ENTROPY AND $H$ THEOREM
THE MATHEMATICAL LEGACY
OF LUDWIG BOLTZMANN

Cambridge/Newton Institute, 15 November 2010

Cédric Villani
University of Lyon
& Institut Henri Poincaré (Paris)
Cutting-edge physics at the end of nineteenth century

Long-time behavior of a (dilute) classical gas

Take many (say $10^{20}$) small hard balls, bouncing against each other, in a box

Let the gas evolve according to Newton’s equations
Prediction by Maxwell and Boltzmann

The distribution function is asymptotically Gaussian

\[ f(t, x, v) \approx a \exp\left(-\frac{|v|^2}{2T}\right) \quad \text{as } t \to \infty \]
Based on four major conceptual advances 1865-1875

- Major modelling advance: Boltzmann equation
- Major mathematical advance: the statistical entropy
- Major physical advance: macroscopic irreversibility
- Major PDE advance: qualitative functional study

Let us review these advances

⇒ journey around centennial scientific problems
The Boltzmann equation

Models rarefied gases (Maxwell 1865, Boltzmann 1872)

\[ f(t, x, v) : \text{density of particles in } (x, v) \text{ space at time } t \]

\[ f(t, x, v) \, dx \, dv = \text{fraction of mass in } dx \, dv \]
The Boltzmann equation (without boundaries)

Unknown = time-dependent distribution \( f(t, x, v) \):

\[
\frac{\partial f}{\partial t} + \sum_{i=1}^{3} v_i \frac{\partial f}{\partial x_i} = Q(f, f) = \\
\int_{\mathbb{R}^3} \int_{S^2} B(v-v_*, \sigma) \left[ f(t, x, v')f(t, x, v_*') - f(t, x, v)f(t, x, v_*) \right] \, dv_* \, d\sigma
\]
The Boltzmann equation (without boundaries)

Unknown = time-dependent distribution $f(t, x, v)$:

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{3} v_i \frac{\partial f}{\partial x_i} = Q(f, f) =$$

$$\int_{\mathbb{R}^3_{v_*}} \int_{S^2} B(v-v_*, \sigma) \left[ f(t, x, v') f(t, x, v'_*) - f(t, x, v) f(t, x, v_*) \right] dv_* d\sigma$$
The Boltzmann equation (without boundaries)

Unknown = time-dependent distribution \( f(t, x, v) \):

\[
\frac{\partial f}{\partial t} + \sum_{i=1}^{3} v_i \frac{\partial f}{\partial x_i} = Q(f, f) = \\
\int_{\mathbb{R}^3_{v_*}} \int_{S^2} B(v-v_*, \sigma) \left[ f(t, x, v') f(t, x, v_*) - f(t, x, v) f(t, x, v_*) \right] dv_* d\sigma
\]
The Boltzmann equation (without boundaries)

Unknown = time-dependent distribution $f(t, x, v)$:

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{3} v_i \frac{\partial f}{\partial x_i} = Q(f, f) =$$

$$\int_{\mathbb{R}^3_v} \int_{S^2} B(v-v_*, \sigma) \left[ f(t, x, v') f(t, x, v_*') - f(t, x, v) f(t, x, v_*) \right] dv_* d\sigma$$

$$\begin{cases}
    v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2} \sigma \\
    v_*' = \frac{v + v_*}{2} - \frac{|v - v_*|}{2} \sigma
\end{cases}$$

$$(v, v_*) \leftrightarrow (v', v_*')$$
The Boltzmann equation (without boundaries)

Unknown = time-dependent distribution $f(t, x, v)$:

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{3} v_i \frac{\partial f}{\partial x_i} = Q(f, f) =$$

$$\int_{\mathbb{R}^3} \int_{S^2} B(v-v_*, \sigma) \left[ f(t, x, v') f(t, x, v_*) - f(t, x, v) f(t, x, v_*) \right] dv_* d\sigma$$
The Boltzmann equation (without boundaries)

Unknown = time-dependent distribution $f(t, x, v)$:

