Università degli Studi di Padova

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Corso di Laurea Magistrale in Matematica

TESI DI LAUREA MAGISTRALE

The PDEs system of first order mean field games

Candidato:
Giovanni Fusco
Matricola 1190306

Relatore:
Martino Bardi

Anno Accademico 2018-2019
27 Settembre 2019

Contents

Introduction

Mean field games is a theory that has been concurrently developed around 2006 by the French mathematicians J.M. Lasry and P.L. Lions and by a research group in Canada led by M. Huang, P.E. Caines and R.P. Malhamé with the aim of analyzing differential games with a very large number N of players (see for example [26, 24]). The main assumption is that the agents are very similar to each other that is, the influence of any player on the overall system is very little. This is in strong analogy with the mean field models in mathematical physics which analyses the behavior of many identical particles. In the applications we can find for instance this situation in the financial markets.

If we followed the differential games theory we should consider a system of N Hamilton-Jacobi-Bellman coupled differential equations. The resolution of that is very hard also from the numerical point of view, due to the high number of equations. So the idea is to find a simplified system of PDEs which describes the overall trend and which is the limit for $N \to \infty$ of the previous one in the following sense: Nash equilibria of the game with N players converge for $N \to \infty$ to the mean field equilibrium. This fact has been proved under suitable hypotheses in the stochastic case or in the deterministic one for open loop controls (see for example [19]), but remains an open problem for feedback controls.

In order to have a complete point of view and better understand the differential case, in the first chapter we investigate classical static games with many symmetric players. More precisely we suppose that all players have the same compact set of strategy Q, and that the cost F_i^N of the player i satisfies $\forall (x_1, \dots, x_N) \in Q^N$:

$$F^N_{\sigma(i)}(x_{\sigma(1)}, \dots, x_{\sigma(N)}) = F_i^N(x_1, \dots, x_N) \qquad \forall \sigma \text{ permutation on } \{1, \dots, N\}.$$

First of all we study limits of symmetric functions, then under suitable assumptions we analyze the behavior of Nash equilibria in pure and mixed strategies when N goes to ∞. In particular the main result is that if there exists a continuous mapping $F : Q \times \mathcal{P}(Q) \to \mathbb{R}^d$ such that $F_i^N(x_1, \ldots, x_N) = F(x_i, \frac{1}{N-1}\sum_{j\neq i}\delta_{x_j})$, $\forall i = 1, \ldots, N$ and $\forall (x_1, \ldots, x_N) \in Q^N$, then the symmetric equilibrium $\bar{\pi}^N$ in the mixed strategies up to a subsequence weakly-* converges to $\bar{m} \in \mathcal{P}(Q)$ that satisfies the *mean field equation*:

$$\int_Q F(y, \bar{m})\ \mathrm{d}\bar{m}(y) = \inf_{m\in\mathcal{P}(Q)} \int_Q F(y, \bar{m})\ \mathrm{d}m(y).$$

This fact is equivalent to saying that the support of $\bar{m}$ in contained in the set of minima of $F(\cdot, \bar{m})$. Furthermore we give sufficient conditions for uniqueness of the aftermentioned $\bar{m}$.

In the second chapter we study the *first order mean field game equations*:

$$\begin{cases} -\partial_t u(x,t) + \frac{1}{2}|D_x u(x,t)|^2 = F(x, m(t)) & \text{in } \mathbb{R}^d \times (0,T) \\ u(x,T) = G(x, m(T)) & \\ \partial_t m(x,t) - \mathrm{div}_x(D_x u(x,t) m(x,t)) = 0 & \text{in } \mathbb{R}^d \times (0,T) \\ m(x,0) = m_0(x) & \end{cases} \tag{1}$$

We suppose that every player lies in $\mathbb{R}^d$ with the dynamics $x(s) = x + \int_t^s \alpha(\tau)\mathrm{d}\tau$ and that he can control his velocity $\alpha(s)$ to minimize his cost

$$J(x,t;\alpha) := \int_t^T \left(\frac{1}{2}L(\alpha(s)) + F(x(s), m(s))\right) \mathrm{d}s + G(x(T), m(T)),$$

where we suppose that L is strictly convex, so that without loss of generality we can assume $L(\alpha) = \frac{|\alpha|^2}{2}$. According to the dynamic programming theory we can associate to this control problem a Hamilton-Jacobi-Bellman equation, which is the first one in (1); there $u(x,t)$ has to be intended as the value function, that is the minimizer of the cost functional over the controls $\alpha \in L^2([t,T], \mathbb{R}^d)$.

On the other hand the second equation in (1) is a continuity equation which involves $m(x,t)$, the density of the Borel probability measure $m(t)$ on $\mathbb{R}^d$ which describes the distribution of the other players at the time t: it is the only knowledge of a typical agent of the overall system at each time. In fact, if all agents apply an optimal strategy, which is given in the feedback form by $\alpha(x,t) = -D_x u(x,t)$, the density of their distribution over the space will evolve in time with the Fokker-Plank equation in (1).

The main result of the second chapter is that under suitable hypotheses on F, G and m_0 there is at least one solution to (1), that is a pair $(u,m) \in W^{1,\infty}_{loc}(\mathbb{R}^d \times (0,T)) \times L^1(\mathbb{R}^d \times (0,T))$ such that the Hamilton-Jacobi-Bellman equation is satisfied in the viscosity sense while the Fokker-Plank equation is satisfied in the sense of distribution. In particular we suppose:

- F, G are continuous functions over $\mathbb{R}^d \times \mathcal{P}_1$, where $\mathcal{P}_1$ is the set of Borel probability measures on $\mathbb{R}^d$ with finite first order moment, endowed with the Kantorovitch-Rubinstein distance;
- There exists a constant $C > 0$ such that $\|F(\cdot,m)\|_{\mathcal{C}^2} \leq C$ and $\|G(\cdot,m)\|_{\mathcal{C}^2} \leq C$ for any $m \in \mathcal{P}_1$, where $\mathcal{C}^2$ is the space of function with continuous second order derivatives endowed with the norm

$$\|f\|_{\mathcal{C}^2} = \sup_{x\in\mathbb{R}^d} \left[|f(x)| + |D_x f(x)| + |D^2_{xx} f(x)| \right];$$

- m_0 is absolutely continuous with respect to the Lebesgue measure, with a density still denoted m_0 which is bounded and has a compact support.

In this situation the semi-concavity of the value function play a key role, for this reason the first section is devoted to the analysis of the properties of this class of functions following the monograph [7]. Other fundamental preliminaries concern the existence and the properties of a minimizer in the problem of Calculus of variations

$$\inf_{\alpha(\cdot)\in L^p([t,T];\mathbb{R}^d)} J(t,x;\alpha(\cdot))$$
$$\text{subjet at: } x(s) = x + \int_t^s \alpha(\tau)\ d\tau.$$

For this reason in the second section of the chapter we study these arguments. Then in the next two sections we analyze separately the Hamilton-Jacobi-Bellman equation and the continuity equation. Finally we apply the Schauder-Tychonoff fixed point Theorem to prove the existence of solutions of (1). In conclusion of the chapter, we prove uniqueness of solutions under monotony assumptions on the costs F and G.

Finally the third chapter is devoted to some examples in the linear quadratic setting. In particular we want to discuss the following cases:

1. We set $G(x, m(T)) = \frac{\tilde{a}}{2}|x - h|^2 + \frac{\tilde{b}}{2}|x - \mathbb{E}[m(T)]|^2$, $F(x, m(t)) = 0$. In this situation $\tilde{a}, \tilde{b} \in \mathbb{R}$ and $h \in \mathbb{R}^d$ are given and we consider the classic dynamics $y(s) = x + \int_t^s \alpha(r)\ \mathrm{d}r$.

2. In the second case we change the dynamics which becomes $\dot{y} = Ay + B\alpha$, where A, B are given matrices.

3. Finally we generalize the previous point by adding a linear quadratic current cost such as $F(x, m(t)) = x'Mx + \mathbb{E}[m(t)]'N\mathbb{E}[m(t)]$, where $M, N \in \mathbb{R}^{d\times d}$ are symmetric given matrices.

In these simplified models the meaning of the parameters is the following: if $\tilde{a} > 0$ (respectively $\tilde{b} > 0$) the population tends to aggregate around h (respectively $\mathbb{E}[m(T)]$) at the final time; whereas as the opposite occurs if $\tilde{a} < 0$ (respectively $\tilde{b} < 0$).

We explicitly solve the first model and the second one when $d = 1$. In the other cases we give sufficient conditions on the parameters to ensure the existence, at least locally, of a mean field solution.

Chapter 1

Static games with a large number of players

The goal of this chapter is to study one-shot game with a large number N of symmetric players.

Definition 1 (Symmetric games)**.** A game is symmetric if the set of strategy Q is the same for each player and the following holds:

$$F^N_{\sigma(i)}(x_{\sigma(1)}, \dots, x_{\sigma(N)}) = F^N_i(x_1, \dots, x_N) \qquad \forall \sigma \text{ permutation on } \{1, \dots, N\},$$

where $F^N_i = F^N_i(x_1, \dots, x_N)$ is the cost of the player i when every player j chooses the strategy $x_j \in Q$.

The interpretation of this definition is that the costs are invariant under permutations of $1, \dots, N$. In particular, if we consider the point of view of a typical agent i, its cost does not change if the same choices are made by different other players.

In the rest of the chapter first of all we will study limits of symmetric functions, then Nash equilibria for symmetric games, before in pure and finally in mixed strategies.

1.1 Limits of symmetric functions

Definition 2 (Symmetric function)**.** Let Q a compact metric space. A function $u_N : Q^N \to \mathbb{R}$ is symmetric if

$$u_N(x_1, \dots, x_N) = u_N(x_{\sigma(1)}, \dots, x_{\sigma(N)}) \qquad \forall \sigma \text{ permutation on } \{1, \dots, N\}.$$

Now we consider a sequence $(u_N)_{N\in\mathbb{N}}$ of symmetric functions and we want to define a limit for it when $N \to \infty$.

Notations. To reach our goal we have to introduce some notations.

- We recall that a modulus of continuity is a non decreasing function $\omega : [0, +\infty[\to [0, +\infty[$ such that $\lim_{r\to 0^+} \omega(r) = 0$.
- We set
$$\|u_N\|_{L^\infty(Q)} := \sup_{q\in Q^N} u_N(q).$$
- Given $X = (x_1, \dots, x_N) \in Q^N$ we define
$$m_X^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}.$$
- Let $\mathcal{P}(Q) = \{m : m \text{ Borel probability measure on } Q\}$ endowed with the topology of weak-* convergence: $(m_N)_{N\in\mathbb{N}} \subset \mathcal{P}(Q)$ weakly-* converges to $m \in \mathcal{P}(Q)$ (in symbols $m_N \rightharpoonup^* m$) if
$$\lim_N \int_Q \varphi(x)\ \mathrm{d}m_N(x) = \int_Q \varphi(x)\ \mathrm{d}m(x) \qquad \forall \varphi \in \mathcal{C}^0(Q),$$
where $\mathcal{C}^0(Q) = \{\varphi : Q \to \mathbb{R} \mid \varphi \text{ continuous}\}$. In other words we say that $(m_N)_{N\in\mathbb{N}} \subset \mathcal{P}(Q)$ weakly-* converges to $m \in \mathcal{P}(Q)$ if and only if
$$\lim_N \mathbb{E}_{m_N}[\varphi(X)] = \mathbb{E}_m[\varphi(X)] \qquad \forall \varphi \in \mathcal{C}^0(Q).$$

Remark 1.1. By the compactness of Q descends that also $\mathcal{P}(Q)$ is a compact metric space for the weak-* topology. In particular it can be metrized with the

Kantorowich-Rubinstein distance:

$$\mathbf{d}_1(\mu,\nu) := \sup\left\{\int_Q f \, \mathrm{d}(\mu-\nu) \mid f : Q \to \mathbb{R} \text{ Lipschitz function with } Lip(f) \leq 1.\right\}$$

This distance is very important for the optimal transport theory, in fact it is studied in detail in [30]. Furthermore we will find it again in the second chapter in an equivalent different formulation.

We are now ready to state and prove the main theorem of this section, which is a sort of Ascoli-Arzelá Theorem for symmetric functions.

Theorem 1.2. *If $(u_N)_{N\in\mathbb{N}}$ is a sequence of symmetric uniformly bounded and uniformly continuous functions, that is:*

- *There exists some $C > 0$ such that*

$$\|u_N\|_{L^\infty(Q)} \leq C \qquad \forall N \in \mathbb{N}. \tag{1.1}$$

- *There exists a modulus of continuity ω independent on N such that*

$$|u_N(X) - u_N(Y)| \leq \omega(\boldsymbol{d}_1(m_X^N, m_Y^N)), \qquad \forall X, Y \in Q^N \;, \forall N \in \mathbb{N}. \tag{1.2}$$

Then there exists a subsequence $(u_{N_k})_{k\in\mathbb{N}}$ of $(u_N)_{N\in\mathbb{N}}$ and a continuous map $U : \mathcal{P}(Q) \to \mathbb{R}$ such that

$$\lim_{k\to\infty} \sup_{X\in\mathbb{Q}^{N_k}} \left|u_{N_k}(X) - U\left(m_X^{N_k}\right)\right| = 0.$$

Proof. Without loss of generality we can assume that the modulus of continuity is concave; in fact if ω is not concave we can consider its concave envelope, which stays above ω by construction. By the concavity of ω we can deduce:

$$\omega(x+y) - \omega(x) \leq \omega(y) \qquad \forall x, y \in [0, +\infty[. \tag{1.3}$$

Then we define the sequence $U^N : \mathcal{P}(Q) \to \mathbb{R}$ as follows:

$$U^N(m) := \inf_{X\in\mathbb{Q}^N} \{u_N(X) + \omega(\mathbf{d}_1(m_X^N, m))\}.$$

By (1.2) we have that $u_N(X) \leq u_N(Y)+\omega(\mathbf{d}_1(m_X^N, m_Y^N))$, $\forall Y \in Q^N$. Therefore the inequality holds also taking the infimum of the second term, so that we obtain

$$U^N(m_X^N) = \inf_{Y\in Q^N}\{u_N(Y) + \omega(\mathbf{d}_1(m_X^N, m_Y^N))\} = u_N(X). \tag{1.4}$$

We want to apply the Ascoli-Arzelá Theorem on the sequence $(U^N)_{N\in\mathbb{N}}$, so we have to show that it is uniformly bounded and equicontinuous.

- $(U^N)_{N\in\mathbb{N}}$ uniformly bounded.

  Since $\mathcal{P}(Q)$ is a compact metric space there exists a constant $D > 0$ such that

  $$\mathbf{d}_1(m, \tilde{m}) \leq D \qquad \forall m, \tilde{m} \in \mathcal{P}(Q). \tag{1.5}$$

  Let $q \in Q$ and $X_q^N := \underbrace{(q, \ldots, q)}_{N \text{ times}}$, then $m^N_{X_q^N} = \delta_q$. Therefore:

  $$|U^N(m)| \leq u_N(X_q^N) + \omega(\mathbf{d}_1(\delta_q, m)) \overset{(1.1)+(1.5)}{\leq} C + \omega(D).$$

- $(U^N)_{N\in\mathbb{N}}$ equicontinuous, that is they have the same modulus of continuity on $\mathcal{P}(Q)$.

  Let $m_1, m_2 \in \mathcal{P}(Q)$ and $X \in Q^N$ ε- optimal in the definition of $U^N(m_2)$, that is:

  $$u_N(X) + \omega(\mathbf{d}_1(m_X^N, m_2)) \leq U^N(m_2) + \varepsilon. \tag{1.6}$$

  Moreover by the axioms of distance, we have that

  $$\mathbf{d}_1(m_X^N, m_1) \leq \mathbf{d}_1(m_X^N, m_2) + \mathbf{d}_1(m_1, m_2),$$

  and the monotony of ω implies:

  $$\omega(\mathbf{d}_1(m_X^N, m_1)) \leq \omega\left(\mathbf{d}_1(m_X^N, m_2) + \mathbf{d}_1(m_1, m_2)\right). \tag{1.7}$$

So we can compute:

$$\begin{aligned} U^N(m_1) &\le u_N(X) + \omega\left(\mathbf{d}_1(m_X^N, m_1)\right) \overset{(1.6)+(1.7)}{\le} \\ &\le U^N(m_2) + \varepsilon + \omega\left(\mathbf{d}_1(m_X^N, m_2) + \mathbf{d}_1(m_1, m_2)\right) - \omega(\mathbf{d}_1(m_X^N, m_2)) \\ &\overset{(1.3)}{\le} U^N(m_2) + \omega(\mathbf{d}_1(m_1, m_2)) + \varepsilon. \end{aligned}$$

By letting $\varepsilon \to 0$ we obtain the claim.

So, by Ascoli-Arzelá Theorem, there exists a subsequence $(U^{N_k})_{k\in\mathbb{N}}$ of $(U^N)_{N\in\mathbb{N}}$ and a map $U : \mathcal{P}(Q) \to \mathbb{R}$ such that $U^{N_k} \rightrightarrows U$ when $k \to \infty$, that is:

$$\lim_k \sup_{m\in\mathcal{P}(Q)} \left|U^{N_k}(m) - U(m)\right| = 0.$$

In particular, being $m_X^N \in \mathcal{P}(Q)$ for each $N \in \mathbb{N}$ and for each $X \in Q^N$ we have:

$$0 = \lim_k \sup_{X\in Q^{N_k}} \left|U^{N_k}\left(m_X^{N_k}\right) - U\left(m_X^{N_k}\right)\right| \overset{(1.4)}{=} \lim_k \sup_{X\in Q^{N_k}} \left|u_{N_k}(X) - U\left(m_X^{N_k}\right)\right|.$$

□

Example. If Q is a compact subset of $\mathbb{R}^d$ and the $(u_N)_{N\in N}$ are differentiable functions, (1.2) is verified if the following Lipschitz condition holds $\forall N \in \mathbb{N}$:

$$\sup_{i=1,\ldots,N} \|D_{x_i}u\|_\infty \le C, \qquad \exists C > 0.$$

In fact it is sufficient to apply the Lagrange Mean Value Theorem and to take $\omega(r) = Cr$.

Remark 1.3. Let us observe that the limit function maintains a dependence on $X \in Q^N$ only through the measure of the empirical average measure of the vector X.

1.2 Nash equilibria in pure strategies

Let us come back to the analysis of classical static games with a large number N of symmetric player as in Definition 1. In addition we suppose that the set of strategies Q is a compact metric space with $\mathcal{P}(Q)$ as in the previous section.

Let us fix the point of view of a typical player i and consider the sequence $(F_i^N)_{N\in\mathbb{N}}$ of his costs as the number of players N changes: by our assumptions this is a sequence of symmetric functions. So if conditions (1.1) and (1.2) hold by Theorem 1.2 we know that the $(F_i^N)_{N\in\mathbb{N}}$ have a limit for $N \to \infty$, which depends both on i and on the empirical average measure of the choices of the other players.

For this reason, since N is very large, we assume from now on that there exists a continuous map $F : Q \times \mathcal{P}(Q) \to \mathbb{R}$ such that, $\forall i \in \{1, \dots, N\}$:

$$F_i^N(x_1, \dots, x_N) = F\left(x_i, \frac{1}{N-1}\sum_{j\neq i}\delta_{x_j}\right) \qquad \forall (x_1, \dots, x_N) \in Q^N. \tag{1.8}$$

Definition 3 (Nash equilibria in pure strategies)**.** We say that $(\bar{x}_1^N, \dots, \bar{x}_N^N) \in Q^N$ is a Nash equilibrium for the game $(F_1^N, \dots, F_N^N)$ if:

$$F_i^N(\bar{x}_1^N, \dots, \bar{x}_N^N) \leq F_i^N(\bar{x}_1^N, \dots, \bar{x}_{i-1}^N, y_i, \bar{x}_{i+1}^N, \dots, \bar{x}_N^N) \qquad \forall y_i \in Q.$$

Note that the symmetry assumption states that if the above definition hold for a typical i, then it holds for any $i \in \{1, \dots, N\}$.

Notations. Let us set $X^N := (\bar{x}_1^N, \dots, \bar{x}_N^N)$ the Nash equilibrium for the game with N symmetric players, and $\bar{m}_N := \frac{1}{N}\sum_{i=1}^N \delta_{\bar{x}_i^N}$ the empirical average measure linked to it. Moreover we indicate with $\mathrm{Spt}\bar{m}$ the support of the measure $\bar{m}$.

We are now ready to explore the *Mean field Theorem* for classical pure symmetric games.

Theorem 1.4. *Let X^N be a Nash equilibrium for the game $(F_1^N, \dots, F_N^N)$ for each $N \in \mathbb{N}$. Then, up to a subsequence, the sequence of measures $(\bar{m}^N)_{N\in\mathbb{N}}$ weakly-* converges to a measure $\bar{m} \in \mathcal{P}(Q)$ such that*

$$\int_Q F(y, \bar{m})\ d\bar{m}(y) = \inf_{m\in\mathcal{P}(Q)} \int_Q F(y, \bar{m})\ dm(y), \tag{1.9}$$

where F is such in (1.8).

Proof. $(\bar{m}^N)_{N\in\mathbb{N}}$ is a sequence in $\mathcal{P}(Q)$ endowed with the weak-* topology. So, by construction, we can extract from $(\bar{m}^N)_{N\in\mathbb{N}}$ a convergent subsequence, still denoted by $(\bar{m}^N)_{N\in\mathbb{N}}$ to simplify the notation. Let us call $\bar{m}$ its limit, we want to show that $\bar{m}$ satisfies (1.9).

We claim that the measure $\delta_{\bar{x}_i^N}$ realizes a minimum of the problem:

$$\inf_{m\in\mathcal{P}(Q)} \int_Q F\left(y, \frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}\right)\,\mathrm{d}m(y). \tag{1.10}$$

In fact when $m = \delta_{\bar{x}_i^N}$ the integral is reduced to $F\left(\bar{x}_i^N, \frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}\right)$ and, by Definition 3 of Nash equilibrium we have that

$$F\left(y, \frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}\right) \geq F\left(\bar{x}_i^N, \frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}\right) \qquad \forall y\in Q.$$

And this implies the claim. Indeed, $\forall m\in\mathcal{P}(Q)$:

$$\begin{aligned}
\int_Q F\left(y, \frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}\right)\,\mathrm{d}m(y) &\geq \int_Q F\left(\bar{x}_i^N, \frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}\right)\,\mathrm{d}m(y)\\
&= F\left(\bar{x}_i^N, \frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}\right)\int_Q \mathrm{d}m(y)\\
&= F\left(\bar{x}_i^N, \frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}\right)\\
&= \int_Q F\left(y, \frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}\right)\,\mathrm{d}\delta_{\bar{x}_i^N}(y).
\end{aligned}$$

Then we note that:

$$\begin{aligned}
\sup_{x\in Q}\left|\bar{m}_N(x) - \frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}(x)\right| &= \sup_{x\in Q}\left|\frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}(x) - \frac{1}{N}\sum_{j}\delta_{\bar{x}_j^N}(x)\right|\\
&\leq \sup_{x\in Q}\frac{1}{N}\left|\delta_{\bar{x}_i^N}\right| + \sup_{x\in Q}\left|\frac{1}{N-1}\sum_{j\neq i}\delta_{\bar{x}_j^N}(x) - \frac{1}{N}\sum_{j\neq i}\delta_{\bar{x}_j^N}(x)\right|\\
&= \frac{1}{N} + \sum_{j\neq i}\left(\frac{1}{N-1} - \frac{1}{N}\right)\\
&= \frac{1}{N} + (N-1)\frac{1}{N(N-1)} = \frac{2}{N} \xrightarrow[N\to\infty]{} 0.
\end{aligned}$$

For this reason and by the continuity of F, we have that $\delta_{\bar{x}_i^N}$ is ε-optimal for the probem (1.10). That is, for N large enough:

$$\begin{aligned} F\left(\bar{x}_i^N, \bar{m}^N\right) &= \int_Q F\left(y, \bar{m}\right)\ \mathrm{d}\delta_{\bar{x}_i^N}(y) \\ &\leq \inf_{m \in \mathcal{P}(Q)} \int_Q F\left(y, \bar{m}^N\right)\ \mathrm{d}m(y) + \varepsilon. \end{aligned} \tag{1.11}$$

Since it holds $\forall i \in \{1, \dots, N\}$, remembering the definition of $\bar{m}$, by linearity we have:

$$\begin{aligned} \int_Q F\left(y, \bar{m}^N\right)\ \mathrm{d}\bar{m}^N(y) &= \frac{1}{N}\left(F\left(\bar{x}_1^N, \bar{m}\right) + \cdots + F\left(\bar{x}_N^N, \bar{m}\right)\right) \\ &\overset{(1.11)}{\leq} \frac{1}{N}\left[N\left(\inf_{m \in \mathcal{P}(Q)} \int_Q F\left(y, \bar{m}^N\right)\ \mathrm{d}m(y) + \varepsilon\right)\right] \\ &= \inf_{m \in \mathcal{P}(Q)} \int_Q F\left(y, \bar{m}^N\right)\ \mathrm{d}m(y) + \varepsilon. \end{aligned}$$

By letting $N \to \infty$ we obtain the thesis. □

Remark 1.5. In the hypotheses of the Theorem 1.4 the existence of a Nash equilibrium is required for each $N \in \mathbb{N}$. This is a very strong assumption, since in general it is not true. Conversely, we will see in the next section that this thing always happens in mixed strategies.

Proposition 1.6. *The static mean field equation (1.9) holds if and only if the support of $\bar{m}$ is contained in the set of minima of $F(y, \bar{m})$, that is:*

$$Spt(\bar{m}) \subseteq \underset{y \in Q}{\operatorname{argmin}} F(y, \bar{m})$$

Proof. $\Longrightarrow$) If (1.9) holds, for each $x \in Q$, choosing $m = \delta_x$ we obtain:

$$\int_Q F(y, \bar{m})\ \mathrm{d}\bar{m} \leq F(x, \bar{m}) \qquad \forall x \in Q.$$

In particular

$$\int_Q F(y, \bar{m})\ \mathrm{d}\bar{m} \leq \min_{x \in Q} F(x, \bar{m}).$$

And this trivially implies that $\mathrm{Spt}(\bar{m}) \subseteq \underset{y \in Q}{\operatorname{argmin}} F(y, \bar{m})$.

$\Longleftarrow$) Conversely, if $\mathrm{Spt}(\bar{m}) \subseteq \underset{y \in Q}{\mathrm{argmin}} F(y, \bar{m})$, then

$$\int_Q F(y, \bar{m}) \,\mathrm{d}\bar{m}(y) \leq \int_Q F(y, \bar{m}) \,\mathrm{d}m(y) \qquad \forall m \in \mathcal{P}(Q),$$

and this provides the thesis. □

This proposition states how the Nash equilibria of the game tend to be arranged when $N \to \infty$. In fact take place in Q so that the support of the limit measure $\bar{m}$ is contained in the set of minima of $F(y, \bar{m})$ over Q.

1.3 Nash equilibria in mixed strategies

In this section we analyze what happens when the players are allowed to randomize their behavior by playing strategies in $\mathcal{P}(Q)$ instead of Q. That is we want to study the trend of Nash equilibria for the same game $F_1^N, \ldots, F_N^N$ in the so called *mixed strategies*.

So, if the agents play the strategy $\pi_1, \ldots, \pi_N \in \mathcal{P}(Q)^N$, the cost of the player i will be the sum of the costs of every single choice multiplied with the probability of playing it under the strategy $\pi_1, \ldots, \pi_N \in \mathcal{P}(Q)^N$. That is:

$$\begin{aligned}\bar{F}_N^i(\pi_1, \ldots, \pi_N) &= \int_{Q^N} F_N^i(x_1, \ldots, x_N) \,\mathrm{d}\pi_1(x_1) \ldots \mathrm{d}\pi_N(x_N) \\ &\overset{(1.8)}{=} \int_{Q^N} F\left(x_i, \frac{1}{N-1} \sum_{j \neq i} \delta_{x_j}\right) \,\mathrm{d}\pi_1(x_1) \ldots \mathrm{d}\pi_N(x_N).\end{aligned}$$

Definition 4 (Nash equilibria in mixed strategies)**.** We say that $\bar{\pi}_1, \ldots, \bar{\pi}_N \in \mathcal{P}(Q)^N$ is a Nash equilibrium in the mixed strategies if, for any $i \in \{1, \ldots, N\}$,

$$\bar{F}_N^i(\bar{\pi}_1, \ldots, \bar{\pi}_N) \leq \bar{F}_N^i\left((\bar{\pi}_j)_{j \neq i}, \pi_i\right) \qquad \forall \pi_i \in \mathcal{P}(Q).$$

Now we want to extend Theorem 1.4 to the mixed strategies case. To do this we need two important results. The first of these is about the existence of symmetric Nash equilibria for symmetric games in mixed strategies.

