)\mu^\tau(dx)\Big)\big]\mu^\tau(dy)$$
$$-\int_{\mathbb{T}^d}\nabla\varphi(\Sigma^1[s,\mu^\tau](t^\tau,y))\cdot[\nabla_q\Sigma^1[s,\mu^\tau](t^\tau,y)w^\tau(y)]\mu^\tau(dy).$$

Since $\Sigma^1_{t^\tau}[s,\mu^\tau]_\#\mu^\tau = \sigma^\tau$ by definition, we obtain (6.2.2) after writing the latter expression as an integral with respect to σ^τ and comparing against the right hand side of (6.2.3).

(ii) This follows by Remark 2.1.1(i) and the boundedness of $\partial_t\Sigma^1$, $\nabla_q\Sigma^1$, $\bar\nabla_\mu\Sigma^1$. □

Remark 6.2.2. It follows that, evaluated at $\Sigma^1_{t^\tau}[s,\mu^\tau](y)$, v^τ has the simpler expression:

$$v^\tau(\Sigma^1_{t^\tau}[s,\mu^\tau](y)) = (t'-t)\partial_t\Sigma^1_{t^\tau}[s,\mu^\tau](y) + \int_{\mathbb{T}^d}\bar\nabla_\mu\Sigma^1_{t^\tau}[s,\mu^\tau](y)(x)w^\tau(x)\mu^\tau(dx)$$
$$-\nabla_q\Sigma^1_{t^\tau}[s,\mu^\tau](y)w^\tau(y), \qquad y\in\mathbb{T}^d. \qquad /\!/$$

Lemma 6.2.3. *Let $t,t'\in(0,T)$, $\mu,\nu\in\mathscr{P}(\mathbb{T}^d)$, $q,q'\in\mathbb{T}^d$ be arbitrary, and put $\sigma_t = \Sigma^1_t[s,\mu]_\#\mu$, $\sigma'_{t'} = \Sigma^1_{t'}[s,\nu]_\#\nu$. Then there exists a constant C such that*

$$\big|F(q',\sigma'_{t'}) - F(q,\sigma_t) - \nabla_q F(q,\sigma_t)\cdot(q'-q)$$
$$-\int_{\mathbb{T}^d}\nabla_\mu F(q,\sigma_t)(\Sigma^1_t[s,\mu](x))\cdot\partial_t\Sigma^1[s,\mu](t,x)\mu(dx)(t'-t)$$
$$-\int_{\mathbb{T}^d\times\mathbb{T}^d}\mathcal{N}^F_t[s,\mu](q)(x)\cdot(y-x)\gamma(dx,dy)\big|$$
$$\leqslant C(|t'-t|^2 + |q'-q|^2 + \mathscr{W}^2(\mu,\nu)),$$

for any $\gamma\in\Gamma_0(\mu,\nu)$. The functions $\mathcal{N}^F_t[s,\mu](q)(x)$ and $(\partial_t)(F(q,\Sigma^1_t[s,\mu]_\#\mu)$ are continuous in all its variables.

Naturally, the same result holds for g in place of F.

Proof. Before we begin the proof, let us note that the continuity assertion follows immediately from the continuity of the functions that enter in the definition of $\mathcal{N}^F_t[s,\mu](q)(x)$ and $(\partial_t)(F(q,\Sigma^1_t[s,\mu]_\#\mu)$.

Let $\gamma\in\Gamma_0(\mu,\nu)$, and define μ^τ, $0\leqslant\tau\leqslant 1$ as in Remark 2.1.1. Let the notation (6.2.1) be in effect, so that $\tau\mapsto\sigma^\tau$ is a continuous path joining σ_t with $\sigma'_{t'}$ and Proposition 6.2.1 holds, with w^τ as defined therein. Denote by E the expression inside the bars on the left-hand side of the inequality of the lemma.

Step 1.

Claim 1.

$$\int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(x)\cdot v^0(x)\sigma_t(dx) = \int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(\Sigma_t^1[s,\mu](x))\cdot \partial_t \Sigma_t^1[s,\mu](x)\mu(dx)(t'-t) + \int_{\mathbb{T}^d\times\mathbb{T}^d} \mathcal{N}_t^F[s,\mu](q)(x)(y-x)\gamma(dx,dy).$$

Proof of Claim 1. Using Remark 6.2.2 with $\tau = 0$,

$$\begin{aligned}
&\int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(x)\cdot v^0(x)\sigma_t(dx) = \\
&= \int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(\Sigma_t^1[s,\mu](x))\cdot \partial_t \Sigma_t^1[s,\mu](y)\mu(dx)(t'-t) \\
&- \int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(\Sigma_t^1[s,\mu](x))\cdot \big[\nabla_q \Sigma_t^1[s,\mu](x) w^0(x) - \int_{\mathbb{T}^d} \bar\nabla_\mu \Sigma_t^1[s,\mu](x)(b)w^0(b)\mu(db)\big]\mu(dx).
\end{aligned} \tag{6.2.4}$$

Note that

$$\begin{aligned}
&\int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(\Sigma_t^1[s,\mu](x))\cdot \big[\int_{\mathbb{T}^d} \bar\nabla_\mu \Sigma_t^1[s,\mu](x)(b)w^0(b)\mu(db)\big]\mu(dx) \\
&= \int_{\mathbb{T}^d}\int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(\Sigma_t^1[s,\mu](x))\cdot \bar\nabla_\mu \Sigma_t^1[s,\mu](x)(b)\mu(dx)w^0(b)\mu(db) \\
&= \int_{\mathbb{T}^d}\int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(\Sigma_t^1[s,\mu](r))\cdot \bar\nabla_\mu \Sigma_t^1[s,\mu](r)(x)\mu(dr)w^0(x)\mu(dx).
\end{aligned}$$

Substituting this identity into (6.2.4), we see that

$$\begin{aligned}
&\int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(x)\cdot v^0(x)\sigma_t(dx) = \\
&= \int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(\Sigma_t^1[s,\mu](x))\cdot \partial_t \Sigma_t^1[s,\mu](y)\mu(dx)(t'-t) + \int_{\mathbb{T}^d} \mathcal{N}_t^F[s,\mu](q)(x)\cdot w^0(x)\mu(dx)
\end{aligned}$$

Finally, we use Remark 2.1.1(ii) for the last integral in the latter expression, and the claim is proved.