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{3} v_i \frac{\partial f}{\partial x_i} = Q(f, f) =$$

$$\int_{\mathbb{R}^3_{v_*}} \int_{S^2} B(v-v_*, \sigma) \left[ f(t, x, v') f(t, x, v_*) - f(t, x, v) f(t, x, v_*) \right] dv_* d\sigma$$
The Boltzmann equation (without boundaries)

Unknown = time-dependent distribution \( f(t, x, v) \):

\[
\frac{\partial f}{\partial t} + \sum_{i=1}^{3} v_i \frac{\partial f}{\partial x_i} = Q(f, f) = \\
\int_{\mathbb{R}^3} \int_{S^2} B(v-v_*, \sigma) \left[ f(t, x, v') f(t, x, v_*) - f(t, x, v) f(t, x, v_*) \right] dv_* d\sigma
\]

Conceptually tricky!

Molecular chaos: No correlations between particles

... 

However correlations form due to collisions!
The Boltzmann equation (without boundaries)

Unknown = time-dependent distribution $f(t, x, v)$:

$$
\frac{\partial f}{\partial t} + \sum_{i=1}^{3} v_i \frac{\partial f}{\partial x_i} = Q(f, f) =
$$

$$
\int_{\mathbb{R}^3_{v_*}} \int_{S^2} B(v-v_*, \sigma) \left[ f(t, x, v') f(t, x, v'_*) - f(t, x, v) f(t, x, v_*) \right] dv_* d\sigma
$$

Conceptually tricky!

**One-sided chaos:** No correlations on pre-collisional configurations, but correlations on post-collisional configurations (so which functional space!?), and this should be preserved in time, although correlations continuously form!

... Valid approximation? 100 years of controversy
Lanford’s Theorem (1973)

- Rigorously derives the Boltzmann equation from Newtonian mechanics of $N$ hard spheres of radius $r$, in the scaling $Nr^2 \to 1$, for distributions with Gaussian velocity decay, on a short but macroscopic time interval.
- Improved (Illner & Pulvirenti): rare cloud in $\mathbb{R}^3$
- Recent work on quantum case (Erdös–Yau, Lukkarinen–Spohn)
Leaves many open issues

- Large data? Box? Long-range interactions?
- Nontrajectorial proof??
- Propagation of one-sided chaos??

Qualitative properties?

- Fundamental properties of Newtonian mechanics: preservation of volume in configuration space
- What can be said about the volume in this infinite-dimensional limit??
Boltzmann’s $H$ functional

$$S(f) = -H(f) := - \int_{\Omega_x \times \mathbb{R}^3_v} f(x, v) \log f(x, v) \, dv \, dx$$

Boltzmann identifies $S$ with the entropy of the gas and proves that $S$ can only increase in time (strictly unless the gas is in a hydrodynamical state) — an instance of the Second Law of Thermodynamics
The content of the \( H \) functional

Mysterious and famous, also appears in Shannon’s theory of information

In Shannon’s own words:

I thought of calling it ‘information’. But the word was overly used, so I decided to call it ‘uncertainty’. When I discussed it with John von Neumann, he had a better idea: (...) “You should call it entropy, for two reasons. In first place your uncertainty has been used in statistical mechanics under that name, so it already has a name.
The content of the $H$ functional

Mysterious and famous, also appears in Shannon’s theory of information

In Shannon’s own words:

I thought of calling it ‘information’. But the word was overly used, so I decided to call it ‘uncertainty’. When I discussed it with John von Neumann, he had a better idea: (...) “You should call it entropy, for two reasons. In first place your uncertainty has been used in statistical mechanics under that name, so it already has a name. In second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage.”
Information theory

The Shannon–Boltzmann entropy $S = -\int f \log f$ quantifies how much information there is in a “random” signal, or a language.

$$H_\mu(\nu) = \int \rho \log \rho \, d\mu; \quad \nu = \rho \mu.$$
Microscopic meaning of the entropy functional

Measures the volume of microstates associated, to some degree of accuracy in macroscopic observables, to a given macroscopic configuration (observable distribution function)

⇒ How exceptional is the observed configuration?
**Microscopic meaning of the entropy functional**

Measures the volume of microstates associated, to some degree of accuracy in macroscopic observables, to a given macroscopic configuration (observable distribution function)

⇒ How exceptional is the observed configuration?

