

Séminaire Laurent Schwartz

EDP et applications

Année 2015-2016


Pierre Cardaliaguet

Mean field games: the master equation and the mean field limit

Séminaire Laurent Schwartz — EDP et applications (2015-2016), Exposé n° XVII, 10 p.

<http://sisedp.cedram.org/item?id=SLSEDP_2015-2016____A17_0>

© Institut des hautes études scientifiques & Centre de mathématiques Laurent Schwartz,
École polytechnique, 2015-2016.

 Cet article est mis à disposition selon les termes de la licence
CREATIVE COMMONS ATTRIBUTION – PAS DE MODIFICATION 3.0 FRANCE.
<http://creativecommons.org/licenses/by-nd/3.0/fr/>

Institut des hautes études scientifiques
Le Bois-Marie • Route de Chartres
F-91440 BURES-SUR-YVETTE
<http://www.ihes.fr/>

Centre de mathématiques Laurent Schwartz
UMR 7640 CNRS/École polytechnique
F-91128 PALAISEAU CEDEX
<http://www.math.polytechnique.fr/>

cedram

Exposé mis en ligne dans le cadre du
Centre de diffusion des revues académiques de mathématiques
<http://www.cedram.org/>

Mean field games: the master equation and the mean field limit

P. Cardaliaguet*

Abstract

We present here results obtained in the joint work with Delarue, Lasry and Lions [4] on the convergence, as N tends to infinity, of a system of N coupled Hamilton-Jacobi equations, the Nash system. This system arises in differential game theory. The limit problem can be expressed in terms of the “Mean Field Game” system (coupling a Hamilton-Jacobi equation with a Fokker-Planck equation), or, alternatively, in terms of the “master equation” (a kind of second order partial differential equation stated on the space of probability measures). We also discuss the behavior of the optimal trajectories, for which we show a propagation of chaos property.

The description of interactions between “rational agents” is often a difficult issue because the agents, being supposed rational, observe each other and react in function of their observation (strategic interaction): this leads to the notion of Nash equilibria, which are often difficult to compute and interpret. Aumann [2] was among the first to notice that this problem simplifies a lot when there are infinitely many “non-atomic” agents. By non-atomic, we mean that the agents have an infinitesimal influence on the global system. For a long time the ideas of non-atomic games have been applied to games in which the action space of the players is relatively simple (one-shot games). The importance of dynamic optimization problems (optimal control) in engineering sciences, economic theory, finance, etc... lead Lasry and Lions [12, 13, 14] and Huang, Caines and Malhamé [8, 9, 10, 11] to develop the counterpart of Aumann’s “non-atomic games” to optimal control. This is the so-called mean field games. It is worth mentioning that similar ideas were discussed in earlier works in the economic literature (heterogeneous agent models).

Based on heuristic considerations, these authors derived a system describing interacting, indistinguishable agents. This system, called Mean Field Game (MFG) system, takes the form of a coupling between a backward Hamilton-Jacobi (HJ) equation and a forward Kolmogorov equation:

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m(t)) & \text{in } [0, T] \times \mathbb{R}^d, \\ \partial_t m - \Delta m - \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } [0, T] \times \mathbb{R}^d, \\ u(T, x) = G(x, m(T)), m(0, \cdot) = m_0 & \text{in } \mathbb{R}^d, \end{cases} \quad (1)$$

The data are the Hamiltonian $H : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, the horizon $T > 0$, the initial measure m_0 and the maps F, G which describe the influence of the second equation on the first one. The function u can be thought of as the value function for an average agent seeking to optimize an optimal control problem, while m represents the time-evolving probability distribution of the state of the players when all players play in an optimal way.

The terminology *Mean Field Games* comes from the analogy with statistical physics, which deals with large populations of particles and derives macroscopic laws from microscopic ones.

