The work of Cédric Villani

Citation

Published Version
doi:10.1142/9789814324359_0004

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:32706715

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
The Work of Cédric Villani

Horng-Tzer Yau
Department of Mathematics,
Harvard University
Cambridge MA 02138
August 10, 2010

1 Introduction

The starting point of Cédric Villani’s work goes back to the introduction of entropy in the nineteenth century by L. Carnot and R. Clausius. At the time, entropy was a vague concept and its rigorous definition had to wait until the fundamental work of L. Boltzmann who introduced nonequilibrium statistical physics and the famous $H$ functional. Boltzmann’s work, though a fundamental breakthrough, did not resolve the question concerning the nature of entropy and time arrow; the debate on this central question continued for a century until today. J. von Neumann, in recommending C. Shannon to use entropy for his uncertainty function, quipped that entropy is a good name because “nobody knows what entropy really is, so in a debate you will always have the advantage”.

The first result of Villani I will report on concerns the fundamental connection between entropy and its dissipation. In this work, we will see that rigorous mathematical analysis is not just a display of powerful analytic skill, but also leads to deep insights into nature. Based on this work, Villani has developed a general theory, hypercoercivity, which applies to broad systems of equations. In a separate direction, entropy was used by Villani as a fundamental tool in optimal transport and the study of curvature in metric spaces. Finally, I will describe Villani’s work on Landau damping, which predicts a very surprising decay (and thus the word damping) of the electric field in a plasma without particle collisions, and therefore without entropy increase. This is in sharp contrast with Boltzmann’s picture that the time irreversibility comes from collision processes.

2 Boltzmann Equation

The Boltzmann equation was derived by L. Boltzmann in 1873 based on his physical intuition of collision processes. The most striking feature of the Boltzmann equation, the time irreversibility, contradicts the reversibility of the Newton equations. This fact is most concisely expressed via the Boltzmann H-theorem stating that the entropy

$$S = - \int \int f \log f \, dv \, dx$$

is always nondecreasing. Furthermore, the entropy production vanishes if and only if the state is spatially homogeneous and Maxwellian in the velocity variable. The Boltzmann H-theorem is semi-rigorous in the sense
that if the solution to the Boltzmann equation is sufficiently smooth then the original proof of Boltzmann is rigorous. The mathematical study of the Boltzmann equation started perhaps from T. Carleman and H. Grad in the middle of the last century. Despite decades of intensive research, most fundamental questions concerning the Boltzmann equation remain open, e.g., 1. Are solutions of the Boltzmann equation smooth if the initial data are sufficiently smooth? 2. The Boltzmann H-theorem states that the entropy increases, but what is the rate? Or, more generally, how fast do solutions to the Boltzmann equation approach the equilibrium (Maxwellian) states?

The first question, the regularity of the Boltzmann equation, is only understood for small perturbation of equilibrium measures. There is a general framework of renormalized solutions developed by R. DiPerna and P.-L. Lions [16], but precise estimates on the solutions remain elusive. The second question, the decay to equilibrium for the Boltzmann equation, is where Villani made his fundamental contribution. Before we describe Villani’s work in some detail, several important recent results concerning the Boltzmann equation should be mentioned here. This incomplete list includes the well-posedness and the approach to equilibrium for small perturbation data by Y. Guo [24], and the recent extension of this approach to long-range interactions and soft potentials by P. Gressman and R. Strain [23], the weak shock solutions by S.-H. Yu [52], the derivation of incompressible Navier-Stokes equations from the Boltzmann equation by C. Bardos, F. Golse, D. Levermore and L. Saint-Raymond [5, 17].

The Boltzmann equation is given by

$$\partial_t f + v \nabla_x f = Q(f, f)$$

where $f(t, x, v)$ is the probability density in the phase space at the time $t$. The nonlinear term $Q$ is the collision operator

$$Q(f, f) = \int \int \int [f(v') f(v'_*) - f(v) f(v_*)] B(v - v_*, \sigma) \, dv_* \, d\sigma,$$

where $v, v_*$ are the incoming velocities, $v', v'_*$ the outgoing velocities, $\sigma$ the collision angle and $B$ is the scattering kernel depending on the details of the microscopic interactions. The Boltzmann H-functional (negative of the entropy) is defined by

$$H(f) = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} f(x, v) \log f(x, v) \, dx \, dv$$

and the Boltzmann H-theorem states that

$$\partial_t H(f(t)) = -D(f(t)) \leq 0, \quad D(f) = \frac{1}{4} \int \int \int [f(v) f(v_*) - f(v') f(v'_*)] \log \frac{f(v) f(v_*)}{f(v') f(v'_*)} \, d\sigma \, dv \, dv_*.$$