Claim 2.

$$E = F(q',\sigma'_{t'}) - F(q,\sigma_t) - \nabla_q F(q,\sigma_t)\cdot(q'-q) - \int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(x)\cdot v^0(x)\sigma_t(dx).$$

Proof of Claim 2. This follows immediately from Claim 1 and the definition of E.

Thus, the lemma will be proved if we show that

$$\begin{aligned}
&|F(q',\sigma'_{t'}) - F(q,\sigma_t) - \nabla_q F(q,\sigma_t)\cdot(q'-q) - \int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(x)\cdot v^0(x)\sigma_t(dx)| \\
&\leqslant C(|t'-t|^2 + |q'-q|^2 + \mathscr{W}^2(\mu,\nu))
\end{aligned} \tag{6.2.5}$$

for some constant C.

Step 2. Before we set out to prove (6.2.5), we transform the expression for E some more. By the chain rule, we have that

$$F(q',\sigma'_{t'}) - F(q,\sigma_t) = \int_0^1 \big[\nabla_q F(q^\tau,\sigma^\tau)\cdot(q'-q) + \nabla_\mu F(q^\tau,\sigma^\tau)(x)\cdot v^\tau(x)\sigma^\tau(dx)\big]d\tau,$$

so

$$E = \int_0^1 \Bigg([\nabla_q F(q^\tau,\sigma^\tau) - \nabla_q F(q,\sigma_t)]\cdot(q'-q) + \int_{\mathbb{T}^d} \nabla_\mu F(q^\tau,\sigma^\tau)\cdot v^\tau(x)\sigma^\tau(dx) - \int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)\cdot v^0(x)\sigma_t(dx)\Bigg)d\tau.$$

With our knowledge of v^τ, that is, from Remark 6.2.2), we may rewrite E as

$$\begin{aligned}E = \int_0^1 \Bigg[&(\nabla_q F(q^\tau,\sigma^\tau) - \nabla_q F(q,\sigma_t))\cdot(q'-q)\\
&+ \int_{\mathbb{T}^d} \nabla_\mu F(q^\tau,\sigma^\tau)(\Sigma^1_{t^\tau}[s,\mu^\tau](x))\cdot\Bigg(\partial_t\Sigma^1_{t^\tau}[s,\mu^\tau](x)(t'-t)\\
&\qquad + \int_{\mathbb{T}^d} \bar{\nabla}_\mu\Sigma^1_{t^\tau}[s,\mu^\tau](x)(r)w^\tau(r)\mu^\tau(dr) - \nabla_q\Sigma^1_{t^\tau}[s,\mu^\tau](x)w^\tau(x)\Bigg)\mu^\tau(dx)\\
&- \int_{\mathbb{T}^d} \nabla_\mu F(q,\sigma_t)(\Sigma^1_t[s,\mu](x))\cdot\Bigg(\partial_t\Sigma^1_t[s,\mu](x)(t'-t)\\
&\qquad + \int_{\mathbb{T}^d} \bar{\nabla}_\mu\Sigma^1_t[s,\mu](x)(r)w^0(r)\mu(dr) - \nabla_q\Sigma^1_t[s,\mu](x)w^0(x)\Bigg)\mu(dx)\Bigg]d\tau.\end{aligned}$$

Now, let

$$\gamma^\tau \in \Gamma(\mu,\mu^\tau), \quad 0<\tau<1.$$

Then E takes the form

$$\begin{aligned}E =& \int_0^1 (\nabla_q F(q^\tau,\sigma^\tau) - \nabla_q F(q,\sigma_t))\cdot(q'-q)d\tau\\
&+ \int_0^1\int_{\mathbb{T}^d\times\mathbb{T}^d} \Bigg[\nabla_\mu F(q^\tau,\sigma^\tau)(\Sigma^1_{t^\tau}[s,\mu^\tau](y))\cdot\Bigg(\partial_t\Sigma^1_{t^\tau}[s,\mu^\tau](y)(t'-t)\\
&\qquad + \int_{\mathbb{T}^d} \bar{\nabla}_\mu\Sigma^1_{t^\tau}[s,\mu^\tau](y)(r)w^\tau(r)\mu^\tau(dr) - \nabla_q\Sigma^1_{t^\tau}[s,\mu^\tau](y)w^\tau(y)\Bigg)\\
&\quad - \nabla_\mu F(q,\sigma_t)(\Sigma^1_t[s,\mu](x))\cdot\Bigg(\partial_t\Sigma^1_t[s,\mu](x)(t'-t)\\
&\qquad + \int_{\mathbb{T}^d} \bar{\nabla}_\mu\Sigma^1_t[s,\mu](x)(r)w^0(r)\mu(dr) - \nabla_q\Sigma^1_t[s,\mu](x)w^0(x)\Bigg)\Bigg]\gamma^\tau(dx,dy)d\tau.\\
=:& E_1 + E_2,\end{aligned}$$

where

$$E_1 = \int_0^1 (\nabla_q F(q^\tau,\sigma^\tau) - \nabla_q F(q,\sigma_t))\cdot(q'-q)d\tau, \quad E_2 = E - E_1.$$