*Ceremade, Université Paris-Dauphine, cardaliaguet@ceremade.dauphine.fr

Here also it is tempting to derivate the MFG system from a model where N agents interact, by letting the number N of agents tend to infinity. Let us recall that the Nash equilibrium configuration for a system of N -interacting players is described by the following nonlinear parabolic system (called here the Nash system):

$$\left\{ \begin{array}{l} -\partial_t v^{N,i}(t, \mathbf{x}) - \sum_{j=1}^N \Delta_{x_j} v^{N,i}(t, \mathbf{x}) - \beta \sum_{j,k=1}^N \text{Tr} D_{x_j, x_k}^2 v^{N,i}(t, \mathbf{x}) + H(x_i, D_{x_i} v^{N,i}(t, \mathbf{x})) \\ \quad + \sum_{j \neq i} D_p H(x_j, D_{x_j} v^{N,j}(t, \mathbf{x})) \cdot D_{x_j} v^{N,i}(t, \mathbf{x}) = F^{N,i}(\mathbf{x}) \\ v^{N,i}(T, \mathbf{x}) = G^{N,i}(\mathbf{x}) \end{array} \right. \quad \begin{array}{l} \text{for } (t, \mathbf{x}) = (t, x_1, \dots, x_N) \text{ in } [0, T] \times (\mathbb{R}^d)^N, \\ \text{in } (\mathbb{R}^d)^N. \end{array} \quad (2)$$

System (2) describes the Nash equilibrium configuration of an N -player differential game. Roughly speaking, for any agent $i \in \{1, \dots, N\}$, the map $v^{N,i}(t, \mathbf{x})$ is the minimal cost of agent i when the initial position of the system at time t is $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{R}^{Nd}$. The cost is described in term of an optimal control problem with running cost $F^{N,i}$ and terminal cost $G^{N,i}$. Such Nash equilibria have been discussed, for instance, by Friedman [7] or Bensoussan and Frehse [3]. Note that (2) is a nonlinear parabolic system in a high dimensional space (\mathbb{R}^{Nd}) with many unknown (N).

Intimately related with the Nash system, the set of optimal trajectories describes the evolution in time of the state of each agent when playing in an optimal way. It solves a system of N coupled stochastic differential equations (SDE):

$$dX_{i,t} = -D_p H(X_{i,t}, D_{x_i} v^{N,i}(t, \mathbf{X}_t)) dt + \sqrt{2} dB_t^i + \sqrt{2\beta} dW_t, \quad t \in [0, T], i \in \{1, \dots, N\}, \quad (3)$$

where $v^{N,i}$ is the solution to (2) and the $((B_t^i)_{t \in [0, T]})_{i=1, \dots, N}$ and $(W_t)_{t \in [0, T]}$ are d -dimensional independent Brownian motions. The Brownian motions $((B_t^i)_{t \in [0, T]})_{i=1, \dots, N}$ correspond to the *individual noises*, while the Brownian motion $(W_t)_{t \in [0, T]}$ is the same for all the equations and, for this reason, is called the *common noise*. These trajectories somehow play the role of (random) characteristics for the Nash system (in a kind of Feynman-Kac formula). It is also interesting to study the limit of the $(X_{i,t})$ as the number of players tends to infinity.

In order to expect a relation between the Nash system and the MFG system, one has of course to assume that the F^N, G^N, F and G are related. As we wish to formalize the fact that the agents are small and indistinguishable, the more natural choice is to require that

$$F^{N,i}(\mathbf{x}) = F(x_i, m_{\mathbf{x}}^{N,i}), \quad G^{N,i}(\mathbf{x}) = G(x_i, m_{\mathbf{x}}^{N,i}), \quad \text{where } m_{\mathbf{x}}^{N,i} := \frac{1}{N-1} \sum_{j \neq i} \delta_{x_j} \quad (4)$$

for any $\mathbf{x} = (x_1, \dots, x_N)$. We will mostly work under this assumption. Another interesting regime is when the $F^{N,i}(\mathbf{x})$ can be written in the same form $F^N(x_i, m_{\mathbf{x}}^{N,i})$, but F^N becomes more and more singular as $N \rightarrow +\infty$: this configuration is discussed in Section 4.

The relation between the MFG system and the Nash system is far from obvious, even heuristically. First the measure m_0 does not appear in the Nash system, while it is a data for the MFG system. Second, each map $v^{N,i}$ depends on Nd -space variables, while the pair (u, m) depends only on d -space variables only. Third there is no parameter β in the MFG system: actually we will see that the MFG system is indeed the limit of the Nash system only in the regime $\beta = 0$. In the regime $\beta > 0$, the MFG system has to be replaced by a stochastic MFG system (a system

of SPDEs) presented below. In order to explain the convergence, Lions introduced (mostly in an informal way) the following ideas in his courses at the Collège de France [15].

