ON THE TREND TO EQUILIBRIUM FOR THE FOKKER-PLANCK EQUATION : AN INTERPLAY BETWEEN PHYSICS AND FUNCTIONAL ANALYSIS

P. A. Markowich **C. Villani** *

Abstract

We present connections between the problem of trend to equilibrium for the Fokker-Planck equation of statistical physics, and several inequalities from functional analysis, like logarithmic Sobolev or Poincaré inequalities, together with some inequalities arising in the context of concentration of measures, introduced by Talagrand, or in the study of Gaussian isoperimetry.

1. The Fokker-Planck equation

The Fokker-Planck equation is basic in many areas of physics. It reads

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \Big(D(\nabla \rho + \rho \nabla V) \Big), \qquad t \geq 0, \quad x \in \mathbb{R}^n, \tag{1}$$

where $D = D(x)$ is a symmetric, locally uniformly positive definite (diffusion) matrix, and $V = V(x)$ a confining potential. Here the unknown $\rho = \rho(t,x)$ stands for the density of an ensemble of particles, and without loss of generality can be assumed to be a probability distribution on $\mathbb{R}^n$ since the equation (1) conserves nonnegativity and the integral of the solution over $\mathbb{R}^n$. The phase space can be a space of position vectors, but also a space of velocities v; in the latter case the potential V is usually the kinetic energy $|v|^2/2$.

We refer to [32] for a phenomenological derivation, and a lot of basic references. The Fokker-Planck equation can be set on any differentiable structure,

*The authors acknowledge support by the EU-funded TMR-network 'Asymptotic Methods in Kinetic Theory' (Contract # ERB FMRX CT97 0157) and from the Erwin-Schrödinger-Institute in Vienna.

in particular on a Riemannian manifold M, rather than on Euclidean space $\mathbb{R}^n$. It can also be considered in a bounded open set, with (say) a vanishing out-flux condition at the boundary.

Far from aiming at a systematic study of equation (1), our intention here is to focus on some tight links between this equation, and several functional inequalities which have gained interest over the last decade, and especially in the last years. In order to simplify the presentation, we restrict to the case when the diffusion matrix is the identity – but in order to keep some generality in (1), we allow any underlying Riemannian structure. Thus we shall study

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\nabla \rho + \rho \nabla V), \qquad t \geq 0, \quad x \in \mathbb{R}^n \text{ or } M. \tag{2}$$

Moreover, we do not address regularity issues, and shall always assume that V is smooth enough, say C^2, perform formal calculations and do not deal with their rigorous justifications in this paper.

As dictated by physical intuition, we mention that the stochastic differential equation underlying (2) is

$$dX_t = dW_t - \nabla V(X_t)\, dt,$$

with W_t a standard Wiener process (or Brownian motion). Thus the Fokker-Planck equation models a set of particles experiencing both diffusion and drift. The interplay between these two processes is at the basis of most of its properties.

2. Trend to equilibrium

Let us begin an elementary study of the Fokker-Planck equation. From (2) we see that there is an obvious stationary state : $\rho = e^{-V}$ (adding a constant to V if necessary, one can always assume that e^{-V} is a probability distribution). It is then natural to change variables by setting $\rho = he^{-V}$. Then we obtain for (2) the equivalent formulation

$$\frac{\partial h}{\partial t} = \Delta h - \nabla V \cdot \nabla h, \qquad t \geq 0, \quad x \in \mathbb{R}^n \text{ or } M. \tag{3}$$

The operator $L = \Delta - \nabla V \cdot \nabla$ is self-adjoint w.r.t. the measure e^{-V}. More precisely,

$$\langle Lh, g\rangle_{e^{-V}} = -\langle \nabla h, \nabla g\rangle_{e^{-V}}. \tag{4}$$

(we use the obvious notation for weighted L^2-scalar products and norms).

In particular,

$$\langle Lh, h\rangle_{e^{-V}} = -\|\nabla h\|^2_{L^2(e^{-V})},$$

so that L is a nonpositive operator, whose kernel consists of constants (since e^{-V} is a positive function). This shows that the only acceptable equilibria for (2) are constant multiples of e^{-V} – the constant being determined by the norm of h in $L^1(e^{-V})$, which is preserved.

Now, consider the Cauchy problem for the Fokker-Planck equation, which is (2) supplemented with an initial condition

$$\rho(t=0,\cdot) = \rho_0; \qquad \rho_0 \geq 0, \quad \int \rho_0 = 1.$$

We expect the solution of the Cauchy problem to converge to the equilibrium state e^{-V}, and would like to estimate the rate of convergence in terms of the initial datum. Let us work with the equivalent formulation (3), with the initial datum $h_0 = \rho_0 e^V$. Since L is a nonpositive self-adjoint operator, we would expect $h(t,\cdot)$ to converge exponentially fast to 1 if L has a spectral gap $\lambda > 0$. This easily follows by elementary spectral analysis, or by noting that the existence of a spectral gap of size λ for L is equivalent to the statement that e^{-V} satisfies a Poincaré inequality with constant λ, i.e

$$\forall\, g \in L^2(e^{-V}), \quad \left[\int g e^{-V}\, dx = 0 \Longrightarrow \int |\nabla g|^2 e^{-V} \geq \lambda \int g^2 e^{-V}\right]. \tag{5}$$

Indeed, knowing (5), and using (3), one can perform the computation

$$-\frac{d}{dt}\int (h-1)^2 e^{-V} = 2\int |\nabla h|^2 e^{-V} \geq 2\lambda \int (h-1)^2 e^{-V}, \tag{6}$$

which entails

$$\int (h-1)^2 e^{-V} \leq e^{-2\lambda t} \int (h_0 - 1)^2 e^{-V}.$$

Thus, if h solves (3) with initial datum h_0,

$$h_0 \in L^2(e^{-V}) \Rightarrow \|h(t,\cdot) - 1\|_{L^2(e^{-V})} \le e^{-\lambda t}\|h_0 - 1\|_{L^2(e^{-V})}.$$

Equivalently, if ρ solves (2) with initial datum ρ_0,

$$\rho_0 \in L^2(e^{V}) \Rightarrow \|\rho(t,\cdot) - e^{-V}\|_{L^2(e^{V})} \le e^{-\lambda t}\|\rho_0 - e^{-V}\|_{L^2(e^{V})}. \tag{7}$$

This approach is fast and effective, but has several drawbacks, which are best understood when one asks whether the method may be generalized :

1) Note that the functional space which is natural at the level of (3) ($h \in L^2(e^{-V})$) is not at all so when translated to the level of (2) ($\rho \in L^2(e^{V})$). For mathematical and physical purposes, it would be desirable to be as close as possible to the space $\rho \in L^1$ (which corresponds to the assumption of finite mass).

2) The physical content of the estimate (7) is quite unclear. Strongly based on the theory of linear operators, this estimate turns out to be very difficult, if not impossible, to generalize to nonlinear diffusion equations (like porous medium equations, or the Fokker-Planck-Landau equations) which arise in many areas of physics, cf. Section 8.

3) Also, it is often quite difficult to find explicit values of the spectral gap of a given linear operator. Many criteria are known, which give existence of a spectral gap, but without estimate on its magnitude the results obtained in this manner are of limited value only.

