Entropy in PDE analysis

Clément Mouhot, University of Cambridge

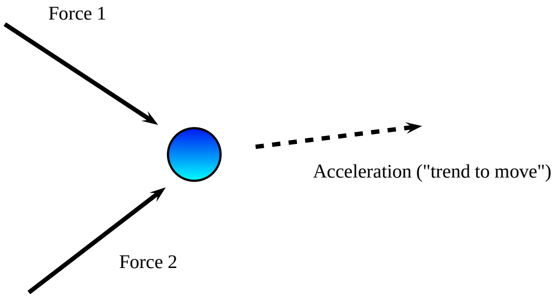
Clay Research Conference and workshop
"About Entropy in Large Classical Particle Systems"
Oxford, September 25th to 29th, 2023

Why is entropy so important in PDE analysis?

- Entropy functionals play a fundamental role in the study of many evolution systems, in particular PDEs (focus of this talk)
- It appears in the Cauchy theory of fluid equations, kinetic equations, (reaction-)diffusion PDEs, Ricci flow...
- It appears in the large dimension (many-body) limit of many systems (Newton system with Coulomb or Newton interaction, hard spheres system, vortex system...)
- It appears in many scaling limits (hydrodynamical limit of kinetic equations, hydrodynamical limit of interacting particle systems, long wave scaling of hyperbolic PDEs...)
- It even appears in disguise in the elliptic/parabolic regularity theory of De Giorgi-Nash-Moser
- To understand why it is so central in PDE analysis we have to go back to the origin of the concept of entropy

Origin of many PDEs (in the classical setting)

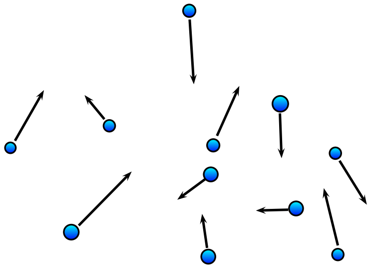
Classical mechanics rests on the fundamental laws of dynamics



Isaac Newton (2d law): *Sum of all forces applied to a body is proportional to its mass and to its acceleration*

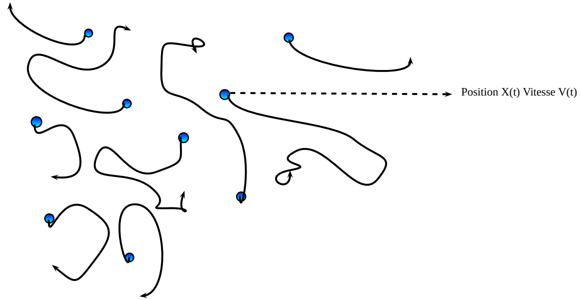
Particle systems

Many bodies (atoms, electrons, grains, stars. . .)
Fluid: $N \sim 10^{24}$ Avogadro number
Plasma of the solar kernel: $N \sim 10^{32}$
Galaxies: $N \sim 10^{11}$ stars in the Milkyway
Interactions: collision, electro-magnetism, gravitation. . .



Microscopic description (Hamiltonian reversible dynamics)

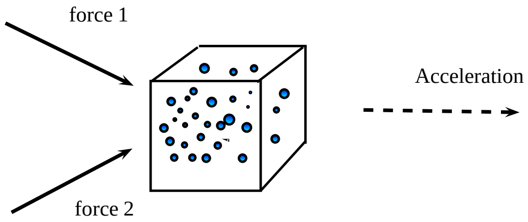
Microscopic description: Fundamental law of dynamics to each particle $\rightarrow$ **trajectories of all particles**.



Resists to analysis and theoretical predictions as soon as $N \geq 3$, and extremely hard to compute for large N

Macroscopic description (the observable world)

Macroscopic description: Fundamental laws of dynamics on infinitesimal volume elements of a continuum (fluid)



$\rightarrow$ **hydrodynamical (partial differential) equations**:
Leonhard Euler (1707–1783) in 1755 for non-viscous fluids
Claude-Louis Navier (1785–1836) in 1821 and George Stokes (1819–1903) in 1845 for viscous fluids

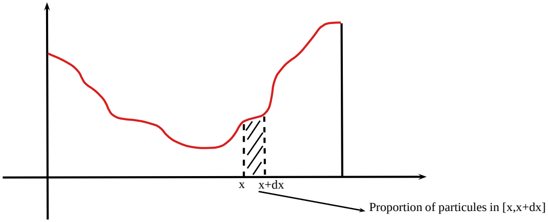
Mesoscopic description (keeping track of velocity statistics)

Between the *microscopic* and *macroscopic* levels of description

→ **Kinetic theory (mesoscopic level)**

Describe proportions of particles with given position and velocity

Important thing is not <u>which</u> particle but <u>how many of them</u>



$f(t, x, v)$ distribution of particles at x and with velocity v (cf. statistics of a population)

Natural intermediate step between microscopic & macroscopic descriptions in Hilbert's 6-th problem: "Axiomatize mechanics"

Thermodynamical entropy: Clausius

- Microscopic laws are time-reversible however the observable world is time-irreversible
- To account for the time-irreversible observations (and motivated by the inventions of the industrial revolution), Clausius pioneers the new field of thermodynamics
- He introduces the idea of entropy in 1865, with a formula for macroscopic systems in equilibrium

$$S_C = S_C(M_{eq}) \quad \text{where } M_{eq} \text{ some observables at equilibrium}$$

- His article introduces the two laws of thermodynamics:
  1. The energy of the world is constant.
  2. The entropy of the world tends to a maximum.

