

FINITENESS OF ENTROPY FOR GRANULAR MEDIA EQUATIONS

BY

JULIAN TUGAUT* (SAINT-ÉTIENNE)

Abstract. The current work deals with the granular media equation whose probabilistic interpretation is the McKean–Vlasov diffusion. It is well known that the Laplacian provides a regularization of the solution. Indeed, for any $t > 0$, the solution is absolutely continuous with respect to the Lebesgue measure. It has also been proved that all the moments are bounded for positive t . However, the finiteness of the entropy of the solution is a new result which will be presented here.

2010 AMS Mathematics Subject Classification: Primary: 35K55; Secondary: 60J60, 60E15.

Key words and phrases: Granular media equation, McKean–Vlasov diffusion, Wasserstein distance, functional inequalities, entropy.

1. INTRODUCTION

Our aim is to show that the entropy of the solution of the granular media equation is finite provided that t is positive.

Indeed, several results are proven under the assumption that the initial entropy is finite. For example, in [13] and [14], we prove the long-time convergence under the following three assumptions:

- the finiteness of some moment,
- the fact that μ_0 is absolutely continuous with respect to the Lebesgue measure,
- the finiteness of the entropy of μ_0 .

The first assumption is necessary to prove the existence of a solution to the self-stabilizing diffusion (whose law is the solution of the granular media equation) so that it cannot be relaxed.

The second assumption may be suppressed. Indeed, we know – see [10] and [11] – that the law at time $t > 0$ is absolutely continuous with respect to the Lebesgue measure.

* I would like to thank Florent Malrieu for his article which led me to simplify the writing of this paper. I would also like to thank the anonymous referee for his precious remarks.

However, to obtain the convergence, we need the initial free energy to be finite so that the initial entropy has to be finite.

In [2], the authors establish a convergence in the Wasserstein distance and the rate of convergence. Nonetheless, they assume the finiteness of entropy of the initial law μ_0 .

The results in [2] have been used in [4] to establish a creation of chaos and a uniform propagation of chaos without global convexity properties. However, the authors need to apply the results in [2] to μ_0 which are discrete probability measures. Consequently, we had to adapt the results and were not able to apply them directly.

This stresses the importance of the finiteness of the entropy for granular media equations.

Such a result has been obtained in [1] and claimed in [12] in the case of a linear partial differential equation which corresponds to a time reversible diffusion.

Now, we present the model. Let us consider $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ a probability measure and X_0 a random variable on \mathbb{R}^d the law of which is μ_0 . We look at the diffusion

$$(1.1) \quad X_t = X_0 + \sigma W_t - \int_0^t \nabla V(X_s) ds - \int_0^t (\nabla F * \mathcal{L}(X_s))(X_s) ds,$$

V and F being two potentials on \mathbb{R}^d , $(W_t)_{t \geq 0}$ being a Brownian motion, and $*$ being used to denote the convolution.

By $\mu_t := \mathcal{L}(X_t)$ we denote the law of the so-called McKean–Vlasov diffusion X , which is the solution of equation (1.1). We know from [11] and [10] that for any $t > 0$, $\mu_t(dx) = u(t, x)dx$. Moreover, the family $\{u(t, x) : t > 0, x \in \mathbb{R}^d\}$ satisfies a nonlinear partial differential equation, the granular media one:

$$(1.2) \quad \frac{\partial u}{\partial t} = \nabla \cdot \left\{ \frac{\sigma^2}{2} \nabla u + u (\nabla V + \nabla F * u) \right\}.$$

We consider the semigroup $(P_t)_{t \geq 0}$ defined by $P_t f(x) := \mathbb{E}_x[f(X_t)]$. The semigroup is associated with the following generator L_t :

$$(1.3) \quad L_t := \frac{\sigma^2}{2} \Delta - (\nabla V + \nabla F * u(t, \cdot)) \cdot \nabla.$$

This generator depends on the time so the semigroup may not be time-reversible.

Let us now formulate the assumptions needed in the paper. We make the same assumptions as the ones in [14].

