Mean Field Games: A Paradigm for Individual-Mass Interactions

Roland Malhamé

Department of Electrical Engineering
École Polytechnique de Montréal and GERAD

AMS Meeting Denver Colorado
January, 2020

Outline

- Origins: foundational papers, motivations and roots
- Initial mathematical results for a continuum of agents:
  (i) The LQG framework
  (ii) The nonlinear stochastic framework
- Issues of existence and uniqueness of solutions
- What it means for finite populations of interacting agents: the direction of approximation dilemma
- An application: Mean Field Games (MFG) and a model of fish schooling dynamics
- Recent extensions and future directions

Foundational Papers

- Control engineering literature: 2003
  - **M. Huang, P.E. Caines, R.P. Malhamé**, "Individual and Mass Behaviour...", Proc. the 42nd IEEE Conference on Decision and Control (CDC).
- Management science literature: 2005
  - **G.Y. Weintraub, C.L. Benkard, B. Van Roy**, "Oblivious Equilibrium: A Mean Field ", Advances in Neural Information Processing Systems, MIT Press.
- Mathematical literature: 2006, 2007
  - **J.-M. Lasry, P.-L. Lions**, "Jeux à champ Moyens I,II", C.Rendus Math (2006), "Mean Field Games", Jpn. J. Math (2007).
  - **M. Huang, R.P. Malhamé, P.E. Caines**, "Large Population Stochastic Dynamic Games: Closed Loop McKean-Vlasov ", Commu. Inf. Syst. (2006).

Mean Field Games: The Modeling Setup

- A large number of interacting dynamic agents with individual cost/utility functions.
- The agents are not necessarily homogeneous but share a certain degree of similarity in dynamics and cost functions.
- Agents interactions are anonymous, indifferent to their particular ordering, and for a given agent are dictated by empirical averages of functions of only a pair of states at a time: the state of the agent and that of another agent in the mass.
- Agents are rational, and share distributional information on dynamical parameters, cost functions and initial state of the mass of agents. Agents with identical parameters are interchangeable in their decision making.

Motivations: Modeling Side

Crowd-following agents:

- Nonmonopolistic pricing (Economics): price is global consequence of actions of a large number of suppliers and consumers. Deviations from average can be costly.
- Collective motion (Biology, Navigation): In herds, fish schools, pursuit of self interest but need for cohesion. In navigating large groups of tiny robots for exploration purposes, autonomy and coherence must be balanced.
- Societal dynamics: the need to balance conformity with originality; analysis of emerging behaviors.

Agoraphobic agents: crowd dynamics in situations of panic evacuation.

Agents interacting through large interconnected systems:

- Internet: when congested, equalize bandwidth share of agents.

Motivations: Why Infinite Populations?

- The case of N non-cooperative dynamic agents: Finite dynamic games are notoriously difficult because actions of individual agents tend to create responses from all other agents. As $N \to \infty$, influence of single individual vanishes $\Longrightarrow$ decoupling and hope for decentralization.
- The case of centralized control of N interconnected dynamic agents (technological paradigm): Deliberately structure as game with freedom over choice over utility function to achieve decentralized control as $N \to \infty$.

Intuition about large numbers making things easier already expressed in Von Neuman and Morgenstern (1944)!

Idealized economic behaviors occur with a continuum of players: Aumann (1964), perfect competition.

Statistical Mechanics Meets Dynamic Game Theory

Why Statistical Mechanics?

- It is a theory that starts from microscopic dynamic descriptions of interacting particles to proceed with the derivation of macroscopic properties.
- It studies limiting behaviors when dealing with a continuum of agents.

What does one learn from Statistical Mechanics?

- Mathematical techniques for the study of the limits of empirical measures as one moves to a continuum of particles.
- The notions of "propagation of chaos" and phase changes.
- That a "generic" microscopic particle process is key to population aggregation.