From Maxwell to Boltzmann

- Maxwell 1867 derived the distribution of a gas at equilibrium (Maxwellian=gaussian) and the form of the collision operator
- Boltzmann 1872: "*it has not yet been proved that, for any initial state of the gas, it must approach the limit distribution discovered by Maxwell*"
- To answer this question, Boltzmann derived the so-called Boltzmann equation (see below) on which he proves (formally) the H-theorem, that is the growth of the entropy
- But he also gives between 1872-1875 a microscopic interpretation of the entropy and of the macroscopic irreversibility, with a modern formulation by Planck in 1900
- Following Schrödinger 1956, Lebowitz 1993 and many other physicists, I consider it to be the correct explanation (rather than quantum microscopic phenomena for instance)

From Gibbs to Shannon

- The Boltzmann entropy writes

$$S_B(M) := k \log |W_M|$$

as we shall see, where $|W_M|$ is the volume of microstates W_M associated to a given macrostate M

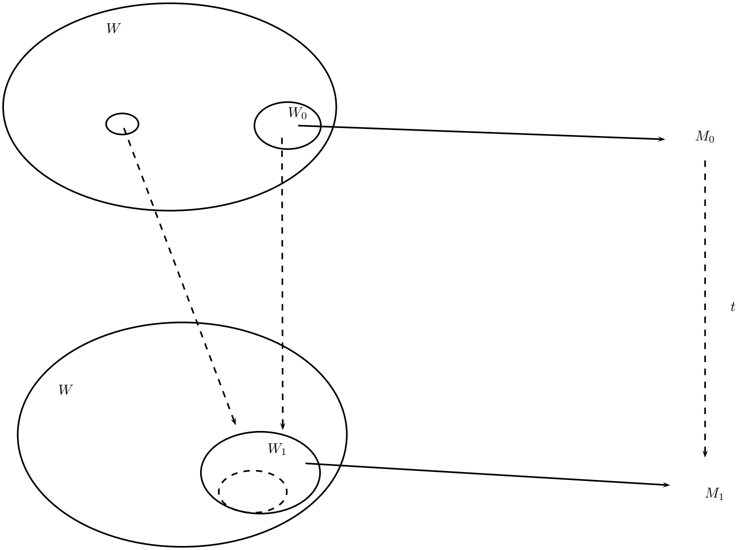
- Soon after, Gibbs 1876 proposed a related slightly different viewpoint, more statistical, and defines the statistical entropy associated to a probability density ρ on the microstates

$$S_G(\rho) := \int_{\mathcal{W}} \rho(X) \log \rho(X) \, \mathrm{d}X$$

- Shannon and Weaver pioneer in 1948 the theory of information and use the Boltzmann entropy, to which they give a related slightly different interpretation: this measures the uncertainty of a given observed macrostate

Irreversibility according to Boltzmann (Lebowitz 1993)

Representation in terms of a “factorization” of a dynamics



$$S = k \log W$$

Irreversibility according to Boltzmann (Lebowitz 1993)

- Phase space $\mathcal{W}$: all possible microscopic configurations (positions and velocities of all particles)
- Right hand side: macroscopic configurations (kinetic distribution, or in fluid mechanics: density, momentum, temperatures. . .)
- Microscopic evolution is **Hamiltonian** & preserves volume
$$\Rightarrow |W_1| \geq |W_0|$$
- Possible that several microscopic states evolves toward the same macroscopic state $\Rightarrow |W_1| > |W_0|$
- Boltzmann's entropy: $S(M) := k \log |W_M|$ non-decreasing
- The logarithm is introduced naturally for preserving additivity when considering several independent systems
- No contradiction with the reversibility of the microscopic dynamics (loss in factorized-hidden degrees of freedom)
- Note that the entropy is not an intrinsic property of the microscopic system but depends on the scale of observation

Irreversibility according to Boltzmann

- So in this theory irreversibility is a product of a (huge) separation of scales: it follows from
  1. Seeing through "macroscopic filtering-blurring glasses"
  2. The microscopic volume preservation (help not obstacle!)
  3. A microscopic evolution path that is highly atypical, going back to the universe having started with a low entropy
- To illustrate the insane separation of scales when a constraint is lifted and irreversible evolution follows: consider what happens when a wall dividing the box is removed; for 1 mole of fluid in a 1-liter container the volume ratio of the unconstrained region to the constrained one is of order $10^{10^{20}}$
- Subtle point: even on W_0 and W_1 the microstates that give the observed time-arrow pointing towards the future are atypical since one could reverse velocities without changing the fluid observables

Boltzmann entropy vs Clausius and Gibbs entropies

- Boltzmann entropy coincides with Clausius entropy for large number of particles N and when the thermodynamical global equilibrium is reached
- It also coincides with Gibbs entropy at a given time (up to a constant) by considering the probability density

$$\rho_W(X) := \frac{1}{|W|}\mathbf{1}_{X \in W}$$

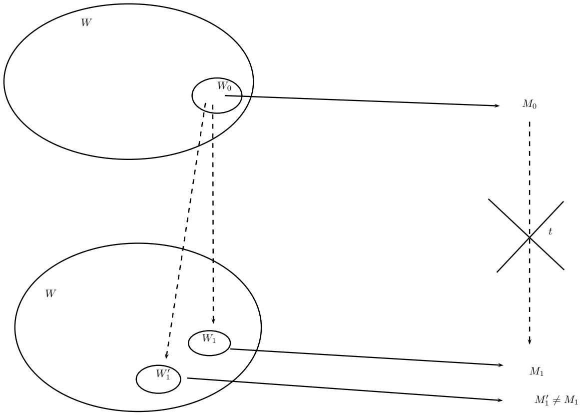
  since then

$$S_G(\rho_W) := \int_{\mathcal{W}} \rho_W(X) \log \rho_W(X) \, \mathrm{d}X = \log |W|$$

- However for a given probability density ρ, the Gibbs statistical entropy is constant due to the volume-preserving evolution
- The specific contribution of Boltzmann's discovery is in understanding the time evolution of the entropy, which is why it has had such indirect impact in PDEs