(A1) *The potential V is a smooth function.*

(A2) *There exists a compact subset \mathcal{K} of \mathbb{R}^d such that $\nabla^2 V(x) > 0$ for all $x \notin \mathcal{K}$. Moreover, $\lim_{\|x\| \rightarrow +\infty} \nabla^2 V(x) = +\infty$.*

(A3) The gradient ∇V is slowly increasing: there exist $m \in \mathbb{N}$, $C > 0$ and a function \mathcal{R} from \mathbb{R}^d to \mathbb{R}^d such that

$$\nabla V(x) = C\|x\|^{2m-2}x + \mathcal{R}(x)$$

for all $x \in \mathbb{R}^d$. Here, the function \mathcal{R} satisfies $\lim_{\|x\| \rightarrow +\infty} \mathcal{R}(x)\|x\|^{-(2m-1)} = 0$.

(A4) There exists an even polynomial function G on \mathbb{R} such that $F(x) = G(\|x\|)$. Also, $\deg(G) =: 2n \geq 2$.

(A5) The function G is convex.

Typically, V is a polynomial function, like $V(x) = x^4/4 - x^2/2$ (in dimension one) and $F(x) = \frac{\alpha}{2}x^2$. Let us point out that the convexity of G is not necessary in this work.

We need to make another assumption:

(A6) The moment of order $8q^2$ of the law μ_0 is finite:

$$\int_{\mathbb{R}^d} x^{8q^2} \mu_0(dx) < \infty, \quad \text{where } q := \max\{m, n\}.$$

Under assumptions (A1)–(A6), we know from Theorem 2.13 in [5] that there exists a unique strong solution X on \mathbb{R}_+ to equation (1.1). Moreover, we have the following uniform boundedness of the moments:

$$\max_{1 \leq j \leq 8q^2} \sup_{t \geq 0} \mathbb{E}[\|X_t\|^j] \leq M_{8q^2} < +\infty.$$

Furthermore (see Proposition A.2 in [13]), for any $k \in \mathbb{N}$ and for any $t_0 > 0$, the quantity $\sup_{t \geq t_0} \mathbb{E}[\|X_t\|^k]$ is finite.

According to [16], there exists an invariant probability μ^σ . Let us note that we may not have uniqueness of this invariant probability, see [6], [16], [15].

Here, our aim is to prove the finiteness of the quantity

$$\int \mu_t \log(\mu_t)$$

for any $t > 0$. In [1], the authors have obtained the finiteness of the relative entropy with respect to the unique invariant probability for a linear diffusion, without assuming any convexity properties.

First, we give the main result – Theorem A – that is to say, the finiteness of the entropy for t positive. Then, we provide two immediate corollaries: one about the simple convergence (Corollary C which comes from [13]) and the other about the convergence in the Wasserstein distance (Corollary D which comes from [2]).

In the last section, we prove Theorem A.

2. MAIN RESULTS

We first give the main result of our current work.

THEOREM A. *Under assumptions (A1)–(A6), for any $t > 0$, we have the finiteness of the entropy of the law $\mu_t = \mathcal{L}(X_t)$. In other words, for any $t > 0$, $\mathcal{L}(X_t)$ is absolutely continuous with respect to the Lebesgue measure and its density f_t satisfies the inequality*

$$\int_{\mathbb{R}^d} f_t(x) \log[f_t(x)] dx < +\infty.$$

REMARK B. *The convexity of the function G is not necessary. In fact, the result still holds for any inhomogeneous diffusion whose diffusion coefficient is constant and whose drift has the form $\nabla V + \nabla_x \mathcal{F}(x, \mu_t)$.*

By Theorem A in [14], we deduce the following assertion.

COROLLARY C. *If assumptions (A1)–(A6) are satisfied and if the set of invariant probabilities is discrete (see [6]–[8] for assumptions such that there are exactly three invariant probabilities), we have the weak convergence in long-time of μ_t toward an invariant probability μ^σ .*

By the results in [2], Theorem A implies the following statement.

