Quantitative Methods for the Mean Field Limit Problem

Zhenfu Wang
Joint work with Didier Bresch and Pierre-Emmanuel Jabin

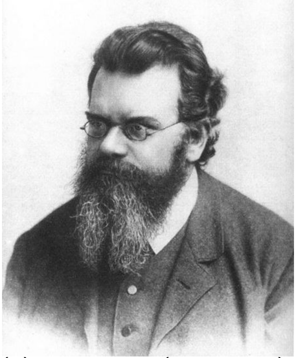
Department of Mathematics
University of Pennsylvania

March 29, 2020
International Workshop on Interacting Particle Systems, SJTU

The Large N limit

From Microscopic Descriptions to Macroscopic Descriptions. Part of Hilbert's 6th Problem.

(a) Maxwell (1831-1879)



(b) Boltzmann (1844-1906)

Deriving Boltzmann/Landau and Vlasov-type Kinetic Equations. Bose-Einstein Condensation, Hydrodynamic Limit, Thermodynamic Limit... And Mean Field Limit.

Newton Dynamics (2nd order system)

Consider the classical Newton dynamics for N indistinguishable point particles in the mean filed scaling in the classical regime. Denote (X_i, V_i) the position and the velocity of particle number i. Then

$$\dot{X}_i = V_i, \quad \dot{V}_i = \frac{1}{N}\sum_{j\neq i} K(X_i - X_j), \quad i = 1, 2, \cdots, N.$$

where $X_i, V_i \in \mathbb{R}^3$.
As $N \to \infty$, the expected PDE is the famous Vlasov(-Poisson) equation

$$\partial_t f + v \cdot \nabla_x f + K \star_x \rho \cdot \nabla_v f = 0,$$

where $\rho(t,x) = \int_{\mathbb{R}^3} f(t,x,v)\,\mathrm{d}v$.
Gravitational or Coulomb force: $K(x) = \pm \frac{x}{|x|^3}$, i.e. the inverse-square law. Still open!
Hauray & Jabin ('07 and '15) Lazarovici & Pickl ('15), Jabin & W. ('16), Serfaty & Duerinckx ('18).

Our Setting: 1st order system

Consider the weakly interacting particle system for N indistinguishable point particles. Denote $X_i \in E = \Pi^d$ (torus) the position of particle number i. The dynamics reads

$$\mathrm{d}X_i = \frac{1}{N} \sum_{j \neq i} K(X_i - X_j)\,\mathrm{d}t + \sqrt{2\sigma_N}\,\mathrm{d}W_t^i, \quad i = 1, 2, \cdots, N, \quad (IPS)$$

where $X_i \in E$, and W^i are N independent Brownian motions which may model random collisions on particles with rate $\sqrt{2\sigma_N}$. In particular, if $\sigma_N = 0$, the system (IPS) is deterministic. The interaction kernels K model 2−body interaction forces between particles.
The expected limit PDE reads (as $N \to \infty$)

$$\partial_t \bar{\rho} + \mathrm{div}_x (\bar{\rho}\, K \star_x \bar{\rho}) = \sigma \Delta_x \bar{\rho}. \qquad (MFD)$$

Goal: Establish and quantify the convergence.

Particle Systems

Individual based models (First Principle) are conceptually simple.
Examples:

- Physics (ions and electrons in plasmas, molecules in a fluid, galaxies in large scale cosmological models);
- Bio-sciences (modeling collective behaviors like flocking/swarming);
- Economics or Social Science (Opinion dynamics, consensus model, mean field games).
- Distribution Sampling Algorithm (Stein Variational Gradient Descent. Sinkhorn Descent (Hassani, Shen and W. ('20)), where $\dot{x}_i = -\nabla f_{\mu_N}(x_i)$.) Neural Networks.

Difficulty:

- The number N of particles are usually very large: Analytically and computationally complicated. Note that $N \sim 10^{25}$ in typical physical settings. **Curse of dimensionality!**

Examples of Kernels

Classical Results: McKean ('67), Braun& Hepp ('77), Dobrusin ('79), Sznitmann ('91)... $K \in W^{1,\infty}$ (K is Lipschitz!) (Coupling Method).