**Note:** “Entropy is an anthropic concept”
Microscopic meaning of the entropy functional

Measures the volume of microstates associated, to some degree of accuracy in macroscopic observables, to a given macroscopic configuration (observable distribution function)

⇒ How exceptional is the observed configuration?

Note: “Entropy is an anthropic concept”

Boltzmann’s formula

\[ S = k \log W \]
How to go from $S = k \log W$ to $S = - \int f \log f$?

**Famous computation by Boltzmann**

- $N$ particles in $k$ boxes
- $f_1, \ldots, f_k$ some (rational) frequencies; $\sum f_j = 1$
- $N_j$ = number of particles in box $\#j$
- $\Omega_N(f)$ = number of configurations such that $N_j/N = f_j$
How to go from $S = k \log W$ to $S = - \int f \log f$ ?

**Famous computation by Boltzmann**

$N$ particles in $k$ boxes

$f_1, \ldots, f_k$ some (rational) frequencies; $\sum f_j = 1$

$N_j$ = number of particles in box $\#j$

$\Omega_N(f)$ = number of configurations such that $N_j/N = f_j$

$\Omega_N(f) = 1$

$f = (0, 0, 1, 0, 0, 0, 0)$
How to go from $S = k \log W$ to $S = -\int f \log f$?

Famous computation by Boltzmann

$N$ particles in $k$ boxes

$f_1, \ldots, f_k$ some (rational) frequencies; $\sum f_j = 1$

$N_j = \text{number of particles in box \#} j$

$\Omega_N(f) = \text{number of configurations such that } N_j/N = f_j$

\[
\Omega_8(f) = \frac{8!}{6! \cdot 2!}
\]

\[f = (0, 0, 3/4, 0, 1/4, 0, 0)\]
How to go from $S = k \log W$ to $S = -\int f \log f$?

Famous computation by Boltzmann

$N$ particles in $k$ boxes

$f_1, \ldots, f_k$ some (rational) frequencies; $\sum f_j = 1$

$N_j = \text{number of particles in box } \#j$

$\Omega_N(f) = \text{number of configurations such that } N_j/N = f_j$

\[
\Omega_N(f) = \frac{N!}{N_1! \ldots N_k!}
\]

\[
f = (0, 1/6, 1/3, 1/4, 1/6, 1/12, 0)
\]
→ How to go from $S = k \log W$ to $S = - \int f \log f$ ?

**Famous computation by Boltzmann**

$N$ particles in $k$ boxes

$f_1, \ldots, f_k$ some (rational) frequencies; $\sum f_j = 1$

$N_j = \text{number of particles in box } #j$

$\Omega_N(f) = \text{number of configurations such that } N_j/N = f_j$

Then as $N \to \infty$

$$
\#\Omega_N(f_1, \ldots, f_k) \sim e^{-N \sum f_j \log f_j}
$$

$$
\frac{1}{N} \log \#\Omega_N(f_1, \ldots, f_k) \sim - \sum f_j \log f_j
$$
Sanov’s theorem

A mathematical translation of Boltzmann’s intuition

\[ x_1, x_2, \ldots \ (“microscopic \ variables”) \text{ independent, law } \mu; \]

\[ \hat{\mu}^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} \] (random measure, “empirical”)

What measure will we observe??
Sanov’s theorem

A mathematical translation of Boltzmann’s intuition

\( x_1, x_2, \ldots \) ("microscopic variables") independent, law \( \mu \);

\[ \hat{\mu}^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} \] (random measure, "empirical")

What measure will we observe??

Fuzzy writing: \( \mathbb{P} [\hat{\mu}^N \simeq \nu] \sim e^{-NH_{\mu}(\nu)} \)
Sanov’s theorem

A mathematical translation of Boltzmann’s intuition

\( x_1, x_2, \ldots \) (“microscopic variables”) independent, law \( \mu; \)

\( \hat{\mu}^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} \) (random measure, “empirical”)

What measure will we observe??