3. Entropy dissipation

Instead of investigating the decay in $L^2(e^{-V})$ norm for h, we could as well consider a variety of functionals controlling the distance between h and 1. Actually, whenever ϕ is a convex function on $\mathbb{R}$, one can check that

$$\int \phi(h)e^{-V}\,dx = \int \phi\left(\frac{\rho}{e^{-V}}\right) e^{-V}\,dx \tag{8}$$

defines a Lyapunov functional for (3), or equivalently for (2). Indeed,

$$\frac{d}{dt}\int \phi(h)e^{-V}\,dx = -\int \phi''(h)|\nabla h|^2 e^{-V}\,dx. \tag{9}$$

Our previous computation in the L^2 norm corresponds of course to the choice $\phi(h) = (h-1)^2$; but to investigate the decay towards equilibrium, we could also decide to consider any strictly convex, nonnegative function ϕ such that $\phi(1) = 0$.

For several reasons, a very interesting choice is

$$\phi(h) = h \log h - h + 1. \tag{10}$$

Indeed, in this case, taking into account the identity $\int (h-1)e^{-V} = 0$, we find

$$\int \phi(h)e^{-V} = \int \rho \log \frac{\rho}{e^{-V}} = \int \rho(\log \rho + V). \tag{11}$$

This functional is well-known. In kinetic theory it is often called the free energy, while in information theory it is known as the (Kullback) relative entropy of ρ w.r.t. e^{-V} (strictly speaking, of the measure $\rho\, dx$ w.r.t. the measure $e^{-V}\, dx$).

As we evoke in section 8, the occurrence of the relative entropy is rather universal in convection-diffusion problems, linear or nonlinear. So an assumption of boundedness of the entropy (for the initial datum) is satisfactory both from the physical and from the mathematical point of view. We shall denote the relative entropy (11) by $H(\rho|e^{-V})$, which is reminiscent of the standard notation of Boltzmann's entropy.

The relative entropy is an acceptable candidate for controlling the distance between two probability distributions, in view of the elementary inequality

$$H(\rho|\tilde{\rho}) \geq \frac{1}{2}\|\rho - \tilde{\rho}\|_{L^1}^2. \tag{12}$$

Inequality (12) is known as Csiszár-Kullback inequality by (many) analysts, and Pinsker inequality by probabilists (cf. [2] for a detailed account).

By (9), if ρ is a solution of the Fokker-Planck equation (2), then

$$\frac{d}{dt}H(\rho|e^{-V}) = -\int \rho \left|\nabla \left(\log \frac{\rho}{e^{-V}}\right)\right|^2 dx \equiv -I(\rho|e^{-V}). \tag{13}$$

The functional I is known in information theory as the (relative) Fisher information, and in the theory of large particle systems as the "Dirichlet form" (since it can be rewritten as $4\int|\nabla\sqrt{h}|^2e^{-V}$). In a kinetic context, it is simply the entropy dissipation functional associated to the Fokker-Planck equation.

Now, let us see how the computation (13) can help us investigating the trend to equilibrium for (2).

4. Logarithmic Sobolev inequalities

Let $\gamma(x) = (2\pi)^{-n/2}e^{-|x|^2/2}$ denote the standard Gaussian on $\mathbb{R}^n$. The Stam-Gross logarithmic Sobolev inequality [23, 33] asserts that for any probability distribution ρ (absolutely continuous w.r.t. γ),

$$H(\rho|\gamma) \le \frac{1}{2}I(\rho|\gamma). \tag{14}$$

By a simple rescaling, if γ_σ denotes the centered Gaussian with variance σ, $(2\pi\sigma)^{-n/2}e^{-|x|^2/(2\sigma)}$, then

$$H(\rho|\gamma_\sigma) \le \frac{\sigma}{2}I(\rho|\gamma_\sigma).$$

In the study of the trend to equilibrium for (2), this inequality plays precisely the role of (5). It implies that (by (13) with $V(x) = \frac{|x|^2}{2\sigma}$), if ρ is a solution of

$$\frac{\partial\rho}{\partial t} = \nabla\cdot\left(\nabla\rho + \rho\frac{x}{\sigma}\right), \tag{15}$$

then ρ satisfies an estimate of *exponential decay in relative entropy*,

$$H(\rho(t,\cdot)|\gamma_\sigma) \le H(\rho_0|\gamma_\sigma)e^{-2t/\sigma}. \tag{16}$$

Why is (14) called a *logarithmic Sobolev inequality* ? because it can be rewritten

$$\int u^2\log u^2\,d\gamma - \left(\int u^2 d\gamma\right)\log\left(\int u^2 d\gamma\right) \le 2\int|\nabla u|^2\,d\gamma.$$

This asserts the embedding $H^1(d\gamma)\subset L^2\log L^2(d\gamma)$, which is an infinite-dimensional version of the usual Sobolev embedding $\dot{H}^1(\mathbb{R}^n)\subset L^{2*}(\mathbb{R}^n)$.

Now, of course, (16) is a quite limited result, because it concerns a very peculiar case (quadratic confinement potential). All the more that the solution of (15) is explicitly computable ! So it is desirable to understand how all of this can be extended to a more general setting.

By definition, we shall say that the probability measure e^{-V} satisfies a logarithmic Sobolev inequality with constant $\lambda > 0$ if for all probability measures ρ,

$$H(\rho|e^{-V}) \leq \frac{1}{2\lambda} I(\rho|e^{-V}). \tag{17}$$

By a computation completely similar to the previous one, we see that as soon as e^{-V} satisfies a logarithmic Sobolev inequality with constant λ, then the solution of the Fokker-Planck equation (with V as confining potential) goes to equilibrium in relative entropy, with a rate $e^{-2\lambda t}$ at least. So the question is now : *which probability measures satisfy logarithmic Sobolev inequalities ?*

5. The Bakry-Emery reversed point of view

In 1985, Bakry and Emery [3] proved the basic following result, which goes a long way towards the solution of the preceding question.

Theorem 1. *Let e^{-V} be a probability measure on $\mathbb{R}^n$ (resp. a Riemannian manifold M), such that $D^2V \geq \lambda I_n$ (resp. $D^2V + \mathrm{Ric} \geq \lambda I_n$), where I_n is the identity matrix of dimension n (and Ric the Ricci curvature tensor on the manifold M). Then, e^{-V} satisfies a logarithmic Sobolev inequality with constant λ.*

(In the Riemannian case, D^2V stands of course for the Hessian of V.)

Moreover, there is room for perturbation in this theorem, as can be seen from the standard Holley-Stroock perturbation lemma [24]: If V is of the form $V_0 + v$, where $v \in L^\infty$ and e^{-V_0} satisfies a logarithmic Sobolev inequality with constant λ, then also e^{-V} satisfies a logarithmic Sobolev inequality, with constant $\lambda e^{-\mathrm{osc}(v)}$, with osc $(v) = \sup v - \inf v$. By combining the Bakry-Emery theorem with the Holley-Strock lemma, one can generate a lot of probability measures satisfying a logarithmic Sobolev inequality. We also refer to [2] for more general statements with a non-constant diffusion matrix D.

But what is striking above all in the Bakry-Emery theorem, is that its proof is obtained by a complete inversion of the point of view (with respect to our approach). Indeed, while our primary goal was to establish the inequality (17) in order to study the equation (2), they used the equation (2) to establish the inequality (17) ! Here is how the argument works, or rather how we can understand it from a physical point of view, developed in [2] (the original paper of Bakry and Emery takes a rather abstract point of view, based on the so-called Γ_2, or *carré du champ itéré*).