Examples of Singular Kernels:

- Biot-Savart Law with $K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2}$.
- The Poisson kernels $K(x) = \pm C_d \frac{x}{|x|^d}$. (Repulsive or attractive.)
- $K = -\nabla V$, with

$$V(x) = \lambda \log |x| + V_e(x), \quad \lambda > 0.$$

The classical methods fail for systems with some singular kernels. But they are still very useful in many applications.

Recent Results (1st order systems)

- 2D Euler: Goodman, Hou and Lowengrub ('90). Schochet ('96), Hauray ('09). Well-prepared initial data. Point vortex system.
- 2D Navier-Stokes: Osada ('87), Fournier, Hauray and Mischer ('14). Long ('98). (The Biot-Savart kernel $K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2}$. Compactness argument.)
- Patlak-Keller-Segel: Haskovec and Schmeiser ('11), Fournier and Jourdain ('15) (very sub-critical regime, no rate...) Similar setting: Liu & Yang ('16), Li, Liu & Yu ('19).
- 1st order systems with $K \in W^{-1,\infty}$ (but also $\mathrm{div} K \in W^{-1,\infty}$). Jabin and W. ('18). (Include 2D Navier-Stokes and 2D Euler).
- Coulomb (like) flows or conservative flows, deterministic case. Serfaty ('18).
- Stochastic systems with a large class of singular interactions. Bresch, Jabin and W. ('19).

The Liouville Equation (Master Equation)

Key object: the coupled law of $N-$particle $\rho_N(t, x_1, \cdots, x_N)$ governed by the Liouville equation

$$\partial_t \rho_N + \frac{1}{N} \sum_{i=1}^{N} \sum_{j \neq i} \mathrm{div}_{x_i} \left(\rho_N K(x_i - x_j) \right) = \sigma_N \sum_{i=1}^{N} \Delta_{x_i} \rho_N.$$

Note: $\rho_N \in \mathcal{P}_{sym}(E^N)$ (**Symmetric probability measures**) but not *experimentally measurable*. The observable (statistical information: temperature, pressure for instance) is contained in the marginals $\rho_{N,k}$ of ρ_N as

$$\rho_{N,k}(t, x_1, \cdots, x_k) = \int_{E^{N-k}} \rho_N(t, x_1, \cdots, x_N) \, \mathrm{d}x_{k+1} \cdots \mathrm{d}x_N,$$

for fixed $k = 1, 2, \cdots$.
The evolution of $\rho_{N,k}$ involves $\rho_{N,k+1}$. BBGKY hierarchy.

Formal Derivation assuming Molecular Chaos

Integrating the Liouville Eq. w.r.t. $x_2, \cdots, x_N$ and using the symmetry of ρ_N,

$$\partial_t \rho_{N,1} + \frac{N-1}{N} \int_E \mathrm{div}_x(\rho_{N,2} K(x-y))\, \mathrm{d}y = \sigma_N \Delta_x \rho_{N,1}.$$

If we assume that $\rho_{N,2}(x,y) = \rho_{N,1}(x)\rho_{N,1}(y)$ (Molecular Chaos), then we obtain the limit PDE (MFD) as $N \to \infty$,

$$\partial_t \rho_{\infty,1} + \mathrm{div}_x(\rho_{\infty,1} K \star_x \rho_{\infty,1}) = \sigma \Delta_x \rho_{\infty,1}.$$

Even initially $\rho_{N,2}(0) = \rho_{N,1}(0)^{\otimes 2}$, as long as you run the dynamics of the particle system, $\rho_{N,2}(t) \neq \rho_{N,1}(t)^{\otimes 2}$. Correlation exists since particles do interact!

Relaxation: **Kac's chaos** ('56). To derive the space homogeneous Boltzmann equation.

Propagation of Chaos

- Tensorized/Chaotic initial law: $\rho_N^0 = \bar{\rho}_0^{\otimes N}$.