Fuzzy writing: \( \mathbb{P} [\hat{\mu}^N \simeq \nu] \sim e^{-NH_\mu(\nu)} \)

Rigorous writing: \( H_\mu(\nu) = \lim_{k \to \infty} \limsup_{\varepsilon \to 0} \limsup_{N \to \infty} \frac{1}{N} \log \mathbb{P}_{\mu^\otimes N} \left[ \left\{ \forall j \leq k, \left| \frac{\varphi_j(x_1) + \ldots + \varphi_j(x_N)}{N} - \int \varphi_j \, d\nu \right| < \varepsilon \right\} \right] \)
Universality of entropy

Lax entropy condition

Entropy is used to select physically relevant shocks in compressible fluid dynamics

(Entropy should increase,
i.e. information be lost, not created!!)
Universality of entropy

Voiculescu’s classification of $\mathcal{II}_1$ factors
Recall: Von Neumann algebras

Initial motivations: Quantum mechanics, group representation theory

\( H \) a Hilbert space

\( \mathcal{B}(H) \): bounded operators on \( H \), with operator norm

Von Neumann algebra \( \mathcal{A} \): a sub-algebra of \( \mathcal{B}(H) \),

- containing \( I \)
- stable by \( A \to A^* \)
- closed for the weak topology (w.r.t. \( A \to \langle A\xi, \eta \rangle \))

The classification of VN algebras is still an active topic with famous unsolved basic problems
States

Type $II_1$ factor := infinite-dimensional VN algebra with trivial center and a tracial state = linear form $\tau$ s.t.

\[ \tau(A^*A) \geq 0 \quad \text{(positivity)} \]
\[ \tau(I) = 1 \quad \text{(unit mass)} \]
\[ \tau(AB) = \tau(BA) \]

($\mathcal{A}, \tau$): noncommutative probability space
States

Type $II_1$ factor := infinite-dimensional VN algebra with trivial center and a **tracial state** = linear form $\tau$ s.t.

\[
\tau(A^*A) \geq 0 \quad \text{(positivity)}
\]
\[
\tau(I) = 1 \quad \text{(unit mass)}
\]
\[
\tau(AB) = \tau(BA)
\]

$(\mathcal{A}, \tau)$: noncommutative probability space

Noncommutative probability spaces as “macroscopic” limits

$X_1^{(N)}, \ldots, X_n^{(N)}$ random $N \times N$ matrices

\[
\tau(P(A_1, \ldots, A_n)) := \lim_{N \to \infty} \frac{1}{N} \mathbb{E} \operatorname{tr} P(X_1^{(N)}, \ldots, X_n^{(N)})
\]

may define a noncommutative probability space.
Voiculescu’s entropy

Think of \((A_1, \ldots, A_n)\) in \((\mathcal{A}, \tau)\) as the observable limit of a family of “microscopic systems” = large matrices

\[
\Omega(N, \varepsilon, k) := \left\{ (X_1, \ldots, X_n), \text{ } N \times N \text{ Hermitian}; \forall P \text{ polynomial of degree } \leq k, \right.
\]

\[
\left. \left| \frac{1}{N} \text{tr} \ P(X_1, \ldots, X_n) - \tau(P(A_1, \ldots, A_n)) \right| < \varepsilon \right\}
\]

\[
\chi(\tau) := \lim_{k \to \infty} \limsup_{\varepsilon \to 0} \limsup_{N \to \infty} \left[ \frac{1}{N^2} \log \text{vol}(\Omega(N, \varepsilon, k)) - \frac{n}{2} \log N \right]
\]
Universality of entropy

Ball–Barthe–Naor’s quantitative central limit theorem

\[ X_1, X_2, \ldots, X_n, \ldots \text{ identically distributed, independent real random variables}; \]

\[ \mathbb{E}X_j^2 < \infty, \quad \mathbb{E}X_j = 0 \]

