

ON MEAN FIELD GAMES

Pierre-Louis LIONS

Collège de France, Paris

(joint project with Jean-Michel LASRY)

*Mathematical and Algorithmic Sciences Lab
France Research Center
Huawei Technologies*

Boulogne-Billancourt, March 27, 2018

- I INTRODUCTION
- II A REALLY SIMPLE EXAMPLE
- III GENERAL STRUCTURE
- IV THREE PARTICULAR CASES
- V OVERVIEW AND PERSPECTIVES
- VI MEANINGFUL DATA

I. INTRODUCTION

- New class of models for the average (Mean Field) behavior of “small” agents (Games) started in the early 2000’s by J-M. Lasry and P-L. Lions.
- Requires new mathematical theories.
- Numerous applications: economics, finance, social networks, crowd motions . .
- Independent introduction of a particular class of MFG models by M. Huang, P.E. Caines and R.P. Malhamé in 2006.
- A research community in expansion: mathematics, economics, finance. Economics: anonymous games, Krusell Smith!, joint projects with Ph. Aghion, J. Scheinkman, B. Moll, P-N. Giraud . . .
- Some written references but most of the existing mathematical material to be found in the Collège de France videotapes (4 × 18h) that can be downloaded . . . !

- Combination of Mean Field theories (classical in Physics and Mechanics) and the notion of Nash equilibria in Games theory.
- Nash equilibria for continua of “small” players: a single heterogeneous group of players (adaptations to several groups. . .).
- Interpretation in particular cases (but already general enough!) like process control of McKean-Vlasov. . .
- Each generic player is “rational” i.e. tries to optimize (control) a criterion that depends on the others (the whole group) and the optimal decision affects the behavior of the group (however, this interpretation is limited to some particular situations. . .).
- Huge class of models: agents \rightarrow particles, no dep. on the group are two extreme particular cases.

II. A REALLY SIMPLE EXAMPLE

- Simple example, not new but gives an idea of the general class of models (other “simple” exs later on).
- E metric space, N players ($1 \leq i \leq N$) choose a position $x_i \in E$ according to a criterion $F_i(X)$ where $X = (x_1, \dots, x_N) \in E^N$.
- Nash equilibrium: $\bar{X} = (\bar{x}_1, \dots, \bar{x}_N)$ if for all $1 \leq i \leq N$ \bar{x}_i min over E of $F_i(\bar{x}_1, \dots, \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, \dots, \bar{x}_N)$.
- Usual difficulties with the notion
- $N \rightarrow \infty$? simpler ?
- Indistinguishable players:

$$F_i(X) = F(x_i, \{x_j\}_{j \neq i}), F \text{ sym. in } (x_j)_{j \neq i}$$

- Part of the mathematical theories is about $N \rightarrow \infty$:

$$F_i = F(x, m) \quad x \in E, \quad m \in \mathcal{P}(E)$$

$$\text{where } x = x_i, \quad m = \frac{1}{N-1} \sum_{j \neq i} \delta_{x_j}$$

- “Thm”: Nash equilibria converge, as $N \rightarrow \infty$, to solutions of

$$\text{(MFG)} \quad \forall x \in \text{Supp } m, F(x, m) = \inf_{y \in E} F(y, m)$$

- Facts: i) general existence and stability results
- ii) uniqueness if $(m \rightarrow F(\bullet, m))$ monotone
- iii) If $F = \Phi'(m)$, then $(\min_{\mathcal{P}(E)} \Phi)$ yields one solution of MFG.