Definition 1 (Kac's chaos)

Let $E = \Pi^d$. A sequence $(\rho_N)_{N\geq 2}$ of symmetric probability measures, *i.e.* $\rho_N \in \mathcal{P}_{Sym}(E^N)$, is said to be $\bar{\rho}-$chaotic for a probability measure $\bar{\rho}$ on E, if for any fixed $k = 1, 2, 3, \cdots$, $\rho_{N,k} \rightharpoonup \bar{\rho}^{\otimes k}$, as $N \to \infty$.

"Asymptotic independence" for a finite group.

Definition 2 (Propagation of (Kac's) chaos)

The diagram commutes.

$$\begin{array}{ccc} \rho_{N,k}(0) & \rightharpoonup & \bar{\rho}^{\otimes k}(0) \\ \Downarrow_{\text{IPS}} & & \Downarrow_{\text{MFD}} \\ \rho_{N,k}(t) & \rightharpoonup & \bar{\rho}^{\otimes k}(t) \end{array}$$

We don't use the hierarchy. We adopt a more straightforward way.

From Relative Entropy to Propagation of Chaos

We use the (scaled) relative entropy to quantify chaos

$$0 \le \mathcal{H}_N(\rho_N|\bar{\rho}^{\otimes N})(t) = \frac{1}{N}\int_{E^N} \rho_N \log \frac{\rho_N}{\bar{\rho}^{\otimes N}} \, \mathrm{d}x_1 \cdots \mathrm{d}x_N.$$

Thanks to the monotonicity of the (scaled) relative entropy

$$\mathcal{H}_k(\rho_{N,k}|\bar{\rho}^{\otimes k}) := \frac{1}{k}\int_{E^N} \rho_{N,k} \log \frac{\rho_{N,k}}{\bar{\rho}^{\otimes k}} \, \mathrm{d}x_1 \cdots \mathrm{d}x_k \le \mathcal{H}_N(\rho_N|\bar{\rho}^{\otimes N})$$

and the classical Csiszár-Kullback-Pinsker inequality

$$\|\rho_{N,k} - \bar{\rho}^{\otimes k}\|_{L^1} \le \sqrt{2k\mathcal{H}_k(\rho_{N,k}|\bar{\rho}^{\otimes k})},$$

one can obtain *propagation of chaos* given a vanishing sequence of $H_N(\rho_N|\bar{\rho}^{\otimes N})$.
Ben Arous & Zeitouni ('99).

The Previous Result

Theorem (Jabin & W. ('18))

Assume that $K \in \dot{W}^{-1,\infty}(\Pi^d)$ *with* $\mathrm{div} K \in \dot{W}^{-1,\infty}$. *Assume that* $\sigma_N \equiv \sigma > 0$. *Assume finally that* $\bar{\rho} \in L^\infty([0,\ T],\ W^{2,p}(\Pi^d))$ *for any* $p < \infty$ *solves (MFD) with* $\inf \bar{\rho} > 0$ *and* $\int_{\Pi^d} \bar{\rho} = 1$. *Then*

$$\mathcal{H}_N(\rho_N \mid \bar{\rho}_N)(t) \leq e^{\bar{M}\,(\|K\|+\|K\|^2)\,t} \left(\mathcal{H}_N(\rho_N^0 \mid \bar{\rho}_N^0) + \frac{1}{N}\right),$$

where we denote $\|K\| = \|K\|_{\dot{W}^{-1,\infty}} + \|\mathrm{div} K\|_{\dot{W}^{-1,\infty}}$ *and* $\bar{M}$ *is a universal constant.*

This result applies to the Biot-Savart law, i.e. $K(x) = \frac{1}{2\pi}\frac{x^\perp}{|x|^2}$, since $K = \mathrm{div} V$ with

$$V = \frac{1}{2\pi}\begin{bmatrix} -\arctan\frac{x_1}{x_2} & 0 \\ 0 & \arctan\frac{x_2}{x_1} \end{bmatrix}.$$

Ideas of the proof

We write the tensorized law $\bar{\rho}_N := \bar{\rho}^{\otimes N}$ and compute the time evolution of the relative entropy