McKean-Vlasov Mean Field Game Formulation for N Agents

Agent i dynamics

$$dx_i = \frac{1}{N}\sum_{j=1}^{N} f_{\theta_i}(x_i, u_i, x_j)dt + \sigma dw_i$$

Agent i cost function

$$J_i(x_i^0, x_{-i}^0, u_i, u_{-i}) := E_{|x_i^0, x_{-i}^0} \int_0^T \frac{1}{N}\sum_{j=1}^{N} L[x_i, u_i, x_j]dt, \quad T < \infty$$

Assumptions

- Noise processes and initial conditions independent.
- θ_i parameters indicate potential non homogeneity and have a limiting empirical distribution.
- Initial conditions have a limiting empirical distribution.

Generic Mean Field Game Formulation for N Agents

Agent i dynamics

$$dx_i = f_{\theta_i}(x_i, u_i, \hat{\mu}_t^{(N)})dt + \sigma dw_i$$

Agent i cost function

$$J_i(x_i^0, x_{-i}^0, u_i, u_{-i}) := E_{|x_i^0, x_{-i}^0} \int_0^T L[x_i, u_i, \hat{\mu}_t^{(N)}]dt, \quad T < \infty$$

Assumptions: Same as before

Specifications: $\hat{\mu}_t^{(N)}$ is the empirical distribution of the agent states at time t. Its limit becomes the mean field.

An Important Special Case: LQG Mean Field Games (HMC'03,'07)

Agent i dynamics

$$dx_i = (A_{\theta_i} x_i + Bu_i + C\bar{x}^{(N)})dt + \sigma dw_i$$

Agent i cost function

$$J_i(x_i^0, x_{-i}^0, u_i, u_{-i}) := E_{|x_i^0, x_{-i}^0} \int_0^\infty e^{-\rho t} \Big\{ \big| x_i - \Phi(\bar{x}^{(N)}) \big|_Q^2 + \big| u_i \big|_R^2 \Big\} dt$$

Specifications

- $\bar{x}^{(N)} := \frac{1}{N} \sum_{j=1}^N x_j$, $\Phi(\bar{x}^{(N)}) := \Gamma \bar{x}^{(N)} + \eta$, $Q \geq 0$, $R > 0$ and $z|_Q^2 := z^T Q z$.
- Cost functions could be modified to include cross terms
- θ_i covers non homogeneous dynamics case
- $\Phi(\cdot)$ could be an arbitrary nonlinear function

Basic Intuitions: Individuals Versus the Mass

- Asymmetric individual-mass influences: Because of the $1/N$ factor, as N goes to infinity, an isolated individual action will not modify mass behavior. However, individuals all share the influence of the mass through the mass mean trajectory $m(t)$.
- The Nash certainty equivalence (NCE) principle as a guessing device (HCM 2007): Treat $m(t)$ as a deterministic given trajectory, and look for optimal response of the individual $=$ individual must solve an optimal tracking problem.
- A necessary fixed point property: if $m(t)$ should exist as a deterministic trajectory characteristic of a Nash equilibrium in an infinite population, then it is not sustainable unless it is collectively replicated as the time average of all agent states when optimally responding to $\Phi(m(\cdot))$ (when $\Phi(\cdot)$ is affine).

Scalar LQG case: An Illustration of the NCE Equations

Computation of optimal control law for agent with parameter a and mass trajectory $m(t)$

$$u_a(t) = -\frac{b}{r}\big(\Pi_a x_a(t) + s_a(t)\big) \text{ with } \rho\Pi_a = 2a\Pi_a - \frac{b^2}{r}\Pi_a^2 + 1 \quad \text{(Riccati)}$$

$$\rho s_a(t) = \frac{ds_a(t)}{dt} + as_a(t) - \frac{b^2}{r}\Pi_a s_a(t) - \Phi(m(t)) \quad \text{(Tracking equation)}$$

Collective replication requirement for $m(t)$ under the optimal feedback strategy

$$dx_a(t) = \Big[(a - \frac{b^2}{r}\Pi_a)x_a(t) - \frac{b^2}{r}s_a(t)\Big]dt + \sigma dw(t) \qquad \text{(the generic process!)}$$

$$\frac{d\bar{x}_a(t)}{dt} = (a - \frac{b^2}{r}\Pi_a)\bar{x}_a(t) - \frac{b^2}{r}s_a(t) \qquad \text{(the mean state conditional on } a\text{)}$$

$$m(t) = \int_{\mathcal{A}} \bar{x}_a(t)dF(a) \qquad \text{(the unconditional mean state)}$$

Scalar Nonlinear Case I: Intuitions Revisited, Explosion of Complexity

Asymmetric individual-mass influences: Because of the $1/N$ factor, as N goes to infinity, an isolated individual action will not modify mass behavior. However, individuals now share the influence of the mass through the evolution of not only the mean mass trajectory, but that of the "entire" population limiting empirical distribution $\mu_t(x)$, $t \geq 0$, i.e., a flow of measures!