1) Recognize the entropy dissipation (in this case, the relative Fisher information) as the relevant object, and analyse its time evolution. For this purpose, compute (under suitable regularity conditions)

$$J(\rho|e^{-V}) \equiv -\frac{d}{dt}I(\rho|e^{-V}); \tag{18}$$

2) Prove that under the assumptions of Theorem 1, the following functional inequality holds

$$I(\rho|e^{-V}) \leq \frac{1}{2\lambda}J(\rho|e^{-V}), \tag{19}$$

so that the entropy dissipation goes to 0 exponentially fast,

$$I(\rho(t,\cdot)|e^{-V}) \leq e^{-2\lambda t}I(\rho_0|e^{-V}). \tag{20}$$

3) Integrate the identity (20) in time, from 0 to $+\infty$. Noting that $\int_0^{+\infty} e^{-2\lambda t}\,dt = 1/(2\lambda)$ and $\int_0^{+\infty} I(\rho|e^{-V})\,dt = H(\rho_0|e^{-V})$, recover

$$H(\rho_0|e^{-V}) \leq \frac{1}{2\lambda}I(\rho_0|e^{-V}),$$

which was our goal.

4) Perform a density argument to establish the logarithmic Sobolev inequality for all probability densities ρ_0 with bounded entropy dissipation, getting rid of regularity constraints occured in performing the steps 1)–3).

The reader may feel that the difficulty has simply been shifted : why should the functional inequality (19) be simpler to prove than (17) ? It turns out that (19) is quite trivial once the calculation of J has been rearranged in a proper way :

$$\begin{aligned} J(\rho|e^{-V}) &= 2\int \rho \,\mathrm{tr}\, \left(\left(D^2 \log \tfrac{\rho}{e^{-V}} \right)^T \left(D^2 \log \tfrac{\rho}{e^{-V}} \right) \right) \\ &\quad +2\int \rho \left\langle (D^2V + \ \mathrm{Ric})\, \nabla \log\left(\tfrac{\rho}{e^{-V}}\right), \nabla \log\left(\tfrac{\rho}{e^{-V}}\right) \right\rangle . \end{aligned} \tag{21}$$

The first term in (21) is obviously nonnegative without any assumptions, while the second one is bounded below by $2\lambda \int \rho |\nabla \log(\rho/e^{-V})|^2 = 2\lambda I(\rho|e^{-V})$, if λ is a lower bound for $D^2V + \mathrm{Ric}$. Of course, the difficulty is to establish (21) ! As we shall see later, there are simple formal ways towards it. Let us only mention at this stage that the Ricci tensor comes naturally through the Bochner formula,

$$-\nabla u \cdot \nabla \triangle u + \triangle \frac{1}{2} |\nabla u|^2 \ = \ \mathrm{tr}\Big((\mathrm{D}^2 u)^T \mathrm{D}^2 u \Big) + \langle \mathrm{Ric} \cdot \nabla u, \nabla u \rangle.$$

6. Log Sobolev $\Rightarrow$ Poincaré

The reader may wonder what price we had to pay for leaving the $L^2(e^V)$ theory in favor of the more general (and physically more natural) framework of data with finite entropy. It turns out that *we lost nothing* : this is the content of the following simple theorem, due to Rothaus and Simon :

Theorem 2. *Assume that* e^{-V} *satisfies the logarithmic Sobolev inequality* (17) *with constant* λ*. Then* e^{-V} *also satisfies the Poincaré inequality* (5) *with constant* λ*.*

Actually the Poincaré inequality is a linearized version of the logarithmic Sobolev inequality : to see this, it suffices to notice that if g is smooth and satisfies $\int g e^{-V}\, dx = 0$, then as $\varepsilon \to 0$,

$$H\Big((1+\varepsilon g) e^{-V} | e^{-V} \Big) \simeq \frac{\varepsilon^2}{2} \|g\|^2_{L^2(e^{-V})}$$

$$I\Big((1+\varepsilon g) e^{-V} | e^{-V} \Big) \simeq \varepsilon^2 \|\nabla g\|^2_{L^2(e^{-V})}.$$

In [2] a more general study is undertaken. Actually one can define a whole family of relative entropy functionals, of the form (8), whose “extremals” are given by $\phi(h) = h\log h - h + 1$ at one end, $\phi(h) = (h-1)^2/2$ at the other end. For each of these entropies one can perform a Bakry-Emery-type argument to prove logarithmic-Sobolev-type inequalities; and they are all the stronger as the nonlinearity in the relative entropy is weaker (the strongest one corresponding to the $h\log h$ nonlinearity). Corresponding variants of the Holley-Stroock perturbation lemma, and of the Csiszár-Kullback-Pinsker inequality are also established in great generality in [2].

7. The nonuniformly convex case

So far we saw that the combination of the Bakry-Emery theorem and the Holley-Stroock perturbation lemma is enough to treat the case of confining potentials which are uniformly convex ($+L^\infty$-perturbations). What happens if $V(x)$ behaves at infinity like, say, $|x|^\alpha$ with $0 < \alpha < 2$? If $1 \le \alpha < 2$, then there is no logarithmic Sobolev inequality, while a Poincaré inequality still holds. This seems to indicate that the linear approach is better for such situations. This is not the case : a simple way to overcome the absence of logarithmic Sobolev inequality is exposed in [37]. There, modified logarithmic Sobolev inequalities are established, in which the degeneracy of the convexity is compensated by the use of *moments* to “localize” the distribution function. For instance, in the situation we are considering,

$$H(\rho|e^{-V}) \le C I(\rho|e^{-V})^{1-\delta} M_s(\rho)^\delta,$$

where $M_s(\rho)$ is the moment of order s of ρ,

$$M_s(\rho) = \int_{\mathbb{R}^N} \rho(x)(1+|x|^2)^{s/2}\,dx \qquad (s>2)$$

and

$$\delta = \frac{2-\alpha}{2(2-\alpha)+(s-2)} \in (0,1/2).$$

Combining this estimate with a separate study of the time-behavior of moments, one can prove convergence to equilibrium with rate $O(t^{-\infty})$ (this means $O(t^{-\kappa})$ for all κ) if the initial datum is rapidly decreasing. A striking feature of the argument is that it is not at all necessary to prove that the moments stay uniformly bounded : it only suffices to show that their growth is *slow enough*.

8. Generalizations to other physical systems

Let us now give a short review of some physical models for which the methodology of entropy dissipation estimates has enabled a satisfactory solution of the problem of trend to equilibrium :

1) nonlinearly coupled Fokker-Planck equation, like the drift-diffusion-Poisson model (see [2], the references therein and [1], [7], [8]). This is a Fokker-Planck equation $\partial_t \rho = \nabla \cdot (\nabla \rho + \rho \nabla V)$, in which the confining potential is equal to the sum of an external potential (say, quadratic), and a self-consistent potential obtained through Poisson coupling. In other words,

$$V(x) = \frac{|x|^2}{2} + W(x), \qquad -\Delta W = \rho.$$

Once uniform in time L^∞ bounds on V are established, the use of logarithmic Sobolev inequalities leads to the conclusion that solutions of this model converge exponentially fast to equilibrium. For the corresponding bipolar problem we refer to [7].