Definition 5 (Upper hemicontinuous correspondence)**.** Given two sets X, Y, a correspondence (or set valued map) $\Phi : X \rightrightarrows Y$ associates to every element

$x \in X$ a subset of Y.

A correspondence $\Phi : X \rightrightarrows Y$ is said to be upper hemicontinuous at the point x if for any open neighborhood V of $\Phi(x)$, there exists a neighborhood U of x such that $\Phi(x) \subseteq V$ for any $x \in U$.

Remark 1.7. It can be proved (see for example [2]) that if Y is compact, then Φ is hemicontinuous if and only if for any sequence $(x_n)_{n\in\mathbb{N}} \subset X$ which converges to $x \in X$ and for any sequence $(y_n)_{n\in\mathbb{N}}$ which converges to $y \in Y$ and such that $y_n \in \Phi(x_n)$ $\forall n \in \mathbb{N}$, then $y \in \Phi(x)$.

We omit the proof of this fact because it goes beyond the intent of these pages.

Theorem 1.8 (Nash equilibria in mixed strategies for symmetric games)**.** *If the game is symmetric then there is a Nash equilibrium in the mixed strategies of the form $(\bar{\pi}, \dots, \bar{\pi})$, where $\bar{\pi} \in \mathcal{P}(Q)$.*

Proof. It is a straightforward application of the Fan's fixed point Theorem (see [18]).

Theorem (Fan's Theorem)**.** *Let X be a non empty, compact and convex subset of a locally convex topological vector space. Let $\Phi : X \rightrightarrows X$ any upper hemicontinuous set valued map such that $\Phi(x) \neq \emptyset$, compact and convex $\forall x \in X$. Then $\exists \bar{x} \in X$ such that $\bar{x} \in \Phi(\bar{x})$.*

We define the best response correspondence $\mathcal{R} : \mathcal{P}(Q) \rightrightarrows \mathcal{P}(Q)$ as follows:

$$\mathcal{R}(\pi) = \left\{ \sigma \in \mathcal{P}(Q) \,:\, \bar{F}_1(\sigma, \pi, \dots, \pi) = \min_{\tilde{\sigma} \in X} \bar{F}_1(\tilde{\sigma}, \pi, \dots, \pi) \right\}.$$

Let us verificate the hypotheses of the Fan's Theorem.

- First of all we check that $\bar{F}_1$ is a continuous map. Since $F : \mathcal{P}(Q) \to \mathbb{R}$ is a continuous function on $\mathcal{P}(Q)$ compact then by Weierstrass Theorem it is bounded, that is there exists $D > 0$ such that $\|F\|_\infty \leq D < \infty$. Then:

$$\begin{aligned}
\left|\bar{F}_1\left((\pi_i)_{i=1,\dots,N}\right) - \bar{F}_1\left((\sigma_j)_{j=1,\dots,N}\right)\right| &\leq \int_{Q^N} |F| \left(\prod_{i=1}^{N} \mathrm{d}\pi_i - \prod_{i=1}^{N} \mathrm{d}\sigma_i\right) \\
&\leq D \int_{Q^N} \left(\prod_{i=1}^{N} \mathrm{d}\pi_i - \prod_{i=1}^{N} \mathrm{d}\sigma_i\right) \\
&\leq D \ \mathbf{d}_1\left(\prod_{i=1}^{N} \pi_i, \prod_{i=1}^{N} \sigma_i\right).
\end{aligned}$$

- $\mathcal{P}(Q)$ is clearly non empty, compact because Q it is, and convex because every set of probability measure it is.

- Since $\mathcal{P}(Q)$ is compact we can use te characterization of upper hemicontinuity of Remark 1.7. So let $(\pi_n) \subset \mathcal{P}(Q)$ convergent to $\pi \in \mathcal{P}(Q)$ and $(\sigma_n) \subset \mathcal{P}(Q)$ such that $\sigma_n \in \mathcal{R}(x_n)$ for any $n \in \mathbb{N}$. This means that

$$\bar{F}_1(\sigma_n, \pi_n, \dots, \pi_n) \leq \bar{F}_1(\nu, \pi_n, \dots, \pi_n) \qquad \forall \nu \in \mathcal{P}(Q),\ \forall n \in \mathbb{N}. \quad (1.12)$$

  If $\sigma \in \mathcal{P}(Q)$ is the limit of (σ_n), then using the continuity of $\bar{F}_1$ we can pass to the limit in (1.12). So we obtain:

$$\bar{F}_1(\sigma, \pi, \dots, \pi) \leq \bar{F}_1(\nu, \pi, \dots, \pi) \qquad \forall \nu \in \mathcal{P}(Q).$$

  And this implies that $\sigma \in \mathcal{R}(\pi)$.

- $\mathcal{R}(\pi)$ is non empty for any $\pi \in \mathcal{P}(Q)$ because by Weierstrass Theorem there exist at least a minimum of the continuous map $\bar{F}_1(\cdot, \pi, \dots, \pi)$ on the compact set $\mathcal{P}(Q)$.

- $\mathcal{R}(\pi)$ is compact for any $\pi \in \mathcal{P}(Q)$. In fact $\mathcal{R}(\pi) \subset \mathcal{P}(Q)$ is a closed set because it is the anti-image of the closed singleton $\{\min_{\sigma \in \mathcal{P}(Q)} F(\sigma, \pi, \dots, \pi)\}$ through the continuous map $\bar{F}_1(\cdot, \pi, \dots, \pi)$. And a closed subset of a compact set is compact.

- Finally $\mathcal{R}(\pi)$ is convex. In fact given $\sigma, \tilde{\sigma} \in \mathcal{R}(\pi)$ and $\lambda \in (0, 1)$, by the linearity of $\bar{F}_1$ we have

$$\begin{aligned}
\bar{F}_1(\lambda\sigma + (1-\lambda)\tilde{\sigma}; \pi, \dots, \pi) &= \lambda \bar{F}_1(\sigma; \pi, \dots, \pi) + (1-\lambda)\bar{F}_1(\tilde{\sigma}; \pi, \dots, \pi) \\
&\overset{\sigma,\tilde{\sigma}\in\mathcal{R}(\pi)}{=} \lambda \min_{\mu\in\mathcal{P}(Q)} \bar{F}_1(\mu; \pi, \dots, \pi) + (1-\lambda) \min_{\mu\in\mathcal{P}(Q)} \bar{F}_1(\mu; \pi, \dots, \pi) \\
&= \min_{\mu\in\mathcal{P}(Q)} \bar{F}_1(\mu; \pi, \dots, \pi).
\end{aligned}$$

This means that $\lambda\sigma + (1-\lambda)\tilde{\sigma} \in \mathcal{R}(\pi)$.

So by Fan's Theorem there exists $\bar{\pi}$ fixed point for $\mathcal{R}$. By construction $\bar{\pi}$ satisfies

$$\bar{F}_1(\bar{\pi}, \dots, \bar{\pi}) \leq \bar{F}_1(\sigma, \bar{\pi}, \dots, \bar{\pi}) \qquad \forall \sigma \in \mathcal{P}(Q).$$

We can conclude by the symmetry assumption. □

The second result we need is about the properties of sequences of symmetric probability measures.

Definition 6. Given a compact set Q, a measure μ on Q^k is said to be symmetric if

$$\pi_\sigma \# \mu = \mu \qquad \forall \sigma \text{ permutation of } \{1, \dots, k\},$$

where $\pi_\sigma(x_1, \dots, x_k) = (x_{\sigma(1)}, \dots, x_{\sigma(k)})$ and $\pi_\sigma \# \mu = \mu(\pi_\sigma^{-1})$.

Therefore μ is symmetric if

$$\mu(A) = \mu\left(\pi_\sigma^{-1}(A)\right) \qquad \forall A \in Q^k.$$

Remark 1.9. Let us observe that Definition 6, if m is a symmetric measure on Q^k then, for any permutation σ of $1, \dots, k$:

$$m(\mathrm{d}y_1, \dots, \mathrm{d}y_k) = m(\mathrm{d}y_{\sigma(1)}, \dots, \mathrm{d}y_{\sigma(k)}).$$

Theorem 1.10 (Hewitt-Savage Theorem)**.** *Let $(m_n)_{n\in\mathbb{N}}$ be a sequence of symmetric probability measure on Q^n such that m_n is the marginal of m_{n+1} with respect the last variable x_{n+1}, that is:*

$$\int_Q dm_{n+1}(x_{n+1}) = m_n \qquad \forall n \in \mathbb{N}. \tag{1.13}$$

Then there is a probability measure μ on $\mathcal{P}(Q)$ such that for any continuous map $f \in \mathcal{C}^0(\mathcal{P}(Q))$:

$$\lim_n \int_{Q^n} f\left(\frac{1}{n}\sum_{i=1}^n \delta_{x_i}\right)\ dm_n(x_1,\dots,x_n) = \int_{\mathcal{P}(Q)} f(m)\ d\mu(m). \tag{1.14}$$

Furthermore $\forall n \in \mathbb{N}$ and for any n-uple $A_1,\dots,A_n$ of Borel sets of Q it holds

$$m_n(A_1 \times \cdots \times A_n) = \int_{\mathcal{P}(Q)} m(A_1)\dots m(A_n)\ d\mu(m). \tag{1.15}$$

Proof. First of all we note that iterating the (1.13) we obtain:

$$\int_{Q^{n-j}} \mathrm{d}m_n(x_{j+1},\dots,x_n) = m_j. \tag{1.16}$$

Let us consider the following sequence of linear and continuous functionals $(L_n)_{n\geq 1}$ of $\mathcal{C}^0(\mathcal{P}(Q))$:

$$L_n(P) := \int_{Q^n} P\left(\frac{1}{n}\sum_{i=1}^n \delta_{y_i}\right) m_n(\mathrm{d}y_1,\dots,\mathrm{d}y_n), \qquad \forall P \in \mathcal{C}^0(\mathcal{P}(Q)).$$

We want to show that it has limit for $n \to \infty$, so we prove that it has limit for any $P \in \mathcal{C}^0(\mathcal{P}(Q))$. Indeed we establish this fact only for P of the form

$$P(m) = \int_{Q^j} \varphi(x_1,\dots,x_j)\ \mathrm{d}m(x_1)\dots\mathrm{d}m(x_j), \tag{1.17}$$

where $\varphi : Q^j \to \mathbb{R}$ is a continuous mapping. In fact this class of functions contains the monomials

$$P(m) = \prod_{i=1}^k \int_Q \psi_i(x)\ \mathrm{d}m(x),$$

and the polynomials generated by them are dense in $\mathcal{C}^0(\mathcal{P}(Q))$ for the Stone-Weierstrass Theorem (see [9, p. 39]).

By (1.17) we have that for any $n \geq j$

$$P\left(\frac{1}{n}\sum_{i=1}^{n}\delta_{y_i}\right) = \frac{1}{n^j}\sum_{(i_1,\ldots,i_j)}\varphi(y_{i_1},\ldots,y_{i_j}),$$

where $(i_1,\ldots,i_j) \in \{1,\ldots,n\}^j$. Therefore:

$$L_n(P) = \frac{1}{n^j}\sum_{i_1,\ldots,i_j}\int_{Q^n}\varphi(y_{i_1},\ldots,y_{i_j})\ m_n(\mathrm{d}y_1,\ldots,\mathrm{d}y_n).$$

If $i_1,\ldots,i_j$ are all different, we can fix the variables $(y_{i_1},\ldots,y_{i_j})$ and integrate on the remaining. Let us denote Q^j with the space of the variables $y_{i_1},\ldots,y_{i_j}$ and let $\sigma = \sigma_{i_1,\ldots,i_j}$ the permutation of $\{1,\ldots,j\}$ that puts in ascending order $i_1,\ldots,i_j$, then

$$\begin{aligned}\int_{Q^n}\varphi(y_{i_1},\ldots,y_{i_j})\ m_n(\mathrm{d}y_1,\ldots,\mathrm{d}y_n) &= \int_{Q^j}\varphi(y_{i_1},\ldots,y_{i_j})\ \int_{Q^{n-j}} m_n(\mathrm{d}y_i,\ldots,\mathrm{d}y_n)\\ &\overset{(1.16)}{=}\int_{Q^j}\varphi(y_{i_1},\ldots,y_{i_j})\ m_j(\mathrm{d}y_{\sigma(i_1)},\ldots,\mathrm{d}y_{\sigma(i_j)})\end{aligned}$$

$$\begin{aligned}&= \int_{Q^j}\varphi(y_{\sigma(i_1)},\ldots,y_{\sigma(i_j)})\ m_j(\mathrm{d}y_{\sigma^2(i_1)},\ldots,\mathrm{d}y_{\sigma^2(i_j)})\\ &= \int_{Q^j(y_1,\ldots,y_j)}\varphi(y_1,\ldots,y_j)\ \mathrm{d}m_j(x_1,\ldots,x_j),\end{aligned} \tag{1.18}$$

where $x_k = y_{\sigma^2_{i_1,\ldots,i_j}(k)}$. In the third step we have ordered the arguments of φ through the change of variables given by $\sigma_{i_1,\ldots,i_j}$, while the last equality holds because $Q^j(y_1,\ldots,y_j) \cong Q^j(y_{i_1},\ldots,y_{i_j})$. Finally we have to observe that the last integral does not depend on the particular permutation $\sigma_{i_1,\ldots,i_j}$ because of the Remark 1.9.

Now, using the Stirling approximation,

$$|\{(i_1,\ldots,i_j) : i_1,\ldots,i_j \text{ distinct}\}| = \frac{n!}{(n-j)!} \underset{n\to\infty}{\longrightarrow} n^j, \tag{1.19}$$

and combining (1.18) and (1.19) we obtain:

$$\lim_{n\to\infty} L_n(P) = \int_{Q^j}\varphi(y_1,\ldots,y_j)\ \mathrm{d}m_j(x_1,\ldots,x_j).$$

So L_n has a limit $L \in (\mathcal{C}^0(\mathcal{P}(Q)))^*$. Since $\mathcal{P}(Q)$ is compact we can apply the Riesz representation Theorem (see [29, Theorem 2.14]), so there exists a unique $\mu \in \mathcal{P}(Q)$ such that:

$$L(P) = \int_{\mathcal{P}(Q)} P(m)\ \mathrm{d}\mu(m).$$

It remains to show that μ satisfies (1.15). Let P such as in (1.17), then:

$$\begin{aligned} L(P) &= \int_{\mathcal{P}(Q)} P(m)\ \mathrm{d}\mu(m) \\ &= \int_{\mathcal{P}(Q)} \left(\int_{Q^j} \varphi(y_1, \ldots, y_j)\ \mathrm{d}m(x_1) \ldots \mathrm{d}m(x_j) \right)\ \mathrm{d}\mu(m). \end{aligned} \tag{1.20}$$

Let now $A_1, \ldots, A_j \subset Q$ closed. We can find a non-increasing subsequence $(\varphi_k)_{k \in \mathbb{N}}$ of continuous functions on Q^j which converges to $\mathbf{1}_{A_1}(y_1) \cdot \cdots \cdot \mathbf{1}_{A_j}(x_j)$. By this fact and (1.20) we have obtained (1.15) for any $A_1, \ldots, A_j \subset Q$, and therefore for any Borel measurable subset $A_1, \ldots, A_j$ of Q. □

Example. Let us retrace the salient steps of the previous proof through an explicit example. Let us suppose that $n = 4$, $j = 3$, $(i_1, i_2, i_3) = (4, 1, 3)$, and $\sigma = \sigma_{4,1,3}$ the permutation that changes $(4, 1, 3)$ in $(1, 3, 4)$. Then:

$$\begin{aligned} \int_{Q^4} \varphi(y_4, y_1, y_3)\ \mathrm{d}m_4(y_1, \ldots, y_4) &= \int_{Q^3(y_4, y_1, y_3)} \varphi(y_4, y_1, y_3) \int_{Q(y_2)} \mathrm{d}m_4(y_1, \ldots, y_4) \\ &= \int_{Q^3} \varphi(y_4, y_1, y_3) \underbrace{\mathrm{d}m_4(y_1, y_3, y_4)}_{= \mathrm{d}m_4(y_{\sigma(4)}, y_{\sigma(1)}, y_{\sigma(3)})} \\ &= \int_{Q^3} \varphi(y_1, y_3, y_4) \underbrace{\mathrm{d}m_4(y_3, y_4, y_1)}_{= \mathrm{d}m_4(y_{\sigma^2(4)}, y_{\sigma^2(1)}, y_{\sigma^2(3)})} \\ &= \int_{Q^3(y_1, y_2, y_3)} \varphi(y_1, y_2, y_3) \underbrace{\mathrm{d}m_4(y_2, y_3, y_1)}_{= \mathrm{d}m_4(y_{\sigma^2(3)}, y_{\sigma^2(2)}, y_{\sigma^2(1)})}. \end{aligned}$$

The last integral does not depend on the particular σ for the symmetry of m_4.

Corollary 1.11. *If $m_0 \in \mathcal{P}(Q)$ and $m_n = \prod_{i=1}^n m_0$ then*

$$\lim_{n \to \infty} \int_{Q^j} f\left(\frac{1}{n} \sum_{i^1}^{n} \delta_{x_i} \right)\ dm_n(x_1, \ldots, x_n) = f(m_0). \tag{1.21}$$

Proof. $(m_n)_{n\in\mathbb{N}}$ satisfies the hypotheses of Hewitt-Savage Theorem, so we have:

$$m_0(A_1)\cdot\dots\cdot m_0(A_n) = m_n(A_1\times\dots\times A_n) \stackrel{(1.15)}{=} \int_{\mathcal{P}(Q)} m(A_1)\cdot\dots\cdot m(A_n)\;\mathrm{d}\mu(m)$$

So we immediately deduce that $\mu=\delta_{m_0}$ and therefore:

$$\lim_{n\to\infty}\int_{Q^j} f\left(\frac{1}{n}\sum_{i^1}^{n}\delta_{x_i}\right)\;\mathrm{d}m_n(x_1,\dots,x_n) \stackrel{(1.10)}{=} \int_{\mathcal{P}(Q)} f(m)\;\mathrm{d}\delta_{m_0}(m) = f(m_0)$$

□

We are now ready to enunciate the main Theorem of this section, on the limit of Nash equilibria in the mixed strategies for symmetric games.

Theorem 1.12. *Let $(\bar{\pi}^N,\dots,\bar{\pi}^N)$ the Nash equilibrium in the mixed strategies of the symmetric game $(\bar{F}_1^N,\dots,\bar{F}_N^N)$, for any $N\geq 2$. Then, up to a subsequence, $(\bar{\pi}^N)$ weakly-* converges to a measure $\bar{m}\in\mathcal{P}(Q)$ which satisfies the static mean field equation (1.9). In particular there is always a solution to the static mean field equation.*

Proof. To simplify the notation we will write Λ instead of $\frac{1}{N-1}\sum_{j\neq i}\delta_{x_j}$.

$(\bar{\pi}^N)_{N\in\mathbb{N}}$ is a sequence on $\mathcal{P}(Q)$, so by construction it has a convergent subsequence to some $\bar{m}\in\mathcal{P}(Q)$, still denoted by $(\bar{\pi}^N)_{N\in\mathbb{N}}$.

Since $F:\mathcal{P}(Q)\to\mathbb{R}$ is a continuous function on $\mathcal{P}(Q)$ compact then by Weierstrass Theorem it is bounded, that is there exists $D>0$ such that $\|F\|_\infty\leq D<\infty$. Therefore, for any $y\in Q$:

$$\begin{aligned}
&\left|\int_{Q^{N-1}} F(y;\Lambda)\prod_{j\neq i}\mathrm{d}\bar{\pi}^N(x_j) - \int_{Q^{N-1}} F(y;\Lambda)\prod_{j\neq i}\mathrm{d}\bar{m}(x_j)\right| \\
&\leq \int_{Q^{N-1}}\left|F(y;\Lambda)\right|\left(\prod_{j\neq i}\mathrm{d}\bar{\pi}^N(x_j) - \prod_{j\neq i}\mathrm{d}\bar{m}(x_j)\right) \qquad (1.22)\\
&\leq D\;\mathbf{d}_1\left(\prod_{j\neq i}\mathrm{d}\bar{\pi}^N(x_j),\prod_{j\neq i}\mathrm{d}\bar{m}(x_j)\right)\underset{N\to\infty}{\longrightarrow} 0.
\end{aligned}$$

In fact the distance is a continuous mapping and $\prod_{j\neq i}\mathrm{d}\bar{\pi}^N(x_j)\rightharpoonup^*\prod_{j\neq i}\mathrm{d}\bar{m}(x_j)$.

Now, by definition of Nash equilibrium we have $\forall m \in \mathcal{P}(Q)$:

$$\int_Q \left(\int_{Q^{N-1}} F(x_i, \Lambda) \prod_{j \neq i} \mathrm{d}\bar{\pi}^N(x_j) \right) \, \mathrm{d}\bar{\pi}^N(x_i) \leq \int_Q \left(\int_{Q^{N-1}} F(x_i, \Lambda) \prod_{j \neq i} \mathrm{d}\bar{\pi}^N(x_j) \right) \, \mathrm{d}m(x_i)$$

On the left we add and remove $\int_Q \left(\int_{Q^{N-1}} F(x_i, \Lambda) \prod_{j \neq i} \mathrm{d}\bar{m}(x_j) \right) \, \mathrm{d}\bar{\pi}^N(x_i)$, while on the right $\int_Q \left(\int_{Q^{N-1}} F(x_i, \Lambda) \prod_{j \neq i} \mathrm{d}\bar{m}(x_j) \right) \, \mathrm{d}m(x_i)$. So we obtain:

$$\begin{aligned}
&\int_Q \left[\int_{Q^{N-1}} F(x_i, \Lambda) \left(\prod_{j \neq i} \mathrm{d}\bar{\pi}^N(x_j) - \prod_{j \neq i} \mathrm{d}\bar{m}(x_j) \right) \right] \, \mathrm{d}\bar{\pi}^N(x_i) + \\
&+ \int_Q \left(\int_{Q^{N-1}} F(x_i, \Lambda) \prod_{j \neq i} \mathrm{d}\bar{m}(x_j) \right) \, \mathrm{d}\bar{\pi}^N(x_i) \leq \\
&\leq \int_Q \left[\int_{Q^{N-1}} F(x_i, \Lambda) \left(\prod_{j \neq i} \mathrm{d}\bar{\pi}^N(x_j) - \prod_{j \neq i} \mathrm{d}\bar{m}(x_j) \right) \right] \, \mathrm{d}m(x_i) + \\
&\qquad + \int_Q \left(\int_{Q^{N-1}} F(x_i, \Lambda) \prod_{j \neq i} \mathrm{d}\bar{m}(x_j) \right) \, \mathrm{d}m(x_i)
\end{aligned}$$

Now we pass to the limit for $N \to \infty$, and we can do this because we are combining strong and weak-* convergence. In fact, by (1.22) and $\bar{\pi}^N \rightharpoonup^* \bar{m}$, both the first term on the left and the first term on the right of the $\leq$ go to 0. Moreover, by (1.21):

$$\lim_{N \to \infty} \int_{Q^{N-1}} F(x_i, \Lambda) \prod_{j \neq i} \mathrm{d}\bar{m}(x_j) = F(x_i, \bar{m})$$

where this convergence is uniform with respect to x_i because of the continuity of F. Therefore:

$$\int_Q F(x_i, \bar{m}) \, \mathrm{d}\bar{m}(x_i) \leq \int_Q F(x_i, \bar{m}) \, \mathrm{d}m(x_i).$$

□

Now we want to give a sufficient condition for the uniqueness of this $\bar{m}$.

Theorem 1.13. *Assume that F satisfies $\forall m_1 \neq m_2 \in \mathcal{P}(Q)$:*

$$\int_Q (F(y, m_1) - F(y, m_2)) \ d(m_1 - m_2)(y) > 0. \tag{1.23}$$

Then there is a unique measure satisfying (1.9).

Proof. Let $\bar{m}_1, \bar{m}_2$ satisfying the static mean field equation (1.9). Since both of them achieve the infimum we have:

$$\begin{cases} \int_Q F(y, \bar{m}_1) \ \mathrm{d}\bar{m}_1(y) \leq \int_Q F(y, \bar{m}_1) \ \mathrm{d}\bar{m}_2(y) \\ \int_Q F(y, \bar{m}_2) \ \mathrm{d}\bar{m}_2(y) \leq \int_Q F(y, \bar{m}_2) \ \mathrm{d}\bar{m}_1(y) \end{cases}$$

By adding these two inequalities we get

$$\int_Q F(y, \bar{m}_1) \ \mathrm{d}\bar{m}_1(y) + \int_Q F(y, \bar{m}_2) \ \mathrm{d}\bar{m}_2(y) \leq \int_Q F(y, \bar{m}_1) \ \mathrm{d}\bar{m}_2(y) + \int_Q F(y, \bar{m}_2) \ \mathrm{d}\bar{m}_1(y)$$

$$\Updownarrow$$

$$\int_Q [F(y, \bar{m}_1) - F(y, \bar{m}_2)] \ \mathrm{d}\bar{m}_1(y) - \int_Q [F(y, \bar{m}_1) - F(y, \bar{m}_2)] \ \mathrm{d}\bar{m}_2(y) \leq 0$$

$$\Updownarrow$$

$$\int_Q [F(y, \bar{m}_1) - F(y, \bar{m}_2)] \ \mathrm{d}(\bar{m}_1 - \bar{m}_2)(y) \leq 0.$$

By (1.23) it must be $\bar{m}_1 = \bar{m}_2$. $\square$

We will discuss in detail in the next chapter the monotony condition (1.23) of the cost function F. Now we conclude this chapter with an example.

Example (Potential games)**.** Let us suppose that there exists $\Phi : Q \times \mathcal{P}(Q) \to \mathbb{R}$ such that

$$\frac{\partial}{\partial m} \Phi(x, m)\Big|_{m=\bar{m}} = F(x, \bar{m}). \tag{1.24}$$

Moreover we assume that $\exists \bar{m} \in \mathcal{P}(Q)$ such that $\bar{m} = \underset{m \in \mathcal{P}(Q)}{\operatorname{argmin}} \int_Q \Phi(x, m) \ \mathrm{d}x$. Then, $\forall m \in \mathcal{P}(Q)$:

$$\int_Q \frac{\partial}{\partial m} \Phi(x, m)\Big|_{m=\bar{m}} (m - \bar{m}) \ \mathrm{d}x \geq 0 \overset{(1.24)}{\iff} \int_Q F(x, \bar{m}) \ \mathrm{d}m \geq \int_Q F(x, \bar{m}) \ \mathrm{d}\bar{m}. \tag{1.25}$$

This shows that $\bar{m}$ satisfies (1.9).