Irreversibility according to Boltzmann: the factorization

Implicit assumption: $X \in W(M) \Rightarrow T_t(X) \in W(F_t(M))$



$S = k \log W$

Necessary for "closure" of macroscopic evolution laws/equations (well-posedness in the sense of Hadamard of the Cauchy problem, i.e. macroscopic causality)

Irreversibility: proving the factorization

- Boltzmann's idea of molecular chaos ("Stosszahlansatz")
- Roughly speaking: for certain initial data with low correlations, the low correlations are preserved with time and the Poincaré recurrence time is "sent to ∞" as $N \to +\infty$
- In fact, even more subtle since time can go forward or backward, and a time arrow is also selected, which corresponds to the special initial data considered, and the specific type of molecular chaos propagated (post or pre-collisional)
- Note that "chaos" here (and irreversiblity) has nothing to do with the chaotic behaviour in dynamical systems, which happens already for small number of degrees of freedom
- Note also that independently of whether the factorization can be proven, the observed correctness and well-posedness of macroscopic evolution laws is enough to infer the existence of a non-decreasing entropy functional

Hilbert's (1862–1943) 6-th problem (ICM Paris 1900)

The investigations on the foundations of geometry suggest the problem: To treat in the same manner, by means of axioms, those physical sciences in which mathematics plays an important part; in the first rank are the theory of probabilities and mechanics. [...]

It is therefore very desirable that the discussion of the foundations of mechanics be taken up by mathematicians also. Thus Boltzmann's work on the principles of mechanics suggests the problem of developing mathematically the limiting processes, there merely indicated, which lead from the atomistic view to the laws of motion of continua.

This question posed by Hilbert emphasizes the hierarchical scaling structure of the so-called fundamental PDEs: they must be derived from first principle, i.e. microscopic dynamics. This explains why the idea of Boltzmann can be applied.

From the concept of entropy to a priori estimates

- Consider a microscopic space of N particles that can distribute themselves along k boxes, with a macroscopic space recording only the proportions $f_i := N_i/N$ of particles in each state
- We assume that all possible microstates have the same measure (uniform counting measure)
- Given a macrostate $M := (f_1, \ldots, f_k) = (N_1/N, \ldots, N_k/N)$, the number of associated microstates is

$$|W_M| = \frac{N!}{N_1! \cdots N_k!}$$

- The Stirling formula yields (for $n \to \infty$)

$$\log n! = n \log n - n + \log \sqrt{2\pi n} + o(1)$$

- We compute

$$\frac{1}{N} \log |W_M| = \frac{1}{N} \left(\log N! - \sum_{i=1}^{k} \log N_i! \right)$$

The continuous limit

- We use Stirling formula to get

$$\frac{1}{N}\log|W_M| = \frac{1}{N}\Bigg(N\log N - N + \log\sqrt{2\pi N} + o(1) \\ - \sum_{i=1}^{k}\left[N_i\log N_i - N_i + \log\sqrt{2\pi N_i} + o(1)\right]\Bigg) \\ = -\sum_{i=1}^{k}\frac{N_i}{N}\log\frac{N_i}{N} + O\left(\frac{k\log N}{N}\right)$$

- Hence for k large but much smaller than N we have a Riemann sum and

$$\frac{1}{N}\log|W_M| \sim -\int f\log f\,\mathrm{d}\nu$$

with respect to the reference measure ν

- This is the Boltzmann relative entropy $H(\mu|\nu)$ of μ with respect to ν with $f := \frac{\mathrm{d}\mu}{\mathrm{d}\nu}$

The continuous limit: large deviations

- In fact another way to recover the Boltzmann entropy is through (here the micro-space is $\mathcal{W} = \mathcal{Y}^{\otimes N}$)

$$H(\mu|\nu) = -\lim_{k\to\infty}\lim_{\varepsilon\to 0}\lim_{N\to\infty}\frac{1}{N}\log\nu^{\otimes N}\Bigg\{\cdots$$

$$(x_1,\dots,x_N)\in\mathcal{Y}^{\otimes N} \quad \text{s.t.} \quad \sup_{j=1}^{k}\left|\int_{\mathcal{Y}}\varphi_j\,\mathrm{d}\mu - \frac{1}{N}\sum_{i=1}^{k}\varphi_i(x_i)\right| \le \varepsilon\Bigg\}$$

- This is Sanov theorem: it estimates the large deviations of the empirical measure $\mu_X^N = \frac{1}{N}\sum\delta_{X_i}$ around its limit ν (fundamental law of statistics, i.e. law of large numbers for empirical measures)
- The general principe relevant to PDE: a large deviation functional is a Lyapunov functional for limit factorized dynamics as $N\to\infty$, e.g. De Roeck-Maes-Netocny 2006

The (Maxwell-)Boltzmann equation (1867, 1872)

Let us now look at concrete a priori estimates, and see how entropy provides particularly useful topology and nonlinear functionals to study the fundamental PDEs:

$$\underbrace{\partial_t f}_{\text{time change}} + \underbrace{v \cdot \nabla_x f}_{\text{space change}} = \underbrace{Q(f,f)}_{\text{collision operator}} \qquad \text{on } f(t,x,v) \geq 0$$

- Partial differential equation: relates infinitesimal changes of several variables
- Transport term $v \cdot \nabla_x$: straight line along velocity v
- Collision operator $Q(f,f)$: bilinear, acting on v only, integral

$$Q(f,f)(v) = \int_{v_*} \int_{collisions} \Big[\underbrace{f(v')f(v'_*)}_{(v',v'_*)\to(v,v_*)} - \underbrace{f(v)f(v_*)}_{(v,v_*)\to\dots} \Big] B$$

Structure of the Boltzmann equation (I)