2) nonlinear diffusion equations of porous-medium or fast diffusion type, with a confining potential, like

$$\frac{\partial \rho}{\partial t} = \Delta \rho^\alpha + \nabla \cdot (\rho x), \qquad t \geq 0, \quad x \in \mathbb{R}^n, \tag{22}$$

where the exponent α satisfies

$$\alpha \geq 1 - \frac{1}{n}.$$

For this model one can establish logarithmic Sobolev-type inequalities of the same type as for the linear Fokker-Planck equation, see [19, 20, 28]. More generally, if one considers the equation

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\nabla P(\rho) + \rho \nabla V(x)), \quad t \geq 0, x \in \mathbb{R}^n \tag{23}$$

under the assumptions

$$D^2V \geq \lambda I_n, \quad \lambda > 0; \qquad \frac{P(\rho)}{\rho^{1-1/n}} \text{ nondecreasing;} \qquad P(\rho) \text{ increasing,}$$

one can prove [17] exponential decay to equilibrium with rate at least $e^{-2\lambda t}$. We shall elaborate on this example in the next section[1]

3) nonlocal diffusion models, like the spatially homogeneous Fokker-Planck-Landau equation in plasma physics. In this model the phase space is a velocity space, and the unknown $f = f(t,v)$ satisfies

$$\frac{\partial f}{\partial t} = \nabla_v \cdot \left(\int_{\mathbb{R}^n} dv_* \, a(v - v_*) \Big[f_* \nabla f - f (\nabla f)_* \Big] \right), \; t \geq 0, \; v \in \mathbb{R}^n \tag{24}$$

where a is a matrix-valued fonction,

$$a(z) = \Psi(|z|)\Pi(z), \qquad \Psi(|z|) \geq 0,$$

and $\Pi(z)$ is the orthogonal projection upon $z^\perp$,

$$\Pi_{ij}(z) = \delta_{ij} - \frac{z_i z_j}{|z|^2}.$$

Standard choices of the function Ψ are the power laws $\Psi(|z|) = |z|^{\gamma+2}$, $-n \leq \gamma \leq 1$. After a careful study of the structure of this equation, and due to the use of the Stam-Gross logarithmic Sobolev inequality, explicit rates of convergence to equilibrium are established in [21] for the case $\gamma > 0$ (like $t^{-2/\gamma}$).

Another nonlocal model linked to the Fokker-Planck equation is

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\nabla \rho + \rho \nabla (W * \rho)),$$

with W strictly convex. This model arises for instance in the context of granular media [6]. A general study will be performed in [18].

4) the spatially homogeneous Boltzmann equation :

$$\frac{\partial f}{\partial t} = Q(f,f) = \int_{\mathbb{R}^n} dv_* \int_{S^{n-1}} d\sigma \, B(v - v_*, \sigma)(f' f'_* - f f_*), \tag{25}$$

[1]**Note added in proof:** equations like (22) naturally arise as rescaled versions of porous-mediun equations, $\partial\rho/\partial t = \Delta\rho^\alpha$. The rescaling method is robust even for more general equations of the type $\partial\rho/\partial t = \Delta P(\rho)$: see recent work by Biler et al.

where $f' = f(v')$ and so on, and

$$\begin{cases} v' = \dfrac{v+v_*}{2} + \dfrac{|v-v_*|}{2}\sigma, \\ \\ v'_* = \dfrac{v+v_*}{2} - \dfrac{|v-v_*|}{2}\sigma. \end{cases} \tag{26}$$

Equations (26) yield a convenient parametrization of the $(n-1)$-dimensional manifold of all the solutions to the equations of elastic collision, where v', v'_* are precollisional velocities and v, v_* postcollisional ones.

Logarithmic-Sobolev-type inequalities do *not* hold for this model; but there are slightly weaker substitutes, which enable to prove convergence to equilibrium with an explicit rate. This was proven in [36], after a careful study of the symmetries of the operator $Q(f, f)$. Let us mention that the proof makes use of two auxiliary diffusion processes : the (non-local) Fokker-Planck-Landau equation, but also the plain Fokker-Planck equation with quadratic confining potential. We refer to [39] for a detailed and elementary review of entropy-dissipation methods in the context of the Boltzmann equation.

5) Finally, we also mention the study of *spatially inhomogeneous kinetic models*, in which both space and velocity variables are introduced, and a collision operator acting only on the velocity variable is coupled with transport and confinement. A typical example is the (kinetic) Fokker-Planck equation

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f - \nabla V(x) \cdot \nabla_v f = \nabla_v \cdot (\nabla_v f + fv), \qquad t \geq 0,\ x \in \mathbb{R}^n,\ v \in \mathbb{R}^n.$$

The fact that the diffusion acts only in the v variable changes drastically the physical situation and makes the study considerably more delicate. We refer to [22] for a physical analysis, complete proofs of trend to equilibrium with an explicit rate (like $O(t^{-\infty})$), and comments. Here we only mention that both the Stam-Gross logarithmic Sobolev inequality (in the v variable) and the logarithmic Sobolev inequality for the potential e^{-V} (in the x variable) are used in the proofs.

9. An example : generalized porous medium equations

In this section we comment on equation (23), and see how a simple approach enables to treat this a priori quite complicated problem.

First of all, there is a variational principle hidden in equation (23). From thermodynamical considerations, it is natural to introduce the "generalized free energy",

$$E_P(\rho) := \int \Big[\rho(V(x) + h(\rho)) - P(\rho)\Big]\, dx,$$

where the enthalpy $h = h(\rho)$ is defined by

$$h(\rho) := \int_1^\rho \frac{P'(\sigma)}{\sigma}\, d\sigma.$$

The free energy is nonincreasing under time-evolution of the equation (23), and one easily computes the dissipation of free energy (or entropy dissipation by abuse of terminology) :

$$I_P(\rho|\rho_\infty) = \int \rho \Big|\nabla(V + h(\rho))\Big|^2\, dx.$$

It is therefore natural to assume that the solution $\rho = \rho(t,x)$ of (23) converges towards the minimizer of E_P, if it is unique. Under the condition of fixed mass $\int \rho = 1$ and nonnegativity, minimizers have to be of the form

$$\rho_\infty(x) := \bar{h}^{-1}(C - V(x)),$$

where C is some normalization constant associated with the mass constraint, and $\bar{h}^{-1}$ is the generalized inverse

$$\bar{h}^{-1}(s) := \begin{cases} 0, & s < h(0+) \\ h^{-1}(s), & h(0+) \le s \le h(\infty) \\ +\infty, & s > h(+\infty) \end{cases}$$

We assume that there is one unique such minimizer (only one possible value of C).

Example : For the case $P(\rho) = \rho^\alpha, V(x) = \frac{|x|^2}{2}$, we compute

$$\rho_\infty(x) = \left(\left(E + \frac{1-\alpha}{2\alpha}|x|^2\right)_+\right)^{\frac{1}{\alpha-1}} \quad \text{for } \alpha \neq 1,$$

with $E > 0$ obtained from normalization in $L^1(\mathbb{R}^n)$. Clearly, $\rho_\infty > 0$ is C^∞ for $1 - \frac{1}{n} < \alpha < 1$ (fast diffusion case) but only Lipschitz continuous with compact support for $1 < \alpha \le 2$ and $C^{0,\frac{1}{\alpha-1}}$ with compact support for $\alpha > 2$ (Barenblatt-Prattle profile, cf. [28], [19], [20]).