Step 3 (estimates). In the following, we will be making use of the boundedness of the quantities displayed in Section 2.2.2. With an abuse of notation, for each fixed $\tau \in (0,1]$, let $\tau(h) = h\tau$, $0 \leqslant h \leqslant 1$, so that $t^{\tau(\cdot)}$ has endpoints t, t^τ, $q^{\tau(\cdot)}$ has endpoints q, q^τ and $\sigma^{\tau(\cdot)}$ has endpoints σ_t, σ^τ. The velocity vector field of the path $h \mapsto v^{\tau(h)}$ is $\tau v^{\tau(h)}$ (see, e.g., [AGS08]). An argument similar to the one in the proof of Proposition 2.1.6, based on the continuity of $\nabla^2_{qq}F$ and $\nabla^2_{\mu q}F$, shows that

$$\nabla_q F(q^\tau, \sigma^\tau) - \nabla_q F(q, \sigma_t) =$$
$$= \int_0^1 \big[\tau \nabla^2_{qq} F(q^{\tau(h)}, \sigma^{\tau(h)})(q' - q) + \int_{\mathbb{T}^d} \tau \nabla^2_{\mu q} F(q^{\tau(h)}, \sigma^{\tau(h)})(x) v^{\tau(h)}(x) \sigma^{\tau(h)}(dx)\big] dh.$$

Therefore, by Proposition (6.2.1)(ii), we get

$$|E_1| \leqslant C|q' - q|\big(|q' - q| + |t' - t| + \mathscr{W}(\mu, \nu)\big) \tag{6.2.6}$$

for some positive constant C. We break E_2 down as follows:

$$E_2 = \int_0^1 \int_{\mathbb{T}^d \times \mathbb{T}^d} (B_1 + B_2 + B_3) \gamma^\tau(dx, dy) d\tau,$$

where

$$B_1 := \big[\nabla_\mu F(q^\tau, \sigma^\tau)(\Sigma^1_{t^\tau}[s, \mu^\tau](y)) - \nabla_\mu F(q, \sigma_t)(\Sigma^1_t[s, \mu](y))\big] \cdot$$
$$\cdot \big[\partial_t \Sigma^1_{t^\tau}[s, \mu^\tau](y)(t' - t) - \nabla_q \Sigma^1_{t^\tau}[s, \mu^\tau](y) w^\tau(y) + \int_{\mathbb{T}^d} \bar{\nabla}_\mu \Sigma^1_{t^\tau}[s, \mu^\tau](y)(r) w^\tau(r) \mu^\tau(dr)\big],$$

$$\begin{aligned} B_2 := \ \nabla_\mu F(q, \sigma_t)(\Sigma^1_t[s, \mu](y)) \cdot \big[&\partial_t \Sigma^1_{t^\tau}[s, \mu^\tau](y)(t' - t) - \nabla_q \Sigma^1_{t^\tau}[s, \mu^\tau](y) w^\tau(y) \\ &+ \int_{\mathbb{T}^d} \bar{\nabla}_\mu \Sigma^1_{t^\tau}[s, \mu^\tau](y)(r) w^\tau(r) \mu^\tau(dr) \\ &- \partial_t \Sigma^1_t[s, \mu](x)(t' - t) + \nabla_q \Sigma^1_t[s, \mu](x) w^0(x) \\ &- \int_{\mathbb{T}^d} \bar{\nabla}_\mu \Sigma^1_t[s, \mu](x)(r) w^0(r) \mu(dr)\big], \end{aligned}$$

and

$$B_3 := \big[\nabla_\mu F(q, \sigma_t)(\Sigma^1_t[s, \mu](y)) - \nabla_\mu F(q, \sigma_t)(\Sigma^1_t[s, \mu](x)\big] \cdot$$
$$\cdot \big[\partial_t \Sigma^1_t[s, \mu](x)(t' - t) - \nabla_q \Sigma^1_t[s, \mu](x) w^0(x) + \int_{\mathbb{T}^d} \bar{\nabla}_\mu \Sigma^1_t[s, \mu](x)(r) w^0(r) \mu(dr)\big].$$

To estimate $\int_0^1 \int_{(\mathbb{T}^d)^2} B_1 \gamma^\tau d\tau$, we address the first square bracket in the definition of B_1. With $\tau w^{\tau(h)}$

being the velocity vector field of the path $h \mapsto w^{\tau(h)}$, we have

$$\nabla_\mu F(q^\tau, \sigma^\tau)(\Sigma^1_{t^\tau}[s,\mu^\tau](y)) - \nabla_\mu F(q,\sigma_t)(\Sigma^1_t[s,\mu](y)) =$$
$$= \int_0^1 \Big[\tau \nabla^2_{q\mu} F(q^{\tau(h)}, \sigma^{\tau(h)})(\Sigma^1_{t^{\tau(h)}}[s,\mu^{\tau(h)}](y))(q'-q)$$
$$+ \int_{\mathbb{T}^d} \tau \nabla^2_{\mu\mu} F(q^{\tau(h)}, \sigma^{\tau(h)})(\Sigma^1_{t^{\tau(h)}}[s,\mu^{\tau(h)}](y))(b) v^{\tau(h)}(b) \sigma^{\tau(h)}(db)$$
$$+ \nabla^2_{x\mu} F(q^{\tau(h)}, \sigma^{\tau(h)})(\Sigma^1_{t^{\tau(h)}}[s,\mu^{\tau(h)}](y)) \big[\tau(t'-t)\partial_t \Sigma^1_{t^{\tau(h)}}[s,\mu^{\tau(h)}](y)$$
$$+ \int_{\mathbb{T}^d} \tau \bar{\nabla}_\mu \Sigma^1_{t^{\tau(h)}}[s,\mu^{\tau(h)}](y)(b) w^{\tau(h)}(b) \mu^{\tau(h)}(db)\big]\Big] dh,$$