- $Q(f,f)$ bilinear integral operator acting on v only (so it is local in t and x), representing interactions between particles:

$$Q(f,f)(v) := \int_{v_* \in \mathbb{R}^d} \int_{\omega \in \mathbb{S}^{d-1}} [\underbrace{f(v'_*)f(v')}_{\text{"appearing"}} - \underbrace{f(v)f(v_*)}_{\text{"disappearing"}}] \underbrace{B(v - v_*, \omega)}_{\text{collision kernel } (\geq 0)}$$

- Velocity collision rule ($(d-1)$ free parameters $\to \omega$):

$$v' := v - (v - v_*, \omega)\omega, \qquad v'_* := v_* + (v - v_*, \omega)\omega$$

- One has (microscopic conservation laws)

$$v' + v'_* = v + v_*, \qquad |v'_*|^2 + |v'|^2 = |v|^2 + |v_*|^2$$

- Given $\omega \in \mathbb{S}^{d-1}$, $(v, v_*) \mapsto (v', v'_*)$ has Jacobian det. -1, and $(v, v_*) \to (v_*, v)$ has Jacobian det. 1

Structure of the Boltzmann equation (II)

- We deduce for a test function $\varphi(v)$

$$\int_{\mathbb{R}^d} Q(f,f)\varphi(v)\,\mathrm{d}v$$
$$= \frac{1}{4}\int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} [f'f'_* - ff_*] B(v-v_*,\omega)(\varphi+\varphi_*-\varphi'-\varphi'_*)\,\mathrm{d}\omega\,\mathrm{d}v_*\,\mathrm{d}v$$

- Choosing $\varphi = 1, v, |v|^2$ we deduce

$$\int_{\mathbb{R}^d} Q(f,f)\begin{pmatrix}1\\ v\\ |v|^2\end{pmatrix}\mathrm{d}v = 0$$

- This implies formally (no boundary for simplification)

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^{2d}} f\begin{pmatrix}1\\ v\\ |v|^2\end{pmatrix}\mathrm{d}x\,\mathrm{d}v = 0$$

Structure of the Boltzmann equation (III)

- Choosing $\varphi = \log f$ we obtain the *H*-theorem

$$\frac{\mathrm{d}}{\mathrm{d}t} H(f) = \frac{\mathrm{d}}{\mathrm{d}t} \int_{x,v} f \log f = -D(f) \leq 0$$

(*Boltzmann wrote "E-theorem" but later Burbury and Gibbs used "H" most likely in reference to the Greek capital "Eta"*)

- The entropy production functional is

$$D(f) = -\int_{x,v} Q(f,f) \log f = \int_{x,v,v_*,\omega} [f' f'_* - f f_*] \log \frac{f' f'_*}{f f_*} B(v - v_*, \omega) \geq 0$$

with cancellation only when $f f_* = f' f'_*$ everywhere:

$$M_f = M[\rho, u, T] = \frac{\rho}{(2\pi T)^{\frac{d}{2}}} e^{-\frac{|v-u|^2}{2T}}$$

- Time-irreversible equation and mathematical basis for studying relaxation to equilibrium (2-d law of thermodynamic)

What is the non-factorized microscopic dynamics?

- Binary interactions through a potential ψ depending only on the distance between two interacting bodies
- External forces with some potential $\phi(t, \text{position})$.
- Hamilton equations (Newton laws)

$$1 \leq i \leq N, \qquad \frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\partial H^N}{\partial v_i} \qquad \frac{\mathrm{d}v_i}{\mathrm{d}t} = -\frac{\partial H^N}{\partial x_i}$$

$$H^N = \underbrace{\sum_{i=1}^{N} \frac{v_i^2}{2}}_{\text{kinetic energy}} + \underbrace{\sum_{i<j} \psi(x_i - x_j)}_{\text{interaction energy}} + \underbrace{\sum_{i=1}^{N} \phi(t, x_i)}_{\text{potential energy}}.$$

- This corresponds to the following ODE's

$$1 \leq i \leq N, \quad \dot{x}_i = v_i, \quad \dot{v}_i = -\sum_{i \neq j} \nabla_x \psi(x_i - x_j) - \nabla_x \phi(x_i).$$

The N-body Liouville equation (I)

Statistical solution to the previous ODEs (continuous superposition of trajectories):

$$\frac{\partial F^N}{\partial t} + \sum_{i=1}^{N} \left(\frac{\partial H^N}{\partial v_i} \cdot \frac{\partial F^N}{\partial x_i} - \frac{\partial H^N}{\partial x_i} \cdot \frac{\partial F^N}{\partial v_i} \right) = 0$$

on the joint microscopic probability distribution function F^N

Liouville theorem

For any $t \in \mathbb{R}$ one has $F^N(t, S_t(X,V)) = F^N(0,X,V)$, where S_t is the flow of the Hamilton equations, and S_t preserves volume.

Consequence: statistical Casimir invariants (for $\Theta : \mathbb{R} \mapsto \mathbb{R}_+$)

$$\int_{\mathbb{R}^{2dN}} \Theta\left(F^N(t,X,V)\right) \mathrm{d}X\,\mathrm{d}V = \int_{\mathbb{R}^{2dN}} \Theta\left(F^N(0,X,V)\right) \mathrm{d}X\,\mathrm{d}V$$

including Boltzmann entropy for $\Theta(r) = r \log r$

The N-body Liouville equation (II)

Proof: Differentiate in time $J(t, X, V) := \det \nabla_{X,V} S_t(X, V)$:

$$\frac{\mathrm{d}}{\mathrm{d}t} J(t, X, V) = \left[\sum_i \left(\frac{\partial^2 H^N}{\partial x_i \partial v_i} - \frac{\partial^2 H^N}{\partial v_i \partial x_i} \right) \right] J(t, X, V) = 0$$

Together with $J(0, X, V) = \det \mathrm{Id} = 1$, it yields $J(t, X, V) \equiv 1$
One deduces by change of variable

$$\int_{\mathbb{R}^{2dN}} \Theta \left(F^N(t, X, V) \right) \mathrm{d}X \, \mathrm{d}V = \int_{\mathbb{R}^{2dN}} \Theta \left(F^N(0, X, V) \right) \mathrm{d}X \, \mathrm{d}V$$