Let $m_t(x)$, $t \geq 0$, be the family of Radon-Nykodim derivatives (probability density functions) associated with these measures.

Scalar Nonlinear Case II: Intuitions Revisited, Explosion of Complexity

- The Nash certainty equivalence (NCE) principle as a guessing device: Treat m_t, $t \geq 0$, as a "deterministic" given flow of pdf's and look for optimal response of the individual $=$ Individual must solve a "backward propagating m_t dependent HJB equation" to produce a local state feedback strategy which depends on the assumed entire flow of pdf's.

- A necessary fixed point property: If m_t, $t \geq 0$, should exist as a deterministic flow of pdf's characteristic of a Nash equilibrium in an infinite population, then it is not sustainable unless it is collectively replicated as the Radon-Nykodim derivative of the limiting flow of empirical measures of the infinite population when optimally responding in the context of the assumed flow of pdf's m_t, $t \geq 0$ $=$ Require that m_t satisfy the "forward Kolmogorov (or Fokker-Planck) equations" associated with the m_t, $t \geq 0$, dependent HJB based closed loop optimally controlled generic process.

Scalar Nonlinear MFG: The NCE Equations (LL'06,HMC'06, LL'07)

Computation of optimal control law for agent (under homogeneous dynamics) with assumed mass pdf flow m_t, $t \geq 0$

$$-\frac{\partial V}{\partial t} = \inf_{u \in \mathcal{U}} \Big[f(x,u,m_t)\frac{\partial V}{\partial x} + L(x,u,m_t) \Big] + \frac{\sigma^2}{2}\frac{\partial^2 V}{\partial x^2} \quad \text{(HJB equation)}$$

$$V(x,T) = 0 \quad \text{(Backwards propagation)} \qquad (t,x) \in [0,T] \times \mathbb{R},$$

$$u^*_m(t,x) := \psi(t,x,m_\tau : \tau \in [0,T]) \quad \text{(Candidate optimal control)}$$

Collective replication requirement for m_t, $t \geq 0$, under the optimal feedback strategy

$$dx_t = f\big(x_t, u^*_m(t,x), m_t\big)dt + \sigma dw_t \qquad \text{(The generic closed loop process)}$$

$$\frac{\partial m_t}{\partial t} = -\frac{\partial \big[f\big(x_t, u^*_m(t,x), m_t\big) m_t \big]}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 m_t}{\partial x^2} \qquad \text{(Fokker-Planck equation)}$$

$$m_t(x)|_{t=0} = m_0(x) \qquad \text{(Forward propagation!)}$$

Existence and Uniqueness: The HCM'07 Approach for LQG Case

MFG (NCE) LQG equations

$$\rho s_a(t) = \frac{ds_a(t)}{dt} + a s_a(t) - \frac{b^2}{r}\Pi_a s_a(t) - \Phi(m(t)) \quad \text{(Tracking equation)}$$

$$\frac{d\bar{x}_a(t)}{dt} = (a - \frac{b^2}{r}\Pi_a)\bar{x}_a(t) - \frac{b^2}{r}s_a(t) \quad \text{(the mean state conditional on } a\text{)}$$

$$m(t) = \int_{\mathcal{A}} \bar{x}_a(t) dF(a) \quad \text{(the unconditional mean state)}$$

Idea: Look at the map that goes from $m(t)$ in the tracking differential equation, back to $m(t)$ in the unconditional mean state equation, and develop sufficient conditions for this map to be a contraction operator from the space of bounded continuous functions of time on itself. Guarantee at the same time existence and uniqueness (Theorems 4.3, 4.4 in Huang, Caines and Malhamé IEEE TAC, 2007)

See also ergodic analysis by Li-Zhang (2008), M. Bardi (2012).