Example: $E = \mathbb{R}^d$, $F_i(X) = f(x_i) + g\left(\frac{\#\{j/|x_i - x_j| < \varepsilon\}}{(N-1)|B_\varepsilon|}\right)$

$g \uparrow$ aversion crowds, $g \downarrow$ like crowds

$$F(x, m) = f(x) + g(m * 1_{B_\varepsilon}(x))(|B_\varepsilon|^{-1})$$

$$\varepsilon \rightarrow 0 \quad F(x, m) = f(x) + g(m(x))$$

$$(MFG) \quad \text{supp } m \subset \text{Arg min} \left(f(x) + g(m(x)) \right)$$

– $g \uparrow$ uniqueness, $g \downarrow$ non uniqueness

$$\min \left\{ \int f m + \int G(m) / m \in \mathcal{P}(E) \right\}, \quad G = \int_0^z f(s) ds$$

– explicit solution if $g \uparrow$: $m = g^{-1}(\lambda - f)$, $\lambda \in \mathbb{R}$ s.t. $\int m = 1$

III. GENERAL STRUCTURE

- Particular case: dynamical problem, horizon T , continuous time and space, Brownian noises (both indep. and common), no intertemporal preference rate, control on drifts (Hamiltonian H), criterion dep. only on m
- $U(x, m, t)$ ($x \in \mathbb{R}^d$, $m \in \mathcal{P}(\mathbb{R}^d)$ or $\mathcal{M}_+(\mathbb{R}^d)$, $t \in [0, T]$ and $H(x, p, m)$ (convex in $p \in \mathbb{R}^d$)
- MFG master equation

$$\begin{cases} \frac{\partial U}{\partial t} - (\nu + \alpha)\Delta_x U + H(x, \nabla_x U, m) + \\ + \langle \frac{\partial U}{\partial m}, -(\nu + \alpha)\Delta m + \operatorname{div}(\frac{\partial H}{\partial p} m) \rangle + \\ -\alpha \frac{\partial U}{\partial m^2} (\nabla m, \nabla m) + 2\alpha \langle \frac{\partial}{\partial m} \nabla_x U, \nabla m \rangle = 0 \end{cases}$$

and $U|_{t=0} = U_0(x, m)$ (final cost)

- ν amount of ind. rand. , α amount of common rand.

- ∞ d problem !

- If $\nu = 0$ (ind): Nash N special case

using $x = x_i, m = \frac{1}{N-1} \sum_{j \neq i} \delta_{x_j}$

- Aggregation/decentralization: IF $H(x, p, m) = H(x, p) + F'(m)$ and $U_0 = \Phi'_0(m)$, then $U = \frac{\partial \Phi}{\partial m}$ solves MFG if Φ solves *HJB* on $\mathcal{P}(E)$ for the optimal control of a *SPDE*

- Particular case: many extensions and variants ...

IV. THREE PARTICULAR CASES

- ∞ d problem in general but reductions to finite d in two cases

1. Indep. noises ($\alpha = 0$)

int. along caract. in m yields

$$(MFGi) \begin{cases} \frac{\partial u}{\partial t} - \nu \Delta u + H(x, \nabla u, m) = 0 \\ u|_{t=0} = U_0(x, m(0)), m|_{t=T} = \bar{m} \\ \frac{\partial m}{\partial t} + \nu \Delta m + \operatorname{div} \left(\frac{\partial H}{\partial p} m \right) = 0 \end{cases}$$

where \bar{m} is given

FORWARD — BACKWARD system !

contains as particular cases: *HJB*, heat, porous media, *FP*, *V*, *B*, Hartree, semilinear elliptic, barotropic Euler ...

2. Finite state space ($i \leq i \leq k$)

$$\text{(MFGf)} \quad \frac{\partial U}{\partial t} + (F(x, U) \cdot \nabla) U = G(x, U), U|_{t=0} = U_0$$

(no common noise here to simplify ...)

$$x \in \mathbb{R}^k, U \rightarrow \mathbb{R}^k, F \text{ and } G : \mathbb{R}^{2k} \rightarrow \mathbb{R}^k$$

non-conservative hyperbolic system

Example: If $F = F(U) = H'(U)$, $G \equiv 0$

and if $U_0 = \nabla \varphi_0$ ($\varphi_0 \rightarrow \mathbb{R}$) then

– solve HJ

$$\frac{\partial \varphi}{\partial t} + H(\nabla \varphi) = 0, \varphi|_{t=0} = \varphi_0$$