Let $\mathcal{P}_{ac}(Q) = \{m \in \mathcal{P}(Q) : m$ absolutely continuous with respect to Lebesgue measure$\}$, $V : Q \to \mathbb{R}$ continuous, $G : (0, +\infty) \to \mathbb{R}$ strictly increasing, continuous and such that $G(0) = 0$ and $G(s) \geq 2cs$ for some $c > 0$. For instance let us assume that

$$F(x, m) = \begin{cases} V(x) + G(m(x)) & \text{if } m \in \mathcal{P}_{ac}(Q) \\ +\infty & \text{otherwise.} \end{cases}$$

So if H is the primitive of G with $H(0) = 0$, then if $m \in \mathcal{P}_{ac}(Q)$:

$$\Phi(x, m) = V(x)m + H(m(x)).$$

By $G(s) \geq 2cs$ we deduce $H(s) \geq cs^2$, with $\ddot{H} = \dot{G} > 0$. Then:

$$\frac{\partial^2}{\partial m \partial m} \int_Q (Vm + H) \, \mathrm{d}x = \int_Q \frac{\partial^2 H}{\partial m \partial m} \, \mathrm{d}x > 0.$$

This implies that the problem $\inf_{m \in \mathcal{P}(Q)} \int_Q \Phi(x, m) \, \mathrm{d}x$ has a unique solution $\bar{m} \in L^2(Q)$. Then, for any $m \in \mathcal{P}_{ac}(Q)$, by (1.25) we obtain:

$$\int_Q [V(x) + G(\bar{m}(x))] \, \mathrm{d}m(x) \geq \int_Q [V(x) + G(\bar{m}(x))] \, \mathrm{d}\bar{m}(x).$$

Therefore $\bar{m}$ satisfies the static mean field equation (1.9). In particular, by Proposition 1.6, for any $x \in \mathrm{Spt}(\bar{m})$:

$$V(x) + G(\bar{m}(x)) = \min_{y \in Q} V(y) + G(\bar{m}(y)).$$

Chapter 2

The mean field first order equations

In this chapter our aim is to study the existence of solutions to the *Mean field first order equations*, which we abbreviate with *MFE's*:

$$\begin{cases} -\partial_t u(x,t) + \frac{1}{2}|D_x u(x,t)|^2 = F(x, m(t)) & \text{in } \mathbb{R}^d \times (0,T) \\ u(x,T) = G(x, m(T)) \\ \partial_t m(x,t) - \operatorname{div}_x(D_x u(x,t) m(x,t)) = 0 & \text{in } \mathbb{R}^d \times (0,T) \\ m(x,0) = m_0(x). \end{cases} \tag{2.1}$$

Let us deduce heuristically this system of partial differential equations with boundary conditions.

We assume that a typical agent can control his velocity α and moves in $\mathbb{R}^d$ with the dynamics

$$\begin{cases} \dot{x}(s) = \alpha(s) =: f(x, \alpha) & s \in]t, T] \\ x(t) = x, \end{cases}$$

where $t \in [0,T]$ and $x \in \mathbb{R}^d$ are given. Moreover we suppose that he wants to minimize his cost functional

$$J(t, x; \alpha) := \int_t^T \underbrace{\left(\frac{1}{2}|\alpha(s)|^2 + F(x(s), m(s))\right)}_{l(x,\alpha,m)} \mathrm{d}s + G(x(T), m(T)),$$

where $m(s)$ is the distribution at the time s of the other agents over $\mathbb{R}^d$; it is his only knowledge of the overall world, and we denote with $m(x,s)$ its density.

By the optimal control theory we know that under suitable hypotheses the value function

$$u(t,x) := \inf_{\alpha \in L^2([t,T])} J(t,x;\alpha)$$

solves in viscosity sense the following backward Hamilton-Jacobi-Belman equation:

$$\begin{cases} -\partial_t u(x,t) - \inf_a \mathcal{H}(D_x u, x, a) = 0 & \text{in } \mathbb{R}^d \times (0,T) \\ u(x,T) = G(x, m(T)), \end{cases}$$

where $\mathcal{H}(p,x,a) := p \cdot f(x,a) + l(x,a)$. Substituting the expressions of f and l and minimizing $\mathcal{H}$ with respect to the variable a, we obtain the first equation of (2.1).

Furthermore by the verification theorems we guess that $\alpha = -D_x u(t,x)$ is the optimal control in the feedback form. Now, if all the agents argue in this way and apply the optimal strategy, the density $m(x,s)$ evolves over time with the Kolmogorov law in $\mathbb{R}^d \times (0,T)$:

$$\partial_t m(x,t) + \operatorname{div}(f(x,\alpha) m(x,t)) = 0.$$

Substituting the expression of f and $\alpha = -D_x u(t,x)$ we obtain the continuity equation in (2.1).

A pair $(u,m) \in W^{1,\infty}_{loc}(\mathbb{R}^d \times (0,T)) \times L^1(\mathbb{R}^d \times (0,T))$ is a solution of the MFE's if the Hamilton-Jacobi-Bellman equation is satisfied in the viscosity sense while the Fokker-Plank equation is satisfied in the sense of distribution that is, m is a weak solution. To show that (2.1) has solutions we need some notations and assumptions.

Notations. $m \in \mathcal{P}_1 :=$ {Borel probability measure on $\mathbb{R}^d$ with finite first order moment}. We endow $\mathcal{P}_1$ with the Kantorovich-Rubinstein distance

$$\mathbf{d}_1(\mu,\nu) := \inf_{\gamma \in \prod(\mu,\nu)} \left[\int_{\mathbb{R}^{2d}} |x-y| \, \mathrm{d}\gamma(x,y) \right],$$

where $\prod(\mu,\nu)$ is the set of Borel probability measures on $\mathbb{R}^{2d}$ such that for any

Borel subset A of $\mathbb{R}^d$ the following hold:

$$\gamma(A \times \mathbb{R}^d) = \mu(A) \qquad \text{and} \qquad \gamma(\mathbb{R}^d \times A) = \nu(A).$$

Finally we denote with $\mathcal{C}^2$ the space of functions with continuous second order derivatives endowed with the norm $\|\cdot\|_{\mathcal{C}^2}$:

$$\|f\|_{C^2} = \sup_{x \in \mathbb{R}^d} \left[|f(x)| + |D_x f(x)| + |D^2_{xx} f(x)| \right].$$

Remark 2.1. In the first chapter we defined the Kantorovich-Rubinstein distance in a different way. However it is possible to prove that the two formulations are equivalent (see for instance [30, Chapter 5]). It is called Kantorovich duality.

Our main hypotheses are:

1. F and G continuous over $\mathbb{R}^d \times \mathcal{P}_1$.
2. $\forall m \in \mathcal{P}_1$, $F(\cdot, m), G(\cdot, m) \in \mathcal{C}^2$; morover $\exists C > 0$ such that $\forall m \in \mathcal{P}_1$:
$$\|F(\cdot, m)\|_{\mathcal{C}^2} \leq C \qquad \text{and} \qquad \|G(\cdot, m)\|_{\mathcal{C}^2} \leq C. \tag{2.2}$$
3. m_0 is absolutely continuous with respect to the Lebesgue measure, with a density still denoted by m_0 which is bounded and has a compact support.

In order to reach our goal we need some preliminaries about properties of semiconcave functions and the existence of a minimizer of the problem:

$$\min_{\alpha \in L^p([t,T])} J(t, x; \alpha) := \int_t^T L(s, x(s), \alpha(s)) \, \mathrm{d}s + g(x(T)),$$

where $x(s) = x + \int_t^s \alpha(r) \, \mathrm{d}r$.

In the following sections we will investigate in detail all these topics, then we will analyze separately the HJB and the continuity equation, then we will conclude thanks to Schauder fixed point Theorem.

2.1 Semiconcave functions

In this section we analyze the property of the class of semiconcave functions. Most of statements and proofs are taken by the monograph [7].

Definition 7 (Semiconcave function with linear modulus)**.** Let $A \subset \mathbb{R}^d$. We say that $u : A \to \mathbb{R}$ is a semiconcave function with linear modulus if $\exists C > 0$ such that

$$\lambda u(y_1) + (1-\lambda)u(y_2) - u(\lambda y_1 + (1-\lambda)y_2) \leq \lambda(1-\lambda)\frac{C}{2}|y_1 - y_2|^2 \tag{2.3}$$

for any $y_1, y_2 \in S$ such that $[y_1, y_2] \subset A$ and for any $\lambda \in (0,1)$.

Let us see other equivalent formulations of this definition.

Proposition 2.2. *Given $A \subset \mathbb{R}^d$ open, $u : A \to \mathbb{R}$ and $C > 0$, the following are equivalent:*

i. (2.3) is satisfied.

ii. The function $y \mapsto u(y) - \frac{C}{2}|y|^2$ is concave in every convex subset of A.

iii. $u \in \mathcal{C}(A)$ and satisfies

$$u(y+h) + u(y-h) - 2u(y) \leq C|h|^2 \tag{2.4}$$

for any $y, h \in \mathbb{R}^d$ such that $[y-h, y+h] \subset A$.

iv. $\forall \nu \in \mathbb{R}^d$ such that $|\nu| = 1$ we have $\partial^2_{\nu\nu} u \leq C$ in A in the sense of distributions, that is

$$\int_A u(x)\partial^2_{\nu\nu}\varphi(x)\ dx \leq C\int_A \varphi(x)\ dx \qquad \forall \varphi \in \mathcal{C}^\infty_c(A), \varphi \geq 0,$$

where by ∂_ν we intend the directional derivative.

Proof.

- *i. ⇒ iii.*) Let $y_1 = y + h$, $y_2 = y - h$ and $\lambda = \frac{1}{2}$. Substituting in (2.3) we obtain the thesis.

- *i.* $\iff$ *ii.*) Let y_1, y_2 and λ such in (2.3). Then

$$\begin{aligned}
&\lambda|y_1|^2 + (1-\lambda)|y_2|^2 - |\lambda y_1 + (1-\lambda)y_2|^2 \\
&= (\lambda - \lambda^2)|y_1|^2 + |y_2|^2[(1-\lambda) - (1-\lambda)^2] - 2\lambda(1-\lambda)\langle y_1, y_2\rangle \\
&= \lambda(1-\lambda)[|y_1|^2 + |y_2|^2 - 2\langle y_1, y_2\rangle] = \lambda(1-\lambda)|y_1 - y_2|^2.
\end{aligned}$$

  Therefore:

$$\begin{aligned}
&u(\lambda y_1 + (1-\lambda)y_2) - \frac{C}{2}|\lambda y_1 + (1-\lambda)y_2|^2 \\
&\qquad \geq \lambda u(y_1) + (1-\lambda)u(y_2) - \frac{C\lambda}{2}|y_1|^2 - \frac{C(1-\lambda)}{2}|y_2|^2
\end{aligned}$$

$$\Updownarrow$$

$$\begin{aligned}
&u(\lambda y_1 + (1-\lambda y_2)) - \lambda u(y_1) - (1-\lambda)u(y_2) \\
&\qquad \geq -\frac{C}{2}\underbrace{[\lambda|y_1|^2 + (1-\lambda)|y_2|^2 - |\lambda y_1 + (1-\lambda)y_2|^2]}_{\lambda(1-\lambda)|y_1-y_2|^2}.
\end{aligned}$$

- *iii.* $\Rightarrow$ *ii.*) Let us consider the function $v(y) = u(y)\frac{C}{2}|y|^2$, which is continuous. Let $y \in A$ and h as in the hypothesis, then:

$$\begin{aligned}
&v(y+h) + v(y-h) - 2v(y) \\
&= u(y+h) - \frac{C}{2}|y+h|^2 + u(y-h) - \frac{C}{2}|y-h|^2 - 2u(y) + C|y|^2 \\
&\overset{(2.4)}{\leq} C|h|^2 - \frac{C}{2}|y+h|^2 - \frac{C}{2}|y-h|^2 + C|y|^2 = 0.
\end{aligned}$$

  So v is concave.

- *ii.* $\iff$ *iv.*) Let us observe that $\partial^2_{\nu\nu}|y|^2 \equiv 2$ for any $\nu \in \mathbb{R}^d$ such that $|\nu| = 1$. Then u satisfies *iv.* if and only if $v(y) = u(y) - \frac{C}{2}|y|^2$ satisfies $\partial^2_{\nu\nu}v \leq 0$ in the sense of distributions. But this last inequality is satisfied if and only if v is concave.

□

Our main assumption on F and G is that they belong to $\mathcal{C}^2$ and that their norm $\|\cdot\|_{\mathcal{C}^2}$ is finite. Let us show that this class of function is semiconcave.

Proposition 2.3. *If $u \in \mathcal{C}^2$ and there exists $C > 0$ such that $\|u\|_{\mathcal{C}^2} \leq C$, then u is semiconcave with linear modulus of constant C.*

Proof. We use the characterization (2.4.iii). Using twice the mean value theorem we have that for some $x^h \in (x, x+h)$, $x_h \in (x-h, x)$ and $\bar{x} \in (x_h, x^h)$:

$$\begin{aligned}\frac{|u(x+h)+u(x-h)-2u(x)|}{|h|^2} &= \left|\frac{u(x+h)-u(x)}{|h|^2} - \frac{u(x)-u(x-h)}{|h|^2}\right| \\ &= \frac{|\langle D_x u(x^h), h\rangle - \langle D_x u(x_h), h\rangle|}{|h|^2} \\ &\leq \frac{|D_x u(x^h) - D_x u(x_h)|\ |h|}{|h|^2} \\ &= \frac{|D^2_{xx}(\bar{x})|\ |x^h - x_h|}{|h|} \overset{|x^h - x_h| \leq |h|}{\leq} C.\end{aligned}$$

□

Now we extend the previous definition of semiconcavity and we introduce the concept of generalized gradients in order to deduce more general properties of this class of functions.

Definition 8 (Semiconcave function)**.** $u : S \subset \mathbb{R}^d \to \mathbb{R}$ is semiconcave if there exists a non-decreasing upper semicontinuous function $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{r\to 0^+} \omega(r) = 0$ and such that:

$$\lambda u(x) + (1-\lambda)u(y) - u(\lambda x + (1-\lambda)y) \leq \lambda(1-\lambda)|x-y|\omega(|x-y|), \quad (2.5)$$

for any $x, y \in S$ such that $[x, y] \subset S$, $\lambda \in (0, 1)$.

Remark 2.4. Semiconcave functions with linear modulus are semiconcave functions with $\omega(r) = \frac{C}{2} r$.

Definition 9 (Sub/Super-differential)**.** For any $x \in A \subset \mathbb{R}^d$ open we define

$$D^+u(x) := \left\{ p \in \mathbb{R}^d : \limsup_{y\to x} \frac{u(y) - u(x) - \langle p, y - x\rangle}{|y - x|} \leq 0 \right\}$$

$$D^-u(x) := \left\{ p \in \mathbb{R}^d : \liminf_{y\to x} \frac{u(y) - u(x) - \langle p, y - x\rangle}{|y - x|} \geq 0 \right\}.$$

They are called respectively the Frechet superdifferential and subdifferential of u at x.

Definition 10 (Dini derivatives)**.** Let $x \in A$ and $\theta \in \mathbb{R}^d \setminus \{0\}$, we define

$$\partial^+ u(x,\theta) := \limsup_{h\to 0\ \bar\theta\to\theta} \frac{u(x+h\bar\theta)-u(x)}{h}$$

$$\partial^- u(x,\theta) := \liminf_{h\to 0\ \bar\theta\to\theta} \frac{u(x+h\bar\theta)-u(x)}{h}.$$

They are called respectively the upper and lower Dini derivatives of u at x in the direction θ.

Theorem 2.5. *If $u : A \to \mathbb{R}$ is semiconcave, than it is locally Lipschitz in the Interior of A.*

Proof. Without loss of generality we can suppose that A is open.

- *Step 1.* We show that u is locally bounded from below.

  Given $x_0 \in A$, we take a closed cube centered in x_0, all contained in A, with diameter L, vertices $x_1, \dots, x_{2^d}$ and $m_0 := \min_i u(x_i)$. If x_i and x_j are consecutive vertices, using that $\lambda(1-\lambda) \le \frac{1}{4}$ if $\lambda \in (0,1)$, we have by (2.5):

$$\begin{aligned} u(\lambda x_i + (1-\lambda)x_j) &\ge \lambda u(x_i) + (1-\lambda)u(x_j) - \lambda(1-\lambda)|x_i - x_j|\omega(|x_i - x_j|) \\ &\ge m_0 - \frac{1}{4}L\omega(L). \end{aligned}$$

  So u is bounded from below on the 1-dimensional face of the cube. Let us show that it is on the 2-dimensional face too. Let x_k, x_l and x_h, x_r pair of consecutive vertices on opposite 1-dimensional faces. Then if we define $y_i := \mu x_k + (1-\mu)x_l$ and $y_j := \nu x_h + (1-\nu)x_r$, with $\mu, \nu \in (0,1)$, we have (remember $u(y_i), u(y_j) \ge m_0 - \frac{1}{4}L\omega(L)$):

$$\begin{aligned} u(\lambda y_i + (1-\lambda)y_j) &\ge \lambda u(y_i) + (1-\lambda)u(y_j) - \lambda(1-\lambda)|y_i - y_j|\omega(|y_i - y_j|) \\ &\ge m_0 - \frac{1}{2}L\omega(L). \end{aligned}$$

  Iterating this procedure one can show that u is bounded from below in all the k-dimensional faces of the cube for $k = 1, \dots, d$. So it is in the whole cube.

- *Step 2.* We show that u is locally bounded from above.

Let $x_0 \in A$ and $R > 0$ such that $\overline{B(x_0, R)} \subset A$. By *Step 1* we know that $\exists m$ such that $u \geq m$ on $\overline{B(x_0, R)}$. Let $z \in \overline{B(x_0, R)}$, now we write (2.5) with $x = x_0 - R\frac{z - x_0}{|z - x_0|}$, $y = z$ and $\lambda = \frac{|z - x_0|}{R + |z - x_0|}$ we get:

$$\frac{|z - x_0|}{R + |z - x_0|} u\left(x_0 - R\frac{z - x_0}{|z - x_0|}\right) + \frac{R}{R + |z - x_0|} u(z) - u(x_0)$$
$$\leq R\frac{|z - x_0|}{R + |z - x_0|}\omega(R + |z - x_0|).$$

Therefore:

$$u(z) \leq -\frac{|z - x_0|}{R} m + \frac{R + |z - x_0|}{R} u(x_0) + R\omega(R + |z - x_0|)$$
$$\overset{|z - x_0| \leq R}{\leq} |m| + 2|u(x_0)| + R\omega(2R).$$

- *Step 3.* Finally we show that u is locally Lipschitz.

We note that if $z \in [x, y] \subset S$, then $\exists \lambda \in (0, 1)$ such that $z = \lambda x + (1 - \lambda)y$; writing $u(z) = \lambda u(z) + (1 - \lambda)u(z)$ in (2.5) and then dividing by $\lambda(1 - \lambda)|x - y|$ we obtain:

$$\frac{u(x) - u(z)}{|x - z|} - \frac{u(z) - u(y)}{|z - y|} \leq \omega(|x - y|). \tag{2.6}$$

Let $x_0 \in A$ and $R > 0$ such that $\overline{B(x_0, R)} \subset A$; by *Step 1* and *Step 2* there exist $m, M > 0$ such that $m \leq u \leq M$ in $\overline{B(x_0, R)}$. Given $x, y \in B\left(x_0, \frac{R}{2}\right)$, let x' and y' the points at distance R from x_0 in the straight line joining x and y, such that $x \in [x', y]$ and $y \in [y', x]$. Then (2.6) implies

$$\frac{u(x') - u(x)}{|x' - x|} - \omega(|x' - y|) \leq \frac{u(x) - u(y)}{|x - y|} \leq \frac{u(y) - u(y')}{|y - y'|} - \omega(|y' - x|).$$

If we analyze the right part we note that $|y - y'| \geq \frac{R}{2}$ by construction, $u(y) \leq M$, $-u(y') \leq -m$ and

$$|y' - x| \leq |y' - x_0| + |x_0 - x| \leq R + \frac{R}{2} \leq 2R.$$

Repeating the same estimates at the left we get:

$$\frac{|u(y)-u(x)|}{|y-x|} \leq \frac{2(M-m)}{R} + \omega(2R).$$

□

Proposition 2.6. *Let $A \subset \mathbb{R}^d$ open, $u : A \subset \mathbb{R}^d \to \mathbb{R}$, $x \in A$. Then:*

i. $D^+u(x) = \{p \in R^d \,:\, \partial^+ u(x,\theta) \leq \langle p,\theta\rangle \forall\theta \in \mathbb{R}^d \setminus \{0\}\}$
$D^-u(x) = \{p \in R^d \,:\, \partial^- u(x,\theta) \geq \langle p,\theta\rangle \forall\theta \in \mathbb{R}^d \setminus \{0\}\}$

ii. $D^+u(x)$ *and* $D^-u(x)$ *are closed and convex sets.*

iii. $D^+u(x)$ *and* $D^-u(x)$ *are both non-empty if and only if* u *is differentiable at* x *and it holds*

$$D^+u(x) = D^-u(x) = \{Du(x)\}$$

Proof. i. We prove the claim only for $D^+u(x)$. The inclusion $\subseteq$ descends immediately by the definitions of superdifferential and Dini derivative.

We show the inverse inclusion by contradiction. Let $p \in \mathbb{R}^d$ such that $\partial^+ u(x,\theta) \leq \langle p,\theta\rangle \forall\theta \in \mathbb{R}^d \setminus \{0\}$, but we assume that $p \notin D^+u(x)$. Then there exist a sequence $(x_k)_{k\in\mathbb{N}} \subset A$ and $\varepsilon > 0$ such that $x_k \to x$ and

$$u(x_k) - u(x) \geq \langle p, x_k - x\rangle + \varepsilon|x_k - x|.$$

Moreover, up to a subsequence, we can assume that $\theta_k := \frac{x_k - x}{|x_k - x|} \to \theta$ unit vector. Then:

$$\varepsilon + \langle p,\theta_k\rangle \leq \frac{u(x_k)-u(x)}{|x_k - x|},$$

and therefore:

$$\limsup_k \varepsilon + \langle p,\theta_k\rangle \leq \limsup_k \frac{u(x_k)-u(x)}{|x_k - x|} = \limsup_k \frac{u(x + |x_k - x|\theta_k) - u(x)}{|x_k - x|}.$$

But the first term is equal to $\varepsilon + \langle p,\theta\rangle$, while the last is $\leq$ of $\partial^+ u(x,\theta)$. So we have obtained:

$$\varepsilon + \langle p,\theta\rangle \leq \partial^+ u(x,\theta)$$

And this is in contradiction with the definition of Dini derivative.

ii. It is a direct consequence of *i.*.

iii. If u is differentiable at x then $Du(x) \in D^+u(x) \cap D^-u(x)$.

Conversely if $p_1 \in D^+u(x)$ and $p_2 \in D^-u(x)$ then by the point *i.* we have

$$\langle p_2, \theta\rangle \leq \partial^- u(x,\theta) \leq \partial^+ u(x,\theta) \leq \langle p_1, \theta\rangle \qquad \forall \theta \in \mathbb{R}^d \setminus \{0\}.$$

This implies $\langle p_1 - p_2, \theta\rangle \geq 0$ for any $\theta \in \mathbb{R}^d \setminus \{0\}$. And this implies $p_1 - p_2 = 0$ because we can take $\theta = -(p_1 - p_2)$. So $D^+u(x)$ and $D^-u(x)$ coincide and reduce to a singleton. Furthermore, combining the definitions of super and subdifferential we get that u is differentiable at x.

□

Proposition 2.7. *Let $A \subset \mathbb{R}^d$ open, $u : A \to \mathbb{R}$ a semiconcave function with modulus ω, and let $x \in A$. Then $p \in \mathbb{R}^d$ belongs to $D^+u(x)$ if and only if*

$$u(y) - u(x) - \langle p, y - x\rangle \leq |y - x|\omega(|y - x|), \tag{2.7}$$

for any $y \in A$ such that $[x, y] \subset A$.

Proof. If $p \in \mathbb{R}^d$ satisfies (2.7) then it trivially satisfies the definition of $D^+u(x)$.
Conversely, if $p \in D^+u(x)$, then dividing (2.3) by $\lambda|y - x|$ we get

$$\frac{u(y) - u(x)}{|y - x|} \leq \frac{u(x + \lambda(y - x)) - u(x)}{\lambda|y - x|} + (1 - \lambda)\omega(|y - x|) \qquad \forall \lambda \in (0, 1).$$

By letting $\lambda \to 0$ we obtain:

$$\frac{u(y) - u(x)}{|y - x|} \leq \frac{\langle p, y - x\rangle}{|y - x|} + (1 - \lambda)\omega(|y - x|).$$

□

Proposition 2.8. *Let $u : A \to \mathbb{R}$ a semiconcave function with modulus ω and let $x, y \in A$ with $[x, y] \subset A$. Then, $\forall p \in D^+u(x)$ and $\forall q \in D^+u(y)$, we have:*

$$\langle q - p, y - x\rangle \leq 2|y - x|\omega(|y - x|)$$

In particular if u has linear modulus:

$$\langle q-p, y-x\rangle \leq C|y-x|^2. \tag{2.8}$$

Proof. By applying the previous proposition for x and y we get:

$$\begin{cases} u(y)-u(x) \leq \langle p, y-x\rangle + |y-x|\omega(|y-x|) \\ u(x)-u(y) \leq \langle q, x-y\rangle + |y-x|\omega(|y-x|) \end{cases}$$

By adding this two inequalities we get:

$$0 \leq \langle q, x-y\rangle - \langle p, x-y\rangle + 2|y-x|\omega(|y-x|)$$

$$\Updownarrow$$

$$\langle q-p, y-x\rangle \leq 2|y-x|\omega(|y-x|).$$

□

Theorem 2.9. *Let $u_n : A \to \mathbb{R}$ a family of semiconcave functions with the same modulus ω. If $B \subset A$ is open and such that the $(u_n)_{n\in\mathbb{N}}$ are uniformly bounded in B, then there exists a subsequence $(u_{n_k})_{k\in\mathbb{N}}$ which converges uniformly to a semiconcave function $u : B \to \mathbb{R}$ still with modulus ω. Moreover $Du_{n_k} \to Du$ almost everywhere in B.*

Proof. By the *Step 3* of Theorem 2.5 we know that the Lipschitz constant of u in B depends only on B, $\sup_B |u|$ and the modulus of semiconcavity ω. Since the (u_n) are uniformly bounded in B and uniformly semiconcave by hypothesis, we deduce that they are uniformly Lipschitz in B. So they are equicontinuous too, and by Ascoli-Arzelá Theorem there exists a subsequence, denoted with $(u_k)_{k\in\mathbb{N}}$, such that $u_k \rightrightarrows u$. Since the inequality (2.3) is preserved by pointwise convergence, u is semiconcave with the same modulus of the (u_k).

By Theorem 2.5 and Rademacher Theorem, u and the (u_k) are differentiable at x for almost all $x \in B$; let $x_0 \in B$ be such a point. Let us suppose by contradiction that $Du_k(x_0) \nrightarrow Du(x_0)$. But by (2.7) we have that for any $k \in \mathbb{N}$ it holds:

$$u_k(y) - u_k(x_0) - \langle Du_n(x_0), y-x_0\rangle \leq |y-x|\omega(|y-x|), \tag{2.9}$$

For any $y \in A$ such that $[x_0, y] \subset A$. However $(Du_n(x_0))_{k\in\mathbb{N}}$ is bounded because the (u_k) are Lipschitz, so it has a convergent subsequence to $p_0 \neq Du(x_0)$, and if we pass to the limit in (2.9), by Proposition 2.7 we find out that $p_0 \in D^+u(x_0)$. And this is a contradiction because by Proposition 2.6.iii we have that $D^+u(x_0) = Du(x_0) = \{p_0\}$ since u is differentiable at x_0. So we have the thesis. □

Definition 11 (Reachable gradients)**.** Let $A \subset \mathbb{R}^d$ open, $u : A \to \mathbb{R}$ be locally Lipschitz. A vector $p \in \mathbb{R}^d$ is called reachable gradient of u at $x \in A$ if there exists a sequence $(x_k)_{k\in\mathbb{N}} \subset A$ such that:

i. u is differentiable at x_k for any $k \in \mathbb{N}$.

ii. $\lim_k x_k = x$.

iii. $\lim_k Du(x_k) = p$.

We denote with $D^*u(x)$ the set of the reachable gradients of u at x.

Proposition 2.10. *Let $u : A \to \mathbb{R}$ be a semiconcave function with modulus ω, and let $x \in A$. Then:*

i. If $(x_k) \subset A$ converges to x and $(p_k) \in D^+u(x_k)$ converges to p, then $p \in D^+u(x)$.

ii. $D^*u(x) \subset \partial D^+u(x)$.

iii. $D^+u(x) \neq \emptyset$.

iv. If $D^+u(x)$ is a singleton, then u is differentiable at x.

v. If $D^+u(y)$ is a singleton $\forall y \in A$, then $u \in \mathcal{C}^1(A)$.