Existence and Uniqueness: The General Case (HMC'06)

MFG finite horizon equations

$$-\frac{\partial V}{\partial t} = \inf_{u\in\mathcal{U}} \left[f(x,u,m_t)\frac{\partial V}{\partial x} + L(x,u,m_t) \right] + \frac{\sigma^2}{2}\frac{\partial^2 V}{\partial x^2} \quad \text{(HJB equation)}$$

$$\frac{\partial m_t}{\partial t} = -\frac{\partial \left[f\left(x_t, u_m^*(t,x), m_t\right) m_t \right]}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 m_t}{\partial x^2} \qquad \text{(Fokker-Planck equation)}$$

$$V(x,T) = 0, \qquad m_t(x)|_{t=0} = m_0(x)$$

Idea (Huang, Malhamé, Caines, CIS 2006):

- Starting with an arbitrary flow of measures with weak but sufficient "continuity" properties, and under assumptions of strict convexity and adequate smoothness of Hamiltonian in HJB equation, show that control law exists uniquely and is Lipschitz continuous.
- Using McKean-Vlasov equations theory and standard assumptions on system dynamics, show existence of unique flow of measures for FKP equation for which the assumed continuity properties persist.
- Consider combined HJB-FKP operator from measure flows, back to measure flows. Under sufficient conditions it is a contraction and existence /uniqueness are guaranteed.

Existence and Uniqueness: The General Case (LL'06)

Idea (Lasry-Lions CR Acad.Sc. 2006, Cardaliaguet MFG notes 2012):

- Existence: Pretty much the same two steps analysis of starting from a suitably continuous arbitrary flow of measures but also living on a convex set feeding into the HJB equation. Existence and uniqueness of optimal control solution of the HJB equation is based on transformation to heat equation for dynamics affine in control and cost quadratic in control.

  Properties of Itô differential equations are used to conclude that the corresponding FPK equation has a unique solution. The combined operator is shown to be continuous. Schauders fixed point theorem is used for existence. Can handle more general Hamiltonians.

- Uniqueness: separable Hamiltonian and agoraphobic agents

$$H(x, m_t, \frac{\partial V}{\partial x}) := S(x, \frac{\partial V}{\partial x}) + T(x, m_t)$$

$$\int_{\mathbb{R}} \big(T(x, m_1) - T(x, m_2)\big) d(m_1 - m_2)(x) > 0, \ \forall m_1, m_2 \in P_1, m_1 \neq m_2$$

What It Means for Games with N Large but Finite

- An approach directly inspired by the methods of statistical mechanics (LL'06,'07):

  (i) Start studying Nash equilibria for finite agent systems.

  (ii) Obtain MFG equations as characteristic of limiting behavior of finite agent systems as $N \to \infty$ and study existence and uniqueness properties.

  (iii) Under "sufficient regularity" of the relevant functions, and uniqueness of solutions to MFG equations, show that difference of performance relative to Nash equilibrium of infinite population can be made arbitrarily small by increasing N sufficiently.

- A more direct approach (HMC'06, HCM'07):

  (i) Through intuitive analysis, obtain form of MFG equations and carry out corresponding existence and uniqueness analysis.

  (ii) Characterize the family of decentralized Nash equilibrium inducing control laws for infinite population.

  (iii) Apply infinite population based control policies to all agents but one in a finite game, and study the limits on its cost improvement.

Discussion and Nash Equilibrium Approximation Concepts

- Statistical mechanics versus large population games:

  In statistical mechanics, no control on dynamics $\Longrightarrow$ Only an approximation problem $\Longrightarrow$ Essential to move from finite to infinite.

  In large population games, one looks for Nash equilibrium inducing feedback control strategies $\Longrightarrow$ Given any set of feedback strategies, possible to study directly distance from Nash equilibrium.