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_N(\rho_N|\bar{\rho}_N)(t) \leq -\frac{\sigma}{N}\int_{\Pi^{dN}} |\nabla \log \frac{\rho_N}{\bar{\rho}_N}|^2 \,\mathrm{d}\rho_N + \int_{\Pi^{dN}} \Big(\frac{1}{N^2}\sum_{i,j=1}^{N} \phi(x_i, x_j)\Big)\,\mathrm{d}\rho_N,$$

where

$$\phi(x, y) = \nabla \log \bar{\rho}(x) \cdot (K \star \bar{\rho}(x) - K(x-y)) + (\,\mathrm{div} K \star \bar{\rho}(x) - \,\mathrm{div} K(x-y)).$$

Using symmetrization, i.e. taking $\frac{1}{2}(\phi(x,y)+\phi(y,x))$ as the new $\phi(x,y)$, one writes

$$\phi(x,y) = -\frac{1}{2}K(x-y)\cdot(\nabla \log \bar{\rho}(x) - \nabla \log \bar{\rho}(y)) - \,\mathrm{div} K(x-y) \\ + \text{Bounded Terms}.$$

Consider the 2D Navier-Stokes and the 2D Euler case. Then the kernel K is the Biot-Savart kernel, which is divergence free, i.e. $\mathrm{div}_x K = 0$. Dropping the Fisher information term,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_N(\rho_N|\bar{\rho}_N)(t) \leq \int_{\Pi^{dN}} \Big(\frac{1}{N^2}\sum_{i,j=1}^{N} \phi(x_i, x_j)\Big)\,\mathrm{d}\rho_N \quad (\sim O(1) \text{ a prior!})$$

where after symmetrization, $\phi \in L^\infty$ and more importantly

$$\int_E \phi(x,y)\bar{\rho}(y)\,\mathrm{d}y = 0, \forall x, \quad \int_E \phi(x,y)\bar{\rho}(x)\,\mathrm{d}x = 0, \forall y.$$

Recall a Jensen-type inequality, i.e. for any parameter $\eta > 0$,

$$\int \rho_N \Phi_N \leq \frac{1}{\eta}\mathcal{H}_N(\rho_N|\bar{\rho}_N) + \frac{1}{\eta}\frac{1}{N}\log\int \bar{\rho}_N \exp\left(\eta N \Phi_N\right).$$

GOAL: Show the 2nd term is $o(1)$ as $N \to \infty$.

Theorem (Uniform in N large deviation type estimate)

We have

$$\sup_{N\geq 2}\int_{\Pi^{dN}}\bar\rho^{\otimes N}\exp\left(\frac{1}{N}\sum_{i,j=1}^{N}\phi(x_i,x_j)\right)\mathrm{d}X^N$$
$$=\sup_{N\geq 2}\int_{\Pi^{dN}}\bar\rho^{\otimes N}\exp\left(N\int_{\Pi^{2d}}\phi(x,y)(\mathrm{d}\mu_N-\mathrm{d}\bar\rho)^{\otimes 2}(x,y)\right)\mathrm{d}X^N\leq C<\infty,$$

provided that $\|\phi\|_{L^\infty}\leq c_0$ *and*

$$\int_E\phi(x,y)\bar\rho(y)\,\mathrm{d}y=0,\forall x,\quad \int_E\phi(x,y)\bar\rho(x)\,\mathrm{d}x=0,\forall y.$$

Ben Arous and Braunaud ('90): with ϕ continuous.
We need the estimate directly for discontinuous ϕ.
Carefully use two cancellation rules. Law of Large Numbers but for "Double Indices". A recent proof using martingales by Lim, Lu and Nolen ('19).

Discussion

Now we focus on gradient flows, i.e. $K = -\nabla V$.

- Relative entropy (Jabin and W., ('18)) : Less structure and less singularity.

  Recall that there is a term

$$-\frac{1}{N^2}\sum_{i\neq j}\int_{\Pi^{dN}} \mathrm{div} K(x_i - x_j)\,\mathrm{d}\rho_N$$

  in the time evolution of the relative entropy.