Proof. i. It follows by Proposition 2.7, passing to the limit.

ii. By *i.* it follows that $D^*u(x) \subset D^+u(x)$. We have to prove that all the reachable gradients are boundary points of $D^+u(x)$. Let $p \in D^*u(x)$ and $(x_k)_{k\in\mathbb{N}} \subset A$ such in Definition 11. Without loss of generality we can assume that $\lim_k \frac{x-x_k}{|x-x_k|} = \theta \in \mathbb{R}^d$ unit vector; we want to show that $p - t\theta \notin D^+u(x)$ for any $t > 0$, because this would mean that $p \in \partial D^+u(x)$. In fact by (2.7) we have:

$$u(x_k) - u(x) - \langle Du(x_k), x_k - x\rangle \geq -|x_k - x|\omega(|x_k - x|),$$

and so

$$\begin{aligned}
&u(x_k) - u(x) - \langle p - t\theta, x_k - x\rangle \\
&= u(x_k) - u(x) - \langle Du(x_k), x_k - x\rangle + \langle Du(x_k) - p, x_k - x\rangle + t\langle \theta, x_k - x\rangle \\
&\geq -|x_k - x|\omega(|x_k - x|) - |Du(x_k) - p||x_k - x| + t\langle \theta, x_k - x\rangle.
\end{aligned}$$

Therefore $p - t\theta \notin D^+u(x)$ because

$$\limsup_k \frac{u(x_k) - u(x) - \langle p - t\theta, x_k - x\rangle}{|x - x_k|} \geq t.$$

iii. It is a straightforward consequence of *ii*. In fact u is locally Lipschitz by Theorem 2.5, and so $D^*u(x) \neq \emptyset$.

iv. Let us suppose $D^+u(x) = \{p\}$ for some $p \in \mathbb{R}^d$ and let $(x_k) \subset A$ be a sequence which converges to x. By *iii.* we can take a sequence $p_k \in D^+u(x_k)$ that admits only p as cluster point by *i*. Then $p_k \to p$. Furthermore by Proposition 2.7 we have that:

$$\begin{aligned}
&u(x_k) - u(x) - \langle p, x_k - x\rangle \\
&= u(x_k) - u(x) + \langle p_k, x - x_k\rangle + \langle p_k - p, x_k - x\rangle \\
&\geq -|x_k - x|\omega(|x_k - x|) - |p_k - p||x_k - x|.
\end{aligned}$$

Therefore:

$$\liminf_k \frac{u(x_k) - u(x) - \langle p, x_k - x\rangle}{|x_k - x|} \geq 0.$$

Then $p \in D^-u(x)$, but by Proposition 2.6.iii u is differentiable at x.

v. Direct consequence of *iv*.

□

Notations. Let $u : A \to \mathbb{R}$ semiconcave. We denote the directional derivate of u at x in the direction θ with the symbol $\partial u(x, \theta)$.

Finally we denote with $\mathrm{Co}(A)$ the closed convex hull of A.

Theorem 2.11. *Let $u : A \to \mathbb{R}$ be a semiconcave function and let $x \in A$. Then $D^+u(x) = Co(D^*u(x))$. Moreover for any $\theta \in \mathbb{R}^d$ it holds*

$$\partial u(x, \theta) = \min_{p \in D^+u(x)} \langle p, \theta\rangle = \min_{p \in D^*u(x)} \langle p, \theta\rangle. \tag{2.10}$$

Proof. First we prove (2.10). By Proposition 2.6.i and Proposition 2.10.ii we deduce that for any $\theta \in \mathbb{R}^d$

$$\partial u(x,\theta) \leq \min_{p\in D^+u(x)} \langle p,\theta\rangle \leq \min_{p\in D^*u(x)} \langle p,\theta\rangle.$$

So it is sufficient to show that $\min_{p\in D^*u(x)}\langle p,\theta\rangle \leq \partial u(x,\theta)$ for any $\theta \in \mathbb{R}^d$.

Let $\theta \in \mathbb{R}^d$ unit vector. Since u is differentiable almost everywhere by Theorem 2.5 and Rademacher Theorem, we can find a sequence $(x_k)_{k\in\mathbb{N}}$ such that u is differentiable at x_k for any k, $\theta_k = \frac{x-x_k}{|x-x_k|} \to \theta$, and $Du(x_k)$ converges to some $p_0 \in D^*u(x)$. Let ω be the modulus of semiconcavity of u, then by Proposition 2.7 we have that

$$\langle Du(x_k),\theta_k\rangle \leq \frac{u(x+|x_k-x|\theta_k)-u(x)}{|x_k-x|} + \omega(|x_k-x|).$$

By letting $k\to\infty$ we get $\langle p_0,\theta\rangle \leq \partial u(x,\theta)$, from which it follows our claim.

Now, since $\min_{p\in D^*u(x)}\langle p,\theta\rangle = \min_{p\in Co(D^*u(x))}\langle p,\theta\rangle$, the (2.10) implies that the two convex sets $D^+u(x)$ and $\mathrm{Co}(D^*u(x))$ must have the same support function, and consequently they must be the same set.

□

2.2 Minimizer of cost functional in L^p

The aim of this section is to study the problem

$$\inf_{\alpha\in L^p([t,T];\mathbb{R}^d)} J(t,x;\alpha) := \int_t^T L(s,x(s),\alpha(s))\ \mathrm{d}s + g(x(T))$$

where $x(s) = x + \int_t^s \alpha(r)\mathrm{d}r$, for a fixed $T \in]0,+\infty[$. The main results are taken by [21].

We set $Q_T :=]0,T[\times\mathbb{R}^d$, so $L : Q_T \times \mathbb{R}^d \to \mathbb{R}$. Our assumptions are:

$$\begin{aligned}
&\text{(a)}\ \ L, g \geq 0. \\
&\text{(b)}\ \ L, g \in \mathcal{C}^{R+1} && \text{for some integer } R \geq 1. \\
&\text{(c)}\ \ L_{vv}(t,x,v) > 0. \\
&\text{(d)}\ \ L(t,x,v) \geq A|v|^p + B && \text{for some } p \geq 1,\ A > 0 \text{ and } B \in \mathbb{R}. \\
&\text{(e)}\ \ |L_x(t,x,v)| + |L_v(t,x,v)| \leq K(1+|v|^p) && \text{for some } K > 0.
\end{aligned} \tag{2.11}$$

Now we prove that under these hypotheses a minimizer always exists.

Theorem 2.12. *Under hypotheses (2.11) for any $(t,x) \in Q_T$, $\exists \alpha^* \in L^p([t,T];\mathbb{R}^d)$ such that:*

$$J(t,x;\alpha^*) = \min_{\alpha\in L^p([t,T];\mathbb{R}^d)} J(t,x,;\alpha).$$

Proof. To simplify the notations let $A = 1$ and $B = 0$, and let q the conjugate exponent of p (that is $\frac{1}{q} + \frac{1}{p} = 1$).

Let $(\alpha_n)_{n\in\mathbb{N}}$ a minimizing sequence in $L^p([t,T];\mathbb{R}^d)$, that is:

$$\lim_n J(t,x;\alpha_n) = \inf_{\alpha\in L^p} J(t,x;\alpha) := V(t,x).$$

By (2.11.a) we have that $V(t,x) \geq 0$; moreover from the definition of limit for any $\varepsilon > 0$ there exists $\bar{n} \in \mathbb{N}$ such that $\forall n \geq \bar{n}$:

$$J(t,x;\alpha_n) \geq V(t,x) + \varepsilon. \tag{2.12}$$

Now we prove that this sequence is bounded in norm L^p, so that by Banach-Alaoglu Theorem we can extract a subsequence which weakly converges to some α^*, that is:

$$\lim_n \int_t^T \alpha_n(s)\phi(s)\ \mathrm{d}s = \int_t^T \alpha^*(s)\phi(s)\ \mathrm{d}s \qquad \forall\phi \in L^q([t,T])$$

In fact if $x_n(s) = x + \int_t^s \alpha_n(r)\ \mathrm{d}r$, $\forall n \geq \bar{n}$:

$$\begin{aligned}
\|\alpha_n\|^p_{L^p([t,T])} &\overset{(2.11.d)}{\leq} \int_t^T L(s,x_n(s),\alpha_n(s))\ \mathrm{d}s \\
&\overset{g\geq 0}{\leq} J(t,x;\alpha_n) \overset{(2.12)}{\leq} V(t,x) + \varepsilon \leq D,
\end{aligned}$$

for some D large. Since $(\alpha_n)_{n\in\mathbb{N}}$ is definitely bounded in the L^p norm, it is bounded too.

Now we want to apply Ascoli-Arzelá Theorem to $(x_n)_{n\in\mathbb{N}}$. In fact the (x_n) are equicontinuous:

$$|x_n(r) - x_n(s)| \leq \int_s^r |\alpha_n(\tau)| \, \mathrm{d}\tau \overset{\text{Holder}}{\leq} \|\alpha_n\|_p (r-s)^{\frac{1}{q}} \leq D(r-s)^{\frac{1}{q}},$$

and also uniformly bounded:

$$|x_n| \leq |x| + \int_t^T |\alpha_n| \, \mathrm{d}s \overset{\text{Holder}}{\leq} |x| + \|\alpha_n\|_p (T-t)^{\frac{1}{q}} \leq |x| + D(T-t)^{\frac{1}{q}}.$$

Consequently we can find a common subsequence (still denoted by $(\alpha_n)_{n\in\mathbb{N}}$ and $(x_n)_{n\in\mathbb{N}}$) such that:

$$\alpha_n \rightharpoonup \alpha^* \qquad \text{and} \qquad x_n \rightrightarrows x^*,$$

for some α^* and x^* such that $\dot{x}^* = \alpha^*$ almost everywhere. Now it remains to prove that:

$$J(t,x;\alpha^*) \leq \liminf_n J(t,x;\alpha_n) = V(t,x).$$

Let us set $\Lambda(s,v) := L(s, x^*(s), v)$. Λ and Λ_v are clearly continuous in $[t,T] \times \mathbb{R}^d$ and moreover $\Lambda(s,\cdot)$ is convex in the variable v because of (2.11.b) and (2.11.c). Therefore, for any $v \in \mathbb{R}^d$ we have:

$$\Lambda(s,v) \geq \Lambda(s,\alpha^*(s)) + (v - \alpha^*(s)) \cdot \Lambda_v(s,\alpha^*(s)).$$

In particular it is true for $v = \alpha_n(s)$. So, if we set $\chi_R(s)$ the indicator function of $\{s : |\alpha^*(s)| \leq R\}$, we obtain:

$$\begin{aligned} &\int_t^T \Lambda(s,\alpha_n(s))\chi_R(s) \, \mathrm{d}s \\ &\geq \int_t^T \Lambda(s,\alpha^*(s))\chi_R(s) \, \mathrm{d}s + \int_t^T (\alpha_n(s) - \alpha^*(s)) \cdot \Lambda_v(s,\alpha^*(s))\chi_R(s) \, \mathrm{d}s. \end{aligned}$$

Now we note that:

$$|\Lambda_v(s,\alpha^*(s))\chi_R(s)| \overset{(2.11.e)}{\leq} K(1 + |\alpha^*(s)|^p)\chi_R(s) \leq K(1 + R^p).$$

So $\Lambda_v(s,\alpha^*(s))\chi_R(s) \in L^\infty([t,T])$, and consequently $\Lambda_v(s,\alpha^*(s))\chi_R(s) \in L^q([t,T])$. This implies, since $\alpha_n \rightharpoonup \alpha^*$, that the rightmost term goes to 0 when $n \to \infty$. Therefore, for any $R > 0$:

$$\liminf_n \int_t^T \Lambda(s,\alpha_n(s))\chi_R(s)\ \mathrm{d}s \geq \int_t^T \Lambda(s,\alpha^*(s))\chi_R(s)\ \mathrm{d}s. \tag{2.13}$$

Since $\Lambda \geq 0$, $\chi_R(s) \leq 1$ and $\chi_R \underset{R\to\infty}{\longrightarrow} 1$, applying Fatou's Lemma twice we get:

$$\begin{aligned}\liminf_n \int_t^T \Lambda(s,\alpha_n(s))\ \mathrm{d}s &\geq \liminf_n \limsup_R \int_t^T \Lambda(s,\alpha_n(s))\chi_R(s)\ \mathrm{d}s\\ &\geq \liminf_n \liminf_R \int_t^T \Lambda(s,\alpha_n(s))\chi_R(s)\ \mathrm{d}s\\ &\overset{(2.13)}{\geq} \liminf_R \int_t^T \Lambda(s,\alpha^*(s))\chi_R(s)\ \mathrm{d}s\\ &\geq \int_t^T \Lambda(s,\alpha^*(s))\ \mathrm{d}s.\end{aligned} \tag{2.14}$$

Now, if we set $x_{\lambda n}(s) = x^*(s) + \lambda(x_n(s) - x^*(s))$, by the Mean value Theorem we have:

$$\begin{aligned}&\left|\int_t^T [L(s,x_n(s),\alpha_n(s)) - L(s,x^*(s),\alpha_n(s))]\ \mathrm{d}s\right|\\ \leq &\int_t^T \int_0^1 |L_x(s,x_{\lambda n}(s),\alpha_n(s))||x_n(s)-x^*(s)|\ \mathrm{d}\lambda\mathrm{d}s\\ \overset{(2.11.e)}{\leq} &K\|x_n - x^*\|_\infty \int_t^T (|\alpha_n(s)|^p + 1)\ \mathrm{d}s\\ \leq &K\|x_n - x^*\|_\infty (D+T).\end{aligned}$$

Since $x_n \rightrightarrows x^*$, we deduce that:

$$\liminf_n \left|\int_t^T [L(s,x_n(s),\alpha_n(s)) - L(s,x^*(s),\alpha_n(s))]\ \mathrm{d}s\right| = 0. \tag{2.15}$$

Finally we observe that

$$g(x_n(T)) \to g(x^*(T)) \tag{2.16}$$

because $x_n(T) \to x^*(T)$ and g is continuous.

In conclusion we have:

$$\begin{aligned}
&\liminf_n J(t,x;\alpha_n) = \\
&= \liminf_n \left[\int_t^T \Lambda(s,\alpha_n)\,\mathrm{d}s + \int_t^T [L(s,x_n,\alpha_n) - L(s,x^*,\alpha_n)]\,\mathrm{d}s + g(x_n(T))\right] \\
&\overset{(2.15)}{=} \liminf_n \left[\int_t^T \Lambda(s,\alpha_n(s))\,\mathrm{d}s + g(x_n(T))\right] \\
&\overset{(2.14)+(2.16)}{\geq} \int_t^T \Lambda(s,\alpha^*(s))\,\mathrm{d}s + g(x^*(T)) = J(t,x;\alpha^*).
\end{aligned}$$

□

Now we show that the minimizer solves the Euler-Lagrange equations in integrated form. Note that it is not trivial because we do not know yet if $\alpha^* \in \mathcal{C}^2([t,T])$.

Lemma 2.13. *Let $\bar{P}(s) := \int_t^s L_x(r,x^*(r),\alpha^*(r))\ dr$ for any $t \leq s \leq T$. Then for almost all $s \in [t,T]$ and some constant $C > 0$:*

$$\bar{P}(s) = L_v(s,x^*(s),\alpha^*(s)) + C. \tag{2.17}$$

Proof. Let $t < \bar{t} < T$ and $\varphi \in \mathcal{C}_c^1([t,\bar{t}];\mathbb{R}^d)$. For δ sufficiently small, $\lambda \in [0,1]$ and $s \in [t,\bar{t}]$ we set $x_\lambda(s) = x^*(s) + \delta\lambda\varphi(s)$, which is an admissible control. Since x^* and α^* are the minimizer, applying the Fundamental Theorem of calculus we have:

$$\begin{aligned}
0 &\leq \int_t^{\bar{t}} [L(s,x_1(s),\dot{x_1}(s)) - L(s,x^*(s),\alpha^*(s))]\,\mathrm{d}s \\
&= \int_t^{\bar{t}} [L(s,x_1(s),\dot{x_1}(s)) - L(s,x_0(s),\dot{x_0}(s))]\,\mathrm{d}s \\
&=\delta \int_t^{\bar{t}} \int_0^1 [L_x(s,x_\lambda(s),\dot{x_\lambda}(s)) \cdot \varphi(s) + L_v(s,x_\lambda(s),\dot{x_\lambda}(s)) \cdot \dot{\varphi}(s)]\,\mathrm{d}\lambda\mathrm{d}s.
\end{aligned}$$

Now we divide by δ and let $\delta \to 0$. Thanks to (2.11.d) and (2.11.e) we can apply the dominated convergence theorem to obtain:

$$0 \leq \int_t^{\bar{t}} [L_x(s,x^*(s),\alpha^*(s)) \cdot \varphi(s) + L_v(,x^*(s),\alpha^*(s)) \cdot \dot{\varphi}(s)]\,\mathrm{d}s.$$

Taking $-\varphi$ instead of φ, we get the converse inequality. Therefore:

$$0 = \int_t^{\bar{t}} [L_x(s, x^*(s), \alpha^*(s)) \cdot \varphi(s) + L_v(, x^*(s), \alpha^*(s)) \cdot \dot{\varphi}(s)] \, \mathrm{d}s.$$

Integrating by part $L_x(s, x^*(s), \alpha^*(s)) \cdot \varphi(s)$, we can conclude by the arbitrariness of $\bar{t}$ and φ and applying the fundamental lemma of the calculus of variations (in fact $\dot{\varphi} \in \mathcal{C}^0$). □

Now we want to prove that the minimizer $\alpha^* \in \mathcal{C}^2([t, T])$. To do this we need some notions about duality relationships.

Given a Lagrangian function $L \in \mathcal{C}^2(Q_T \times \mathbb{R}^d)$ convex in v and superlinear (that is $\lim_{|v|\to\infty} \frac{L(t,x,v)}{|v|} = \infty$ for any $(x, t) \in Q_T$), we can define the Hamiltonian function:

$$H(t, x, p) := \max_{v \in \mathbb{R}^d} [-v \cdot p - L(t, x, v)].$$

Proposition 2.14. *$H \in \mathcal{C}^2(\bar{Q}_T \times \mathbb{R}^d)$.*

Proof. We set $\xi = (t, x)$. By definition of H:

$$H(\xi, p) \geq -v \cdot p - L(\xi, p) \qquad \forall v \in \mathbb{R}^d, \tag{2.18}$$

and equality holds in (2.18) if and only if v maximizes the right side. Since L is convex and superlinear, for any $(\xi, p) \in \bar{Q}_T \times \mathbb{R}^d$ it happens in the unique $v \in \mathbb{R}^d$ such that $p = -L_v(\xi, v)$.

Now we rewrite (2.18) as

$$L(\xi, v) \geq -v \cdot p - H(\xi, p) \qquad \forall p \in \mathbb{R}^d.$$

Given $(\xi, v) \in \bar{Q}_T \times \mathbb{R}^d$ if we choose $p = -L_v(\xi, v)$, we note that it realizes the equality in (2.18). This gives the dual formula

$$L(t, x, v) = \max_p [-v \cdot p - H(t, x, p)].$$

These arguments show that the map $v \mapsto -L_v(t, x, v)$ is injective and surjective on $\mathbb{R}^d$ for any $\xi \in \bar{Q}_T$. Using the convexity of L in v and the implicit function theorem, there exists $\Gamma \in \mathcal{C}^2(\bar{Q}_T \times \mathbb{R}^d)$ such that:

$$\begin{cases} p = -L_v(\xi, \Gamma(\xi, p)) \\ H(\xi, p) = -\Gamma(\xi, p) \cdot p - L(\xi, \Gamma(\xi, p)). \end{cases} \tag{2.19}$$

This clearly gives that $H \in \mathcal{C}^2(\bar{Q}_T \times \mathbb{R}^d)$. □

By differentiating H in (2.19) with respect to p we get:

$$H_p(t, x, p) = -\Gamma - \Gamma_p \cdot p - L_v \cdot \Gamma_p \overset{p=-L_v \text{ in } (2.19)}{=} -\Gamma(t, x, p) = -v$$

So we got the Legendre transformation

$$\begin{cases} p = -L_v(t, x, v) \\ v = -H_p(t, x, p). \end{cases} \tag{2.20}$$

We are now ready to return to the discussion of the properties of the minimizer.

Lemma 2.15. $x^* \in \mathcal{C}^2([t, T])$.

Proof. By the Pontryagin's maximum principle (see [20, Chapter II]) we know that $\dot{x}^*$ must minimizes

$$-P(s) \cdot v - L(s, x^*(s), v),$$

where $P(s)\mathcal{C}^1$ is the solution of the adjoint equation

$$\dot{P}(s) = -L_x(s, x^*(s), P(s)).$$

This means that

$$P(s) = -L_v(s, x^*(s), \dot{x^*}(s)).$$

By the Legendre transformation (2.20) we get:

$$\dot{x^*} = -H_p(s, x^*(s), P(s)). \tag{2.21}$$

Since the right term is a continuous function of the variable s, we deduce that $\dot{x^*}$ is continuous. By differentiating (2.21) with respect to s, we obtain the thesis. □

Corollary 2.16 (Euler-Lagrange equation)**.** *x^* solves the Euler-Lagrange equation*

$$\begin{cases} \frac{d}{ds} L_v(s, x^*(s), \dot{x^*}(s)) = L_x(s, x^*(s), \dot{x^*}(s)) \\ L_v(T, x^*(T), \alpha^*(T)) = -D_x g(x^*(T)) \end{cases} \tag{2.22}$$

Proof. The transversality condition $L_v(T, x^*(T), \alpha^*(T)) = -D_x g(x^*(T))$ comes from one of the necessary conditions given by Pontryagin's maximum principle.

Moreover since $x^* \in \mathcal{C}^2$ we can differentiate (2.17) with respect to s, obtaining the thesis. □

2.3 Analysis of the HJB equation

This section is devoted to the study of the Hamilton-Jacobi-Bellman equation

$$\begin{cases} -\partial_t u(x,t) + \frac{1}{2}|D_x u(x,t)|^2 = f(x,t) & \text{in } \mathbb{R}^d \times (0,T) \\ u(x,T) = g(x) & \text{in } \mathbb{R}^d. \end{cases} \tag{2.23}$$

Here the hypotheses on f and g are the same on F and G at the beginning of the chapter. As we have already discussed, this partial differential equation is linked to the optimal control problem

$$\inf_{\alpha \in L^2([t,T];\mathbb{R}^d)} J(t,x;\alpha) = \inf_{\alpha \in L^2([t,T];\mathbb{R}^d)} \int_t^T \left[\frac{1}{2}|\alpha(s)|^2 + f(x(s),s)\right] \mathrm{d}s + g(x(T)), \tag{2.24}$$

where $x(s) = x + \int_t^s \alpha(r)\, \mathrm{d}r$. So our first task is to see what the results of Section 2.2 become in this particular case.

Remark 2.17. Let us start noting that the Lagrangian of the problem (2.24) is $L(t,x,v) := f(x,t) + \frac{1}{2}|v|^2$, and that hypotheses (2.11) are satisfied:

a) $L, g \geq 0$ because they represent costs.

b) $L, g \in \mathcal{C}^2$ by our assumptions (hence $R = 1$).

c) $L_{vv} \equiv 1 > 0$.

d) It is sufficient to take $p = 2$, $A = \frac{1}{2}$ and $B = -\|f\|_{\mathcal{C}^2}$.

e) Since $L_v = v$ and $L_x = D_x f$, then $|L_x| + |L_v| \leq K(1 + |v|^2)$ if

$$|v| + \|f\|_{\mathcal{C}^2} \leq K(|v|^2 + 1),$$

that is a second degree equation in $|v|^2$ which is always satisfied as soon as $K \geq \frac{\|f\|_{\mathcal{C}^2} + \sqrt{1+\|f\|^2_{\mathcal{C}^2}}}{2}$.

Hence for any $(t, x) \in Q_T$ there exists a minimizer α^* of the problem (2.24). Furthermore $\alpha^* \in \mathcal{C}^1([t, T])$ and solves the Euler-Lagrange equation:

$$\begin{cases} \dot{\alpha^*}(s) = D_x f(x^*(s), s) \\ \alpha^*(T) = -D_x g(x^*(T)) \end{cases} \tag{2.25}$$

Corollary 2.18. *The minimizer α^* of (2.24) is bounded by a constant M which depends only on C and T and not on the initial condition. Therefore it holds:*

$$\left|\frac{1}{2}\alpha(s)^2 + f(x(s), s)\right| \leq \frac{M^2}{2} + C \qquad \forall s \in [0, T]. \tag{2.26}$$

Proof. The boundedness of α descends from (2.25) because of our assumption (2.2). It follows (2.26). □

We are now ready to enunciate and demonstrate the main result of this section, about the solution of (2.23).

Theorem 2.19. *If $f : Q_T \to \mathbb{R}$ and $g\mathbb{R}^d \to \mathbb{R}$ are continuous and such that*

$$\|f(\cdot, t)\|_{\mathcal{C}^2} \leq C \ \forall t \in [0, T], \qquad \|g\|_{\mathcal{C}^2} \leq C \tag{2.27}$$

then the value function $u(x, t) := \inf_{\alpha \in L^2([t,T])} J(t, x; \alpha)$ is the unique bounded and uniformly continuous viscosity solution of (2.23). Moreover $\exists C_1 = C_1(C, T)$ such that:

$$\|D_{t,x}u\|_\infty \leq C_1 \qquad \text{and} \qquad D^2_{xx}u \leq C_1, \tag{2.28}$$

where the last inequality holds in the sense of distributions.

Proof. By Corollary 2.18 we deduce that the image of a minimizer of (2.24) is contained in $[-M, M]$. Thererefore:

$$\inf_{\alpha \in L^2([t,T];\mathbb{R}^d)} J(t, x; \alpha) = \inf_{\alpha \in L^2([t,T];[-M,M])} J(t, x; \alpha).$$

Moreover, from Corollary 2.18 and our assumptions on f and g descends that both L and g are bounded and Lipschitz. It is a well known fact (see [17, Chapter 10]) that under these hypotheses the value function u is the unique bounded and uniformly continuous viscosity solution of (2.23).

Hence we have only to prove (2.28), so we start looking for the Lipschitz constant of u. First we do it for the x variable. Let $(x_1, x_2, t) \in \mathbb{R}^d \times \mathbb{R}^d \times [0, T]$ and let $\alpha \in L^2([t, T])$ ε-optimal for $u(x_1, t)$, that is

$$\int_t^T \left[\frac{1}{2}|\alpha(s)|^2 + f(x(s), s)\right] \, \mathrm{d}s + g(x(T)) \leq u(t, x_1) + \varepsilon, \tag{2.29}$$

where $x(s) = x_1 + \int_t^s \alpha(r) \, \mathrm{d}r$. Clearly $x(t) = x_1$, if we wanted an admissible state with initial position x_2 we should consider $\tilde{x}(s) := x(s) + x_2 - x_1$. Using the minimality of $u(x_2, t)$ for $J(t, x; \cdot)$ we have:

$$\begin{aligned}
u(x_2, t) &\leq \int_t^T \left[\frac{1}{2}|\alpha(s)|^2 + f(\tilde{x}(s), s)\right] \, \mathrm{d}s + g(\tilde{x}(T)) \\
&= \int_t^T \left[\frac{1}{2}|\alpha|^2 + f(x(s) + x_2 - x_1, s) - f(x(s), s) + f(x(s), s)\right] \, \mathrm{d}s + \\
&\qquad + g(x(T) + x_2 - x_1) - g(x(T)) + g(x(T)) \\
&\overset{(2.27)}{\leq} \int_t^T \left[\frac{1}{2}|\alpha|^2 + f(x(s), s) + C|x_2 - x_1|\right] \, \mathrm{d}s + g(x(T)) + C|x_2 - x_1| \\
&\overset{(2.29)}{\leq} u(x_1, t) + \varepsilon + C(T + 1)|x_2 - x_1|.
\end{aligned}$$

By letting $\varepsilon \to 0$ we conclude that:

$$\|D_x u\|_\infty \leq C(T + 1) \tag{2.30}$$

Now we do the same for the time variable, but with a different strategy. By the dynamic programming principle we know that for any $t \leq s \leq T$ it holds:

$$u(x, t) = \int_t^s \left[\frac{1}{2}|\alpha(r)|^2 + f(x(r), r)\right] \, \mathrm{d}r + u(x(s), s), \tag{2.31}$$

where α is optimal for $u(x,t)$ and $x(\cdot)$ is the state associated to it. Therefore:

$$
\begin{aligned}
|u(x,s)-u(x,t)| &\le |u(x,t)-u(x(s),s)| + |u(x(s),s)-u(x,s)| \\
&\overset{(2.31)+(2.30)}{\le} \left| \int_t^s \left[\frac{1}{2}|\alpha(r)|^2 + f(x(r),r) \right] \, \mathrm{d}r \right| + C(T+1)|x(s)-x| \\
&\le \int_t^s \left[\frac{1}{2}|\alpha|^2 + |f| \right] \, \mathrm{d}r + C(T+1)\int_t^s |\alpha(r)| \, \mathrm{d}r \\
&\overset{(2.26)+(2.18)}{\le} (s-t)\left[\frac{1}{2}M^2 + C + C(T+1)M \right].
\end{aligned}
$$

If we set $C_1 := \max\left\{C(T+1), \frac{1}{2}M^2 + C + C(T+1)M\right\}$, we get $D_{t,x}u \le C_1$ as we wanted. Note that C_1 depends only on C and T, so that M.