so

$$\begin{aligned} &|\nabla_\mu F(q^\tau, \sigma^\tau)(\Sigma^1_{t^\tau}[s,\mu^\tau](y)) - \nabla_\mu F(q,\sigma_t)(\Sigma^1_t[s,\mu](y))| \\ &\leqslant C|q'-q| + C(|t'-t| + \mathscr{W}(\mu,\nu)) + C(|t'-t| + \mathscr{W}(\mu,\nu)) \end{aligned}$$

and thus,

$$\nabla_\mu F(q^\tau, \sigma^\tau)(\Sigma^1_{t^\tau}[s,\mu^\tau](y)) - \nabla_\mu F(q,\sigma_t)(\Sigma^1_t[s,\mu](y))| \leqslant C(|t'-t| + |q'-q| + \mathscr{W}(\mu,\nu))$$

for some constant C. Invoking the boundedness of $\partial_t \Sigma^1$, $\nabla_q \Sigma^1$, $\bar{\nabla}_\mu \Sigma^1$ and Remark 2.1.1(i), we get

$$\Big| \int_0^1 \int_{\mathbb{T}^d \times \mathbb{T}^d} B_1 \gamma^\tau(dx,dy) d\tau \Big| \leqslant C(|t'-t| + |q'-q| + \mathscr{W}(\mu,\nu))(|t'-t| + \mathscr{W}(\mu,\nu)). \tag{6.2.7}$$

Recall that we denote by $\nabla^2_{x\mu} F(q,\mu)(x)$ the gradient at x of the mapping $x \mapsto \nabla_\mu F(q,\mu)(x)$, and $\nabla^2_{x\mu} F(q,\mu)(x)$ is uniformly bounded, by assumption. Thus,

$$\begin{aligned} |\nabla_\mu F(q,\sigma_t)(\Sigma^1_t[s,\mu](y)) - \nabla_\mu F(q,\sigma_t)(\Sigma^1_t[s,\mu](x)| &\leqslant \|\nabla^2_{x\mu} F(q,\sigma_t)\|_\infty |\Sigma^1_t[s,\mu](y) - \Sigma^1_t[s,\mu](x)| \\ &\leqslant 2\|\nabla^2_{x\mu} F(q,\sigma_t)\|_\infty A_1 |y-x|, \end{aligned}$$

for any $x, y \in \mathbb{T}^d$. Therefore, for some constant C,

$$\begin{aligned} \Big| \int_0^1 \int_{\mathbb{T}^d \times \mathbb{T}^d} B_3 \gamma^\tau(dx,dy) d\tau \Big| &\leqslant C \int_0^1 \int_{\mathbb{T}^d \times \mathbb{T}^d} |x-y| (|t'-t| + \mathscr{W}(\mu,\nu)) \gamma^\tau(dx,dy) d\tau \\ &\leqslant C|x-y|(|t'-t| + \mathscr{W}(\mu,\nu)). \end{aligned} \tag{6.2.8}$$

Next, we are going to estimate $\int_0^1 \int_{(\mathbb{T}^d)^2} B_2 \gamma^\tau d\tau$. To ease notation, let us make the abbreviations

$$\phi_2 = \Sigma^1_{t^\tau}[s,\mu^\tau], \quad \phi_1 = \Sigma^1_t[s,\mu], \quad \psi_1 = \nabla_\mu F(q,\sigma_t) \circ \phi_1.$$

Then $\int_{(\mathbb{T}^d)^2} B_2 \gamma^\tau(dx,dy)$ reads:

$$\int_{\mathbb{T}^d \times \mathbb{T}^d} \psi_1(y) \cdot \big[\partial_t \phi_2(y)(t'-t) - \nabla_q \phi_2(y) w^\tau(y) + \int_{\mathbb{T}^d} \nabla_\mu \phi_2(y)(r_2) w^\tau(r_2) \mu^\tau(dr_2)$$
$$- \partial_t \phi_1(x)(t'-t) + \nabla_q \phi_1(x) w^0(x) - \int_{\mathbb{T}^d} \nabla_\mu \phi_1(x)(r_1) w^0(r_1) \mu(dr_1)\big] \gamma^\tau(dx,dy).$$

Therefore,

$$\Big|\int\limits_{\mathbb{T}^d\times\mathbb{T}^d} B_2\gamma^\tau(dx,dy)\Big| \leqslant \|\psi_1\|_{L^\infty(\gamma^\tau)}\|D\|_{L^1(\gamma^\tau)},$$

where, by applying Remark 2.1.1(iii),

$$D = (\partial_t\phi_2(y) - \partial_t\phi_1(x))(t'-t) - (\nabla_q\phi_2(y) - \nabla_q\phi_1(x))\frac{y-x}{\tau} + \int_{\mathbb{T}^d\times\mathbb{T}^d}(\nabla_\mu\phi_2(y)(b) - \nabla_\mu\phi_1(x)(a))\frac{b-a}{\tau}\gamma^\tau(da,db),$$

and we are left with estimating the $L^1(\gamma^\tau)$-norm of D. We write

$$\tau D = \tau D + (\phi_2(y) - \phi_1(x)) + (\phi_1(x) - \phi_2(y)),$$

to apply (6.1.22), once with $\mu=\mu$, $\nu=\mu^\tau$, $t=t, t'=t^\tau$, $q=x$, $q'=y$, then with $\mu=\mu^\tau$, $\nu=\mu$, $t=t^\tau$, $t'=t$, $q=y$, $q'=x$, and obtain:

$$|\tau D| \leqslant 2C(\tau^2|t'-t|^2 + \tau^2\mathscr{W}^2(\mu,\nu) + |x-y|^2).$$