We now have all the elements required to perform a study by means of entropy dissipation, generalized logarithmic Sobolev inequalities, and Bakry-Emery arguments. In order to make the analogy more apparent, we work with the relative free energy, or, by abuse of terminology, "relative entropy",

$$H_P(\rho|\rho_\infty) := E_P(\rho) - E_P(\rho_\infty),$$

and we also use the notation $I_P(\rho|\rho_\infty)$ for the entropy dissipation.

A lenghty calculation then shows that the time-derivative of the entropy dissipation, under evolution by (23), is given by the functional

$$J_P(\rho|\rho_\infty) = 2\int \left[\Big(P'(\rho)\rho - P(\rho) \Big) (\nabla \cdot y)^2 + P(\rho)\,\mathrm{tr}\Big((Dy)^T Dy \Big) \right] dx$$
$$+2 \int \rho \, \langle D^2 V \cdot y, y \rangle \, dx$$

with $y := \nabla(V + h(\rho))$, and of course $Dy = D^2(V + h(\rho))$.

As in the case of linear diffusion, the first integral is nonnegative, and the second integral is bounded below by $2\lambda I_P(\rho|\rho_\infty)$. Using the condition "$P(\rho)/\rho^{1-1/n}$ is nondecreasing" and

$$\nabla \cdot y = \mathrm{tr}(Dy), \quad \mathrm{tr}(Z)^2 \le n\,\mathrm{tr}(Z^2)$$

for all symmetric $n \times n$-matrices Z, it is immediate that the first integral is nonnegative. Thus,

$$E_P(\rho(t)|\rho_\infty) \le E_P(\rho_0|\rho_\infty) e^{-2\lambda t}$$

follows by proceeding as in the linear diffusion case. Moreover the Bakry-Emery approach gives the generalized Sobolev-inequality:

$$E_P(\rho|\rho_\infty) \le \frac{1}{2\lambda} I_P(\rho|\rho_\infty)$$

for all probability densities ρ (after a somewhat involved approximation/density argument, cf. [17]). Obviously this inequality is identical to (17) in the linear case $P(\rho) = \rho$.

10. Gaussian isoperimetry

In this and the next sections, we continue our review of links between logarithmic Sobolev inequalities, Fokker-Planck equations and some areas of functional analysis.

It is well-known that the usual Sobolev embedding $W^{1,1}(\mathbb{R}^n) \subset L^{n/(n-1)}(\mathbb{R}^n)$ can be interpreted in geometrical terms as an isoperimetric statement. Now, what about the logarithmic Sobolev inequality ? It turns out that the Stam-Gross logarithmic Sobolev inequality can be viewed as a consequence of the *Gaussian isoperimetry.* The Gaussian isoperimetry [11, 34] states that in Gauss space, for fixed volume, half-spaces have maximal surface. We recall the definition of the surface : if B is a measurable set, define

$$B_t = \Big\{x \in \mathbb{R}^n; \quad d(x,B) \le t\Big\}. \tag{27}$$

Here $d(x,B) = \inf_{y \in B} \|x-y\|_{\mathbb{R}^n}$. The gaussian surface $S(B)$ of B is defined in a natural way as

$$S(B) = \liminf_{t \downarrow 0} \frac{\gamma(B_t) - \gamma(B)}{t}.$$

A functional version of the Gaussian isoperimetry was established by Bobkov [9]; it can be stated as follows. Let $\mathcal{U}$ denote the Gaussian isoperimetric function, i.e. $\mathcal{U} = \varphi \circ \Phi^{-1}$, where $\varphi(x) = (2\pi)^{-1/2} e^{-|x|^2/2}$ is the one-dimensional standard Gaussian, and $\Phi(x) = \int_{-\infty}^{x} \varphi(s)\, ds$. A few moments of reflexion show that $\mathcal{U}(x)$ is the Gaussian surface of the half-space with Gaussian volume x. Then, for all function $h : \mathbb{R}^n \to [0,1]$, the inequality holds

$$\mathcal{U}\left(\int_{\mathbb{R}^n} h\, d\gamma\right) \le \int_{\mathbb{R}^n} \sqrt{\mathcal{U}^2(h) + |\nabla h|^2}\, d\gamma, \tag{28}$$

where γ still denotes the n-dimensional Gauss measure. It is not very difficult to see that (28) is equivalent to the isoperimetric statement. For instance, if in (28) we replace h by (an approximation of) the characteristic function of a set B, we find

$$\mathcal{U}(\gamma(B)) \le S(B),$$

which is precisely the gaussian isoperimetry.

This inequality was generalized by Bakry and Ledoux [4]. They prove that under the assumption $D^2V \geq \lambda I_n$, then

$$\mathcal{U}\left(\int_{\mathbb{R}^n} h\, e^{-V}\, dx\right) \leq \int_{\mathbb{R}^n} \sqrt{\mathcal{U}^2(h) + \frac{1}{\lambda}|\nabla h|^2}\, e^{-V}\, dx. \tag{29}$$

The proof is again based on the Fokker-Planck equation ! Actually, it is a direct consequence of the (non-trivial) observation that the right-hand side of (29) defines a Lyapunov functional under the action of (3).

Now, as was noticed by Beckner, from (29) one can recover (17) by a simple limiting procedure (just like the Poincaré inequality follows from the log Sobolev inequality). Namely, it suffices to replace h by εh, expand it for ε close to 0, and use the fact that, for x close to 0,

$$\mathcal{U}(x) \sim x\sqrt{2\log(1/x)}.$$

11. Talagrand inequalities and concentration of the Gauss measure

Let us now turn to other inequalities with an information content. Most of the material for the remaining sections follows [29].

M being still a given Riemannian manifold, we define the Wasserstein distance between two probability measures by

$$W(\mu,\nu) = \sqrt{\inf_{\pi\in\Pi(\mu,\nu)} \int_{M\times M} d(x,y)^2\, d\pi(x,y)}, \tag{30}$$

where $\Pi(\mu,\nu)$ denotes the set of probability measures on $M\times M$ with marginals μ and ν, i.e. such that for all bounded continuous functions φ and ψ on M,

$$\int_{M\times M} d\pi(x,y)\,\Big[\varphi(x) + \psi(y)\Big] = \int_M \varphi\, d\mu + \int_M \psi\, d\nu. \tag{31}$$

Equivalently,

$$W(\mu,\nu) = \inf\left\{\sqrt{E\, d(X,Y)^2}, \quad \mathrm{law}(X) = \mu, \quad \mathrm{law}(Y) = \nu\right\},$$

where the infimum is taken over arbitrary random variables X and Y on M.

It is a well-known fact that the Wasserstein distance (actually known under many names : Monge-Kantorovich-Fréchet-Höffding-Gini-Tanaka...) metrizes the weak-* topology on $P_2(M)$, the set of probability measures on M with finite second moments. More precisely, if (μ_k) is a sequence of probability measures on M such that for some (and thus any) $x_0 \in M$,

$$\lim_{R\to\infty} \sup_k \int_{d(x_0,x)\geq R} d(x_0,x)^2 \, d\mu_k(x) = 0,$$

then $W(\mu_k,\mu) \longrightarrow 0$ if and only if $\mu_k \longrightarrow \mu$ in weak measure sense.