This reflects the time-reversibility of the Liouville equation:
invariance under the change of variable $(t, X, V) \mapsto (-t, X, -V)$
Cf. reversibility of Newton laws at microscopic level

Factorization of the dynamics: the BBGKY hierarchy (I)

- N-particle Liouville equation allows for considering superpositions of all trajectories at the same time, still contains same amount of information as the Newton equations
- Simplify description of the system by throwing away information: (Hopefully) the system is described by a one particle distribution (first marginal):

$$f_1^N(t,x,v) := \int_{\mathbb{R}^{2d(N-1)}} F^N(t,X,V)\,\mathrm{d}x_2\,\mathrm{d}x_3\ldots\,\mathrm{d}x_N\,\mathrm{d}v_2\ldots\,\mathrm{d}v_N$$

  (Observe that it still depends on N!)
- Why the marginal according to the first variable? Consider F^N symmetric (invariant under permutations) by indistinguability of the particles

Factorization of the dynamics: the BBGKY hierarchy (II)

- How can we obtain an equation for how f_1^N evolves?
- Integrate the N-body Liouville equation

$$\frac{\partial F^N}{\partial t} + \sum_{i=1}^{N} \left(\frac{\partial H^N}{\partial v_i} \cdot \frac{\partial F^N}{\partial x_i} - \frac{\partial H^N}{\partial x_i} \cdot \frac{\partial F^N}{\partial v_i} \right) = 0$$

according to $X^{(1)} := (x_2, x_3, \dots, x_N),\ V^{(1)} = (v_2, v_3, \dots, v_N)$

- We obtain the following equation for the one marginal distribution

$$\frac{\partial f_1^N}{\partial t} + v \frac{\partial f_1^N}{\partial x} - \frac{\partial \phi}{\partial x} \frac{\partial f_1^N}{\partial v} - (N-1) \int_{\mathbb{R}^2} \frac{\partial}{\partial x}(\psi(x-y)) \frac{\partial f_2^N}{\partial v}(x, y, v, w) \, \mathrm{d}y \, \mathrm{d}w = 0$$

with the substitutions $x_1 \to x, x_2 \to y, v_1 \to v, v_2 \to w$

Factorization of the dynamics: the BBGKY hierarchy (III)

- How can we interpret this equation?
- Binary collisions $\Rightarrow$ evolution of first marginal (f_1^N) depends on second marginal f_2^N: interactions = correlations!
- Similarly f_2^N's evolution depends on f_3^N and so on:

$$\begin{aligned} \frac{\partial f_1^N}{\partial t} &= \mathcal{L}_1(f_1^N) + \mathcal{B}_1(f_2^N) \\ \ldots \frac{\partial f_k^N}{\partial t} &= \mathcal{L}_k(f_k^N) + \mathcal{B}_k(f_{k+1}^N) \\ \ldots \frac{\partial f_N}{\partial t} &= \frac{\partial F^N}{\partial t} = \left\{ H^N, F^N \right\} \end{aligned}$$

- This is the BBGKY hierarchy (Bogoliubov, Born, Green, Kirkwood, Yvon) or "Bogoliubov approach" for

$$f_1^N, f_3^N, \ldots, f_k^N, \ldots, f_N^N = F_N$$

The Many-particle or "Thermodynamic" Limit

- Goal of thermodynamical limit: perform $N \to \infty$ and recover closed equations on reduced distributions, ideally on the first marginal $f_1^N \sim f_1$ as $N \sim \infty$
- In view of the equation on f_1^N and assuming low correlations it is natural to ask whether

$$f_2^N = f_1^N \otimes f_1^N := f_1^N(t,x,v) f_1^N(t,y,w)?$$

- However the probability independence assumption is always false for interacting particle systems, due to interactions!
- Idea of Boltzmann (formulated mathematically by Kac 1956) is that one can hope in the limit as $N \to \infty$ (similar to propagating fundamental of statistics)

$$f_2^N \sim f_1^N \otimes f_1^N \quad \text{as } N \to +\infty \qquad (\text{"near-product structure"})$$

Mathematical formulation of molecular chaos of Boltzmann (in fact more subtle with pre / post-collisional chaos)

Weak coupling / mean-field / Vlasov limit (I)

- Discovery Jeans 1915 (galaxies) and Vlasov 1938 (plasmas)
- Describe binary interactions through their collective effect
- Adapted to long-range interactions: Coulomb or Newton fields, but not hard spheres!
- Mathematically, we let $\psi_N(z) = \bar{\psi}(z)/N$ and $r = r(N) \to 0$ such that $\frac{Nr^3}{\mathcal{V}} \ll 1$ (dilute gas)
- Force between two particles is $O(1/N)$, hence the action of one particle becomes negligible in the limit
- However a given particle feels the interaction of $N-1$ other particles, hence it feels a force of $O(\frac{N-1}{N}) = O(1)$.
- This "mean-field approach" has found much wider application in many areas (biology, sociology, etc.)

Weak coupling / mean-field / Vlasov limit (II)

- Vlasov equation for $f = \lim f_1^N$ as $N \to \infty$

$$\frac{\partial f}{\partial t}+v\frac{\partial f}{\partial x}-\frac{\partial \phi}{\partial x}\frac{\partial f}{\partial v}-\underbrace{\int \frac{\partial}{\partial x}(\psi(x-y))\frac{\partial f}{\partial v}(x,v)f(y,w)dydw}_{(\nabla_x\psi*\rho_f)\cdot\nabla_v f}=0$$

  obtained from equation on f_1^N as $N \to \infty$
- Vlasov-Poisson equation when ψ Coulomb or Newton potential: $\Delta\Psi_f = \pm(\rho_f - \rho^0)$ with $\psi_f = \psi * \rho_f$
- Proof of the mean-field limit known when ψ regular: Braun-Hepp-Dobrushin 1970s, 1980s
- For Coulomb-Newton interaction potential major open problem, best result so far Hauray-Jabin, Pickl & collaborators, Jabin-Wang, etc.
- Related problem slightly less difficult (first order dynamics): convergence of the vortex system to 2D incompressible Euler in vorticity formulation: Marchioro-Pulvirenti, Serfaty & co...

