FINITENESS OF ENTROPY FOR GRANULAR MEDIA EQUATIONS

BY

JULIAN TUGAUT∗ (SAINT-ÉTIENNE)

Abstract. The current work deals with the granular media equation, which 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 proven that all the moments are bounded for positive \( t \). However, the finiteness of the entropy of the solution is a new result, that we present 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

We aim 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], [14], we prove the long-time convergence under three assumptions:

• the finiteness of some moment,
• the fact that \( \mu_0 \) is absolutely continuous with respect to the Lebesgue measure,
• the finiteness of the entropy of \( \mu_0 \).

The first hypothesis is necessary to prove the existence of a solution to the self-stabilizing diffusion (which law is the solution of the granular media equation) so that this hypothesis can not be relaxed.

The second hypothesis may be suppressed. Indeed, we know - see [McK66], [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 leads me to simplify the writing of this paper. I would 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 Wasserstein distance and the rate of convergence. But, they assume the finiteness of entropy of the initial law $\mu_0$.

The results in [2] have been used in [3] in order 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 $\mu_0$ which are discrete probability measures. Consequently, we had to adapt the results and were not able to apply it 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.

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

\begin{equation}
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,
\end{equation}

$V$ 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], [McK66] 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:

\begin{equation}
\frac{\partial u}{\partial t} = \nabla \left\{ \frac{\sigma^2}{2} \nabla u + u (\nabla V + \nabla F * u) \right\}.
\end{equation}

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

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

This generator does depend on the time so the semi-group may not be time-reversible.

Let us now give the assumptions of the paper. We take the same hypotheses than the ones in [13]:

(A-1): The potential $V$ is a smooth function.

(A-2): There exists a compact subset $K$ of $\mathbb{R}^d$ such that $\nabla^2 V(x) > 0$ for all $x \notin K$. Moreover, $\lim_{||x|| \to +\infty} \nabla^2 V(x) = +\infty$. 
(A-3): The gradient $\nabla V$ is slowly increasing: there exist $m \in \mathbb{N}$, $C > 0$ and a function $R$ from $\mathbb{R}^d$ to $\mathbb{R}^d$ such that

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

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

(A-4): There exists an even polynomial function $G$ on $\mathbb{R}$ such that $F(x) = G(||x||)$. And, $\text{deg}(G) =: 2n \geq 2$.

(A-5): The function $G$ is convex.

Typically, $V$ is a polynomial function, like $V(x) = x^4 - x^2$ (in dimension 1) and $F(x) = \frac{x^2}{2}$. Let us point out that the convexity of $G$ is not necessary on this work.

We need to assume another hypothesis:

(A-6): The moment of order $8q^2$ of the law $\mu_0$ is finite:

$$\int_{\mathbb{R}^d} x^{8q^2} \mu_0(dx) < \infty,$$

where $q := \max \{m; n\}$.

Under hypothesis (A-1)–(A-6), 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:

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

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

According to [16], there exists an invariant probability $\mu^\sigma$. Let us note that we may not have uniqueness of this invariant probability, see [3], [10], [15].

We here aim to prove the finiteness of the quantity

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

for any $t > 0$. In [3], 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 one about the convergence in Wasserstein distance (Corollary D which comes from [2]).

In a last section, we prove Theorem A.
2. MAIN RESULTS

We first give the main result of the current work.

**Theorem A:** Under assumptions (A-1)–(A-6), 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 \( \nabla V + \nabla_x F(x, \mu_t) \).

Thanks to Theorem A in [14], we deduce the following result.

**Corollary C:** Under the hypotheses of the article, if moreover the set of invariant probabilities is discrete (see [6], [7], [8] for assumptions such that there are exactly three invariant probabilities), we have the weak convergence in long-time of \( \mu_t \) toward an invariant probability \( \mu^\frac{\star}{\star} \).

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

**Corollary D:** Under the assumptions of the article, if moreover \( V \) is strictly convex (but not necessarily uniformly strictly convex), \( \mu_t \) converges, for 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 \( t > 0 \), \( \mu_t \) is absolutely continuous with respect to the Lebesgue measure so that there exists \( f_t \) such that

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

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 \( g_t := P_t g \) where the semi-group \( (P_t)_{t \geq 0} \) is generated by

\[
L_t = \frac{\sigma^2}{2} \Delta - (\nabla V + \nabla F \ast \mu_t) \cdot \nabla.
\]

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

Let us recall Proposition 2.1 in [10]:
Lemma 3.1. For any $\sigma > 0$, there exists an invariant probability $\mu^\sigma$ to Diffusion (1.1).

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

We will consider the relative entropy with respect to $\mu^\sigma$:

\[
H(\nu \mid \mu^\sigma) = \int \frac{f}{\mu^\sigma} \log \left( \frac{f}{\mu^\sigma} \right) \mu^\sigma(dx),
\]

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

Let us remind the reader that the probability measure $\mu^\sigma$ - see [3], [16] - satisfies:

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

Consequently, we get

\[
H(\nu \mid \mu^\sigma) = \int_{\mathbb{R}^d} f(x) \log(f(x)) dx + \frac{2}{\sigma^2} \int_{\mathbb{R}^d} (V(x) + F \ast \mu^\sigma(x)) f(x) dx
\]

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

However, according to Theorem 2.13 in [3], 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 obtain that the quantity

\[
\int_{\mathbb{R}^d} (V(x) + F \ast \mu^\sigma(x)) g_t(x) dx
\]

is bounded by a constant $C_0$.