– take $U = \nabla \varphi$, “ U solves” (MFGf) in this case

3. Another point of view

(Ω, \mathcal{F}, P) a “rich enough” proba. space \mathcal{H} Hilbert space of L^2 Random Variables

$$\Phi(m) = \Phi(X) \text{ if } \mathcal{L}(X) = m(X \rightarrow \mathbb{R}^d)$$

Then MFG may be written as

$$\frac{\partial U}{\partial t} + (\mathcal{F}(X, U) \cdot D)U = \mathcal{G}(X, U) + \alpha \Delta_d U$$

$$(+\nu D^2 U(G, G) \quad G \perp F_X)$$

$$\Delta_d U = \Delta_Z U(\cdot + Z)|_{Z=0} (Z \in \mathbb{R}^d),$$

where $U : \mathcal{H} \rightarrow \mathcal{H}$

Remarks: 1) MFG $U(X) \in F_X, \mathcal{L}(U(X)) = \mathcal{L}(U(Y))$
if $\mathcal{L}(X) = \mathcal{L}(Y)$

$$2) U(X) = \nabla_x U(x, \mathcal{L}(X))|_{x=X}$$

Allows to prove that the problem is well-posed in the “small”.

V. OVERVIEW AND PERSPECTIVES

Lots of questions, partial results exist, many open problems

- Existence/regularity:
 - (MFGi) “simple” if H “smooth” in m (or if H almost linear ...), OK if monotone (Zoom 1)
 - (MFGf) OK if (G, F) mon. on \mathbb{R}^{2k} or small time (Zoom 2)
- Uniqueness: OK if “monotone” or T small ...
- Non existence, non uniqueness, non regularity (!)
- Qualitative properties, stationary states and stability, comparison, cycles ...
- $N \rightarrow \infty$ (see above)
- Numerical methods (currently, 3 “general” methods and some particular cases)
- Variants: other noises, several populations ...
- random heterogeneity, partial info ...
- applications (MFG Labs ...)

- optimal stopping, impulsive controls
- intertemporal preference rates ($+\lambda \rightarrow \infty$ effective models)
- macroscopic limits
- ? Beyond MFG ? (fluctuations, LD, transitions)
- Two more $S.$ examples:
 - at which time will the meeting start ?
 - the (mexican) wave

ZOOM 1

$$(\text{MFGi}) \begin{cases} \frac{\partial u}{\partial t} - \nu \Delta u + H(x, \nabla u) = f(x, m) \\ u|_{t=0} = U_0(x), m|_{t=T} = \bar{m} \\ \frac{\partial m}{\partial t} + \nu \Delta m + \operatorname{div} \left(\frac{\partial H}{\partial p} m \right) = 0 \end{cases}$$

- $m \mapsto f(\bullet, m)$ smoothing operator
 \exists regular solution
- uniqueness if operator monotone or if T small
- $f(m(x)) \uparrow$: $\exists !$ regular solution $\nu > 0$
- $f(m(x)) \uparrow$: if $\nu = 0$ $m = f^{-1}(\frac{\partial u}{\partial t} + H(x, \nabla u))$

equation in m becomes quasilinear elliptic equation of second order ($x \in Q, t \in [0, T]$) with “elliptic” boundary conditions

$$u|_{t=0} = U_0(x), \quad \frac{\partial u}{\partial t} + H(\nabla u) = f(\bar{m}) \quad \text{if } t = T$$

ZOOM 2

$$(\text{MFGf}) \begin{cases} \frac{\partial u}{\partial t} + (F(x, U) \cdot \nabla) U = G(x, U) & x \in \mathbb{R}^d \\ U \rightarrow \mathbb{R}^d, U|_{t=0} = U_0(x) \end{cases}$$