COROLLARY D. *If assumptions (A1)–(A6) are satisfied and if V is strictly convex (but not necessarily uniformly strictly convex), then μ_t converges, in the Wasserstein distance, toward the unique invariant probability. Moreover, the rate of convergence is exponential.*

3. PROOF OF THEOREM A

First of all, provided that $t > 0$, μ_t is absolutely continuous with respect to the Lebesgue measure so that there exists f_t such that

- $f_t \geq 0$,
- $\int_{\mathbb{R}^d} f_t(x) dx = 1$,
- $\mu_t(dx) = f_t(x) dx$.

For the moment, nothing ensures us that

$$\int_{\mathbb{R}^d} f_t(x) \log(f_t(x)) dx < +\infty.$$

Let g_0 be a nonnegative function with integral equal to one. We put $g_t := P_t g_0$, where the semigroup $(P_t)_{t \geq 0}$ is generated by

$$(3.1) \quad L_t = \frac{\sigma^2}{2} \Delta - (\nabla V + \nabla F * \mu_t) \cdot \nabla.$$

It is sufficient to show that $\int g_t \log(g_t) < +\infty$ for any $t > 0$.

Let us recall Proposition 2.1 in [16]:

LEMMA 3.1. *For any $\sigma > 0$, there exists an invariant probability μ^σ for diffusion (1.1).*

We can also find a proof of this statement in [2].

We will consider the relative entropy with respect to μ^σ :

$$(3.2) \quad H(\nu | \mu^\sigma) = \int \frac{f}{\mu^\sigma} \log \left(\frac{f}{\mu^\sigma} \right) \mu^\sigma(dx),$$

where $\nu(dx) := f(x)dx$. Indeed, we will have the time reversibility by starting from μ^σ .

Let us remind the reader that the probability measure μ^σ (see [6] and [16]) satisfies

$$(3.3) \quad \mu^\sigma(dx) = \frac{\exp \left[-\frac{2}{\sigma^2} (V(x) + F * \mu^\sigma(x)) \right]}{\int_{\mathbb{R}^d} \exp \left[-\frac{2}{\sigma^2} (V(y) + F * \mu^\sigma(y)) \right] dy} dx.$$

Consequently, we get

$$(3.4) \quad H(\nu | \mu^\sigma) = \int_{\mathbb{R}^d} f(x) \log(f(x)) dx + \frac{2}{\sigma^2} \int_{\mathbb{R}^d} (V(x) + F * \mu^\sigma(x)) f(x) dx \\ + \log \left\{ \int_{\mathbb{R}^d} \exp \left[-\frac{2}{\sigma^2} (V(y) + F * \mu^\sigma(y)) \right] dy \right\}.$$

However, according to Theorem 2.13 in [5], we have the uniform boundedness of the moments from 1 to $8q^2$, where, roughly speaking, $2q$ is defined as the maximum of the degrees of V and F .

We thus infer that the quantity

$$\left| \int_{\mathbb{R}^d} (V(x) + F * \mu^\sigma(x)) g_t(x) dx \right|$$

is bounded by a constant C_0 .

Moreover, $F * \mu^\sigma$ is a polynomial function with parameters which only depend on σ (through the moments of the fixed measure μ^σ). We also know that it is convex. Since V has polynomial behavior and is convex at infinity, the function $x \mapsto \exp \left[-\frac{2}{\sigma^2} (V(x) + F * \mu^\sigma(x)) \right]$ is integrable with respect to the Lebesgue measure. As a consequence, the quantity

$$\log \left\{ \int_{\mathbb{R}^d} \exp \left[-\frac{2}{\sigma^2} (V(y) + F * \mu^\sigma(y)) \right] dy \right\}$$

is finite. Since σ is fixed, it is bounded with respect to the time.

We deduce that it is sufficient to prove the finiteness of $H(P_t g | \mu^\sigma)$.