Assume that the F^N, G^N are given by (4). Then, by the symmetry properties of the maps $v^{N,i}$, one expects that there exists a map $U^N = U^N(t, x, m)$, such that the following representation holds:

$$v^{N,i}(t, \mathbf{x}) = U^N(t, x_i, m_{\mathbf{x}}^{N,i}).$$

Note that, a priori, U^N is defined only on discrete measures taking at most N values. However, if one could be able to extend U^N to the full space of probability measure, it is natural to guess that the sequence of maps U^N should converge, as $N \rightarrow +\infty$, to a map $U = U(t, x, m)$. In order to guess the equation satisfied by U , let us suppose that U^N is sufficiently smooth and converge smoothly to U . Here, by smooth, we mean that U^N is sufficiently derivable in all variable, in particular with respect to the measure m : this point is detailed in Section 1. Then one can show that

$$\begin{aligned} D_{x_j} u^{N,i}(t, \mathbf{x}) &= \frac{1}{N-1} D_m U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j) & (j \neq i), \\ D_{x_j, x_j}^2 v^{N,i}(t, \mathbf{x}) &= \frac{1}{N-1} D_y [D_m U^N](t, x_i, m_{\mathbf{x}}^{N,i}, x_j) \\ &\quad + \frac{1}{(N-1)^2} D_{mm}^2 U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j, x_j) & (j \neq i) \\ \text{while, if } j \neq k, \quad D_{x_j, x_k}^2 v^{N,i}(t, \mathbf{x}) &= \frac{1}{(N-1)^2} D_{mm}^2 U^N(t, x_i, m_{\mathbf{x}}^{N,i}, x_j, x_k) & (i, j, k \text{ distinct}). \end{aligned} \tag{5}$$

Plugging these expressions into the equation satisfied by $v^{N,1}$ and assuming that U^N converges to U in a sufficiently smooth way leads to the so-called *master equation* that the limit map U should satisfy:

$$\left\{ \begin{array}{l} -\partial_t U - (1 + \beta) \Delta_x U + H(x, D_x U) \\ \quad - (1 + \beta) \int_{\mathbb{R}^d} \operatorname{div}_y [D_m U] dm(y) + \int_{\mathbb{R}^d} D_m U \cdot D_p H(y, D_x U) dm(y) \\ \quad - 2\beta \int_{\mathbb{R}^d} \operatorname{div}_x [D_m U] dm(y) - \beta \int_{\mathbb{R}^{2d}} \operatorname{Tr} [D_{mm}^2 U] dm \otimes dm = F(x, m) \\ \quad \text{in } [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \\ U(T, x, m) = G(x, m) \quad \text{in } \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \end{array} \right. \tag{6}$$

The master equation is not—by far—the first equation stated in the space of probability measures: for instance, in a pioneering work, Otto [16] described the porous medium equation as a gradient flow in the space of measures. However the difficulty with the master equation is that it is nonlinear, nonlocal and of second order in the measure variable (and not a gradient flow).

There is a strong relation between the master equation and the MFG system (1), at least when $\beta = 0$. Indeed, when $\beta = 0$, the solutions of the MFG system play the role of characteristics for the master equation (see Section 3). In this sense the MFG system can indeed be understood as the limit of the Nash system. When β is positive, the MFG system has to be replaced by the stochastic MFG system.

Making the above argument rigorous took some time. The main issue is the lack of estimate for the Nash system which prevents to use simple compactness arguments. Actually one only has uniform L^∞ -bounds on the $v^{N,i}$, which is far from sufficient to justify the computation above. The first step has been the construction of a solution for the master equation (see the references in [4, 6]). The paper [4], presented in this notes and co-authored with Delarue, Lasry and Lions, has been the first to derive the existence of a solution to the master equation with common noise ($\beta > 0$). It also shows for the first time the convergence of the Nash system to

the master equation, and thus to the MFG system. The aim of these notes is to present these results, as well as their (partial) extension to the case of singular coupling ([5]).

To fit with the presentation in [4, 5], we work here with functions which are periodic in the space variable: although this might seem artificial, this simplifies somewhat the analysis, avoiding the issue of the boundary conditions. So the ambient space is $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$. An extension of the results to \mathbb{R}^d is discussed in the forthcoming monograph by Carmona and Delarue [6].