The dissipation $D$ vanishes if and only if the state is Maxwellian:

$$D(f) = 0 \quad \text{if and only if} \quad f = M_{\rho, u, T} := \rho(x) \frac{e^{-\frac{|v-u(x)|^2}{2T(x)}}}{(2\pi T(x))^{3/2}}$$

where $M$ is any local equilibrium state of the Boltzmann equation with density $\rho$, velocity $u$ and temperature $T$ which can depend on the space variable. If $\rho, u, T$ are independent of the space variables, $M$ is called a global equilibrium.

To understand the approach to equilibrium via the Boltzmann H-theorem, we first consider the spatially homogeneous case, i.e., the function $f(x, v)$ depends only on $v$. C. Cercignani in 1983 [11] conjectured that,
under suitable assumptions on the collision kernel $B$, there is a constant $K$ such that

$$D(f) \geq K(f)H(f|M), \quad H(f|M) := \int dx \, dv \, f \log \frac{f}{M}$$

where $M$ is a global equilibrium and $H(f|M)$ is the entropy of $f$ relative to a global equilibrium $M$. The Cercignani conjecture is similar to the logarithmic Sobolev inequality for the diffusion process, but the dissipation operator $D$ is nonlinear in the function $f$. If the Cercignani conjecture holds, then the decay to global equilibrium would be exponentially fast. Through counterexamples, A.V. Bobylev and C. Cercignani [7] proved that this conjecture is false if the constant $K$ depends on the function $f$ only through finite Sobolev norms and moments. On the other hand, it was shown by E. Carlen and M. Carvalho [12] that $D(f) \geq \Theta_f(H(f|M))$, where the function $\Theta$ is not explicit but depends on $f$ only through its moments and some derivatives. The conjecture was finally settled in a joint work by G. Toscani and Villani [45] and the subsequent work by Villani [47]. The conclusion is as surprising as it is beautiful: The Cercignani conjecture is in general false, but it is always almost correct in the following sense.

**Theorem 2.1** For a reasonable physical scattering kernel $B$, if $f$ is smooth and with certain decay property in high momentum regime, then for any $\varepsilon > 0$ we have

$$D(f) \geq K_\varepsilon(f)H(f|M)^{1+\varepsilon}$$

where $K_\varepsilon(f)$ depends on the smoothness and moments of $f$.

This inequality then implies that the entropic convergence rate is faster than $C_\varepsilon(f_0) t^{-1/\varepsilon}$ for any initial smooth data $f_0$. This is a much deeper inequality than the logarithmic Sobolev inequality, as the operator $D$ is nonlinear and the inequality fails for $\varepsilon = 0$ except in certain nonphysical situations, such as when the collision kernel is quadratic at large velocities.

Villani’s next project in this direction is the very ambitious extension of this theorem to the spatially inhomogeneous case. One’s immediate reaction to this question is that this is beyond reach since there is no global existence theory for the Boltzmann equation. The key physical question, however, is to understand the mechanism that leads to relaxation in the space variable. If we assume that good smooth solutions are given, the intrinsic difficulties are immediately visible: The identity $D(f) = 0$ implies that $f$ is a local Maxwellian, but not necessarily a global one, i.e., the density, temperature, and velocity parameters in the Maxwellian (2.2) depend on the space variable. Therefore, the relaxation to the global Maxwellian requires an additional mechanism different from the consideration of the entropy production. The only control on the space variable in the Boltzmann equation is the first order operator $v \nabla_x$. Now we have a formidable problem: It is analogue of a hypoelliptic problem, but the elliptic part is a nonlinear integral operator! Numerically, the entropy does decay very fast in the spatially inhomogeneous case, but the entropy production is far from monotonic. The main result in this direction is the following theorem by L. Desvillettes and Villani [15].