LEMMA 3.2. *Let X_0 be a random variable which follows the law μ^σ . Then, for any $t \geq 0$, for any functions f and g , we have*

$$(3.5) \quad \mathbb{E}[f(X_t)g(X_0)] = \mathbb{E}[f(X_0)g(X_t)].$$

Proof. By definition, we have

$$X_t = X_0 + \sigma W_t - \int_0^t (\nabla V + \nabla F * \mu_s)(X_s) ds.$$

However, since $\mu_0 = \mu^\sigma$, we deduce that $\mu_s = \mu^\sigma$ for any $s \geq 0$. Then

$$X_t = X_0 + \sigma W_t - \int_0^t (\nabla V + \nabla F * \mu^\sigma)(X_s) ds.$$

Consequently, X is a Kolomogorov diffusion and thus we have the time reversibility (3.5). ■

We put $f_0 := g_0/\mu^\sigma$. We will work with f_0 and $(f_t)_{t \geq 0}$. We proceed like in [1]. We have

$$(3.6) \quad H(\nu_t | \mu^\sigma) = \int_{\mathbb{R}^d} P_t f \log(P_t f) \mu^\sigma,$$

where $\nu_t(dx) = f_t(x)\mu^\sigma(dx)$. We apply equality (3.5) and obtain

$$(3.7) \quad H(\nu_t | \mu^\sigma) = \int_{\mathbb{R}^d} f P_t \log(P_t f) \mu^\sigma.$$

We will now bound $P_t \log(P_t f)$ by $\log(P_{2t} f)$.

Let x and y be in \mathbb{R}^d . We set $x(s) := y + \frac{s}{t}(x - y)$ for any $s \in [0, t]$. We also consider a function h from $[0, t]$ to $[0, t]$ which is \mathcal{C}^1 -continuous such that $h(0) = 0$ and $h(t) = t$.

We consider the trajectory $\gamma(s) := x(h(s))$. We remark that $\gamma(0) = y$ and $\gamma(t) = x$. This function γ plays the role of a geodesic between x and y with respect to the Riemannian metric of the diffusion.

We now introduce

$$(3.8) \quad \xi(s) := (P_s \log(P_{2t-s} f))(\gamma(s)).$$

LEMMA 3.3. *We have the following derivative:*

$$\frac{d\xi}{ds} = -P_s \frac{|\nabla P_{2t-s} f|^2}{(P_{2t-s} f)^2}(\gamma(s)) + \frac{h'(s)}{t} \langle \nabla P_s \log(P_{2t-s} f)(\gamma(s)); x - y \rangle.$$

P r o o f. We put $g := P_{2t-s}f$. Thus, we have

$$\begin{aligned}\xi'(s) = & P_s (L_s \log g) (\gamma(s)) - P_s \frac{L_s g}{g} (\gamma(s)) \\ & + \langle \nabla P_s (\log g) (\gamma(s)) ; \gamma'(s) \rangle.\end{aligned}$$

By the diffusion property, we have

$$L_s \log g = \frac{1}{g} L_s g - \frac{1}{g^2} \Gamma(g, g).$$

Here, $\Gamma(f, g)$ is equal to $\langle \nabla f ; \nabla g \rangle$. Consequently, we obtain

$$\begin{aligned}\xi'(s) = & -P_s \frac{\Gamma(g, g)}{g^2} (\gamma(s)) + \langle \gamma'(s) ; \nabla P_s (\log g) (\gamma(s)) \rangle \\ = & -P_s \frac{|\nabla P_{2t-s}f|^2}{(P_{2t-s}f)^2} (\gamma(s)) + \frac{h'(s)}{t} \langle \nabla P_s \log (P_{2t-s}f) (\gamma(s)) ; x - y \rangle,\end{aligned}$$

which completes the proof. ■

We have the immediate upper bound:

$$(3.9) \quad \xi'(s) \leq -P_s \frac{|\nabla P_{2t-s}f|^2}{(P_{2t-s}f)^2} (\gamma(s)) + \frac{|h'(s)|}{t} |x - y| |\nabla P_s \log (P_{2t-s}f) (\gamma(s))|.$$

Now we give a crucial result.

LEMMA 3.4. *For any $s \geq 0$, we have the upper bound*

$$|\nabla P_s (\log P_{2t-s}f)| \leq e^{Ks} P_s \frac{|\nabla P_{2t-s}f|}{P_{2t-s}f},$$

with $K := -\inf_{\mathbb{R}^d} \nabla^2 V > 0$.

We do not provide the proof. It is sufficient to adapt Lemma 1.3 in [9]. Moreover, it is a particular case of Lemma 3.7 in [3].