It remains to show the second one in (2.28), that by Proposition 2.2 is equivalent to saying that u is semiconcave with linear modulus of constant C_1 in the variable x for any $t \in [0,T]$. Furthermore by Proposition 2.3 we have that f and g are semiconcave with linear modulus of constant C.

Let $(x,y,t) \in \mathbb{R}^d \times \mathbb{R}^d \times [0,T]$, $\lambda \in (0,1)$ and $x_\lambda := \lambda x + (1-\lambda)y$. Moreover let $\alpha \in L^2([t,T])$ ε-optimal for $u(x_\lambda,t)$, with $x_\lambda(s) = x_\lambda + \int_t^s \alpha(r)\,\mathrm{d}r$. By definition of semiconcavity we have that:

$$
\begin{aligned}
&\lambda f(\underbrace{x_\lambda(s)+x-x_\lambda}_{a}, s) + (1-\lambda) f(\underbrace{x_\lambda(s)+y-x_\lambda}_{b}, s) \\
&\le f(\lambda a + (1-\lambda)b, s) + \lambda(1-\lambda)\frac{C}{2}|x-y|^2 \\
&= f(x_\lambda(s), s) + \lambda(1-\lambda)\frac{C}{2}|x-y|^2.
\end{aligned}
\tag{2.32}
$$

And the same for g. Therefore:

$$
\begin{aligned}
&\lambda u(x,t) + (1-\lambda)u(y,t) \le \\
&\le \lambda \left(\int_t^T \left[\frac{1}{2}|\alpha|^2 + f(x_\lambda(s)+x-x_\lambda, s) \right] \, \mathrm{d}s + g(x_\lambda(T)+x-x_\lambda) \right) + \\
&\qquad + (1-\lambda) \left(\int_t^T \left[\frac{1}{2}|\alpha|^2 + f(x_\lambda(s)+y-x_\lambda, s) \right] \, \mathrm{d}s + g(x_\lambda(T)+y-x_\lambda) \right) \\
&= \int_t^T \frac{1}{2}|\alpha|^2 \, \mathrm{d}s + \int_t^T \left[\lambda f(a,s) + (1-\lambda) f(b,s) \right] \, \mathrm{d}s + \lambda g(a) + (1-\lambda) g(b)
\end{aligned}
$$

$$\overset{(2.32)}{\leq} \int_t^T \left[\frac{1}{2}|\alpha|^2 + f(x_\lambda(s), s)\right] \, \mathrm{d}s + g(x(T)) + \frac{1}{2}C(T+1)\lambda(1-\lambda)|x-y|^2$$
$$\leq u(x_\lambda, t) + \varepsilon + \frac{1}{2}C_1\lambda(1-\lambda)|x-y|^2.$$

By letting $\varepsilon \to 0$ we get that the constant of semiconcavity of u is C_1. □

Remark 2.20. It is possible to weaken our assumptions on f and g, but we would get only the local semiconcavity of u, while in the following discussion we will need its global semiconcavity.

Otherwise, stressing the hypotheses on f and g, we should be content with local results. See for instance [27, Th. 3.2, Cor. 9.2, Th. 2.2] and overall [7, Chapter 6].

Now we can demonstrate a last property of the minimizers of (2.24).

Notations. For any $(t, x) \in Q_T$ we denote with $\mathcal{A}(t, x)$ the set of all optimal controls which realizes the minimum in (2.24).

Lemma 2.21 (Stability of optimal solutions)**.** *If $(t_n, x_n) \to (t, x)$ then, up to a subsequence, $\alpha_n \in \mathcal{A}(t, x)$ weakly converges in $L^2([t, T])$ to some $\alpha \in \mathcal{A}(t, x)$*

Proof. Clearly $\mathcal{A}(t, x)$ and $\mathcal{A}(t_n, x_n)$ are non-empty for any $n \in N$ by Theorem 2.12. Hence we set:

$$\tilde{\alpha}_n(s) = \begin{cases} \alpha_n(s) & s \in [t_n, T] \\ 0 & s \in [0, t_n[. \end{cases}$$

Since $\alpha_n \in L^2([t_n, T])$ for any $n \in \mathbb{N}$, then $\tilde{\alpha}_n \in L^2([0, T])$ for any $n \in \mathbb{N}$. Moreover it is bounded in the L^2 norm, in fact:

$$\|\tilde{\alpha}_n\|_2^2 = \int_{t_n}^T |\alpha_n(s)|^2 \, \mathrm{d}s \overset{\text{Cor. 2.18}}{\leq} TM^2.$$

Therefore by Banach-Alaoglu's Theorem there exists a subsequence of $(\tilde{\alpha}_n)_{n\in\mathbb{N}}$, still denoted with $(\tilde{\alpha}_n)_{n\in\mathbb{N}}$, which weakly converges to some $\alpha \in L^2([t, T])$; we have to show that $\alpha \in \mathcal{A}(t, x)$. Using the dominated convergence theorem thanks to (2.26) and the continuity of the value function u we have:

$$\begin{aligned} u(x,t) &= \lim_n u(x_n,t_n) \\ &= \lim_n \int_{t_n}^T \left[\frac{1}{2}|\alpha_n(s)|^2 + f(x_n(s),s)\right] \, \mathrm{d}s \\ &= \int_t^T \left[\frac{1}{2}|\alpha(s)|^2 + f(x(s),s)\right] \mathrm{d}s, \end{aligned}$$

where $x_n(s) = x_n + \int_t^s \alpha_n(r) \, \mathrm{d}r$. Hence $\alpha \in \mathcal{A}(t,x)$. $\square$

Remark 2.22. By the previous Lemma we have that if we endow $L^2([t,T];\mathbb{R}^d)$ with the weak topology, then the correspondence $\Lambda(x,t) := \mathcal{A}(x,t)$ has a closed graph. Hence Λ is measurable with nonempty closed values, so that it has a Borel selection $\bar{\alpha} \in \mathcal{A}(x,t)$ for any $(x,t) \in Q_T$ (see [2]).

We would conclude that $\alpha^* = -D_x u$, but this is not trivial because we can not use the verification theorem since first $D_x u$ is not said to be continuous. Hence we have to demonstrate the regularity of the value function along optimal solutions, which gives a sufficient and necessary condition for being an optimal trajectory.

Lemma 2.23 (Uniqueness of optimal control along optimal trajectories)**.** *Let $(x,t) \in Q_T$, $\alpha \in \mathcal{A}(x,t)$ and $x(s) = x + \int_t^s \alpha(s) \, dr$. Then for any $s \in (t,T)$ we have $\mathcal{A}(x(s),s) = \{\alpha_{|_{[s,T]}}\}$.*

Proof. Let $\alpha_1 \in \mathcal{A}(x(s),s)$ and $x_1(\cdot)$ its trajectory, we want to show that $\alpha_1 = \alpha_{|_{[s,T]}}$.

Now, for any $h > 0$ small, we build the admissible control:

$$\alpha_h(r) := \begin{cases} \alpha(r) & r \in [t, s-h[\\ \dfrac{x_1(s+h) - x(s-h)}{2h} & r \in [s-h, s+h[\\ \alpha_1(r) & r \in [s+h, T]. \end{cases}$$

By easy calculations we get that its linked state $x_h(r) = x + \int_t^r \alpha_h(\tau) \, \mathrm{d}\tau$ is

$$x_h(r) = \begin{cases} x(r) & r \in [t, s-h[\\ x(s-h) + (r - (s-h))\dfrac{x_1(s+h) - x(s-h)}{2h} & r \in [s-h, s+h[\\ x_1(r) & r \in [s+h, T]. \end{cases}$$

Now, if we concatenate $\alpha_{|_{[t,s]}}$ and $\alpha_1 \in A(x(s), s)$ we obtain a control $\alpha_0 \in \mathcal{A}(x, t)$, to which we associate the state

$$x_0(r) = x + \int_t^r \alpha_0(\tau)\, \mathrm{d}\tau = \begin{cases} x(r) & r \in [t, s] \\ x_1(r) & r \in [s, T]. \end{cases}$$

Then, comparing the payoff of α_0 and α_h we have:

$$\begin{aligned}
&J(t, s; \alpha_0) = \\
&= \int_t^s \left[\frac{1}{2}|\alpha(r)|^2 + f(x(r), r))\right] \mathrm{d}r + \int_s^T \left[\frac{1}{2}|\alpha_1(r)|^2 + f(x_1(r), r))\right] \mathrm{d}r + g(x_1(T)) \\
&\overset{\alpha_0 \in \mathcal{A}(x,t)}{\leq} \int_t^T \left[\frac{1}{2}|\alpha_h(r)|^2 + f(x_h(r), r))\right] \mathrm{d}r + g(x_h(T)) = \\
&= \int_t^{s-h} \left[\frac{1}{2}|\alpha(r)|^2 + f(x(r), r))\right] \mathrm{d}r + \int_{s+h}^T \left[\frac{1}{2}|\alpha_1(r)|^2 + f(x_1(r), r))\right] \mathrm{d}r + \\
&+ \int_{s-h}^{s+h} \left[\frac{1}{2}\left|\frac{x_1(s+h) - x(s-h)}{2h}\right|^2 + f(x_h(r), r))\right] \mathrm{d}r + g(x_1(T)).
\end{aligned}$$

Therefore:

$$\begin{aligned}
0 \geq \int_{s-h}^s & \left[\frac{1}{2}|\alpha(r)|^2 + f(x(r), r))\right] \mathrm{d}r + \int_s^{s+h} \left[\frac{1}{2}|\alpha_1(r)|^2 + f(x_1(r), r))\right] \mathrm{d}r - \\
& - \int_{s-h}^{s+h} \left[\frac{1}{2}\left|\frac{x_1(s+h) - x(s-h)}{2h}\right|^2 + f(x_h(r), r))\right] \mathrm{d}r.
\end{aligned} \tag{2.33}$$

We divide this inequality by h and then we let $h \to 0$. Using the mean value theorem for integrals we get:

$$\begin{aligned}
&\frac{1}{2}|\alpha(s)|^2 + f(x(s), s) + \frac{1}{2}|\alpha_1(s)|^2 + f(x_1(s), s) - \\
&\qquad - \frac{1}{2} \lim_{h \to 0} \frac{1}{h} \int_{s-h}^{s+h} \frac{1}{4} \left[\left|\frac{x_1(s+h) - x(s-h)}{h}\right|^2 + 2f(x_h(r), r))\right] \mathrm{d}r \leq 0.
\end{aligned}$$

Now $x_1(s) = x(s)$, so:

$$\lim_h \frac{x_1(s+h) - x(s-h)}{h} = \lim_h \frac{x_1(s+h) - x_1(s)}{h} + \lim_h \frac{x(s) - x(s-h)}{h} =$$

$$= \dot{x_1}(s) + \dot{x}(s) = \alpha_1(s) + \alpha(s).$$

Hence:

$$\frac{1}{2} \lim_{h\to 0} \frac{1}{h} \int_{s-h}^{s+h} \frac{1}{4} \left[\left| \frac{x_1(s+h) - x(s-h)}{h} \right|^2 + 2f(x_h(r), r)) \right] \, \mathrm{d}r =$$

$$= \frac{1}{4}(|\alpha_1(s) + \alpha(s)|^2) + \lim_{h} \frac{1}{h} \int_{s-h}^{s+h} f(x_h(r), r) \, \mathrm{d}r.$$

Since $x_h(s) = \frac{x_1(s+h)+x(s-h)}{2} \underset{h\to 0}{\longrightarrow} x(s)$ the last term tends to $f(x(s), s)$. Combining this facts with (2.33) we deduce:

$$\begin{aligned} 0 &\geq \frac{1}{2}|\alpha(s)|^2 + \frac{1}{2}|\alpha_1(s)|^2 - \frac{1}{4}|\alpha(s) + \alpha_1(s)|^2 \\ &= \frac{1}{2}|\alpha(s)|^2 + \frac{1}{2}|\alpha_1(s)|^2 - \frac{1}{4}|\alpha(s)|^2 - \frac{1}{4}|\alpha_1(s)|^2 - \frac{1}{2}\langle \alpha(s), \alpha_1(s) \rangle \\ &= \frac{1}{4}|\alpha(s)|^2 + \frac{1}{4}|\alpha_1(s)|^2 - \frac{1}{2}\langle \alpha(s), \alpha_1(s) \rangle = \frac{1}{4}|\alpha(s) - \alpha_1(s)|^2. \end{aligned}$$

So this implies $\alpha(s) = \alpha_1(s)$. But by (2.25) this means that both $x(\cdot)$ and $x_1(\cdot)$ solve the ODE

$$\ddot{y}(r) = D_x f(y(r), r) \qquad r \in [s, T],$$

with the same initial conditions $x(s) = x_1(s)$ and $\dot{x}(s) = \alpha(s) = \alpha_1(s) = \dot{x_1}(s)$. Therefore $x(\cdot) = x_1(\cdot)$ on $[s, T]$, and this implies $\alpha_1 = \alpha_{|[s,T]}$, that is the optimal solution for $u(x(s), s)$ is unique. □

Lemma 2.24 (Uniqueness of the optimal trajectories)**.** *Under the same notations of the previous Lemma, $D_x u(x, t)$ exists if and only if $\mathcal{A}(x, t)$ is reduced to a singleton. In this case $D_x u(x, t) = -\alpha(t)$, where $\mathcal{A}(x, t) = \{\alpha\}$.*

Proof.

- $\Longrightarrow$) Let $\alpha \in \mathcal{A}(x, t)$ and $x(\cdot)$ its trajectory. By minimality, for any $v \in \mathbb{R}^d$ we have:

  $$u(x + v, t) \leq \int_t^T \left[\frac{1}{2}|\alpha(s)|^2 + f(x(s) + v, s) \right] \, \mathrm{d}s + g(x(T) + v),$$

  and equality holds when $v = 0$. Since we suppose that $D_x u(x, t)$ exists, if we differentiate with respect to v at $v = 0$ we get:

$$D_x u(x,t) = \int_t^T D_x f(x(s),s)\ \mathrm{d}s + D_x g(x(T))$$

$$\stackrel{(2.25)}{=} \int_t^T \dot{\alpha}(s)\ \mathrm{d}s - \alpha(T) = -\alpha(t).$$

In particular $x(\cdot)$ must be the unique solution of the Cauchy problem

$$\begin{cases} \ddot{x}(s) = D_x f(x(s),s) & s \in]t,T] \\ \dot{x}(t) = -D_x u(x,t) \\ x(t) = x. \end{cases} \tag{2.34}$$

So $\alpha = \dot{x}$ is unique.

- $\Longleftarrow$) If $p \in D_x^* u(x,t)$ then the solution $x(\cdot)$ of

$$\begin{cases} \ddot{x}(s) = D_x f(x(s),s) & s \in]t,T] \\ \dot{x}(t) = -p \\ x(t) = x \end{cases}$$

is optimal. In fact by definition of $D_x^* u(x,t)$ there exists $(x_n)_{n\in\mathbb{N}}$ such that $x_n \to x$, $u(\cdot,t)$ is differentiable at x_n and $D_x u(x_n,t) \to p$. By the inverse implication of this Lemma, the unique solution $x_n(\cdot)$ of

$$\begin{cases} \ddot{y}(s) = D_x f(y(s),s) & s \in]t,T] \\ \dot{y}(t) = -D_x u(x_n,t),\ y(t) = x_n \end{cases}$$

is optimal. Since $D_x f$ is Lipschitz in x uniformly in t by (2.27), $x(\cdot)$ is the uniform limit of the $(x_n(\cdot))_{n\in\mathbb{N}}$ and by Lemma 2.21 also $x(\cdot)$ is optimal.

Now, since $\mathcal{A}(x,t)$ is a singleton, this implies that $D_x^* u(x,t) = \{p\}$. Therefore:

$$D_x^+ u(x,t) \stackrel{\text{Th. } 2.11}{=} \mathrm{Co}(D_x^* u(x,t)) = \{p\}.$$

That is $D_x^+ u(x,t)$ is a singleton, and by Proposition 2.6.iii we have that $u(\cdot,t)$ is differentiable at x.

$\square$

Corollary 2.25. *$u(\cdot, s)$ is always differentiable in $x(s)$ for any $s \in (t, T)$, with $D_x u(x(s), s) = -\alpha(s)$. In particular $x(\cdot)$ is solution of the differential equation*

$$\dot{x}(s) = -D_x u(x(s), s).$$

Proof. By Lemma 2.23 we have $\mathcal{A}(x(s), s) = \{\alpha_{|_{[s,T]}}\}$ for any $s \in]t, T]$, that is a singleton; so by Lemma 2.24 $u(\cdot, s)$ is differentiable at $x(s)$ and $\alpha(s) = \dot{x}(s) = -D_x u(x(s), s)$. □

So we have proved that the optimal trajectory solves the previous differential equation with the initial condition $x(t) = x$. Now we want to prove that the reverse is also true.

Lemma 2.26 (Optimal synthesis)**.** *Let $(x, t) \in Q_T$ and let $x(\cdot)$ be a solution of*

$$\begin{cases} \dot{x}(s) = -D_x u(x(s), s) & a.e.\ in\ [t, T] \\ x(t) = x. \end{cases} \tag{2.35}$$

Then $\alpha(\cdot) = \dot{x}(\cdot) \in \mathcal{A}(x, t)$. In particular, if $u(\cdot, t)$ is differentiable at x, (2.35) has unique solution corresponding to the optimal trajectory.

Proof. Let $s \in]t, T[$ such that (2.35) is satisfied and the map $r \mapsto u(x(r), r)$ is differentiable at s. Since u is Lipschitz continuous by Lebourg's mean value Theorem (see [13, Th. 2.3.7]) for any $h > 0$ small there exists $(y_h, s_h) \in [(x(s), s), (x(s + h), s + h)]$ and $(\xi_x^h, \xi_t^h) \in \mathrm{Co}(D^*_{t,x} u(y_h, s_h)) \overset{\text{Th. 2.11}}{=} D^+_{t,x} u(y_h, s_h)$ such that:

$$u(x(s + h), s + h) - u(x(s), s) = \langle \xi_x^h, x(s + h) - x(s) \rangle + \xi_t^h h. \tag{2.36}$$

From Caratheodory Theorem there are some $(\lambda^{h,i}, \xi_x^{h,i}, \xi_t^{h,i})_{i=1,\ldots,d+2}$ with $\lambda^{h,i} \geq 0$, $\sum_i \lambda^{h,i} = 1$ and $(\xi_t^{h,i}, \xi_x^{h,i}) \in D^*_{t,x} u(y_h, s_h)$ such that

$$(\xi_t^h, \xi_x^h) = \sum_{i=1}^{d+2} \lambda^{h,i} (\xi_t^{h,i}, \xi_x^{x,i}).$$

By Proposition 2.10.i any cluster point of $(\xi_x^{h,i})_h$ belongs to $D^*_x u(x(s), s) \overset{\text{Th. 2.6.iii}}{=} D_x u(x(s), s)$. Hence

$$\xi_x^{h,i} \underset{h \to 0}{\longrightarrow} D_x(x(s), s). \tag{2.37}$$

Since u is viscosity solution of (2.23) and $(\xi_t^{h,i}, \xi_x^{h,i}) \in D^*_{t,x}u(y_h, s_h)$, from the consistency properties and passing to the limit we have that:

$$-\xi_t^{h,i} + \frac{1}{2}|\xi_x^{h,i}|^2 = f(y_h, s_h).$$

Therefore, using (2.37):

$$\xi_t^h = \sum_i \lambda^{h,i}\xi_t^{h,i} = \frac{1}{2}\sum_i \lambda^{h,i}|\xi_x^{h,i}|^2 - f(y_h, s_h) \underset{h\to 0}{\longrightarrow} \frac{1}{2}|D_x u(x(s), s)|^2 - f(x(s), s).$$

Now, dividing (2.36) by h and letting $h \to 0$ we get

$$\frac{\mathrm{d}}{\mathrm{d}s}u(x(s), s) = \langle D_x u(x(s), s), \dot{x}(s)\rangle + \frac{1}{2}|D_x u(x(s), s)|^2 - f(x(s), s).$$

Using $\dot{x}(s) = D_x u(x(s), s)$ a.e. in $[t, T]$ we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s}u(x(s), s) = -\frac{1}{2}|D_x u(x(s), s)|^2 - f(x(s), s) \qquad \text{a.e. in }]t, T[.$$

Integrating this last equality over $[t, T]$ we have

$$\underbrace{u(x(T), T)}_{g(x(T))} - u(\underbrace{x(t)}_{x}, t) = \int_t^T \left[-\frac{1}{2}|\underbrace{D_x u(x(s), s)}_{\dot{x}(s)}|^2 - f(x(s), s)\right] \mathrm{d}s$$

$$\Updownarrow$$

$$u(x, t) = \int_t^T \left[\frac{1}{2}|\dot{x}(s)|^2 + f(x(s), s)\right] \mathrm{d}s + g(x(T)).$$

Therefore $\alpha(\cdot) = \dot{x}(\cdot)$ is optimal.

The last part of the statement descends from Lemma 2.24. □

Definition 12 (Flow)**.** For any $(x, t) \in Q_T$ let $\bar{\alpha}(x, t) \in \mathcal{A}(x, t)$ as in Remark 2.22. Then we define the flow

$$\Phi(x, t, s) = x + \int_t^s \bar{\alpha}(x, t)(r)\, \mathrm{d}r \qquad \forall s \in [t, T].$$

Now we study some properties of the flow that will be useful in the next section.

Proposition 2.27. *The following hold:*

i. *The flow has the semigroup property, that is*

$$\Phi(x,t,s') = \Phi(\Phi(x,t,s),s,s') \qquad \text{for any } t \le s \le s' \le T. \tag{2.38}$$

ii. $\partial_s \Phi(x,t,s) = -D_x u(\Phi(x,t,s),s)$ *for any* $x \in \mathbb{R}^d$ *and* $s \in (t,T)$*.*

iii. $|\Phi(x,t,s') - \Phi(x,t,s)| \le \|D_x u\|_\infty |s'-s|$ *for any* $x \in \mathbb{R}^d$*,* $t \le s \le s' \le T$*.*

Proof. i. It is a direct consequence of Lemma 2.23, from which we know that $\mathcal{A}(\Phi(x,t,s),s) = \{\bar{\alpha}(x,t)_{|[s,T]}\}$.

ii. From Corollary 2.25 we know that $u(\cdot,s)$ is differentiable at $\Phi(x,t,s)$ with $D_x u(\Phi(x,t,s),s) = -\bar{\alpha}(x,t)(s)$. We get the thesis noting that by definition $\partial_s \Phi(x,t,s) = \bar{\alpha}(x,t)(s)$.

iii. Since the optimal trajectory solves (2.35) and u is Lipschitz continuous, it descends that Φ is Lipschitz with respect to the variable s. We reach the thesis by the previous point since $\|\partial_s \Phi\|_\infty = \|D_x u\|_\infty$.

□

Lemma 2.28 (Contraction property)**.** *Let C be as in (2.27), then there exists some constant $C_2 = C_2(C,T)$ such that if u is solution of (2.23), then:*

$$|x-y| \le C_2 |\Phi(x,t,s) - \Phi(y,t,s)| \qquad \forall x,y \in \mathbb{R}^d \text{ and } \forall\ 0 \le t < s \le T. \tag{2.39}$$

In particular the map $x \mapsto \Phi(x,t,s)$ has a Lipschitz continuous inverse on the set $\Phi(\mathbb{R}^d,t,s)$.

Proof. Clearly $\Phi(\cdot,t,s)$ is injective and then it is invertible on the image. Let us define $x(r) := \Phi(x,t,s-r)$ and $y(r) := \Phi(y,t,s-r)$ for any $r \in [0,s-t]$. From the previous proposition we have

$$\begin{cases} \dot{x}(r) = -\partial_s \Phi(x,t,s-r) = D_x u(\Phi(x,t,s-r),s-r) = D_x u(x(r),s-r) \\ x(0) = \Phi(x,t,s) \end{cases}$$

Similarly $\dot{y}(r) = D_x u(y(r), s-r)$ with initial condition $y(0) = \Phi(y,t,s)$. Using the semiconcavity of u and (2.8) we get

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{1}{2}|(x-y)(r)|^2\right) = \langle(\dot{x}-\dot{y})(r),(x-y)(r)\rangle \leq C_1|(x-y)(r)|^2.$$

Hence we derive:

$$\int_0^r \left[\frac{\frac{\mathrm{d}}{\mathrm{d}\tau}|(x-y)(\tau)|^2}{|(x-y)(\tau)|^2}\right]\,\mathrm{d}\tau \leq \int_0^r 2C_1\,\mathrm{d}r \qquad \forall r\in[0,s-t]$$

Then

$$\log\left(\frac{|(x-y)(r)|^2}{|(x-y)(0)|^2}\right) \leq 2C_1 r \leq 2C_1 T \qquad \forall r\in[0,s-t],$$

from which we deduce, taking $r = s-t$:

$$|x(0)-y(0)|e^{C_1T} \geq |x(s-t)-y(s-t)| \iff |\Phi(x,t,s)-\Phi(y,t,s)|e^{C_1T} \geq |x-y|.$$

If we set $C_2 = e^{C_1T}$ we have the thesis. □

2.4 Analysis of the continuity equation

In this section we study the Kolmogorov equation:

$$\begin{cases}\partial_s\mu(x,s) - \mathrm{div}_x(D_x u(x,s)\mu(x,s)) = 0 & \text{in } \mathbb{R}^d\times]0,T]\\ \mu(x,0) = m_0(x).\end{cases} \tag{2.40}$$

We want to show that it has unique solution under our assumptions on f and g and once given u solution of (2.23). Being (2.40) a continuity equation one can guess that the solution is the density of the push-forward measure of m_0 through the flow Φ.

Definition 13. For any $s\in[0,T]$, we set the push forward measure $\mu_s := \Phi(\cdot,0,s)\#m_0$, which satisfies for any borel set $A\in\mathbb{R}^d$:

$$\mu_s(A) = m_0(\Phi^{-1}(A,0,s))$$

The interpretation is that the measure of A at the time s is equal to the measure at time 0 of that set which after a time s evolves in A through the dynamics of the system.

In the next we first demonstrate that the density $\mu(s,x)$ of μ_s is a weak solution of (2.40), then we prove that uniqueness holds with suitable regularity conditions on $D_x u$ and finally, through regularizations, we show that we can extend uniqueness to the general case. In the following we will assume without loss of generality $t = 0$ in (2.24).

First of all we need a result which will often be used later about change of variables in integrals with push forward measure.