1 Derivatives in the space of probability measures

Throughout these notes we denote by $\mathcal{P}(\mathbb{T}^d)$ the set of Borel probability measures on the torus $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$. It is endowed with the Monge-Kantorovitch distance:

$$\mathbf{d}_1(m, m') = \sup_{\phi} \int_{\mathbb{T}^d} \phi(y) d(m - m')(y),$$

where the supremum is taken over all 1-Lipschitz continuous maps $\phi : \mathbb{T}^d \rightarrow \mathbb{R}$.

Our aim is now to introduce the notion of derivatives for functions defined on $\mathcal{P}(\mathbb{T}^d)$. There is a large literature on this topic: see, for instance, the monograph [1] and the references therein. Here we use two notions. The first one fits with the linear structure of the space signed measures. The second one has more to do with the ‘‘Riemannian’’ geometry of the space $\mathcal{P}(\mathbb{T}^d)$.

Definition 1.1. *We say that $U : \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ is \mathcal{C}^1 if there exists a continuous map $\frac{\delta U}{\delta m} : \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \rightarrow \mathbb{R}$ such that, for any $m, m' \in \mathcal{P}(\mathbb{T}^d)$,*

$$\lim_{s \rightarrow 0^+} \frac{U((1-s)m + sm') - U(m)}{s} = \int_{\mathbb{T}^d} \frac{\delta U}{\delta m}(m, y) d(m' - m)(y).$$

The derivative $\frac{\delta U}{\delta m}$ is defined up to an additive constant: we use the convention

$$\int_{\mathbb{T}^d} \frac{\delta U}{\delta m}(m, y) dm(y) = 0. \quad (7)$$

Note also that

$$\forall m, m' \in \mathcal{P}(\mathbb{T}^d), \quad U(m') - U(m) = \int_0^1 \int_{\mathbb{T}^d} \frac{\delta U}{\delta m}((1-s)m + sm', y) d(m' - m)(y) ds. \quad (8)$$

In particular, the Lipschitz continuity of U can be estimated by:

$$|U(m') - U(m)| \leq \sup_{m''} \left\| D_y \frac{\delta U}{\delta m}(m'', \cdot) \right\|_{\infty} \mathbf{d}_1(m, m').$$

This leads us to define the ‘‘intrinsic derivative’’ of U .

Definition 1.2. *If $\frac{\delta U}{\delta m}$ is of class \mathcal{C}^1 with respect to the second variable, the intrinsic derivative $D_m U : \mathcal{P}(\mathbb{T}^d) \times \mathbb{T}^d \rightarrow \mathbb{R}^d$ is defined by*

$$D_m U(m, y) := D_y \frac{\delta U}{\delta m}(m, y)$$

The expression $D_m U$ can be understood as a derivative of U along vector fields: indeed, if $\phi : \mathbb{T}^d \rightarrow \mathbb{R}^d$ is a smooth and bounded vector field, then one can show that

$$\lim_{h \rightarrow 0} \frac{U((id + h\phi)\#m) - U(m)}{h} = \int_{\mathbb{T}^d} D_m U(m, y) \cdot \phi(y) dm(y).$$

Second order derivatives can be defined in a similar way (see [4] for details).

2 Main results

We discuss here the well-posedness of the master equation and the convergence of the Nash system when the coupling functions F^N and G^N can be represented as in (4) where F and G are nonlocal and smooth.

To simplify the discussion, we assume that all the data are sufficiently regular. We also suppose (which is a very strong condition) that the Hamiltonian is globally Lipschitz continuous and that all its derivatives are bounded. Finally, and this is a key condition, we assume that F and G are monotone: for F , for instance, this means that

$$\int_{\mathbb{T}^d} (F(x, m) - F(x, m')) d(m - m')(x) \geq 0 \quad \forall m, m' \in \mathcal{P}(\mathbb{T}^d).$$

Under these assumptions, the master equation (6) is well-posed:

Theorem 2.1 ([4]). *Under the above conditions, equation (6) has a unique classical solution.*

Our main convergence result is based on the existence of a solution to the master equation. It can be expressed in two different ways.