Lemma 3.4 together with (3.9) yields

$$(3.10) \quad \xi'(s) \leq -P_s \frac{|\nabla P_{2t-s}f|^2}{(P_{2t-s}f)^2} (\gamma(s)) + \frac{|h'(s)|}{t} |x - y| e^{Ks} P_s \frac{|\nabla P_{2t-s}f|}{P_{2t-s}f} (\gamma(s)).$$

By putting

$$X := \frac{|\nabla P_{2t-s}f|}{P_{2t-s}f} (\gamma(s)) \quad \text{and} \quad Y := \frac{|h'(s)|}{2t} |x - y| e^{Ks},$$

we have

$$\xi'(s) \leq -P_s(X^2 + 2XY) \leq P_s Y^2 = \frac{|h'(s)|^2}{4t^2} |x - y|^2 e^{2Ks}.$$

Consequently, we have the inequality

$$(3.11) \quad \xi(t) - \xi(0) \leq \int_0^t \frac{|h'(s)|^2}{4t^2} |x - y|^2 e^{2Ks} ds.$$

However, $\xi(t) = P_t \log P_t f(x)$ and $\xi(0) = \log P_{2t} f(y)$. Thus, we have the log-Harnack inequality

$$(3.12) \quad P_t \log P_t f(x) \leq \log P_{2t} f(y) + \frac{|x - y|^2}{4t^2} \int_0^t |h'(s)|^2 e^{2Ks} ds.$$

From inequality (3.12) we have the finiteness of the relative entropy with respect to the probability measure μ^σ . However, we will give a better result by linking the entropy to the Wasserstein distance between ν_0 and μ^σ .

To obtain the best inequality, we take

$$h(s) := t \frac{e^{-2Ks} - 1}{e^{-2Kt} - 1}.$$

Thus, we obtain

$$(3.13) \quad P_t \log P_t f(x) \leq \log P_{2t} f(y) + \frac{|x - y|^2}{2S(t)}$$

with

$$(3.14) \quad \frac{1}{S(t)} = K \left[1 - \frac{1}{1 - e^{2Kt}} \right].$$

We take the infimum for y running over \mathbb{R}^d and obtain

$$(3.15) \quad P_t \log P_t f(x) \leq \frac{1}{S(t)} \min_{y \in \mathbb{R}^d} \left\{ S(t) \varphi(y) + \frac{1}{2} |x - y|^2 \right\}$$

with $\varphi(y) := \log P_{2t} f(y)$. However, by Jensen's inequality, we have

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(y) \mu^\sigma(dy) &= \int_{\mathbb{R}^d} \log(P_{2t} f)(y) \mu^\sigma(dy) \\ &\leq \log \int_{\mathbb{R}^d} P_{2t} f \mu^\sigma(dy) = \log \int_{\mathbb{R}} \nu_{2t}(dy) = 0. \end{aligned}$$

Consequently, inequality (3.15) becomes

$$(3.16) \quad P_t \log P_t f(x) \leq \frac{1}{S(t)} \left[\min_{y \in \mathbb{R}^d} \left\{ S(t) \varphi(y) + \frac{1}{2} |x - y|^2 \right\} - \int_{\mathbb{R}^d} S(t) \varphi(y) \mu^\sigma(dy) \right].$$

We take the supremum for φ bounded and measurable, and then we carry out the integration over $x \in \mathbb{R}^d$. We obtain

$$(3.17) \quad H(\nu_t | \mu^\sigma) \leq \frac{1}{S(t)} \sup_{\varphi} \left[\int_{\mathbb{R}^d} \min_{y \in \mathbb{R}^d} \left(\varphi(y) + \frac{1}{2} |x - y|^2 \right) \nu_0(dx) - \int_{\mathbb{R}^d} \varphi(y) \mu^\sigma(dy) \right].$$

We now use the Monge–Kantorovich duality. For any measure ν , we have

$$\mathbb{W}_2^2(\nu; \mu^\sigma) = \sup_{\varphi} \left[\int_{\mathbb{R}^d} \min_{y \in \mathbb{R}^d} \left(\varphi(y) + \frac{1}{2} |x - y|^2 \right) \nu(dx) - \int_{\mathbb{R}^d} \varphi(y) \mu^\sigma(dy) \right].$$

We refer the reader to page 678 in [1]. This yields

$$(3.18) \quad H(\nu_t | \mu^\sigma) \leq \frac{1}{S(t)} \mathbb{W}_2^2(\nu_0; \mu^\sigma),$$

which completes the proof.