**Theorem 2.2** Suppose that $f_t(x,v)$ is a regular solution to the Boltzmann equation and $f_t$ satisfies some lower bound estimate in the large velocity region. Under some assumptions on the collision kernel $B$, for any $\varepsilon > 0$ there is a $C_\varepsilon$ such that

$$H(f_t|M) \leq C_\varepsilon(f_0) t^{-1/\varepsilon}$$

where $f_0$ is the initial value of the Boltzmann equation.
This result assumes that the regularity of \( f_t(x, v) \) is given, but is a large data theorem in the sense that there is no smallness condition on the initial function \( f_0 \). With a smallness condition, i.e., if the initial data is near a global Maxwellian, the assumptions of Theorem 2.2 can be verified, see, e.g., [24, 25]. Furthermore, significant progress was made in this direction for soft and long range potentials [23] and the decay rates can be exponentially fast for certain collision operators [22]. The method introduced to prove Theorem 2.2 is a very powerful one; Villani later developed a general theory, hypocoercivity [50], to estimate the large time asymptotics of a general class of hypoelliptic operators.

This program was also continued by younger mathematicians, in particular in the series of papers by C. Mouhot, C. Baranger, R. Strain, M. Gualdani and S. Mischler, on the spectral gap for the linearized Boltzmann operator [6, 34] and on the matching of the nonlinear convergence to equilibrium with the linearized theory, in a homogeneous setting [33] and in an inhomogeneous hypocoercive setting [19].

Finally, we mention that Villani’s other work related to the Boltzmann equation includes a series of papers on the influence of grazing collisions, mainly with L. Desvillettes and R. Alexandre: existence of renormalized solutions (with defect measure) for the Boltzmann equation without cutoff [2], the rigorous derivation of the Fokker–Planck–Landau equation from the Boltzmann equation in the grazing collision limit [46, 3], and sharp regularity bounds associated with entropy production [1].

3 Optimal Transportation and Curvature

The optimal transport problem, also known as Monge-Kantorovich problem, is an ancient engineering problem seeking to minimize the cost to transport mass. For our purpose, the initial and final mass distributions are given by two probability measures \( \mu \) and \( \nu \) on a compact measurable metric space \( X \). The goal is to find a measurable map \( T : X \to X \) with \( T\# \mu = \nu \) to minimize the transportation cost

\[
\int c(x, T(x)) \, d\mu(x).
\]

The square root of the minimal transportation cost with the squared distance transportation cost function \( c(x, y) = d(x, y)^2 \) is called the 2-Wasserstein distance, \( W_2 \), between these two measures. The minimizer \( T \) is called the optimal transport map. The existence and uniqueness of the optimal transport map was proved in the Euclidean space by Y. Brenier [9] and in the Riemannian manifolds by R. McCann [32].

The probability measures on the metric space \( X \) with the Wasserstein distance constitute a compact metric space, called the Wasserstein space \( (P(X), W_2) =: P_2(X) \) on \( X \). We now take \( X \) to be a compact manifold \( M \) with metric tensor \( g \) which in turns generates a geodesic distance \( d \) and the normalized volume measure \( \nu = \text{dvol}_M / \text{vol}_M \). The information (negative of the entropy) \( H(\mu) \) of a measure \( \mu = \rho \nu \) absolutely continuous w.r.t. \( \nu \) is defined by

\[
H(\mu) = \int \rho \log \rho \, d\nu.
\]

In a study of nonlinear heat equations, F. Otto [38] defined a formal Riemannian structure on \( P_2(M) \) and interpreted these equations as gradient flows on the Wasserstein space \( P_2(M) \) with this formal Riemannian structure. Subsequently, Otto and Villani [39] found the remarkable property that the entropy, viewed as a functional on the Wasserstein space \( P_2(M) \), is concave if the Ricci curvature of the manifold \( M \) is nonnegative. This provided the first link between the concavity of entropy on the Wasserstein space and the Ricci curvature of the underlying manifold. This relation was subsequently established rigorously in [13], partly motivated by the earlier work [31]. Otto and Villani [39] also argued that the converse should hold, and it was rigorously established in [41].
If we replace the convexity of entropy by a lower bound $K$ on the Hessian of entropy on $P_2(M)$, the corresponding condition on the Ricci curvature becomes the lower bound

$$\text{Ric} \geq Kg.$$

Furthermore, the volume measure can be generalized to the weighted volume measure $e^{-\Phi}d\text{vol}$ provided the Ricci curvature is replaced by the Bakry-Émery tensor

$$\text{Ric}_\infty := \text{Ric} + \text{Hess} (\Phi) \geq Kg.$$

Using this heuristic idea, Otto and Villani [39] then provided a unified approach to a wide range of inequalities in analysis and geometry including the logarithmic Sobolev inequality and Talagrand’s concentration inequality.