Theorem 2.2 ([4]). *Let $(v^{N,i})$ be the solution to the Nash system (2) and U be the classical solution to the second order master equation (6). Fix $N \geq 1$ and $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$.*

(i) *For any $\mathbf{x} \in (\mathbb{T}^d)^N$, let $m_{\mathbf{x}}^N := \frac{1}{N} \sum_{i=1}^N \delta_{x_i}$. Then*

$$\sup_{i \in \{1, \dots, N\}} |v^{N,i}(t_0, \mathbf{x}) - U(t_0, x_i, m_{\mathbf{x}}^N)| \leq CN^{-1}.$$

(ii) *For any $i \in \{1, \dots, N\}$ and $x \in \mathbb{T}^d$, let us set*

$$w^{N,i}(t_0, x_i, m_0) := \int_{\mathbb{T}^d} \dots \int_{\mathbb{T}^d} v^{N,i}(t_0, \mathbf{x}) \prod_{j \neq i} m_0(dx_j) \quad \text{where } \mathbf{x} = (x_1, \dots, x_N).$$

Then

$$\|w^{N,i}(t_0, \cdot, m_0) - U(t_0, \cdot, m_0)\|_{\infty} \leq \begin{cases} CN^{-1/d} & \text{if } d \geq 3 \\ CN^{-1/2} \log(N) & \text{if } d = 2 \end{cases} \quad (9)$$

In (i) and (ii), the constant C does not depend on i , t_0 , m_0 , i nor N .

The result can actually be proved in more general frameworks and, in particular, as soon as one knows the existence of a classical solution to the master equation (see [6]). We explain below that the second statement holds (in a weaker form) when the coupling functions are local.

Theorem 2.2 is the first to show the link between the Nash system and the master equation (and, as we explain in the next section, with the MFG system). However there remains many

open problems. The main open question concerns the convergence without the monotonicity assumptions on F and G . Indeed, when F or G are not monotone, the classical solution of the master equation does not exist, it is expected to develop shocks in finite time. Then the method of proof completely breaks down and nothing is known after the onset of singularities. Another open question concerns more realistic models (where H is not globally Lipschitz continuous, when there are boundaries, etc...).

The above convergence is also strongly related with the mean field analysis. Namely, let $t_0 \in [0, T]$, $m_0 \in \mathcal{P}(\mathbb{T}^d)$ and let (Z_i) be an i.i.d family of N random variables of law m_0 . We set $\mathbf{Z} = (Z_1, \dots, Z_N)$. Let also $((B_t^i)_{t \in [0, T]})_{i \in \{1, \dots, N\}}$ be a family of N independent Brownian motions which is also independent of (Z_i) and let $(W_t)_{t \in [0, T]}$ be a Brownian motion independent of the (B^i) and (Z_i) . We consider the optimal trajectories $(\mathbf{Y}_t = (Y_{1,t}, \dots, Y_{N,t}))_{t \in [t_0, T]}$ for the N -player game:

$$\begin{cases} dY_{i,t} = -D_p H(Y_{i,t}, D_{x_i} v^{N,i}(t, \mathbf{Y}_t)) dt + \sqrt{2} dB_t^i + \sqrt{2\beta} dW_t, & t \in [t_0, T] \\ Y_{i,t_0} = Z_i \end{cases}$$

and the solution $(\tilde{\mathbf{X}}_t = (\tilde{X}_{1,t}, \dots, \tilde{X}_{N,t}))_{t \in [t_0, T]}$ of stochastic differential equation of McKean-Vlasov type:

$$\begin{cases} d\tilde{X}_{i,t} = -D_p H(\tilde{X}_{i,t}, D_x U(t, \tilde{X}_{i,t}, \mathcal{L}(\tilde{X}_{i,t}|W))) dt + \sqrt{2} dB_t^i + \sqrt{2\beta} dW_t, \\ \tilde{X}_{i,t_0} = Z_i. \end{cases}$$

The next result says that the solutions of the two systems are close:

Theorem 2.3. *Let the assumption of Theorem 2.2 be in force. Then, for any $N \geq 1$ and any $i \in \{1, \dots, N\}$, we have*

$$\mathbb{E} \left[\sup_{t \in [t_0, T]} |Y_{i,t} - \tilde{X}_{i,t}| \right] \leq CN^{-\frac{1}{d+8}}$$

for some constant $C > 0$ independent of t_0 , m_0 and N .

In particular, since the $(\tilde{X}_{i,t})$ are independent conditioned on W , the above result is a (conditional) propagation of chaos.

3 Ideas of proofs

The convergence results in Theorems 2.2 and Theorem 2.3 rely in a crucial way on the existence of a classical solution to the master equation. Let us first discuss this point first.