Structure of the mean-field Vlasov equation

- Still time-reversible: if $f = f(t, x, v)$ solution, then $g(t, x, v) := f(-t, x, -v)$ also solution
- (Mean-field) Hamiltonian structure

$$\frac{\partial f}{\partial t} + \left(\frac{\partial E_f}{\partial v} \cdot \frac{\partial f}{\partial x} - \frac{\partial E_f}{\partial x} \cdot \frac{\partial f}{\partial v} \right) = \partial_t f + \{E_f, f\} = 0$$

  with the microscopic *mean-field* Hamiltonian function

$$E_f(t, x, v) := \frac{|v|^2}{2} + \Psi_f(t, x).$$

- Energy conservation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(f) = \frac{\mathrm{d}}{\mathrm{d}t} \left(\int_{\mathbb{R}^{2d}} f \frac{|v|^2}{2} \,\mathrm{d}x\,\mathrm{d}v \pm \int_{\mathbb{R}^d} \frac{|\nabla_x \Psi_f|^2}{2} \,\mathrm{d}x \right) = 0$$

- Conservation of entropy and of all "Casimir functionals"

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^{2d}} G(f_t) \,\mathrm{d}x\,\mathrm{d}v = 0$$

Boltzmann-Grad / collisional limit (I)

- Short-range interactions and no external force or boundary
- Assume $N \to \infty$ (infinite number of particles)
- Finite volume $\mathcal{V}$ and radius satisfies $r(N) \to 0$
- Each particle performs $O(1)$ collisions per unit of time

$$Nr(N)^2 = O(1)$$

- Mean-free path $\ell(N) = \mathcal{V}/(Nr(N)^2) = O(1) \gg r(N)$
- Mass $m(N) \to 0$ with $Nm(N) = O(1)$: average density $O(1)$
- Assume some form of molecular chaos $f_2^N \sim f_1^N \otimes f_1^N$: however here this cannot be true before and after collision, and this notion refines into pre-collisional and post-collisional chaos

Boltzmann-Grad / collisional limit (II)

- More difficult than mean-field limit, open beyond small time
- Start from the N-body Liouville equation

$$\partial_t F^N + V \cdot \nabla_X F^N = 0$$

on the domain

$$\Omega_N := \{\forall i \neq j, \quad |x_i - x_j| \geq 2r(N)\}$$

- Corresponds to a "wall" potential: mathematically involves boundary integral terms (additional difficulty)
- We then consider again the one-particle distribution

$$f_1^N(t, x_1, v_1) := \int_{\mathbb{R}^{2d(N-1)}} F^N(t, X, V) \, \mathrm{d}x_2 \, \mathrm{d}x_3, \dots \, \mathrm{d}x_N \, \mathrm{d}v_2 \dots \, \mathrm{d}v_N$$

Boltzmann-Grad / collisional limit (III)

- Search for an evolution equation on it by integrating

$$\partial_t f_1^N + v_1 \cdot \nabla_{x_1} f_1^N = -\sum_{j=2}^{N} \int_{X^{(1)}, V^{(1)} \in \Omega_N^{(1)}} v_j \cdot \nabla_{x_j} F^N$$

$$= (N-1) r(N)^2 \Big[\int_+ f_2^N(x_1, x_2, v_1, v_2) |(v_1 - v_2) \cdot \omega_{12}| \, \mathrm{d}\sigma_{12} \, \mathrm{d}v_2$$

$$- \int_- f_2^N(x_1, x_2, v_1, v_2) |(v_1 - v_2) \cdot \omega_{12}| \, \mathrm{d}\sigma_{12} \Big] + \text{cancelling or negligeable terms}$$

(factor $r(N)^2$ comes from surface element of the sphere)

- ω_{12} outer normal to the sphere $|x_1 - x_2| = 2r(N)$
- $\mathrm{d}\sigma_{12}$ surface element on the same sphere
- $\int_+$ surface term for outgoing collisions $(v_1 - v_2) \cdot \omega_{12} \geq 0$
- $\int_-$ surface term for ingoing collisions $(v_1 - v_2) \cdot \omega_{12} \leq 0$

Boltzmann-Grad / collisional limit (IV)

- Multiple collisions (more than binary) negligeable in the limit
- Cancellation of surface terms not involving x_1 (micro-reversibility)
- Express outgoing velocities in $\int_+$ in terms of the ingoing velocities in $\int_-$ (time arrow)
- Choice arbitrary at microscopic level but cannot be reversed after the limit $N \to +\infty$ has been taken
- For similar reasons in the limit binary collision trajectories no more well-defined between point particles: statistical outcomes
- Other choice (expressing pre-collisional velocities in terms of post-collisional ones) would lead to a backward Boltzmann equation, with a minus in front of the collision operator

Boltzmann-Grad / collisional limit (V)

- Using previous assumptions go back to

$$\partial_t f_1^N + v_1 \cdot \nabla_{x_1} f_1^N$$
$$= (N-1) r(N)^2 \left[\int_+ f_2^N(x_1, x_2, v_1, v_2) |(v_1 - v_2) \cdot \omega_{12}| \, \mathrm{d}\sigma_{12} \, \mathrm{d}v_2 \right.$$
$$\left. - \int_- f_2^N(x_1, x_2, v_1, v_2) |(v_1 - v_2) \cdot \omega_{12}| \, \mathrm{d}\sigma_{12} \right] \; + \; o(N^{-1})$$

- Use in $\int_+$ with $\omega_{12} = (x_1 - x_2)/(2r(N))$

$$v_1^+ := v_1 - \omega_{12} \left(\omega_{12} \cdot (v_1 - v_2) \right), \quad v_2^+ := v_1 + \omega_{12} \left(\omega_{12} \cdot (v_1 - v_2) \right)$$

- If propagation of pre-collisional chaos holds $f_2^N \sim f_1^N \otimes f_1^N$ and scaling $N r(N)^2 = O(1)$: hard spheres Boltzmann eq.