Developing an idea of Marton [25], Talagrand [35] showed how to use the Wasserstein distance to obtain rather sharp concentration estimates in a Gaussian setting, with a completely elementary method, which works as follows. With γ still denoting the n-dimensional Gauss measure, Talagrand proved the functional inequality

$$W(\mu,\gamma) \leq \sqrt{2H(\mu|\gamma)}. \tag{32}$$

Now, let $B \subset \mathbb{R}^n$ be a measurable set with positive measure $\gamma(B)$, and define B_t as in (27). Moreover, let $\gamma|_B$ denote the restriction of γ to B, i.e. the measure $(1_B/\gamma(B))d\gamma$. A straightforward calculation, using (32) and the triangle inequality for W, yields the estimate

$$W\Big(\gamma|_B\,,\,\gamma|_{\mathbb{R}^n\setminus B_t}\Big) \leq \sqrt{2\log\frac{1}{\gamma(B)}} + \sqrt{2\log\frac{1}{1-\gamma(B_t)}}.$$

Since, obviously, this distance is bounded below by t, this entails

$$\gamma(B_t) \geq 1 - e^{-\frac{1}{2}\left(t-\sqrt{2\log\frac{1}{\gamma(B)}}\right)^2} \quad \text{for } t \geq \sqrt{2\log\frac{1}{\gamma(B)}}. \tag{33}$$

Thus, the measure of B_t goes rapidly to 1 as t grows : this is a standard result in the theory of the concentration of the measure in Gauss space, which can also be derived from the Gaussian isoperimetry.

Other applications of inequalities like (32) can be found in statistics or in statistical mechanics.

12. Log Sobolev $\Rightarrow$ Talagrand $\Rightarrow$ Poincaré

Inequality (32) was generalized in the recent work [29], solving (unwillingly) a conjecture of Bobkov and Götze [10], namely that *a logarithmic Sobolev inequality implies a Talagrand inequality.* More precisely,

Theorem 3. *Assume that e^{-V} satisfies a logarithmic Sobolev inequality with constant λ, eq.* (17). *Then it also satisfies a Talagrand inequality with constant λ, namely*

$$W(\rho, e^{-V}) \le \sqrt{\frac{2H(\rho|e^{-V})}{\lambda}}. \tag{34}$$

As the reader may have guessed, the proof in [29] is based on the Fokker-Planck equation... The basic tool is the following estimate; we denote by $(d/dt)^+$ the right-upper derivative (which is introduced only because the Wasserstein distance is not a priori differentiable).

Proposition 4. *If $\rho(t,x)$ is a (smooth) solution of the Fokker-Planck equation $\partial_t \rho = \nabla \cdot (\nabla\rho + \rho\nabla V)$, with initial datum $\rho(t=0) = \rho_0$, then*

$$\left.\frac{d}{dt}\right|^{+} W(\rho_0, \rho) \le \sqrt{I(\rho|e^{-V})}. \tag{35}$$

Interestingly enough, the proof of Proposition 4 relies on considering (in a somewhat non-natural manner) the linear, diffusive-type Fokker-Planck equation as a linear *transport equation*, i.e of the form

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho\xi) = 0,$$

with a nonlinearly coupled velocity field, $\xi = -\nabla(\log\rho + V)$. Then one solves this equation a posteriori by the methods of characteristics[2]

Theorem 3 follows readily from Proposition 4 and a few technical lemmas (that we skip). We present the proof because it is so short.

[2]**Note added in proof:** this procedure is very much reminiscent of the "diffusion velocity method" in numerical analysis (see works by Mas-Gallic and coworkers).

Assume that e^{-V} satisfies a logarithmic Sobolev inequality, and consider the solution $\rho(t,x)$ of the Fokker-Planck equation starting from some initial datum ρ_0. Then, from (35) and (17),

$$\frac{d}{dt}\bigg|^{+} W(\rho_0,\rho) \le \frac{I(\rho|e^{-V})}{\sqrt{2\lambda H(\rho|e^{-V})}}. \tag{36}$$

But, applying (13), we see that

$$\frac{I(\rho|e^{-V})}{\sqrt{2\lambda H(\rho|e^{-V})}} = -\frac{d}{dt}\sqrt{\frac{2H(\rho|e^{-V})}{\lambda}}. \tag{37}$$

Thus,

$$\frac{d}{dt}\bigg|^{+} \varphi(t) \equiv \frac{d}{dt}\bigg|^{+} \left[W(\rho_0,\rho) + \sqrt{\frac{2H(\rho|e^{-V})}{\lambda}}\right] \le 0.$$

In particular $W(\rho_0, e^{-V}) = \varphi(+\infty) \le \varphi(0) = \sqrt{\frac{2H(\rho_0|e^{-V})}{\lambda}}$ since $\rho(0,\cdot) = \rho_0$ and $\rho(t,\cdot) \to e^{-V}$ as $t \to +\infty$. Disregarding rigorous justification here, we just proved (34).

It is interesting to note that the inequality (34) is still stronger than the Poincaré inequality. Namely, as shown in [29],

Theorem 5. *Assume that e^{-V} satisfies the Talagrand inequality* (34) *with constant λ. Then e^{-V} also satisfies the Poincaré inequality* (5) *with constant λ.*

We remark that the Talagrand inequality (34) allows in a canonical way to generalize the concentration estimate (33) to probability densities e^{-V} which satisfy a logarithmic Sobolev inequality with constant λ (e.g. log-concave densities with log-concavity constant $-\lambda < 0$). Proceeding as in Section 11 one obtains the inequality (already proven by Bobkov and Götze [10])

$$e^{-V}(B_t) \ge 1 - \exp\left(-\frac{\lambda}{2}\left(t - \sqrt{\frac{2}{\lambda}\log\frac{1}{e^{-V}(B)}}\right)^2\right), \quad t \ge \sqrt{\frac{2}{\lambda}\log\frac{1}{e^{-V}(B)}}$$

for all measurable sets $B \subseteq \mathbb{R}^n$ with positive measure $e^{-V}(B)$.

13. Related PDE's

Before going further, let us introduce a few new tools. The theory of the Wasserstein distance is related to other famous PDE's. For the following we consider the n-dimensional Euclidean case.

13.1. The Monge-Ampère equation. One can prove (see for instance [13, 26, 31] and the references therein) that if the measure μ is absolutely continuous with density ρ, then the optimal π in the (Monge-Kantorovich) minimization problem (30) has to be of the form

$$d\pi(x,y) = \rho(x)\,dx\,\delta(y - \nabla\varphi(x)), \tag{38}$$

where φ is a convex function. If we insert (38) into (31), and assume $d\nu = \widetilde{\rho}(x)\,dx$, we find $\int \rho(x)\psi(\nabla\varphi(x))dx = \int \psi(x)\widetilde{\rho}(x)dx$ for all bounded and continuous test-functions ψ, which means that φ is a weak solution (actually solution in Brenier sense) of the Monge-Ampère equation

$$\rho(x) = \widetilde{\rho}(\nabla\varphi(x))\det D^2\varphi(x). \tag{39}$$

See Caffarelli [16, 15], and Urbas [38] for a study of regularity properties : in particular, one can prove $C^{2,\alpha}$ smoothness of φ (for some $\alpha \in (0,1)$) if ρ and $\widetilde{\rho}$ are $C^{0,\alpha}$ and positive everywhere.

13.2. The Hamilton-Jacobi equation. Starting from the relation (39), a natural interpolation $(\rho_t)_{0\le t\le 1}$ between two probability distributions ρ and $\widetilde{\rho}$, introduced by McCann [27], is given by

$$\rho(x) = \rho_t\Big((1-t)x + t\nabla\varphi(x)\Big)\det\Big((1-t)I_n + tD^2\varphi(x)\Big), \tag{40}$$

so that $\rho_0 = \rho$, $\rho_1 = \widetilde{\rho}$.