Then

\[ \frac{X_1 + \ldots + X_N}{\sqrt{N}} \xrightarrow{N \to \infty} \text{Gaussian random variable} \]

Ball–Barthe–Naor (2004): Irreversible loss of information

Entropy \( \left( \frac{X_1 + \ldots + X_N}{\sqrt{N}} \right) \) increases with \( N \)

(some earlier results: Linnik, Barron, Carlen–Soffer)
Back to Boltzmann: The $H$ Theorem

Boltzmann computes the rate of variation of entropy along the Boltzmann equation:

$$\frac{1}{4} \int \left( f(v)f(v_*) - f(v')f(v'_*) \right) \log \frac{f(v)f(v_*)}{f(v')f(v'_*)} B \, d\sigma \, dv \, dv_* \, dx$$

Obviously $\geq 0$

Moreover, $= 0$ if and only if $f(x, v) = \rho(x) \frac{e^{-|v-u(x)|^2/2T(x)}}{(2\pi T(x))^{3/2}} = \text{Hydrodynamic state}$
Reversibility and irreversibility

Entropy increase, but Newtonian mechanics is reversible!?

Ostwald, Loschmidt, Zermelo, Poincaré… questioned Boltzmann’s arguments
Reversibility and irreversibility

Entropy increase, but Newtonian mechanics is reversible!?

Ostwald, Loschmidt, Zermelo, Poincaré... questioned Boltzmann’s arguments
Reversibility and irreversibility

Entropy increase, but Newtonian mechanics is reversible!?

Ostwald, Loschmidt, Zermelo, Poincaré... questioned Boltzmann’s arguments
Reversibility and irreversibility

Entropy increase, but Newtonian mechanics is reversible!?

Ostwald, Loschmidt, Zermelo, Poincaré... questioned
Boltzmann’s arguments
Reversibility and irreversibility

Entropy increase, but Newtonian mechanics is reversible!?

Ostwald, Loschmidt, Zermelo, Poincaré... questioned Boltzmann’s arguments
Loschmidt’s paradox

At time $t$, reverse all velocities (entropy is preserved), start again (entropy increases) but go back to initial configuration (and initial entropy!)
Loschmidt’s paradox

At time $t$, reverse all velocities (entropy is preserved), start again (entropy increases) but go back to initial configuration (and initial entropy!)

Boltzmann: “Go on, reverse them!”
**Loschmidt’s paradox**

At time $t$, reverse all velocities (entropy is preserved), start again (entropy increases) but go back to initial configuration (and initial entropy!)

**Boltzmann:** “Go on, reverse them!”

... Subtleties of one-sided chaos

... Importance of the prepared starting state:
Loschmidt’s paradox

At time $t$, reverse all velocities (entropy is preserved), start again (entropy increases) but go back to initial configuration (and initial entropy!)

Boltzmann: “Go on, reverse them!”

... Subtleties of one-sided chaos

... Importance of the prepared starting state:

low macroscopic entropy (unlikely observed state),

high microscopic entropy (no initial correlations)
Mathematicians chose their side

All of us younger mathematicians stood by Boltzmann’s side.

Arnold Sommerfeld (about a 1895 debate)

Boltzmann’s work on the principles of mechanics suggest the problem of developing mathematically the limiting processes (...) which lead from the atomistic view to the laws of motion of continua.

David Hilbert (1900)

Boltzmann summarized most (but not all) of his work in a two volume treatise Vorlesungen über Gastheorie. This is one of the greatest books in the history of exact sciences and the reader is strongly advised to consult it.

Mark Kac (1959)
Recall: The most important nonlinear models of classical mechanics have not been solved (compressible and incompressible Euler, compressible and incompressible Navier–Stokes, Boltzmann, even Vlasov–Poisson to some extent)

But Boltzmann’s $H$ Theorem was the first qualitative estimate!