Therefore

$$\begin{aligned}\int_{\mathbb{T}^d\times\mathbb{T}^d}|D|\gamma^\tau(dx,dy) &\leqslant 2C\big(\tau|t'-t|^2 + \tau\mathscr{W}^2(\mu,\nu) + \frac{1}{\tau}\int_{\mathbb{T}^d\times\mathbb{T}^d}|x-y|^2\gamma^\tau(dx,dy)\big)\\ &\leqslant 2C\big(\tau|t'-t|^2 + \tau\mathscr{W}^2(\mu,\nu) + \frac{1}{\tau}\tau^2\mathscr{W}^2(\mu,\nu)\big).\end{aligned}$$

Consequently, for some constant C,

$$\int_0^1\int_{\mathbb{T}^d\times\mathbb{T}^d}|B_2|\gamma^\tau(dx,dy)d\tau \leqslant C(|t'-t|^2 + \mathscr{W}^2(\mu,\nu)). \tag{6.2.9}$$

Step 4. Note that all the estimates derived in the previous step, namely, (6.2.6), (6.2.7), (6.2.8), (6.2.9), are quadratic in the increments. Hence,

$$|E| \leqslant\ C(|t'-t|^2 + |q'-q|^2 + \mathscr{W}^2(\mu,\nu)),$$

which is (6.2.5), for some constant C. This concludes the proof. □

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

We collect now the results on differentiability in μ of the functions g, F, $H(q,\mathcal{V})$ that constitute the full value function u, with the following definition and corollary. Define the $\mathbb{R}^d$-valued function

$$\Upsilon[s,\mu](q,y) := \mathcal{N}_0^g[s,\mu](q)(y) + \int_0^s[\bar\nabla_\mu H(q,\mathcal{V}_t[s,\mu](q,y)) + \mathcal{N}_t^F[s,\mu](q)(y)]dt, \tag{6.2.10}$$

where $s\in[0,T]$, $q\in\mathbb{T}^d$, $y\in\mathbb{R}^d$, $\mu\in\mathscr{P}(\mathbb{T}^d)$.

Corollary 6.2.4. *The function Υ, just defined, is continuous on $[0,T] \times \mathbb{T}^d \times \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d)$, and $u(s,q,\cdot)$, defined by (6.0.1), is differentiable on $\mathscr{P}(\mathbb{T}^d)$, in the sense that there exists a constant C such that*

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

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

Proof. The continuity of Υ is a consequence of the continuity of its parts, and combining Lemma 6.1.14 with Lemma 6.2.3 produces the stated estimate. □

We refer back to Section 2.1.3 for the definition of $\mathscr{T}_\mu\mathscr{P}(\mathbb{T}^d)$. Since we do not know whether $\Upsilon[s,\mu](q,\cdot)$ belongs to the $L^2(\mu)$ closure of $\{\nabla\varphi \mid \varphi \in C_c^\infty(\mathbb{T}^d)\}$, we make the following definition.

Definition 6.2.5. *Let $u = u(s,q,\mu)$ be as in (6.0.1), for $s \in [0,T]$, $q \in \mathbb{T}^d$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, and Υ be as in (6.2.10). At every s,q,μ, by*

$$\nabla_\mu u(s,q,\mu)$$

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

We need to note that the velocity vector fields $v[s,\mu](t,\cdot)$ are not necessarily elements of $\mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d)$, even though this is true in the case $H(q,p) = \frac{1}{2}|p|^2$ (see [GŚ15, Theorem 5.1]). This leads to the following definition.

Definition 6.2.6. *At every $s,t \in [0,T]$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, by*

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

we will mean the projection of $v[s,\mu](t,\cdot)$ onto $\mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d)$, where $\sigma_t = \Sigma_t^1[s,\mu](\cdot)_\#\mu$.

Note that, if $w \in L^2(\mathbb{T}^d,\mu)$ is arbitrary and $\bar{w}$ is its projection onto $\mathscr{T}_\mu\mathscr{P}(\mathbb{T}^d)$, then

$$\int_{\mathbb{T}^d} \Upsilon[s,\mu](q,x)\cdot\bar{w}(x)\mu(dx) = \int_{\mathbb{T}^d} \nabla_\mu u(s,q,\mu)(x)\cdot w(x)\mu(dx). \tag{6.2.11}$$

This follows from (2.1.3).

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

Theorem 6.2.7. *Let H, F, g be as in Sections 2.2.1, 2.2.2, and Σ be the unique solution to the system (3.1.3) obtained in Corollary 3.1.7(ii). Let $u = u(s,q,\mu)$ be defined as in (6.0.1). Then:*

(i) For any $s \in [0,T]$, $\mu \in \mathscr{P}(\mathbb{T}^d)$, there exists $\sigma \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ such that $\sigma_s = \mu$ and the continuity equation

$$\partial_t\sigma_t + div(\nabla_p H(q,\nabla_q u(t,q,\sigma_t)\sigma_t) = 0 \quad \text{in } \mathcal{D}'((0,T)\times\mathscr{P}(\mathbb{T}^d))$$

holds;

(ii) The function u is a classical solution to the master equation

$$\begin{cases} \partial_s u(s,q,\mu) + \int_{\mathbb{T}^d} \nabla_\mu u(s,q,\mu)(x) \cdot \nabla_p H(x, \nabla_q u(s,x,\mu))\mu(dx) \\ \qquad + H(q, \nabla_q u(s,q,\mu)) + F(q,\mu) = 0 \qquad \text{in } (0,T) \times \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d), \\ \qquad\qquad u(0,q,\mu) = g(q,\mu) \qquad \text{on } \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d), \end{cases} \tag{1.3.6}$$

in the sense explained in Section 2.3.