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = \int_{\mathbb{S}^{d-1} \times \mathbb{R}^d} \left(f(v') f(v'_*) - f(v) f(v_*) \right) |(v - v_*) \cdot \omega| \, \mathrm{d}v_* \, \mathrm{d}\omega$$

Proving the Boltzmann-Grad limit (I)

Cercignani 1972: The apparently paradoxical connection between the reversible nature of the basic equations of classical mechanics and the irreversible features of the gross description of large systems of classical particles satisfying those equations, came under strong focus with the celebrated H-theorem of Boltzmann and the related controversies between Boltzmann on one side and Loschmidt and Zermelo on the other. [. . .]
In particular, it is not clear whether an averaging is taking place during the duration and over the region of a molecular collision. This averaging is related to another controversial point, i.e., **whether irreversibility can appear only through the intervention of a stochastic or random model or can be a consequence of the progressive weakening of the property of continuous dependence on initial conditions**.

Open difficult problem. However, as I have tried to argue: for the purpose of PDE analysis, it is enough to expect a closed factorized dynamics for infering the existence of the entropy.

The hydrodynamic limit from kinetic theory

- We now continue climbing the hierarchy of scales
- Small Knudsen number limit of the Boltzmann equation

$$\partial_t f^\varepsilon + v \cdot \nabla_x f^\varepsilon = \frac{1}{\varepsilon} Q(f^\varepsilon, f^\varepsilon)$$

- Hilbert or Chapman-Enskog expansions:

$$f^\varepsilon(t, x, v) \sim_{\varepsilon \to 0} M[\rho_t, u_t, T_t]$$

  where the density, momentum and temperature fields (ρ_t, u_t, T_t) satisfy the compressible Euler equations
- First-order correction in ε is the Navier-Stokes viscosity
- Hence for such systems of conservation laws we indeed expect entropy conditions compatible with a kinetic limit or simply a vanishing viscosity limit (first-order correction)
- Justification of entropy conditions on Euler from kinetic entropy inequality only partially done for the Riemann shock in 1d to my knowledge (Liu-Yu, Cuesta-Hittmeir-Schmeiser...)

The incompressible limit

- The incompressible limit consists in studying small fluctuations in the small Knudsen number regime
- Yields following eq on the fluctuation $h_t^\varepsilon = \varepsilon^{-1}(f_t - M_{eq})$:

$$\partial_t h^\varepsilon + \frac{1}{\varepsilon} v \cdot \nabla_x h^\varepsilon = \frac{1}{\varepsilon^2} L h^\varepsilon + \frac{1}{\varepsilon} Q(h^\varepsilon, h^\varepsilon)$$

  where L is the linearized collision operator
- The fluctuation $h_t^\varepsilon \in L^2(M^{-1})$ and we expect

$$h_t^\varepsilon \sim_{\varepsilon \to 0} \left[\rho_t + u_t \cdot v + T_t \left(\frac{|v|^2 - d}{2} \right) \right] M_{eq}$$

  where (ρ_t, u_t, T_t) satisfy the incompressible Navier-Stokes eqs

$$\rho + T = 0, \quad \partial_t u + u \cdot \nabla_x u + \nabla_x p = \nu \Delta_x u,$$
$$\nabla_x \cdot u = 0, \quad \partial_t T + u \cdot \nabla_x T = \kappa \Delta_x T$$

Linearization of the entropy

- As observed by Hilbert, if $f = M + \varepsilon h$ then

$$\int f \log \frac{f}{M} = \int (M + \varepsilon h) \log \left(1 + \varepsilon \frac{h}{M} \right)$$

$$= \underbrace{\varepsilon \int h}_{=0 \text{ (mass cons.)}} + \frac{\varepsilon^2}{2} \underbrace{\int h^2 M^{-1}}_{\text{linearized entropy}} + O(\varepsilon^3)$$

- This weighted L^2 norm is the norm of symmetry where the linearized collision operator is self-adjoint
- It appears also in the study of Fokker-Planck (Ornstein-Uhlenbeck) operators, which can be seen as diffusive linearization of the Boltzmann operator
- But the original Boltzmann theory was the motivation of Hilbert when he initiated the theory of non-local operators
- Since the incompressible hydrodynamic limit is also a linearized fluctuation regime, L^2 norm naturally appears

Entropy and the Cauchy pb for incomp Navier-Stokes

- Consider the case ρ and T constant, then INS reduces to

$$\partial_t u + u \cdot \nabla_x u + \nabla_x p = \Delta_x u, \quad \nabla_x \cdot u = 0$$

  on $u : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ (normalizing the viscosity), with $d \geq 2$
- The seminal work of Leray 1934 initiated the method of a priori estimates in PDEs, and the notion of weak solutions
- The core of Leray's result is the entropy (aka energy) inequality

$$\frac{\mathrm{d}}{\mathrm{d}t} \int |u|^2 = - \int |\nabla_x u|^2 \leq 0$$

  which is the limit of the kinetic entropy inequality (see Bardos, Golse, Levermore, Masmoudi, Saint-Raymond)
- The corresponding a priori estimate $u \in L^\infty_t L^2_x \cap L^2_t H^1_x$ is sufficient to deduce strong L^2 compactness on an approximation sequence