Moreover, $F \ast \mu^\sigma$ is a polynomial function with parameters which only depends on $\sigma$ (through the moments of the fixed measure $\mu^\sigma$). We also know that it is convex. Since $V$ has polynomial behaviour and is convex at infinity, the function

\[
x \mapsto \exp \left[ -\frac{2}{\sigma^2} (V(x) + F \ast \mu^\sigma(x)) \right]
\]

is integrable with respect to the Lebesgue measure. As a consequence, the quantity

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

is finite. Since $\sigma$ is fixed, it is bounded with respect to the time.

We deduce that it is sufficient to prove the finiteness of $H(P_t g \mid \mu^\sigma)$. 
LEMMA 3.2. Let $X_0$ be a random variable which follows the law $\mu^\sigma$. Then, for any $t \geq 0$, for any functions $f$ and $g$, we have

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

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 Kolmogorov diffusion and we thus have the time reversibility (3.5).

We put $f_0 := \frac{\partial}{\partial \mu^\sigma}$. We will work with $f_0$ and $(f_t)_{t \geq 0}$.

We proceed like in [11]. We have

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

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

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

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 $x(s) := y + \frac{t}{t} (x - y)$ for any $s \in [0; t]$. We also consider a function $h$ from $[0; t]$ to $[0; t]$ which is $C^1$-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 $\gamma$ plays the role of a geodesic between $x$ and $y$ with respect to the Riemannian metric of the diffusion.

We now introduce

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

LEMMA 3.3. We have the following derivative:

$$\frac{d \xi}{ds} = -P_s \frac{\nabla P_{2t - s} f}{(P_{2t - s} f)^2} (\gamma (s)) + \frac{h'(s)}{t} \langle \nabla P_s \log (P_{2t - s} f) (\gamma (s)) ; x - y \rangle.$$
**Proof.** We put \( g := P_{2t-s} f \). Thus, we have:

\[
\xi'(s) = P_s ( L_s \log g ) (\gamma(s)) - P_s \frac{L_s g}{g} (\gamma(s)) \\
+ \langle \nabla P_s (\log g) (\gamma(s)) ; \gamma'(s) \rangle.
\]

Thanks to 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

\[
\xi'(s) = - P_s \frac{\Gamma (g, g)}{g^2} (\gamma(s)) + \langle \gamma'(s) ; \nabla P_s (\log g) (\gamma(s)) \rangle \\
= - P_s \frac{\nabla P_{2t-s} f}{(P_{2t-s} f)^2} (\gamma(s)) + \frac{h'(s)}{t} \langle \nabla P_s \log (P_{2t-s} f) (\gamma(s)) ; x - y \rangle.
\]

We have the immediate upper-bound:

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

We now give a crucial result.

**Lemma 3.4.** For any \( s \geq 0 \), we have the upper-bound:

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

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].

This lemma, together with (3.9) yields:

\[
(3.10) \\
\xi'(s) \lessgtr - P_s \frac{\nabla P_{2t-s} f}{(P_{2t-s} f)^2} (\gamma(s)) + \frac{|h'(s)|}{t} |x - y| e^{K_s} P_s \frac{|\nabla P_{2t-s} f|}{P_{2t-s} f} (\gamma(s)).
\]

By putting \( X := \frac{\nabla P_{2t-s} f}{P_{2t-s} f} (\gamma(s)) \) and \( Y := \frac{|h'(s)|}{2t} |x - y| e^{K_s} \), we have

\[
\xi'(s) \lessgtr - P_s (X^2 + 2XY) \leq P_s Y^2 = \frac{|h'(s)|^2}{4t^2} |x - y|^2 e^{2K_s}.
\]
Consequently, we have the inequality

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

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

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

¿From Inequality (3.12), we have the finiteness of the relative entropy with respect to the probability measure \( \mu^\sigma \). However, we will give a better result by linking the entropy to the Wasserstein distance between \( \nu_0 \) and \( \mu^\sigma \).

To obtain the best inequality, we take \( h(s) := \frac{e^{-2Ks} - 1}{e^{-2Kt} - 1} \). We thus obtain

\[ P_t \log P_t f(x) \leq \log P_{2t} f(y) + \frac{|x - y|^2}{2S(t)}, \]

with

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

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

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

with \( \varphi(y) := \log P_{2t} f(y) \). However, by Jensen inequality, we have the following:

\[ \int_{\mathbb{R}^d} \varphi(y) \mu^\sigma(dy) = \int_{\mathbb{R}^d} \log(P_{2t} f(y) \mu^\sigma(dy) \leq \log \int_{\mathbb{R}^d} P_{2t} f \mu^\sigma = \log \int_{\mathbb{R}^d} \nu_0(dy) = 0. \]

Consequently, Inequality (3.15) becomes

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

We take the supremum for \( \varphi \) bounded and measurable then we take the integration over \( x \in \mathbb{R}^d \). We obtain

\[ H (\nu_t, \mu^\sigma) \leq \frac{1}{S(t)} \sup_{\varphi} \left[ \int_{\mathbb{R}^d} \min_{y \in \mathbb{R}^d} \left( \varphi(y) + \frac{1}{2} |x - y|^2 \right) \nu_0(dx) - \int_{\mathbb{R}^d} \varphi(y) \mu^\sigma(dy) \right]. \]
We now give the Monge-Kantorovitch duality. For any measure $\nu$, 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 the 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 achieves the proof.

Let us observe that $\frac{1}{S(t)}$ goes to infinity as $t$ goes to 0.

**REFERENCES**


Université Jean Monnet, Institut Camille Jordan
23, rue du docteur Paul Michelon, CS 82301,
42023 Saint-Étienne Cedex 2, France
E-mail: tugaut@math.cnrs.fr

Received on 2.12.2016;
revised version on 10.9.2017