Lemma 2.29. *Let X and Y measurable spaces, μ a nonnegative measure on X and $f : X \to Y$ a measurable function. Then a measurable function g on Y is integrable with respect to the measure $f\#\mu$ precisely when the function $g \circ f$ is integrable with respect to μ. In addition one has*

$$\int_Y g(y)\ f\#\mu(dy) = \int_X g(f(x))\mu(dx) \tag{2.41}$$

Proof. Let $\mathscr{B}$ the σ-algebra on Y. For the indicators of sets in $\mathscr{B}$ formula (2.41) is just the definition of push forward measure and by linearity it holds also for simple simple functions. Next we can extend this formula to bounded $\mathscr{B}$-measurable functions because they are uniform limits of simple ones. If g is a nonnegative $\mathscr{B}$-measurable function which is integrable with respect to $f\#\mu$, then for the functions $g_n := \min(g, n)$ (2.41) already holds. By the monotone convergence theorem it remains true for g, since the integrals of the functions $g_n \circ f$ against the measure μ are uniformly bounded.

These arguments show the necessity of the μ-integrability of $g \circ f$ for the integrability of $g \geq 0$ with respect $f\#\mu$. By the linearity of (2.41) in g we obtain the general case. □

Now we prove that $\mu_s = \Phi(\cdot, 0, s)\#m_0$ is absolutely continuous with respect the Lebesgue measure and Lipschitz.

Lemma 2.30. *Let C as in (2.27) and such that m_0 is absolutely continuous with respect to Lebesgue measure with a density $m_0(x)$ which satisfies $\|m_0\|_\infty \leq C$ and with support contained in $B(0, C)$. Let us set $\mu_s = \Phi(\cdot, 0, s)\#m_0$ for any $s \in [0, T]$. Then there exists a constant $C_3 = C_3(C, T)$ such that for any*

$s \in [0,T]$ the measure μ_s is absolutely continuous with respect to the Lebesgue measure with a density which satisfies $\|\mu\|_\infty \leq C_3$ and with support contained in $B(0, C_3)$. Moreover μ_s is Lipschitz with constant C_1, that is:

$$\boldsymbol{d}_1(\mu_{s'}, \mu_s) \leq C_1|s' - s| \qquad \forall t \leq s \leq s' \leq T. \tag{2.42}$$

Proof. First we prove the lipschitzianity of μ_s. Let $t \leq s \leq s' \leq T$ and consider consider the measure π on $\mathbb{R}^{2d}$ defined as follows:

$$\pi = (x \mapsto \Phi(x, 0, s'), x \mapsto \Phi(x, 0, s))\#m_0.$$

Since $\pi \in \prod(\mu_{s'}, \mu_s)$ by construction, from the dual definition of Kantorovich-Rubinstein distance we have:

$$\begin{aligned}
\mathbf{d}_1(\mu_{s'}, \mu_s) &\leq \int_{\mathbb{R}^{2d}} |x - y| \ \pi(\mathrm{d}x, \mathrm{d}y) \\
&\overset{(2.41)}{=} \int_{\mathbb{R}^d} |\Phi(x, 0, s') - \Phi(x, 0, s)| \ \mathrm{d}m_0(x) \\
&\overset{\text{Prop. 2.27.}iii}{\leq} \|D_x u\|_\infty |s' - s| \overset{\text{Th. 2.19}}{\leq} C_1|s' - s|.
\end{aligned}$$

Now, since $\mathrm{Spt}(m_0) \subseteq B(0, C)$ and $\mathbf{d}_1(\mu_s, m_0) \leq C_1 s \leq C_1 T$, we have that $\mathrm{Spt}(\mu_s) \subseteq B(0, C + C_1 T)$ for any $s \in [0, T]$.

Now we fix $s \in [0, T]$. By Lemma 2.28 we know that the map $x \mapsto \Phi(x, 0, s)$ has an inverse Ψ. Therefore, for any borel set $E \subseteq \mathbb{R}^d$ it holds $\mu_s(E) = m_0(\Phi^{-1}(E, 0, s)) = m_0(\Psi(E))$. Since m_0 is absolutely continuous with respect to the Lebesgue measure $\mathscr{L}$ on $\mathbb{R}^d$ and using that Ψ is C_2-Lipschitz continuous we have:

$$\mu_s(E) = m_0(\Psi(E)) \leq \|m_0\|_\infty \mathscr{L}(\Psi(E)) \leq C\mathscr{L}(C_2 E) = CC_2\mathscr{L}(E). \tag{2.43}$$

Hence $\mu_s(E)$ is absolutely continuous with respect to the Lebesgue measure for any $s \in [0, T]$ with a density $\mu_s(x)$ which satisfies:

$$\|\mu_s\|_\infty \leq CC_2.$$

In fact if it existed some $\bar{x}$ such that $\mu_s(\bar{x}) > CC_2$ we could find a borel set $E_{\bar{x}}$ such that $\mu_s(x) > CC_2$ for any $x \in E_{\bar{x}}$. And this would be in contradiction with (2.43).

By taking $C_3 := \max\{C + C_1T, CC_2\}$ we get the thesis. $\square$

Definition 14 (Weak solution)**.** $m \in L^1([0,T], \mathcal{P}_1)$ is a weak solution of (2.40) if for any $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d \times [0,T[)$ test function we have:

$$\int_{\mathbb{R}^d} \varphi(x,0)\ \mathrm{d}m_0(x) + \int_0^T \int_{\mathbb{R}^d} [\partial_s\varphi(x,s) - \langle D_x\varphi(x,s), D_xu(x,s)\rangle]\ \mathrm{d}m_s(x)\ \mathrm{d}s = 0. \tag{2.44}$$

Lemma 2.31. *The map $s \mapsto \mu(s) := \Phi(\cdot,0,s)\#m_0$ is a weak solution of (2.40).*

Proof. Let $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d \times [0,T[)$, in particular

$$\varphi(x,T) = 0 \qquad \forall x \in \mathbb{R}^d. \tag{2.45}$$

Since $s \mapsto \mu_s$ is Lipschitz continuous then it is differentiable almost everywhere; so by the fundamental theorem of calculus the map $s \mapsto \int_{\mathbb{R}^d} \phi(x,s)\mu(x,s)\ \mathrm{d}s$ is absolutely continuous. Writing Φ instead of $\Phi(x,0,s)$ we have:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}s}\int_{\mathbb{R}^d} \varphi(x,s)\mu(x,s)\ \mathrm{d}s \overset{(2.41)}{=} \frac{\mathrm{d}}{\mathrm{d}s}\int_{\mathbb{R}^d} \varphi(\Phi,s)m_0(x)\ \mathrm{d}x \\
&= \int_{\mathbb{R}^d} [\partial_s\varphi(\Phi,s) + \langle D_x\varphi(\Phi,s), \partial_s\Phi\rangle]\, m_0(x)\ \mathrm{d}x \\
&\overset{\text{Prop. 2.27.}ii}{=} \int_{\mathbb{R}^d} [\partial_s\varphi(\Phi,s) - \langle D_x\varphi(\Phi,s), D_xu(\Phi,s)\rangle]\, m_0(x)\ \mathrm{d}x \\
&\overset{(2.41)}{=} \int_{\mathbb{R}^d} [\partial_s\varphi(y,s) - \langle D_x\varphi(y,s), D_xu(y,s)\rangle]\ \mathrm{d}\mu_s(y).
\end{aligned}$$

By integrating the above equality between 0 and T and using (2.45) we get (2.44). $\square$

Now we want to show the uniqueness of the weak solution, where the problem is in the discontinuity of D_xu. Hence, we first prove this fact when D_xu is locally Lipschitz continuous.

Lemma 2.32. *Let $b \in L^\infty(\mathbb{R}^d \times (0,T); \mathbb{R}^d)$ such that for any $R > 0$ and for almost all $t \in [0,T]$ there exists $L = L(R)$ with $b(\cdot,t)$ is L-Lipschitz continuous on $B(0,R)$. Then the continuity equation*

$$\begin{cases} \partial_t\mu(x,s) + div_x(b(x,s)\mu(x,s)) = 0 & \text{in } \mathbb{R}^d \times (0,T) \\ \mu(x,0) = m_0(x) \end{cases} \tag{2.46}$$

has a unique weak solution $\mu(t) = \Phi(\cdot,t)\#m_0$, where Φ is the flow of the differential equation

$$\begin{cases} \dot{y}(s) = b(y(s),s) & \textit{in } \mathbb{R} \\ y(0) = x \end{cases} \tag{2.47}$$

Proof. Note that the boundedness and the local lipschitzianity in space of b are the standard requirements to guarantee global existence and uniqueness of the solution of (2.47) in $\mathbb{R}$. Thanks to the the semigroup property (2.38) this implies that the function $\Phi(\cdot,t)$ is surjective on $\mathbb{R}^d$ for any $t \in \mathbb{R}$, in particular for any $t \in [0,T]$. In particular, if $\Psi(t,x) = \Phi(-t,x)$ is the locally Lipschitz continuous inverse of the map $x \mapsto \Phi(t,x)$ as in Lemma 2.28, for any $x \in \mathbb{R}^d$ it holds

$$x = \Psi(\Phi(x,s),s) \qquad \forall s \in \mathbb{R}. \tag{2.48}$$

The proof of the fact that $\mu(t) = \Phi(\cdot,t)\#m_0$ is weak solution of (2.46) is the same as the previous lemma. So it remains to prove the uniqueness.

Let μ be a solution of (2.46) and $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$, then $w(x,t) := \varphi(\Psi(t,x)) \in \mathcal{C}_c^0(\mathbb{R}^{d+1})$ by construction. Now, for almost all $(x,t) \in \mathbb{R}^{d+1}$:

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}t}\varphi(x) &\overset{2.48}{=} \frac{\mathrm{d}}{\mathrm{d}t} w(\Phi(x,t),t) \\ &\overset{(2.47)}{=} \partial_t w(\Phi(x,t),t) + \langle D_x w(\Phi(x,t),t), b(\Phi(x,t),t)\rangle. \end{aligned}$$

Since $x \mapsto \Phi(x,t)$ is surjective this implies that

$$\partial_t w(y,t) + \langle D_x w(y,t), b(y,t)\rangle = 0 \qquad \text{for almost all } (t,y) \in \mathbb{R}^{d+1}. \tag{2.49}$$

Moreover $\mathrm{div}_x(b\mu w) = \mathrm{div}_x(b\mu)w + \langle D_x w, b\mu\rangle$. Therefore:

$$-w\mathrm{div}_x(b\mu) = \langle D_x w, b\rangle\mu - \mathrm{div}_x(b\mu w). \tag{2.50}$$

Furthermore, since w has compact support contained in some $K \subset \mathbb{R}^{d+1}$, by the divergence theorem we have:

$$\int_{\mathbb{R}^d} \mathrm{div}_x(b\mu w)\ \mathrm{d}y = \int_{K \supset \mathrm{Spt} w} \mathrm{div}_x(b\mu w)\ \mathrm{d}y = \int_{\partial K \supset \partial \mathrm{Spt} w} b\mu w \cdot \nu\ \mathrm{d}\sigma(y) = 0. \tag{2.51}$$

Hence we have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} w(y,t)\mu(y,t)\ \mathrm{d}y = \int_{\mathbb{R}^d}[\mu\partial_t w + w\partial_t\mu]\ \mathrm{d}y$$
$$\overset{(2.46)}{=} \int_{\mathbb{R}^d}[\mu\partial_t w - w\ \mathrm{div}_y(b\mu)]\ \mathrm{d}y \overset{(2.50)+(2.51)}{=} \int_{\mathbb{R}^d}[\partial_t w + \langle D_y w, b\rangle]\mu(y)\ \mathrm{d}y \overset{(2.49)}{=} 0.$$

So the function $t \mapsto \int_{\mathbb{R}^d} w(y,t)\mu(y,t)\mathrm{d}y$ is constant. Therefore:

$$\int_{\mathbb{R}^d} w(y,t)\ \mathrm{d}\mu_t(y) = \int_{\mathbb{R}^d} w(y,0)\ \mathrm{d}\mu_0(y)$$
$$= \int_{\mathbb{R}^d}\varphi(y)\ \mathrm{d}m_0(y) \overset{(2.48)}{=} \int_{\mathbb{R}^d} w(\Phi(y,t),t)\ \mathrm{d}m_0(y).$$

Since φ is arbitrary, changing it and consequently w, we deduce that $\forall\psi \in \mathcal{C}^0_c(\mathbb{R}^{d+1})$:

$$\int_{\mathbb{R}^d}\psi(y,t)\ \mathrm{d}\mu_t(y) = \int_{\mathbb{R}^d}\psi(\Phi(y,t),t)\ \mathrm{d}m_0(y) \overset{(2.41)}{=} \int_{\mathbb{R}^d}\psi(y,t)\ \mathrm{d}(\Phi(\cdot,t)\#m_0)(y) \tag{2.52}$$

Now, for any borel set $B \in \mathbb{R}^d$ we can build a sequence of functions $(\psi_n) \subset \mathcal{C}^0_c(K)$ with $B \subseteq K \subseteq \mathbb{R}^{d+1}$ which converges to indicator function of B. By the dominated convergence theorem:

$$\mu_t(B) = \int_B \mathrm{d}\mu_t(y) = \lim_n \int_K \psi_n(y,t)\ \mathrm{d}\mu_t(y)$$
$$\overset{(2.52)}{=} \lim_n \int_K \psi_n(y,t)\ \mathrm{d}(\Phi(\cdot,t)\#m_0)(y) = \int_B \mathrm{d}(\Phi(\cdot,t)\#m_0)(y) = \Phi(\cdot,t)\#m_0(B)$$

We can conclude $\mu_t = \Phi(\cdot,t)\#m_0$ by the arbitrariness of B. □

Now we come back to the continuity equation (2.40) and let μ a weak solution. If we regularize through a nonnegative smooth kernel ρ^ε of $\mathbb{R}^d$ (for instance the Gaussian kernel) we can define the following

$$\mu^\varepsilon(x,t) := \mu(x,t) * \rho^\varepsilon(x) \qquad \text{and} \qquad b^\varepsilon(x,t) := -\frac{(D_x u(x,t)\mu(x,t)) * \rho^\varepsilon(x)}{\mu^\varepsilon(x,t)}.$$

Note that b^ε satisfies the hypotheses of Lemma 2.32 for any $\varepsilon > 0$: it is locally Lipschitz continuous because it is $\mathcal{C}^\infty$ and it is also bounded by C_1:

$$|b^\varepsilon| = \frac{\left|\int_{\mathbb{R}^d} D_x u(v,t)\mu(v,t)\rho^\varepsilon(x-v,t)\ \mathrm{d}v\right|}{|\mu^\varepsilon|} \leq \frac{\int_{\mathbb{R}^d} |D_x u(v,t)|\mu(v,t)\rho^\varepsilon(x-v,t)\ \mathrm{d}v}{\mu^\varepsilon}$$
$$\overset{(2.28)}{\leq} C_1 \frac{\int_{\mathbb{R}^d} \mu(v,t)\rho^\varepsilon(x-v,t)\ \mathrm{d}v}{\mu^\varepsilon} = C_1 \frac{\mu^\varepsilon}{\mu^\varepsilon} = C_1.$$

By the differentiation properties of the convolution we have:

$$\partial_t \mu^\varepsilon = \partial_t(\mu * \rho^\varepsilon) = \partial_t \mu * \rho^\varepsilon$$
$$\mathrm{div}(b^\varepsilon \mu^\varepsilon) = \mathrm{div}(-D_x u\ \mu * \rho^\varepsilon) = -\mathrm{div}(D_x u\ \mu) * \rho^\varepsilon.$$

Then it immediately follows that $\mu^\varepsilon(x,t)$ solves for any $\varepsilon > 0$

$$\begin{cases} \partial_t \mu^\varepsilon + \mathrm{div}(b^\varepsilon \mu^\varepsilon) = 0 \\ \mu^\varepsilon(x,0) = m^\varepsilon(x) := m_0 * \rho^\varepsilon \end{cases} \tag{2.53}$$

Hence, by Lemma 2.32, $\mu_t^\varepsilon = \Phi^\varepsilon(\cdot,t)\#m^\varepsilon$ is the unique weak solution of (2.53), where Φ^ε is the flow of the differential equation

$$\begin{cases} \partial_s \Phi^\varepsilon(x,s) = b^\varepsilon(\Phi^\varepsilon(x,s),s) \\ \Phi^\varepsilon(x,0) = x. \end{cases} \tag{2.54}$$

Now we would let $\varepsilon \to 0$ and conclude that consequently $\mu_t = \Phi(\cdot,t)\#m_0$ is the unique solution of (2.40), but this transition to the limit is very touchy and we will adopt the following strategy.

Idea. Let $\Gamma_T := \mathcal{C}^0([0,T];\mathbb{R}^d)$. We will define a family η^ε of measures on $\mathbb{R}^d \times \Gamma_T$ such that:

- If $(e_t)_{t\in[0,T]}$ is the family of the evaluation map defined by $e_t(\gamma) = \gamma(t)$ for any $\gamma \in \Gamma_T$ and for any $t \in [0,T]$, then it holds:

$$\int_{\mathbb{R}^d \times \Gamma_T} f(e_t(\gamma))\ \mathrm{d}\eta^\varepsilon(x,\gamma) = \int_{\mathbb{R}^d} f(x)\ \mathrm{d}\mu_t^\varepsilon(x) \qquad \forall f \in \mathcal{C}_0^b(\mathbb{R}^d). \tag{2.55}$$

- $(\eta^\varepsilon)_{\varepsilon>0}$ is tight so that it has a subsequence which narrowly converges to some measure η on $\mathbb{R}^d \times \Gamma_T$ which we can be disintegrated in the following way: $\mathrm{d}\eta(x,\gamma) = \mathrm{d}\eta_x(\gamma)\mathrm{d}m_0(x)$.

Then we will show the superposition principle:

$$\int_{\mathbb{R}^d\times\Gamma_T} \left|\gamma(t) - x + \int_0^t D_x u(\gamma(s), s)\ \mathrm{d}s\right|\ \mathrm{d}\eta(x,\gamma) = 0,$$

from which we will deduce that m_0-a.e. η_x-a.e. γ is given by $\Phi(x, 0, \cdot)$.

Finally, letting $\varepsilon \to 0$ in (2.55) and using Lemma 2.29 we will conclude that $\mu_t = \Phi(\cdot, 0, t)\#m_0$, that is the solution of (2.40) is unique.

Definition 15. Let η^ε the measure on $\mathbb{R}^d \times \Gamma_T$ defined by:

$$\int_{\mathbb{R}^d\times\Gamma_T} \varphi(x,\gamma)\ \mathrm{d}\eta^\varepsilon(x,\gamma) = \int_{\mathbb{R}^d} \varphi(x, \Phi^\varepsilon(x,\cdot))\ \mathrm{d}m^\varepsilon(x) \qquad \forall\varphi \in \mathcal{C}^0_b(\mathbb{R}^d \times \Gamma_T). \tag{2.56}$$

Applying the previous definition with the family of the evaluation maps $(e_t)_{t\in[0,T]} : \Gamma_T \to \mathbb{R}$ defined by $e_t(\gamma) = \gamma(t)$, we obtain (2.55). In fact, for any $f \in \mathcal{C}^b_0(\mathbb{R}^d; \mathbb{R})$ it holds:

$$\int_{\mathbb{R}^d\times\Gamma_T} f(e_t(\gamma))\ \mathrm{d}\eta^\varepsilon(x,\gamma) \overset{(2.56)}{=} \int_{\mathbb{R}^d} f(\Phi^\varepsilon(x,t))\ \mathrm{d}m^\varepsilon(x) \overset{(2.41)}{=} \int_{\mathbb{R}^d} f(x)\ \mathrm{d}\mu^\varepsilon_t(x).$$

Proposition 2.33. *The sequence $(\eta^\varepsilon)_{\varepsilon>0}$ is tight in $\mathbb{R}^d \times \Gamma_T$. That is, for any $\delta > 0$ there exists some compact $K_\delta \subset \mathbb{R}^d \times \Gamma_T$ such that $\eta^\varepsilon(K_\delta) \geq 1 - \delta$ for any $\varepsilon > 0$.*

Proof. $(m^\varepsilon)_\varepsilon$ converges to m_0 when $\varepsilon \to 0$, so it is tight and for any $\delta > 0$ there exists some compact $\tilde{K}_\delta \subset \mathbb{R}^d$ such that $m_\varepsilon(\tilde{K}_\delta) \geq 1 - \delta$ for any $\varepsilon > 0$. Now we consider

$$K_\delta := \{(x,\gamma) \in \tilde{K}_\delta \times \Gamma_T : \gamma(0) = x,\ \gamma \text{ Lipschitz with } \|\dot{\gamma}\|_\infty \leq C_1\}.$$

If $K_\Gamma := \{\gamma \in \Gamma_T : \exists x \in \tilde{K}_\delta$ such that $\gamma(0) = x,\ \gamma$ Lipschitz with $\|\dot{\gamma}\|_\infty \leq C_1\}$, then $K_\delta = \tilde{K}_\delta \times K_\Gamma$. We want to prove that K_δ is compact, so it is sufficient to show that both $\tilde{K}_\delta$ and K_Γ are: $\tilde{K}_\delta$ it is by construction, while K_Γ is precompact by Ascoli-Arzelá Theorem and clearly closed, so it is compact. In fact K_Γ is equicontinuous because uniformly Lipschitz and for any $t \in [0, T]$:

$$|\gamma(t)| \leq |x| + t \sup_t |\dot{\gamma}(t)| \leq \operatorname{diam}(\bar{K}_\delta) + C_1 T.$$

Therefore, by definition of η^ε, it holds:

$$\eta^\varepsilon(K_\delta) = m^\varepsilon(\tilde{K}_\delta) \geq 1 - \delta \qquad \forall \varepsilon > 0.$$

□

Given a set X, we recall that a sequence $(\mu)_n \in \mathcal{P}(X)$ narrowly converges to some measure $\mu \in \mathcal{P}(X)$ if and only if

$$\lim_n \int_X f \, \mathrm{d}\mu_n = \int_X f \, \mathrm{d}\mu \qquad \text{for any } f \in \mathcal{C}^0_b(X).$$

Proposition 2.34. *$(\eta^\varepsilon)_\varepsilon$ converges up to a subsequence to some measure η on $\mathbb{R}^d \times \Gamma_T$ which has m_0 as its first marginal.*

Proof. By Prokhorov Theorem (see [5, Section 8.6]) we have that there exists a subsequence still denoted by $(\eta^\varepsilon)_\varepsilon$ which narrowly converges to some η measure over $\mathbb{R}^d \times \Gamma_T$. Therefore, if we pass to the limit for $\varepsilon \to 0$ in (2.55) we get, for any $t \in [0, T]$ and for any $\varphi \in \mathcal{C}^0_b(\mathbb{R}^d; \mathbb{R})$:

$$\int_{\mathbb{R}^d \times \Gamma_T} \varphi(e_t(\gamma)) \, \mathrm{d}\eta(x, \gamma) = \int_{\mathbb{R}^d} \varphi(x) \, \mathrm{d}\mu_t(x). \tag{2.57}$$

And letting $\varepsilon \to 0$ in (2.56) we have, for any $f \in \mathcal{C}^0_b(\mathbb{R}^d; \mathbb{R})$:

$$\int_{\mathbb{R}^d \times \Gamma_T} f(x) \, \mathrm{d}\eta(x, \gamma) = \int_{\mathbb{R}^d} f(x) \, \mathrm{d}m_0(x). \tag{2.58}$$

We can conclude such in Lemma 2.32. □

Proposition 2.35. *Let $c \in \mathcal{C}^0_c(Q_T; \mathbb{R}^d)$. Then*

$$\int_{\mathbb{R}^d \times \Gamma_T} \left| \gamma(t) - x - \int_0^t c(\gamma(s), s) \; ds \right| \; d\eta(x, \gamma) \leq \int_{Q_T} |c(x,t) + D_x u(x,t)| \mu(x,t) \; dx\, dt. \tag{2.59}$$

Proof. For any $\varepsilon > 0$ small:

$$\begin{aligned} &\int_{\mathbb{R}^d \times \Gamma_T} \left| \gamma(t) - x - \int_0^t c(\gamma(s), s) \; \mathrm{d}s \right| \; \mathrm{d}\eta^\varepsilon(x, \gamma) \\ &\overset{(2.56)}{=} \int_{\mathbb{R}^d} \left| \Phi^\varepsilon(x, t) - x - \int_0^t c(\Phi^\varepsilon(x, s), s) \mathrm{d}s \right| \; \mathrm{d}m^\varepsilon(x) \end{aligned}$$

$$\overset{(2.54)}{=} \int_{\mathbb{R}^d} \left| \int_0^t [b^\varepsilon(\Phi^\varepsilon(x,s),s) - c(\Phi^\varepsilon(x,s),s)] \, \mathrm{d}s \right| \, \mathrm{d}m^\varepsilon(x)$$
$$\leq \int_0^t \int_{\mathbb{R}^d} |b^\varepsilon(\Phi^\varepsilon(x,s),s) - c(\Phi^\varepsilon(x,s),s)| \, m^\varepsilon(x) \, \mathrm{d}x\mathrm{d}s$$
$$\overset{(2.41)}{=} \int_0^t \int_{\mathbb{R}^d} |b^\varepsilon(y,s) - c(y,s)| \, \mu^\varepsilon(x,s) \, \mathrm{d}x\mathrm{d}s.$$

Now we set $c^\varepsilon := \frac{(c\mu)*\rho^\varepsilon}{\mu^\varepsilon}$. Then, by adding and removing c^ε in the last step we get:

$$\begin{aligned}\int_{\mathbb{R}^d\times\Gamma_T} \left| \gamma(t) - x - \int_0^t c(\gamma(s),s) \, \mathrm{d}s \right| \, \mathrm{d}\eta^\varepsilon(x,\gamma) \leq \\ \leq \int_0^t \int_{\mathbb{R}^d} |b^\varepsilon - c^\varepsilon| \, \mu^\varepsilon(x,s) \, \mathrm{d}x\mathrm{d}s + \int_0^t \int_{\mathbb{R}^d} |c^\varepsilon - c| \, \mu^\varepsilon(x,s) \, \mathrm{d}x\mathrm{d}s.\end{aligned} \tag{2.60}$$

The left term in (2.60) converges to the left term in (2.59), while the rightmost term in (2.60) goes to 0 because c is continuous, so that $c^\varepsilon \rightrightarrows c$. So we miss to analyze $\int_0^t \int_{\mathbb{R}^d} |b^\varepsilon - c^\varepsilon| \, \mu^\varepsilon(x,s)\mathrm{d}x\mathrm{d}s$; to do this we need the fact that

$$\|f * \rho^\varepsilon\|_{L^p} \leq \|f\|_{L^p} \qquad \forall f \in L^p(\mathbb{R}^d) \text{ and } \forall p \geq 1. \tag{2.61}$$

Hence

$$\int_0^t \int_{\mathbb{R}^d} |b^\varepsilon - c^\varepsilon| \, \mu^\varepsilon(x,s) \, \mathrm{d}x\mathrm{d}s = \int_0^t \int_{\mathbb{R}^d} |-(D_x u \; \mu) * \rho^\varepsilon - (c\mu) * \rho^\varepsilon| \, \mathrm{d}x\mathrm{d}s$$
$$\int_0^t \int_{\mathbb{R}^d} |(\mu D_x u + \mu c) * \rho^\varepsilon| \, \mathrm{d}x\mathrm{d}s \overset{(2.61)}{\leq} \int_0^t \int_{\mathbb{R}^d} |D_x u \; \mu + c\mu| \, \mathrm{d}x\mathrm{d}s$$
$$\leq \int_0^T \int_{\mathbb{R}^d} |c(x,t) + D_x u(x,t)| \mu(x,t) \, \mathrm{d}x\mathrm{d}t.$$

□

Lemma 2.36 (Superposition principle)**.** *For any $t \in [0,T]$ it holds*

$$\int_{\mathbb{R}^d\times\Gamma_T} \left| \gamma(t) - x + \int_0^t D_x u(\gamma(s),s) \; ds \right| \, d\eta(x,\gamma) = 0. \tag{2.62}$$

Proof. Let $(c_n)_{n\in\mathbb{N}}$ s sequence of uniformly bounded and continuous vector fields which converges almost everywhere to $-D_x u$. Replacing c by c_n in (2.59) and letting $n \to \infty$ we obtain the thesis thanks to the dominated convergence theorem. □

Lemma 2.36 states that density of η is concentrated on solutions of the differential equation

$$\begin{cases} \dot{y}(s) = -D_x u(y(s), s) \\ y(0) = x. \end{cases}$$

Let us state a fundamental disintegration Theorem for measures. For the proof see [14, Ch. III].