For a start, we focus on the case $\beta = 0$. The key idea is that the solution of the MFG system is a kind of characteristic of the solution of the master equation. More precisely, let $(t_0, m_0) \in [0, T] \times \mathcal{P}(\mathbb{T}^d)$ be an initial position and (u, m) be a solution to the MFG system

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m(t)) & \text{in } [t_0, T] \times \mathbb{R}^d, \\ \partial_t m - \Delta m - \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } [t_0, T] \times \mathbb{R}^d, \\ u(T, x) = G(x, m(T)), m(t_0, \cdot) = m_0 & \text{in } \mathbb{R}^d, \end{cases}$$

Under the monotony condition on F and G this solution is known to exist and to be unique (see Lasry and Lions [14]). If one sets

$$U(t_0, x, m_0) := u(t_0, x),$$

then we claim that the map U is (at least if it is smooth), a solution of the master equation (6). Indeed, note first that, for any $h \in [0, T - t_0]$,

$$u(t_0 + h, \cdot) = U(t_0 + h, \cdot, m(t_0 + h)).$$

Hence

$$\begin{aligned} \partial_t u(t_0, x) &= \partial_t U(t_0, x, m_0) + \int_{\mathbb{T}^d} \frac{\delta U}{\delta m}(t_0, x, m_0, y) \partial_t m(t_0, y) dy \\ &= \partial_t U(t_0, x, m_0) + \int_{\mathbb{T}^d} \frac{\delta U}{\delta m}(m_0, y) (\Delta m + \operatorname{div}(m D_p H(x, Du))) dy \\ &= \partial_t U(t_0, x, m_0) + \int_{\mathbb{T}^d} \Delta_y \left[\frac{\delta U}{\delta m} \right] (x, m_0, y) m_0(y) dy \\ &\quad - \int_{\mathbb{T}^d} D_y \left[\frac{\delta U}{\delta m} \right] (x, m_0, y) \cdot D_p H(x, Du) m_0(y) dy \\ &= \partial_t U(t_0, x, m_0) + \int_{\mathbb{T}^d} \operatorname{div}_y [D_m U] (x, m_0, y) m_0(y) dy \\ &\quad - \int_{\mathbb{T}^d} D_m U(x, m_0, y) \cdot D_p H(x, Du) m_0(y) dy. \end{aligned}$$

As

$$\begin{aligned} \partial_t u(t_0, x) &= -\Delta u + H(x, Du) - F(x, m_0) \\ &= -\Delta_x U(t_0, x, m_0) + H(x, D_x U(t_0, x, m_0)) - F(x, m_0), \end{aligned}$$

the map U satisfies the master equation (with $\beta = 0$).

Now of course the difficult part is to justify the assumed regularity of U . For this we argue as for standard transport equations, by showing that the smoothness of the characteristics with respect to the initial conditions propagate to the map U : to do so, we linearize the MFG system with respect to the measure m . The difficult part here is to use in a clever way the monotonicity condition to ensure the well-posedness of (and to get estimates on) the linearized system.

Note finally that, in view of the above discussion, one can rewrite (9) in Theorem 2.2 as

$$\|w^{N,i}(t_0, \cdot, m_0) - u(t_0, \cdot)\|_\infty \leq \begin{cases} CN^{-1/d} & \text{if } d \geq 3 \\ CN^{-1/2} \log(N) & \text{if } d = 2 \end{cases}$$

where (u, m) is the solution of the MFG system. This explains why the MFG system is indeed the limit of the Nash system.

The case $\beta > 0$ is much more involve. The general principle is the same as for $\beta = 0$, but the system of characteristics becomes a stochastic MFG system:

$$(MFGs) \begin{cases} d_t u_t = \left\{ -(1 + \beta) \Delta u_t + H(x, Du_t) - F(x, m_t) - \sqrt{2\beta} \operatorname{div}(v_t) \right\} dt \\ \quad + v_t \cdot \sqrt{2\beta} dW_t & \text{in } [t_0, T] \times \mathbb{T}^d, \\ d_t m_t = \left[(1 + \beta) \Delta m_t + \operatorname{div}(m_t D_p H(m_t, Du_t)) \right] dt - \sqrt{2\beta} \operatorname{div}(m_t dW_t) \\ & \text{in } [t_0, T] \times \mathbb{T}^d \\ m_{t_0} = m_0, u_T(x) = G(x, m_T) & \text{in } \mathbb{T}^d. \end{cases}$$