If $P_2(M)$ is a regular Riemannian manifold, a lower bound $K$ on the Hessian of the entropy functional is equivalent to the displacement convexity inequality

$$H(\mu_t) \leq (1 - t)H(\mu_0) + tH(\mu_1) - Kt(1 - t)W_2(\mu_0, \mu_1)^2$$

for any Wasserstein geodesic $\mu_t$. Notice that this definition depends only on the concept of geodesic on the Wasserstein space which can be defined on any metric space. There is no need for a Riemannian structure on $M$ if we take (3.2) as the definition that the Ricci curvature on a metric space $X$ is bounded below by $K$. With this definition of a lower bound on the Ricci curvature, J. Lott and Villani [36] proved the fundamental stability result that the lower bound on the Ricci curvature is stable under the Gromov-Hausdorff convergence. A closely related definition of Ricci curvature lower bounds, and similar stability results were obtained independently by K.-T. Sturm [43, 44]. The main statement of Lott-Villani’s results can be stated as follows.

**Theorem 3.1** Let $\{(X_i, d_i, \nu_i)\}$ be a sequence of compact measured length spaces and $\lim_{i \to \infty} (X_i, d_i, \nu_i) = (X, d, \nu)$ in the measured Gromov-Hausdorff topology. If the Ricci curvature of $(X_i, d_i, \nu_i)$ is bounded below by $K$ then the Ricci curvature of $(X, d, \nu)$ is also bounded below by $K$.

This theorem demonstrates the robustness of this definition of Ricci curvature lower bounds. On the other hand, the definition is also a very effective notion since it allows one to generalize many theorems in Riemannian geometry to the setting of metric spaces, including the Bishop-Gromov theorem, logarithmic Sobolev inequality and Bonnet-Myers theorem. Moreover, the definition can easily be discretized [8]. We note that there are other notions and definitions of curvatures on metric spaces or graphs. This includes the work of Y. Ollivier [37] and F. R. Chung and S.-T. Yau’s definitions of curvatures on graphs [18, 28].

To summarize, Villani has brought the tools of entropy and its time-evolution from the study of convergence to equilibrium in the Boltzmann equation to a geometric setting involving the Wasserstein space. In addition to the Ricci curvature, Villani has explored connections with other geometric or analytic problems, such as the Sobolev inequality, for which he has provided a new proof based on optimal transport and entropy-type functionals, in collaboration with D. Cordero-Erausquin and B. Nazaret [14]. Like [36], this paper was a starting point for other developments, including the work of A. Figalli, F. Maggi and A. Pratelli [21] on quantitative anisotropic isoperimetric inequalities. Villani also wrote a series of papers with A. Figalli, G. Loeper and L. Rifford relating the smoothness of optimal transport with the shape of the cut locus in Riemannian geometry [30, 20].
4 Landau Damping

The last theorem of Villani I will describe is a rigorous proof of Landau damping in the nonlinear setting. The fundamental equation governing plasma dynamics is the Vlasov-Poisson equation, which, for periodic data is given by

$$\partial_t f + v \nabla_x f + E \nabla_v f = 0, \quad f(t, x, v) \geq 0$$

where $f(t, x, v)$ is the density of charged particles with velocity $v \in \mathbb{R}^3$ at $x \in \mathbb{T}^3$, the unit torus in $\mathbb{R}^3$. The electric field $E$ is related to the density of charged particles $\rho(t, x) = \int f(t, x, v) dv$ via the Poisson equation

$$E := E[\rho] := -\nabla \phi, \quad -\Delta \phi = \rho(x) - 1$$

where the constant 1 is the density of background charges normalized to be one. This equation describes the dynamics of galaxies if we make a sign change in the electric field due to the sign difference between the Coulomb and gravitational forces. The result I will describe is valid with both signs, but I will use the language of plasma physics.

The Vlasov-Poisson equation describes collisionless dynamics and is time reversible. It is well known that dissipative dynamics often approach equilibrium exponentially fast, but for reversible dynamics the state $f_t$ at any given time carries the same information as the initial data and decay to equilibrium can only be valid after certain averaging. On the other hand, fast relaxation to equilibrium in nature is ubiquitous even for systems governed by Newtonian dynamics. The common explanation has been that the relaxation is due to collision processes which produce dissipation. In 1946, L. Landau [27] revolutionized this common belief by arguing that the electric field in the Vlasov-Poisson equation, which is a collisionless equation, decays exponentially fast. He computed this rate of convergence for the linearized Vlasov-Poisson equation. This astonishing discovery is thus termed Landau damping. Despite intensive studies, the understanding of Landau damping for the Vlasov-Poisson equation is very limited.