Theorem (Disintegration). *Let X, Y be Radon separable metric spaces, $\mu \in \mathcal{P}(X)$, $\pi : X \to Y$ a Borel map and let $\nu = \pi \# \mu \in \mathcal{P}(Y)$. Then there exists a $\nu -$ a.e. uniquely determined Borel family of probability measures $(\mu_y)_{y \in Y} \subset \mathcal{P}(X)$ such that*

$$\mu_y(X \setminus \pi^{-1}(y)) = 0 \qquad \text{for } \nu\text{-a.e. } y \in Y.$$

and

$$\int_X f(x)\ d\mu(x) = \int_Y \left(\int_{\pi^{-1}(y)} f(x)\ d\mu_y(x) \right)\ d\nu(y) \tag{2.63}$$

for every Borel map $f : X \to [0, +\infty]$.

In particular we can disintegrate η with respect to its first marginal m_0; in fact $m_0 = \pi \# \eta$, where $\pi : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$ is the canonical projection. So we get $\mathrm{d}\eta(x, \gamma) = \mathrm{d}\eta_x(\gamma)\mathrm{d}m_0$. Then, from (2.62), we have for any $t \in [0, T]$:

$$\begin{aligned} 0 &= \int_{\mathbb{R}^d \times \Gamma_T} \left| \gamma(t) - x + \int_0^t D_x u(\gamma(s), s)\ \mathrm{d}s \right|\ \mathrm{d}\eta(x, \gamma) \\ &\overset{(2.63)}{=} \int_{\mathbb{R}^d} \left[\int_{\Gamma_T} \left| \gamma(t) - x + \int_0^t D_x u(\gamma(s), s)\ \mathrm{d}s \right|\ \mathrm{d}\eta_x(\gamma) \right]\ \mathrm{d}m_0(x). \end{aligned}$$

This implies that for m_0-a.e. $x \in \mathbb{R}^d$, η_x-a.e. γ is a solution of

$$\begin{cases} \dot{\gamma}(s) = -D_x u(\gamma(s), s) \qquad s \in [0, T] \\ y(0) = x. \end{cases}$$

But u is Lipschitz, so $u(\cdot, 0)$ is differentiable for almost all $x \in \mathbb{R}^d$, and we know by Lemma 2.26 that the equation above has a unique solution given by

$\Phi(x,0,\cdot)$. Hence for m_0-a.e. $x \in \mathbb{R}^d$, η_x-a.e. γ is given by $\Phi(x,0,\cdot)$, and using it in (2.57) we get that for any $\varphi \in \mathcal{C}_b^0(\mathbb{R}^d;\mathbb{R})$ and for any $t \in [0,T]$:

$$\int_{\mathbb{R}^d} \varphi(x)\mu(x,t)\,\mathrm{d}x = \int_{\Gamma_T \times \mathbb{R}^d} \varphi(e_t(\gamma))\,\mathrm{d}\eta(x,\gamma) = \int_{\mathbb{R}^d}\int_{\Gamma_T} \varphi(e_t(\gamma))\,\mathrm{d}\eta_x(\gamma)\mathrm{d}m_0(x)$$
$$= \int_{\mathbb{R}^d} \varphi(\Phi(x,0,t))\,\mathrm{d}m_0(x) \overset{(2.41)}{=} \int_{\mathbb{R}^d} \varphi(x)\,\mathrm{d}(\Phi(\cdot,0,t)\#m_0)(x).$$

Similarly to what we have done in the end of Lemma 2.32 we can conclude that $\mu_t = \Phi(\cdot,0,t)\#m_0$. Hence we have proved the following theorem.

Theorem 2.37. *Given a solution u of (2.23) and under hypotheses (2.27) the map $s \mapsto \mu_s := \Phi(\cdot,0,s)\#m_0$ is the unique weak solution of (2.40).*

2.5 Existence and uniqueness for MFEs

Let us come back to the study of the system (2.1). We note that F and G depend on t through m, so that (2.2) directly implies (2.27). Though, just because of this dependence on m, what we have done so far is not sufficient to prove the existence of a solution of (2.1). In fact, given a map $m \in \mathcal{C}([0,T];\mathcal{P}_1)$ such that $m(0) = m_0$ and considered the solution u of

$$\begin{cases} -\partial_t u + \frac{1}{2}|D_x u|^2 = F(x,m(t)) \\ u(x,T) = G(x,m(T)), \end{cases}$$

then the solution μ of

$$\begin{cases} \partial_t \mu - \mathrm{div}_x(D_x u\ \mu) = 0 \\ \mu(0,x) = m_0(x) \end{cases}$$

could not coincide with the initial m, so that (2.1) would not have solutions. So it is clear that we need a sort of fixed point theorem which connects (2.23) with (2.40), but before we need a stability lemma for the system.

Lemma 2.38 (Stability)**.** *Let $(m_n)_{n\in\mathbb{N}}$ a sequence in $\mathcal{C}([0,T];\mathcal{P}_1)$ which uniformly converges to $m \in \mathcal{C}([0,T];\mathcal{P}_1)$. Let $(u_n)_{n\in\mathbb{N}}$ the sequence of solutions* *of*

$$\begin{cases} -\partial_t u_n + \frac{1}{2}|D_x u_n|^2 = F(x, m_n) \\ u_n(x,T) = G(x, m_n(T)) \end{cases}$$

and u the solution of

$$\begin{cases} -\partial_t u + \frac{1}{2}|D_x u|^2 = F(x, m) \\ u(x,T) = G(x, m(T)). \end{cases}$$

Furthermore let Φ_n (respectively Φ) the flow associated to u_n (respectively to u) and let us set $\mu_n(s) = \Phi_n(\cdot, 0, s)\#m_0$ and $\mu(s) = \Phi(\cdot, 0, s)\#m_0$.

Then the sequence $(u_n)_n$ locally uniformly converges to u in Q_T while $(\mu_n)_n$ converges to μ in $\mathcal{C}([0,T]; \mathcal{P}_1)$.

Proof. From our assumptions of continuity on F and G we have that $(F(x, m_n))_n$ and $(G(x, m_n))_n$ locally uniformly converges to $F(x, m)$ and $G(x, m)$ respectively. Hence the locally uniformly convergence of $(u_n)_n$ to u is a directly consequence of the stability of viscosity solutions. Furthermore the $(u_n)_{n\in\mathbb{N}}$ are uniformly bounded:

$$\begin{aligned} |u_n(x,t)| &\leq \int_t^T \left[\frac{1}{2}|\alpha_n^*|^2 + |F(x, m(s))|\right] \, \mathrm{d}s + |G(x, m(T))| \\ &\overset{\text{Cor. 2.18+(2.2)}}{\leq} T(M + C) + C \end{aligned}$$

and also uniformly semiconcave since in (2.28) the semiconcave constant C_1 depends only on C and T. Consequently we can apply Theorem 2.9 and deduce $D_x u_n$ converges to $D_x u$ almost everywhere.

Moreover $\|D_x u_n\|_\infty \leq C_1$ for any $n \in \mathbb{N}$ because this estimate depends only on C and T and consequently by Lemma 2.30 we know that the $(\mu_n)_{n\in\mathbb{N}}$ have support in $K := \overline{B(0, C_3)}$, are uniformly Lipschitz of constant C_1 (hence equicontinuous) and with densities uniformly bounded by C_3. Therefore by Ascoli-Arzelá Theorem and Banach-Alaoglu Theorem $(\mu_n)_{n\in\mathbb{N}}$ has a subsequence which locally uniformly and L^∞-weakly* converges to some $\bar{m}$ with support in $K \times [0,T]$ and which belongs to $L^\infty(Q_T)$ and to $\mathcal{C}([0,T]; \mathcal{P}_1(K))$.

But μ_n solves the continuity equation for u_n for any $n \in \mathbb{N}$ and passing to the limit in (2.44) thanks to dominated convergence Theorem we deduce that $\bar{m}$ solves it for u. By the uniqueness of solution of (2.40) stated in Theorem

2.37 we have that $\mu = \bar{m}$. □

Remark 2.39. In this proof we have used the so called Ascoli-Arzelá Theorem and a classical formulation of it is for an equicontinuous and uniformly bounded family of functions in $\mathcal{C}^0(X;\mathbb{R})$, where X is a compact metric space. Though, if you see the proof in [29], you can see that the only properties of $\mathbb{R}$ which are used are the completeness and the Bolzano's Theorem for bounded sequences. Hence, if we consider an equicontinuous and uniformly bounded family of functions in $\mathcal{C}^0(X;Y)$ with Y compact (and therefore complete), everything still works.

So we can apply it to $(\mu_n) \subset \mathcal{C}([0,T];\mathcal{P}_1(K))$ since we have just seen that $(\mathcal{P}_1(K), \mathbf{d}_1)$ is a compact metric space because $K := \overline{B(0,C_3)}$ is compact for the weak-* topology.

Theorem 2.40 (Existence of solutions of the First order mean field system)**.** *Under the three hypotheses on m_0, F and G of page 31 , there exists at least one solution of (2.1).*

Proof. Let $\mathcal{M} := \{m \in \mathcal{C}([0,T];\mathcal{P}_1)\colon\ m(0) = m_0\}$, which is clearly convex. For any $m \in \mathcal{M}$ we can consider the unique solution u_m of

$$\begin{cases} -\partial_t u_m + \frac{1}{2}|D_x u_m|^2 = F(x, m(t)) \\ u_m(x,T) = G(x, m(T)), \end{cases}$$

and then the unique solution μ_n of

$$\begin{cases} \partial_t \mu_m - \mathrm{div}_x(D_x u_m\ \mu_m) = 0 \\ \mu_m(0) = m_0 \end{cases}$$

So we can consider the operator $\mathcal{T} : \mathcal{M} \to \mathcal{M}$ defined as $\mathcal{T}(m) = \mu_m$, which is continuous from what we have seen in Lemma (2.38). So to apply the Schauder-Tychonoff fixed point Theorem (see [15]) it remains to show that $\mathcal{T}(\mathcal{M})$ is contained into a compact subset of $\mathcal{M}$. In fact $(\mu_m)_{m\in\mathcal{M}}$ is relatively compact in $\mathcal{M}$ for the Ascoli-Arzelá Theorem because by Lemma 2.30 they are uniformly Lipschitz (and hence equicontinuous) of constant C_1 and are uniformly bounded by C_3. □

Now we prove a uniqueness result for the system of the first order mean field games (2.1).

Theorem 2.41 (Sufficient condition for uniqueness)**.** *Under the same hypotheses of Theorem 2.40 and assuming also that*

$$\int_{\mathbb{R}^d} [F(x,m_1) - F(x,m_2)] \; d(m_1 - m_2)(x) > 0 \qquad \forall m_1 \neq m_2 \in \mathcal{P}_1 \tag{2.64}$$

$$\int_{\mathbb{R}^d} [G(x,m_1) - G(x,m_2)] \; d(m_1 - m_2)(x) \geq 0 \qquad \forall m_1, m_2 \in \mathcal{P}_1, \tag{2.65}$$

then (2.1) has a unique solution.

Proof. Let (u_1, m_1) and (u_2, m_2) two solutions of (2.1) and let us set $\bar{u} := u_1 - u_2$ and $\bar{m} := m_1 - m_2$. Then:

$$-\partial_t \bar{u} + \frac{1}{2}(|D_x u_1|^2 - |D_x u_2|^2) - (F(x,m_1) - F(x,m_2)) = 0, \tag{2.66}$$

$$\partial_t \bar{m} - \mathrm{div}_x(m_1 D_x u_1 - m_2 D_x u_2) = 0. \tag{2.67}$$

Setting $\Theta := m_1 D_x u_1 - m_2 D_x u_2$ we note that

$$\begin{cases} (\partial_t \bar{m})\bar{u} = \partial_t(\bar{m}\bar{u}) - (\partial_t \bar{u})\bar{m} \\ \bar{u} \; \mathrm{div}_x \Theta = \; \mathrm{div}(\bar{u}\Theta) - \langle D_x \bar{u}, \Theta \rangle \\ \int_{Q_T} \mathrm{div}_x(\bar{u}\Theta) \mathrm{d}x\mathrm{d}t = 0 \end{cases} \tag{2.68}$$

where in the last equality we have used the divergence theorem and the fact that m_1 and m_2 have compact support by Lemma 2.30.

Furthermore, multiplying (2.66) by $\bar{m}$ and remembering the initial condition satisfied by m_1, m_2, u_1 and u_2 we have:

$$\begin{cases} (\partial_t \bar{u})\bar{m} = \frac{\bar{m}}{2}(|D_x u_1|^2 - |D_x u_2|^2) - \bar{m}(F(x,m_1) - F(x,m_2)) \\ \bar{m}(0) = 0 \\ \bar{u}(x,T) = G(x, m_1(T)) - G(x, m_2(T)). \end{cases} \tag{2.69}$$

A convex function $L(p)$ satisfies

$$L(p_1) - L(p_2) \geq \nabla L(p_2) \cdot (p_1 - p_2) \qquad \text{for any } p_1, p_2.$$

In particular for $L(\alpha) = \frac{|\alpha|^2}{2}$ we have

$$\begin{cases} \frac{|D_x u_1|^2}{2} - \frac{|D_x u_2|^2}{2} \geq D_x u_2 \cdot (D_x u_1 - D_x u_2) \\ \frac{|D_x u_1|^2}{2} - \frac{|D_x u_2|^2}{2} \leq D_x u_1 \cdot (D_x u_1 - D_x u_2) \end{cases} \tag{2.70}$$

If we multiply the first equation in (2.70) by $-m_2$ and the second one by m_1 and then we add them up we obtain:

$$-\frac{\bar{m}}{2}(|D_x u_1|^2 - |D_x u_2|^2) + D_x \bar{u} \cdot \Theta \geq 0. \tag{2.71}$$

By the density of $\mathcal{C}_c^\infty$ in $\mathcal{C}^0$, we can use $\bar{u}$ as test function in (2.67). In fact $\bar{u}$ is Lipschitz continuous (so they are u_1 and u_2) and we can apply the dominated convergence theorem because the density of $\bar{m}$ is bounded by Lemma 2.30. Therefore we have

$$\begin{aligned} 0 &= \int_{Q_T} [\bar{u}\partial_t \bar{m} - \bar{u} \operatorname{div}_x \Theta] \, \mathrm{d}x\mathrm{d}t \\ &\overset{(2.68)}{=} \int_{\mathbb{R}^d} (\bar{m}\bar{u})(T) \, \mathrm{d}x - \int_{\mathbb{R}^d} (\bar{m}\bar{u})(0) \, \mathrm{d}x - \int_{Q_T} [(\partial_t \bar{u})\bar{m} - \langle D_x \bar{u}, \Theta \rangle] \, \mathrm{d}x\mathrm{d}t \\ &\overset{(2.69)}{=} \int_{\mathbb{R}^d} [G(x, m_1(T)) - G(x, m_2(T))] \, \mathrm{d}\bar{m}_T(x) + \\ &\quad + \int_{Q_T} \left[-\frac{\bar{m}}{2}(|D_x u_1|^2 - |D_x u_2|^2) + \bar{m}(F(x, m_1) - F(x, m_2)) + \langle D_x \bar{u}, \Theta \rangle \right] \, \mathrm{d}x\mathrm{d}t \\ &\overset{(2.65)+(2.71)}{\geq} \int_{Q_T} [F(x, m_1) - F(x, m_2)]\bar{m} \, \mathrm{d}x\mathrm{d}t. \end{aligned}$$

Hence, by assumption (2.64) we deduce $\bar{m} = 0$, so that $m_1 = m_2$. But now u_1 and u_2 solve the same differential equation, therefore $u_1 = u_2$. □

Remark 2.42. Uniqueness still holds with convex lagrangians, substituting (2.64) with

$$\int_{\mathbb{R}^d} [F(x, m_1) - F(x, m_2)] \, \mathrm{d}(m_1 - m_2)(x) \geq 0 \qquad \forall m_1 \neq m_2 \in \mathcal{P}_1.$$

In this case we assume by contradiction that there exists $(\bar{x}, \bar{t}) \in \{m_1 > 0\} \cup \{m_2 > 0\}$ such that $D_x u_1(\bar{x}, \bar{t}) \neq D_x(\bar{x}, \bar{t})$. Repeating the same proof as before, we have that (2.70) and (2.71) hold with strict inequalities because

$L(\alpha)$ is strictly convex, and consequently $\int_{Q_T}[F(x,m_1) - F(x,m_2)]\bar{m}\ \mathrm{d}x\mathrm{d}t < 0$. A contradiction. Therefore for any $(x,t) \in \{m_1 > 0\} \cup \{m_2 > 0\}$, then $D_x u_1 = D_x u_2$; hence $m_1 = m_2$ because they solve the same Kolmogorov equation which has unique solution by Theorem 2.37, consequently $u_1 = u_2$.

Remark 2.43. You can note that condition (2.64) is very similar to condition (1.23) which ensures uniqueness of solution of (1.9) in the static case. It is a monotony condition that occurs if the cost function F increases as soon as the population aggregates. Clearly it does not include every real situation, in fact uniqueness in more general case is still an open problem.

For instance it is easy to check that (2.64) holds in the following cases:

- $F(x,m) = f(x, m(x))$, where $\frac{\partial}{\partial_m} f \geq 0$.
- $F(x,m) = f(\cdot, m * \rho) * \rho$, with $\rho \in \mathcal{C}^1_c(\mathbb{R}^d)$ even and $\frac{\partial}{\partial_m} f \geq 0$.

Chapter 3

Linear quadratic mean field games

In this chapter we analyze three models of linear quadratic mean field games. Sometimes we will find explicit solutions, in other occasions we will provide existence results.

3.1 First model

We suppose:

$$
\begin{aligned}
&i. \quad J(t,x;\alpha) = \int_t^T \left[\frac{|\alpha(s)|^2}{2}\right] \, \mathrm{d}s + \underbrace{\frac{\tilde{a}}{2}|y(T)-h|^2 + \frac{\tilde{b}}{2}|y(T)-\mathbb{E}[m(T)]|^2}_{G(y(T),m(T))} \\
&ii. \quad \begin{cases} \dot{y}(s) = \alpha(s) = l_0(x,\alpha) \\ y(t) = x \end{cases} \\
&iii. \; m_0(\cdot) \text{ absolutely continuous, bounded and with compact support,}
\end{aligned}
\tag{3.1}
$$

where $\tilde{a} \in \mathbb{R}$, $\tilde{b} \in \mathbb{R}$, t, T and x are given and $y(\cdot) \in \mathbb{R}^d$. Since a typical agent wants to minimize the cost functional, if $\tilde{a} > 0$ then the population tends to aggregate around h; vice versa, if $\tilde{a} < 0$, then it tends to move away from h. The same applies to $\tilde{b}$ and $\mathbb{E}[m(T)]$.

By easy calculations, one can rewrite (3.1.i) as

$$G(y(T), m(\cdot, T)) = \frac{a}{2}|y(T)|^2 - b \cdot y(T) + c,$$

where

$$\begin{cases} a = \tilde{a} + \tilde{b} \\ b = \tilde{a}h + \tilde{b}\mathbb{E}[m(T)] \\ c = \frac{\tilde{a}}{2}|h|^2 + \frac{\tilde{b}}{2}|\mathbb{E}[m(T)]|^2. \end{cases} \tag{3.2}$$

Remark 3.1. We note that the parameter b is not known, because it depends on $\mathbb{E}[m(T)]$ and the density $m(x,t)$ is an unknown of the problem. Hence, after finding m, we will be able to determine b through a fixed point equation.

The Hamilton-Jacobi-Bellman equation linked to this control problem is

$$\begin{cases} -u_t + \frac{|D_x u|^2}{2} = 0 & \text{in } [0,T) \times \mathbb{R}^d \\ u(x,T) = \frac{a}{2}|x|^2 - b \cdot x + c. \end{cases} \tag{3.3}$$

Since the feedback control that minimizes the pre-hamiltonian function $\mathcal{H}(p,x,\alpha) = p \cdot l_0(x,\alpha) + l(x,\alpha)$ is $\alpha^* = -D_x u$, the optimal dynamics is $\dot{y}(s) = -D_x u(y(s), s)$, so that the continuity equation for the density $m(x,t)$ becomes

$$\begin{cases} m_t - \operatorname{div}_x(m D_x u) = 0 & \text{in } (0,T] \times \mathbb{R}^d \\ m(x,0) = m_0(x). \end{cases} \tag{3.4}$$

Remembering $\operatorname{div}_x(fX) = f \operatorname{div}_x(X) + D_x f \cdot X$, we find out that the mean field system of this linear quadratic model is

$$\begin{cases} -u_t + \frac{|D_x u|^2}{2} = 0 & \text{in } [0,T) \times \mathbb{R}^d \\ m_t - D_x m \cdot D_x u - m \operatorname{div}_x(D_x u) = 0 & \text{in } (0,T] \times \mathbb{R}^d \\ u(x,T) = \frac{a}{2}|x|^2 - b \cdot x + c \\ m(x,0) = m_0(x). \end{cases} \tag{3.5}$$

We first consider separately (3.3) and (3.4), then we give a solution for (3.5) after solving a fixed point equation for b.

We try to solve (3.3) with the Hopf-Lax formula

$$u(x,t) = \min_{y\in\mathbb{R}^d}\left\{(T-t)\frac{|x-y|^2}{2(T-t)^2} + \frac{a}{2}|x|^2 - b\cdot x + c\right\}.$$

The quadratic function $Q(y) = \frac{|x-y|^2}{2(T-t)} + \frac{a}{2}|x|^2 - b\cdot x + c$ has a minimum if and only if the coefficient of $|y|^2$ is positive, that is $\frac{a}{2} + \frac{1}{2(T-t)} > 0$. Hence we assume that

$$\begin{cases} a > 0 \\ t < T \end{cases} \qquad \text{or} \qquad \begin{cases} a < 0 \\ T + \frac{1}{a} < t < T. \end{cases} \tag{3.6}$$

This condition states that if $a < 0$, then system (3.5) has a meaning only for $T < -\frac{1}{a}$. Under (3.6), we get

$$u(x,t) = \frac{a|x|^2 - 2x\cdot b - (T-t)|b|^2}{2[1+a(T-t)]} \qquad \forall x,\ \forall t \text{ s.t. (3.6) holds.} \tag{3.7}$$

One can directly check that (3.7) solves (3.3).

Now we calculate $D_x u = \frac{ax-b}{1+a(T-t)} =: f(x,t)$ and $div(D_x u) = \frac{ad}{1+a(T-t)}$, so that we can write explicitly (3.4) and solve it with the method of characteristics: hence we have to suppose at least $m_0(\cdot) \in \mathcal{C}^1(\mathbb{R}^d)$. The system of ODEs that we obtain is:

$$\begin{cases} \dot{X} = -\frac{a}{1+a(T-t)}X + \frac{b}{1+a(T-t)} \\ X(0) = y \\ \dot{z} = z\ \mathrm{div}(f) \\ z(0) = m_0(y). \end{cases}$$

The solutions of this system are:

$$X(t,y) = \frac{b}{a} + \frac{1+a(T-t)}{1+aT}\left(y - \frac{b}{a}\right), \qquad z(t,y) = m_0(y)\left(\frac{1+aT}{1+a(T-t)}\right)^d$$

and we immediately note that $X(t,\cdot)$ is invertible, so that the characteristics never cross. Since

$$Y(t,x) = \frac{x(1+aT) - bt}{1+a(T-t)} \tag{3.8}$$

is the inverse of $X(t,y)$, we know that the solution of (3.4) is

$$m(x,t) = z(t, Y(t,x)) = m_0\left(\frac{x(1+aT)-bt}{1+a(T-t)}\right)\left(\frac{1+aT}{1+a(T-t)}\right)^d, \qquad (3.9)$$

for any $x \in \mathbb{R}^d$ and for any t such that (3.6) is satisfied.

Now we want to determine b. Formula (3.9) gives an expression for $m(x,T)$ and that allows to calculate $\mathbb{E}[m(T)] = \frac{1}{1+aT}(\mathbb{E}[m_0] + bT)$. Substituting it in (3.2) we get

$$\begin{cases} b = \frac{\tilde{a}(1+aT)h+\tilde{b}\mathbb{E}[m_0]}{1+\tilde{a}T} \\ c = \frac{\tilde{a}}{2}|h|^2 + \frac{\tilde{b}}{2(1+aT)^2}|\mathbb{E}[m_0]+bT|^2. \end{cases} \qquad (3.10)$$

By the verification theorem of the dynamic programming, $u(x,t) =$ (3.7) is the value function of the optimal control (3.1), while $m(x,t) =$ (3.9) is the density of the optimizing agents. In conclusion, we have proved the following result.

Proposition 3.2. *Under assumption (3.6), if b and c are given by (3.10), then $u(x,t) =$ (3.7) and $m(x,t) =$ (3.9) solve the mean filed games system (3.5) linked to the mean field game model (3.1).*

3.2 Second model

In this second model we assume:

$$\begin{aligned} &i. \quad J(t,x;\alpha) = \int_t^T \left[\frac{|\alpha(s)|^2}{2}\right] \mathrm{d}s + \frac{\tilde{a}}{2}|y(T)-h|^2 + \frac{\tilde{b}}{2}|y(T) - \mathbb{E}[m(T)]|^2 \\ &ii. \quad \begin{cases} \dot{y}(s) = Ay(s) + B\alpha(s) \\ y(t) = x \end{cases} \\ &iii. \; m_0 \text{ absolutely continuous, bounded and with compact support,} \end{aligned} \qquad (3.11)$$

where $y(\cdot) \in \mathbb{R}^d$, $\alpha(\cdot) \in \mathbb{R}^k$ and $A \in \mathbb{R}^{d\times d}$, $B \in \mathbb{R}^{d\times k}$, t, T, x, $\tilde{a}$ and $\tilde{b}$ are given. In particular, model (3.11) is equal to model (3.1) except for the dynamics. We will study first the case $d = 1$ and $k = 1$, where the solutions are explicit, then we will provide an existence Theorem for the general case.

Let us start with the dimensions $d = 1$ and $k = 1$. Then the Hamilton-Jacobi-Bellman equation of the control problem is

$$\begin{cases} -u_t - Ax\ u_x + \frac{B^2}{2}u_x^2 = 0 & (x,t) \in \mathbb{R} \times [0,T) \\ u(x,T) = \frac{a}{2}x^2 - bx + c, \end{cases} \tag{3.12}$$

where a, b and c are given by (3.2). Since the feedback control which minimizes the pre-hamiltonian function is $\alpha^*(s) = -Bu_x$, then we can deduce the continuity equation for the density $m(x,t)$ of the agents:

$$\begin{cases} m_t + \ \mathrm{div}_x(mf) = 0 & (x,t) \in \mathbb{R} \times (0,T] \\ m(x,0) = m_0(x), \end{cases} \tag{3.13}$$

where $f(x,t) = Ax - B^2 u_x$.

So the mean field system of model (3.11) is

$$\begin{cases} -u_t - Ax\ u_x + \dfrac{B^2}{2}u_x^2 = 0 & (x,t) \in \mathbb{R} \times [0,T) \\ u(x,T) = \dfrac{a}{2}x^2 - bx + c \\ m_t + \ \mathrm{div}_x\big(m(Ax - B^2 u_x)\big) = 0 & (x,t) \in \mathbb{R} \times (0,T] \\ m(x,0) = m_0(x). \end{cases} \tag{3.14}$$

As in the previous section, we will analyze separately (3.12) and (3.13), and then we will give a solution of (3.14) after solving a fixed point equation for b.