- Finiteness of the entropy production prevents clustering, providing compactness, crucial in the Arkeryd and DiPerna–Lions stability theories
- The $H$ Theorem is the main “explanation” for the hydrodynamic approximation of the Boltzmann equation
Large-time behavior

The state of maximum entropy given the conservation of total mass and energy is a Gaussian distribution

... and this Gaussian distribution is the only one which prevents entropy from growing further
If this is true, then $f$ becomes Gaussian as $t \to \infty$!
What do we want to prove?

Prove that a “nice” solution of the Boltzmann equation approaches Gaussian equilibrium: short and easy (basically Boltzmann’s argument)

Get quantitative estimates like “After a time ......, the distribution is close to Gaussian, up to an error of 1 %”: tricky and long (1989–2004 starting with Cercignani, Desvillettes, Carlen–Carvalho)
Conditional convergence theorem (Desvillettes – V)

Let \( f(t, x, v) \) be a solution of the Boltzmann equation, with appropriate boundary conditions. \textbf{Assume} that

(i) \( f \) is very regular (uniformly in time): all moments (\( \int f|v|^k\ dv\ dx \)) are finite and all derivatives (of any order) are bounded;

(ii) \( f \) is strictly positive: \( f(t, x, v) \geq Ke^{-A|v|^q} \).

Then \( f(t) \longrightarrow f_\infty \) at least like \( O(t^{-\infty}) \) as \( t \rightarrow \infty \).
Conditional convergence theorem (Desvillettes – V)

Let \( f(t, x, v) \) be a solution of the Boltzmann equation, with appropriate boundary conditions. Assume that

(i) \( f \) is very regular (uniformly in time): all moments (\( \int f|v|^k \, dv \, dx \)) are finite and all derivatives (of any order) are bounded;

(ii) \( f \) is strictly positive: \( f(t, x, v) \geq Ke^{-A|v|^q} \).

Then \( f(t) \rightarrow f_\infty \) at least like \( O(t^{-\infty}) \) as \( t \rightarrow \infty \)

Rks: (i) can be relaxed, (ii) removed by regularity theory.

- Gualdani–Mischler–Mouhot complemented this with a rate \( O(e^{-\lambda t}) \) for hard spheres
The proof uses differential inequalities of first and second order, coupled via many inequalities including:

- precised **entropy production inequalities** (information theoretical input)
- Instability of hydrodynamical description (fluid mechanics input)
- decomposition of entropy in **kinetic** and **hydrodynamic** parts
- functional inequalities coming from various fields (information theory, quantum field theory, elasticity theory, etc.)

It led to the discovery of **oscillations** in the entropy production
Numerical simulations by Filbet

These oscillations of the entropy production slow down the convergence to equilibrium (fluid mechanics effect) and are related to current research on the convergence for nonsymmetric degenerate operators ("hypocoercivity")
Other quantitative use of entropy a priori estimates

- Instrumental in Nash’s proof of continuity of solutions of linear diffusion equations with nonsmooth coefficients
- Allowed Perelman to prove the Poincaré conjecture
- Used by Varadhan and coworkers/students for (infinite-dimensional) hydrodynamic problems
- Basis for the construction of the heat equation in metric-measure spaces (... Otto ... V ... Savaré, Gigli ...)
Other quantitative use of entropy a priori estimates

- Instrumental in Nash’s proof of continuity of solutions of linear diffusion equations with nonsmooth coefficients
- Allowed Perelman to prove the Poincaré conjecture
- Used by Varadhan and coworkers/students for (infinite-dimensional) hydrodynamic problems
- Basis for the construction of the heat equation in metric-measure spaces (... Otto ... V ... Savaré, Gigli ...)
- ...

A final example of universality of entropy

Lott–Sturm–Villani’s synthetic “Ricci ≥ 0”
Ricci curvature

One of the three most popular notions of curvature. At each $x$, $\text{Ric}$ is a quadratic form on $T_xM$.

$$(\text{Ric})_{ij} = (\text{Riem})^k_{kij}$$

It measures the rate of separation of geodesics in a given direction, in the sense of volume (Jacobian)
Because of positive curvature effects, the observer overestimates the surface of the light source; in a negatively curved world this would be the contrary.

\[ \text{Distortion coefficients always } \geq 1 \] \iff \[ \text{Ric } \geq 0 \]
The connection between optimal transport of probability measures and Ricci curvature was recently studied (McCann, Cordero-Erausquin, Otto, Schmuckenschläger, Sturm, von Renesse, Lott, V).