Again: the full value function $u = u(s,q,\mu)$ is the value of the solution U of the MFG system's Hamilton-Jacobi equation (1.3.2) at the time $(t = s)$ at which the terminal condition $\sigma_{t=s} = \mu$ is prescribed for the continuity equation (1.3.3).

Proof. (i) Let $s \in [0,T]$, $\mu \in \mathscr{P}(\mathbb{T}^d)$. Set $\sigma_t := \Sigma_t^1[s,\mu]_{\#}\mu$. Then the statement follows from Proposition 3.2.10, Corollary 5.0.3, formula (5.0.9) and Lemma 6.0.1.

(ii) The regularity of u in q is the same as the regularity of U in q, which was discussed in Lemma 5.0.2. Fix $0 < s < T$, $q \in \mathbb{T}^d$, $\mu \in \mathscr{P}(\mathbb{T}^d)$. As usual, $\sigma_t = \Sigma_t^1[s,\mu]_{\#}\mu$, and $v_t = \partial_t \Sigma_t^1[s,\mu] \circ X_t[s,\mu]$, $0 \leqslant t \leqslant T$. Set

$$\hat{\sigma}_t := (id + (t-s)v_s)_{\#}\mu, \quad \bar{\sigma}_t := (id + (t-s)\bar{v}_s)_{\#}\mu,$$

where $\bar{v}_s$ is the projection of v_s to $\mathscr{T}_\mu \mathscr{P}(\mathbb{T}^d)$. Through $\sigma_{s+h} \times \hat{\sigma}_{s+h}$, we estimate $\mathscr{W}(\sigma_{s+h}, \hat{\sigma}_{s+h})$:

$$\begin{aligned} \mathscr{W}^2(\sigma_{s+h}, \hat{\sigma}_{s+h}) &\leqslant \int_{\mathbb{T}^d \times \mathbb{T}^d} |x-y|^2 (\sigma_{s+h} \times \hat{\sigma}_{s+h})(dx,dy) \\ &= \int_{\mathbb{T}^d} |\Sigma_{s+h}^1[s,\mu](y) - \Sigma_s^1[s,\mu](y) - hv_s[s,\mu](y)|^2 \mu(dy). \end{aligned}$$

Note that $v_s[s,\mu](q) = \partial_t \Sigma_t^1[s,\mu](q)|_{t=s}$, since $X_s[s,\mu] = id$. Therefore,

$$\mathscr{W}(\sigma_{s+h}, \hat{\sigma}_{s+h}) \leqslant |h|^2 \|\partial_{tt}^2 \Sigma^1\|_\infty^2. \tag{6.2.12}$$

Let

$$\gamma_h := (id \times (id + hv_s))_{\#}\mu \ \in \ \Gamma(\mu, \hat{\sigma}_{s+h}).$$

Since, by definition, $\nabla_\mu u(s+h,q,\mu) \in \mathscr{T}_\mu \mathscr{P}(\mathbb{T}^d)$, we apply Lemma 2.1.5 to write

$$|u(s+h,q,\hat{\sigma}_{s+h}) - u(s+h,q,\mu) - \int_{\mathbb{T}^d \times \mathbb{T}^d} \nabla_\mu u(s+h,q,\mu)(x) \cdot (y-x)\gamma_h(dx,dy)| = o(\|\pi^2 - \pi^1\|_{\gamma_h}),$$

which is the same as

$$|u(s+h,q,\hat{\sigma}_{s+h}) - u(s+h,q,\mu) - h \int_{\mathbb{T}^d \times \mathbb{T}^d} \nabla_\mu u(s+h,q,\mu)(x) \cdot v_s(x)\mu(dx)| = o(|h|),$$

because $o(\|\pi^2 - \pi^1\|_{\gamma_h}) = o(|h|)$ as can be easily checked. Recall now (6.2.11). Formula (6.2.10) shows that $\Upsilon[\cdot,\mu](q,y)$ is continuous, so there is a modulus of continuity ω such that

$$\begin{aligned}
&\int_{\mathbb{T}^d} \nabla_\mu u(s+h,q,\mu)(x)\cdot v_s(x)\mu(dx) = \int_{\mathbb{T}^d} \Upsilon[s+h,\mu](q,x)\cdot \bar{v}_s(x)\mu(dx)\\
&= \int_{\mathbb{T}^d} \Upsilon[s,\mu](q,x)\cdot \bar{v}_s(x)\mu(dx) + \omega(|h|)\\
&= \int_{\mathbb{T}^d} \nabla_\mu u(s,q,\mu)(x)\cdot v_s(x)\mu(dx) + \omega(|h|).
\end{aligned}$$

Therefore

$$\begin{aligned}
&|u(s+h,q,\hat{\sigma}_{s+h}) - u(s+h,q,\mu) - h\int_{\mathbb{T}^d} \nabla_\mu u(s,q,\mu)(x)\cdot v_s(x)\mu(dx)|\\
&\quad = o(|h|) + |h|\omega(|h|).
\end{aligned} \tag{6.2.13}$$

Corollary 6.2.4 shows that $u(s,q,\cdot)$ is κ_1-Lipschitz for some constant κ_1, because $[0,T]\times\mathbb{T}^d$ is compact. Using the bound (6.2.12), we then have

$$|u(s+h,q,\hat{\sigma}_{s+h}) - u(s+h,q,\sigma_{s+h})| \leqslant \kappa_1 h^2 \|\partial^2_{tt}\Sigma^1\|^2_\infty. \tag{6.2.14}$$