Let us observe that $1/S(t)$ goes to infinity as t goes to zero.

REFERENCES

- [1] S. G. Bobkov, I. Gentil, and M. Ledoux, *Hypercontractivity of Hamilton–Jacobi equations*, J. Math. Pures Appl. (9) 80 (7) (2001), pp. 669–696.
- [2] F. Bolley, I. Gentil, and A. Guillin, *Uniform convergence to equilibrium for granular media*, Arch. Ration. Mech. Anal. 208 (2) (2013), pp. 429–445.
- [3] J.-F. Collet and F. Malrieu, *Logarithmic Sobolev inequalities for inhomogeneous semi-groups*, ESAIM Probab. Stat. 12 (2008), pp. 492–504.
- [4] P. Del Moral and J. Tugaut, *Uniform propagation of chaos and creation of chaos for a class of nonlinear diffusions* (2013), <https://hal.archives-ouvertes.fr/hal-00798813>
- [5] S. Herrmann, P. Imkeller, and D. Peithmann, *Large deviations and a Kramers’ type law for self-stabilizing diffusions*, Ann. Appl. Probab. 18 (4) (2008), pp. 1379–1423.
- [6] S. Herrmann and J. Tugaut, *Non-uniqueness of stationary measures for self-stabilizing processes*, Stochastic Process. Appl. 120 (7) (2010), pp. 1215–1246.
- [7] S. Herrmann and J. Tugaut, *Stationary measures for self-stabilizing processes: Asymptotic analysis in the small noise limit*, Electron. J. Probab. 15 (2010), pp. 2087–2116.
- [8] S. Herrmann and J. Tugaut, *Self-stabilizing processes: Uniqueness problem for stationary measures and convergence rate in the small-noise limit*, ESAIM Probab. Stat. 16 (2012), pp. 277–305.
- [9] M. Ledoux, *The geometry of Markov diffusion generators*, Ann. Fac. Sci. Toulouse Math. (6) 9 (2) (2000), pp. 305–366.
- [10] H. P. McKean, Jr., *A class of Markov processes associated with nonlinear parabolic equations*, Proc. Natl. Acad. Sci. USA 56 (1966), pp. 1907–1911.

- [11] H. P. McKean, Jr., *Propagation of chaos for a class of nonlinear parabolic equations*, in: *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, Air Force Office Sci. Res., Arlington, VA, 1967, pp. 41–57.
- [12] F. Otto and C. Villani, *Comment on: “Hypercontractivity of Hamilton–Jacobi equations”*, *J. Math. Pures Appl.* (9) 80 (7) (2001), pp. 697–700.
- [13] J. Tugaut, *Convergence to the equilibria for self-stabilizing processes in double-well landscape*, *Ann. Probab.* 41 (3A) (2013), pp. 1427–1460.
- [14] J. Tugaut, *Self-stabilizing processes in multi-wells landscape in \mathbb{R}^d – Convergence*, *Stochastic Process. Appl.* 123 (5) (2013), pp. 1780–1801.
- [15] J. Tugaut, *Phase transitions of McKean–Vlasov processes in double-wells landscape*, *Stochastics* 86 (2) (2014), pp. 257–284.
- [16] J. Tugaut, *Self-stabilizing processes in multi-wells landscape in \mathbb{R}^d – Invariant probabilities*, *J. Theoret. Probab.* 27 (1) (2014), pp. 57–79.

Julian Tugaut
 Université Jean Monnet, Saint-Étienne
 and Institut Camille Jordan, Lyon
 23, rue du docteur Paul Michelon, CS 82301
 42023 Saint-Étienne Cedex 2, France
E-mail: julian.tugaut@univ-st-etienne.fr

*Received on 2.12.2016;
 revised version on 10.9.2017*