where (v_t) is a vector field which ensures (u_t) to be adapted to the filtration $(\mathcal{F}_t)_{t \in [t_0, T]}$ generated by the M.B. $(W_t)_{t \in [0, T]}$. As an intermediate result, we prove the well-posedness of the stochastic master equation, for which little was known so far (here again we refer to [4, 6] for

references). The main issue is that, in contrast with the deterministic MFG system, compactness (and Schauder fixed point) arguments are not adapted and one has instead to rely on continuation methods, which make the work much more technical.

We now turn to the proof of the convergence of the Nash system. As pointed out in the introduction, we have basically no estimates on the solution of the system (except in a few particular situations, such as the stationary case or the short time horizon, see [14]). To pass to the limit, the difficult terms to deal with in the Nash system are the crossed derivatives $D_p H(x_j, D_{x_j} v^{N,j}(t, \mathbf{x})) \cdot D_{x_j} v^{N,i}(t, \mathbf{x})$. The first key idea is to get rid of these terms by looking at the solutions along suitable “characteristics”. Not surprisingly, these characteristics are the optimal trajectories associated with the Nash system. The second key idea is that the classical solution U of the master equation furnishes an approximate solution to the Nash system. Namely, let

$$u^{N,i}(t, \mathbf{x}) := U(t, x_i, m_{\mathbf{x}}^{N,i}), \quad (10)$$

then it is not difficult to check that the $(u^{N,i})$ solve the Nash system (up to an error of $1/N$ in each equation). The key point is that the $(u^{N,i})$ enjoy very good estimates—actually all the estimates that are missing for the $(v^{N,i})$. For instance, in view of expressions like (5), one has

$$\|D_{x_j} u^{N,i}\|_{\infty} \leq C/N, \quad \|D_{x_j, x_k} u^{N,i}\|_{\infty} \leq C/N^2 \quad (11)$$

for any i, j, k distinct. Comparing the $v^{N,i}$ and the $u^{N,i}$ along the optimal solutions for $v^{N,i}$ and using Gronwal type arguments (thanks to the regularity of the $(u^{N,i})$) then yield the result after some computation (that would be a little long to reproduce here).

4 Convergence for local couplings

We now turn to the case where, in the Nash system, the maps F^N becomes increasingly singular: Namely we suppose that there exists a smooth (local) function $F : \mathbb{T}^d \times [0, +\infty) \rightarrow \mathbb{R}$ such that

$$\lim_{N \rightarrow +\infty} F^N(x_i, m dx) = F(x_i, m(x_i)), \quad (12)$$

for any sufficiently smooth probability density $m dx = m(x) dx$. In view of the above convergence result, one expects that the limit system is a MFG system with local interactions:

$$\begin{cases} -\partial_t u - \Delta u + H(x, Du) = F(x, m(t, x)) & \text{in } [0, T] \times \mathbb{T}^d, \\ \partial_t m - \Delta m - \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } [0, T] \times \mathbb{T}^d, \\ u(T, x) = G(x), m(0, \cdot) = m_0 & \text{in } \mathbb{T}^d \end{cases} \quad (13)$$

(here we have to assume that $G^N = G(x)$ does not depend on m).

There is a major difference between the non-local and the local coupling. Indeed, when F is a local couplings, the meaning of the master equation is not clear: obviously one cannot expect U to be a smooth solution to (6), if only because the coupling blows up at singular measures. As a consequence, the definition of the maps $u^{N,i}$ through (10) is dubious and, even if such a definition could make sense, there is no hope that the $u^{N,i}$ satisfy the regularity properties (11) needed in the proof of convergence.

Nevertheless we can still prove the convergence of the Nash system:

Theorem 4.1 ([5]). Assume that $\beta = 0$ and that $F^N = F^{\epsilon_N}$ where, for any $\epsilon > 0$, F^ϵ is given by

$$F^\epsilon(x, m) = F(\cdot, \xi^\epsilon \star m(\cdot)) \star \xi^\epsilon(x)$$

and where $\xi^\epsilon(x) = \epsilon^{-d}\xi(x/\epsilon)$, ξ being a symmetric smooth nonnegative kernel with compact support. If one chooses $\epsilon_N = N^{-\beta}$, with $\beta \in (0, (3d(d+1))^{-1})$, then there exists $\gamma \in (0, 1)$ such that

$$\|w^{N,i}(t_0, \cdot, m_0) - u(t_0, \cdot)\|_{L^1(m_0)} \leq CN^{-\gamma}.$$

where the $w^{N,i}$ are defined from the solution $(v^{N,i})$ of the Nash system (2) as in Theorem 2.2, part (ii), while (u, m) is the solution of the MFG system with local coupling (13).