Actually, equation (40) is a Lagrangian way of interpolating between ρ and $\widetilde{\rho}$: from a physical viewpoint, it means that the mass of ρ is transported onto the mass of $\widetilde{\rho}$, with all particles describing straight lines. The equivalent Eulerian point of view is given by a transport equation coupled with a Hamilton-Jacobi equation :

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \nabla \cdot (\rho \nabla u) = 0, \qquad 0 \le t \le 1, \\ \\ \dfrac{\partial u}{\partial t} + \dfrac{1}{2} |\nabla u|^2 = 0, \end{cases} \tag{41}$$

supplemented with the initial condition $\rho(t=0) = \rho_0$, $u_0(x) = \varphi(x) - |x|^2/2$. A simple calculation gives $\rho(t=1) = \tilde{\rho}$. For this approach consult in particular [5].

It should be noted that this procedure still yields the "right" interpolation equation when considering the Wasserstein distance on a Riemannian manifold (see [29, 30]).

13.3. The sticky particles system. Formally (in $\mathbb{R}^n$), the system (41) is equivalent to the system of sticky particles (pressureless gas dynamics)

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0, \quad 0 < t \le 1, \\ \\ \dfrac{\partial (\rho v)}{\partial t} + \nabla \cdot (\rho v \otimes v) = 0; \end{cases} \tag{42}$$

with $v = \nabla u$ (see [14, 12] and the references included).

14. HWI inequalities

The interpolation (40), or equivalently (41), applied with $\widetilde{\rho} = e^{-V}$, is used in [29] to establish a partial converse to Theorem 3. In particular, it is proven that if e^{-V} satisfies a Talagrand inequality with constant λ, and V is convex, then e^{-V} also satisfies a logarithmic Sobolev inequality, but with constant $\lambda/2$.

This converse statement actually arises as a consequence of what is called HWI inequalities in [29]. These inequalities mix the relative entropy, the Wasserstein distance, and the relative Fisher information; a particular case of it was first proven in [28], and used in the study of porous-medium type equations. Let us give a general statement :

Theorem 6. *Let V be a confining potential on $\mathbb{R}^n$, satisfying $D^2V \geq \lambda I_n$ (λ is not necessarily nonnegative). As usual, assume that e^{-V} is a probability distribution. Then, for any two probability distributions ρ_0 and ρ_1, the HWI inequality holds*

$$H(\rho_0|e^{-V}) \leq H(\rho_1|e^{-V}) + W(\rho_0,\rho_1)\sqrt{I(\rho_0|e^{-V})} - \frac{\lambda}{2}W(\rho_0,\rho_1)^2. \tag{43}$$

Inequality (43) is quite powerful. Note in particular that if $\lambda > 0$, the choice $\rho_0 = e^{-V}$ yields

$$W(\rho_1, e^{-V}) \leq \sqrt{\frac{2H(\rho_1|e^{-V})}{\lambda}},$$

which is essentially what one obtains by combining the Bakry-Emery Theorem with Theorem 3. On the other hand, the choice $\rho_1 = e^{-V}$ yields

$$H(\rho_0|e^{-V}) \leq W(\rho_0, e^{-V})\sqrt{I(\rho_0|e^{-V})} - \frac{\lambda}{2}W(\rho_0, e^{-V})^2. \tag{44}$$

Then, if $\lambda > 0$,

$$W(\rho_0, e^{-V})\sqrt{I(\rho_0|e^{-V})} \leq \frac{\lambda}{2}W(\rho_0, e^{-V})^2 + \frac{1}{2\lambda}I(\rho_0|e^{-V}),$$

so that (44) implies

$$H(\rho_0|e^{-V}) \leq \frac{1}{2\lambda}I(\rho_0|e^{-V}),$$

and we recover the Bakry-Emery theorem !

15. Displacement convexity

Even if the Wasserstein distance is not present in the formulation of logarithmic Sobolev inequalities, it generates a very appealing way to understand them, through the concept of displacement convexity, introduced by McCann [27] and developed in [28, 29]. We still use the notation $(\rho_t)_{0\leq t\leq 1}$ for the interpolant defined in subsection 13.2, between two probability distributions ρ_0 and ρ_1. By definition, a functional J is displacement convex if $t \mapsto J(\rho_t)$ is convex on $[0,1]$, for all probability distributions ρ_0 and ρ_1. We also say that J is uniformly displacement convex, with constant λ, if

$$\frac{d^2}{dt^2}J(\rho_t) \geq \lambda\, W(\rho_0,\rho_1)^2, \qquad 0 < t < 1.$$

Let us give some examples of displacement convex functionals; here the probability distributions are defined on $\mathbb{R}^n$:

$$V(\rho) = \int \rho(x)V(x)\,dx, \qquad V \text{ convex;}$$

$$W(\rho,\rho) = \int \rho(x)\rho(y)W(x-y)\,dx\,dy, \qquad W \text{ convex;}$$

$$U(\rho) = \int A(\rho), \qquad P(\rho)/\rho^{1-1/n} \text{ nondecreasing}, \quad A(0) = 0,$$

where $P(\rho) = \rho A'(\rho) - A(\rho)$. What is more, in the first example above, if the potential V is uniformly convex with constant λ, then also $\rho \mapsto V(\rho)$ is uniformly displacement convex with the same constant. As a consequence, the relative entropy, $H(\rho|e^{-V}) = \int \rho \log \rho + \int \rho V$ also defines a (uniformly) displacement convex functional. As for $W(\rho,\rho)$, see [18].

These examples can also be translated to a Riemannian setting, under some assumptions on the Ricci curvature of the manifold (again, coming naturally through the Bochner formula). For instance, $\rho \mapsto \int \rho V$ is displacement convex if $D^2V + \mathrm{Ric} \geq 0$.

Like standard convexity, the notion of displacement convexity can be formulated in a differentiable manner. This construction was performed in Otto [28] : a formal Riemannian structure is introduced on the set of probability measures, in such a way that displacement convexity (resp. uniform displacement convexity) of the functional J is equivalent to the nonnegativity (resp. the uniform positivity) of the Hessian of J. We refer to [28], or to [29, section 3] for details. We only mention that the formal calculus exposed in these references gives a very efficient way to perform computations like the one which yields (21).

Given a functional E on the set of probability measures, Otto's formal calculus enables to define in a natural way a functional $|\mathrm{grad}E|^2$, by

$$|\mathrm{grad}E|^2(\rho) = \int \rho|\nabla\Phi|^2, \tag{45}$$

where $\Phi[= \delta E/\delta\rho]$ is the solution of

$$\int \Phi g = E'(\rho)\cdot g \qquad \text{[functional derivative].}$$

In the case when $E(\rho) = \int \rho \log \rho + \int \rho V$, i.e $E(\rho) = H(\rho|e^{-V})$, then one easily checks that $|\mathrm{grad} E|^2(\rho) = I(\rho|e^{-V})$. Assuming now that $D^2V \geq \lambda I_n$, the Bakry-Emery result can be rewritten as

$$|\mathrm{grad} H(\cdot|e^{-V})|^2 \geq 2\lambda H(\cdot|e^{-V}).$$

But, formally, this inequality is a very easy consequence of the fact that the relative entropy is uniformly displacement convex with constant λ.