- Nash equilibrium approximation concepts: ϵ-Nash equilibrium

$$J_i(u_i^*, u_{-i}^*) - \varepsilon \leq \inf_{u_i} J_i(u_i, u_{-i}^*) \leq J_i(u_i^*, u_{-i}^*)$$

where $\epsilon \to 0$ as $N \to \infty$, with

(i) Full information: u_i depends on $(t, x_1, x_2, \cdots, x_N)$

(ii) Local information: u_i depends on (t, x_i).

Results for MFG Induced Decentralized Control Laws

- Full information ϵ-Nash property holds for:

  (i) Infinite horizon exponentially discounted LQG case with linear or nonlinear coupling in the cost and dynamics (HMC 2006, HCM 2007)

  (ii) Finite horizon nonlinear case with dynamic interactions of agents with mean field of the form $[f(x,u)g(m) + h(m)]$, and arbitrary individual running cost $L(x,u,m)$ (HMC 2006)

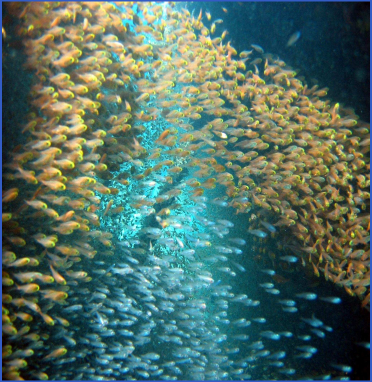
  (iii) Ergodic MFG, i.e. considering time average cost, motion occurring on a toroidal structure, and all functions being periodic, and under sufficient conditions for uniqueness (LL 2006, Cardaliaguet MFG notes 2012)

- Local information ϵ-Nash property holds for:

  Finite horizon general nonlinear case (HMC 2006, LL 2006, Cardaliaguet MFG notes 2012).

MFG and a Model of Fish Schooling Dynamics

- **M. Nourian, R.P. Malhamé, M. Huang, P.E. Caines**, "Mean field (NCE) formulation of estimation based leader-follower collective dynamics", International J. Robotics and Automation (2011).



MFG and a Model of Fish Schooling Dynamics (NMHC'11)

Individual dynamics:

$$dx_i = (Ax_i + Bu_i)dt + \sigma dw_i, \quad 1 \le i \le N.$$

Cost for the i-th leader:

$$J_i(x_i^0, x_{-i}^0, u_i, u_{-i}) = E\int_0^\infty e^{-\rho t}\Big\{\Big|x_i - \Big(\lambda h + (1-\lambda)\frac{1}{N}\sum_{k=1}^{N} x_k\Big)\Big|^2 + |u_i|_R^2\Big\}dt$$

Cost for the i-th follower:

$$J_i(x_i^0, x_{-i}^0, u_i, u_{-i}) = E\int_0^\infty e^{-\rho t}\Big\{\Big|x_i - \frac{1}{N_1}\sum_{k\in\mathcal{L}} x_k\Big|^2 + |u_i|_R^2\Big\}dt$$

$\mathcal{L}$: The set of leaders with cardinal N_1

MFG and a Model of Fish Schooling Dynamics

- $h(\cdot)$ is intended for a certain reference trajectory assumed to be known to both leaders and followers (the case of reference trajectory unknown to the followers: NCMH, IEEE TAC 2012).
- Leader-follower architecture is of interest in animal behaviour analysis and also in group motion of multiple autonomous robots.
- Direct tracking of h may result in poor cohesiveness during transient phase due to different initial conditions.
- In many problems of interest (e.g., swarming, flocking of animals or schooling of fish), group cohesiveness is important at all stages (Couzin et. al. Nature, Feb. 2005).

MFG and a Model of Fish Schooling Dynamicse

Assumption: There exists $\alpha \in (0,1)$ such that $\lim_{N\to\infty} \frac{N_1}{N} = \alpha$, i.e., the proportion of leaders is fixed.