The Vlasov-Poisson equation has infinitely many stationary solutions. In fact, any density $g(v)$ satisfying the normalization condition $\int g(v) dv = 1$ is a stationary solution. The stability analysis of the linearized Vlasov-Poisson equation was mainly due to the work of O. Penrose [40] in the sixties. It states that $f^0$ is stable if for any $\sigma \in S^2$ and $f_\sigma(v) = \int_{v_\sigma + \sigma^\perp} f^0(z) dz$, then for any $w$ such that $f'_\sigma(w) = 0$ one has

$$\int f'_\sigma(v) \frac{dv}{v - w} dv < 1.$$  

The Landau damping for the linearized Vlasov-Poisson was already understood in the sixties by the work of A. Saenz [42]; for the quasi-linear case only nonrigorous results were available. On the other hand, it was pointed out by G. Backus [4] that the linear approximation is not expected to be valid for the full nonlinear equation in the large time regime. For the nonlinear Landau damping, the only partial results available were examples of solution to the Vlasov-Poisson equation that exhibit Landau damping [10, 26]. Last year Mouhot and Villani [35] proved that, for any analytic data near an analytic linearly stable stationary state, the electric field decay exponentially fast. Notice that the analyticity assumption is not an artifact of the proof, it is in fact necessary [29]. This resolves the long standing problem of Landau damping. We will not be able to state their theorem precisely nor in its general form, but the following limited version in the physical dimension $d = 3$ gives a flavor of the depth of the full theorem.

**Theorem 4.1 (nonlinear Landau damping for general interaction)** There is an analytic norm $\| \cdot \|_a$ on functions of the phase space such that the following holds: Let $f^0 : \mathbb{R}^d \to \mathbb{R}_+$ be an analytic stationary
state satisfying Penrose’s linear stability criterion. Suppose the initial profile \( f_i \geq 0 \) is near the analytic stationary state in the sense that
\[
\|f_i - f^0\|_a \leq \varepsilon, \quad (4.1)
\]
for some \( \varepsilon \) sufficiently small. Then there are analytic profiles \( f_{+\infty}(v), f_{-\infty}(v) \) such that
\[
f(t, \cdot) \xrightarrow{t \to \pm \infty} f_{\pm \infty} \quad \text{weakly}
\]
exponentially fast. Furthermore, the marginal density of the unique solution of the nonlinear Vlasov equation with initial value \( f(0, \cdot) = f_i \) converges exponentially fast as time \( t \to \pm \infty \), i.e., there exists \( \rho_\infty \) and \( \lambda > 0 \) such that for all integer \( r \) we have
\[
\|\rho(t, \cdot) - \rho_\infty\|_{C^r(\mathbb{R})} \leq C e^{-\lambda|t|}, \quad t \to \pm \infty, \quad (4.2)
\]
where \( C^r \) denotes the \( L_\infty \) norm of the derivatives up to order \( r \).

Since the Vlasov-Poisson equation is time reversible, the profiles \( f_i \) keep the memory of the initial datum for all time. The fast relaxations in Mouhot-Villani’s theorem only refer to averaged quantities such as density in the position space or in the weak sense. This is due to the fact that weak convergence preserves only the information of low frequency modes; the information at low frequencies was transferred to high frequencies to “maintain the constant total information for all times”. Although \( f_i \) carries all information of the initial data for all time, more and more information is stored at high frequency modes. Hence if we only look at low frequency modes (such as weak convergence), there is a loss of information and this is responsible for the fast relaxation of various averaged quantities. This resembles the phenomena in turbulence and it requires very precise understanding of this transfer of information to yield a mathematical proof. Mouhot-Villani’s theorem is the first rigorous result to establish a fast decay to equilibrium, a time irreversible behavior, in confined collisionless time-reversible dynamics.

## 5 Conclusion

In Villani’s work, we have seen not only rigorous mathematical analysis providing deep insights into physical behavior, but also important new mathematics emerging from the study of natural phenomena, in the spirit of Maxwell and Boltzmann. Besides his research articles, Villani has written extensive surveys and books [48, 50, 49, 51], and, through these, as well as the insights of his work, he has inspired a generation of young mathematicians with deep, rich, physically motivated mathematical questions. We are witnessing the beginning of Villani’s spectacular career and influence in shaping the directions of analysis and mathematics.

## References