The result actually holds under general assumption on the coupling function F^N : these assumptions explain how the regularity of F^N is allowed to deteriorate in function of the distance of F^N to F , see [5]. One can also show that the optimal trajectories of the Nash system converge to optimal trajectories associated with the MFG system (leading to a propagation of chaos property).

As the master equation for the limit problem does not seem to make much sense, we cannot follow the approach of Theorem 2.2 or 2.3. The new idea consists in comparing directly the solution of the Nash system to the solution of the MFG system *without using the master equation*. Let us point out, however, that we do not recover all the above convergence results and that the convergence rate is also much sharper in Theorem 2.2. In order to compare directly the solution of the Nash system $v^{N,i}$ and the u component of the MFG system, we build *different* and well chosen paths along which these functions behave in a same way. Then we overcome the difficulty that the paths are different (as well as the lack of estimate for $v^{N,i}$) by using the structure of the equation (convexity of the Hamiltonian and monotonicity of the map F), somehow reproducing the uniqueness argument for the MFG system [14] at the level of the difference $v^{N,i} - u$.

References

- [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 Zürich. Birkhäuser Verlag, Basel, 2008.
- [2] Aumann R. *Markets with a continuum of traders*. Econometrica, 32(1/2), 1964.
- [3] Bensoussan, A., Frehse, J. (2002). *Smooth solutions of systems of quasilinear parabolic equations*. ESAIM: Control, Optimisation and Calculus of Variations, 8, 169-193.
- [4] Cardaliaguet P., Delarue F., Lasry J.-M., and Lions P.-L. (2015) *The master equation and the convergence problem in mean field games*. Preprint.
- [5] Cardaliaguet P (2016) *The convergence problem in mean field games with a local coupling*. Preprint.
- [6] Carmona, R., and Delarue, F. PROBABILISTIC THEORY OF MEAN FIELD GAMES WITH APPLICATIONS. To appear.
- [7] Friedman, A. (2013). DIFFERENTIAL GAMES. Courier Corporation.

- [8] Huang, M., Malhamé, R.P. Caines, P.E. (2006). *Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*. Communication in information and systems. Vol. 6, No. 3, pp. 221-252.
- [9] Huang, M., Caines, P.E., Malhamé, R.P. (2007). *Large-Population Cost-Coupled LQG Problems With Nonuniform Agents: Individual-Mass Behavior and Decentralized ϵ -Nash Equilibria*. IEEE Transactions on Automatic Control, 52(9), p. 1560-1571.
- [10] Huang, M., Caines, P.E., Malhamé, R.P. (2007). *The Nash Certainty Equivalence Principle and McKean-Vlasov Systems: an Invariance Principle and Entry Adaptation*. 46th IEEE Conference on Decision and Control, p. 121-123.
- [11] Huang, M., Caines, P.E., Malhamé, R.P. (2007). *An Invariance Principle in Large Population Stochastic Dynamic Games*. Journal of Systems Science & Complexity, 20(2), p. 162-172.
- [12] Lasry, J.-M., Lions, P.-L. *Jeux à champ moyen. I. Le cas stationnaire*. C. R. Math. Acad. Sci. Paris 343 (2006), no. 9, 619–625.
- [13] Lasry, J.-M., Lions, P.-L. *Jeux à champ moyen. II. Horizon fini et contrôle optimal*. C. R. Math. Acad. Sci. Paris 343 (2006), no. 10, 679–684.
- [14] Lasry, J.-M., Lions, P.-L. *Mean field games*. Jpn. J. Math. 2 (2007), no. 1, 229–260.
- [15] Lions, P.-L. *Cours au Collège de France*. www.college-de-france.fr.
- [16] Otto F., *The geometry of dissipative evolution equations: the porous medium equation*. Communications in Partial Differential Equations, 26 (1-2), 101-174, 2001.