A theorem obtained by these tools (Lott–V; Sturm):

A limit of manifolds with nonnegative Ricci curvature, is also of nonnegative curvature.
The connection between **optimal transport of probability measures** and Ricci curvature was recently studied (McCann, Cordero-Erausquin, Otto, Schmuckenschläger, Sturm, von Renesse, Lott, V).

A theorem obtained by these tools (Lott–V; Sturm):

**A limit of manifolds with nonnegative Ricci curvature, is also of nonnegative curvature.**

Limit in which sense? **Measured Gromov–Hausdorff topology** (very weak: roughly speaking, ensures convergence of distance and volume)
The connection between optimal transport of probability measures and Ricci curvature was recently studied (McCann, Cordero-Erausquin, Otto, Schmuckenschläger, Sturm, von Renesse, Lott, V)

A theorem obtained by these tools (Lott–V; Sturm):

A limit of manifolds with nonnegative Ricci curvature, is also of nonnegative curvature.

Limit in which sense? Measured Gromov–Hausdorff topology (very weak: roughly speaking, ensures convergence of distance and volume)

A key ingredient: Boltzmann’s entropy!
The lazy gas experiment

Describes the link between Ricci and optimal transport

\[ S = - \int \rho \log \rho \]
1946: Landau’s revolutionary paradigm

Landau studies equations of plasma physics: collisions essentially negligible (Vlasov–Poisson equation)

Landau suggests that there is dynamic stability near a stable homogeneous equilibrium $f^0(v)$

... even if no collisions, no diffusion, no $H$ Theorem, no irreversibility. Entropy is constant!!

Landau and followers prove relaxation only in the linearized regime, but numerical simulations suggest wider range. Astrophysicists argue for relaxation with constant entropy = violent relaxation
force $F[f](t, x) = - \int \int \nabla W(x - y) f(t, y, w) \, dy \, dw$
But ... Is the linearization reasonable?? \( f = f^0 + h \)

\[
\frac{\partial h}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} h + F[h] \cdot \left( \nabla_{\mathbf{v}} f^0 + \nabla_{\mathbf{v}} h \right) = 0 \quad \text{(NLin \ V)}
\]

\[
\frac{\partial h}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} h + F[h] \cdot \left( \nabla_{\mathbf{v}} f^0 + 0 \right) = 0 \quad \text{(Lin \ V)}
\]
But ... Is the linearization reasonable??  \( f = f^0 + h \)

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \left( \nabla_v f^0 + \nabla_v h \right) = 0 \quad \text{(NLin V)}
\]

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \left( \nabla_v f^0 + 0 \right) = 0 \quad \text{(Lin V)}
\]

- OK if \( |\nabla_v h| \ll |\nabla_v f^0| \), but \( |\nabla_v h(t, \cdot)| \geq \varepsilon t \to +\infty \)

“destroying the validity of the linear theory”  (Backus 1960)
But ... Is the linearization reasonable??  \( f = f^0 + h \)

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \left( \nabla_v f^0 + \nabla_v h \right) = 0 \quad \text{(NLin V)}
\]

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \left( \nabla_v f^0 + 0 \right) = 0 \quad \text{(Lin V)}
\]

- OK if \(|\nabla_v h| \ll |\nabla_v f^0|\), but \(|\nabla_v h(t, \cdot)| \geq \varepsilon t \rightarrow +\infty\) “destroying the validity of the linear theory” (Backus 1960)

- Natural nonlinear time scale = \(1/\sqrt{\varepsilon}\) (O’Neil 1965)
But ... Is the linearization reasonable?? \( f = f^0 + h \)

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \left( \nabla_v f^0 + \nabla_v h \right) = 0 \quad \text{(NLin V)}
\]

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \left( \nabla_v f^0 + 0 \right) = 0 \quad \text{(Lin V)}
\]