Similar considerations also apply to the inequalities of sections 12 and 14; we refer to [29] for a detailed exposition. The study of displacement convexity is still at an early stage; more applications of this notion are to be expected.

References

[1] Arnold, A.; Markowich, P. A.; Toscani, G., *On large time asymptotics for drift diffusion Poisson systems*, To appear in Transp. Theo. Stat. Phys.

[2] Arnold, A.; Markowich, P.; Toscani, G.; Unterreiter, A., *On logarithmic Sobolev inequalities, Csiszár-Kullback inequalities, and the rate of convergence to equilibrium for Fokker-Planck type equations*, Preprint, 1998.

[3] Bakry, D.; Emery, M., *Diffusions hypercontractives*, Sém. Proba. XIX, no. 1123 in Lect. Notes in Math. Springer, 1985, pp. 177–206.

[4] Bakry, D.; Ledoux, M., *Lévy-Gromov's isoperimetric inequality for an infinite-dimensional diffusion generator*, Invent. Math. 123, 2 (1996), 259–281.

[5] Benamou, J.-D.; Brenier, Y., *A numerical method for the optimal time-continuous mass transport problem and related problems*, In Monge Ampère equation: applications to geometry and optimization (Deerfield Beach, FL, 1997). Amer. Math. Soc., Providence, RI, 1999, pp. 1–11.

[6] Benedetto, D.; Caglioti, E.; Carrillo, J. A.; Pulvirenti, M., *A non-Maxwellian steady distribution for one-dimensional granular media*, J. Statist. Phys. 91, 5-6 (1998), 979–990.

[7] Biler, P.; Dolbeault, J.; Markowich, P. A., *Large time asymptotics of nonlinear drift diffusion systems*, Submitted (1999).

[8] Biler, P.; Dolbeault, J. *Long time behavior of solutions to Nernst-Planck and Debye-Hückel drift-diffusion systems*, Preprint CEREMADE no. 9915 (1999).

[9] Bobkov, S., *An isoperimetric inequality on the discrete cube and an elementary proof of the isoperimetric inequality in Gauss space*, Ann. Probab., 25 (1997), 206–214.

[10] Bobkov, S.; Götze, F., *Exponential integrability and transportation cost related to logarithmic Sobolev inequalities*, J. Funct. Anal. 163, 1 (1999), 1–28.

[11] Borell, C., *The Brunn-Minkowski inequality in Gauss space*, Invent. Math. 30, 2 (1975), 207–216.

[12] Bouchut, F.; James, F., *Solutions en dualité pour les gaz sans pression*, C. R. Acad. Sci. Paris Sér. I Math. 326, 9 (1998), 1073–1078.

[13] Brenier, Y., *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. 44, 4 (1991), 375–417.

[14] Brenier, Y.; Grenier, E., *Sticky particles and scalar conservation laws*, SIAM J. Numer. Anal. 35, 6 (1998), 2317–2328 (electronic).

[15] Caffarelli, L. A., *Boundary regularity of maps with convex potentials*, Comm. Pure Appl. Math. 45, 9 (1992), 1141–1151.

[16] Caffarelli, L. A., *The regularity of mappings with a convex potential*, J. Amer. Math. Soc. 5, 1 (1992), 99–104.

[17] Carrillo, J.; Jüngel, A.; Markowich, P. A.; Toscani, G.; Unterreiter, A., *Entropy dissipation methods for degenerate parabolic systems and generalized Sobolev inequalities*, Preprint (1999).

[18] Carrillo, J.; McCann, R.; Villani, C., Work in preparation.

[19] Carrillo, J.; Toscani, G., *Asymptotic L^1-decay of solutions of the porous medium equation to self-similarity*, To appear in Indiana Math. J., 1999.

[20] Del Pino, M.; Dolbeault, J., *Generalized Sobolev inequalities and asymptotic behaviour in fast diffusion and porous medium problems*, Preprint of the University Paris IX-Dauphine, CEREMADE, n. 9905, 1999.

[21] Desvillettes, L.; Villani, C., *On the spatially homogeneous Landau equation with hard potentials. Part II : H-Theorem and applications*, Comm. P.D.E. 25, 1-2 (2000), 261-298.

[22] Desvillettes, L.; Villani, C., *On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems. Part I : The linear Fokker-Planck equation*, To appear in Comm. Pure Appl. Math.

[23] Gross, L., *Logarithmic Sobolev inequalities*, Amer. J. Math. 97 (1975), 1061–1083.

[24] Holley, R.; Stroock, D., *Logarithmic Sobolev inequalities and stochastic Ising models*, J. Stat. Phys. 46, 5–6 (1987), 1159–1194.

[25] Marton, K., *A measure concentration inequality for contracting Markov chains*, Geom. Funct. Anal. 6 (1996), 556–571.

[26] McCann, R. J., *Existence and uniqueness of monotone measure-preserving maps*, Duke Math. J. 80, 2 (1995), 309–323.

[27] McCann, R. J., *A convexity principle for interacting gases*, Adv. Math. 128, 1 (1997), 153–179.

[28] Otto, F., *The geometry of dissipative evolution equations: the porous medium equation*, To appear in Comm. P.D.E.

[29] Otto, F.; Villani, C., *Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality*, J. Funct. Anal. 173 (2) (2000), 361-400.

[30] Otto, F.; Villani, C., *Differential tools for the optimal mass transportation problem on a Riemannian manifold*, Work in preparation.

[31] Rachev, S.; Rüschendorf, L., *Mass Transportation Problems*, Probability and its applications. Springer-Verlag, 1998.

[32] Risken, H., *The Fokker-Planck equation*, second ed. Springer-Verlag, Berlin, 1989. Methods of solution and applications.

[33] Stam, A., *Some inequalities satisfied by the quantities of information of Fisher and Shannon*, Inform. Control 2 (1959), 101–112.

[34] Sudakov, V. N.; Cirel'son, B. S., *Extremal properties of half-spaces for spherically invariant measures*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 41 (1974), 14–24, 165. Problems in the theory of probability distributions, II.

[35] Talagrand, M., *Transportation cost for Gaussian and other product measures*, Geom. Funct. Anal. 6, 3 (1996), 587–600.

[36] Toscani, G.; Villani, C., *Sharp entropy dissipation bounds and explicit rate of trend to equilibrium for the spatially homogeneous Boltzmann equation*, Comm. Math. Phys. 203, 3 (1999), 667–706.

[37] Toscani, G.; Villani, C., *On the trend to equilibrium for some dissipative systems with slowly increasing a priori bounds*, J. Stat. Phys. 98, 5-6 (2000), 1279-1309

[38] Urbas, J., *On the second boundary value problem for equations of Monge-Ampère type*, J. Reine Angew. Math. 487 (1997), 115–124.

[39] Villani, C., *On the trend to equilibrium for solutions of the Boltzmann equation : quantitative versions of Boltzmann's H-theorem*, Proceedings of the Applied mathematics seminar of the Institut Henri Poincaré. Intented for publication in the Research Notes in Mathematics series, Pitman.

Universität Wien
Institut für Mathematik
AUSTRIA
e-mail: Peter.Markowich@univie.ac.at

École Normale Supérieure, DMI
45, rue d'Ulm 75230 Paris Cedex 05
FRANCE
e-mail: cvillani@umpa.ens-lyon.fr