- shocks (discontinuities of U) in finite time in general
- well-posed problem on $[0, T_{\max})$ ($T_{\max} \leq +\infty$)
- \exists !regular solution monotone in x if U_0 monotone and (G, F) monotone of $\mathbb{R}^{2,k}$ in $\mathbb{R}^{2k} (+ \dots)$
- + change of unknown functions:

$$\text{ex.: } \frac{\partial U}{\partial t} + (F(U) \cdot \nabla) U = 0$$

then $V = F(U)$ solves

$$\frac{\partial V}{\partial t} + (V \cdot \nabla)V = 0$$

max class of regularity

$$\forall \delta > 0, \inf_{x \in \mathbb{R}^d} \text{dist}(Sp(DV_0(x)), (-\infty, \delta]) > 0$$

($V_0 = F(U_0)$ gives the maximum class of regularity \approx composed of 2 monotone applications)

Remark: gives new results of regularity for Hamilton-Jacobi equations of the first order.

VI. MEANINGFUL DATA

- MFG Labs
- Practical expertise and models mainly for “big” data involving “people”
- New models that include classical clustering models in M.L. (K-mean, EM . . .), then algorithms
- No need for euclidean structures or for “a priori” distances

- Why “PEOPLE”

Ex. 1: Taxis

Ex. 2: Movies and Fb

People that are “close” will say they like movies that are “close”

→ consistency distance - like on items/people

- Even for “pure data” models make sense: data points become agents ... (in fact lots of terminology from Game Theory in M.L./Data Analysis)

- Clustering: classical K-Mean

set of points $\{x_1, \dots, x_v\}$ in $\mathbb{R}^d (N \gg 1)$

Find K points y_1, \dots, y_k s.t. \exists partition (A_1, \dots, A_K) of $\{1, \dots, N\}$ for which

$$\text{i) } |y_i - x_j| \leq |y_{i'} - x_j|, \forall j \in A_i, \forall i' \neq i$$

$$\text{ii) } y_i = (\#A_i)^{-1} \sum_{j \in A_i} x_j$$

MFG INTERPRETATION :

INTRODUCE • A GLOBAL CRITERION

$$F(u_1, \dots, u_k)$$

Ex: $\min(u_1, \dots, u_k)$

- K value functions (u_1, \dots, u_k)
- K “densities” (m_1, \dots, m_k)

f being the initial density of “data” (no need to restrict to “discrete” data)

SOLVE MFG: EXAMPLE

$$\rho u_i - \nu \Delta u_i + \frac{1}{2}(\nabla u_i)^2 = F_i(x; m_i)$$

$$\rho m_i - \nu \Delta m_i - \operatorname{div}(\nabla u_i m_i) = \rho \frac{\partial F}{\partial u_i} f$$

ex. $1_{(u_i < \min_{j \neq i} u_j)}$

BACK TO K-Mean

$$F_i = \frac{1 + \rho}{2} \left| x - \frac{\int x m_i}{\int m_i} \right|^2 - \nu d$$

then indeed : $u_i = \frac{1}{2} |x - y_i|^2$, $y_i = \frac{\int x m_i}{\int m_i}$ and

$$\int m_i = \int f \mathbf{1}_{(u_i < \min_{j \neq i} u_j)}, \quad \int x m_i = \int x f_i \mathbf{1}_{(u_i < \min_{j \neq i} u_j)}$$

Next, this allows to

- create lots of new models

- smooth clustering if needed, clusters within clusters, overlaps. . .
- no need for distances, no need for euclidean structure (choose criterion F , class criteria $F_i \rightarrow u_i \dots$)
- transposition to graphs easy (ODE's, massively //)

Remark : $-\Delta u + |\nabla u|^2 = e^u(+\Delta)e^{-u}$

$$e^{u_i} \sum_j (e^{-u_j} - e^{-u_i}) = \sum_j (e^{u_i - u_j} - 1)$$

- social networks equilibria: “distance on items” \longleftrightarrow
 “distance on users” \longleftrightarrow preferences \longleftrightarrow “distances on items” . . .
- interpretation of “deep learning” . . .