Entropy and the Cauchy pb for the Boltzmann equation

- In spite of being closer to first principles the Boltzmann equation was introduced later than the Euler and Navier-Stokes equations, and consequently its mathematical study started later
- Equivalent of Leray theorem proved by DiPerna-Lions in 1989: weak (renormalised) solutions in $L^1_{x,v}(1+|v|^2) \cap L \log L$
- It uses the kinetic entropy inequality two times

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{x,v} f \log f = - \int_{x,v,v_*,\omega} [f' f'_* - f f_*] \log \frac{f' f'_*}{f f_*} B(v - v_*, \omega) \leq 0$$

- $L \log L \cap L^1(1+v^2) \Rightarrow$ strong L^1 compactness Dunford-Pettis
- It is also used to control the positive part of the collision operator in the renormalisation proceedure

$$\int \frac{1}{1+f} Q^+(f,f) \lesssim \underbrace{\int \frac{1}{1+f} Q^-(f,f)}_{\lesssim \int f(1+|v|^2)} + D(f)$$

Entropy and the Cauchy pb for conservation laws (I)

- The Cauchy theory of conservation laws (compressible fluid dynamics) is unfortunately mathematically still largely open
- We have however three notable settings with important results, and they all make use of some notion of entropy
- For systems of conservations laws in one dimension $x \in \mathbb{R}$

  $$\partial_t \vec{u} + A(u) \partial_x \vec{u} = 0$$

  Glimm 1965 proved the first existence theorem for initial data with small total variation
- The core idea is to decompose the solution into Riemann shocks and build a functional keeping track of the total variation when shocks interact; in this last estimate the entropy conditions are crucially used.
- This existence was later refined by Bressan into existence and uniqueness in the 1990s, and Bianchini-Bressan 2005 showed that the Glimm solutions are obtained as a vanishing viscosity limit (see also Chen-Perepelitsa 2010)

Entropy and the Cauchy pb for conservation laws (II)

- For scalar conservation laws in any dimension $x \in \mathbb{R}^d$

  $$\partial_t \rho + \nabla_x \cdot [\vec{F}(\rho)] = 0 \quad \text{where } \vec{F} \text{ is the flux function}$$

  the Kruzkhov theory 1970 exploits the total order of the real line and entropy conditions to prove by the doubling of variable argument a powerful L^1 stability estimate

  $$\int |u_t - v_t| \leq \int |u_0 - v_0|$$

  which implies global well-posedness and BV propagation
- Isolated incursion into systems in the large

  $$\begin{aligned} &\partial_t \rho + \partial_x(\rho u) = 0 \\ &\partial_t(\rho u) + \partial_x \left(\rho u^2 + p(\rho)\right) = 0 \\ &p(\rho) = \kappa \rho^\gamma \end{aligned}$$

  Isentropic 1d Euler (p-system)

Entropy and the Cauchy pb for conservation laws (III)

- DiPerna 1983 breakthrough: proved existence in the large by combining appropriate families of entropies and compensated compactness (inspired by works of Tartar) to show compactness of approximation sequence and prove existence of weak solutions
- Result of DiPerna was for $\gamma = (N-2)/N$, $N > 3$ integer, later extended by Chen 1990, Lions-Perthame-Tadmor 1994 and many others
- Note that an open question related to the entropy itself is to derive rigorously the entropy conditions used in the theory of conservation laws from the kinetic entropy inequality in the hyperbolic fluid limit

Entropy and Perelman's proof of the Poincaré conjecture

- The Poincaré conjecture is whether all closed smooth 3d simply connected Riemannian manifolds (M, g) are topologically equivalent to $\mathbb{S}^3$
- Richard Hamilton introduced the so-called Ricci flow $(M, g(t))$ in 1982 (quasilinear parabolic PDE)

$$\partial g_{\alpha\beta} = -2\text{Ric}_{\alpha\beta}$$

- Local existence was proved by Hamilton and improved by others, but singularities appear in finite time and the difficulty was to continue the flow beyond such singularities
- Perelman constructs a Ricci flow with surgery to go beyond the singularities, that extincts in finite time, and deduce from properties of the flow at late times the Poincaré conjecture
- Discovery of a new monotone quantity scale-invariant and coercive which gives estimates beyond singularity, i.e. an "entropy formula" (in the words of Perelman)

Entropy production to measure relaxation to equilibrium

- Immense field of study nowadays
- In the linearised setting the entropy is a weighted L^2 norm and the entropy production is the Dirichlet form; relating linearly both is a spectral gap (Poincaré ineq for diffusion operators)
- Gross, Bakry-Emery... initiated the proof of logarithmic Sobolev inequalities which are the nonlinear counterpart of Poincaré inequalities, taking advantage of the first and second time variations of the Ornstein-Uhlenbeck flow
- Combining degenerate entropy production with mixing conservative flow can result in fast relaxation rates: this is theory of hypocoercivity (Desvillettes-Villani)
- Other example of recent work I was involved in: Gualdani-Mischler-CM 2017 proof of H-theorem with exponential rate for Boltzmann hard spheres in a periodic box (a priori solutions)

Relative entropy as a nonlinear distance to limit regimes

- Due to its microscopic origin, the entropy is naturally well-behaved in large dimensions: hence it is a natural tool for proving mean-field limit, modulating it with an artefact microscopic solutions built from the target macroscopic solution, see Jabin-Wang on Vlasov mean-field limit, Serfaty & co on vortex mean-field limit...
- Due to its Lyapunov nature for the dynamics, it is naturally useful in scaling limits which mix fast and slow scales, by allowing to get at least partly rid of the fast scale error terms; hence it is used in the hydrodynamic limit from kinetic equations to fluid mechanics (see review of Saint-Raymond)
- Entropy and log-Sob inequality used to capture the fast local thermalisation in interacting particle systems on lattices (e.g. exclusion process, zero-range process, Ginzburg-Landau process...), see for instance Kipnis-Landim