- Modulated Energy (Serfaty (with an appendix with Duerinckx) ('18)): More structure and also more singular (Riesz potentials+ possible perturbation!). Deterministic flows.

  Serfaty's modulated (potential) energy is defined as

$$\int_{x\neq y} V(x-y)(\,\mathrm{d}\mu_N(x) - \bar{\rho}(x))(\,\mathrm{d}\mu_N(y) - \bar{\rho}(y)).$$

Modulated Free Energy

Idea: introducing weights G_N and $G_{\bar{\rho}_N}$ in the relative entropy to cancel the term $\mathrm{div} K$ in its time evolution

$$E_N(\rho_N|\bar{\rho}_N) = \frac{1}{N} \int_{\Pi^{dN}} \rho_N \log \left(\frac{\rho_N/G_N}{\bar{\rho}_N/G_{\bar{\rho}_N}} \right) \mathrm{d}x_1 \cdots \mathrm{d}x_N,$$

where G_N is the Gibbs measure, $G_{\bar{\rho}_N}$ is a tilted Gibbs measure by the limit $\bar{\rho}$.
In an equivalent way

$$E_N(\rho_N|\bar{\rho}_N) = \mathcal{H}_N(\rho_N|\bar{\rho}_N) + \mathcal{K}_N(\rho_N|\bar{\rho}_N),$$

with

$$\mathcal{K}_N(\rho_N|\bar{\rho}_N) = \frac{1}{2\sigma} \mathbb{E}_{\rho_N} \int_{x \neq y} V(x-y)(\mathrm{d}\mu_N(x) - \mathrm{d}\bar{\rho}(x))(\mathrm{d}\mu_N(y) - \mathrm{d}\bar{\rho}(y)).$$

Time Evolution of E_N

Suppose that V is an even function. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}E_N(\rho_N|\bar{\rho}_N) \leq -\frac{\sigma}{N}\int_{\Pi^{dN}} \left|\nabla \log \frac{\rho_N}{\bar{\rho}_N} - \nabla \log \frac{G_N}{G_{\bar{\rho}_N}}\right|^2 \mathrm{d}\rho_N$$

$$-\frac{1}{2}\int_{\Pi^{dN}} \mathrm{d}\rho_N \int_{x\neq y} \nabla V(x-y)\cdot\left(\nabla \log \frac{\bar{\rho}}{G_{\bar{\rho}}}(x) - \nabla \log \frac{\bar{\rho}}{G_{\bar{\rho}}}(y)\right)(\mathrm{d}\mu_N - \mathrm{d}\bar{\rho})^{\otimes 2},$$

where $G_{\bar{\rho}}(x) = \exp\left(-\frac{1}{\sigma}V\star\bar{\rho}(x) + \frac{1}{2\sigma}\int_{\Pi^d} V\star\bar{\rho}\bar{\rho}\right)$ and hence

$$\psi(x) := \nabla \log \frac{\bar{\rho}}{G_{\bar{\rho}}}(x) = \nabla \log \bar{\rho}(x) + \frac{1}{\sigma}\nabla V\star\bar{\rho}(x).$$

Derivation of the Patlak-Keller-Segel system

Take $K = -\nabla V$ and $V = \lambda \log |x| + V_e(x)$ where $\lambda > 0$. Then (MFD) is the famous Patlak-Keller-Segel (PKS) model, which is one of the first models of chemotaxis for micro-organisms. Note that in 2D, V is the **attractive Poisson** potential.

Theorem (Bresch, Jabin & W. ('19))

Given the potential V and $K = -\nabla V$. Assume that $\rho_N \in L^\infty(0, T; L^1(\Pi^{dN}))$ is an entropy solution to the Liouville equation, with initial condition that $\rho_N(0) = \bar\rho^{\otimes N}(0)$. Assume that $\bar\rho \in L^\infty(0, T; W^{2,\infty}(\Pi^d))$ solves (MFD) with $\inf \bar\rho > 0$. Assume further that $\lambda < 2d\sigma$. Then there exists a constant $C > 0$ and an exponent $\theta > 0$, independent of N, such that for any fixed k,

$$\|\rho_{N,k} - \bar\rho^{\otimes k}\|_{L^\infty(0,T;L^1(\Pi^{kd}))} \leq \frac{Ck^{1/2}}{N^\theta}.$$

The optimal constant 4σ in 2D corresponds to the critical mass $8\pi\sigma$ for which we have blow-up in finite time for PKS. (Blanchet,Dolbeault and Perthame ('06)).