Invoking (6.0.6) of Lemma 6.0.1, we write

$$|u(s+h,q,\sigma_{s+h}) - u(s,q,\mu) + h\big[H(q,\nabla_q u(s,q,\sigma_s)) + F(q,\sigma_s)\big]| = o(|h|). \tag{6.2.15}$$

Finally, (6.2.13), (6.2.14) and (6.2.15) are needed to obtain:

$$\begin{aligned}
&|u(s+h,q,\mu) - u(s,q,\mu) + h\int_{\mathbb{T}^d} \nabla_\mu u(s,q,\mu)(x)\cdot v_s(x)\mu(dx) + h\big[H(q,\nabla_q u(s,q,\sigma_s)) + F(q,\sigma_s)\big]|\\
&= |u(s+h,q,\mu) - u(s+h,q,\hat{\sigma}_{s+h}) + h\int_{\mathbb{T}^d} \nabla_\mu u(s,q,\mu)(x)\cdot v_s(x)\mu(dx)\\
&\quad + u(s+h,q,\hat{\sigma}_{s+h}) - u(s+h,q,\sigma_{s+h})\\
&\quad + u(s+h,q,\sigma_{s+h}) - u(s,q,\mu) + h\big[H(q,\nabla_q u(s,q,\sigma_s)) + F(q,\sigma_s)\big]|\\
&= o(|h|) + |h|\omega(|h|) + \kappa_1 h^2\|\partial^2_{tt}\Sigma^1\|^2_\infty + o(|h|) = o(|h|).
\end{aligned}$$

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

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

Let us check the continuity of $s\mapsto \partial_s u(s,q,\mu)$. Due to (6.0.5), $\nabla_q u(s,q,\mu) = \mathcal{V}[s,\mu](s,q) = \Sigma^2[s,\mu](s,q)$, 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 (6.2.11). Let $s'\in(0,T)$. Then

$$\begin{aligned}
&|\int_{\mathbb{T}^d}\Upsilon[s,\mu](q,x)\cdot\bar{v}_s(x)\mu(dx) - \int_{\mathbb{T}^d}\Upsilon[s',\mu](q,x)\cdot\bar{v}_{s'}(x)\mu(dx)|\\
&\leqslant |\int_{\mathbb{T}^d}\Upsilon[s,\mu](q,x)\cdot\bar{v}_s(x)\mu(dx) - \int_{\mathbb{T}^d}\Upsilon[s,\mu](q,x)\cdot\bar{v}_{s'}(x)\mu(dx)|\\
&\quad + |\int_{\mathbb{T}^d}\Upsilon[s,\mu](q,x)\cdot\bar{v}_{s'}(x)\mu(dx) - \int_{\mathbb{T}^d}\Upsilon[s',\mu](q,x)\cdot\bar{v}_{s'}(x)\mu(dx)|\\
&\leqslant \|\Upsilon[s,\mu](q,\cdot)\|_{L^2(\mu)}\|\bar{v}_s - \bar{v}_{s'}\|_{L^2(\mu)} + \|\Upsilon[s,\mu](q,\cdot) - \Upsilon[s',\mu](q,\cdot)\|_{L^2(\mu)}\|\bar{v}_{s'}\|_{L^2(\mu)}.
\end{aligned}$$

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 $\|\bar{v}_s - \bar{v}_{s'}\|_{L^2(\mu)} \leqslant \|v_s - v_{s'}\|_{L^2(\mu)}$. 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. $\square$

Remark 6.2.8. We do not claim that the function $\nabla_\mu u(s, q, \cdot)$ is continuous on $\mathscr{P}(\mathbb{T}^d)$, which is true in the case [GŚ15] of $H(q, p) = \frac{1}{2}|p|^2$. The reason is that we have had to define $\nabla_\mu u$ as the projection of a vector field (Definition 6.2.5) that, in general, is not in the tangent space $\mathscr{T}_\mu \mathscr{P}(\mathbb{T}^d)$, whereas for the quadratic Hamiltonian, $\Upsilon[s, \mu](q, \cdot)$ and $\nabla_\mu u(s, q, \mu)$ are the same. //

6.3 Link with the Nash system

As we said in the Introduction, we can recover the N-player Nash system if we let $\mu = \mu^x = \frac{1}{N}\sum_{j=1}^N \delta_{x_j}$ in the master equation. We let $v^N(s, q, x) := u(t, q, \mu^x)$. Since

$$\partial_{x_j} v^N(s, q, x) = \frac{1}{N} u(s, q, \mu^x)(x_j),$$

the integral term that appears in (1.3.6) becomes

$$\sum_{j=1}^N \partial_{x_j} v^N(s, q, x) \cdot \nabla_p H(x_j, \nabla_q v^N(s, q, x_j)).$$

Then (1.3.6) becomes

$$\begin{aligned} \partial_s v^N(s, q, x) + \sum_{j=1}^N \partial_{x_j} v^N(s, q, x) \cdot \nabla_p H(x_j, \nabla_q v^N(s, q, x_j)) \\ + H(q, \nabla_q v^N(s, q, x)) + F(q, \mu^x) = 0 \qquad \text{in } (0, T) \times \mathbb{T}^d \times (\mathbb{T}^d)^N, \\ u(0, q, \mu) = g(q, \mu^x) \qquad \text{on } \mathbb{T}^d \times (\mathbb{T}^d)^N, \end{aligned}$$

which is essentially the same as (1.2.8-1.2.9).

Chapter 7

BIBLIOGRAPHY

Bibliography

[AGS08] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. 2nd. Birkhäuser, 2008.

[Aum64] R. Aumann. "Markets with a continuum of traders". In: *Econometrica* 32.1-2 (1964), pp. 39–50.