- OK if \( |\nabla_v h| \ll |\nabla_v f^0| \), but \( |\nabla_v h(t, \cdot)| \geq \varepsilon t \to +\infty \) “destroying the validity of the linear theory” (Backus 1960)

- Natural nonlinear time scale = \( 1/\sqrt{\varepsilon} \) (O’Neil 1965)

- Neglected term \( \nabla_v h \) is dominant-order!
But ... Is the linearization reasonable?? $f = f^0 + h$

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot (\nabla_v f^0 + \nabla_v h) = 0 \quad \text{(NLin V)}
\]

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot (\nabla_v f^0 + 0) = 0 \quad \text{(Lin V)}
\]

- OK if $|\nabla_v h| \ll |\nabla_v f^0|$, but $|\nabla_v h(t, \cdot)| \geq \varepsilon t \to +\infty$ “destroying the validity of the linear theory” (Backus 1960)

- Natural nonlinear time scale $= 1/\sqrt{\varepsilon}$ (O’Neil 1965)

- Neglected term $\nabla_v h$ is dominant-order!

- Linearization removes entropy conservation
But ... Is the linearization reasonable?? \( f = f^0 + h \)

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \left( \nabla_v f^0 + \nabla_v h \right) = 0 \quad \text{(NLin V)}
\]

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \left( \nabla_v f^0 + 0 \right) = 0 \quad \text{(Lin V)}
\]

• OK if \(|\nabla_v h| \ll |\nabla_v f^0|\), but \(|\nabla_v h(t, \cdot)| \geq \varepsilon t \to +\infty\)
  “destroying the validity of the linear theory” (Backus 1960)

• Natural nonlinear time scale = \(1/\sqrt{\varepsilon}\) (O’Neil 1965)

• Neglected term \(\nabla_v h\) is dominant-order!

• Linearization removes entropy conservation

• Isichenko 1997: approach to equilibrium only \(O(1/t)\)
But ... Is the linearization reasonable?? \( f = f^0 + h \)

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \left( \nabla_v f^0 + \nabla_v h \right) = 0 \quad \text{(NLin V)}
\]

\[
\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \left( \nabla_v f^0 + 0 \right) = 0 \quad \text{(Lin V)}
\]

- OK if \(|\nabla_v h| \ll |\nabla_v f^0|\), but \(|\nabla_v h(t, \cdot)| \geq \varepsilon t \to +\infty\) “destroying the validity of the linear theory” (Backus 1960)
- Natural nonlinear time scale = \(1/\sqrt{\varepsilon}\) (O’Neil 1965)
- Neglected term \(\nabla_v h\) is dominant-order!
- Linearization removes entropy conservation
- Isichenko 1997: approach to equilibrium only \(O(1/t)\)
- Caglioti–Maffei (1998): at least some nontrivial solutions decay exponentially fast
Theorem (Mouhot–V. 2009)

Landau damping holds for the nonlinear Vlasov equation: If i.d. very close to a stable homogeneous equilibrium $f^0(v)$, then convergence in large time to a stable homogeneous equilibrium.

- Information still present, but not observable (goes away in very fast velocity oscillations)
- Relaxation by regularity, driven by confined mixing.
- Proof has strong analogy with K-A-M theory
- First steps in a virgin territory? Universal problem of constant-entropy relaxation in particle systems
Theorem (Mouhot–V. 2009)

Landau damping holds for the nonlinear Vlasov equation: If i.d. very close to a stable homogeneous equilibrium \( f^0(v) \), then convergence in large time to a stable homogeneous equilibrium.

- Information still present, but not observable (goes away in very fast velocity oscillations)
- Relaxation by regularity, driven by confined mixing.
- Proof has strong analogy with K-A-M theory
- First steps in a virgin territory? Universal problem of constant-entropy relaxation in particle systems
- ... Will entropy come back as a selection principle??
Goal of mathematical physics?

- Not to rigorously prove what physicists already know
- Through mathematics get new insights about physics
- From physics identify new mathematical problems