Let us make the *ansatz* that the solution of (3.12) is of the form

$$u(x,t) = h(t)x^2 + l(t)x + s(t),$$

where $h(\cdot)$, $l(\cdot)$ and $s(\cdot)$ are unknown function that by construction satisfy the terminal condition

$$\begin{cases} h(T) = \frac{a}{2} \\ l(T) = -b \\ s(T) = c. \end{cases}$$

Substituting $u_x = 2hx + l$ and $u_t = \dot{h}x^2 + \dot{l}x + \dot{s}$ in (3.12), we find out that such a $u(x,t)$ solves (3.12) if and only if the functions h, l, s satisfy

$$\begin{cases} \dot{h} = 2B^2h^2 - 2Ah; & h(T) = \frac{a}{2} \\ \dot{l} = (2B^2h - A)l; & l(T) = -b \\ \dot{s} = \frac{B^2}{2}l^2; & s(T) = c. \end{cases} \tag{3.15}$$

Through the substitution $w(t) = \frac{1}{h(t)}$ we can transform the first Riccati equation in (3.15) in a linear ODE in the variable w, so that we obtain

$$h(t) = \frac{aA}{aB^2 + (2A - aB^2)e^{2A(t-T)}}, \tag{3.16}$$

for any t such that $aB^2 + [2A - aB^2]e^{2A(t-T)} \neq 0$. So we have to assume

$$\begin{cases} \delta := \frac{1}{2A} \log\left(\frac{aB^2}{aB^2-2A}\right) > 0 \\ t < T \end{cases} \quad \text{or} \quad \begin{cases} \delta < 0 \\ T + \delta < t < T. \end{cases} \tag{3.17}$$

This is the analogue of condition (3.6) for model (3.1).

Now we note that we can rewrite the first one in (3.15) as follows

$$\frac{\dot{h}}{h} = (2B^2h - A) - A,$$

so that the second one becomes

$$\frac{\dot{l}}{l} = \frac{\dot{h}}{h} + A \iff (\dot{\log l}) = (\dot{\log h}) + A.$$

By integration we get

$$l(t) = \frac{-2Abe^{A(t-T)}}{aB^2 + (2A - aB^2)e^{2A(t-T)}}. \tag{3.18}$$

Finally we find by integration

$$s(t) = \frac{AB^2b^2}{aB^2 - 2A} \frac{1}{aB^2 + (2A - aB^2)e^{2A(t-T)}} + \frac{b^2B^2}{2(2A - aB^2)} + c. \tag{3.19}$$

Now we use again the method of characteristics to solve (3.13), so that we get

$$m(x,t) = m_0(Y(t,x))[aB^2 + (2A - aB^2)e^{-2AT}]\frac{e^{At}}{aB^2 + (2A - aB^2)e^{2A(t-T)}}, \tag{3.20}$$

where

$$Y(t,x) = \left(x + \frac{bB^2}{(2A - aB^2)e^{A(t-T)}}\frac{2A - aB^2}{e^{2AT}}(1 - e^{2At})\right)e^{At}\frac{aB^2 + (2A - aB^2)e^{-2AT}}{aB^2 + (2A - aB^2)e^{2A(t-T)}}.$$

Hence we calculate

$$\mathbb{E}[m(T)] = \frac{\mathbb{E}[m_0]}{C} - \frac{Db}{C}$$

where

$$C = e^{AT}\frac{aB^2 + (2A - aB^2)e^{-2AT}}{2A}, \qquad D = \frac{B^2 e^{AT}}{2A}(e^{-2AT} - 1). \tag{3.21}$$

Substituting it in (3.2) and solving the fixed point equation for b, we find

$$\begin{cases} b = \frac{C\tilde{a}}{C + D\tilde{b}}h + \frac{\tilde{b}}{C + D\tilde{b}}\mathbb{E}[m_0] \\ c = \frac{\tilde{a}}{2}h^2 + \frac{\tilde{b}}{2C^2}(\mathbb{E}[m_0] - Db)^2. \end{cases} \tag{3.22}$$

Thanks to the verification theorem of the dynamic programming we have proved the following result:

Proposition 3.3. *Let us assume that $d = 1$ and $k = 1$. Furthermore we suppose that b and c are given by (3.22), C and D by (3.21) and that the functions $h(\cdot)$, $l(\cdot)$ and $s(\cdot)$ are given respectively by (3.16), (3.18) and (3.19). Then, under condition (3.17), $u(x,t) = h(t)x^2 + l(t)x + s(t)$ and $m(x,t) =$ (3.20) solve the mean filed system (3.14) linked to the mean field game model (3.11). Moreover, $u(x,t)$ is the unique solution quadratic in x because h, l and s are the unique solutions of (3.15).*

Now we analyze the general case $d \geq 1$, $k \geq 1$.

Notations. In the following we denote with A' the transpose of the matrix A, with e^A the matrix exponential of A and with $tr(A)$ the trace of the matrix A. Finally we suppose $a \neq 0$.

In this case (3.14) becomes

$$\begin{cases} -u_t - (Ax)' D_x u + \frac{1}{2}(D_x u)' S D_x u = 0 & (x,t) \in \mathbb{R}^d \times [0,T) \\ u(x,T) = \frac{a}{2}|x|^2 - b \cdot x + c & \\ m_t + \operatorname{div}_x (m(Ax - SD_x u)) = 0 & (x,t) \in \mathbb{R}^d \times (0,T] \\ m(x,0) = m_0(x). & \end{cases} \tag{3.23}$$

where $S := BB' \in Sym(d)$ and the parameters a, b and c are given by (3.2).

As in the case $d = 1$, $u(x,t) = x'H(t)x + L(t) \cdot x + s(t)$ with $H(\cdot) \in Sym(d)$ is a solution of the Hamilton-Jacobi-Bellman equation in (3.23) if and only if $H(\cdot)$, $L(\cdot)$ and $s(\cdot)$ satisfy

$$\begin{cases} \dot{H} = 2HSH - A'H - HA =: r(H); & H(T) = \frac{a}{2}\mathbb{I} \\ \dot{L} = 2HSL - A'L; & L(T) = -b \\ \dot{s} = \frac{1}{2}L'SL; & s(T) = c. \end{cases} \tag{3.24}$$

Remark 3.4. We note that $r(\cdot) \in \mathcal{C}^\infty(\mathbb{R}^{d\times d})$, then it is locally Lipschitz; hence, by the theory on ODEs, we know that the equation for H has local solution. That is, there exists some $\tilde{T} > 0$ such that for any $T \in]0, \tilde{T}[$ there exists a unique $H \in \mathcal{C}^1(]0,T[; Sym(d))$ which solves the first one in (3.24). In fact if we transpose the first equation in (3.24) we find out that H' solves the same equation of H, so that $H = H'$.

Furthermore $H(\cdot)$ is invertible in $]0,T[$ for any $T \in]0, \tilde{T}]$ because $det(H(T)) = (\frac{a}{2})^d \neq 0$. Hence, from now on we will suppose $T \leq \tilde{T}$.

Given that $H(\cdot)$ which solves the first equation in (3.24), one can directly check that

$$\begin{cases} L(t) = -\frac{2}{a}H(t)e^{A(t-T)}b \\ s(t) = c + \int_0^t \frac{1}{2}L'(r)SL(r)\ \mathrm{d}r \end{cases} \tag{3.25}$$

solve the second and the third equation of (3.24) in $]0,T[$. In fact:

$$\begin{aligned} \dot{L} &= -\frac{2}{a}[\dot{H}e^{A(t-T)} + HAe^{A(t-T)}]b \\ &= -\frac{2}{a}[2HSH - A'H]e^{A(t-T)}b = \end{aligned}$$

$$= (2HS - A')\left[-\frac{2}{a}He^{A(t-T)}b\right] = (2HS - A')L$$

Through the method of characteristics we find out that a solution in $]0,T[$ of the continuity equation in (3.23) is

$$m(x,t) = m_0(Y(t,x))e^{M(t)}, \tag{3.26}$$

where

$$\begin{cases} Y(t,x) = H^{-1}(0)e^{A't}H(t)x - \frac{2}{a}H^{-1}(0)\left(\int_0^t e^{A's}H(s)SH(s)e^{A(s-T)}\,\mathrm{d}s\right)b \\ M(t) = \int_0^t \mathrm{tr}(2SH(s) - A)\,\mathrm{d}s. \end{cases} \tag{3.27}$$

Note that Remark 3.4 states that $Y(t,x)$ is well defined for $T \leq \tilde{T}$. By (3.26), we can calculate

$$\mathbb{E}[m(T)] = C^{-1}(\mathbb{E}[m(0)] - Db),$$

where

$$\begin{cases} C = \frac{a}{2}H^{-1}(0)e^{A'T} \\ D = -\frac{2}{a}H^{-1}(0)\left(\int_0^T e^{A's}H(s)SH(s)e^{A(s-T)}\,\mathrm{d}s\right). \end{cases} \tag{3.28}$$

In fact from $1 = \int_{\mathbb{R}^d} m_T(x)\,\mathrm{d}x$ we deduce $det(C) = e^{M(T)}$. Note that C^{-1} is well defined if $T \leq \tilde{T}$ by Remark 3.4.

So if we write the fixed point equation for b obtained by (3.2), we have

$$b = \tilde{a}h + \tilde{b}C^{-1}(\mathbb{E}[m_0] - Db),$$

which has unique solution if and only if the matrix

$$Q := \mathbb{I} + \tilde{b}C^{-1}D \tag{3.29}$$

is invertible; and this is true for small times since $D(T) \underset{T\to 0}{\longrightarrow} \mathbb{O}$ and $C(T) \underset{T\to 0}{\longrightarrow} \mathbb{I}$.

So we have:

$$\begin{cases} b = Q^{-1}\left(\tilde{a}h + \tilde{b}C^{-1}\mathbb{E}[m_0]\right) \\ c = \frac{\tilde{a}}{2}|h|^2 + \frac{\tilde{b}}{2}\left|C^{-1}\mathbb{E}[m(0)] - C^{-1}Db\right|^2. \end{cases} \tag{3.30}$$

Therefore, only applying the verification theorem of the dynamic programming, we have proved the following Theorem.

Theorem 3.5. *There exists $\tau > 0$ sufficiently small such that for any $T < \tau$ the matrix $Q =$ (3.29) is invertible; furthermore, for any $T < \tau$, there exists a unique $H \in \mathcal{C}^1(]0,T[; Sym(d))$ continuous in 0 and T which solves the first equation in (3.24) in $[0,T]$ and that it is invertible for any $t \in [0,T]$. Moreover, if C and D are given by (3.28), b and c are given by (3.30) and $L(t)$ and $s(t)$ are given by (3.25), then $u(x,t) = x'H(t)x + L(t)\cdot x + s(t)$ and $m(x,t) =$ (3.26) are solutions in $\mathbb{R}^d \times (0,T)$ of the mean field games system (3.23) linked to the mean field model (3.11).*

Remark 3.6 (Consistency of the two models)**.** Model (3.1) differs from model (3.11) for the dynamics of a typical agent: in the first case we have $\dot{y} = \alpha$, while in the second one we have $\dot{y} = Ay + B\alpha$. If we consider a solution $\left(u^{(1)}, m^{(1)}\right)$ of (3.5) in the sense of Proposition 3.2 and a solution $\left(u^{(2)}, m^{(2)}\right)$ of (3.23) in the sense of Theorem 3.5, we would like to state:

$$u^{(2)}(x,t) \underset{A\to\mathbb{O},\ B\to\mathbb{I}}{\longrightarrow} u^{(1)}(x,t) \qquad \text{and} \qquad m^{(2)}(x,t) \underset{A\to\mathbb{O},\ B\to\mathbb{I}}{\longrightarrow} m^{(1)}(x,t),$$

for any $(x,t) \in \mathbb{R}^d \times (0,T)$, with uniform convergence in the compact sets contained in $\mathbb{R}^d \times (0,T)$. In order to reach our goal we have to enunciate Kamke's Theorem.

Theorem (Kamke's Theorem)**.** *Let $f, f_k : \Omega \subseteq \mathbb{R}\times\mathbb{R}^d$ continuous, $k = 1, 2, \dots$ and $(t_k, y_k) \in \Omega$ for any $k \in \mathbb{N}$. Let us suppose that the Cauchy problem*

$$\begin{cases} \dot{y} = f(t,y) \\ y(t_0) = y_0 \end{cases}$$

has a unique solution $y(t)$ in the compact set I. If $(t_k, y_k) \to (t_0, y_0)$ and $f_k \rightrightarrows f$ on the compact sets contained in Ω, then $y_k(t) \rightrightarrows y(t)$ uniformly on I, where $y_k(\cdot)$ is a solution of the Cauchy problem

$$\begin{cases} \dot{y} = f_k(t,y) \\ y(t_k) = y_k. \end{cases}$$

Now, if we apply to model (3.1) the same resolution strategy of model (3.11), we find that $u(x,t) = x'H(t)x + L(t)\cdot x + s(t)$ is a solution of the Hamilton-Jacobi-Bellman equation of system (3.5) if and only if the functions H, L and s satisfy

$$\begin{cases} \dot{H} = 2H^2; & H(T) = \frac{a}{2}\mathbb{I} \\ \dot{L} = 2HL; & L(T) = -b \\ \dot{s} = \frac{1}{2}|L|^2; & s(T) = c. \end{cases} \tag{3.31}$$

Given a solution $\left(H^{(2)}, L^{(2)}, s^{(2)}\right)$ of the system (3.24) and the solution

$$H^{(1)} = \frac{a}{2[1+a(T-t)]}\mathbb{I}, \qquad L^{(1)} = -\frac{b}{1+a(T-t)}\mathbb{I}, \qquad s^{(1)} = -\frac{(T-t)|b|^2}{2[1+a(T-t)]}$$

of (3.31), by Kamke's Theorem we have that $H^{(2)} \rightrightarrows H^{(1)}$ on $[0,T]$ when $A \to \mathbb{O}$ and $B \to \mathbb{I}$. In particular $H^{(2)}(0) \to H^{(1)}(0) = \frac{a}{2(1+aT)}\mathbb{I}$.

Hence we can calculate the limit for $A \to \mathbb{O}$ and $B \to \mathbb{I}$ in (3.28), so that we get

$$C \underset{A\to\mathbb{O},\ B\to\mathbb{I}}{\longrightarrow} (1+aT)\mathbb{I} \qquad \text{and} \qquad D \underset{A\to\mathbb{O},\ B\to\mathbb{I}}{\longrightarrow} -T\mathbb{I}.$$

Using that in (3.29) we have

$$\lim_{A\to\mathbb{O},\ B\to\mathbb{I}} Q = \left(1 - \frac{\tilde{b}T}{1+aT}\right)\mathbb{I} \overset{a=\tilde{a}+\tilde{b}}{=} \left(\frac{1+\tilde{a}T}{1+aT}\right)\mathbb{I}.$$

Finally, if we use that to calculate the same limit in (3.30), we immediately get that the expressions for b and c in (3.30) tends to formula (3.10) when $A \to \mathbb{O}$ and $B \to \mathbb{I}$:

$$\begin{aligned} &\lim_{A\to\mathbb{O},\ B\to\mathbb{I}} Q^{-1}\left(\tilde{a}h + \tilde{b}C^{-1}\mathbb{E}[m_0]\right) = \frac{\tilde{a}(1+aT)h + \tilde{b}\mathbb{E}[m_0]}{1+\tilde{a}T} \\ &\lim_{A\to\mathbb{O},\ B\to\mathbb{I}} \frac{\tilde{a}}{2}|h|^2 + \frac{\tilde{b}}{2}\left|C^{-1}\mathbb{E}[m(0)] - C^{-1}Db\right|^2 = \frac{\tilde{a}}{2}|h|^2 + \frac{\tilde{b}}{2(1+aT)^2}|\mathbb{E}[m_0] + bT|^2. \end{aligned} \tag{3.32}$$

Therefore, by the Kamke's Theorem we deduce that $L^{(2)} \rightrightarrows L^{(1)}$ and $s^{(2)} \rightrightarrows s^{(1)}$

on $[0,T]$ when $A \to \mathbb{O}$ and $B \to \mathbb{I}$. Consequently $u^{(2)} \rightrightarrows u^{(1)}$ on the compact sets contained in $\mathbb{R}^d \times (0,T)$, so that there is pointwise convergence in $\mathbb{R}^d \times (0,T)$ too.

It follows from the facts above that also $e^{M(t)} \rightrightarrows \left(\frac{1+aT}{1+a(T-t)}\right)^d$ and $Y^{(2)}(t,x) \rightrightarrows Y^{(1)}(t,x)$, where $Y^{(2)}$ and $Y^{(1)}$ are respectively given by (3.27) and (3.8). By the method of characteristics we have supposed that $m_0(\cdot) \in \mathcal{C}^1(\mathbb{R}^d)$, so that we deduce $m^{(2)}(x,t) \rightrightarrows m^{(1)}(x,t)$ on the compact sets contained in $\mathbb{R}^d \times (0,T)$, and consequently there is pointwise convergence in $\mathbb{R}^d \times (0,T)$ when $A \to \mathbb{O}$ and $B \to \mathbb{I}$.

3.3 Third model

The hypotheses of this third model are:

$$\begin{aligned}
&i. \quad J(t,x,\alpha) = \int_t^T \left[\frac{|\alpha(s)|^2}{2} + y(s)'My(s) + \mathbb{E}[m(s)]'N\mathbb{E}[m(s)]\right] \mathrm{d}s + \\
&\qquad\qquad + \frac{\tilde{a}}{2}|y(T)-h|^2 + \frac{\tilde{b}}{2}|y(T)-\mathbb{E}[m(T)]|^2 \\
&ii. \quad \begin{cases} \dot{y}(s) = Ay(s) + B\alpha(s) \\ y(t) = x \end{cases}
\end{aligned} \tag{3.33}$$

iii. $m_0(\cdot)$ absolutely continuous, bounded and with compact support,

where $M \in Sym(d)$, $N \in Sym(d)$ and the rest of parameters and matrices are as in the previous section. The discussion is very similar to that done in the general case of the previous section. The new mean field game system is

$$\begin{cases}
-u_t - (Ax)'D_x u - \mathbb{E}[m(t)]'N\mathbb{E}[m(t)] + \\
\qquad + \frac{1}{2}(D_x u)'SD_x u - x'Mx = 0 & (x,t) \in \mathbb{R}^d \times [0,T) \\
u(x,T) = \frac{a}{2}|x|^2 - b \cdot x + c \\
m_t + \operatorname{div}_x(m(Ax - SD_x u)) = 0 & (x,t) \in \mathbb{R}^d \times (0,T] \\
m(x,0) = m_0(x).
\end{cases} \tag{3.34}$$

where the parameters a, b and c are given by (3.2). A function $u(x,t) = x'H(t)x + L(t)\cdot x + s(t)$ solves the Hamilton-Jacobi-Bellman equation in (3.34) if and only if H, L, s satisfy

$$\begin{cases} \dot{H}(t) = 2H(t)SH(t) - A'H(t) - H(t)A - M; & H(T) = \frac{a}{2}\mathbb{I} \\ \dot{L}(t) = 2H(t)SL(t) - A'L(t); & L(T) = -b \\ \dot{s}(t) = \frac{1}{2}L'(t)SL(t) - \mathbb{E}[m(t)]'N\mathbb{E}[m(t)]; & s(T) = c. \end{cases} \tag{3.35}$$

Here again there exists some $\tilde{T} > 0$ such that for any $T \leq \tilde{T}$ the Riccati equation has local unique solution $\tilde{H} \in \mathcal{C}^1(]0,T[; Sym(d))$ and such that $\tilde{H}(t)$ is invertible for any $t \in [0,T]$. Then it is easy to check that

$$\tilde{L}(t) = -\frac{2}{a}\tilde{H}(t)e^{A(t-T)}b \tag{3.36}$$

solve the second equation of (3.35) in $]0,T[$.

Momentarily we can not solve the equation for $s(\cdot)$ because $\mathbb{E}[m(\cdot)]$ is unknown. So we first solve the continuity equation, which depends only on $\tilde{H}$ and $\tilde{L}$, then we analyze the fixed point equation for b and finally we determine $s(\cdot)$.

Through the method of characteristics we find out that a solution in $]0,T[$ of the continuity equation in (3.34) is

$$m(x,t) = m_0(\tilde{Y}(t,x))e^{\tilde{M}(t)}, \tag{3.37}$$

where

$$\begin{cases} \tilde{Y}(t,x) = \tilde{H}^{-1}(0)e^{A't}\tilde{H}(t)x - \frac{2}{a}\tilde{H}^{-1}(0)\left(\int_0^t e^{A's}\tilde{H}(s)S\tilde{H}(s)e^{A(s-T)}\,\mathrm{d}s\right)b \\ \tilde{M}(t) = \int_0^t \mathrm{tr}(2S\tilde{H}(s) - A)\,\mathrm{d}s. \end{cases}$$

If we define

$$\begin{cases} \tilde{C} = \frac{a}{2}\tilde{H}^{-1}(0)e^{A'T} \\ \tilde{D} = -\frac{2}{a}\tilde{H}^{-1}(0)\left(\int_0^T e^{A's}\tilde{H}(s)S\tilde{H}(s)e^{A(s-T)}\,\mathrm{d}s\right), \end{cases} \tag{3.38}$$

we have that for T small the matrix

$$\tilde{Q} := \mathbb{I} + \tilde{b}\tilde{C}^{-1}\tilde{D} \tag{3.39}$$

is invertible. After determining $\mathbb{E}[m(T)]$, by (3.2) we deduce

$$\begin{cases} b = \tilde{Q}^{-1}\left(\tilde{a}h + \tilde{b}\tilde{C}^{-1}\mathbb{E}[m_0]\right) \\ c = \frac{\tilde{a}}{2}|h|^2 + \frac{\tilde{b}}{2}\left|\tilde{C}^{-1}\mathbb{E}[m(0)] - \tilde{C}^{-1}\tilde{D}b\right|^2. \end{cases} \tag{3.40}$$

Finally

$$\tilde{s}(t) = c + \int_0^t \left[\frac{1}{2}L'(s)SL(s) - \mathbb{E}[m(s)]'N\mathbb{E}[m(s)]\right] \, \mathrm{d}s \tag{3.41}$$

solves the third equation of (3.35) in $]0,T[$.

In conclusion we have proved the following theorem.

Theorem 3.7. *There exists $\tau > 0$ sufficiently small such that for any $T < \tau$ the matrix $\tilde{Q} = (3.39)$ is invertible; furthermore, for any $T < \tau$, there exists a unique $H \in \mathcal{C}^1(]0,T[; Sym(d))$ continuous in 0 and T which solves the first equation in (3.35) in $[0,T]$ and that is invertible for any $t \in [0,T]$. Moreover, if $\tilde{C}$ and $\tilde{D}$ are given by (3.38), b and c are given by (3.40) and $\tilde{L}(t) = (3.36)$ and $\tilde{s}(t) = (3.41)$, then $u(x,t) = x'\tilde{H}(t)x + \tilde{L}(t)\cdot x + \tilde{s}(t)$ and $m(x,t) = (3.37)$ are solutions in $\mathbb{R}^d \times (0,T)$ of the mean field games system (3.34) linked to the mean field model (3.33).*

Remark 3.8. For the same reasons of Remark 3.6, there is consistency between model (3.33) and models (3.11) and (3.1). In particular, with the same notations of the aforementioned remark, if $\left(u^{(3)}, m^{(3)}\right)$ is a solutions of system (3.34) in the sense of Theorem 3.7, then:

$$u^{(3)}(x,t) \underset{M,N\to\mathbb{O}}{\rightarrow} u^{(2)}(x,t) \qquad\qquad m^{(3)}(x,t) \underset{M,N\to\mathbb{O}}{\rightarrow} m^{(2)}(x,t)$$

and

$$u^{(3)}(x,t) \underset{A,M,N\to\mathbb{O};\ B\to\mathbb{I}}{\rightarrow} u^{(1)}(x,t) \qquad\qquad m^{(3)}(x,t) \underset{A,M,N\to\mathbb{O};\ B\to\mathbb{I}}{\rightarrow} m^{(1)}(x,t),$$

for any $(x,t) \in \mathbb{R}^d \times (0,T)$.

Furthermore, there is uniform convergence on the compact sets contained in $\mathbb{R}^d \times (0,T)$.

Bibliography

[1] Ambrosio L., Gigli N., Savaré G., *Gradient Flows in metric spaces and in the space of probability measures.* Second edition. Lectures in Mathematics ETH Zurich. Birkhauser Verlag, Basel, (2008).

[2] Aubin J.P. and Frankowska H., *Set-valued analysis.* Boston: Birkhauser, (1990).

[3] Bardi M., Capuzzo-Dolcetta I., *Optimal Control and Viscosity solutions of Hamilton-Jacobi-Bellman Equations*, Birkhauser, (1997).

[4] Bogachev V.I., *Measure Theory Volume I*, Springer, (2007).

[5] Bogachev V.I., *Measure Theory Volume II*, Springer, (2007).

[6] Cannarsa P., *Funzioni semiconcave, singolaritá e pile di sabbia*, Bollettino dell'Unione Matematica Italiana, Serie 8, Vol. 8-b (2005), n.3, p. 549-567.

[7] Cannarsa P. and Sinestrari C. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control.* Progress in Nonlinear Differential Equations and their Applications, 58. Birkhauser Boston Inc., Boston, MA, 2004.

[8] Cannarsa P., Soner H.M., *Generalized one-sided estimates for solutions of Hamilton-Jacobi equations and applications*, Nonlinear Analysis, Theory, Methods and Applications, Vol. 13, No. 3, pp. 305-323, (1989)

[9] Cardaliaguet P, *Notes on Mean Field Games*, `www.science.unitn.it/~bagagiol/NotesByCardaliaguet.pdf`, (2009).

[10] Cardaliaguet P., *Introduction to differential games*, `www.ceremade.dauphine.fr/~cardaliaguet/CoursJeuxDiff.pdf`, (2010).

[11] Cardaliaguet P. and Hadikhanloo S., *Learning in mean field games: the fictitious play*, COCV Volume 23, Number 2, April-June 2017, 569-591.

[12] Cheng S.F., Reeves D.M., Wellman M.P., *Notes on Equilibria in Symmetric Games*, Proceedings of the 6th International Workshop On Game Theoretic And Decision Theoretic Agents GTDT 2004. 71-78. Research Collection School Of Information Systems. ink.library.smu.edu.sg/sis_research/1213/.

[13] Clarke F.H., *Optimization and nonsmooth analysis*, Second edition, Classics in Applied Mathematics, 5. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, (1990).

[14] Dellacherie C., Meyer P.A., *Probabilities and potential*, vol. 29 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam, (1978).

[15] Diaz L., Naulin R., *A Proof of the Schauder-Tychonov Theorem*, Divulgaciones Matematicas Vol. 14 No. 1(2006), pp. 47-57.

[16] Engwerda J.C., *LQ Dynamic optimization and Differential Games*, Chichester: Wiley, (2005).

[17] Evans L.C., *Partial differential equations*, American Mathematical Society, (1998).

[18] Fan K., *Fixed-point and Minimax Theorems in Locally Convex Topology Linear Spaces*. Pro. Nat. Acad. Sci. U.S.A. 38(2) (1952), 121-126.

[19] Fisher M. and Silva F.J, *On the asymptotic nature of first order mean field game*, arXiv:1903.03602v1 [math.OC], March 8, 2019.

[20] Fleming W.H. and Rishel R.W., *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York, (1975).

[21] Fleming W.H. and Rishel R.W., *Controlled Markov processes and viscosity solutions*, Springer Verlag Berlin, 1993.

[22] Green J.W. and Valentine F.A., *On the Arzelá-Ascoli Theorem*, Mathematics Magazine, Vol. 34, No. 4 (Mar.-Apr., 1961), pp. 199-202.

[23] Hewitt E., Savage L.J. *Symmetric measure on Cartesian products*, Transaction of the American Mathematical Society, 80:470-501.

[24] M. Huang, R.P. Malhamé, and P.E. Caines, *Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Syst., 6(3):221-252, 2006.

[25] Ladyzenskaja O.A., Solinnokov V.A., Ural'ceva N.N., *Linear and quasilinear equations of parabolic type*, Translations of Mathematical Monographs, Vol. 23 American Mathematical Society, Providence, R.I., (1967).

[26] Lasry J.M. and Lions P.L. *Mean field games*, Jpn. J. Math. 2 (2007), no. 1, 229-260.

[27] Lions P.L., *Generalized solutions of Hamilton-Jacobi equations*, Boston, Pitman, (1982).

[28] Peters H., *Game Theory, a Multi-Leveled Approach*, Springer, (2008).

[29] Rudin W., Real and Complex Analysis, Third edition, McGraw-Hill, (1987).

[30] Villani C. *Optimal transport, old and new*, Springer, Berlin, 2006.