[Bes16] U. Bessi. "Existence of solutions of the master equation in the smooth case". In: *SIAM J. Math. Anal* 48.1 (2016), pp. 204–228. DOI: 10.1137/15M1018782.

[BF83] A. Bensoussan and J. Frehse. "Nonlinear elliptic systems in stochastic game theory". In: *Journal für die reine und angewandte Mathematik* 1984.350 (1983), pp. 23–67. DOI: 10.1515/crll.1984.350.23.

[BFY15] A. Bensoussan, J. Frehse, and P. Yam. "The master equation in mean field theory". In: *J. Math. Pures Appl.* 103.6 (2015), pp. 1441–1474. DOI: 10.1016/j.matpur.2014.11.005.

[BFY17] A. Bensoussan, J. Frehse, and P. Yam. "On the interpretation of the master equation". In: *Stochastic Processes and their Applications* 127.7 (2017), pp. 2093–2137. DOI: 10.1016/j.spa.2016.10.004.

[Bog07] V. Bogachev. *Measure Theory*. Vol. 2. Springer, 2007.

[Car+19] P. Cardaliaguet et al. *The Master Equation and the Convergence Problem in Mean Field Games*. Princeton University Press, 2019.

[Car04] G. Carmona. "Nash equilibria of games with a continuum of players". FEUNL working paper No. 466. 2004.

[Car12] P. Cardaliaguet. "Notes on Mean Field Games". https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf. 2012.

[CD14] R. Carmona and F. Delarue. "The master equation for large population equlibriums". In: *Stochastic Analysis and Applications, Springer Proceedings in Mathematics & Statistics*. Vol. 100. Springer, Cham, 2014. DOI: 10.1007/978-3-319-11292-3_4.

[CD18] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications*. Vol. I, II. Springer, 2018.

[CG15] P. Cardaliaguet and P. J. Graber. "Mean field games systems of first order". In: *ESAIM: COCV* 21.3 (2015), pp. 690–722. DOI: 10.1051/cocv/2014044.

[CPT15] P. Cardaliaguet, A. Porretta, and D. Tonon. "Sobolev regularity for the first order Hamilton-Jacobi equation". In: *Calc. Var.* 54.3 (2015), pp. 3037–3065. DOI: 10.1007/s00526-015-0893-3.

[Dac08] B. Dacorogna. *Direct Methods in the Calculus of Variations*. 2nd ed. Springer, 2008.

[Eva10] L. C. Evans. *Partial Differential Equations*. 2nd. AMS, 2010.

[FG95] I. Fonseca and W.Gangbo Gangbo. *Degree Theory in Analysis and its Applications*. Oxford University Press, 1995.

[FM93] W. Fleming and H. Mete Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer, 1993.

[Fri75] A. Friedman. *Stochastic Differential Equations and Applications*. Academic Press, 1975.

[Gan18] W. Gangbo. "On some analytical aspects of Mean Field Games". https://math.berkeley.edu/~wgangbo/math_278/MFG-1-18.pdf. 2018.

[GC17] W. Gangbo and Y. T. Chow. "A partial Laplacian as an infinitesimal generator on the Wasserstein space". arXiv:1710.10536v1. 2017.

[Gho17] N. Ghoussoub. "Optimal ballistic transport and Hopf-Lax formulae on Wasserstein space". arXiv:1705.05951. 2017.

[GM18] P. J. Graber and A. Mészáros. "Sobolev regularity for first order mean field games". In: *Ann. I. H. Poincaré* 14.1 (2018), pp. 1557–1576. DOI: 10.1016/j.anihpc.2018.01.002.

[GPV16] D. Gomes, E. Pimentel, and V. Voskanyan. *Regularity Theory for Mean-Field Game Systems*. Springer, 2016.

[GŚ15] W. Gangbo and A. Święch. "Existence of a solution to an equation arising from the theory of mean field games". In: *Journal of Differential Equations* 259.11 (2015), pp. 6573–6643. DOI: 10.1016/j.jde.2015.08.001.

[GT14] W. Gangbo and A. Tudorascu. "Weak KAM theory on the Wasserstein torus with multidimensional underlying space". In: *Journal of Differential Equations* 67.3 (2014), pp. 408–463. DOI: doi.org/10.1002/cpa.21492.

[GT18] W. Gangbo and A. Tudorascu. "On differentiability in the Wasserstein space and well-posedness for Hamilton-Jacobi Equations". In: *J. Math. Pures Appl.* (2018). In press.

[HCM07] M. Huang, P. E. Caines, and R. P. Malhamé. "Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ-Nash equilibria". In: *IEEE Trans. Automat. Control* 52.9 (2007), pp. 1560–1571. DOI: 10.1109/TAC.2007.904450.

[LL07] J-M. Lasry and P. L. Lions. "Mean field games". In: *Jpn. J. Math.* 2.1 (2007), pp. 229–260. DOI: 10.1007/s11537-007-0657-8.

[LS17] H. Lavenant and F. Santambrogio. "Optimal density evolution with congestion: L^∞ bounds via flow interchange techniques and applications to variational mean field game". arXiv:1705.05658v1. 2017.

[May18] S. Mayorga. “Short time solution to the master equation of a first order mean field game system”. arXiv:1811.08964v2. 2018.

[Rud76] W. Rudin. *Principles of Mathematical Analysis*. 3rd. Mc-Graw Hill, 1976.

[San15] F. Santambrogio. *Optimal Transport for Applied Mathematicians*. Birkhäuser, 2015. DOI: 10.1007/978-3-319-20828-2.

[San18] F. Santambrogio. “Regularity via duality in calculus of variations and degenerate elliptic PDEs”. In: *J. Math. Anal. Appl* 457.2 (2018), pp. 1649–1674. DOI: 10.1016/j.jmaa.2017.01.030.