Comments

- The modulated free energy E_N "effectively" control the distance between ρ_N and $\bar{\rho}_N$. **Goal:**
$$E_N(\rho_N|\bar{\rho}_N) \geq \frac{1}{C}\mathcal{H}_N(\rho_N|\bar{\rho}_N) - \frac{C}{N^\theta}.$$
- Control $\frac{\mathrm{d}}{\mathrm{d}t}E_N$ above by E_N or $\mathcal{H}_N$ or $\mathcal{K}_N + C/N^\theta$. The PKS case is okay by the previous large deviation estimates, since now for $|\nabla V(x)| \leq C/|x|$.
- For more singular kernels, we need to establish that
$$-\frac{1}{2\sigma}\int_{\Pi^{dN}} \mathrm{d}\rho_N \int_{x\neq y} \nabla V(x-y)(\psi(x)-\psi(y))(\mathrm{d}\mu_N - \mathrm{d}\bar{\rho})^{\otimes 2}$$
$$\leq C\mathcal{K}_N(\rho_N|\bar{\rho}_N) + C\mathcal{H}_N(\rho_N|\bar{\rho}_N) + \varepsilon(N),$$

  where $\varepsilon(N) \to 0$ as $N \to \infty$.

General Results

Theorem (Bresch, Jabin & W. ('19))

We establish the Mean Field limit from (IPS) towards (MFD) for the following cases:

- *The case* $\sigma_N = \sigma > 0$. *Let* $K = -\nabla V$, *where* V *is an even potential with* $V = V_a + V_r$ *and*

$$V_a,\ V_r \in L^p(\Pi^d) \cap C^2(\Pi^d \setminus \{0\}), \quad \text{for } p > 1;$$

$$V_a(x) \geq \gamma \log|x| + C, \quad \text{for } 0 \leq \gamma < 2d\sigma, \quad |\nabla V_a(x)| \leq \frac{C}{|x|};$$

$$V_r \geq 0,\ |\nabla V_r(x)| \leq \frac{C}{|x|^k} \text{ for } k > 0, \quad |\nabla_\xi \hat{V}_r(\xi)| \leq C\left(\frac{\hat{V}_r(\xi)}{1+|\xi|} + \frac{1}{1+|\xi|^{d+1}}\right).$$

- *The case* $\sigma_N \to 0$. *Just choose* $K = -\nabla V_r$, *where* V_r *is a repulsive potential specified above.*

We can cover the Riesz potentials with possible perturbations.

Further Discussion

1. What's next for our program?
2. General large N limit problems.
3. Distribution sampling algorithm based on interacting particle system. As a new solver for PDEs? As a new framework for learning theory?

References

P.-E. Jabin and Z. Wang, Mean field limit and propagation of chaos for Vlasov Systems with bounded forces. **J. Funct. Anal. 271** (2016) 3588–3627.

P.-E. Jabin and Z. Wang, Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. **Invent. Math. 214** (2018) 523–591.

D. Bresch, P.E. Jabin and Z. Wang, On mean field limit and quantitative estimates with a large class of singular kernels: Application to the Patlak-Keller-Segel Model. **Comptes Rendus Mathematique 357** (2019) 708–720.

D. Bresch, P.E. Jabin and Z. Wang, Modulated free energy and mean field limit. **Séminaire Laurent Schwartz — EDP et applications** (2019) Talk no. 2, 22 p.

D. Bresch, P.E. Jabin and Z. Wang, Mean field limit and quantitative estimate with a large class of singular kernels. In preparation.

H. Hassani, Z. Shen and Z. Wang, Sinkhorn Barycenter via Functional Gradient Descent. Submitted to **ICML 2020**.

Thank you!