Let the mean field centroid trajectory for leaders and followers be $\bar{x}_L(\cdot)$ and $\bar{x}_F(\cdot)$

MFG (NCE) LQG equations for the leaders

$$\rho s_L(t) = \frac{ds_L(t)}{dt} + A^T s_L(t) - \Pi B R^{-1} B^T s_L(t) - x_L^*(t)$$

$$\frac{d\bar{x}_L(t)}{dt} = (A - BR^{-1}B^T\Pi)\bar{x}_L(t) - BR^{-1}B^T s_L(t)$$

$$x_L^*(t) = \lambda h(t) + (1-\lambda)\big(\alpha \bar{x}_L(t) + (1-\alpha)\bar{x}_F(t)\big)$$

- Individual control action for the leaders

$$u_i = -R^{-1}B^T(\Pi x_i + s_L)$$

MFG and a Model of Fish Schooling Dynamics

MFG (NCE) LQG equations for the followers

$$\rho s_F(t) = \frac{ds_F(t)}{dt} + A^T s_F(t) - \Pi B R^{-1} B^T s_F(t) - \bar{x}_L(t)$$
$$\frac{d\bar{x}_F(t)}{dt} = (A - BR^{-1}B^T\Pi)\bar{x}_F(t) - BR^{-1}B^T s_F(t)$$

- Individual control action for the followers

$$u_i = -R^{-1}B^T(\Pi x_i + s_F)$$

More recent developments I: MFG Cooperative Games

Towards mean field LQG cooperative games: Decentralized socially optimal control laws (HCM, IEEE TAC July 2012)

Agent i dynamics: $dx_i = (A_{\theta_i} x_i + Bu_i + C\bar{x}^{(N)})dt + \sigma dw_i$

Social cost to be optimized:

$$J(x^0, u) := \sum_{i=1}^{N} J_i(x_i^0, x_{-i}^0, u_i, u_{-i})$$

$$= \sum_{i=1}^{N} E_{|x_i^0, x_{-i}^0} \int_0^\infty e^{-\rho t} \Big\{ \big| x_i - \Phi(\bar{x}^{(N)}) \big|_Q^2 + \big| u_i \big|_R^2 \Big\} dt$$

- As $N \to \infty$, small individual agent actions are reflected in the individual costs of all other agents (one way coupling of individual and mass property of non cooperative MFGs is lost!).
- Consider person by person by person optimization as in team theory together with MFG collective replication requirement.
- Individual asymptotic optimal control laws are still decentralized and are full information ϵ-optimal!

More recent developments II

- Risk sensitive MFG games: H.Tembine, Q. Zhu, T. Basar, "Risk -Sensitive Mean Field Stochastic Differential Games", Proceedings 2011 IFAC world Congress.
- Towards oligopolies within MFG games:

  - LQG case: M. Huang, "Large-Population LQG Games Involving a Major Player: the Nash Certainty Equivalence Principle", SIAM J. on Control and Optimization, 2010

  - Nonlinear case: M. Nourian, P. E. Caines, "ϵ-Nash Mean Field Game Theory for Nonlinear Stochastic Dynamical Systems with Major and Minor agents", preprint, 2012.
- Nonlinear MFG models:

  - O. Guéant, J.-M. Lasry, P.-L. Lions. "Mean Field Games and Oil Production", preprint, 2010.

  - H. Yin, P. G. Mehta, S. Meyn, U. V. Shanbhag, "Synchronization of Coupled Oscillators is a Game", IEEE TAC May 2012.

  - V. N. Kolokoltsov, J. Li, W. Yang, "Mean Field Games and Nonlinear Markov Processes", Arxiv, 2011.

Summary

- MFG's are a convenient approach for crystallizing the interaction of an individual agent with a large population of similar agents which individually carry a vanishing decision weight.

- In general, they lead to ϵ-Nash or ϵ-optimal decentralized feedback control strategies with the degree of error typically decaying as $1/\sqrt{N}$.

- Decentralization is possible only if the necessary a priori data is available i.e. typically, initial agent states empirical distributions as well as dynamic parameters empirical distributions (identification and on line learning issues (LQG case: A. C. Kizilkale, P. E. Caines, IEEE TAC 2012)).

Summary (cnt)

- MFGs exploit to the fullest extent the predictability stemming from the law of large numbers under conditions of propagation of chaos. They do so by constructing a generic stochastic agent process.
- With LQG-MFGs you almost get a free lunch because the mass is aggregated into a single agent. Many applications are expected.
- Numerical methods are bound to become a very important issue for the viability of nonlinear MFG analysis in general.